HU e-ISSN 1787-2413

Miskolc Mathematical Notes

A Publication of the University of Miskolc

VOLUME 7 (2006), NUMBER 1

MISKOLC UNIVERSITY PRESS

Miskolc Mathematical Notes

VOLUME 7 (2006), NUMBER 1

CONTENTS

Doi: 10.18514/MMN.2006.94

Miskolc Mathematical Notes
Vol. 7 (2006), No. 1, pp. 3–11

HU ISSN 1787-2413
electronic version

ON UPPER AND LOWER D_δ-SUPERCONTINUOUS MULTIFUNCTIONS

METIN AKDAĞ

[Received: March 4, 2004]

ABSTRACT. In this paper, we define a multifunction $F : X \rightsquigarrow Y$ as an upper (lower) D_δ-supercontinuous if $F^+(V)$ $(F^-(V))$ is d_δ-open in X for every open set V of Y. We obtain some characterizations and several properties concerning upper (lower) D_δ-supercontinuous multifunctions.

Mathematics Subject Classification: Primary 54C10; Secondary 54C60

Keywords: Multifunction, D-supercontinuous multifunction, Z-supercontinuous multifunction, supercontinuous multifunction and strongly θ-continuous multifunctions

1. INTRODUCTION

Several weak and strong variants of continuity of multifunctions occur in the literature. The strong variants of continuity of multifunctions we shall be dealing with in this paper are treated in [1–4]. Certain of these strong forms of continuity of multifunctions coincide with the continuity of multifunctions if the domain or range space is suitably augmented. In 2003, J. K. Kohli [5] introduced the concept of D_δ-supercontinuous functions and gave some properties of D_δ-supercontinuous functions. In this paper we introduce a new strong form of continuity of multifunctions called the "upper (lower) D_δ-supercontinuity," which coincides with the upper (lower) continuity if the domain or range is a D_δ-completely regular space. Characterizations and basic properties of upper (lower) D_δ-supercontinuous multifunctions are elaborated in Section 3. In Section 4, we show that if the domain of a upper (lower) D_δ-supercontinuous multifunction F is retopologized in an appropriate way, then F is simply a continuous multifunction.

A multifunction F of a set X into a set Y is a set-valued function mapping set of X into $2^Y \setminus \{\varnothing\}$, where 2^Y stands for the power set of Y. Let A be a subset of a topological space (X, τ). $\mathring{A}$ and $\bar{A}$ denote the interior and closure of A respectively. A multifunction F of a set X into Y is a correspondence such that $F(x)$ is a nonempty subset of Y for each $x \in X$. We will denote such a multifunction by $F : X \rightsquigarrow Y$. For a multifunction F, the upper and lower inverse set of a set B of Y will be denoted by $F^+(B)$ and $F^-(B)$ respectively that is $F^+(B) = \{x \in X : F(x) \subseteq B\}$ and $F^-(B) = \{x \in X : F(x) \cap B \neq \varnothing\}$. The graph $G(F)$ of the multifunction $F : X \rightsquigarrow Y$ is strongly closed [4] if for each $(x, y) \notin G(F)$, there exist open sets U and V containing x and containing y respectively such that $(U \times \overline{V}) \cap G(F) = \emptyset$ [6]. A multifunction $F : X \rightsquigarrow Y$ is said to be upper semi continuous (briefly u. s. c.) at a point $x \in X$ if for each open set V in Y with $F(x) \subseteq V$, there exists an open set U containing x such that $F(U) \subseteq V$; lower semi continuous (briefly l. s. c.) at a point $x \in X$ if for each open set V in Y with $F(x) \cap V \neq \emptyset$, there exists an open set U containing x such that $F(z) \cap V \neq \emptyset$ for every $z \in U$. A set G in a topological space X is said to be z-open if for each $x \in G$ there exists a cozero set H such that $x \in H \subset G$, or equivalently, if G is expressible as the union of cozero sets. The complement of a z-open set will be referred to as a z-closed set [7]. A subset H in a space X is said to be a regular G_δ-set if H is an intersection of a sequence of closed sets whose interior contain H. The complement of a regular G_δ-set is called a regular F_σ-set [8].

Throughout this paper, the spaces (X, τ) and (Y, σ) (or simply X and Y) always mean topological spaces, and $F : X \rightsquigarrow Y$ (resp., $f : X \to Y$) is a multivalued (resp., single valued) function.

2. Preliminaries and basic properties

Definition 1. (a) A multifunction $F : X \rightsquigarrow Y$ is called upper D_δ-supercontinuous (u. D_δ-super c.) at a point $x \in X$ if for any open set $V \subset Y$ such that $F(x) \subset V$ there exists a regular F_σ-set $U \subset X$ containing x such that $F(U) \subset V$.

(b) A multifunction $F : X \rightsquigarrow Y$ is called lower D_δ-supercontinuous (l. D_δ-super c.) at a point $x \in X$ if for any open set $V \subset Y$ such that $F(x) \cap V \neq \emptyset$ there exists a regular F_σ-set $U \subset X$ containing x such that $F(u) \cap V \neq \emptyset$ for every $u \in U$.

(c) F is said to be D_δ-supercontinuous (briefly D_δ-super c.) at $x \in X$, if it is both u. D_δ-super c. and l. D_δ-super c. at $x \in X$.

(d) F is said to be u. D_δ-super c. (l. D_δ-super c., D_δ-super c.) on X, if it has this property at each point $x \in X$.

Example 1. Let X denote the real line endowed by the cofinite topology, $Y = \{a, b\}$ with the topology $\tau = \{\emptyset, Y, \{a\}\}$ and we define the multifunction as follows: $F : X \rightsquigarrow Y$, $F(x) = \{a\}$ if x is irrational and $F(x) = \{b\}$ if x is rational. Then F is u. D_δ-supercontinuous (l. D_δ-super c.).

Definition 2 ([3])**.** A multifunction $F : X \rightsquigarrow Y$ is said to be (a) upper Z-supercontinuous (briefly, u. Z-super c.) at a point $x \in X$ if for every open set V with $F(x) \subset V$, there exists a cozero set U containing x such that $F(U) = \bigcup\{F(u) : u \in U\} \subset V$;

(b) lower Z-supercontinuous (l. Z-super c.) at a point $x \in X$ if for every open set V with $F(x) \cap V \neq \emptyset$, there exists a cozero set U containing x such that $F(u) \cap V \neq \emptyset$ for every $u \in U$.

Example 2. Let X denote the real line endowed with usual topology and we define the multifunction as follows: $F : X \rightsquigarrow X$, $F(x) = \{x\}$ for each $x \in X$. Then F is u. Z-supercontinuous (l. Z-super c.).

Definition 3 ([4])**.** (a) A multifunction $F : X \rightsquigarrow Y$ is called strongly θ-upper semi continuous (s. θ-u. s. c.) at a point $x \in X$ if for any open set $V \subset Y$ such that $F(x) \subset V$ there exists an open set $U \subset X$ containing x such that $F(\overline{U}) \subset V$.

(b) A multifunction $F : X \rightsquigarrow Y$ is called strongly θ-lower semi continuous (s. θ-l. s. c.) at a point $x \in X$ if for any open set $V \subset Y$ such that $F(x) \cap V \neq \varnothing$ there exists an open set $U \subset X$ containing x such that $F(u) \cap V \neq \varnothing$ for every $u \in \overline{U}$.

Definition 4 ([2])**.** (a) A multifunction $F : X \rightsquigarrow Y$ is called upper supercontinuous (u. super c.) at a point $x \in X$ if for any open set $V \subset Y$ such that $F(x) \subset V$ there exists an open set $U \subset X$ containing x such that $F(\overset{\circ}{\overline{U}}) \subset V$.

(b) A multifunction $F : X \rightsquigarrow Y$ is called lower supercontinuous (l. super c.) at a point $x \in X$ if for any open set $V \subset Y$ such that $F(x) \cap V \neq \varnothing$ there exists an open set $U \subset X$ containing x such that $F(u) \cap V \neq \varnothing$ for every $u \in \overset{\circ}{\overline{U}}$.

Definition 5 ([1])**.** (a) A multifunction $F : X \rightsquigarrow Y$ is called upper D-supercontinuous (u. D-super c.) at a point $x \in X$ if for any open set $V \subset Y$ such that $F(x) \subset V$ there exists an open F_σ-set $U \subset X$ containing x such that $F(U) \subset V$.

(b) A multifunction $F : X \rightsquigarrow Y$ is called lower D-supercontinuous (l. D-super c.) at a point $x \in X$ if for any open set $V \subset Y$ such that $F(x) \cap V \neq \varnothing$ there exists an open F_σ-set $U \subset X$ containing x such that $F(u) \cap V \neq \varnothing$ for every $u \in U$.

The following diagram well illustrates the relations that exist between u. D_δ-supercontinuous (l. D_δ-supercontinuous) multifunctions and other strong variants of continuity defined above.

$$
\begin{tikzcd}
\text{u. } Z\text{-super c. (l. } Z\text{-super c.)} \arrow[d, Rightarrow] & \text{u. super c. (l. super c.)} \\
\text{u. } D_\delta\text{-super c. (l. } D_\delta\text{-super c.)} \arrow[d, Rightarrow] & \text{u. strongly } \theta\text{-c. (l. strongly } \theta\text{-c.)} \arrow[u, Rightarrow] \arrow[d, Rightarrow] \\
\text{u. } D\text{-super c. (l. } D\text{-super c.)} & \text{u. semi c. (l. semi c.)}
\end{tikzcd}
$$

None of the above implications in general is reversible, as will be shown the sequel. We give Example 3, Example 4, and Example 5 to show that a u. strongly θ-c. (l. strongly θ-c.) multifunction need not be u. Z-super c. (l. Z-super c.) and that u. D-super c. (l. D-super c.) multifunction need not be u. Z-super c. (l. Z-super c.) and u. D_δ-super c. (l. D_δ-super c.) multifunction need not be u. Z-super c. (l. Z-super c.) and that u. D-super c. (l. D-super c.) multifunction.

Example 3 ([5]). Let $X = Y$ be the Mountain chain space due to Helderman [7], which is a regular space but not a D_δ-completely regular space [10]. Then the multifunction $F : X \rightsquigarrow X$, $F(x) = \{x\}$ for each $x \in X$ is a u. strongly θ-continuous (l. strongly θ-continuous) but not u. D_δ-supercontinuous (l. D_δ-supercontinuous).

Example 4. Let X denote the set of positive integers endowed with cofinite topology. Then the multifunction $F : X \rightsquigarrow X$, $F(x) = \{x\}$ for each $x \in X$ is u. D-supercontinuous (l. D-supercontinuous) but neither u. supercontinuous (l. supercontinuous) nor u. strongly θ-continuous (l. strongly θ-continuous) and hence not u. Z-supercontinuous (l. Z-supercontinuous).

Example 5. Consider the space A defined by E. Hewitt in [10, p. 504], which is a D_δ-completely regular but not completely regular. It turns out that the multifunction $F : A \rightsquigarrow A$, $F(x) = \{x\}$ for each $x \in X$ is u. D_δ-supercontinuous (l. D_δ-supercontinuous) but not u. Z-supercontinuous (l. Z-supercontinuous.).

3. Characterizations

Definition 6. A set G in a topological space X is said to be d_δ-open if for each $x \in G$ there exists a regular F_σ-set H such that $x \in H \subset G$. The complement of a d_δ-open set will be referred to as a d_δ-closed set [5].

Theorem 1. *The following statements are equivalent for a multifunction $F : X \rightsquigarrow Y$:*

(a) *F is u. D_δ-super c. (l. D_δ-super c.).*
(b) *For each open set $V \subseteq Y$, $F^+(V)$ ($F^-(V)$) is a d_δ-open set in X.*
(c) *For each closed set $K \subseteq Y$, $F^-(K)$ ($F^+(K)$) is a d_δ-closed set in X.*
(d) *For each x of X and for each open set V with $F(x) \subset V (F(x) \cap V \neq \emptyset)$, there is a d_δ-open set U containing x such that the implication $y \in U \Rightarrow F(y) \subset V$ holds ($F(y) \cap V \neq \emptyset$).*

Proof. (a)⇒(b): Let V be an open set of Y and $x \in F^+(V)$. Then there exist a regular F_σ-set U containing x such that $F(U) \subset V$. Then $U \subset F^+(V)$. Since U is regular F_σ-set, we have $x \in U \subset F^+(V)$.

(b)⇒(c): Let K be a closed set of Y. Then $Y - K$ is an open set and $F^+(Y - K) = X - F^-(K)$ is d_δ-open. Thus, $F^-(K)$ is d_δ-closed in X.

(c)⇒(b): Obvious.

(b)⇒(a): Let V be an open set of Y containing $F(x)$. Then $F^+(V)$ is d_δ-open and $x \in F^+(V)$. Since $F^+(V)$ is a d_δ-open set there exists a regular F_σ-set U containing x such that $U \subset F^+(V)$. Thus, $F(U) \subset F(F^+(V)) \subset V$.

(a)⇔(d): Clear. □

The proof for the case where F is l. D_δ-super c. is similar.

Definition 7. Let X be a topological space and let $A \subset X$. A point $x \in X$ is said to be a d_δ-adherent point of A if every regular F_σ-set containing x intersects A. Let $[A]_{d_\delta}$ denote the set of all d_δ-adherent points of A. Clearly the set A is d_δ-closed if and only if $[A]_{d_\delta} = A$ [5].

Theorem 2. *A multifunction $F : X \rightsquigarrow Y$ is l. D_δ-super c. if and only if $F([A]_{d_\delta}) \subset \overline{F(A)}$ for every $A \subset X$.*

Proof. Suppose that F is l. D_δ-super c. Since $\overline{F(A)}$ is closed in Y, by Theorem 1, $F^+(\overline{F(A)})$ is d_δ-closed in X. Also, since $A \subset F^+(\overline{F(A)})$, $[A]_{d_\delta} \subset [F^+(\overline{F(A)})]_{d_\delta} = F^+F([A]_{d_\delta})$. Thus, $F([A]_{d_\delta}) \subset F(F^+(\overline{F(A)})) \subset \overline{F(A)}$.

Conversely, suppose $F([A]_{d_\delta}) \subset \overline{F(A)}$ for every $A \subset X$. Let K be any closed set in Y. Then $F([F^+(K)]_{d_\delta}) \subset \overline{F(F^+(K))}$ and $\overline{F(F^+(K))} \subset \overline{K} = K$. Hence, $[F^+(K)]_{d_\delta} \subset F^+(K)$ which shows that F is l. D_δ-super c. □

Theorem 3. *A multifunction F from a space X into a space Y is l. D_δ-super c. if and only if $[F^+(B)]_{d_\delta} \subset F^+(\overline{B})$ for every $B \subset Y$.*

Proof. Suppose F is l. D_δ-super c. By Theorem 1, $F^+(\overline{B})$ is d_δ-closed in X for every $B \subset Y$ and $F^+(\overline{B}) = [F^+(\overline{B})]_{d_\delta}$. Hence, $[F^+(B)]_{d_\delta} \subset F^+(\overline{B})$.

Conversely, let K be any closed set in Y. Then $[F^+(K)]_{d_\delta} \subset F^+(\overline{K}) = F^+(K) \subset [F^+(K)]_{d_\delta}$. Thus, we have $F^+(K) = [F^+(K)]_{d_\delta}$, which in turn implies that F is l. D_δ-super c. □

Definition 8. A filter base $\mathcal{F}$ is said to d_δ-converge to a point x (written as $\mathcal{F} \xrightarrow{d_\delta} x$) if for every regular F_σ-set containing x contains a member of $\mathcal{F}$ [5].

Theorem 4. *A multifunction $F : X \rightsquigarrow Y$ is l. D_δ-super c. a point x of X if and only if for each $x \in X$ and each filter base $\mathcal{F}$ that D_δ-converges to x, $F(\mathcal{F}) \to F(x)$.*

Proof. Assume that F is l. D_δ-super c. and let $\mathcal{F} \xrightarrow{d_\delta} x$. Let W be an open set containing $F(x)$. Then $x \in F^-(W)$ and $F^-(W)$ is d_δ-open. Let H be a regular F_σ-set in X such that $x \in H \subset F^-(W)$. Since $\mathcal{F} \xrightarrow{d_\delta} x$, there exists $U \in \mathcal{F}$ such that $U \subset H$ and so $F(U) \subset F(H) \subset W$. Thus, $F(\mathcal{F}) \to F(x)$.

Conversely, let W be an open subset of Y containing $F(x)$. Now, the filter $\mathcal{F}$ generated by the filterbase $\aleph_x$ consisting of all regular F_σ-sets containing x, d_δ-converges to x and so by hypothesis $F(\mathcal{F}) \to F(x)$. Hence, there exists a member $F(N)$ of $F(\aleph_x)$ such that $F(N) \subset W$. Since $N \in \aleph_x$, N is a regular F_σ-set containing x. Thus, F is l. D_δ-super c. at x. □

Theorem 5. *If $F : X \rightsquigarrow Y$ is u. D_δ-super c. (l. D_δ-super c.) and $F(X)$ is endowed with subspace topology, then $F : X \rightsquigarrow F(X)$ is u. D_δ-super c. (l. D_δ-super c.)*

Proof. Since $F : X \rightsquigarrow Y$ is u. D_δ-super c. (l. D_δ-super c.), for every open subset V of Y, $F^+(V \cap F(X)) = F^+(V) \cap F^+(F(X)) = F^+(V)(F^-(V \cap F(X)) = F^-(V) \cap F(F(X)) = F^-(V))$ is d_δ-open. Hence, $F : X \rightsquigarrow F(X)$ is u. D_δ-super c. (l. D_δ-super c.) □

Theorem 6. *If $F : X \rightsquigarrow Y$ is u. D_δ-super c. (l. D_δ-super c.) and $G : Y \rightsquigarrow Z$ u. s. c. (l. s. c.), then $G \circ F$ is u. D_δ-super c. (l. D_δ-super c.).*

Proof. Let V be an open subset of Z. Since G is u. s. c. (l. s. c.), it follows that $G^+(V)$ ($G^-(V)$) is open subset of Y, and since F is u. D_δ-super c. (l. D_δ-super c.), one has that $F^+(G^+(V))$ $(F^-(G^-(V)))$ is d_δ-open in X. Thus, $G \circ F$ is u. D_δ-super c. (l. D_δ-super c.). □

Theorem 7. *Let $\{F_\alpha : X \rightsquigarrow X_\alpha, \alpha \in \Delta\}$ be a family of multifunctions and let $F : X \rightsquigarrow \prod_{\alpha\in\Delta} X_\alpha$ be defined by $F(x) = (F_\alpha(x))$. Then F is u. D_δ-super c. if and only if each $F_\alpha : X \rightsquigarrow X_\alpha$ is u. D_δ-super c.*

Proof. Let G_{α_0} be an open set of X_{α_0}. Then $(P_{\alpha_0} \circ F)^+(G_{\alpha_0}) = F^+(P^+_{\alpha_0}(G_{\alpha_0})) = F^+(G_{\alpha_0} \times \prod_{\alpha\neq\alpha_0} X_\alpha)$. Since F is u. D_δ-super c. $F^+(G_{\alpha_0} \times \prod_{\alpha\neq\alpha_0} X_\alpha)$ is d_δ-open in X. Thus, $P_{\alpha_0} \circ F = F_\alpha$ is u. D_δ-super c. Here, P_α denotes the projection of X onto the α-coordinate space X_α.

Conversely, let us suppose that each $F_\alpha : X \rightsquigarrow X_\alpha$ is u. D_δ-super c. To show that the multifunction F is u. D_δ-super c., in view of Theorem 1, it is sufficient to show that $F^+(V)$ is d_δ-open for each open set V in the product space $\prod_{\alpha\in\Delta} X_\alpha$. Since the finite intersections and arbitrary unions of d_δ-open sets are d_δ-open, it suffices to prove that $F^+(S)$ is d_δ-open for every subbasic open set S in the product space $\prod_{\alpha\in\Delta} X_\alpha$. Let $U_\beta \times \prod_{\alpha\neq\beta} X_\alpha$ be a subbasic open set in $\prod_{\alpha\in\Delta} X_\alpha$. Then $F^+(U_\beta \times \prod_{\alpha\neq\beta} X_\alpha) = F^+(P^+_\beta(U_\beta)) = F^+_\beta(U_\beta)$ is d_δ-open. Hence, F is u. D_δ-super c. □

Theorem 8. *Let $F : X \rightsquigarrow Y$ be a multifunction and $G : X \rightsquigarrow X \times Y$ defined by $G(x) = (x, F(x))$ for each $x \in X$ be the graph function. Then G is u. D_δ-super c. if and only if F is u. D_δ-super c. and X is D_δ-completely regular space.*

Proof. To prove the necessity, suppose that G is D_δ-super c. By Theorem 6, $F = P_Y \circ G$ is D_δ-super c., where P_Y is the projection from $X \times Y$ onto Y. Let U be any open set in X and let $U \times Y$ be an open set containing $G(x)$. Since G is D_δ-super c., there exists a regular open F_σ-set W containing x such that the implication $x' \in W \Rightarrow G(x') \subset U \times Y$ holds. Thus, $x \in W \subset U$, which shows that U is d_δ-open and so X is a D_δ-completely regular space.

To prove the sufficiency, let $x \in X$ and let W be an open set containing $G(x)$. There exists open sets $U \subset X$ and $V \subset Y$ such that $(x, F(x)) \subset U \times V \subset W$. Since X is D_δ-completely regular space, there exists a regular F_σ-set G_1in X containing x such that $x \in G_1 \subset V$. Since F is D_δ-super c., there exists a cozero set G_2 in X containing x such that the implication $x' \in G_2 \Rightarrow F(x') \subset V$. Let $G_1 \cap G_2 = H$. Then H is a regular F_σ-set containing x and $G(H) \subset U \times V \subset W$ which implies that G is u. D_δ-super c. □

Definition 9. Let $F : X \rightsquigarrow Y$ be a multifunction.

(a) F is said to be upper D_δ-continuous (briefly u. D_δ-c.) at $x \in X$, if for each regular F_σ-set V with $F(x) \subset V$, there exists an open U set containing x such that the implication $x' \in U \Rightarrow F(x') \subset V$ is hold.

(b) F is said to be lower D_δ-continuous (briefly l. D_δ-c.) at $x \in X$, if for each regular F_σ-set V with $F(x) \cap V \neq \emptyset$, there exists an open set U containing x such that the implication $x' \in U \Rightarrow F(x') \cap V \neq \emptyset$ holds.

(c) F is said to be D_δ-continuous (briefly D_δ-c.) at $x \in X$, if it is both u. D_δ-c. and l. D_δ-c. at $x \in X$.

(d) F is said to be u. D_δ-c. (l. D_δ-c., D_δ-c.) on X if it has this property at every point $x \in X$.

Theorem 9. *For a multifunction $F : X \rightsquigarrow Y$, the following statements are equivalent:*

(a) *F is u. D_δ-c. (l. D_δ-c.).*
(b) *For every d_δ-open set $V \subseteq Y$, $F^+(V)$ $(F^-(V))$ is an open set in X.*
(c) *For every d_δ-closed set $K \subseteq Y$, $F^-(K)$ $(F^+(K))$ is a closed set in X.*

Lemma 1. For a multifunction $F : X \rightsquigarrow Y$, the following statements are equivalent:

(a) F is u. D_δ-c.
(b) $F(\overline{A}) \subset [F(A)]_{d_\delta}$ for all $A \subseteq X$
(c) $\overline{F^+(B)} \subseteq F^+([B]_{d_\delta})$ for all $B \subseteq X$
(d) For every d_δ-closed set $K \subseteq Y$, $F^+(K)$ is closed
(e) For every d_δ-open set $G \subseteq Y$, $F^+(G)$ is open

Proof. (a)$\Rightarrow$(b): Let $y \in F(\overline{A})$. Choose $x \in \overline{A}$ such that $y \in F(x)$. Let V be a regular F_σ-set containing $F(x)$ and thus y. Since F is u. D_δ-c., $F^+(V)$ is an open set containing x. This gives $F^+(V) \cap A \neq \emptyset$ which in turn implies that $V \cap F(A) \neq \emptyset$ and consequently $y \in [F(A)]_{d_\delta}$. Hence, $F(\overline{A}) \subset [F(A)]_{d_\delta}$.

(b)$\Rightarrow$(c): Let B be any subset of Y. Then $F(\overline{F^+(B)}) \subseteq [F(F^+(B))]_{d_\delta} \subseteq [B]_{d_\delta}$ and consequently $\overline{F^+(B)} \subseteq F^+([B]_{d_\delta})$.

(c)$\Rightarrow$(d): Since a set K is d_δ-closed if and only if $K = [K]_{d_\delta}$, therefore the implication (c)$\Rightarrow$(d) is obvious.

(d)$\Rightarrow$(e): Obvious.

(e)$\Rightarrow$(a): Since every cozero set is d_δ-open and since a multifunction is u. D_δ-c. if and only if the inverse image of every regular F_σ-set is open. Hence, (e)$\Rightarrow$(a). □

Theorem 10. *Let X, Y and Z be topological spaces and let the function $F : X \rightsquigarrow Y$ be u. D_δ-c. and $G : Y \rightsquigarrow Z$ be u. D_δ-super c. Then $G \circ F : X \rightsquigarrow Z$ is u. s. c.*

Proof. Since $(G \circ F)^+(V) = F^+(G^+(V))$, it is immediate in view of Lemma 1 and Theorem 1. □

Theorem 11. *Let $F : X \rightsquigarrow Y$ be a u. s. c. (l. s. c.) multifunction defined on a D_δ-completely regular space. Then F is u. D_δ-super c. (l. D_δ-super c.).*

Proof. In a D_δ-completely regular space, every open set is d_δ-open. □

Definition 10 ([5])**.** Let $f : X \to Y$ be a surjection from a topological space X onto a set$Y$. The topology on Y for which a subset $A \subset Y$ is open if and only if $f^{-1}(A)$ is d_δ-open in X is called the D_δ-quotient topology and the map f is called the D_δ-quotient map.

Theorem 12. *Let F be a multifunction from a topological space (X, τ_1) onto a topological the space (Y, τ_2), where τ_2 is D_δ-quotient topology on Y. Then F is l. D_δ-super c. Moreover, τ_2 is the finest topology on Y which makes the map $F : X \rightsquigarrow Y$ l. D_δ-super c.*

Proof. The l. D_δ-super continuity of F follows from the definition of the D_δ-quotient topology. □

Theorem 13. *Let $f : X \to Y$ be a D_δ-quotient map. Then a multifunction $F : Y \rightsquigarrow Z$ is l. s. c. if and only if $F \circ f$ is l. D_δ-super c.*

Proof. If U is an open set in Z and $F \circ f$ is l. D_δ-super c. then $(F \circ f)^+(U) = f^+(F^+(U)) = f^{-1}(F^+(U))$ which is d_δ-open in X. Since f is D_δ-quotient map, $F^+(U)$ is open in Y. Thus, F is l. s. c. Conversely, let $F : Y \rightsquigarrow Z$ be u. s. c. Let U be an open set in Z. By the l. D_δ-super continuity of $F \circ f$, $(F \circ f)^+(U) = f^{-1}(F^+(U))$ is d_δ-open in X. □

4. D_δ-complete regularization

In this Section we show that if the domain of upper-lower D_δ-supercontinuous multifunction F is retopologized in an appropriate way, then F is simply a upper-lower semi continuous multifunction.

Let (X, τ) be a topological space and let β denote the collection of all regular F_σ-subsets of (X, τ). Since the intersection of two regular F_σ-sets is a regular F_σ-set, the collection β is a base for a topology τ^* on X called the D_δ-complete regularization of τ. Clearly $\tau^* \subset \tau$. The space (X, τ) is D_δ-completely regular if and only if $\tau^* = \tau$ [5] .

Throughout the section, the symbol τ^* will have the same meaning as in the above paragraph.

Theorem 14. *A multifunction $F : (X, \tau) \rightsquigarrow (Y, \sigma)$ is u. D_δ-super c. if and only if $F : (X, \tau^*) \rightsquigarrow (Y, \sigma)$ is u. s. c.*

Theorem 15. *Let (X, τ) be topological space.Then the following are equivalent.*

(a) *The space (X, τ) is D_δ-completely regular.*
(b) *Every upper-lower semi continuous multifunction from (X, τ) into a space (Y, σ) is upper-lower D_δ-supercontinuous.*

Proof. (a)⇒(b): Obvious.

(b)⇒(a): Take $(Y, \sigma) = (X, \tau)$. Then the identity multifunction I_X on X is upper-lower semi continuous and hence upper-lower D_δ-supercontinuous. Thus, by Theorem 11, $1_X : (X, \tau^*) \to (X, \tau)$ is upper-lower semi continuous.Since $U \in \tau$ implies $1_X^{-1}(U) = U \in \tau^*$, $\tau \subset \tau^*$. Therefore, $\tau = \tau^*$ and so (X, τ) is D_δ-completely regular. □

Theorem 16. *Let $F : (X, \tau) \rightsquigarrow (Y, \sigma)$ be a multifunction. Then F is upper-lower D_δ-continuous multifunction from (X, τ) to a space (Y, σ) if and only if $F : (X, \tau) \rightsquigarrow (Y, \sigma^*)$ is upper-lower semi continuous.*

References

[1] Akdağ, M.: *On the upper and lower super d-continuous multifunctions*, İstanbul Univ., Sciences Faculty, Journal of Math., **60** (2001), 101–109.

[2] Akdağ, M.: *On supercontinuous multifunctions*, Acta Math. Hung., **99** (2003), No. 1–2, 143–153.

[3] Akdağ, M.: *On upper and lower Z-supercontinuous multifunctions*, submitted.

[4] Küçük, Y.: *On strongly θ-continuous multifunctions*, Pure And Appl. Math. Sci., **40** (1994), 43–54.

[5] Kohli, J. K. and Singh, D.: *D_δ-supercontinuous functions*, Indian J. Pure Appl. Math., **34** (2003), No. 7, 1089–1100.

[6] Ponomarev, V. I.: *Properties of topological spaces preserved under multivalued continuous mappings of compacts*, Amer. Math. Soc. Transl., **38** (1964), No. 2, 119–140.

[7] Kohli, J. K. and Kumar, R.: *Z-supercontinuous functions*, Indian J. Pure Appl. Math., **33** (2002), No. 7, 1097–1108.

[8] Mack, J.: *Countable paracompactness and weak normality properties*, Trans. Amer. Math. Soc., **148** (1970), 265–272.

[9] Helderman, N. C.: *Developability and some new regularity axioms*, Can. J. Math., **33** (1981), 641–668.

[10] Hewitt, E.: *On two problems of Urysohn*, Annals of Math., **47** (1946), No. 3, 503–509.

Author's address

Metin Akdağ:

Cumhuriyet University Science and Art Faculty, Department of Mathematics, 58140 Sivas, Turkey

E-mail address: makdag@cumhuriyet.edu.tr

Doi: 10.18514/MMN.2006.137

Miskolc Mathematical Notes
Vol. 7 (2006), No. 1, pp. 13–26

HU ISSN 1787-2413
electronic version

GENERAL QUASILINEARIZATION METHOD FOR SYSTEMS OF DIFFERENTIAL EQUATIONS WITH A SINGULAR MATRIX

T. JANKOWSKI

[Received: August 26, 2005]

ABSTRACT. The method of quasilinearization coupled with the method of lower and upper solutions has been very useful in providing an analytical approach to obtaining approximate solutions of non-linear differential equations. In this paper, it is applied to systems of non-linear differential equations with a singular matrix. Sequences of approximate solutions are convergent to the solution and the convergence is quadratic or semiquadratic.

Mathematics Subject Classification: 34A45

Keywords: quasilinearization, monotone iterations, quadratic and semiquadratic convergence

1. Introduction

Let $y_0, z_0 \in C^1(J, \mathbb{R}^m)$ with $y_0(t) \le z_0(t)$, $y_0'(t) \le z_0'(t)$ on J and define the following set

$$\Omega = \{(t,u,v) : \; y_0(t) \le u \le z_0(t), \quad y_0'(t) \le v \le z_0'(t), \quad t \in J, \; u, v \in \mathbb{R}^m\}.$$

In this paper, the vectorial inequalities mean that the same inequalities hold between their corresponding components.

Assume that A is a singular square matrix of order m and $f \in C(\Omega, \mathbb{R}^m)$. In this paper we shall study the following system of differential equations

$$Ax'(t) = f(t, x(t), x'(t)), \quad t \in J = [0, b] \tag{1.1}$$

with the initial condition

$$x(0) = x_0 \in \mathbb{R}^m. \tag{1.2}$$

The method of quasilinearization offers an approach for obtaining approximate solutions to non-linear differential equations. It has been generalized in recent years by Lakshmikantham and various coauthors to apply to a wide variety of problems, (see, for example, [5–12] and [3, 4]). In this paper, we apply this technique to problems of type (1.1)–(1.2). We show that it is possible to construct monotone sequences that converge to the solution if f is replaced by $f + g$ with $f + \Phi$ convex and $g + \Psi$ concave for some convex function Φ and for some concave function Ψ. This convergence is quadratic or semiquadratic. This paper generalizes the results of [4]. If f does not depend on the third variable with a unit matrix in the place of A, then problem (1.1)–(1.2) is considered in [8].

2. Assumptions

In the place of (1.1)–(1.2), we consider the system of differential equations of the form:

$$\begin{aligned} &Ax'(t) = f(t, x(t), x'(t)) + g(t, x(t), x'(t)) \equiv \mathscr{F}(t, x, x'), \quad t \in J, \\ &x(0) = x_0 \in \mathbb{R}^m, \end{aligned} \tag{2.1}$$

where $J = [0, b]$ and $f, g \in C(J \times \mathbb{R}^m \times \mathbb{R}^m, \mathbb{R}^m)$. Note that problem (2.1) is identical with the problem

$$\begin{aligned} &x'(t) = (A + B)^{-1}[\mathscr{F}(t, x, x') + Bx'(t)], \quad t \in J, \\ &x(0) = x_0 \end{aligned}$$

provided that B is an $m \times m$ matrix such that $(A + B)^{-1}$ exists.

A function $v \in C^1(J, \mathbb{R})$ is said to be a lower solution of problem (2.1) if

$$v'(t) \le (A + B)^{-1}[\mathscr{F}(t, v(t), v'(t)) + Bv'(t)], \quad t \in J, \quad v(0) \le x_0,$$

and an upper solution of (2.1) if the inequalities in these relations are reversed.

Let us introduce the following assumptions:

H_1. There exists a square matrix B of order m such that the matrix $A + B$ is non-singular and $(A + B)^{-1}B \ge 0$; moreover, for $f, g \in C(\Omega, \mathbb{R}^m)$, function $\mathscr{F} = f + g$ satisfies the Lipschitz condition with respect to the last variable, so for $u, \alpha, \bar{\alpha} \in \mathbb{R}^m$ such that $y_0(t) \le u \le z_0(t)$, $y_0'(t) \le \alpha, \bar{\alpha} \le z_0'(t)$ on J, the condition

$$|(A + B)^{-1}[\mathscr{F}(t, u, \alpha) - \mathscr{F}(t, u, \bar{\alpha})]| \le (A + B)^{-1}B|\alpha - \bar{\alpha}|$$

holds, where $|\alpha| = (|\alpha_1|, \ldots, |\alpha_m|)^T$ for $\alpha \in \mathbb{R}^m$.

H_2. $f_x, g_x, \Phi, \Phi_x, \Phi_y, \Psi, \Psi_x, \Psi_y \in C(\Omega, \mathbb{R}^m)$; here x and y denote the second and third variable, respectively.

H_3. The matrices $(A + B)^{-1}F_x$, $(A + B)^{-1}\Phi_x$ are non-decreasing with respect to the second variable, and $(A + B)^{-1}G_x$, $(A + B)^{-1}\Psi_x$ are non-increasing with respect to the second variable on Ω with $F = f + \Phi$, $G = g + \Psi$.

H_4. $(A + B)^{-1}F_x$, $(A + B)^{-1}\Phi_x$ are non-decreasing in the third variable, and $(A + B)^{-1}G_x$, $(A + B)^{-1}\Psi_x$ are non-increasing in the third variable on Ω.

H_5. $(A + B)^{-1}V(t, y_0, z_0) \ge 0$, $t \in J$ for some function V defined later [$(A + B)^{-1}V(t, y_0, z_0) \ge 0$ means that the entries of the matrix $(A + B)^{-1}V(t, y_0, z_0)$ are non-negative].

H_6. There exist $m\times m$ matrices C_1, C_2, C_3, C_4 with non-negative entries such that

$$\left|(A+B)^{-1}\left[f_x(t,u,v)-f_x(t,\bar{u},\bar{v})\right]\right| \le C_1 \sum_{i=1}^{m}\left[|u_i-\bar{u}_i|+|v_i-\bar{v}_i|\right],$$

$$\left|(A+B)^{-1}\left[g_x(t,u,v)-g_x(t,\bar{u},\bar{v})\right]\right| \le C_2 \sum_{i=1}^{m}\left[|u_i-\bar{u}_i|+|v_i-\bar{v}_i|\right],$$

$$\left|(A+B)^{-1}\left[\Phi_x(t,u,v)-\Phi_x(t,\bar{u},\bar{v})\right]\right| \le C_3 \sum_{i=1}^{m}\left[|u_i-\bar{u}_i|+|v_i-\bar{v}_i|\right],$$

$$\left|(A+B)^{-1}\left[\Psi_x(t,u,v)-\Psi_x(t,\bar{u},\bar{v})\right]\right| \le C_4 \sum_{i=1}^{m}\left[|u_i-\bar{u}_i|+|v_i-\bar{v}_i|\right]$$

for $y_0(t)\le u\le z_0(t)$, $y_0'(t)\le v,\bar{v}\le z_0'(t)$, $t\in J$ with $u,\bar{u},v,\bar{v}\in\mathbb{R}^m$.

3. Main results

The next lemma is a special case of Theorem 1.1.4 from [8].

Lemma 1. *Assume that $s_{ij}(t)\ge 0$, $t\in J$ for $i\ne j$, where $S=[s_{ij}]$ is a continuous square matrix of order m. Let $p\in C^1(J,\mathbb{R}^m)$ and*

$$p'(t)\le S(t)p(t),\quad t\in J,$$
$$p(0)\le 0=[\underbrace{0,\ldots,0}_{m}]^T.$$

Then $p(t)\le 0$ on J.

Lemma 2. *Let assumptions H_1 and H_3 be satisfied. Then, for $u,v,\bar{u},\bar{v}\in\mathbb{R}^m$ such that $y_0(t)\le u\le\bar{u}\le z_0(t)$, $y_0'(t)\le v\le\bar{v}\le z_0'(t)$, $t\in J$, we have*

$$\begin{aligned}(A+B)^{-1}[\mathscr{F}(t,u,v)-\mathscr{F}(t,\bar{u},\bar{v})] \le (A+B)^{-1}\{&[-F_x(t,u,v)-G_x(t,\bar{u},v)\\ &+\Phi_x(t,\bar{u},v)+\Psi_x(t,u,v)](\bar{u}-u)+B(\bar{v}-v)\}.\end{aligned}$$

Proof. The mean value theorem and assumption H_1 yield

$$\begin{aligned}&(A+B)^{-1}[\mathscr{F}(t,u,v)-\mathscr{F}(t,\bar{u},\bar{v})]\\
&\quad=(A+B)^{-1}[\mathscr{F}(t,u,v)-\mathscr{F}(t,\bar{u},v)+\mathscr{F}(t,\bar{u},v)-\mathscr{F}(t,\bar{u},\bar{v})]\\
&\quad\le(A+B)^{-1}\left\{\left[\int_0^1\mathscr{F}_x(t,su+(1-s)\bar{u},v)ds\right](u-\bar{u})+B(\bar{v}-v)\right\}\\
&\quad=(A+B)^{-1}\left\{\int_0^1[F_x(t,su+(1-s)\bar{u},v)+G_x(t,su+(1-s)\bar{u},v)\right.\\
&\qquad\left.-\Phi_x(t,su+(1-s)\bar{u},v)-\Psi_x(t,su+(1-s)\bar{u},v)]ds(u-\bar{u})+B(\bar{v}-v)\right\}.\end{aligned}$$

Hence, we have the assertion of Lemma 2, by using assumption H_3. □

Now we are in a position to prove the following result:

Theorem 1. *Assume that* $f, g \in C(\Omega, \mathbb{R}^m)$, *and*

(i) $y_0, z_0 \in C^1(J, \mathbb{R}^m)$ *are lower and upper solutions of problem* (2.1) *and such that* $y_0(t) \le z_0(t)$ *and* $y_0'(t) \le z_0'(t)$ *on* J,

(ii) *Assumptions* H_1–H_6 *hold with*

$$V(t, y, z) = F_x(t, y, y') + G_x(t, z, z') - \Phi_x(t, z, z') - \Psi_x(t, y, y').$$

(iii) *Problem* (2.1) *has at most one solution.*

Then, there exist monotone sequences $\{y_n\}$, $\{z_n\}$ *which converge uniformly on* J *to the unique solution* x *of problem* (2.1). *Moreover, the convergence is quadratic with respect to* u *and it is semiquadratic with respect to* u' *for* $u = y_n$ *and* $u = z_n$.

Proof. Let y_{n+1} and z_{n+1} be the solutions of the linear initial value problems

$$\begin{aligned} &y_{n+1}'(t) = (A+B)^{-1}\{\mathscr{F}(t, y_n, y_n') + By_n'(t) + V(t, y_n, z_n)[y_{n+1}(t) - y_n(t)]\},\\ &y_{n+1}(0) = x_0, \end{aligned}$$

and

$$\begin{aligned} &z_{n+1}'(t) = (A+B)^{-1}\{\mathscr{F}(t, z_n, z_n') + Bz_n'(t) + V(t, y_n, z_n)[z_{n+1}(t) - z_n(t)]\},\\ &z_{n+1}(0) = x_0, \end{aligned}$$

for $n = 0, 1, \ldots$. Note that the sequences $\{y_n\}, \{z_n\}$ are well defined.

First of all, we shall prove that

$$\begin{aligned} &y_0(t) \le y_1(t) \le z_1(t) \le z_0(t), \quad t \in J,\\ &y_0'(t) \le y_1'(t) \le z_1'(t) \le z_0'(t), \quad t \in J. \end{aligned} \tag{3.1}$$

Let us put $p = y_0 - y_1$, so $p(0) \le 0$. Then we see that

$$\begin{aligned} p'(t) \le (A+B)^{-1}\{&\mathscr{F}(t, y_0, y_0') + By_0'(t) - \mathscr{F}(t, y_0, y_0') - By_0'(t)\\ &- V(t, y_0, z_0)[y_1(t) - y_0(t)]\} = (A+B)^{-1}V(t, y_0, z_0)p(t), \quad t \in J. \end{aligned}$$

Assumption H_5 and Lemma 1 yield $p(t) \le 0$ on J proving that $y_0(t) \le y_1(t)$ on J. Since $(A+B)^{-1}V(t, y_0, z_0) \ge 0$, and $p(t) \le 0$ on J, then $p'(t) \le 0$, so $y_0'(t) \le y_1'(t)$ on J. By the same way we can show that $z_1(t) \le z_0(t)$ and $z_1'(t) \le z_0'(t)$, $t \in J$. Put

$p = y_1 - z_1$. Then, by Lemma 2 and assumption H_4, we have

$$\begin{aligned}
p'(t) &= (A+B)^{-1}\{\mathscr{F}(t,y_0,y_0') - \mathscr{F}(t,z_0,z_0') + B[y_0'(t) - z_0'(t)] \\
&\quad + V(t,y_0,z_0)[y_1(t) - y_0(t) - z_1(t) + z_0(t)]\} \\
&\le (A+B)^{-1}\{[-F_x(t,y_0,y_0') - G_x(t,z_0,y_0') + \Phi_x(t,z_0,y_0') \\
&\quad + \Psi_x(t,y_0,y_0')][z_0(t) - y_0(t)] + B[z_0'(t) - y_0'(t)] \\
&\quad + V(t,y_0,z_0)[y_1(t) - y_0(t) - z_1(t) + z_0(t)] + B[y_0'(t) - z_0'(t)]\} \\
&= (A+B)^{-1}\{[G_x(t,z_0,z_0') - G_x(t,z_0,y_0') + \Phi_x(t,z_0,y_0') \\
&\quad - \Phi_x(t,z_0,z_0')][z_0(t) - y_0(t)] + V(t,y_0,z_0)p(t)\} \\
&\le (A+B)^{-1}V(t,y_0,z_0)p(t)
\end{aligned}$$

with $p(0) = 0$. Hence, we have $p(t) \le 0$, and then $p'(t) \le 0$ on J which shows that $y_1(t) \le z_1(t)$, $y_1'(t) \le z_1'(t)$, $t \in J$. This means that (3.1) holds.

In the next step we need to show that y_1 and z_1 are lower and upper solutions of problem (2.1), respectively. By Lemma 2 and assumptions H_3 and H_4, we obtain

$$\begin{aligned}
y_1'(t) &= (A+B)^{-1}\{\mathscr{F}(t,y_0,y_0') + By_0'(t) + V(t,y_0,z_0)[y_1(t) - y_0(t)]\} \\
&\le (A+B)^{-1}\{\mathscr{F}(t,y_1,y_1') + By_1'(t) + [-F_x(t,y_0,y_0') - G_x(t,y_1,y_0') \\
&\quad + \Phi_x(t,y_1,y_0') + \Psi_x(t,y_0,y_0')][y_1(t) - y_0(t)] + V(t,y_0,z_0)[y_1(t) - y_0(t)]\} \\
&= (A+B)^{-1}\{\mathscr{F}(t,y_1,y_1') + By_1'(t) + [G_x(t,z_0,z_0') - G_x(t,y_1,y_0') \\
&\quad + \Phi_x(t,y_1,y_0') - \Phi_x(t,z_0,z_0')][y_1(t) - y_0(t)]\} \\
&\le (A+B)^{-1}[\mathscr{F}(t,y_1,y_1') + By_1'(t)],
\end{aligned}$$

and

$$\begin{aligned}
z_1'(t) &= (A+B)^{-1}\{\mathscr{F}(t,z_0,z_0') + Bz_0'(t) + V(t,y_0,z_0)[z_1(t) - z_0(t)]\} \\
&\ge (A+B)^{-1}\{\mathscr{F}(t,z_1,z_1') + Bz_1'(t) + [F_x(t,z_1,z_1') + G_x(t,z_0,z_1') \\
&\quad - \Phi_x(t,z_0,z_1') - \Psi_x(t,z_1,z_1')][z_0(t) - z_1(t)] + V(t,y_0,z_0)[z_1(t) - z_0(t)]\} \\
&= (A+B)^{-1}\{\mathscr{F}(t,z_1,z_1') + Bz_1'(t) + [F_x(t,z_1,z_1') - F_x(t,y_0,y_0') \\
&\quad + G_x(t,z_0,z_1') - G_x(t,z_0,z_0') + \Phi_x(t,z_0,z_0') - \Phi_x(t,z_0,z_1') \\
&\quad + \Psi_x(t,y_0,y_0') - \Psi_x(t,z_1,z_1')][z_0(t) - z_1(t)]\} \\
&\ge (A+B)^{-1}[\mathscr{F}(t,z_1,z_1') + Bz_1'(t)]
\end{aligned}$$

which shows that y_1 and z_1, respectively, are lower and upper solutions of problem (2.1). Let us assume that

$$\begin{aligned}
&y_{k-1}(t) \le y_k(t) \le z_k(t) \le z_{k-1}(t), \quad t \in J, \\
&y_{k-1}'(t) \le y_k'(t) \le z_k'(t) \le z_{k-1}'(t), \quad t \in J,
\end{aligned}$$

and let y_k, z_k be lower and upper solutions of problem (2.1) for some $k \geq 1$. We shall prove that

$$\begin{aligned} y_k(t) \leq y_{k+1}(t) \leq z_{k+1}(t) \leq z_k(t), \quad t \in J, \\ y'_k(t) \leq y'_{k+1}(t) \leq z'_{k+1}(t) \leq z'_k(t), \quad t \in J. \end{aligned} \tag{3.2}$$

Put $p = y_k - y_{k+1}$. Then

$$\begin{aligned} p'(t) \leq (A+B)^{-1}\{&\mathscr{F}(t, y_k, y'_k) + By'_k(t) - \mathscr{F}(t, y_k, y'_k) - By'_k(t) \\ &- V(t, y_k, z_k)[y_{k+1}(t) - y_k(t)]\} = (A+B)^{-1}V(t, y_k, z_k)p(t) \end{aligned}$$

with $p(0) = 0$. Note that, by assumptions H_3–H_5,

$$\begin{aligned} (A+B)^{-1}V(t, y_k, z_k) &= (A+B)^{-1}[F_x(t, y_k, y'_k) + G_x(t, z_k, z'_k) - \Phi_x(t, z_k, z'_k) \\ &\quad - \Psi_x(t, y_k, y'_k)] \\ &\geq (A+B)^{-1}[F_x(t, y_0, y'_0) + G_x(t, z_0, z'_0) - \Phi_x(t, z_0, z'_0) \\ &\quad - \Psi_x(t, y_0, y'_0)] \\ &= (A+B)^{-1}V(t, y_0, z_0) \geq 0, \quad t \in J. \end{aligned}$$

Hence, by Lemma 1, $p(t) \leq 0$, $p'(t) \leq 0$, $t \in J$, which shows that $y_k(t) \leq y_{k+1}(t)$ and $y'_k(t) \leq y'_{k+1}(t)$, $t \in J$. Using the same argument we can prove that $z_{k+1}(t) \leq z_k(t)$, $z'_{k+1}(t) \leq z'_k(t)$, $t \in J$.

Let $p = y_{k+1} - z_{k+1}$. Then $p(0) = 0$. Using Lemma 2 and assumption H_4, we get

$$\begin{aligned} p'(t) &= (A+B)^{-1}\{\mathscr{F}(t, y_k, y'_k) - \mathscr{F}(t, z_k, z'_k) + B[y'_k(t) - z'_k(t)] \\ &\quad + V(t, y_k, z_k)[y_{k+1}(t) - y_k(t) - z_{k+1}(t) + z_k(t)]\} \\ &\leq (A+B)^{-1}\{[-F_x(t, y_k, y'_k) - G_x(t, z_k, y'_k) + \Phi_x(t, z_k, y'_k) \\ &\quad + \Psi_x(t, y_k, y'_k)][z_k(t) - y_k(t)] \\ &\quad + V(t, y_k, z_k)[y_{k+1}(t) - y_k(t) - z_{k+1}(t) + z_k(t)]\} \\ &= (A+B)^{-1}\{[G_x(t, z_k, z'_k) - G_x(t, z_k, y'_k) + \Phi_x(t, z_k, y'_k) \\ &\quad - \Phi_x(t, z_k, z'_k)][z_k(t) - y_k(t)] + V(t, y_k, z_k)p(t)\} \\ &\leq (A+B)^{-1}V(t, y_k, z_k)p(t), \quad t \in J. \end{aligned}$$

This proves that $y_{k+1}(t) \leq z_{k+1}(t)$, and $y'_{k+1}(t) \leq z'_{k+1}(t)$, $t \in J$, so relation (3.2) holds. Hence, by induction, for all n, we have

$$\begin{aligned} y_0(t) \leq y_1(t) \leq \cdots \leq y_n(t) \leq z_n(t) \leq \cdots \leq z_1(t) \leq z_0(t), \quad t \in J, \\ y'_0(t) \leq y'_1(t) \leq \cdots \leq y'_n(t) \leq z'_n(t) \leq \cdots \leq z'_1(t) \leq z'_0(t), \quad t \in J. \end{aligned}$$

Employing standard techniques (using the Arzeli theorem and the Lebesgue theorem), it can be shown that $y_n \to y$, $y'_n \to y'$, $z_n \to z$, $z'_n \to z'$, $y, z \in C^1(J, \mathbb{R}^m)$, where y and z are solutions of problem (2.1). Hence, by assumption (iii), we have $y = z = x$ on J is the unique solution of (2.1).

The order of convergence of sequences $\{y_n\}$, $\{z_n\}$, $\{y_n'\}$, $\{z_n'\}$ is considered in the next part of our considerations. For this purpose, we put

$$p_{n+1} = x - y_{n+1} \geq 0, \qquad q_{n+1} = z_{n+1} - x \geq 0 \quad \text{on} \quad J,$$

and note that $p_{n+1}(0) = q_{n+1}(0) = 0$ for $n \geq 0$. Using the integral mean value theorem and assumptions H_1, H_3, H_6, we get

$$\begin{aligned}
p_{n+1}'(t) &= (A+B)^{-1}\{\mathscr{F}(t,x,x') + Bx'(t) - \mathscr{F}(t,y_n,x') + \mathscr{F}(t,y_n,x') \\
&\quad - \mathscr{F}(t,y_n,y_n') - V(t,y_n,z_n)[y_{n+1}(t) - x(t) + x(t) - y_n(t)] - By_n'(t)\} \\
&\leq (A+B)^{-1}\left\{\left[\int_0^1 \mathscr{F}_x(t,sx+(1-s)y_n,x')ds\right]p_n(t) + 2B|p_n'(t)| \right. \\
&\quad \left. + V(t,y_n,z_n)[p_{n+1}(t) - p_n(t)]\right\} \\
&= (A+B)^{-1}\left\{\int_0^1 [F_x(t,sx+(1-s)y_n,x') + G_x(t,sx+(1-s)y_n,x') \right. \\
&\quad - \Phi_x(t,sx+(1-s)y_n,x') - \Psi_x(t,sx+(1-s)y_n,x')]ds\ p_n(t) \\
&\quad \left. + 2B|p_n'(t)| + V(t,y_n,z_n)[p_{n+1}(t) - p_n(t)]\right\} \\
&\leq (A+B)^{-1}\{[F_x(t,x,x') - F_x(t,y_n,y_n') + G_x(t,y_n,x') - G_x(t,z_n,z_n') \\
&\quad + \Phi_x(t,z_n,z_n') - \Phi_x(t,y_n,x') + \Psi_x(t,y_n,y_n') - \Psi_x(t,x,x')]p_n(t) \\
&\quad + V(t,y_n,z_n)p_{n+1}(t) + 2B|p_n'(t)|\} \\
&\leq \left\{(C_1 + C_2 + 2C_3 + 2C_4)\sum_{i=1}^m p_{ni}(t)\right. \\
&\quad + (C_2 + C_3 + C_4)\sum_{i=1}^m [q_{ni}(t) + |q_{ni}'(t)|] \\
&\quad \left. +(C_1 + C_3 + C_4)\sum_{i=1}^m |p_{ni}'(t)|\right\} p_n(t) \\
&\quad + (A+B)^{-1}\left\{2B|p_n'(t)| + V(t,y_n,z_n)p_{n+1}(t)\right\}.
\end{aligned}$$

Note that

$$\begin{aligned}
\sum_{i=1}^m p_{ni}(t)p_n(t) &\leq \frac{m}{2}p_n^2(t) + \frac{1}{2}Wp_n^2(t), \\
\sum_{i=1}^m q_{ni}(t)p_n(t) &\leq \frac{m}{2}p_n^2(t) + \frac{1}{2}Wq_n^2(t),
\end{aligned} \tag{3.3}$$

where $p_n^2 = [p_{1,n}^2, \dots, p_{m,n}^2]^T$, $W = [w_{ij}]$, $w_{ij} = 1$, $i, j = 1, \dots, m$. This and previous calculations give

$$p'_{n+1}(t) \le Kp_{n+1}(t) + A_1 p_n^2(t) + A_2 q_n^2(t) + A_3|p'_n(t)|^2 + A_4|q'_n(t)|^2 + A_5|p'_n(t)| \quad (3.4)$$

with $(A+B)^{-1} f_x \le K_1$, $(A+B)^{-1} g_x \le K_2$, $K = K_1 + K_2$ on Ω. Here, K_1, K_2 are $m \times m$ non-negative matrices and

$$\begin{aligned} A_1 &= \frac{1}{2}(C_1 + C_2 + 2C_3 + 2C_4)(mI + W) + (C_2 + C_3 + C_4)m \\ &\quad + (C_1 + C_3 + C_4)\frac{m}{2}, \\ A_2 &= \frac{1}{2}(C_2 + C_3 + C_4)W, \\ A_3 &= \frac{1}{2}(C_1 + C_3 + C_4)W, \\ A_4 &= A_2, \\ A_5 &= 2(A+B)^{-1}B. \end{aligned}$$

There is no loss of generality assuming that K^{-1} exists such that $k_{ij} \ge 0$, where k_{ij} represents the components of this matrix. Hence, for $t \in J$, we have

$$p_{n+1}(t) \le \int_0^t e^{K(t-s)} \Big[A_1 p_n^2(s) + A_2 q_n^2(s) + A_3|p'_n(s)|^2 + A_4|q'_n(s)|^2 + A_5|p'_n(s)|\Big] ds.$$

This implies

$$\begin{aligned} \max_{t\in J} \|p_{n+1}(t)\| \le B_1 \max_{t\in J} \|p_n(t)\|^2 + B_2 \max_{t\in J} \|q_n(t)\|^2 + B_3 \max_{t\in J} \|p'_n(t)\|^2 \\ + B_4 \max_{t\in J} \|q'_n(t)\|^2 + B_5 \max_{t\in J} \|p'_n(t)\|, \quad (3.5) \end{aligned}$$

where $\|v\|^2 = \big[|v_1|^2, \dots, |v_m|^2\big]^T$, $v \in \mathbb{R}^m$, and

$$A_0 = K^{-1} e^{Kb}, \qquad B_i = A_0 A_i,$$

for $i = \overline{1,5}$. Combining (3.4) and (3.5) we obtain

$$\begin{aligned} \max_{t\in J} \|p'_{n+1}(t)\| \le \bar{A}_1 \max_{t\in J} \|p_n(t)\|^2 + \bar{A}_2 \max_{t\in J} \|q_n(t)\|^2 + \bar{A}_3 \max_{t\in J} \|p'_n(t)\|^2 \\ + \bar{A}_4 \max_{t\in J} \|q'_n(t)\|^2 + \bar{A}_5 \max_{t\in J} \|p'_n(t)\| \end{aligned}$$

with $\bar{A}_i = A_i + KB_i$, $i = \overline{1,5}$.

Similarly we have

$$
\begin{aligned}
q'_{n+1}(t) &= (A+B)^{-1}\{\mathscr{F}(t,z_n,z'_n) + Bz'_n(t) - \mathscr{F}(t,x,z'_n) \\
&\quad + \mathscr{F}(t,x,z'_n) - \mathscr{F}(t,x,x') + V(t,y_n,z_n)[q_{n+1}(t) - q_n(t)] - Bx'(t)\} \\
&\le (A+B)^{-1}\left\{\left[\int_0^1 \mathscr{F}_x(t, sz_n + (1-s)x, z'_n)ds\right] q_n(t) + 2B|q'_n(t)| \right. \\
&\quad \left. + V(t,y_n,z_n)[q_{n+1}(t) - q_n(t)]\right\} \\
&\le (A+B)^{-1}\{[F_x(t,z_n,z'_n) - F_x(t,y_n,y'_n) + G_x(t,x,z'_n) - G_x(t,z_n,z'_n) \\
&\quad + \Phi_x(t,z_n,z'_n) - \Phi_x(t,x,z'_n) + \Psi_x(t,y_n,y'_n) - \Psi_x(t,z_n,z'_n)]q_n(t) \\
&\quad + V(t,y_n,z_n)q_{n+1}(t) + 2B|q'_n(t)|\} \\
&\le \left\{(C_1 + C_2 + 2C_3 + 2C_4)\sum_{i=1}^m q_{ni}(t) \right. \\
&\quad \left. + (C_1 + C_3 + C_4)\sum_{i=1}^m [p_{ni}(t) + |p'_{ni}(t)| + |q'_{ni}(t)|]\right\} q_n(t) \\
&\quad + Kq_{n+1}(t) + A_5|q'_n(t)| \\
&\le D_1 p_n^2(t) + D_2 q_n^2(t) + D_1|p'_n(t)|^2 + D_1|q'_n(t)|^2 + Kq_{n+1}(t) + A_5|q'_n(t)|,
\end{aligned}
$$

where

$$
\begin{aligned}
D_1 &= \frac{1}{2}(C_1 + C_3 + C_4)W, \\
D_2 &= \frac{3}{2}m(C_1 + C_3 + C_4) + \frac{1}{2}(C_1 + C_2 + 2C_3 + 2C_4)(mI + W).
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\max_{t\in J}\|q_{n+1}(t)\| \le \overline{B}_1 \max_{t\in J}\|p_n(t)\|^2 + \overline{B}_2 \max_{t\in J}\|q_n(t)\|^2 + \overline{B}_1 \max_{t\in J}\|p'_n(t)\|^2 \\
+ \overline{B}_1 \max_{t\in J}\|q'_n(t)\|^2 + \overline{B}_3 \max_{t\in J}\|q'_n(t)\|,
\end{aligned}
$$

where $\overline{B}_1 = A_0 D_1$, $\overline{B}_2 = A_0 D_2$, and $\overline{B}_3 = B_5 A$.

Combining this with the last relation for q'_{n+1} we get

$$
\begin{aligned}
\max_{t\in J}\|q'_{n+1}(t)\| \le \overline{L}_1 \max_{t\in J}\|p_n(t)\|^2 + \overline{L}_2 \max_{t\in J}\|q_n(t)\|^2 + \overline{L}_1 \max_{t\in J}\|p'_n(t)\|^2 \\
+ \overline{L}_1 \max_{t\in J}\|q'_n(t)\|^2 + \overline{L}_3 \max_{t\in J}\|q'_n(t)\|,
\end{aligned}
$$

with $\overline{L}_1 = D_1 + K\overline{B}_1$, $\overline{L}_2 = D_2 + K\overline{B}_2$, and $\overline{L}_3 = A_5 + K\overline{B}_3$. This completes the proof. □

Let us introduce the following assumptions:

H_{17} (i) $(A+B)^{-1}F_x$ is non-decreasing in the third variable on Ω and $V_1 = F_x(t,y,y')$, or
(ii) $(A+B)^{-1}F_x$ is non-increasing in the third variable on Ω and $V_1 = F_x(t,y,z')$.

H_{27} (i) $(A+B)^{-1}G_x$ is non-increasing in the third variable on Ω and $V_2 = G_x(t,z,z')$, or
(ii) $(A+B)^{-1}G_x$ is non-decreasing in the third variable on Ω and $V_2 = G_x(t,z,y')$.

H_{37} (i) $(A+B)^{-1}\Phi_x$ is non-decreasing in the third variable on Ω and $V_3 = \Phi_x(t,z,z')$, or
(ii) $(A+B)^{-1}\Phi_x$ is non-increasing in the third variable on Ω and $V_3 = \Phi_x(t,z,y')$.

H_{47} (i) $(A+B)^{-1}\Psi_x$ is non-increasing in the third variable on Ω and $V_4 = \Psi_x(t,y,y')$, or
(ii) $(A+B)^{-1}\Psi_x$ is non-decreasing in the third variable on Ω and $V_4 = \Psi_x(t,y,z')$.

The set of all assumptions from H_{17} to H_{47} will be denoted by H_7. Since in any assumptions H_{17}–H_{47} we have two cases (i) or (ii), so we have 16 possibilities for constructing assumption H_7. Note that if we choose case (i) in any assumptions H_{17}–H_{47}, then assumption H_7 is identical with assumption H_4.

Now we can formulate the following

Theorem 2. *Assume that the assumptions of Theorem 1 are satisfied with assumption H_7 instead of H_4 and for*

$$V = V_1 + V_2 - V_3 - V_4.$$

Then the conclusion of Theorem 1 is true.

Proof. Since the proof can be constructed on the basis of the proof of the previous theorem, we shall only indicate the necessary changes. We should create assumption H_7. Let H_7 be produced from assumptions H_{17}(ii), H_{27}(ii), H_{37}(ii), and H_{47}(ii). Note that the sequences $\{y_n\}$, $\{z_n\}$ are constructed as before with

$$V(t,y,z) = F_x(t,y,z') + G_x(t,z,y') - \Phi_x(t,z,y') - \Psi_x(t,y,z').$$

Based on the assumption

$$(A+B)^{-1}V(t,y_0,z_0) \geq 0$$

and Lemma 1, it is quite easy to show that $y_0(t) \leq y_1(t)$, $y_0'(t) \leq y_1'(t)$, $z_1(t) \leq z_0(t)$ and $z_1'(t) \leq z_0'(t)$ on J. If we put $p = y_1 - z_1$, then, by Lemma 2 and assumptions

H_{17}(ii), H_{47}(ii), we have

$$\begin{aligned}
p'(t) &\le (A+B)^{-1}\{[-F_x(t,y_0,y_0') - G_x(t,z_0,y_0') + \Phi_x(t,z_0,y_0') \\
&\quad + \Psi_x(t,y_0,y_0')][z_0(t)-y_0(t)] + B[z_0'(t)-y_0'(t)] \\
&\quad + V(t,y_0,z_0)[y_1(t)-y_0(t)-z_1(t)+z_0(t)] + B[y_0'(t)-z_0'(t)]\} \\
&= (A+B)^{-1}\{[F_x(t,y_0,z_0') - F_x(t,y_0,y_0') + \Psi_x(t,y_0,y_0') \\
&\quad - \Psi_x(t,y_0,z_0')][z_0(t)-y_0(t)] \\
&\quad + V(t,y_0,z_0)p(t)\} \le (A+B)^{-1}V(t,y_0,z_0)p(t), \\
p(0) &= 0.
\end{aligned}$$

Hence, by Lemma 1, we have $p(t) \le 0$, and therefore $p'(t) \le 0$ on J which shows that $y_1(t) \le z_1(t)$, $y_1'(t) \le z_1'(t)$, $t \in J$. It means that (3.1) holds.

In the next step we need to show that y_1 and z_1 are lower and upper solutions of problem (2.1), respectively. Note that, using Lemma 2 and assumptions H_3 and H_7, we get

$$\begin{aligned}
y_1'(t) &\le (A+B)^{-1}\{\mathscr{F}(t,y_1,y_1') + By_1'(t) + [-F_x(t,y_0,y_0') - G_x(t,y_1,y_0') \\
&\quad + \Phi_x(t,y_1,y_0') + \Psi_x(t,y_0,y_0')][y_1(t)-y_0(t)] + V(t,y_0,z_0)[y_1(t)-y_0(t)]\} \\
&= (A+B)^{-1}\{\mathscr{F}(t,y_1,y_1') + By_1'(t) + [F_x(t,y_0,z_0') - F_x(t,y_0,y_0') \\
&\quad + G_x(t,z_0,y_0') - G_x(t,y_1,y_0') + \Phi_x(t,y_1,y_0') - \Phi_x(t,z_0,y_0') + \Psi_x(t,y_0,y_0') \\
&\quad - \Psi_x(t,y_0,z_0')][y_1(t)-y_0(t)]\} \le (A+B)^{-1}[\mathscr{F}(t,y_1,y_1') + By_1'(t)],
\end{aligned}$$

and

$$\begin{aligned}
z_1'(t) &\ge (A+B)^{-1}\{\mathscr{F}(t,z_1,z_1') + Bz_1'(t) + [F_x(t,z_1,z_1') + G_x(t,z_0,z_1') \\
&\quad - \Phi_x(t,z_0,z_1') - \Psi_x(t,z_1,z_1')][z_0(t)-z_1(t)] + V(t,y_0,z_0)[z_1(t)-z_0(t)]\} \\
&= (A+B)^{-1}\{\mathscr{F}(t,z_1,z_1') + Bz_1'(t) + [F_x(t,z_1,z_1') - F_x(t,y_0,z_0') \\
&\quad + G_x(t,z_0,z_1') - G_x(t,z_0,y_0') + \Phi_x(t,z_0,y_0') - \Phi_x(t,z_0,z_1') \\
&\quad + \Psi_x(t,y_0,z_0') - \Psi_x(t,z_1,z_1')][z_0(t)-z_1(t)]\} \\
&\ge (A+B)^{-1}[\mathscr{F}(t,z_1,z_1') + Bz_1'(t)],
\end{aligned}$$

which shows that y_1 and z_1 are lower and upper solutions of problem (2.1), respectively.

By induction in n, we can show that

$$\begin{aligned}
&y_0(t) \le y_1(t) \le \cdots \le y_n(t) \le z_n(t) \le \cdots \le z_1(t) \le z_0(t), \quad t \in J, \\
&y_0'(t) \le y_1'(t) \le \cdots \le y_n'(t) \le z_n'(t) \le \cdots \le z_1'(t) \le z_0'(t), \quad t \in J
\end{aligned}$$

for all n.

Employing standard techniques, it is easy to conclude that $y_n \to y$, $y_n' \to y'$, $z_n \to z$, $z_n' \to z'$, $y, z \in C^1(J, \mathbb{R}^m)$, where y and z are solutions of problem (2.1). Hence, by assumption (iii), we have $y = z = x$ on J is the unique solution of (2.1).

To show the quadratic and semiquadratic convergence, we set

$$p_{n+1} = x - y_{n+1} \geq 0, \qquad q_{n+1} = z_{n+1} - x \geq 0$$

on J. Note that $p_{n+1}(0) = q_{n+1}(0) = 0$ for $n \geq 0$. The beginning for p_{n+1}' is the same as in the proof of Theorem 1, so

$$\begin{aligned} p_{n+1}'(t) \leq (A+B)^{-1}\bigg\{ \int_0^1 & [F_x(t, sx + (1-s)y_n, x') + G_x(t, sx + (1-s)y_n, x') \\ & - \Phi_x(t, sx + (1-s)y_n, x') - \Psi_x(t, sx + (1-s)y_n, x')]ds \ p_n(t) \\ & + 2B|p_n'(t)| + V(t, y_n, z_n)[p_{n+1}(t) - p_n(t)] \bigg\}. \end{aligned}$$

Now, using the same argument as in the proof of Theorem 1, we can prove that

$$\begin{aligned} \max_{t\in J} \|p_{n+1}(t)\| \leq \alpha_1 \max_{t\in J} \|p_n(t)\|^2 + \alpha_2 \max_{t\in J} \|q_n(t)\|^2 + \alpha_3 \max_{t\in J} \|p_n'(t)\|^2 \\ + \alpha_4 \max_{t\in J} \|q_n'(t)\|^2 + \alpha_5 \max_{t\in J} \|p_n'(t)\| \end{aligned}$$

and

$$\begin{aligned} \max_{t\in J} \|p_{n+1}'(t)\| \leq \bar{\alpha}_1 \max_{t\in J} \|p_n(t)\|^2 + \bar{\alpha}_2 \max_{t\in J} \|q_n(t)\|^2 + \bar{\alpha}_3 \max_{t\in J} \|p_n'(t)\|^2 \\ + \bar{\alpha}_4 \max_{t\in J} \|q_n'(t)\|^2 + \bar{\alpha}_5 \max_{t\in J} \|p_n'(t)\| \end{aligned}$$

where

$$\begin{aligned} \delta_1 &= \frac{1}{2}C_1(2mI + W) + \frac{1}{2}C_2(3mI + W) + \frac{1}{2}(C_3 + C_4)(5mI + 2W), \\ \delta_2 &= \frac{1}{2}(C_2 + C_3 + C_4)W, \\ \delta_3 &= \frac{1}{2}(C_3 + C_4)W, \\ \delta_4 &= \frac{1}{2}(C_1 + C_2 + C_3 + C_4)W, \\ \delta_5 &= A_5, \end{aligned}$$

and $\alpha_i = A_0\delta_i$, $\bar{\alpha}_i = \delta_i + K\alpha_i$, $i = \overline{1,5}$, with A_0 and K defined as in the proof of Theorem 1.

Similarly, we can show that

$$\max_{t\in J}\|q_{n+1}(t)\| \le \beta_1 \max_{t\in J}\|p_n(t)\|^2 + \beta_2 \max_{t\in J}\|q_n(t)\|^2 + \beta_3 \max_{t\in J}\|p_n'(t)\|^2 + \beta_4 \max_{t\in J}\|q_n'(t)\|^2 + \beta_5 \max_{t\in J}\|q_n'(t)\|$$

and

$$\max_{t\in J}\|q_{n+1}'(t)\| \le \overline{\beta}_1 \max_{t\in J}\|p_n(t)\|^2 + \overline{\beta}_2 \max_{t\in J}\|q_n(t)\|^2 + \overline{\beta}_3 \max_{t\in J}\|p_n'(t)\|^2 + \overline{\beta}_4 \max_{t\in J}\|q_n'(t)\|^2 + \overline{\beta}_5 \max_{t\in J}\|q_n'(t)\|,$$

with

$$\begin{aligned}
\eta_1 &= \frac{1}{2}(C_1 + 2C_3 + C_4)W,\\
\eta_3 &= \eta_4 = \frac{1}{2}(C_2 + C_3 + C_4)W,\\
\eta_5 &= A_5,\\
\eta_2 &= \frac{1}{2}C_1(2mI + W) + \frac{1}{2}C_2(3mI + W) + \frac{1}{2}(C_3 + C_4)(5mI + W),
\end{aligned}$$

and

$$\beta_i = A_0\eta_i, \quad \overline{\beta_i} = \eta_i + K\beta_i$$

for $i = \overline{1,5}$.

It is now easy to construct the proofs of the assertions corresponding to the remaining cases of assumption H_7 following the proof of Theorem 1 and the proof given above. We omit the details. The proof of this theorem is therefore complete. □

Remark 1. Note that if v is a lower solution of problem (2.1) and $(A + B)^{-1} \ge 0$, then v satisfies the relation

$$Av'(t) \le \mathscr{F}(t, v(t), v'(t)), \quad t \in J, \quad v(0) \le x_0.$$

Here, $(A + B)^{-1} \ge 0$ means that some entries of $(A + B)^{-1}$ may be equal to zero.

References

[1] Bellman, R.:*Methods of Nonlinear Analysis*, Vol. II, Academic Press, New York, 1973.

[2] Bellman, R. and Kalaba, R.: *Quasilinearization and Nonlinear Boundary Value Problems*, American Elsevier, New York, 1965.

[3] Jankowski, T.: *Generalization of the method of quasilinearization for differential problems with a parameter*, Dynam. Systems and Appl., **8** (1999), 53–72.

[4] Jankowski, T.: *Systems of differential equations with a singular matrix*, Commun. Appl. Anal., **6** (2002), 209–218.

[5] Ladde, G.S., Lakshmikantham, V. and Vatsala, A. S.: *Monotone Iterative Techniques for Nonlinear Differential Equations*, Pitman, 1985.

[6] Lakshmikantham, V. and Vatsala, A. S.: *The method of generalized quasilinearization and its applications*, Neural, Parallel and Scientific Computations, **7** (1999), 119–140.

[7] Lakshmikantham, V., Leela, S., and Sivasundaram, S.: *Extensions of the method of quasilinearization*, J. Optimization Theory Appl., **87** (1995), 379–401.

[8] Lakshmikantham, V. and Vatsala, A. S.: *Generalized Quasilinearization for Nonlinear Problems*, Kluwer Academic Publishers, Derdrecht, 1998.

[9] Lakshmikantham, V., Leela, S., and McRae, F. A.: *Improved generalized quasilinearization method*, Nonlinear Anal., **24** (1995), 1627–1637.

[10] Lakshmikantham, V. and Shahzad, N.: *Further generalization of generalized quasilinearization method*, J. Appl. Math. Stoch. Anal., **7** (1994), 545–552.

[11] Lakshmikantham, V. and Malek, S.: *Generalized quasilinearization*, Nonlinear World, **1** (1994), 59–63.

[12] Vajravelu, K. and Vatsala, A. S.: *Extended method of quasilinearization for systems of nonlinear differential equations*, Appl. Anal., **74** (2000), 333–349.

Author's address

T. Jankowski:

Gdansk University of Technology, Department of Differential Equations, 11/12 Narutowicz St., 80-952 Gdańsk, Poland

E-mail address: `tjank@mifgate.mif.pg.gda.pl`

Doi: 10.18514/MMN.2006.89

Miskolc Mathematical Notes
Vol. 7 (2006), No. 1, pp. 27–42

HU ISSN 1787-2413
electronic version

CRITICAL CASE FOR SINGULARLY PERTURBED LINEAR BOUNDARY-VALUE PROBLEMS OF ORDINARY DIFFERENTIAL EQUATIONS

LYUDMIL KARANDZHULOV

[Received: May 12, 2003 and, in revised form, March 21, 2006]

ABSTRACT. The conditions under which a unique asymptotic representation of the solution of boundary-value problems exists for singularly perturbed systems of ordinary differential equations are shown in the work. The solution is obtained with the help of boundary functions and pseudo-inverse matrices.

Mathematics Subject Classification: 34B15

Keywords: boundary-value problems, singularly perturbation, asymptotic solution, boundary functions

1. Statement of the problem

We consider the singularly perturbed differential system

$$\varepsilon \dot{x} = Ax + \varepsilon A_1(t)x + \varphi(t), \quad t \in [a,b], \quad 0 < \varepsilon \ll 1, \tag{1.1}$$

$$lx(\cdot) = h, \quad h \in \mathbb{R}^m, \tag{1.2}$$

where the coefficients of system (1.1) and equation (1.2) satisfy the conditions:

(H1) A is a constant $(n \times n)$ matrix. If λ_i are eigenvalues of A, then $\lambda_i = 0$, $i = \overline{1,k}$, $k < n$, $\operatorname{Re}\lambda_i < 0$, $i = \overline{k+1,n}$, as p, $p < k$, linear independent eigenvectors of matrix A correspond to the zero eigenvalue;

(H2) $A_1(t)$ is an $(n \times n)$ matrix, $A_1(t) \in C^\infty[a,b]$, $\varphi(t)$ is an n-dimensional vector-function $\varphi(t) \in C^\infty[a,b]$;

(H3) $l : C[a,b] \to \mathbb{R}^m$ is an m-dimensional linear bounded vector-functional, $l = \operatorname{col}(l^1,\dots,l^m)$;

(H4) The degenerate ($\varepsilon = 0$) system (1.1), $Ax_0 + \varphi(t) = 0$, is solvable with respect to x_0.

We look for an n-dimensional vector-function $x(t,\varepsilon)$: $x(\cdot,\varepsilon) \in C^1[a,b]$, $x(t,\cdot) \in C(0,\varepsilon_0]$, satisfying (1.1), (1.2) and following relation $\lim_{\varepsilon\to 0} x(t,\varepsilon) = x_0(t)$, $t \in (a,b]$.

We shall consider the case $m \neq n$ and $p < k$. We use an asymptotic method of the boundary functions and construct an asymptotic series for the boundary-value problem (1.1), (1.2) with $\det A = 0$ (the critical case [10]).

In the case $m = n$ and $p = k$ an asymptotic solution of the Cauchy problem and two point boundary-value problem for linear and quasilinear systems is studied in [10] on the basis of the method of boundary functions. In the non-critical case $m \neq n$ and $\det A \neq 0$ the system is studied in [5]. When $m \neq n$ and $p = k$, the problem (1.1), (1.2) is considered in [8].

The construction of an asymptotic solution of (1.1), (1.2) in this work $m \neq n$, $p < k$ is represented on the basis of generalized inverse matrices and projectors [1, 4, 7] and central canonical form [2, 3].

We denote by $n_1, n_2, \dots, n_p$ ($\sum_{i=1}^{p} n_i = k$) the lengths of the Jordan cells. We will consider the case where $n_1 > \dots > n_s$, $n_{s+1} = n_{s+2} = \dots = n_{p-1} = n_p = 1$, i. e., the matrix A has a block diagonal representation

$$A = \operatorname{diag}(\overline{A},\ J_1,\ J_2, \dots, J_s,\ \Theta_{p-s}), \tag{1.3}$$

where $\overline{A}$ is a $((n-k)\times(n-k))$ matrix and has eigenvalues with negative real parts, J_i, $i = \overline{1, s}$, are $(n_i \times n_i)$ Jordan cells, and Θ_{p-s} is the $((p-s)\times(p-s))$ zero matrix.

By $A^\dagger$, we denote the unique Moore–Penrose pseudo-inverse $(n \times n)$ matrix of the matrix A [4, 7]. Denote by P_A and P_{A^*} orthoprojectors $P_A : \mathbb{R}^n \to \ker A$, $P_{A^*} : \mathbb{R}^n \to \ker A^*$, $A^* = A^T$. According to (H1) we find $\operatorname{rank} A = n - p$ and $\operatorname{rank} P_A = \operatorname{rank} P_{A^*} = n - (n - p) = p$. Let P_{A_p} be a $(n \times p)$ matrix with p linear independent columns from the matrix P_A, and let $P_{A_p^*}$ be a $(p \times n)$ matrix with kp linear independent rows of the matrix P_{A^*}.

Let $C = P_{A_p^*} P_{A_p}$ be an $(m \times n)$-constant matrix.

Lemma 1. $\operatorname{rank} C = p - s$.

Proof. The proof is based on the equalities $J_i J_i^\dagger = \operatorname{diag}(1, 1, \dots, 1, 0)$ and $J_i^\dagger J_i = \operatorname{diag}(0, 1, \dots, 1, 1)$. Keeping in mind the representation

$$A^\dagger = \operatorname{diag}(\overline{A}^{-1}, J_1^\dagger, J_2^\dagger, \dots, J_s^\dagger, \Theta_{p-s})$$

and the equalities $P_A = E_n - A^\dagger A$, $P_{A^*} = E_n - AA^\dagger$, we get that $C = P_{A_p^*} P_{A_p} = \operatorname{diag}(0, E_{p-s})$, i. e., $\operatorname{rank} C = p - s$. □

We consider the degenerate differential system

$$C \frac{d}{dt} z(t) = B(t) z(t) + l(t), \quad t \in [a, b], \tag{1.4}$$

where C is the matrix from Lemma 1.1, $B(t) = P_{A_p^*} A_1(t) P_{A_p}$ is $(p \times p)$ matrix, and l is a p-dimensional vector-function, $l(t) \in C^\infty[a, b]$.

Let the matrix $B(t)$ have the block representation

$$\begin{pmatrix} B_{11}(t) & B_{12}(t) \\ B_{21}(t) & B_{22}(t) \end{pmatrix},$$

where the matrices B_{11}, B_{12}, B_{21}, and B_{22} have dimensions $((p-s)\times(p-s))$, $((p-s)\times s)$, $(s\times(p-s))$, and $s\times s$, respectively.

Lemma 2. *System* (1.4) *takes the central canonical form if and only if* $\det B_{11}\neq 0$ $\forall t\in[a,b]$.

Proof. The proof of Lemma 2 is based on Lemma 1 and the work [3]. □

In accordance with Lemma 1 under $p\neq s$ and Lemma 2 $(p\times p)$, matrices $P(t)$ and $Q(t)$ exist such that substituting $z(t)=Q(t)y(t)$ and multiplying by $P(t)$ on the left, the system (1.4) takes central canonical form

$$\begin{pmatrix} E_{p-s} & 0 \\ 0 & \Theta_s \end{pmatrix}\frac{dy(t)}{dt}=\begin{pmatrix} L(t) & 0 \\ 0 & E_s \end{pmatrix}y(t)+\begin{pmatrix}\mu(t)\\ \nu(t)\end{pmatrix}, \tag{1.5}$$

where Θ_s is the $(s\times s)$ zero matrix, $L(t)$ is a $((p-s)\times(p-s))$ matrix, E_{p-s} and E_s are $((p-s)\times(p-s))$ and $(s\times s)$ unit matrices, respectively, and $\mu(t)$ and $\nu(t)$ are $(p-s)$ and s-dimensional vector-functions such that

$$P(t)g(t)=\begin{pmatrix}\mu(t)\\ \nu(t)\end{pmatrix}. \tag{1.6}$$

Let the $(p-s)$-dimensional vector-function $u(t)$ and s-dimensional vector-function $v(t)$ are such that $y(t)=\begin{pmatrix}u(t)\\ v(t)\end{pmatrix}$. Then the system (1.5) takes the form

$$\begin{aligned}\dot u_i(t)&=L(t)u_i(t)+\mu_i(t),\\ 0&=v_i(t)+\nu_i(t).\end{aligned} \tag{1.7}$$

We denote by $\Phi(t)$ a normal fundamental matrix of the solutions of the system $\dot x=L(t)x$. Then system (1.7) has a generalized solution

$$\begin{aligned}u(t)&=\Phi(t)\Phi^{-1}(t)\eta+\overline{u}(t),\quad \eta\in\mathbb{R}^{\mathbf{p-s}},\\ v(t)&=-\nu(t),\end{aligned} \tag{1.8}$$

where $\overline{u}(t)=\Phi(t)\int_a^t\Phi^{-1}(s)\mu(s)ds$.

Let the matrix $Q(t)$ be reduced to the block form $Q(t)=[Q_1(t), Q_2(t)]$, where $Q_1(t)$ is a $(p\times(p-s))$ matrix and $Q_2(t)$ is a $(p\times s)$ matrix. Keeping in mind the substitution $z(t)=Q(t)y(t)$, where $y(t)=[u(t),v(t)]^T$, we obtain

$$z(t)=Q_1(t)u(t)+Q_2(t)v(t).$$

In the last equality we substitute solution (1.8). Thus,

$$z(t)=\Phi(t,a)\eta+\overline{z}(t),\ \ t\in[a,b],\ \eta\in\mathbb{R}^{p-s}, \tag{1.9}$$

where

$$\bar{\Phi}(t,a) = Q_1(t)\Phi(t)\Phi^{-1}(a) \quad \text{is a } (p \times (p-s)) \text{ matrix and}$$

$$\bar{z}(t) = Q_1(t)\bar{u}(t) - Q_2(t)v(t) \tag{1.10}$$

The following lemma is needed.

Lemma 3. *Let the matrix A satisfy condition* (H1)*, and let the vector-function $f(\tau) \in C[0,+\infty)$ and satisfy the inequality $\|f(\tau)\| < c_1 e^{-\alpha_1 \tau}$, where $\tau \geq 0$, $c_1 > 0$, and $\alpha_1 > 0$. Then there exist positive constants c and γ such that the system $dx/d\tau = Ax + f(\tau)$ has a particular solution of the form*

$$x(\tau) = \int_0^\infty K(\tau,s) f(s) ds,$$

satisfying the inequality $\|x(\tau)\| \leq c \exp(-\gamma\tau)$, $\tau \geq 0$, where

$$K(\tau,s) = \begin{cases} X(\tau)PX^{-1}(s) & \text{for } 0 \leq s \leq \tau < \infty, \\ -X(\tau)(I-P)X^{-1}(s) & \text{for } 0 < \tau < s \leq \infty, \end{cases}$$

and P is the spectral projector of the matrix A to the left semi-plane.

The lemma is proved analogously to a similar lemma in [5].

2. Formally asymptotic expansion

We shall seek for a formally asymptotic expansion of the solution of problem (1.1), (1.2) in the form of the regular and singular series

$$x(t,\varepsilon) = \sum_{i=0}^{\infty} \varepsilon^i (x_i(t) + \Pi_i(\tau)), \quad \tau = \frac{t-a}{\varepsilon}, \tag{2.1}$$

where $x_i(t)$ and $\Pi_i(\tau)$ are unknown n vector functions. By $\Pi_i(\tau)$ (see [10]) we denote the boundary function in a neighbourhood of the point $t = a$. They will be constructed so that when $0 < \varepsilon \leq \varepsilon_0$, the inequalities

$$\|\Pi_i(\tau)\| \leq \gamma_i \exp(-\alpha_i \tau), \tag{2.2}$$

where γ_i and α_i are positive constants for $i = 0,1,2,\dots$ and $\tau \geq 0$, hold in $[a,b]$.

Formally, by substituting (2.1) in (1.1), (1.2), for $x_i(t)$ we obtain the systems

$$Ax_i(t) = f_i(t), \quad t \in [a,b], \quad i = 0,1,\dots, \tag{2.3}$$

where

$$f_i(t) = \begin{cases} -\varphi(t) & \text{for } i = 0, \\ L_1(x_{i-1}(t)) & \text{for } i = 1,2,\dots, \end{cases}$$

and L_1 is the differential operator $L_1(x(t)) = \frac{dx(t)}{dt} - A_1(t)x$. The boundary functions $\Pi_i(\tau)$ are solutions of the boundary problems

$$\frac{d}{d\tau}\Pi_i(\tau) = A\Pi_i(\tau) + \psi_i(\tau), \quad \tau \in [0, \tau_b],\ \tau_b = \frac{b-a}{\varepsilon}, \tag{2.4}$$

$$l(x_i(\cdot)) + l\left(\Pi_i\left(\frac{(\cdot)-a}{\varepsilon}\right)\right) = \begin{cases} h & \text{for } i = 0, \\ 0 & \text{for } i = 1, 2, \ldots, \end{cases} \tag{2.5}$$

where

$$\psi_i(\tau) = \begin{cases} 0 & \text{for } i = 0, \\ \sum_{q=i-1}^{0} \frac{1}{q!}\tau^q A_1^{(q)}(a)\Pi_{i-1-q}(\tau) & \text{for } i = 1, 2, \ldots, \end{cases} \tag{2.6}$$

We denote the normal fundamental matrix of the solutions of the homogeneous system $\frac{dx}{d\tau} = Ax$, $\tau \in [0, \tau_b]$, by $X(\tau) = \exp(\tau A)$. Let $X_{n-k}(\tau)$ be an $(n \times (n-k))$ matrix with $(n-k)$ columns from the matrix $X(\tau)$, consisting of exponentially small functions (see [8]).

2.1. Obtaining the coefficients $x_0(t)$ and $\Pi_0(\tau)$. Consider systems (2.3)–(2.6) for $i = 0$. Then the degenerate system

$$Ax_0(t) + \varphi(t) = 0 \tag{2.7}$$

is solvable with respect to $x_0(t)$ (according to (H4)) if and only if $P_{A^*}\varphi(t) = 0$ for all $t \in [a, b]$, and it has a solution

$$x_0(t) = P_{A_p}\alpha_0(t) - A^\dagger\varphi(t), \tag{2.8}$$

where $\alpha_0(t)$ is an arbitrary p-dimensional vector-function.

The general solution of system (2.4) has the form

$$\Pi_0(\tau) = X_{n-k}(\tau)c_0, \quad c_0 \in \mathbb{R}^{n-k}. \tag{2.9}$$

We define the vector-function $\alpha_0(t)$ by obtaining of $x_1(t)$. Consider the system $Ax_1(t) = f_1(t)$, where $f_1(t) = L_1(x_0(t))$. The latter system has a solution

$$x_1(t) = P_{A_p}\alpha_1(t) + A^\dagger L_1(x_0(t)) \tag{2.10}$$

if and only if $P_{A_p^*}L_1(x_0(t)) = 0$ for all $t \in [a, b]$. Keeping in mind the representation $x_0(t)$ from (2.8) and L_1, we obtain the differential system for $\alpha_0(t)$,

$$C\frac{d}{dt}\alpha_0(t) = B(t)\alpha_0(t) + g_0(t), \quad t \in [a, b], \tag{2.11}$$

where $g_0(t) = -P_{A_p^*}L_1(A^\dagger\varphi(t))$. System (2.11) coincides with system (1.4) at $l(t) \equiv g_0(t)$, $t \in [a, b]$. Then, according to Lemma 2 and the equality (1.9), we obtain

$$\alpha_0(t) = \Phi(t, a)\eta_0 + \overline{\alpha}_0(t), \quad t \in [a, b], \quad \eta_0 \in \mathbb{R}^{p-s}, \tag{2.12}$$

where $\bar{\alpha}_0(t) = Q_1(t)\bar{u}_0(t) - Q_2(t)\nu_0(t)$,

$$\bar{u}_0(t) = \Phi(t)\int_a^t \Phi^{-1}(s)\mu(s)_0 ds,$$

$$P(t)g_0(t) = \begin{pmatrix}\mu_0(t)\\ \nu_0(t)\end{pmatrix},$$

and $\Phi(t,a)$, $Q(t) = [Q_1(t),\ \ Q_2(t)]$, and $P(t)$ are the matrices from Section 1. The vector-functions $u_0(t)$ and $v_0(t)$ are solutions of the following system (see (1.7)):

$$\dot{u}_0(t) = L(t)u_0(t) + \mu_0(t),$$
$$0 = v_0(t) + \nu_0(t).$$

We substitute (2.12) into equality (2.8), and for $x_0(t)$ we obtain

$$x_0(t) = P_{A_p}\Phi(t,a)\eta_0 + P_{A_p}\bar{\alpha}_0(t) - A^{\dagger}\varphi(t). \tag{2.13}$$

Finally, for obtaining the functions $x_0(t)$ and $\Pi_0(t)$ it is sufficient to determine the vectors $\eta_0 \in \mathbb{R}^{p-s}$ and $c_0 \in \mathbb{R}^{n-k}$. In this connection, we use the boundary condition (2.5) for $i = 0$, where we substitute (2.13) and (2.9). We obtain the vectors η_0 and c_0 by the system

$$D_0(\varepsilon)c_0 + S_0\eta_o = h_0, \tag{2.14}$$

where $D(\varepsilon) = lX_{n-k}(\cdot)$ is an $(m \times (n-k))$ matrix, $S_0 = lA_p\Phi(\cdot,a)$ is an $(n \times (p-s))$ matrix, $h_0 = h - l(A_p\bar{\alpha}_0(\cdot)) - l(A^{\dagger}\varphi(\cdot))$ is an m-dimensional vector.

Keeping in mind the expression of the matrix $X_{n-k}(\tau)$ and the form of the functional $l(x)$, we assume that $D_0(\varepsilon) = \bar{D}_0 + O(\varepsilon^s \exp(-\alpha/\varepsilon))$, where $\alpha > 0$, $s \in N$, $\bar{D}_0$ is a $(m\times(n-k))$-constant matrix, and $O(\varepsilon^s \exp(-\alpha/\varepsilon))$ we denote a matrix consisting of elements infinitely small with respect to ε. Because the elements of the matrix $\bar{D}_0$ are continuous for all $\varepsilon \in (0,\varepsilon_0]$ and $\lim_{\varepsilon\to 0} D_0(\varepsilon) = \bar{D}_0$, then we determine the matrix $D_0(\varepsilon)$ for $\varepsilon = 0$, putting $D_0(0) = \bar{D}_0$. We neglect the exponentially small elements in the matrix $D_0(\varepsilon)$ and system (2.14) takes the form

$$M\begin{pmatrix}c_0\\ \eta_0\end{pmatrix} = h_0, \tag{2.15}$$

where $M = [\bar{D}_0,\ S_0]$ is a $(m \times (n+p-k-s))$ constant matrix.

Let the following condition hold:

(H5) rank $M = m = n - k + p - s$.

Then $\det M \neq 0$ and system (2.15) is always solvable and

$$\begin{aligned} c_0 &= [M^{-1}]_{n-k}h_0\\ \eta_0 &= [M^{-1}]_{p-s}h_0, \end{aligned} \tag{2.16}$$

where $[M^{-1}]_{n-k}$ and $[M^{-1}]_{p-s}$ are the first $(n-k)$ and last $(p-s)$ rows of the matrix M^{-1}. We should note that in this case $n - m = k - p + s > 0$, i. e., $n > m$.

We substitute (2.16) into (2.13) and (2.9) and get

$$\begin{aligned} x_0(t) &= P_{A_p}\Phi(t,a)[M^{-1}]_{p-s}\,h_0 + \overline{x}_0(t), \\ \Pi_0(\tau) &= X_{n-k}(\tau)[M^{-1}]_{n-k}\,h_0, \end{aligned} \tag{2.17}$$

where $\overline{x}_0(t) = P_{A_p}\overline{\alpha}_0(t) - A^{\dagger}\varphi(t)$.

2.2. Obtaining the coefficients $x_1(t)$ and $\Pi_1(\tau)$. To obtain the coefficient $x_1(t)$ from (2.10), it is sufficient to determine the function $\alpha_1(t)$. This will be realized in terms of the coefficient $x_2(t)$. System (2.3) under $i = 2$ has a solution $x_2(t) = P_{A_p}\alpha_2(t) + A^{\dagger}L_1(x_1(t))$ if and only if

$$P_{A_p^*}L_1(x_1(t)) = 0$$

for all $t \in [a,b]$. In the last equation we substitute $x_1(t)$ from (2.10). Keeping in mind the form of the operator L_1, for determining the function $\alpha_1(t)$, we obtain the degenerate differential system

$$C\frac{d}{dt}\alpha_1(t) = B(t)\alpha_1(t) + g_1(t), \quad t \in [a,b], \tag{2.18}$$

where $g_1(t) = -P_{A_p^*}L_1(A^{\dagger}L_1x_0(t))$.

System (2.18) coincides with system (2.3), (1.4) at $l(t) \equiv g_1(t)$, $t \in [a,b]$ and in accordance with Lemma 1.2 and equation (1.9), we obtain

$$\alpha_1(t) = \Phi(t,a)\eta_1 + \overline{\alpha}_1(t), \quad t \in [a,b], \quad \eta_1 \in \mathbb{R}^{p-s}, \tag{2.19}$$

where $\overline{\alpha}_1(t) = Q_1(t)\overline{u}_1(t) - Q_2(t)v_1(t)$,

$$\overline{u}_1(t) = \Phi(t)\int_a^t \Phi^{-1}(s)\mu_1(s)_0\,ds,$$

and $P(t)g_1(t) = \binom{\mu_1(t)}{\nu_1(t)}$. The vector-functions $u_1(t)$ and $v_1(t)$ are solutions of system (1.7), where $u(t) = u_1(t)$, $v(t) = v_1(t)$.

We substitute (2.19) into $x_1(t)$ from (2.10) and obtain

$$x_1(t) = P_{A_p}\Phi(t,a)\eta_1 + P_{A_p}\overline{\alpha}_1(t) + A^{\dagger}L_1(x_0(t)). \tag{2.20}$$

In accordance with Lemma 3, the general solution of the system (2.4) at $i = 1$ is

$$\Pi_1(\tau) = X_{n-k}(\tau)c_1 + \int_0^{+\infty} K(\tau,s)\psi_1(s)ds, \quad c_1 \in \mathbb{R}^{n-k}. \tag{2.21}$$

We substitute (2.20) and (2.21) into (2.5) at $i = 1$. The constant vectors η_1 and c_1 are obtained by the system

$$D_0(\varepsilon)c_1 + S_0\eta_1 = h_1(\varepsilon), \tag{2.22}$$

where

$$h_1(\varepsilon) = -l\left(\int_0^{+\infty} K\left(\frac{\cdot - a}{\varepsilon}, s\right)\psi_1(s)ds\right) - l(P_{A_p}\overline{\alpha}_1(\cdot)) - l(A^{\dagger}L_1(x_0(\cdot))).$$

Obviously,

$$h_1(\varepsilon) = h_{10} + O(\varepsilon^{s_1} \exp(-\alpha/\varepsilon)),$$

i. e., $h_1(\varepsilon)$ is with continuous elements for all $\varepsilon \in (\varepsilon_0]$ and $\lim_{\varepsilon\to 0} h_1(\varepsilon) = h_{10}$. Then we determine $h_1(\varepsilon)$ for $\varepsilon = 0$, putting $h_1(0) = h_{10}$. Since $D_0(0) = D_0$, in system (2.22) we neglect the exponentially small elements and obtain

$$M\begin{pmatrix} c_1 \\ \eta_1 \end{pmatrix} = h_{10}, \tag{2.23}$$

where M is the matrix from Section 2.1.

In accordance with condition (H5), the solution of the system (2.23)

$$c_1 = [M^{-1}]_{n-k} h_{10}, \qquad \eta_1 = [M^{-1}]_{p-s} h_{10},$$

we substitute into (2.20) and (2.21). Consequently, the coefficients $x_1(t)$ and $\Pi_1(\tau)$ have the form

$$\begin{aligned} x_1(t) &= P_{A_p}\Phi(t,a)[M^{-1}]_{p-s} h_{10} + \bar{x}_1(t), \\ \Pi_1(\tau) &= X_{n-k}(\tau)[M^{-1}]_{n-k} h_{10} + \bar{\Pi}_1(\tau), \end{aligned} \tag{2.24}$$

where $\bar{x}_1(t) = P_{A_p}\bar{\alpha}_1(t) + A^\dagger L_1(x_0(t))$ and $\bar{\Pi}_1(\tau) = \int_0^{+\infty} K(\tau,s)\psi_1(s)ds$.

2.3. Determining the coefficients $x_q(t)$ and $\Pi_q(\tau)$, $q > 1$. The inductive approach shows that the coefficients $x_q(t)$ and $\Pi_q(\tau)$ $(q > 1)$ have the form

$$\begin{aligned} x_q(t) &= P_{A_p}\Phi(t,a)[M^{-1}]_{p-s} h_{q0} + \bar{x}_q(t), \\ \Pi_q(\tau) &= X_{n-k}(\tau)[M^{-1}]_{n-k} h_{q0} + \bar{\Pi}_q(\tau), \end{aligned} \tag{2.25}$$

where

$$\begin{aligned} h_{q0} &= \lim_{\varepsilon\to 0} h_q(\varepsilon), \\ h_q(\varepsilon) &= -l\left(\int_0^{+\infty} K\left(\frac{\cdot - a}{\varepsilon}, s\right)\psi_q(s)ds\right) - l(P_{A_p}\bar{\alpha}_q(\cdot)) - l(A^\dagger L_1(x_{q-1}(\cdot))), \\ \bar{x}_q(t) &= P_{A_p}\bar{\alpha}_q(t) + A^\dagger L_1(x_{q-1}(t)), \\ \bar{\Pi}_q(\tau) &= \int_0^{+\infty} K(\tau,s)\psi_q(s)ds. \end{aligned} \tag{2.26}$$

Assume that the coefficients $x_i(t)$ and $\Pi_i(\tau)$ $i = \overline{1, q-1}$ are determined. System (2.3) for $i = q$ has a solution

$$x_q(t) = P_{A_p}\alpha_q(t) + A^\dagger L_1(x_{q-1}(t)) \tag{2.27}$$

if and only if $P_{A_p^*} L_1(x_{q-1}(t)) = 0$ for all $t \in [a,b]$. However, this equality is fulfilled because it is used in obtaining the function $\alpha_{q-1}(t)$. This solution $\alpha_{q-1}(t)$ participates in $x_{q-1}(t)$, which with respect to the induction hypothesis is determined completely.

The function $\alpha_q(t)$ is obtained from the solvability condition $P_{A_p^*}L_1(x_q(t)) = 0$, $\forall t \in [a,b]$ of system (2.3) for $i = q+1$ $Ax_{q+1}(t) = L_1(x_q(t))$. Thus, we obtain the following differential system (see (1.4)):

$$C\frac{d}{dt}\alpha_q(t) = B(t)\alpha_q(t) + g_q(t), \quad t \in [a,b],$$

where $g_q(t) = -P_{A_p^*}L_1(A^\dagger L_1(x_q(t)))$.

By Lemma 2 and equation (1.9) we find

$$\alpha_q(t) = \Phi(t,a)\eta_q + \bar{\alpha}_q(t), \quad t \in [a,b], \quad \eta_q \in \mathbb{R}^{p-s}, \tag{2.28}$$

where $\bar{\alpha}_q(t) = Q_1(t)\bar{u}_q(t) - Q_2(t)v_q(t)$,

$$\bar{u}_q(t) = \Phi(t)\int_a^t \Phi^{-1}(s)\mu_q(s)_0 ds,$$

and $P(t)g_q(t) = \binom{\mu_q(t)}{v_q(t)}$.

The vector-functions $u_q(t)$ and $v_q(t)$ are solutions of system (1.7), where $u(t) = u_q(t)$, $v(t) = v_q(t)$.

We substitute (2.28) into (2.27) and obtain

$$x_q(t) = P_{A_p}\Phi(t,a)\eta_q + P_{A_p}\bar{\alpha}_q(t) + A^\dagger L_1(x_{q-1}(t)), \quad \eta_q \in \mathbb{R}^{p-s}. \tag{2.29}$$

The general solution of system (2.4) for $i = q$ is

$$\Pi_q(\tau) = X_{n-k}(\tau)c_q + \int_0^{+\infty} K(\tau,s)\psi_q(s)ds, \quad c_q \in \mathbb{R}^{n-k}. \tag{2.30}$$

We substitute (2.29) and (2.30) in the boundary condition (2.5) for $i = q$ and get the system

$$D_0(\varepsilon)c_q + S_0\eta_q = h_q(\varepsilon),$$

where

$$h_q(\varepsilon) = -l\left(\int_0^{+\infty} K\left(\frac{\cdot - a}{\varepsilon}, s\right)\psi_q(s)ds\right) - l(P_{A_p}\bar{\alpha}_q(\cdot)) - l(A^\dagger L_1(x_{q-1}(\cdot))).$$

Since

$$D_0(\varepsilon) = \bar{D}_0 + O(\varepsilon^s \exp(-\alpha/\varepsilon)),$$

$\bar{D}_0 = \lim_{\varepsilon\to 0} D_0(\varepsilon)$, $h_q(\varepsilon) = h_{q0} + O(\varepsilon^{s_1}\exp(-\alpha/\varepsilon))$, and $h_{q0} = \lim_{\varepsilon\to 0} h_q(\varepsilon)$, after ignoring the exponentially small elements, the last system takes the form

$$M\begin{pmatrix} c_q \\ \eta_q \end{pmatrix} = h_{q0}$$

with the solution (see (H5))

$$c_q = [M^{-1}]_{n-k}h_{q0}, \quad \eta_q = [M^{-1}]_{p-s}h_{q0}. \tag{2.31}$$

We substitute (2.31) in (2.29) and (2.30) and obtain the equations (2.25), (2.26).

All the boundary functions $\Pi_i(\tau)$ satisfy inequalities (2.2). This follows from Lemma 3 and the inequality

$$||X_{n-k}(\tau)|| \leq c_1 \exp(-\beta_1\tau),$$

where $c_1 > 0, \beta_1 > 0$, and $\tau > 0$. After sequential analysis we get

$$||\Pi_0(\tau)|| \leq ||X_{n-k}(\tau)||||[M^{-1}]_{n-k}||||h_0|| \leq c_1 \exp(-\beta_1\tau)c_2c_3 = \gamma_0 \exp(-\alpha_0\tau),$$

where $||[M^{-1}]_{n-k}|| \leq c_2$, $||h_0|| \leq c_{30}$, $\gamma_0 = c_1c_2c_{30}$, $\alpha_0 = \beta_1$, and

$$\begin{aligned} ||\Pi_1(\tau)|| &\leq ||X_{n-k}(\tau)||||[M^{-1}]_{n-k}||||h_{10}|| + ||\overline{\Pi}_1(\tau)|| \leq \\ &\leq c_1 \exp(-\beta_1\tau)c_2c_{31} + \overline{c}_1 \exp(-\overline{\beta}_1\tau) \\ &\leq (c_1c_2c_{31} + \overline{c}_1)\exp(-\alpha_1\tau) = \gamma_1 \exp(-\alpha_1\tau), \end{aligned}$$

where $|h_{10}|| \leq c_{31}$, $||\overline{\Pi}_1(\tau)|| \leq \overline{c}_1 \exp(-\overline{\beta}_q\tau)$, and $\alpha_1 = \max(\beta_1, \overline{\beta}_1)$. Finally,

$$\begin{aligned} ||\Pi_q(\tau)|| &\leq ||X_{n-k}(\tau)||||[M^{-1}]_{n-k}||||h_{q0}|| + ||\overline{\Pi}_q(\tau)|| \leq \\ &\leq c_1 \exp(-\beta_1\tau)c_2c_{3q} + \overline{c}_q \exp(-\overline{\beta}_q\tau) \\ &\leq (c_1c_2c_{3q} + \overline{c}_q)\exp(-\alpha_q\tau) = \gamma_q \exp(-\alpha_q\tau), \end{aligned}$$

where $|h_{q0}|| \leq c_{3q}$, $||\overline{\Pi}_q(\tau)|| \leq \overline{c}_q \exp(-\overline{\beta}_q\tau)$, and $\alpha_q = \max(\beta_1, \overline{\beta}_q)$. Thus, the following theorem is true.

Theorem 1. *Let conditions* (H1)–(H5) *hold and let* $\det B_{11}(t) \neq 0$. *Then the boundary-value problems* (1.1), (1.2) *have a formally asymptotic solution of form* (2.1). *The coefficients of the regular and singular series have representations* (2.17) *and* (2.25) *for* $q = 1, 2, \ldots$. *For the boundary functions, the following estimate holds:*

$$||\Pi_q(\tau)|| \leq \gamma_q \exp(-\alpha_q\tau), \quad q = 0, 1, 2, \ldots,$$

where γ_q and α_q are positive constants. Moreover, the equality

$$\lim_{\varepsilon\to 0} x(t,\varepsilon) = x_0(t)$$

holds for $t \in (a, b]$.

Remark 1. The case where $\operatorname{rank} M = n_1 < \min(m, n-k+p-s)$ and $p = s$ is of independent interest.

3. A bound of the remainder term of the asymptotic series

The solution of the boundary-value problem (1.1), (1.2) we seek in the form

$$x(t,\varepsilon) = X_n(t,\varepsilon) + u_n(t,\varepsilon), \tag{3.1}$$

where $X_n(t,\varepsilon) = \sum_{i=0}^{n} \varepsilon^i(x_i(t) + \Pi_i(\tau))$, $\tau = \frac{t-a}{\varepsilon}$, $t \in [a, b]$.

We shall prove that, for $t \in [a, b]$ and $\varepsilon \in (0, \varepsilon_0]$, the function $u_n(t,\varepsilon)$ satisfies the inequality $||u_n(t,\varepsilon)|| \leq K\varepsilon^{n+1}$, where $K > 0$ and $\lim_{\varepsilon\to 0} x(t,\varepsilon) = x_0(t)$.

Let the smoothness degree of the elements of the matrix $A_1(t)$ and the function $\varphi(t)$ is $n+2$.

If $u_n(t,\varepsilon) = \varepsilon^{n+1}(x_{n+1}(t)+\Pi_{n+1})+u_{n+1}(t,\varepsilon)$ and we should prove that $\|u_{n+1}(t,\varepsilon)\| \le \bar{K}\varepsilon^{n+1}$, $\bar{K} > 0$, then there would exist a positive constant K such that $\|u_n(t,\varepsilon)\| \le K\varepsilon^{n+1}$.

Substituting $x(t,\varepsilon) = X_{n+1}(t,\varepsilon)+u_{n+1}(t,\varepsilon)$ in problem (1.1), (1.2), for the determination of $u_{n+1}(t,\varepsilon)$, we get the boundary-value problem

$$\varepsilon\frac{du_{n+1}(t,\varepsilon)}{dt} = Au_{n+1}(t,\varepsilon) + G(t,u_{n+1},\varepsilon), \tag{3.2}$$

$$l(u_{n+1}(\cdot,\varepsilon)) = 0. \tag{3.3}$$

The function $G(t,u_{n+1},\varepsilon)$ has the form

$$G(t,u_{n+1}(t,\varepsilon),\varepsilon) = AX_{n+1}(t,\varepsilon) + \varepsilon A_1(t,\varepsilon)[X_{n+1}(t,\varepsilon) + u_{n+1}(t,\varepsilon)]+ \\ + \varphi(t) - \varepsilon\frac{dX_{n+1}(t,\varepsilon)}{dt}$$

and satisfies the following conditions:

I. $\|G(t,0,\varepsilon)\| \le \xi\varepsilon^{m+2}$, where $\xi > 0$;

II. For all $\eta > 0$, a exists $\delta = \delta(\eta)$ and $\varepsilon_0 = \varepsilon_0(\eta)$ such that if $\|u'_{n+1}\| \le \delta$ and $\|u''_{n+1}\| \le \delta$, then

$$\|G(t,u'_{n+1},\varepsilon) - G(t,u''_{n+1},\varepsilon)\| \le \eta\|u'_{n+1} - u''_{n+1}\|$$

for $t \in [a,b]$ and $0 < \varepsilon \le \varepsilon_0$.

Let $A = \operatorname{diag}(\bar{A}, \bar{\bar{A}})$, $\bar{\bar{A}} = \operatorname{diag}(J,\Theta_{p-s})$ is a $(k\times k)$ matrix, $J = \operatorname{diag}(J_1,\dots,J_s)$ is a $((k-p+s)\times(k-p+s))$ matrix. Then we represent u_{n+1} in the form

$$u_{n+1}(t,\varepsilon) = (\omega_1(t,\varepsilon),\ \omega_2(t,\varepsilon),\omega_3(t,\varepsilon))^T\ ,$$

where $\omega_1(t,\varepsilon)$ is a $(n-k)$-dimensional vector, $\omega_2(t,\varepsilon)$ is a $(k-p+s)$-dimensional vector, and $\omega_3(t,\varepsilon)$ is a $(p-s)$-dimensional vector.

We introduce the following notation:

$$A_1(t,\varepsilon) = \begin{pmatrix} A_{111}(t,\varepsilon) & A_{112}(t,\varepsilon) \\ A_{121}(t,\varepsilon) & A_{122}(t,\varepsilon) \end{pmatrix},$$

where $A_{111}(t,\varepsilon)$ is a $((n-k)\times(n-k))$ matrix, $A_{112}(t,\varepsilon)$ is a $((n-k)\times k)$ matrix, $A_{121}(t,\varepsilon)$ is a $(k\times(n-k))$ matrix, $A_{122}(t,\varepsilon)$ is a $(k\times k)$ matrix;

$$A_{112}(t) = \begin{pmatrix} B_1(t) & B_2(t) \end{pmatrix}, \quad A_{121}(t) = \begin{pmatrix} C_1(t) \\ C_2(t) \end{pmatrix}, \quad A_{122}(t) = \begin{pmatrix} D_{11}(t) & D_{12}(t) \\ D_{21}(t) & D_{22}(t) \end{pmatrix},$$

where $B_1(t)$ is a $((n-k)\times(k-p+s))$ matrix, $B_2(t)$ is a $((n-k)\times(p-s))$ matrix, $C_1(t)$ is a $((k-p+s)\times(n-k))$ matrix, $C_2(t)$ is a $((p-s)\times(n-k))$ matrix, $D_{11}(t)$ is a $((k-p+s)\times(k-p+s))$ matrix, $D_{12}(t)$ is a $((k-p+s)\times(p-s))$ matrix, $D_{21}(t)$ is a $((p-s)\times(k-p+s))$ matrix, $D_{22}(t)$ is a $((p-s)\times(p-s))$ matrix;

$$G(t,0,0,\varepsilon)=\begin{pmatrix}G_1(t,0,0,0,\varepsilon)\\G_2(t,0,0,0,\varepsilon)\\G_3(t,0,0,0,\varepsilon)\end{pmatrix},$$

where $G_1(t,0,0,0,\varepsilon)$ is a $(n-k)$-dimensional vector, $G_2(t,0,0,0,\varepsilon)$ is a $(k-p+s)$-dimensional vector, $G_3(t,0,0,0,\varepsilon)$ is a $(p-s)$-dimensional vector.

System (3.2) takes the form

$$\varepsilon\frac{d\omega_1}{dt}=\bar{A}\omega_1+\varepsilon A_{111}(t)\omega_1+\varepsilon B_1(t)\omega_2+\varepsilon B_2(t)\omega_3+G_1(t,0,0,0,\varepsilon),\tag{3.4}$$

$$\varepsilon\frac{d\omega_2}{dt}=(J+\varepsilon D_{11}(t))\,\omega_2+\varepsilon D_{12}(t)\omega_3+\varepsilon C_1(t)\omega_1+G_2(t,0,0,0,\varepsilon),\tag{3.5}$$

$$\varepsilon\frac{d\omega_3}{dt}=\varepsilon D_{21}(t)\omega_2+\varepsilon D_{22}(t)\omega_3+\varepsilon C_2(t)\omega_1+G_3(t,0,0,0,\varepsilon).\tag{3.6}$$

Obviously, the inequalities $\|G_i(t,0,0,0,\varepsilon)\|\le c_{1i}\varepsilon^{n+2}$, $c_{1i}>0$, $i=1,2,3$, hold on $[a,b]$.

Let $W(t,s,\varepsilon)$ and $V(t,s)$ be the fundamental matrices for the homogeneous systems $\varepsilon\dot{x}=\bar{A}x$ and $\dot{x}=D_{22}x$. Here, $W(s,s,\varepsilon)=E_{n-k}$ and $V(s,s)=E_{p-s}$ are the unit matrices.

Let the Cauchy problem for the homogeneous system $\varepsilon\dot{x}=(J+\varepsilon D_{11}(t))x$ have only a trivial solution, and system (3.4) has the particular solution

$$\omega_2(t,\varepsilon)=\int_a^b K(t,s,\varepsilon)\left[\varepsilon D_{12}(s)\omega_3+\varepsilon C_1(s)\omega_1+G_2(s,0,0,0,\varepsilon)\right]ds,\quad t\in[a,b],$$

where

$$K(t,s,\varepsilon)=\begin{cases}\frac{1}{\varepsilon}\bar{X}(t,\varepsilon)\,\bar{X}^{-1}(s,\varepsilon), & \tau_{i-1}\le s\le t,\\ 0, & \tau_{i-1}\le t\le s,\end{cases}$$

if the eigenvalues of the matrix $J+\varepsilon D_{11}(t)$ are purely imaginary and

$$K(t,s,\varepsilon)=\begin{cases}\frac{1}{\varepsilon}\bar{X}(t,\varepsilon)\,P\bar{X}^{-1}(s,\varepsilon), & \tau_{i-1}\le s\le t,\\[2ex] -\frac{1}{\varepsilon}\bar{X}(t,\varepsilon)\,(I-P)\bar{X}^{-1}(s,\varepsilon), & \tau_{i-1}\le t\le s,\end{cases}$$

if the eigenvalues are with a positive or negative real part. The matrix P is a spectral projector of the matrix $J+\varepsilon D_{11}(t)$ on the left half-plane, and $\bar{X}(t,\varepsilon)$ is a normal fundamental matrix for the system $\varepsilon\dot{x}=(J+\varepsilon D_{11}(t)x$.

Obviously, $\int_a^b\|K(t,s,\varepsilon)\|ds\le\xi_1,\ \xi_1>0$, for $t\in[a,b]$, $\varepsilon\in(0,\varepsilon_0]$.

Lemma 4 ([6, 10]). *For the matrix $W(t,s,\varepsilon)$, when $a < s \le t \le b$, $0 < \varepsilon \le \varepsilon_0$, the exponential estimate*

$$\|W(t,s,\varepsilon)\| \le \beta \exp\left(-\alpha\left(\frac{t-s}{\varepsilon}\right)\right), \quad a \le s \le t \le b,$$

is fulfilled, where $\alpha > 0$, $\beta > 0$.

It is clear that $\|V(t,s,\varepsilon)\| \le \beta_1$, where $a \le s \le t \le b$, $\beta_1 > 0$.

Lemma 5. *Any continuous solution of system* (3.4)–(3.6) *is a solution of the system of integral equations*

$$\begin{aligned}\omega_1(t,\varepsilon) = W(t,a,\varepsilon)\omega_1(a,\varepsilon) + \int_a^t \frac{1}{\varepsilon}\,[\varepsilon A_{111}(s)\omega_1(s,\varepsilon)+ \\ +\varepsilon B_1(s)\omega_2(s,\varepsilon) + \varepsilon B_2(s)\omega_3(s,\varepsilon) + G_1(s,0,0,0,\varepsilon)]\,ds,\end{aligned} \tag{3.7}$$

$$\begin{aligned}\omega_2(t,\varepsilon) = \int_a^b K(t,s,\varepsilon)\,[\varepsilon D_{12}(s)\omega_3(s,\varepsilon) + \varepsilon C_1(s)\omega_1(s,\varepsilon) \\ +G_2(s,0,0,0,\varepsilon)]\,ds,\end{aligned} \tag{3.8}$$

$$\begin{aligned}\omega_3(t,\varepsilon) = V(t,a)\omega_3(a,\varepsilon) + \int_a^t V(t,s)\frac{1}{\varepsilon}\,[\varepsilon D_{21}(t)\omega_2(s,\varepsilon)+ \\ +\varepsilon D_{22}(s)\omega_3 + \varepsilon C_2(s)\omega_1(s,\varepsilon) + G_3(s,0,0,0,\varepsilon)]\,ds.\end{aligned} \tag{3.9}$$

We substitute $u_{n+1}(t,\varepsilon) = (\omega_1(t,\varepsilon), \omega_2(t,\varepsilon), \omega_3(t,\varepsilon))^T$ into the boundary condition (3.3) and obtain

$$\overline{l_1}\omega_1((\cdot),\varepsilon) + \overline{l_2}\omega_2((\cdot),\varepsilon) + \overline{l_3}\omega_3((\cdot),\varepsilon) = 0,$$

where $\overline{l_i}$, $i = 1,2,3$ are linear m-dimensional bounded functionals. After transformations using (3.7)–(3.9), we obtain

$$\omega_1(t,\varepsilon) = W_i(t,a,\varepsilon)\omega_1(a,\varepsilon) + V_i(t,a,\varepsilon)\omega_3(a,\varepsilon) + S_i(t,\omega_1,\omega_3,a,\varepsilon), \tag{3.10}$$

$i = 1,2,3$, where $W_1(t,a,\varepsilon) = W(t,a,\varepsilon)$, and W_i, $i = 1,2$, V_i, S_i, $i = 1,2,3$, are functions such that, for all $t \in [a,b]$ and $\varepsilon \in (0,\varepsilon_0]$,

$$\begin{aligned}\|W_i(t,a,\varepsilon)\| &\le \varepsilon k_i, \quad k_i > 0,\ i = 1,2,\\ \|V_i(t,a,\varepsilon) &\le \varepsilon d_i, \quad d_i > 0,\ i = 1,2,\\ \|V_3(t,a,\varepsilon)\| &\le \beta_2 + \varepsilon d_3, \quad \beta_2 > 0,\ d_3 > 0,\\ \|S_i(t,0,0,0,a,\varepsilon)\| &\le c_i\varepsilon^{n+1}, \quad c_i > 0,\ i = 1,2,3,\end{aligned} \tag{3.11}$$

and

$$\begin{aligned}\|S_i(t,\omega_1^2,\omega_3^2,a,\varepsilon) - S_i(t,\omega_1^1,\omega_3^1,a,\varepsilon)\| \le \\ \le \varepsilon r_i \max_{t\in[a,b]}\left(\|\omega_1^2(t,\varepsilon) - \omega_1^1(t,\varepsilon)\| + \|\omega_3^2(t,\varepsilon) - \omega_3^1(t,\varepsilon)\|\right),\end{aligned} \tag{3.12}$$

where $r_i > 0$, $i = 1, 2, 3$. It follows from relation (3.10) that the vector $\omega(a,\varepsilon) = (\omega_1(a,\varepsilon), \omega_3(a,\varepsilon))^T$ is determined by the equation

$$R(\varepsilon)\omega(a,\varepsilon) = q(\varepsilon, \omega_1, \omega_3), \tag{3.13}$$

where $R(\varepsilon) = [R_1(\varepsilon) \quad R_2(\varepsilon)]$ is an $(m\times(n+p-k-s))$ matrix, $R_1(\varepsilon) = \overline{l_1}W_1((\cdot), a, \varepsilon) + \overline{l_2}W_2((\cdot), a, \varepsilon) + \overline{l_3}W_3((\cdot), a, \varepsilon)$ is an $(m \times (n-k))$ matrix, $R_2(\varepsilon) = \overline{l_1}V_1((\cdot), a, \varepsilon) + \overline{l_2}V_2((\cdot), a, \varepsilon) + \overline{l_3}V_3((\cdot), a, \varepsilon)$ is an $(m \times (p-s))$ matrix, and

$$q\,(\varepsilon, \omega_1, \omega_3) = -\overline{l_1}S_1\,(\cdot, \omega_1, \omega_3, a, \varepsilon) - \overline{l_2}S_2\,(\cdot, \omega_1, \omega_3, a, \varepsilon) - \overline{l_3}S_3\,(\cdot, \omega_1, \omega_3, a, \varepsilon)$$

is an m-dimensional vector. Also, one has $\|q(\varepsilon, 0, 0)\| \le c_4\varepsilon^{n+1}$, $c_4 > 0$, and

$$\|q(\varepsilon, \omega_1^2, \omega_3^2) - q(\varepsilon, \omega_1^1, \omega_3^1)\| \le \varepsilon r_4 \max_{t\in[a,b]} \left(\|\omega_1^2 - \omega_1^1\| + \|\omega_3^2 - \omega_3^1\|\right),$$

where $r_4 > 0$. Since

$$R(\varepsilon) = R_0 + O\left(\exp\left(-\frac{\alpha}{\varepsilon}\right)\right),$$

where R_0 is a constant matrix, then the following condition is fulfilled:

(H6) $m = n + p - k - s$; $\det R(\varepsilon) \ne 0\ \forall \varepsilon \in [0, \varepsilon_0]$.

System (3.13) is always solvable and

$$\begin{aligned} \omega_1(a,\varepsilon) &= [R^{-1}]_{n-k}\, q(\varepsilon, \omega_1, \omega_3), \\ \omega_3(a,\varepsilon) &= [R^{-1}]_{p-s}\, q(\varepsilon, \omega_1, \omega_3). \end{aligned} \tag{3.14}$$

We shall substitute (3.14) into (3.7)–(3.9) and obtain a system which will be solved by the method of successive approximations. Let

$$\begin{aligned} \omega_i^0(t,\varepsilon) &= 0, \\ \omega_i^{s+1}(t,\varepsilon) &= W_i(t,a,\varepsilon)[R^{-1}]_{n-k}\, q(\varepsilon, \omega_1^s, \omega_3^s) + \\ &\quad + V_i(t,a,\varepsilon)[R^{-1}]_{p-s}\, q(\varepsilon, \omega_1^s, \omega_3^s) + S_i(t, \omega_1^s, \omega_3^s, a, \varepsilon), \quad i = 1, 2, 3, \end{aligned} \tag{3.15}$$

be the Picard successive approximations.

Theorem 2. *Let the conditions of Theorem 1 and assumption* (H6) *be fulfilled. If* $\|R^{-1}\| \le c_R$, *then there exists a positive constant* K *such that the asymptotic solution of the boundary-value problem* (1.1), (1.2) *has representation* (3.1), *where* $u_n(t,\varepsilon)$ *satisfies the inequality*

$$\|u_n(t,\varepsilon)\| \le K\varepsilon^{n+1}.$$

Moreover, $x(t,\varepsilon)$ *approaches the generating system when* $\varepsilon \to 0$ *and* $t \in (a, b]$.

Proof. By virtue of (3.10), (3.11), and (3.12), for the first approximation, we have

$$\max_{t\in[a,b]} \|\omega_i^1(t,\varepsilon) - \omega_i^0(t,\varepsilon)\| \le K_{i1}, \quad K_{i1} > 0,$$

where the constant $K_{i1}\varepsilon^{n+1}$ is determined by the constants c_R, k_i, d_i, c_i, and r_i.

Let $K^1 = \max_i(K_{i1})$ and $K^1\varepsilon^{n+1} = \delta$. For the last approximation we have

$$\max_{t\in[a,b]} \|\omega_i^2(t,\varepsilon) - \omega_i^1(t,\varepsilon)\| \le \varepsilon K_{i2}\delta, \quad K_{i2} > 0,\ i = 1,2,3.$$

Let $\varepsilon_0 = \frac{1}{2}\min_i(1/K_{i2})$. Then

$$\max_{t\in[a,b]} \|\omega_i^2(t,\varepsilon) -_i^1 (t,\varepsilon)\| \le \frac{1}{2}\delta = \frac{1}{2^2}2\delta.$$

Inductively we obtain

$$\max_{t\in[a,b]} \|\omega_i^{k+1}(t,\varepsilon) - \omega_i^k(t,\varepsilon)\| \le \frac{1}{2^{k+1}}2\delta.$$

This reveals that in the segment $[a,b]$, when ε is sufficiently small, the successive approximations (3.15) are absolutely and uniformly convergent. In addition, we have

$$\|\omega_i^{k+1}(t,\varepsilon)\| \le \sum_{j=1}^{k+1} \|\omega_i^j(t,\varepsilon) - \omega_i^{j-1}(t,\varepsilon)\| \le \left(1 + \frac{1}{2} + \cdots + \frac{1}{2^k}\right)\delta \le$$
$$\le \left(1 + \frac{1}{2} + \cdots + \frac{1}{2^k} + \frac{1}{2^{k+1}} + \dots\right)\delta = 2\delta.$$

Let

$$\lim_{k\to\infty} \omega_i^k(t,\varepsilon) = \omega_i(t,\varepsilon)$$

satisfy (3.10) identically. Then, on the interval $[a,b]$, for $\varepsilon \to 0$, the inequality

$$\|\omega_i(t,\varepsilon)\| \le 2\delta$$

is fulfilled. Consequently, system (3.10) has an unique continuous solution, which does not escape from the domain $\{(t,\omega) \mid a \le t \le b,\ \|\omega\| \le 2\delta\}$. Then, for all $t \in [a,b]$ and $\varepsilon \in (0,\varepsilon_0]$,

$$\|u_{n+1}(t,\varepsilon)\| \le \sum_{i=1}^{3} \|\omega_i(t,\varepsilon)\| \le 6\delta = 6K^1\varepsilon^{n+1}\ ,$$

i. e., there exists a positive constant K such that the inequality

$$\|u_n(t,\varepsilon)\| \le K\varepsilon^{n+1}$$

is fulfilled and

$$\lim_{\varepsilon\to 0} x(t,\varepsilon) = x_0(t)$$

for all $t \in [a,b]$. $\square$

References

[1] Boichuk, A. A. and Samoilenko, A. M.: *Generalized Inverse Operators and Fredholm Boundary Value Problems*, VSP, Utrecht-Boston, 2004.

[2] Campbell, S. L.: *A general form for solvable linear time varying singular systems of differential equations*, SIAM J. Math. Anal. **4** (1987), No. 1101–1115.

[3] Chistyakov, V. F.: *On the methods for solving singularity linear systems ordinary differential equations*, In: *Singularly Perturbed Systems of Ordinary Differential Equations* [in Russian], Nauka, Novosibirsk, 1982, pp. 37–65.

[4] Nashed, M. Z.: *Generalized Inverses and Applications*, Academic Press, New York, San Francisco, London, 1967.

[5] Karandzhulov, L. I.: *Asymptotic solution of definite class of singularly perturbed linear boundary-value problems for ordinary differential equations*, Annuaire de l'Université de Sofia "St. Kl. Ohridski", Faculté de mathématiques et informatique, Livre 1 – Mathématiques et Mécanique, **91** (1997), 79–95.

[6] Levinson, N.: *A boundary value problem for a singularly perturbed differential equation*, Duke Math. J., **25** (1958), 331–342.

[7] Penrose, R.: *A generalize inverse for matrices*, Proc. Cambridge Phil. Soc., **51** (1955), 406–413.

[8] Samoilenko, A. M., Boichuk, A. A., and Karandzhulov, L. I.: *Noether boundary value problem with singular perturbation*, Diferents. Uravn., **37** (2001), No. 9, 1186–1193.

[9] Samoilenko, A. M. and Yakovetz, V. P.: *About transformation of linear systems into central canonical form*, Dokl. Acad. Nauk Ukr. (1993), No. 4, 10–15.

[10] Vasil'eva, A. B. and Butuzov, V. F.: *Singularly Perturbed Equations in the Critical Case* [in Russian], Moscow State University, Moscow, 1978.

Author's address

Lyudmil Karandzhulov:

Technical University-Sofia, Faculty of Applied Mathematics and Informatics, P. O. Box 384, Sofia 1000, Bulgaria

E-mail address: likar@tu-sofia.bg

Doi: 10.18514/MMN.2006.141

Miskolc Mathematical Notes
Vol. 7 (2006), No. 1, pp. 43–52

HU ISSN 1787-2413
electronic version

ONE-STEP ITERATIVE METHODS AND THEIR QUALITATIVE ANALYSIS

MIKLÓS MINCSOVICS

[Received: October 20, 2005]

ABSTRACT. We analyze some qualitative properties of the one-step iterative methods which serve as a mathematical model for the discretized heat conduction problem. These properties are a discrete analogues of the qualitative properties of continuous problems, and we give algebraic conditions of the step-matrix under which the above basic properties are also preserved at the discrete level. We also construct the corresponding step-matrices.

Mathematics Subject Classification: 65F10

Keywords: qualitative analysis, one-step method, heat conduction problem

1. Introduction

In this paper, we investigate the qualitative properties of the sequence $\{y^k\}$, generated by the linear algebraic iterative process

$$My^{k+1} = Ny^k + b, \quad k = 0, 1, \ldots \tag{1.1}$$

where $M, N \in \mathbb{R}^{n\times n}$, $b, y^k \in \mathbb{R}^n$ and y^0 is a given vector, y^k denotes the successive iterates. Model (1.1) is called one-step iterative method.

The discretisation of many physical and engineering problems leads to one-step methods of the above form (1.1). In this model, as a necessary condition, we have to guarantee the convergence of the iteration. On the other hand, we must also investigate the model from the point of view of the preservation of the qualitative properties of the continuous solution, like conservation of the non-negativity and the concavity of the initial vector y^0 (the discretization of the initial function), monotonicity in time, etc. There are several papers which deal with the second problem (see, e. g., [3–8]). But all those papers investigate this problem as a preservation property of a given matrix splitting of some fixed matrix A. In this paper, we approach this question from the other side, that is, the iterative model (1.1) is given *a priori* and we investigate its qualitative properties. As a physical model, we consider the heat conduction problem in dimension 1.

The paper is organized as follows. In Section 2 we give those basic mathematical preliminaries which are used in the paper. Then we collect the important properties of the continuous solution of the physical problem in Section 3. In the next section we investigate the algebraic properties of the step-matrix in the iteration, which guarantee the preservation of the basic qualitative properties of the continuous solution. Namely, in Section 4 we consider the invariant subsets, in Section 5 the monotonicity property and in Section 6 we define the relation between the different subspaces. Finally, in Section 7 we construct a corresponding step-matrix by using the symmetric Latin squares.

2. Mathematical preliminaries

Throughout the paper all matrices are real, square matrices in the vector space $\mathbb{R}^{n\times n}$, i. e., $A, B \in \mathbb{R}^{n\times n}$. The ordering relation is defined in the usual way, i. e., element-wise. This means that A is non-negative (in notation: $A \geq 0$) when all elements of A are non-negative. The partial ordering between two matrices is defined as $A \geq B$ when $A - B \geq 0$. The strict ordering is the following: A is positive ($A > 0$) when each of its elements is positive. Hence, $A > B$ if $A - B > 0$. For a vector in the vector space $\mathbb{R}^n$, the definitions are similar. This means that a vector $x \in \mathbb{R}^n$ is called non-negative ($x \geq 0$) if all the components of x are non-negative. We say that $x \geq y$ if $x - y \geq 0$. Analogously, x is positive ($x > 0$) if each of its elements is positive. Therefore, $x > y$ when $x - y > 0$.

A matrix $A \in \mathbb{R}^{n\times n}$ is said to be positively diagonally dominant, or shortly PDD (resp., strictly positively diagonally dominant, or shortly SPDD) when the relation $a_{i,i} \geq \sum_{j\neq i} |a_{i,j}|$ (resp., $a_{i,i} \geq \sum_{j\neq i} |a_{i,j}|$) is fulfilled for every $i = 1, 2, \ldots, n$. This means that they are diagonally dominant or strictly diagonally dominant with non-negative diagonal elements.

If $A \in \mathbb{R}^{n\times n}$, then $\varrho(A)$ denotes the spectral radius of A; $\varrho(A) = \max_i |\lambda_i|$, where λ_i, $i = 1, 2, \ldots, n$ are the eigenvalues of A.

The matrix $A \in \mathbb{R}^{n\times n}$, $A = (a_{i,j})$, $i, j = 1, 2, \ldots, n$ is reducible if there exists a non-empty set $R \subsetneq \{1, 2, .., n\}$ such that $a_{i,j} = 0$ for $i \in R$, $j \in \{1, 2, \ldots, n\} \setminus R$, otherwise the matrix A is irreducible. Consequently, every positive matrix is irreducible.

In the sequel, $I \in \mathbb{R}^{n\times n}$ is the unit matrix. $E \in \mathbb{R}^{n\times n}$ denotes the matrix with all elements equal to 1. Similarly, $e \in \mathbb{R}^n$, $e = (1, 1, \ldots, 1)^T$.

Let $A \in \mathbb{R}^{n\times n}$ be a matrix with the property

$$A = M - N. \tag{2.1}$$

The iterative scheme (1.1) is convergent to the unique solution $y = A^{-1}b$ for each y^0 if and only if M is nonsingular, and the corresponding step-matrix of the iteration $H = M^{-1}N$ has the property $\varrho(H) < 1$. Relation (1.1) can be rewritten in the following form: $y^{k+1} = Hy^k + M^{-1}b$ or

$$x^{k+1} = Hx^k, \tag{2.2}$$

where $x^k = y^k - A^{-1}b$ is the so-called defect vector.

3. Some basic qualitative properties of the continuous solution of the heat conduction problem

In the following we list the most important qualitative properties of the following one-dimensional heat conduction problem:

$$\begin{aligned} \frac{\partial u(x,t)}{\partial t} &= \frac{\partial^2 u(x,t)}{\partial x^2}, \quad x \in (0,1), \quad t > 0, \\ u(x,0) &= u_0(x), \\ u(0,t) = u(1,t) &= 0, \quad t > 0, \end{aligned} \tag{3.1}$$

with a given (sufficiently smooth) initial function $u_0(x)$. We denote the stationary solution of problem (3.1) by $u_{st}(x)$, i. e., $u_{st}(x) = \lim_{t\to\infty} u(x,t)$. We denote by $h(x,t)$ the difference of the solution of the problem (3.1) and the stationary solution $u_{st}(x)$, i. e., $h(x,t) = u(x,t) - u_{st}(x)$. Then this function $h(x,t)$ is expected to possess the following qualitative properties.

(1) $h(x,t)$ exists, and $\lim_{t\to\infty} h(x,t) = 0$.
(2) If $E(0) = \int_0^1 h(x,0)dx \geq 0$, then $E(t) = \int_0^1 h(x,t)dx \geq 0$ for every t.
(3) If $h(x,t^*) \geq 0$, then $h(x,t) \geq 0$ for every $t \geq t^*$.
(4) If $h(x,t') \geq h(x,t)$ for every $t \geq t'$, then for every $t'' \geq t'$ the inequality $h(x,t'') \geq h(x,t^*)$ holds for every $t^* \geq t''$.
(5) If $h(x,t') \geq h(x,t)$ for every $t \geq t'$, then $h(x,t') \geq 0$. If $h(x,t'') \geq 0$, then $E(t'') = \int_0^1 h(x,t'')dx \geq 0$.
(6) If $E(0) = \int_0^1 h(x,0)dx \geq 0$, then $E(t)$ is monotonically decreasing.
(7) $M(t) = \max_x |h(x,t)|$ is monotonically decreasing.
(8) $\int_0^\infty |h(x,t)|dt < \infty$ for every given x.
(9) If $E(0) = \int_0^1 h(x,0)dx > 0$, then there exist $t', t'', t' \leq t''$, such that $h(x,t') > 0$ and $h(x,t'') \geq h(x,t)$ for every $t \geq t''$.

In the sequel, under iteration we understand formula (2.2).

4. Invariant subsets

We start the investigation of the iteration with a definition.

Definition 1. A subset $S \subset \mathbb{R}^n$ is said to be invariant with respect to the iteration if the relation $x^k \in S$ implies that $x^{k+1} \in S$ for all $k = 0, 1, \ldots$.

Let $Y \in \mathbb{R}^{n\times n}$ be a matrix, then we define the subset $S(Y) \subset \mathbb{R}^n$ as follows:

$$S(Y) := \{x \in \mathbb{R}^n : Yx \geq 0\}.$$

We note that in the analysis of the qualitative properties of the iteration (2.2) the subsets $S(E)$, $S(I)$, $S(I - H)$, and $S(A)$ are of a special importance (see [3,4]).

Lemma 1. *If the step-matrix H is non-negative, and H and Y commute, then the subset $S(Y)$ is invariant with respect to the iteration.*

Proof. Assume that $x^k \in S$. Then by use of the definition of the subset $S(Y)$ and the non-negativity of H, the relation $HYx^k \geq 0$ holds. Hence using the commutativity assumption and the definition of the iteration (2.2), we get $Yx^{k+1} \geq 0$, which proves the statement. □

Corollary 1. *Assume that the step-matrix H is non-negative.*

(a) *Then the subset $S(p(H))$, where $p(H)$ is a polynomial of H, is invariant with respect to the iteration. Consequently $S(I)$ and $S(I-H)$ are invariant with respect to the iteration.*
(b) *If H and E/A commute, then the subset $S(E)/S(A)$ is invariant with respect to the iteration.*

In the following we investigate the relation between the subsets $S(I), S(I-H), S(E)$ and $S(A)$.

Lemma 2. *The following relations hold between the subsets $S(I), S(I-H), S(E)$, and $S(A)$:*

(a) *If $T \geq 0$, then $S(T) \supset S(I)$. Consequently, $S(E) \supset S(I)$.*
(b) *If $T^{-1} \geq 0$, then $S(I) \supset S(T)$. Consequently, if $A^{-1} \geq 0$, then $S(I) \supset S(A)$.*
(c) *If $M^{-1} \geq 0$, then $S(I-H) \supset S(A)$.*
(d) *If $(I-H)^{-1} \geq 0$, then $S(I) \supset S(I-H)$.*

Proof. (a) $T \geq 0$, therefore $x \geq 0$ clearly results in the relation $Tx \geq 0$.

The proof of (b) and (d) is straightforward.

The statement (c) follows directly from the fact that (2.1) can be rewritten in the following form: $A = M(I-H)$. □

Remark 1. Assume that the step-matrix H is a non-negative convergent matrix, i. e., $H \geq 0$ and $\varrho(H) < 1$. Then $(I-H)^{-1} = \sum_{k=1}^{\infty} H^k$, and hence $(I-H)^{-1} \geq 0$. Therefore $S(I)$ and $S(I-H)$ are invariant with respect to the iteration, see part (a) in Corollary 1, and $S(I) \supset S(I-H)$, according to assertion (d) of Lemma 2.

Let us investigate the set of the matrices which commute with some given matrix E.

We define the subspace $C_E \subset \mathbb{R}^{n\times n}$ as follows:

$$C_E := \{X \in \mathbb{R}^{n\times n} : EX = XE\}.$$

Remark 2. The subset C_E forms both a vector space and a ring.

The properties of the operations are obviously true. Moreover, the subset C_E is closed for the operations, because for any $A, B \in C_E$ the following relations are valid: $(A \pm B)E = AE \pm BE = EA \pm EB = E(A \pm B)$, $(\lambda A)E = \lambda EA = E(\lambda A)$, $ABE = AEB = EAB$. Clearly, the matrices 0 and $-A$ also belong to C_E.

Lemma 3. *The following statements are equivalent:*

(a) $X \in C_E$;

(b) *For all fixed* $1 \le i, j \le n$, *the relation*

$$\sum_{k=1}^{n} x_{k,i} = \sum_{k=1}^{n} x_{j,k}$$

holds.

Proof. The statement follows by direct calculation of the elements of the matrices XE and EX, respectively. □

5. Monotonicity

First we define monotonicity in the subset S.

Definition 2. Let $Z \in \mathbb{R}^{n\times n}$ be a given matrix. The iteration is said to be Z-monotone in a subset $S \subset \mathbb{R}^n$ if the following two conditions are fulfilled:*

(I) The subset $S \subset \mathbb{R}^n$ is invariant with respect to the iteration;

(II) For any $x^k \in S$, the relation $Zx^k \ge Zx^{k+1}$ holds.

Lemma 4. *For* $H \ge 0$ *the iteration is* I*-monotone in a subset* $S(I - H)$.

Proof. The invariance of $S(I - H)$ follows from Corollary 1 (a).

In order to show the second property, assume that $x^k \in S(I-H)$, i. e., $(I-H)x^k \ge 0$. Then $x^k \ge Hx^k$, which, according to (2.2) means the required relation $x^k \ge x^{k+1}$. □

Remark 3. The above proof also shows that for the iteration the relation $x^k \ge x^{k+1}$ implies the inclusion $x^k \in S(I - H)$. This means that the I-monotonicity property of the iteration is valid only in the subset $S(I - H)$.

Corollary 2. *Let us suppose that* $A^{-1} \ge 0$, $H \ge 0$, $M^{-1} \ge 0$ *and* H *and* A *commute. Then the iteration is* I*-monotone in the subset* $S(A)$.

The first condition follows from the commutativity of H and A, see Lemma 1, about the second condition statement (c) in Lemma 2 and Lemma 4.

Lemma 5. *Let us suppose that the step-matrix* H *is non-negative,* H *and* E *commute and* $I - H$ *is PDD. Then the iteration is* E*-monotone in the subset* $S(E)$.

Proof. The invariance of the subset $S(E)$ follows from the statement (b) in Corollary 1.

For all $x \in S(E)$ the inequality $(I - H)Ex \ge 0$ is fulfilled, because $Ex \ge 0$, and for any $1 \le i, j \le n$ the relation $(Ex)_i = (Ex)_j = \sum_{l=1}^{n} x_l$ holds. Moreover, $I - H$ is PDD. Hence, for any $x^k \in S(E)$ we have $(I - H)Ex^k \ge 0$. So, by use of the commutativity property, we have $E(I - H)x^k \ge 0$, which implies the required relation $Ex^k \ge Ex^{k+1}$. □

*For more details, we refer to [3, 4].

Lemma 6. *If the diagonal elements of H are non-negative, then the following two statements are equivalent:*

(a) *$I - H$ is PDD (resp., SPDD);*
(b) *$\|H\|_\infty \le 1$ (resp., $\|H\|_\infty < 1$).*

Proof. Using the definition of PDD (resp., SPDD) and the non-negativity of the diagonal elements of H, we see that the properties $1 - a_{i,i} \ge \sum_{j \ne i} |a_{i,j}|$ (resp., $1 - a_{i,i} > \sum_{j \ne i} |a_{i,j}|$) and $1 \ge \sum_{j=1}^n |a_{i,j}|$ (resp., $1 > \sum_{j=1}^n |a_{i,j}|$) are equivalent, which proves the statement. □

Lemma 7. *Let $x^0 \in \mathbb{R}^n$, $x^0 \ne 0$. Then the following statements are equivalent:*

(a) *$\|H\|_\infty \le 1$ (resp., $\|H\|_\infty < 1$);*
(b) *$\|x^k\|_\infty \ge \|x^{k+1}\|_\infty$ (resp., $\|x^k\|_\infty > \|x^{k+1}\|_\infty$) for every $x^0 \in \mathbb{R}^n$, $x^0 \ne 0$ and all $k = 0, 1, \dots$.*

Proof. The relation $\|x^{k+1}\|_\infty = \|Hx^k\|_\infty \le \|H\|_\infty \|x^k\|_\infty \le$ (resp., $<$) $\|x^k\|_\infty$ shows that (a) implies (b).

In order to obtain the converse implication, assume that (b) is true. Then, for any $x^0 \in \mathbb{R}^n$ with the property $x^0 \ne 0$, the inequality $\|x^0\|_\infty \ge \|x^1\|_\infty$ is true. Since $x^1 = Hx^0$, one has

$$\frac{\|Hx^0\|_\infty}{\|x^0\|_\infty} \le 1.$$

Consequently,

$$\max_{x \in \mathbb{R}^n,\, x \ne 0} \frac{\|Hx^0\|_\infty}{\|x^0\|_\infty} \le 1,$$

which yields (a). The proof of the strict inequality is similar. □

Lemma 8. *Assume that H is a non-negative convergent matrix, i. e., $H \ge 0$ and $\varrho(H) < 1$. Then $\sum_{k=0}^\infty |(x^k)_i| < \infty$ for every $1 \le i \le n$.*

Proof. Under the assumptions made we have

$$\sum_{k=0}^\infty \left|(x^k)_i\right| = \sum_{k=0}^\infty \left|\left(H^k x^0\right)_i\right| \le \sum_{k=0}^\infty \left(H^k \left\|x^0\right\|_\infty e\right)_i = \left\|x^0\right\|_\infty \left(\sum_{k=0}^\infty H^k e\right)_i$$
$$= \|x^0\|_\infty \left(\left(\sum_{k=0}^\infty H^k\right) e\right)_i = \|x^0\|_\infty \left((I - H)^{-1} e\right)_i = \|x^0\|_\infty \|(I - H)^{-1}\|_\infty.$$

The right-hand side here is finite, and this proves the statement. □

6. Relation between subspaces

In this Section we analyze the relations between the subspaces $S(E)$, $S(I)$, and $S(I - H)$ with respect to the iteration. The analysis is based on the so-called power method.

Lemma 9. *If the step-matrix H is non-negative, irreducible and $H \in C_E$, then*
(1) *the vector e is an eigenvector of H;*
(2) *$\varrho(H)$ is an eigenvalue of H;*
(3) *$\varrho(H)$ corresponds to the eigenvector e;*
(4) *$\varrho(H)$ is in absolute value a simple eigenvalue.*

Proof. Let us note that $He = \|H\|_\infty e$, hence $\|H\|_\infty$ is an eigenvalue of H. Since $\|H\|_\infty \geq \varrho(H)$, it follows that $\|H\|_\infty$ is, in absolute value, the largest eigenvalue with the eigenvector e. For non-negative and irreducible matrices the Perron–Frobenius theorem (see [9]) guarantees that this eigenvalue is in absolute value simple. □

Theorem 1. *Let us suppose that H is a non-negative, irreducible matrix, $H \in C_E$ and H has an orthonormal system of eigenvectors. Assume that $x^0, y^0 \in \mathbb{R}^n$ are two arbitrary vectors with the property $(x^0, e) = (y^0, e) = z \neq 0$. Then*

$$\lim_{k\to\infty} \frac{\max_{1\le i\le n}\{(x^k)_i\}}{\min_{1\le i\le n}\{(y^k)_i\}} = 1.$$

Proof. We denote the eigenvalues of H by λ_m, $m = 1, 2, \ldots, n$. Using the proof of Lemma 9 and the notation $\|H\|_\infty = s$, we can write $s = \lambda_1 > |\lambda_2| \geq \cdots \geq |\lambda_n|$. Furthermore, we denote the corresponding orthonormal eigenvectors by v^m, $m = 1, 2 \ldots, n$, respectively. Then, using again the Lemma 9 and the relation $\left\|\frac{1}{\sqrt{n}}e\right\|_2 = 1$, we get $v^1 = \frac{1}{\sqrt{n}}e$. Considering the decomposition with respect to the system (v^m), we get $x^0 = \sum_{m=1}^n a_m v^m$, where $a_m = (x^0, v^m)$, and $y^0 = \sum_{m=1}^n b_m v^m$, where $b_m = (y^0, v^m)$. Note that $a_1 = b_1 = \frac{z}{\sqrt{n}}$. Hence,

$$\frac{(x^k)_i}{(y^k)_j} = \frac{(H^k x^0)_i}{(H^k y^0)_j} = \frac{\left(H^k \sum_{m=1}^n a_m v^m\right)_i}{(H^k \sum_{m=1}^n b_m v^m)_j} = \frac{\left(\sum_{m=1}^n \lambda_m^k a_m v^m\right)_i}{\left(\sum_{m=1}^n \lambda_m^k b_m v^m\right)_j}$$
$$= \frac{\left(s^k \frac{z}{\sqrt{n}} e + \sum_{m=2}^n \lambda_m^k a_m v^m\right)_i}{\left(s^k \frac{z}{\sqrt{n}} e + \sum_{m=2}^n \lambda_m^k b_m v^m\right)_j} = \frac{\left(e + \frac{\sqrt{n}}{z} \sum_{m=2}^n \left(\frac{\lambda_m}{s}\right)^k a_m v^m\right)_i}{\left(e + \frac{\sqrt{n}}{z} \sum_{m=2}^n \left(\frac{\lambda_m}{s}\right)^k b_m v^m\right)_j}.$$

Since $|\lambda_m| < s$, for $m = 2, \ldots, n$, therefore $\lim_{k\to\infty} \left(\frac{\lambda_m}{s}\right)^k = 0$, for $m = 2, \ldots, n$. This completes the proof. □

We denote the maximum element of H by $\max(H)$, and the minimum element of H by $\min(H)$, respectively.

Corollary 3. Let us apply the Theorem 1 to the columns of the step-matrix H. Since H is non-negative, irreducible, $H \in C_E$ and H has an orthonormal system of eigenvectors, therefore

$$\lim_{k\to\infty} \frac{\max(H^k)}{\min(H^k)} = 1.$$

Theorem 2. *Suppose that the step-matrix H is non-negative, irreducible, $H \in C_E$ and H has an orthonormal system of eigenvectors. Let $x^0 \in \mathbb{R}^n, x^0 \in S(E)$ and $Ex^0 \neq 0$. Then there exists an index $k_0 \in \mathbb{N}_0^+$ such that for every $k \in \mathbb{N}_0^+$, $k \geq k_0$, the inequality $x^k \geq 0$ holds (in other words $x^k \in S(I)$).*

Proof. Corollary 3 means that for all $\varepsilon \in \mathbb{R}^+$, there exists an index $k_0 \in \mathbb{N}_0^+$ such that for each $k \in \mathbb{N}_0^+, k \geq k_0$. the inequality

$$1 \leq \frac{\max(H^k)}{\min(H^k)} \leq 1 + \varepsilon$$

holds. We introduce the notation p, p^+, and p^- as follows. We put $Ex^0 = pe$, where $p \in \mathbb{R}^+$; in other words, p denotes the sum of the elements of x^0. We denote the sum of the non-negative elements of x^0 by p^+, and the sum of the negative elements of x^0 by p^-. Hence, $p^+ + p^- = p$. We consider the following estimate:

$$\begin{aligned}(x^k)_i = (H^k x^0)_i \geq \min(H^k)\, p^+ + \max(H^k)\, p^- \geq \min(H^k)\, p^+ + (1+\varepsilon)\min(H^k)\, p^- \\ = \min(H^k)\,(p^+ + (1+\varepsilon)p^-) = \min(H^k)\,(p + \varepsilon p^-).\end{aligned}$$

The number $\min(H^k)$ is non-negative. The choice of an index k_0 large enough (which depends on $\varepsilon = -p/p^-$) guarantees the non-negativity of x^k for each $k \geq k_0$. □

Theorem 3. *Suppose that the step-matrix H is non-negative, irreducible, $H \in C_E$, $I - H$ is PDD, and H has an orthonormal system of eigenvectors. Let $x^0 \in \mathbb{R}^n, x^0 \in S(I)$. Then there exists a $k_0 \in \mathbb{N}_0^+$ such that the inequality $(I - H)x^k \geq 0$ holds for every $k \in \mathbb{N}_0^+$, $k \geq k_0$ (in other words, $x^k \in S(I - H)$ for such indices).*

Proof. In the same way as proof of the Theorem 2, and finally using the proof of Lemma 5. □

We summarize the results in what follows.

Corollary 4. *Suppose that the step-matrix H is non-negative, irreducible, $H \in C_E$, $\|H\|_\infty < 1$, and H has an orthonormal system of eigenvectors. Let $x^0 \in S(E)$ and $Ex^0 \neq 0$. Then*

(a) *the iteration is convergent and* $\lim_{k\to\infty} x^k = 0$*;*
(b) *$S(E)$, $S(I)$, and $S(I - H)$ are invariant with respect to the iteration;*
(c) *$S(E) \supset S(I) \supset S(I - H)$;*
(d) *the iteration is E-monotone (resp., I-monotone) in $S(E)$ (resp., $S(I - H)$);*
(e) *the iteration is strictly monotonically decreasing in maximum norm;*
(f) $\sum_{k=0}^{\infty} |(x^k)_i| < \infty$, *for every* $1 \leq i \leq n$*;*
(g) $\exists k_0, l_0 \in \mathbb{N}_0^+$, $k \leq l$, $\forall k \geq k_0$, $x^k \in S(I)$, $\forall l \geq l_0$, $x^l \in S(I - H)$.

The above-listed properties (a)–(g) are the discrete analogues of (1)–(9) from Section 3.

7. Constructing corresponding H

A Latin square is an $\mathbb{R}^{n\times n}$ matrix which consists of n sets of n numbers arranged in such a way that no orthogonal (row or column) contains the same number twice.

Lemma 10. *There exists a symmetric Latin square.*

Proof. Assume that $a_1, \dots, a_n$ are arbitrary numbers. The choice

$$h_{i,j} := a_k \quad \text{if } \; i + j = k + 1 \; (\operatorname{mod} n)$$

for the elements of the matrix H creates a symmetric Latin square. □

Corollary 5. *Let a_i, $i = 1, 2, \dots, n$, be arbitrary positive numbers with the property $\sum_{i=1}^{n} a_i < 1$. With the help of a_i we construct the step-matrix H, which is a symmetric Latin square. Then the properties (a)–(g) are satisfied because such a matrix H is non-negative, irreducible, $H \in C_E$, $\|H\|_\infty < 1$, and H has an orthonormal system of eigenvectors.*

8. Concluding remarks

In this paper we listed the basic qualitative properties of the continuous solution of the heat conduction problem (3.1), arising from the physical process. After that we investigated the algebraic properties of the step-matrix in the corresponding one-step iteration, and we have determined conditions under which the above qualitative properties of the continuous solution are preserved. Finally we constructed a suitable corresponding step-matrix.

However, there are different open problems and possible extensions which can be considered in our future works:

(1) Our aim is to analyze the independence of the important qualitative properties of the continuous solution and the important properties of the discrete solution as well.
(2) We constructed a corresponding step-matrix. But naturally it would be useful to define the necessary (and sufficient) conditions for the step-matrix in order to fulfil the qualitative properties (a)–(g).
(3) We analyzed the heat conduction problem in 1D. We can extend this analysis both for higher dimensions and for other time-dependent (parabolic type) physical problems as well.
(4) Our goal is to investigate this method from the point of view of the matrix splitting theory, namely how we can get a corresponding step-matrix from some given matrix A with some splitting procedure (regular matrix splitting, weak regular matrix splitting). These splitting procedures can be found in several papers, e. g., [1, 2, 10–12]. However, their qualitative properties are less investigated.

References

[1] Amedjoe, E. O.: *Splitting and multisplitting of matrices*, Annales Univ. Sci. Budapest, Sect. Comp. **13** (1992), 195–224.

[2] Berman, A. and Plemmons, R. J.: *Nonnegative Matrices in the Mathematical Sciences*, Academic Press, New York, 1979.

[3] Faragó, I. and Tarvainen, P.: *Qualitative analysis of one-step algebraic models with tridiagonal Toeplitz matrices*, Periodica Mathematica Hungarica, **35** (1997), No. 3, 177–192.

[4] Farkas, H., Faragó, I.. and Simon, P.: *Qualitative properties of conductive heat transfer*, In: *Thermodynamics of Energy Conversion and Transport* (eds. S. Sienuitycz and A. De Vos), Springer Verlag, 2000, pp. 199–239.

[5] Faragó, I. and Pfeil, T.: *Preserving concavity in initial-boundary value problems of parabolic type and its numerical solution*, Periodica Mathematica Hungarica, **30** (1995), No. 2, 135–139.

[6] Faragó, I.: *Qualitative properties of the numerical solution of linear parabolic problems with nonhomogeneous boundary conditions*, Computers Math. Appl., **31** (1996), No. 4/5, 143–150.

[7] Faragó I., Haroten H., Komáromi N., Pfeil T.: *A hővezetési egyenlet és numerikus megoldásának kvalitatív tulajdonságai I. Az elsőfokú közelítések nemnegativitása*, Alkalmazott Matematikai Lapok, **17** (1993), 101–121.

[8] Faragó I., Haroten H., Komáromi N., Pfeil T.: *A hővezetési egyenlet és numerikus megoldásának kvalitatív tulajdonságai II. A másodfokú közelítés nemnegativitása, a maximum elv és az oszcillációmentesség*, Alkalmazott Matematikai Lapok, **17** (1993), 123–141.

[9] Rózsa, P.: *Linear Algebra and Its Applications*, Műszaki Kiadó, Budapest, 1976.

[10] Varga, R.: *Matrix Iterative Analysis*, Prentice Hall, New Jersey, 1962.

[11] Woźnicki, I.: *Nonnegative splitting theory*, Japan J. Indust. Appl. Math., **11** (1994), No. 2, 289–342.

[12] Woźnicki, I.: *Comparison theorems for regular splittings on block partitions*, Linear Algebra and Its Applications, **253** (1997), 199–207.

Author's address

Miklós Mincsovics:

Eötvös Loránd University, Department of Applied Analysis, 1/c P. Pázmány St., H-1117, Budapest, Hungary

E-mail address: mikka76@freemail.hu

Doi: 10.18514/MMN.2006.154

Miskolc Mathematical Notes
Vol. 7 (2006), No. 1, pp. 53–89

HU ISSN 1787-2413
electronic version

BOUNDS FOR THE SPECTRUM OF COMPACT LINEAR OPERATORS IN A PREORDERED BANACH SPACE

ANDREI RONTÓ

[Received: December 20, 2005]

ABSTRACT. We establish new efficient upper bounds for the spectral radius of a completely continuous operator in a Banach space equipped by a suitable preordering. The operator considered is assumed to admit a majorant preserving the preordering and, generally speaking, may not leave the given wedge invariant.

Mathematics Subject Classification: 46B40, 34K30, 46N20

Keywords: Wedge, cone, spectral radius, eigenvalue, upper bound

CONTENTS

The research was supported in part by AS CR, Institutional Research Plan AV0Z10190503, and GA CR, Grant No. 201/06/0254. A draft version of this paper had been prepared in 2003 and appeared in 2005 in the form of the preprint [12].

1. Introduction

Many concrete problems of various nature, where the question of uniqueness of a solution is essential, as is well-known, lead one to the study of regular values of a certain bounded linear operator. Since it is always natural to try to get some useful information on the base of as few initial data as possible, is appears that estimates of the spectral radius of an operator which are derived from certain relations involving its value on a single element only, should be of the best imaginable efficiency.

There is a vast literature devoted to this kind of estimates of spectra of linear operators that are positive with respect to a cone in a Banach space (see, e. g., [6, 7]). The main idea of such statements, dating back to O. Perron, P. Jentzsch, P. Urysohn, L. Collatz, and M. Krein, was developed by many authors (see, e. g., [3, 6, 7, 10, 13, 15]).

The conditions imposed on the operator and the space where it acts vary as well as the assumed properties of the chosen element do. For example, the spectral radius of A admits the estimate

$$r(A) \geq \gamma, \tag{1.1}$$

where γ is a given positive constant, on the assumption [6] that A is a bounded linear operator leaving invariant a cone K and such that the relation $Ag - \gamma g \in K$ is true for a certain element $g \in (K - K) \setminus (-K)$. On the other hand, the number $r(A)$ satisfies the inequality

$$r(A) \leq \gamma \tag{1.2}$$

provided that $A(K) \subseteq K$, K is a solid normal cone, and the inclusion

$$\gamma g - Ag \in K \tag{1.3}$$

is true for some interior element g of K [7]. It is natural to find out that obtaining the upper bounds for the spectral radius is more difficult, that a relation of type (1.3) implies (1.2) only under additional conditions on A and K, and that these additional conditions are stronger than those guaranteeing a similar estimate (1.1) from below.

In this paper, we establish a new theorem of the kind indicated for linear mappings majorised by linear operators preserving a preordering which may not be a partial ordering (Section 4). More precisely, we obtain an efficient upper bound for the spectral radius of a completely continuous linear mapping $A : X \to X$ representable in the form

$$A = A_1 - A_2, \tag{1.4}$$

where X is a Banach space with a wedge K (which, generally speaking, may not be a cone), and the operators A_1 and A_2 leave K invariant. The proof of the result mentioned (namely, Theorem 4.1) uses an inequality satisfied by the so-called K-substantial eigenvalues of A and established in Section 3. Note that, in the theorem mentioned, the property expressing a certain "strong positivity" of a test element g from the corresponding relation (1.3) depends upon the structure of the image space of the operator A (see Remark 4.6).

Theorem 4.1 is estabished under the condition that operator (1.4) is completely continuous. This assumption is not unnatural because our present study of completely continuous linear operators in a preordered Banach space is motivated mainly by the related problems arising in the theory of functional differential equations, and to a boundary value problem for a linear functional differential equation, a compact linear operator is usually associated. For example, the set of absolutely continuous solutions of the homogeneous initial value problem

$$u'(t) = p(t)u\,(\omega(t))\,, \qquad t \in [a, b],$$
$$u(\tau) = 0,$$

obviously, coincides with that of the equation

$$u(t) = \int_{\tau}^{t} p(s)u\,(\omega(s))\,ds, \qquad t \in [a, b].$$

The linear operator

$$C\,([a, b], \mathbb{R}) \ni u \longmapsto \int_{\tau}^{\cdot} p(s)u\,(\omega(s))\,ds,$$

as is easy to show, is a compact self-mapping of the space $C\,([a, b], \mathbb{R})$ of continuous functions on the interval $[a, b]$, whenever $p : [a, b] \to \mathbb{R}$ is integrable and $\omega : [a, b] \to [a, b]$ is measurable.

The paper is organised as follows. Section 2 contains some definitions and preliminary results. In Section 3, we establish an estimate for the so-called K-substantial eigenvalues (see Definition 2.44) of a bounded linear operator in a Banach space with a wedge K. The main Theorem 4.1 of Section 4 provides a convenient upper bound for the spectral radius of a completely continuous linear operator vanishing on the blade of the wedge K containing elements that are strongly positive in the sense of Definition 2.12. Finally, in the last Section 7, Theorem 4.1 is applied to obtain conditions sufficient for the solvability of certain integro-functional equations.

2. Wedges in Banach spaces and linear operators leaving them invariant

In this section, we recall the basic definitions related to wedges in Banach spaces, introduce some notation and definitions, and establish a number of statements relied upon in the subsequent sections. Throughout the rest of this paper, X is a Banach space over the field $\mathbb{R}$.

2.1. Wedges and related preorderings. A closed subset K of X is said to be a *wedge* (see, e. g., [6]) if

$$\alpha_1 K + \alpha_2 K \subseteq K \tag{2.1}$$

for all $\{\alpha_1, \alpha_2\} \subset [0, +\infty)$, where, as usual, $\alpha_1 K + \alpha_2 K := \{\alpha_1 x_1 + \alpha_2 x_2 \mid \{x_1, x_2\} \subset K\}$.

In what follows, we assume implicitly that the wedge K is proper, i. e., is different from both the singleton $\{0\}$ and the entire space X, for there is no meaningful theory in those two extreme cases.

Remark 2.1. In the original terminology introduced by M. Krein [10], a closed set satisfying condition (2.1) is called a *linear semigroup.*

The following standard definition introduces a natural preordering in a space X with a wedge K.

Definition 2.2. The relation $x_1 \leqq_K x_2$ is said to be satisfied if, and only if $x_2 - x_1 \in K$.

We also write $x_1 \geqq_K x_2$ if, and only if $x_2 \leqq_K x_1$. Note that the relations $x_1 \leqq_K x_2$ and $x_1 \geqq_K x_2$, generally speaking, do not imply the equality $x_1 = x_2$.

Definition 2.3. The set $K \cap (-K)$ is referred to as the *blade* [6] of the wedge K.

For the sake of brevity, we shall denote the blade of the wedge K by the symbol $K^{\diamond}$:

$$K^{\diamond} := \{x \in X \mid x \geqq_K 0 \wedge x \leqq_K 0\}. \tag{2.2}$$

Remark 2.4. It is obvious from condition (2.1) and definition (2.2) that the blade of an arbitrary wedge K is a closed linear subset of K. One can readily show that $K^{\diamond}$ coincides with the maximal linear subspace contained in K.

Definition 2.5. We write $x_1 \square_K x_2$ if, and only if either $x_1 \leqq_K x_2$ or $x_1 \geqq_K x_2$.

The relation $\square_K$ is obviously reflexive and symmetric.

2.2. Measurable elements of a Banach space. Let f be an element from X, β be a real constant, and $X_{K,\beta}(f)$ be the set defined as follows:

$$X_{K,\beta}(f) := \{x \in X \mid -\beta f \leqq_K x \leqq_K \beta f\}. \tag{2.3}$$

2.2.1. Basic properties of the sets $X_{K,\beta}(f)$. In the sequel, we need some properties of sets (2.3).

Lemma 2.6. *Let β be a fixed real number. Then an element x from X belongs to the set $X_{K,\beta}(f)$ if, and only if $-x \in X_{K,\beta}(f)$.*

Proof. Due to the symmetry of the left-hand and right-hand terms, the inequality

$$-\beta f \leqq_K x \leqq_K \beta f \tag{2.4}$$

is equivalent to the relation

$$-\beta f \leqq_K -x \leqq_K \beta f, \tag{2.5}$$

whereas the latter means that $-x \in X_{K,\beta}(f)$. □

Lemma 2.7. *The following assertions are true:*

(i) $X_{K,0}(f) = K^{\diamond}$ *for all* $f \in X$;

(ii) $X_{K,\beta}(0) = K^\diamond$ *for any* $\beta \in \mathbb{R}$*;*
(iii) *The set* $X_{K,\beta}(f)$*, where* $\beta \neq 0$ *and* $f \neq 0$*, is non-empty if, and only if* $\beta f \geqq_K 0$*;*
(iv) *If* $\beta f \geqq_K 0$*, then* $X_{K,\beta}(f) \supseteq K^\diamond$*;*
(v) *If* $f \in K^\diamond$*, then* $X_{K,\beta}(f) = K^\diamond$ *for all* $\beta \in \mathbb{R}$*;*
(vi) $X_{K,\beta}(f) \setminus K^\diamond \neq \varnothing$ *if, and only if* $\beta f \geqq_K 0$ *and* $\beta f \notin K^\diamond$.

Proof. Assertions (i) and (ii) are obvious from (2.3). Let us verify assertion (iii). Indeed, let x belong to $X_{K,\beta}(f)$. This is true if, and only if (2.4) holds or, which is the same (see Lemma 2.6), relation (2.5) is satisfied. Combining (2.4) and (2.5) and using property (2.1) of K, we obtain

$$-2\beta f \leqq_K 0 \leqq_K 2\beta f,$$

i. e., $\beta f \geqq_K 0$. Conversely, if $\beta f \geqq_K 0$, then, in particular,

$$-\beta f \leqq_K \beta f \leqq_K \beta f.$$

This means that (2.4) is satisfied with $x = \beta f$, i. e., $\beta f \in X_{K,\beta}(f)$.

To prove assertion (iv), it is sufficient to note that if $\beta f \geqq_K 0$, then (2.4) is true for all elements x satisfying the relation $0 \leqq_K x \leqq_K 0$.

Let $f \in K^\diamond$ be arbitrary. By (iv), we have $K^\diamond \subset X_{K,\beta}(f)$ for all $\beta \in \mathbb{R}$. On the other hand, if $x \in X_{K,\beta}(f)$, then, according to (2.4), we obtain $x \in K^\diamond$ because βf is also an element of $K^\diamond$. Thus, assertion (v) is true.

Finally, assertion (vi) is obvious from (iii), (iv), and (v). □

Assertions (i) and (ii) of Lemma 2.7 show that there is no much sense to consider the sets $X_{K,\beta}(f)$ with $\beta f = 0$ because, in that case, they consist solely of those elements of X which are 0-measurable with respect to K in the sense of Definition 2.8 given below.

2.2.2. The definition of *f*-measurability. For the sake of brevity, we shall use the following

Definition 2.8. An element x from X is said to be *f-measurable with respect to K* if there exists a real constant β such that $x \in X_{K,\beta}(f)$.

In other words, x is f-measurable with respect to the wedge K whenever (2.4) holds for some β.

Remark 2.9. Definition 2.8 differs from a similar notion introduced in [10] because the negative values of β are allowed in (2.4). For the purposes of this paper, the definition mentioned seems to be advantageous due to the need to consider complexifications (see Section 2.5 below). Note also that, according to Definition 2.8, the set of f-measurable elements is never empty (see Proposition 2.11).

Definition 2.10. For every fixed $f \in X$, the set of all the elements of X that are f-measurable with respect to K will be denoted by $X_K(f)$.

Clearly, $X_K(f) := \bigcup_{\beta\in\mathbb{R}} X_{K,\beta}(f)$. Moreover, it follows from Lemma 2.7 that, in fact,

$$X_K(f) = \bigcup_{\beta\in\mathbb{R}:\ \beta f \geqq_K 0} X_{K,\beta}(f) \tag{2.6}$$

for any $f \in X$.

Proposition 2.11. *For any $f \in X$, the set $X_K(f)$ is a linear manifold containing $K^{\diamond}$. Furthermore, $X_K(f) \neq K^{\diamond}$ if, and only if the element f is such that $f \sqsupset_K 0$ and $f \notin K^{\diamond}$.*

Proof. The set $X_K(f)$ obviously satisfies the condition

$$\alpha_1 X_K(f) + \alpha_2 X_K(f) \subset X_K(f)$$

for all $\{\alpha_1, \alpha_2\} \subset [0, +\infty)$ and, therefore, Lemma 2.6 guarantees that it is a linear manifold.

According to Definition 2.10 and assertion (i) of Lemma 2.7, we have $X_K(f) \supset X_{K,0}(f) = K^{\diamond}$. Furthermore, equality (2.6) yields

$$X_K(f) \setminus K^{\diamond} = \bigcup_{\beta\in\mathbb{R}:\ 0\neq\beta f \geqq_K 0} X_{K,\beta}(f) \setminus K^{\diamond}. \tag{2.7}$$

However, assertion (vi) of Lemma 2.7 guarantees that the condition $f \notin K^{\diamond}$ is necessary and sufficient for the union in the right-hand side of (2.7) to contain non-empty sets. □

2.3. Strict inequalities with respect to a wedge. Given a wedge $K \subseteq X$ and a linear manifold H in X, we introduce the following binary relation on X.

Definition 2.12. For $\{f_1, f_2\} \subset X$, we write $f_1 \boxminus_{K;H} f_2$ if, and only if the inclusion

$$X_K(f_2 - f_1) \supseteq H$$

is satisfied.

One can readily verify that the equality

$$X_K(-f) = X_K(f)$$

holds for any f and, hence, the relation introduced by Definition 2.12 is symmetric, i. e., $f_1 \boxminus_{K;H} f_2$ if, and only if $f_2 \boxminus_{K;H} f_1$.

Lemma 2.13. *If $f_1 \boxminus_{K;H_1} f_2$ and $H_1 \supseteq H_2$, then $f_1 \boxminus_{K;H_2} f_2$.*

The assertion of the last lemma is easily established by using Definition 2.12.

Lemma 2.14. *For an arbitrary f from X, the relation*

$$f \boxminus_{K;X_K(f)} 0 \tag{2.8}$$

is true.

Proof. By Proposition 2.11, the set $X_K(f)$ is a linear manifold in X. According to Definition 2.12, relation (2.8) is equivalent to the inclusion $X_K(f) \supseteq X_K(f)$ and, hence, is always satisfied. □

In the case where $H = X$, we drop the corresponding subscript in the above notation and, instead of $f_1 \boxbar_{K;X} f_2$, we write $f_1 \boxbar_K f_2$:

Definition 2.15. For $\{f_1, f_2\} \subset X$, we write $f_1 \boxbar_K f_2$ if, and only if $X_K(f_2 - f_1) = X$.

The above definition allows one to introduce the following

Definition 2.16. Two elements f_1 and f_2 are said to be in the relation $f_1 \mathrel{\varepsilon\!\!-}_{K;H} f_2$ (resp., $f_1 \mathrel{-\!\!\ni}_{K;H} f_2$) if they satisfy the conditions $f_1 \boxbar_{K;H} f_2$ and $f_1 \geqq_K f_2$ (resp., $f_1 \leqq_K f_2$).

By analogy with Definition 2.15, we introduce

Definition 2.17. Two elements f_1 and f_2 are said to be in the relation $f_1 \mathrel{\varepsilon\!\!-}_K f_2$ (resp., $f_1 \mathrel{-\!\!\ni}_K f_2$) if they satisfy the conditions $f_1 \square_K f_2$ and $f_1 \geqq_K f_2$ (resp., $f_1 \leqq_K f_2$).

The fulfilment of the relations described by Definition 2.17 is verified most easily in the case of a solid wedge.

Definition 2.18. A wedge is said to be *solid* [10] if its interior is non-empty.

Following [10], we write $x_1 \ll_K x_2$ (resp., $x_1 \gg_K x_2$) if, and only if the difference $x_2 - x_1$ (resp., $x_1 - x_2$) lies in the interior of K.

Lemma 2.19. *If K is a solid wedge in X and an element $f \in X$ is such that $f \gg_K 0$, then f satisfies the relation*

$$f \mathrel{\varepsilon\!\!-}_K 0. \tag{2.9}$$

Proof. A statement equivalent to equality (2.14) for f lying in the interior of K is well-known, e. g., from [10, 14]. □

Remark 2.20. When K is a minihedral cone in X [10] (and, hence, the partial ordering $\leqq_K$ makes X into a vector lattice [1]), an element u possessing the property $u \mathrel{\varepsilon\!\!-}_K 0$ is called a *strong unit* (see Definition XIII.1.5 in [14]). In this case, the condition $X_K(u) = X$ means that the element u satisfies Axiom V from [9].

Remark 2.21. Relation (2.9), generally speaking, does not imply that $f \gg_K 0$. For example, in the space $L_\infty([0, 1])$ of essentially bounded functions endowed with the usual norm and partial ordering [14], relation (2.9) is true, e. g., for f equal almost everywhere to 1. However, the set of functions non-negative almost everywhere on $[0, 1]$ has empty interior in $L_\infty([0, 1])$.

For suitable linear manifolds H, the condition

$$f \mathrel{\varepsilon\!\!-}_{K;H} 0$$

may be regarded as a certain “strong positivity” an element $f \geqq_K 0$. The word “suitable” here means that, roughly speaking, there should not be too many strongly positive elements. For instance, there is no much sense to study the case where

$$H \subseteq K^{\diamond} \tag{2.10}$$

because, by virtue of Proposition 2.11, the inclusion

$$\bigcap_{f \in X} X_K(f) \supseteq K^{\diamond}$$

is always true and, hence, under condition (2.10), the relation $f \multimapinv_{K;K^{\diamond}} 0$ is satisfied by an arbitrary element f from X. On the other hand, certain undesirable classes of vectors f (e. g., $f = 0$ or, more generally, f satisfying the relation $0 \leqq_K f \leqq_K 0$), that are unlikely candidates for strongly positive elements, should also be excluded from consideration. These considerations lead us to the following

Proposition 2.22. *Let H be a linear manifold in X such that*

$$H \nsubseteq K^{\diamond} \tag{2.11}$$

and f be an element of X such that either the relation $f \Box_K 0$ is not true or $0 \leqq_K f \leqq_K 0$. Then the relation

$$f \boxbar_{K;H} 0 \tag{2.12}$$

is not satisfied.

Proof. Indeed, let, on the contrary, relation (2.12) holds. According to Definition 2.12, this means that $H \subseteq X_K(f)$ and, therefore, in view of condition (2.11), the set $X_K(f)$ contains some elements not belonging to $K^{\diamond}$. It then follows from Proposition 2.11 that f should satisfy the relations $f \Box_K 0$ and $f \notin K^{\diamond}$, contrary to the assumption. □

In other words, Proposition 2.22 means that a strongly positive element should always be comparable with zero and cannot be positive and negative simultaneously. This agrees well with the intuitive idea of the strict inequality.

2.4. The mappings $n_{K,f} : X_K(f) \to [0, +\infty)$. Taking a glance at Definition 2.10, we see that the non-negative number

$$n_{K,f}(x) := \inf \left\{ |\beta| \mid \beta \in (-\infty, +\infty) \text{ and } x \in X_{K,\beta}(f) \right\} \tag{2.13}$$

is well-defined for an arbitrary x from $X_K(f)$. It is also convenient to put $n_{K,f}(x) := +\infty$ for all $x \in X \setminus X_K(f)$. Thus, $n_{K,f}(x) < +\infty$ if, and only if x is f-measurable with respect to K.

Remark 2.23. One can show that, for any $f \in X$, the mapping $n_{K,f} : X_K(f) \to [0, +\infty)$ is a seminorm on the linear manifold $X_K(f)$. This seminorm is a norm if, and only if K is a cone [5, 10], i. e., if the blade of K is trivial. The mapping mentioned is defined on the entire space X if, and only if

$$X_K(f) = X, \tag{2.14}$$

which property, in contrast to the poorest case where

$$X_K(f) = K^\diamond,$$

may be regarded as a reflection of a reasonable choice of an element $f \geqq_K 0$. By Lemma 2.19, condition (2.14) is satisfied if $f \gg_K 0$. It may happen, however, that (2.14) does not hold for any f from X (e. g., if X is the Banach space of the Lebesgue integrable functions on a bounded interval $[a, b]$ and K is the cone of integrable functions $[a, b] \to \mathbb{R}$ that are non-negative almost everywhere on $[a, b]$).

In the case where K is a solid cone and $f \gg_K 0$, formula (2.13) determines the so-called f-norm [7, 10]

$$\|x\|_f = \inf \{\beta \in [0, +\infty) \mid \text{ relation (2.4) is true}\} \tag{2.15}$$

of an arbitrary element x from X. Functional (2.15) is also used in [9] in studies of vector lattices.

It is clear from (2.13) that $n_{K,f}(0) = 0$ independently of the choice of f. Moreover, the following lemma holds.

Lemma 2.24. *Let $f \in X$. Then an element $x \in X$satisfies the equality*

$$n_{K,f}(x) = 0 \tag{2.16}$$

if, and only if $x \in K^\diamond$.

Proof. Let $f \in X$ and let x be an element from the corresponding (non-empty) set $X_K(f)$. In view of Proposition 2.11, we can suppose that $f \,\square_K\, 0$. Then, clearly, $\sigma f \geqq_K 0$ for some $\sigma \in \{-1, 1\}$.

Let $x \in K^\diamond$. The element x belongs to the blade of K if, and only if

$$0 \leqq_K x \leqq_K 0, \tag{2.17}$$

which means that (2.4) is true with an arbitrary constant β such that $\operatorname{sign}\beta = \sigma$. In particular,

$$-\frac{\sigma}{k} f \leqq_K x \leqq_K \frac{\sigma}{k} f$$

for all $k \in \mathbb{N}$. Taking (2.13) into account, we conclude that $0 \le n_{K,f}(x) \le \inf_{k\in\mathbb{N}} k^{-1} = 0$, i. e., relation (2.16) holds.

Conversely, if x satisfies equality (2.16), then there exists a sequence $(\beta_k)_{k=1}^{+\infty} \subset (-\infty, +\infty)$ such that $\lim_{k\to+\infty} \beta_k = 0$ and, for all $k \ge 1$,

$$-\beta_k f \leqq_K x \leqq_K \beta_k f. \tag{2.18}$$

Passing to the limit as $k \to +\infty$ in relation (2.18) or, which is the same, in the inclusion

$$\{\beta_k f - x,\ \beta_k f + x\} \subset K$$

and taking into account the fact that K is a closed set, we arrive at relation (2.17). $\square$

Remark 2.25. The existence of an element f satisfying equality (2.14) implies, in particular, that the wedge K is reproducing [10], i. e., $K - K = X$. The converse implication, generally speaking, is not true (in particular, the cone mentioned in Remark 7.4 is reproducing but relation (2.14) is never satisfied there).

2.5. Complexification of a wedge and the related objects. In the sequel, the complex counterparts of some of the notions defined above will be needed. Throughout this section, where the related notions are introduced, we fix a real Banach space X and wedge K in X.

2.5.1. Basic issues. The complexification (see, e. g., [4], Chapt. XIII, §2) of a real Banach space $\langle X, \|\cdot\|\rangle$ is convenient to be interpreted as the complex Banach space $\hat{X}$ of formal sums $x + \mathrm{i}y$, $\{x, y\} \subset X$, $\mathrm{i}^2 = -1$, equipped with the linear operations

$$\begin{aligned} (x_1 + \mathrm{i}y_1) + (x_2 + \mathrm{i}y_2) &:= x_1 + x_2 + \mathrm{i}\,(y_1 + y_2)\,, \\ (\nu + \mathrm{i}\mu)\,(x + \mathrm{i}y) &:= \nu x - \mu y + \mathrm{i}\,(\mu x + \nu y)\,, \end{aligned} \tag{2.19}$$

where $\{x_1, x_2, y_1, y_2, x, y\} \subset X$, $\{\nu, \mu\} \subset \mathbb{R}$, and with the norm

$$\|x + \mathrm{i}y\| := \max_{\vartheta \in [-\pi, \pi]} \|x \cos\vartheta + y \sin\vartheta\|, \qquad \{x, y\} \subset X. \tag{2.20}$$

The same technique allows one to define a natural complexification of an arbitrary wedge in a real Banach space.

Definition 2.26. The set

$$\hat{K} := \{x + \mathrm{i}y \mid x \in K \wedge y \in K\} \tag{2.21}$$

will be referred to as the *complexification* of a wedge K in a Banach space X over $\mathbb{R}$.

It is easy to verify that the set $\hat{K}$, represented alternatively as $\hat{K} = K + \mathrm{i}K$, is closed with respect to norm (2.20) and forms a wedge in $\hat{X}$ in the sense that

$$\alpha_1 \hat{K} + \alpha_2 \hat{K} \subset \hat{K} \quad \text{for all } \{\alpha_1, \alpha_2\} \subset [0, +\infty). \tag{2.22}$$

By analogy with Sections 2.1 and 2.3, one can extend the binary relations $\leqq_K$ and $\ll_K$ to $\hat{X}^2$ in a natural way. More precisely, given two elements $\{z_1, z_2\} \subset \hat{X}$, we shall write $z_1 \geqq_{\hat{K}} z_2$ (resp., $z_1 \gg_{\hat{K}} z_2$) if, and only if $z_1 - z_2 \in \hat{K}$ (resp., $z_1 - z_2$ is an interior element of $\hat{K}$). Similarly, the relation $\square_{\hat{K}}$ is natural to be defined by putting $z_1 \square_{\hat{K}} z_2$ if, and only if the elements z_1 and z_2 satisfy at least one of the relations $z_1 \geqq_{\hat{K}} z_2$ and $z_1 \leqq_{\hat{K}} z_2$. The blade $(\hat{K})^{\diamond}$ of $\hat{K}$ is natural to be defined as the set of all those z from $\hat{X}$ for which both relations $z \geqq_{\hat{K}} 0$ and $z \leqq_{\hat{K}} 0$ are true, i. e.,

$$(\hat{K})^{\diamond} = \hat{K} \cap (-\hat{K}).$$

It is obvious that

$$(\hat{K})^{\diamond} = K^{\diamond} + \mathrm{i}K^{\diamond}. \tag{2.23}$$

The complexification $\hat{K}$ of a real wedge K inherits its main characteristic properties. For example, $\hat{K}$ is solid if, and only if K possesses this property.

2.5.2. Measurability of complex elements. Let $g \in \hat{X}$ and $\lambda \in \mathbb{C}$. Similarly to formula (2.3), one can define the set $\hat{X}_{\hat{K},\lambda}(g) \subset \hat{X}$ by putting

$$\hat{X}_{\hat{K},\lambda}(g) := \left\{ z \in \hat{X} \mid z \geqq_{\hat{K}} -\lambda g \wedge z \leqq_{\hat{K}} \lambda g \right\} \tag{2.24}$$

and introduce the following

Definition 2.27. An element $z \in \hat{X}$ is said to be *g-measurable with respect to the wedge $\hat{K}$* if, and only if it belongs to the set

$$\hat{X}_{\hat{K}}(g) := \bigcup_{\lambda \in \mathbb{C}} \hat{X}_{\hat{K},\lambda}(g). \tag{2.25}$$

Definition 2.27 may be regarded as a natural extension of Definition 2.8 to the complex case. For example, analogues of Lemma 2.6 and Proposition 2.11 are true for sets (2.24) and, just as in the real case, zero belongs to the set $\hat{X}_{\hat{K},\lambda}(g)$ for arbitrary $\lambda \in \mathbb{C}$ and $g \in \hat{X}$. Further properties of sets (2.24) are described by Lemma 2.33 below.

Remark 2.28. Analogues of sets (2.24) and the related objects can also be introduced in the case where $\hat{K}$ is replaced by some other set possessing property (2.22), not necessarily constructed according to formula (2.21). Such more general complex wedges are however not needed for our purposes.

A convenient characterisation of the property introduced in Definition 2.27 is provided by the following

Lemma 2.29. *Let $\{x, y, f\} \subset X$. Then the element $x + \mathrm{i}y$ is $\hat{f}$-measurable with respect to $\hat{K}$ if, and only if there exist some $r \in [0, +\infty)$ and $\omega \in [-\pi, \pi]$ such that the relations*

$$-r f \sin\omega \leqq_K x \leqq_K r f \sin\omega, \tag{2.26}$$

$$-r f \cos\omega \leqq_K y \leqq_K r f \cos\omega \tag{2.27}$$

are satisfied.

Here and everywhere in the sequel, we write $\hat{f} = f + \mathrm{i} f$ for any f from X.

Proof. By virtue of relations (2.24) and (2.25), the element $x + \mathrm{i}y$ is $\hat{f}$-measurable if, and only if there exist some $\varrho \in [0, +\infty)$ and $\vartheta \in [-\pi, \pi]$ for which

$$-\varrho\, \mathrm{e}^{\mathrm{i}\vartheta} \hat{f} \leqq_{\hat{K}} x + \mathrm{i}y \leqq_{\hat{K}} \varrho\, \mathrm{e}^{\mathrm{i}\vartheta} \hat{f}. \tag{2.28}$$

According to (2.19), we have

$$\begin{aligned} \mathrm{e}^{\mathrm{i}\vartheta} \hat{f} &= (\cos\vartheta + \mathrm{i}\sin\vartheta)(f + \mathrm{i} f) = (\cos\vartheta - \sin\vartheta) f + \mathrm{i}\,(\sin\vartheta + \cos\vartheta) f \\ &= \sqrt{2}\left[\sin\left(\frac{\pi}{4} - \vartheta\right) + \mathrm{i}\cos\left(\frac{\pi}{4} - \vartheta\right)\right] f. \end{aligned} \tag{2.29}$$

Therefore, in view of definition (2.21) of the set $\hat{K}$, the relation (2.28) is equivalent to the system of order inequalities

$$-\varrho\sqrt{2}f\sin\left(\frac{\pi}{4}-\vartheta\right)\leqq_K x\leqq_K \varrho\sqrt{2}f\sin\left(\frac{\pi}{4}-\vartheta\right), \tag{2.30}$$

$$-\varrho\sqrt{2}f\cos\left(\frac{\pi}{4}-\vartheta\right)\leqq_K y\leqq_K \varrho\sqrt{2}f\cos\left(\frac{\pi}{4}-\vartheta\right), \tag{2.31}$$

which, obviously, has form (2.26), (2.27) with $r := \varrho\sqrt{2}$ and

$$\omega := \begin{cases} \frac{\pi}{4}-\vartheta & \text{if } -\pi\le\vartheta\le-\frac{3\pi}{4},\\ -\frac{7\pi}{4}-\vartheta & \text{if } -\frac{3\pi}{4}<\vartheta\le\pi. \end{cases}$$

It is clear that the above relation between the pairs (ϱ,ϑ) and (r,ω) is one-to-one. □

Remark 2.30. Definition 2.27 reduces to Definition 2.8 in the real case. Indeed, let $\sigma f\in\hat{K}$ with some $\sigma\in\{-1,1\}$. Lemma 2.29 characterises the $\hat{f}$-measurability of the element $x = x+\mathrm{i}0$ with respect to $\hat{K}$ in terms of the existence of $(r,\omega)\in[0,+\infty)\times[-\pi,\pi]$ such that $\sigma\cos\omega\ge 0$ and relation (2.26) is true. However, the property mentioned means that (2.4) is satisfied with $\beta := r\sin\omega$.

It is natural to find out that the $\hat{f}$-measurability of an element $x+\mathrm{i}y$ with respect to $\hat{K}$ is equivalent to the f-measurability of its real and imaginary parts, x and y.

Lemma 2.31. *Let $\{x,y,f\}\subset X$. Then the element $x+\mathrm{i}y$ is $\hat{f}$-measurable with respect to $\hat{K}$ if, and only if both x and y are f-measurable with respect to K.*

Proof. The f-measurability of x and y, on the assumption that $x+\mathrm{i}y\in\hat{X}_{\hat{K}}(\hat{f})$, is a consequence of Lemma 2.29. Conversely, if $\{x,y\}\subset X_K(f)$, then, according to Definition 2.8, there exist some real α and β such that

$$-\alpha f\leqq_K x\leqq_K \alpha f, \tag{2.32}$$

$$-\beta f\leqq_K y\leqq_K \beta f. \tag{2.33}$$

Let us put

$$\omega := \begin{cases} \frac{\pi}{4}\operatorname{sign}\alpha & \text{if } \beta\ge 0,\\ \frac{3\pi}{4}\operatorname{sign}\alpha & \text{if } \beta<0 \end{cases}$$

and $r := \sqrt{2}\max\{|\alpha|,|\beta|\}$. Then, as is easy to see,

$$\sin\omega=\frac{\operatorname{sign}\alpha}{\sqrt{2}},\qquad \cos\omega=\frac{\operatorname{sign}\beta}{\sqrt{2}},$$

and, therefore, relations (2.32) and (2.33) imply that (2.26) and (2.27) are satisfied with the above values of ω and r. It remains to refer to Lemma 2.29. □

It turns out that all the sets $\hat{X}_{\hat{K}}(\hat{f})$, where $\hat{f} = f+\mathrm{i}f$, are invariant under rotations. More precisely, the following statement is true.

Lemma 2.32. *Let $f \in X$ and $\{x, y\} \subset X_K(f)$. Then, for an arbitrary $\varphi \in [-\pi, \pi]$, the element $\mathrm{e}^{\mathrm{i}\varphi}(x + \mathrm{i}y)$ is $\hat{f}$-measurable with respect to $\hat{K}$.*

Proof. It will suffice to consider the case where $f \sqsupset_K 0$. By assumption, $\{x, y\} \subset X_K(f)$ and, hence, in view of Lemma 2.31, the element $x+\mathrm{i}y$ is $\hat{f}$-measurable with respect to $\hat{K}$, where $\hat{f} = f + \mathrm{i} f$. Lemma 2.29 guarantees the existence of an $\omega \in [-\pi, \pi]$ such that relations (2.26) and (2.27) are satisfied. Multiplying both parts of (2.26) by $|\cos\varphi|$ and $|\sin\varphi|$ and taking Lemma 2.6 into account, we obtain, respectively, the relations

$$-fr|\cos\varphi|\sin\omega \leqq_K \pm x\cos\varphi \leqq_K fr|\cos\varphi|\sin\omega$$

and

$$-fr|\sin\varphi|\sin\omega \leqq_K \pm x\sin\varphi \leqq_K fr|\sin\varphi|\sin\omega,$$

where the symbol "±" means that the inequality is satisfied with both signs of the corresponding term. Similarly, multiplying both parts of (2.27) by $|\cos\varphi|$ and $|\sin\varphi|$, we get

$$-fr|\cos\varphi|\cos\omega \leqq_K \pm y\cos\varphi \leqq_K fr|\cos\varphi|\cos\omega$$

and

$$-fr|\sin\varphi|\cos\omega \leqq_K \pm y\sin\varphi \leqq_K fr|\sin\varphi|\cos\omega.$$

Therefore, by choosing the appropriate signs in the relations above and summing the corresponding terms, we obtain

$$\begin{aligned} x\cos\varphi - y\sin\varphi &\leqq_K rf[|\cos\varphi|\sin\omega + |\sin\varphi|\cos\omega], \\ x\cos\varphi - y\sin\varphi &\geqq_K -rf[|\cos\varphi|\sin\omega + |\sin\varphi|\cos\omega] \end{aligned} \tag{2.34}$$

and

$$\begin{aligned} y\cos\varphi + x\sin\varphi &\leqq_K rf[|\cos\varphi|\cos\omega + |\sin\varphi|\sin\omega], \\ y\cos\varphi + x\sin\varphi &\geqq_K -rf[|\cos\varphi|\cos\omega + |\sin\varphi|\sin\omega]. \end{aligned} \tag{2.35}$$

It is supposed that $f \sqsupset_K 0$, and, therefore, $\sigma f \geqq_K 0$ for some $\sigma \in \{-1, 1\}$. Since neither $|\cos\varphi|\sin\omega + |\sin\varphi|\cos\omega$ nor $|\cos\varphi|\cos\omega + |\sin\varphi|\sin\omega$ takes values outside the interval $[-2, 2]$, relations (2.34) and (2.35) yield

$$2\sigma rf \leqq_K x\cos\varphi - y\sin\varphi \leqq_K 2\sigma rf, \tag{2.36}$$

$$2\sigma rf \leqq_K y\cos\varphi + x\sin\varphi \leqq_K 2\sigma rf. \tag{2.37}$$

Let us put

$$\vartheta_\sigma := \begin{cases} \frac{\pi}{4} & \text{if } \sigma = 1, \\ \frac{-3\pi}{4} & \text{if } \sigma = -1. \end{cases}$$

Then $\sin\vartheta_\sigma = \cos\vartheta_\sigma = \sigma 2^{-\frac{1}{2}}$ and, therefore, relations (2.36) and (2.37) can be brought to the form

$$\begin{aligned} \varrho f\sin\vartheta_\sigma &\leqq_K x\cos\varphi - y\sin\varphi \leqq_K \varrho f\sin\vartheta_\sigma, \\ \varrho f\cos\vartheta_\sigma &\leqq_K y\cos\varphi + x\sin\varphi \leqq_K \varrho f\cos\vartheta_\sigma, \end{aligned}$$

where $\varrho := 2r\sqrt{2}$. Applying now Lemma 2.29 and taking into account the formula

$$e^{i\varphi}(x + iy) = x\cos\varphi - y\sin\varphi + i(y\cos\varphi + x\sin\varphi), \tag{2.38}$$

we conclude that the element $e^{i\varphi}(x + iy)$ is $\hat{f}$-measurable. □

The next lemma summarises several properties of sets (2.24) referred to in the sequel.

Lemma 2.33. *The following assertions are true:*

(i) $\hat{X}_{\hat{K},0}(\hat{f}) = \hat{X}_{\hat{K},\lambda}(0) = K^{\diamond} + iK^{\diamond}$ *for all* $f \in X$ *and* $\lambda \in \mathbb{C}$*;*

(ii) $\hat{X}_{\hat{K},\lambda}(g) = -\hat{X}_{\hat{K},\lambda}(g)$ *for any* $g \in \hat{X}$*;*

(iii) $\alpha^{-1}\hat{X}_{\hat{K},|\alpha|\lambda}(g) = \hat{X}_{\hat{K},\lambda}(g)$ *for any* $g \in \hat{X}$ *and* $\alpha \in \mathbb{R} \setminus \{0\}$*;*

(iv) *For* $\lambda \neq 0$ *and* $g \neq 0$*, the set* $\hat{X}_{\hat{K},\lambda}(g)$ *is non-empty if, and only if* $\lambda g \in \hat{K}$*;*

(v) *If* $\lambda g \in \hat{K}$*, then* $\hat{X}_{\hat{K},\lambda}(g) \supseteq K^{\diamond} + iK^{\diamond}$*;*

(vi) $\hat{X}_{\hat{K},\lambda}(g) \setminus (K^{\diamond} + iK^{\diamond}) \neq \emptyset$ *if, and only if* $\lambda g \in \hat{K} \setminus (K^{\diamond} + iK^{\diamond})$*.*

(vii) $\bigcap_{g\in\hat{X}} \hat{X}_{\hat{K}}(g) \supseteq K^{\diamond} + iK^{\diamond}$*;*

(viii) $\hat{X}_{\hat{K}}(g) \neq K^{\diamond} + iK^{\diamond}$ *if, and only if* $g \in [\hat{K} \cup (-\hat{K})] \setminus (K^{\diamond} + iK^{\diamond})$*.*

Proof. This statement is established similarly to Lemmata 2.6 and 2.7 and Propositions 2.11 and 2.11 from Section 2.2.

Let us prove, e. g., assertion (iii). Indeed, let $\alpha \neq 0$. By virtue of (ii), an element z belongs to the set $\hat{X}_{\hat{K},|\alpha|\lambda}$ if, and only if

$$-|\alpha|\lambda g \leqq_{\hat{K}} z \operatorname{sign}\alpha \leqq_{\hat{K}} |\alpha|\lambda g,$$

or, which is the same,

$$-\lambda g \leqq_{\hat{K}} \frac{z}{\alpha} \leqq_{\hat{K}} \lambda g. \tag{2.39}$$

However, (2.39) means nothing but the inclusion $\alpha^{-1}z \in \hat{X}_{\hat{K},\lambda}(g)$. □

2.5.3. The mappings $n_{\hat{K},g} : \hat{X} \to [0, +\infty]$. Similarly to the case of the original real space X, the $\hat{f}$-measurability of elements of $\hat{X}$ with respect to the complexification $\hat{K}$ of a wedge K in X can be characterised by a certain non-linear functional. More precisely, given $z \in \hat{X}$ and $g \in \hat{X}$, we put

$$n_{\hat{K},g}(z) := \sqrt{2}\inf\left\{|\lambda| \mid \lambda \in \mathbb{C} \wedge z \in \hat{X}_{\hat{K},\lambda}(g)\right\} \tag{2.40}$$

if z is g-measurable with respect to $\hat{K}$, and $n_{\hat{K},g}(z) := +\infty$ for $z \notin \hat{X}_{\hat{K}}(g)$. Here, we retain the same letter, n, as in the real case (cf. Section 2.2) in order not to complicate the notation unnecessarily.

Lemma 2.34. *For any* $g \in \hat{X}$*, the functional* $n_{\hat{K},g} : \hat{X}_{\hat{K}}(g) \to [0, +\infty)$ *is homogeneous in the sense that*

$$n_{\hat{K},g}(\gamma z) = |\gamma|\, n_{\hat{K},g}(z)$$

for all $z \in \hat{X}_{\hat{K}}(g)$ *and* $\gamma \in \mathbb{R}$*.*

Proof. Let us fix some $z \in \hat{X}_{\hat{K}}(g)$ and $\gamma \in \mathbb{R}$, $\gamma \neq 0$. According to formula (2.40), we have

$$n_{\hat{K},g}(\gamma z) = \sqrt{2}\inf\left\{|\lambda| \mid \lambda \in \mathbb{C} \wedge \gamma z \in \hat{X}_{\hat{K},\lambda}(g)\right\}. \tag{2.41}$$

Applying assertion (iii) of Lemma 2.33 with $\alpha = 1/\gamma$, we conclude that an element γz belongs to $\hat{X}_{\hat{K},\lambda}(g)$ if, and only if $z \in \hat{X}_{\hat{K},\lambda|\gamma|^{-1}}(g)$. Therefore, equality (2.41) can be rewritten as

$$\begin{aligned} n_{\hat{K},g}(\gamma z) &= \sqrt{2}\inf\left\{|\lambda| \mid \lambda \in \mathbb{C} \wedge \gamma z \in \hat{X}_{\hat{K},\lambda}(g)\right\} \\ &= \sqrt{2}\inf\left\{|\lambda| \mid \lambda \in \mathbb{C} \wedge z \in \hat{X}_{\hat{K},\lambda/|\gamma|}(g)\right\} \\ &= |\gamma|\sqrt{2}\inf\left\{\frac{\lambda}{|\gamma|} \mid \lambda \in \mathbb{C} \wedge z \in \hat{X}_{\hat{K},\lambda/|\gamma|}(g)\right\} \\ &= |\gamma|\sqrt{2}\inf\left\{|\mu| \mid \mu \in \mathbb{C} \wedge z \in \hat{X}_{\hat{K},\mu}(g)\right\} = |\gamma|\, n_{\hat{K},g}(z), \end{aligned}$$

as required. ◻

We are interested mainly in elements of $\hat{X}$ that are $\hat{f}$-measurable with respect to $\hat{K}$ for a suitably chosen f from X (actually, from $[K \cup (-K)] \setminus K^{\diamond}$ because otherwise, by Proposition 2.11, there are no f-measurable elements outside $K^{\diamond}$). In this case, it is convenient to use the following formulae for computation of value (2.40).

Lemma 2.35. *Let $f \in X$ and let $\{x, y\} \subset X$ be some elements f-measurable with respect to K. Then the formulae*

$$n_{\hat{K},\hat{f}}(x + \mathrm{i}y) = \sqrt{2}\inf\{\varrho \in [0,+\infty) \mid \exists\vartheta \in [-\pi,\pi] : (2.30) \textit{ and } (2.31) \textit{ hold}\} \tag{2.42}$$

and

$$n_{\hat{K},\hat{f}}(x + \mathrm{i}y) = \inf\{r \in [0,+\infty) \mid \exists\omega \in [-\pi,\pi] : (2.26) \textit{ and } (2.27) \textit{ hold}\} \tag{2.43}$$

are true.

Proof. By Lemma 2.31, the element $x + \mathrm{i}y$ is $\hat{f}$-measurable with respect to $\hat{K}$ and, therefore, the value of $n_{\hat{K},g}(x + \mathrm{i}y)$ is finite. In view of formula (2.29) established in the proof of Lemma 2.29, the relation

$$-\lambda\hat{f} \leqq_{\hat{K}} x + \mathrm{i}y \leqq_{\hat{K}} \lambda\hat{f}$$

with $\lambda = \varrho\,\mathrm{e}^{\mathrm{i}\vartheta}$ is equivalent to the system of order inequalities (2.30), (2.31). Therefore, definition (2.40) of the mapping $n_{\hat{K},\hat{f}}$ yields the required equality (2.42).

Formula (2.43) is a consequence of (2.42). Indeed, as is shown in the proof of Lemma 2.29, there is a one-to-one correspondence between systems (2.30), (2.31) and (2.26), (2.27), with

$$\frac{r}{\varrho} = \sqrt{2},$$

and it suffices to use Lemma 2.34. ◻

The best constant r in relations (2.26) and (2.27) satisfied by the respective components x and y of an $\hat{f}$-measurable element $x + \mathrm{i}y$ is determined by the value of functional (2.40). More precisely, we have

Lemma 2.36. *Let* $\{f, x, y\} \subset X$ *and*

$$n_{\hat{K},\hat{f}}(x + \mathrm{i}y) =: r < +\infty.$$

Then there exists an $\omega \in [-\pi, \pi]$ *such that relations* (2.26) *and* (2.27) *are satisfied.*

Proof. The definition of the functional $n_{\hat{K},\hat{f}}$ and Lemmata 2.29 and 2.35 yield the existence of sequences $(r_k)_{k=1}^{+\infty} \subset [0, +\infty)$ and $(\omega_k)_{k=1}^{+\infty} \subset [-\pi, \pi]$ such that $\lim_{k\to+\infty} r_k = r$ and the relations

$$-r_k f \sin\omega_k \leqq_K x \leqq_K r_k f \sin\omega_k, \tag{2.44}$$

$$-r_k f \cos\omega_k \leqq_K y \leqq_K r_k f \cos\omega_k \tag{2.45}$$

are true for all $k \in \mathbb{N}$. The compact real sequence $(\omega_k)_{k=1}^{+\infty}$ contains a subsequence $(\omega_{k_j})_{j=1}^{+\infty}$ convergent to a number $\omega \in [-\pi, \pi]$. Putting $k = k_j$ in (2.44) and (2.45), passing to the limit as $j \to +\infty$, and taking into account the fact that K is a closed subset of X, we arrive at relations (2.26) and (2.27). □

The following statement is an extension of Lemma 2.24 to the complex case.

Lemma 2.37. *Let* $f \in X$ *and* $z \in \hat{X}$. *Then* $n_{\hat{K},\hat{f}}(z) = 0$ *if, and only if* $z \in K^{\diamond} + \mathrm{i}K^{\diamond}$.

Proof. Let us suppose that $z = x + \mathrm{i}y$, where $\{x, y\} \subset K^{\diamond}$. According to formula (2.43) of Lemma 2.35, we have

$$\begin{aligned}
n_{\hat{K},\hat{f}}(x + \mathrm{i}y) &= \inf\{r \in [0, +\infty) \mid \exists\omega \in [-\pi, \pi] : (2.26) \text{ and } (2.27) \text{ hold}\}\\
&\le \inf\{r \in [0, +\infty) \mid \exists\{\sigma, \kappa\} \subset \{-1, 1\} : -\sigma r f \leqq_K x\sqrt{2} \leqq_K \sigma r f\\
&\qquad\qquad \text{and } -\kappa r f \leqq_K y\sqrt{2} \leqq_K \kappa r f\}\\
&\le \inf\left\{r \in [0, +\infty) \mid -\sigma r f \leqq_K x\sqrt{2} \leqq_K \sigma r f \text{ with some } \sigma \in \{-1, 1\}\right\}\\
&= \inf\left\{|\alpha| \mid \alpha \in \mathbb{R} \wedge -\alpha f \leqq_K x\sqrt{2} \leqq_K \alpha f\right\}\\
&= n_{K,f}(x\sqrt{2}).
\end{aligned} \tag{2.46}$$

By virtue of Lemma 2.34, we have

$$n_{K,f}(x\sqrt{2}) = \sqrt{2}\, n_{K,f}(x).$$

However, in view of Lemma 2.24, $n_{K,f}(x) = 0$ and, therefore, by (2.46), the non-negative number $n_{\hat{K},\hat{f}}(x + \mathrm{i}y)$ is equal to 0.

Assume now that $n_{\hat{K},\hat{f}}(x + \mathrm{i}y) = 0$. By virtue of Lemma 2.36, there exists an $\omega \in [-\pi, \pi]$ such that relations (2.26) and (2.27) are satisfied with $r = 0$, i. e., $0 \leqq_K x \leqq_K 0$ and $0 \leqq_K y \leqq_K 0$. This means that $\{x, y\} \subset K^{\diamond}$. □

Formula (2.40) allows one to construct a natural extension $n_{\hat{K},\hat{f}} : \hat{X} \to [0,+\infty]$ of the mapping $X \ni x \mapsto n_{K,f}(x)$ given by relation (2.13). More precisely, the following statement is true.

Proposition 2.38. *Let $f \in X$. Then the equality*

$$n_{\hat{K},\hat{f}}(x) = n_{K,f}(x) \tag{2.47}$$

is true for all $x \in X$.

Proof. First of all, we note that it suffices to consider the case where $f \,\square_K\, 0$ because in the contrary case, by Proposition 2.11, we have $X_K(f) = K^{\diamond}$ and, therefore, in view of Lemmata 2.24 and 2.37, both $n_{K,f}$ and $n_{\hat{K},\hat{f}}$ vanish on the set $X_K(f)$.

For the sake of definiteness, we assume that $f \geqq_K 0$. Setting $y = 0$ in formula (2.43) of Lemma 2.35, we obtain

$$\begin{aligned} n_{\hat{K},\hat{f}}(x) &= \inf\{r \in [0,+\infty) \mid \exists \omega \in [-\pi,\pi] : \cos\omega \geq 0 \text{ and (2.26) holds}\} \\ &= \inf\left\{r \in [0,+\infty) \;\middle|\; \exists \omega \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] : \text{(2.26) holds}\right\}. \end{aligned} \tag{2.48}$$

Since the mapping $\sin : [-\frac{\pi}{2}, \frac{\pi}{2}] \to [-1,1]$ is a bijection, we see that (2.48) can be rewritten in the form

$$\begin{aligned} n_{\hat{K},\hat{f}}(x) &= \inf\{r \in [0,+\infty) \mid \exists h \in [-1,1] : -rhf \leqq_K x \leqq_K rhf\} \\ &= \inf\{|\alpha| \in [0,+\infty) \mid -\alpha f \leqq_K x \leqq_K \alpha f\}, \end{aligned}$$

which, by virtue of (2.13), proves that equality (2.47) is true for all x from $X_K(f)$. In the case where x is not f-measurable with respect to K, by Lemma 2.31, both values are equal to $+\infty$. □

Proposition 2.38 is, in fact, a particular case of a more general result. Namely, the following statement is true.

Proposition 2.39. *Let $f \in X$. For arbitrary $\{x, y\} \subset X$ which are f-measurable with respect to the wedge K, the equality*

$$n_{\hat{K},\hat{f}}(x + \mathrm{i}y) = \sqrt{(n_{K,f}(x))^2 + (n_{K,f}(y))^2} \tag{2.49}$$

is true.

Formula (2.49) resembles to some extent the Pythagorean formula, with $n_{K,f}(x)$, $n_{K,f}(y)$, and $n_{\hat{K},\hat{f}}(x + \mathrm{i}y)$, respectively, playing the roles of catheti and hypotenuse. The proof of Proposition 2.39 is omitted.

2.5.4. Measuring rotated elements. In the sequel, we need to compute the values of the functional $n_{\hat{K},\hat{f}}$ on elements of the form $\mathrm{e}^{\mathrm{i}t}(x + \mathrm{i}y)$, where $t \in [-\pi, \pi]$ is arbitrary and $\{x, y\} \subset X_K(f)$ with some f satisfying the condition $f \,\square_K\, 0$.

Definition 2.40. Given an $f \in X$, we put

$$R_{\hat{K},\hat{f}}(x + \mathrm{i}y) := \inf_{t\in[-\pi,\pi]} n_{\hat{K},\hat{f}}(\mathrm{e}^{\mathrm{i}t}(x + \mathrm{i}y)) \tag{2.50}$$

if $\{x, y\} \subset X$ are f-measurable with respect to K, and set formally $R_{\hat{K},\hat{f}}(x + \mathrm{i}y) := +\infty$ in the contrary case.

It follows from Lemma 2.32 that the right-hand side of (2.50) is finite for arbitrary $\{x, y\} \subset X_K(f)$ and $t \in [-\pi, \pi]$ and, thus, Definition 2.40 makes sense.

Lemma 2.41. *Let $f \in X$ and $\{x, y\} \subset X_K(f)$. Then, for an arbitrary $\varphi \in [-\pi, \pi]$, the equality*

$$R_{\hat{K},\hat{f}}(\mathrm{e}^{\mathrm{i}\varphi}(x + \mathrm{i}y)) = R_{\hat{K},\hat{f}}(x + \mathrm{i}y) \tag{2.51}$$

is true.

Proof. According to formula (2.50), we have

$$\begin{aligned} R_{\hat{K},\hat{f}}(\mathrm{e}^{\mathrm{i}\varphi}(x + \mathrm{i}y)) &= \inf_{t\in[-\pi,\pi]} n_{\hat{K},\hat{f}}(\mathrm{e}^{\mathrm{i}\varphi}\mathrm{e}^{\mathrm{i}t}(x + \mathrm{i}y)) \\ &= \inf_{t\in[-\pi,\pi]} n_{\hat{K},\hat{f}}(\mathrm{e}^{\mathrm{i}(t+\varphi)}(x + \mathrm{i}y)). \end{aligned} \tag{2.52}$$

Let us take an arbitrary $\varphi \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ and put

$$\varphi_t := \begin{cases} t + \varphi + \pi & \text{if } t + \varphi < -\pi, \\ t + \varphi & \text{if } -\pi \le t + \varphi \le \pi, \\ t + \varphi - \pi & \text{if } t + \varphi > \pi \end{cases}$$

for all t from $[-\pi, \pi]$. It is clear that $\{\mathrm{e}^{\mathrm{i}\varphi_t} \mid t \in [-\pi, \pi]\} = \{\lambda \in \mathbb{C} \mid |\lambda| = 1\}$ for any φ. Therefore, equality (2.52) yields

$$\begin{aligned} R_{\hat{K},\hat{f}}(\mathrm{e}^{\mathrm{i}\varphi}(x + \mathrm{i}y)) &= \inf_{t\in[-\pi,\pi]} n_{\hat{K},\hat{f}}(\mathrm{e}^{\mathrm{i}\varphi_t}(x + \mathrm{i}y)) \\ &= \inf_{t\in[-\pi,\pi]} n_{\hat{K},\hat{f}}(\mathrm{e}^{\mathrm{i}t}(x + \mathrm{i}y)) = R_{\hat{K},\hat{f}}(x + \mathrm{i}y). \end{aligned}$$

Applying formula (2.51) sequentially, we prove that it is true with arbitrary values of φ from $[-\pi, \pi]$. □

Together with Lemma 2.41, the next statement is a basic tool in the proof of Theorem 3.1 from Section 3.

Lemma 2.42. *Let $f \in X$, $\{x, y\} \subset X_K(f)$, and*

$$R := R_{\hat{K},\hat{f}}(x + \mathrm{i}y). \tag{2.53}$$

Then there exist some $\{\vartheta_*, \omega_*\} \subset [-\pi, \pi]$ such that the relations

$$-Rf \sin\omega_* \leqq_K x \cos\vartheta_* - y \sin\vartheta_* \leqq_K Rf \sin\omega_*, \tag{2.54}$$

$$-Rf \cos\omega_* \leqq_K x \sin\vartheta_* + y \cos\vartheta_* \leqq_K Rf \cos\omega_* \tag{2.55}$$

hold. Moreover, if $R > 0$, *then there do not exist any numbers* $\varepsilon \in (0, R)$ *and* $\{\tilde{\vartheta}, \tilde{\omega}\} \subset [-\pi, \pi]$ *for which the inequalities*

$$-(R-\varepsilon) f \sin\tilde{\omega} \leqq_K x\cos\tilde{\vartheta} - y\sin\tilde{\vartheta} \leqq_K (R-\varepsilon) f \sin\tilde{\omega}, \tag{2.56}$$

$$-(R-\varepsilon) f \cos\tilde{\omega} \leqq_K x\sin\tilde{\vartheta} + y\cos\tilde{\vartheta} \leqq_K (R-\varepsilon) f \cos\tilde{\omega} \tag{2.57}$$

would be satisfied.

Proof. Let us fix some $\{x, y\} \subset X_K(f)$ and define R by (2.53). Then

$$R = \inf_{t\in[-\pi,\pi]} r_t, \tag{2.58}$$

where $r_t := n_{\hat{K},\hat{f}}(\mathrm{e}^{\mathrm{i}t}(x + \mathrm{i}y))$ for all $t \in [-\pi, \pi]$. By virtue of Lemma 2.32, we have $0 \leq R < +\infty$.

Taking Lemma 2.36 and formula (2.38) into account, we conclude that, with any $t \in [-\pi, \pi]$, one can associate an $\omega_t \in [-\pi, \pi]$ for which

$$-r_t f \sin\omega_t \leqq_K x\cos t - y\sin t \leqq_K r_t f \sin\omega_t \tag{2.59}$$

and

$$-r_t f \cos\omega_t \leqq_K y\cos t + x\sin t \leqq_K r_t f \cos\omega_t. \tag{2.60}$$

By virtue of (2.58), there exists a sequence $(t_m)_{m=1}^{+\infty} \subset [-\pi, \pi]$ such that

$$\lim_{m\to+\infty} r_{t_m} = R. \tag{2.61}$$

Being bounded, this sequence contains a subsequence convergent to a certain $\vartheta_* \in [-\pi, \pi]$. We can assume, without loss of generality, that such a subsequence has already been selected and, thus, in addition to (2.61), we have

$$\lim_{m\to+\infty} t_m = \vartheta_*. \tag{2.62}$$

On the other hand, the sequence $(\omega_{t_m})_{m=1}^{+\infty} \subset [-\pi, \pi]$ is also bounded and, therefore, there exists a sequence $(m_j)_{j=1}^{+\infty} \subset \mathbb{N}$ such that $\lim_{j\to+\infty} \omega_{t_{m_j}} = \omega_*$ with a certain $\omega_* \in [-\pi, \pi]$. Setting $t = t_{m_j}$ in (2.59) and (2.60), passing to the limit as j tends to $+\infty$, and using (2.61), (2.62), and the fact that K is a closed set, we arrive at relations (2.54) and (2.55) with the above values of ϑ_* and ω_*.

Assume now that $R > 0$ and relations (2.56) and (2.57) are satisfied with some ε, $0 \leq \varepsilon < R$, and $\{\tilde{\vartheta}, \tilde{\omega}\} \subset [-\pi, \pi]$. Due to formula (2.43) of Lemma 2.35 and equality (2.38) from the proof of Lemma 2.32, relations (2.56) and (2.57) imply that $r_t \leq R - \varepsilon$ for all $t \in [-\pi, \pi]$, whence the estimate

$$\inf_{t\in[-\pi,\pi]} r_t \leq R - \varepsilon \tag{2.63}$$

follows. However, by virtue of (2.58), inequality (2.63) yields $R \leq R - \varepsilon$ and, therefore, $\varepsilon = 0$. □

The property described by Lemma 2.37 is also true for functional (2.50).

Lemma 2.43. *Let $f \in X$ and $z \in \hat{X}$. Then $R_{\hat{K},\hat{f}}(z) = 0$ if, and only if $z \in K^{\diamond} + \mathrm{i}K^{\diamond}$.*

Proof. The inclusion $z \in K^{\diamond} + \mathrm{i}K^{\diamond}$ means that the element z is 0-measurable with respect to $\hat{K}$. In this case, Lemma 2.32 guarantees that so does the element $e^{\mathrm{i}t}z$ with any t from $[-\pi, \pi]$ and, hence, by Lemma 2.37, it follows that $n_{\hat{K},\hat{f}}(e^{\mathrm{i}t}z) = 0$ for all $t \in [-\pi, \pi]$. Relation (2.50) then yields $R_{\hat{K},\hat{f}}(z) = 0$.

Conversely, if $R_{\hat{K},\hat{f}}(x + \mathrm{i}y) = 0$, then, by Lemma 2.42, there exists some ϑ from $[-\pi, \pi]$ such that

$$0 \leqq_K x\cos\vartheta - y\sin\vartheta \leqq_K 0,$$
$$0 \leqq_K x\sin\vartheta + y\cos\vartheta \leqq_K 0.$$

According to formula (2.38), this means that the element $e^{\mathrm{i}\vartheta}(x + \mathrm{i}y)$ is 0-measurable with respect to $\hat{K}$ and, thus, by Lemma 2.32, so does the element $x + \mathrm{i}y$. □

2.6. Linear operators vanishing on the blade of a wedge. For the sake of brevity, we introduce the following definition [11].

Definition 2.44. We say that an eigenvalue λ of a bounded linear operator $A : X \to X$ is *substantial* with respect to the wedge K (or, shortly, *K-substantial*) if λ is non-zero and at least one eigenvector not belonging to $K^{\diamond} + \mathrm{i}K^{\diamond}$ corresponds to it.

As usual (see, e. g., [5]), by a complex eigenvalue $\lambda \in \mathbb{C}$ of a bounded linear operator $A : X \to X$ acting in a real Banach space X, the eigenvalue of its complexification $\hat{A} = A + \mathrm{i}A : \hat{X} \to \hat{X}$ is meant, where

$$\hat{A}(x + \mathrm{i}y) := Ax + \mathrm{i}Ay \tag{2.64}$$

for all $\{x, y\} \subset X$.

Example 2.45. All the eigenvalues of a bounded linear operator $A : X \to X$ are substantial with respect to an arbitrary cone in X.

We devote our present study mostly to the linear operators $A : X \to X$ vanishing on the blade of a proper wedge K, i. e., such that

$$K^{\diamond} \subseteq \ker A. \tag{2.65}$$

Example 2.46. If K is a cone, then condition (2.65) is satisfied in an obvious way for every linear operator $A : X \to X$.

In the general case, the restrictiveness of condition (2.65) imposed on the operator A grows with the “width” of $K^{\diamond}$.

Example 2.47. Let us consider the set

$$C_{\Omega,\sigma}([a, b], \mathbb{R}) = \{x \in C([a, b], \mathbb{R}) \mid \sigma x([a, b] \setminus \Omega) \subseteq [0, +\infty)\},$$

where Ω is a certain subset of $[a, b]$ such that $[a, b] \setminus \Omega$ is closed, and $\sigma \in \{-1, 1\}$. The set $C_{\Omega,\sigma}([a, b], \mathbb{R})$ is obviously a closed wedge in the Banach space $C([a, b], \mathbb{R})$ of all the continuous scalar functions on the bounded interval $[a, b]$. This wedge is solid because, as one can show, its interior is constituted by the continuous functions $x : [a,b] \to \mathbb{R}$ such that $\sigma x([a,b] \setminus \Omega) \subseteq (0,+\infty)$.

Consider the operator $A : C([a,b],\mathbb{R}) \to C([a,b],\mathbb{R})$ given by the formula

$$(Ax)(t) = \int_\tau^t k(t,s)\, x(\omega(s))\, ds, \qquad t \in [a,b], \tag{2.66}$$

in which $\omega : [a,b] \to [a,b]$ is a measurable function, whereas the function $k : [a,b] \times [a,b] \to \mathbb{R}$ is continuous in the first variable and Lebesgue integrable in the second one. The operator A vanishes on the blade of the wedge $C_{\Omega,\sigma}([a,b],\mathbb{R})$ when ω satisfies the condition

$$\omega([a,b]) \subseteq [a,b] \setminus \Omega. \tag{2.67}$$

Indeed, the blade of $C_{\Omega,\sigma}([a,b],\mathbb{R})$ consists of those continuous functions $x : [a,b] \to \mathbb{R}$ such that

$$x(t) = 0 \quad \text{for all } t \in [a,b] \setminus \Omega. \tag{2.68}$$

If ω is such that condition (2.67) holds, then Ax is equal identically to zero for every function x satisfying condition (2.68), i. e., the relation $x \in (C_{\Omega,\sigma}([a,b],\mathbb{R}))^\diamond$ implies that $Ax = 0$. This means that (2.65) is true for $K = C_{\Omega,\sigma}([a,b],\mathbb{R})$ and A given by (2.66).

Our interest to the property described by condition (2.65) is motivated by the following statement.

Lemma 2.48. *Assume that $A : X \to X$ is a linear operator vanishing on the blade of a wedge $K \subseteq X$. Then every non-zero eigenvalue of A is K-substantial.*

Proof. Let λ be an arbitrary non-zero eigenvalue of A. Then there exists some non-zero element w from $\hat{X}$ such that

$$\lambda w = \hat{A} w. \tag{2.69}$$

Assume that, on the contrary, λ is not K-substantial and, therefore, according to Definition 2.44, every eigenvector w in (2.69) belongs to $K^\diamond + \mathrm{i}K^\diamond$. By virtue of inclusion (2.65), this yields $\hat{A}w = 0$, and, hence, by (2.69), $w = 0$, which is impossible because w is an eigenvector of $\hat{A}$. The contradiction obtained proves our lemma. □

Assumption (2.65) may seem to be unnecessarily strong because, in fact, it guarantees that not only *some* eigenvectors corresponding to non-zero eigenvalues of A do not belong to the blade of $\hat{K}$ but *all* such eigenvectors possess this property. Note however that, in the theorems of Sections 3 and 4, condition (2.65) cannot be dropped even in the two-dimensional case (see Example 5.1).

3. Upper bounds for *K*-substantial eigenvalues

The following theorem provides an upper bound for K-substantial eigenvalues of a sufficiently wide class of linear operators in a real Banach space X.

Theorem 3.1. *Let K be a proper wedge in X and $A_1 : X \to X$, $A_2 : X \to X$ be bounded linear operators such that*

$$A_1(K) \cup A_2(K) \subseteq K. \tag{3.1}$$

Assume also that the relation

$$A_1 f + A_2 f \leqq_K \alpha f \tag{3.2}$$

is true with some $\alpha \in [0, +\infty)$ and $f \in X$ for which

$$f \mathrel{\varepsilon\!\!-}_{K;H} 0, \tag{3.3}$$

where H is a certain linear manifold in X satisfying the inclusion

$$H \supseteq \operatorname{im}(A_1 - A_2). \tag{3.4}$$

Then every K-substantial eigenvalue λ of the operator $A_1 - A_2$ admits the estimate

$$|\lambda| \leq \alpha. \tag{3.5}$$

In (3.1), (3.4), and similar relations, we use the standard notation

$$A(M) := \{Ax \mid x \in M\},$$

where $M \subseteq X$. Recall that the binary relation "$\mathrel{\varepsilon\!\!-}_{K;H}$" appearing in (3.3) is introduced on X by Definition 2.16.

Proof of Theorem 3.1. Let $\lambda = \varrho\, e^{i\vartheta}$, $\varrho \in (0, +\infty)$, be a K-substantial eigenvalue of the complexification $\hat{A} = \hat{A}_1 - \hat{A}_2$ of the operator

$$A := A_1 - A_2. \tag{3.6}$$

In view of Definition 2.44 and equality (2.23), there exists an element $w = x + iy$ such that $\{x, y\} \subset X$, $\{x, y\} \not\subset K^{\diamond}$, and equality (2.69) holds.

We divide the present proof into several parts.

Claim 1. *The element w is $\hat{f}$-measurable with respect to $\hat{K}$.*

Indeed, equality (2.69) means that

$$\varrho w = e^{-i\vartheta} \hat{A} w. \tag{3.7}$$

According to formulae (2.38) and (2.64), we have

$$e^{-i\vartheta} \hat{A} w = Ax \cos\vartheta + Ay \sin\vartheta + i(Ay \cos\vartheta - Ax \sin\vartheta)$$

and, therefore, (3.7) can be rewritten as the system

$$\varrho x = Ax \cos\vartheta + Ay \sin\vartheta, \tag{3.8}$$

$$\varrho y = Ay \cos\vartheta - Ax \sin\vartheta. \tag{3.9}$$

By virtue of assumption (3.4), it follows from (3.8) and (3.9) that x and y both lie in H (to prove this, it suffices to use the linearity of the set H). However, according to Definition 2.12, condition (3.3) means that all the elements from H are f-measurable with respect to K and, hence, by Lemma 2.31, the element $x+\mathrm{i}y$ is $(f+\mathrm{i}f)$-measurable with respect to $\hat{K}$.

CLAIM 2. *The number*

$$R := R_{\hat{K},\hat{f}}(x+\mathrm{i}y) \tag{3.10}$$

is strictly positive.

Indeed, by Claim 1 and Lemma 2.41, the right-hand side of (3.10) is a finite number. Since $\{x, y\} \not\subset K^{\diamond}$, Lemma 2.37 yields $R > 0$.

CLAIM 3. *The elements x and y satisfy the equalities*

$$Ax = \varrho\,(x\cos\vartheta - y\sin\vartheta), \tag{3.11}$$

$$Ay = \varrho\,(x\sin\vartheta + y\cos\vartheta). \tag{3.12}$$

According to formula (2.38), system (3.11), (3.12) is an equivalent form of relation (2.69) satisfied by w.

CLAIM 4. *There exist some ω_* and t_* from $[-\pi,\pi]$ such that* (2.54) *and* (2.55) *are true for x and y, and there do not exist any $\{\tilde{\omega},\tilde{\vartheta}\} \subset [-\pi,\pi]$ for which the relations*

$$-rf\sin\tilde{\omega} \leqq_K x\cos\tilde{\vartheta} - y\sin\tilde{\vartheta} \leqq_K rf\sin\tilde{\omega}, \tag{3.13}$$

$$-rf\cos\tilde{\omega} \leqq_K x\sin\tilde{\vartheta} + y\cos\tilde{\vartheta} \leqq_K rf\cos\tilde{\omega} \tag{3.14}$$

would be satisfied with some $r \in (0, R)$.

This statement is an immediate consequence of formula (3.10) and Lemma 2.42.

CLAIM 5. *There is an Ω from $[-\pi,\pi]$ such that the relations*

$$-\alpha Rf\sin\Omega \leqq_K Ax\cos\vartheta_* - Ay\sin\vartheta_* \leqq_K \alpha Rf\sin\Omega, \tag{3.15}$$

$$-\alpha Rf\cos\Omega \leqq_K Ax\sin\vartheta_* + Ay\cos\vartheta_* \leqq_K \alpha Rf\cos\Omega, \tag{3.16}$$

are true, where A is given by (3.6).

In view of assumption (3.1), both operators A_1 and A_2 preserve order inequalities. Therefore, relations (2.54) and (2.55), together with Lemma 2.6, yield

$$-RA_jf\sin\omega_* \leqq_K \sigma[A_jx\cos\vartheta_* - A_jy\sin\vartheta_*] \leqq_K RA_jf\sin\omega_*, \tag{3.17}$$

$$-RA_jf\cos\omega_* \leqq_K \varkappa[A_jx\sin\vartheta_* + A_jy\cos\vartheta_*] \leqq_K RA_jf\cos\omega_* \tag{3.18}$$

for all $j = 1, 2$ and $\{\sigma, \varkappa\} \subset \{-1, 1\}$. Summing the two relations obtained from (3.17) with $j = 1$, $\sigma = 1$ and $j = 2$, $\sigma = -1$, respectively, we obtain

$$-R(A_1+A_2)f\sin\omega_* \leqq_K (A_1-A_2)x\cos\vartheta_* - (A_1-A_2)y\sin\vartheta_* \leqq_K R(A_1+A_2)f\sin\omega_*,$$

i. e.,

$$-R(A_1+A_2)f\sin\omega_* \leqq_K Ax\cos\vartheta_* - Ay\sin\vartheta_* \leqq_K R(A_1+A_2)f\sin\omega_*. \tag{3.19}$$

In a similar manner, putting in (3.18) $j = 1$, $\varkappa = 1$ and $j = 2$, $\varkappa = -1$ and summing the resulting two relations, we get

$$-R(A_1+A_2)f\cos\omega_* \leqq_K Ax\sin\vartheta_* + Ay\cos\vartheta_* \leqq_K R(A_1+A_2)f\cos\omega_*. \tag{3.20}$$

Let us consider the following four cases.

Case 1. $\sin\omega_* \geq 0$ *and* $\cos\omega_* \geq 0$.

Using assumption (3.2) in relations (3.19), (3.20) and putting $\Omega := \omega_*$, we arrive immediately at (3.15), (3.16).

Case 2. $\sin\omega_* \geq 0$ *and* $\cos\omega_* < 0$.

Recall that, by assumption, $f \geqq_K 0$ and, due to condition (3.1), $A_1 f + A_2 f \geqq_K 0$. In view of assertion (iii) of Lemma 2.7, relation (3.20) and Claim 2 imply that, in this case,

$$0 \leqq_K Ax\sin\vartheta_* + Ay\cos\vartheta_* \leqq_K 0. \tag{3.21}$$

whereas the term $(A_1 + A_2)f\sin\omega_*$, by virtue of (3.2), admits the estimate

$$(A_1 + A_2)f\sin\omega_* \leqq_K \alpha f.$$

Therefore, (3.15) and (3.16) are satisfied with $\Omega := \frac{\pi}{2}$.

Case 3. $\sin\omega_* < 0$ *and* $\cos\omega_* \geq 0$.

Relation (3.19) now yields

$$0 \leqq_K Ax\cos\vartheta_* - Ay\sin\vartheta_* \leqq_K 0 \tag{3.22}$$

and, similarly to Case 2, we conclude that (3.15) and (3.16) hold with $\Omega := 0$.

Case 4. $\sin\omega_* < 0$ *and* $\cos\omega_* < 0$.

A reasoning analogous to those presented above show that, in this case, system (3.19), (3.20) has form (3.21), (3.22) and, therefore, relations (3.15) and (3.16) are satisfied both with $\Omega = \frac{\pi}{2}$ and $\Omega = 0$. This proves our Claim 5.

Having established the facts above, we now turn to the proof of estimate (3.5).

According to Claim 3, the components x and y of the eigenvector w of $\hat{A}$ satisfy equalities (3.11) and (3.12). Therefore,

$$\begin{aligned}
Ax\cos\vartheta_* - Ay\sin\vartheta_* &= \varrho(x\cos\vartheta - y\sin\vartheta)\cos\vartheta_* - \varrho(x\sin\vartheta + y\cos\vartheta)\sin\vartheta_* \\
&= \varrho[\cos\vartheta\sin\vartheta_* - \sin\vartheta\sin\vartheta_*]x \\
&\quad - \varrho[\sin\vartheta\cos\vartheta_* + \cos\vartheta\sin\vartheta_*]y \\
&= \varrho x\cos(\vartheta + \vartheta_*) - \varrho y\sin(\vartheta + \vartheta_*)
\end{aligned}$$

and, similarly,

$$\begin{aligned}
Ax\sin\vartheta_* + Ay\cos\vartheta_* &= \varrho(x\cos\vartheta - y\sin\vartheta)\sin\vartheta_* + \varrho(x\sin\vartheta + y\cos\vartheta)\cos\vartheta_* \\
&= \varrho[\cos\vartheta\sin\vartheta_* + \sin\vartheta\cos\vartheta_*]x \\
&\quad + \varrho[\cos\vartheta\cos\vartheta_* - \sin\vartheta\sin\vartheta_*]y \\
&= \varrho x\sin(\vartheta + \vartheta_*) + \varrho y\cos(\vartheta + \vartheta_*).
\end{aligned}$$

Applying these formulae to the corresponding expressions in (3.15) and (3.16) and taking the inequality $\varrho > 0$ into account, we obtain

$$-\frac{\alpha R}{\varrho} f \sin\Omega \leqq_K x\cos(\vartheta + \vartheta_*) - y\sin(\vartheta + \vartheta_*) \leqq_K \frac{\alpha R}{\varrho} f \sin\Omega, \tag{3.23}$$

$$-\frac{\alpha R}{\varrho} f \cos\Omega \leqq_K x\sin(\vartheta + \vartheta_*) + y\cos(\vartheta + \vartheta_*) \leqq_K \frac{\alpha R}{\varrho} f \cos\Omega. \tag{3.24}$$

System (3.23), (3.24), obviously, has form (3.13), (3.14) with $\tilde{\vartheta} := \vartheta + \vartheta_*$, $\tilde{\omega} := \Omega$, and $r := \alpha R/\varrho$. In view of Claim 4, it now follows that

$$\frac{\alpha R}{\varrho} \geq R,$$

whence, by Claim 2, we arrive at the inequality $\varrho \leq \alpha$. Recalling that $\varrho = |\lambda|$, we conclude that the required estimate (3.5) holds. □

The assumption that f should lie in K and be different from zero in Theorem 3.1 is motivated by Proposition 2.22.

Remark 3.2. A linear operator $A : X \to X$ admits representation in form (1.4), where A_1 and A_2 are linear mappings preserving K, if and only if there exists a linear operator $B : X \to X$ such that

$$B(K) \subseteq K \tag{3.25}$$

and

$$Ax \leqq_K Bx \qquad \text{for all } x \in K. \tag{3.26}$$

Indeed, (1.4) implies that

$$Ax \leqq_K A_1 x \leqq_K A_1 x + A_2 x$$

for all x such that $x \geqq_K 0$ and, therefore, one can set $B := A_1 + A_2$. Conversely, it follows from (3.25) and (3.26) that the operator $A_2 := B - A$ preserves the wedge K and, thus, it remains to put $A_1 := B$ in (1.4).

It should be noted that, in the case where the space X is infinite-dimensional, one cannot claim that every bounded linear operator $A : X \to X$ admits representation (1.4) with bounded linear mappings $A_1 : X \to X$ and $A_2 : X \to X$ preserving K. In particular, in the case where K is a cone which does not possess the property of normality, the classical Theorem 2 of [8] ensures the existence of a continuous (even finite-dimensional) linear operator $A : X \to X$ that cannot be represented in form (1.4) with bounded $A_k : X \to X$, $k = 1, 2$, satisfying condition (3.1).

4. A THEOREM ON THE SPECTRAL RADIUS OF A COMPACT OPERATOR

Theorem 3.1 of Section 3 can be applied to prove the following statement which seems to be useful in studies of the solvability of various linear equations with compact operators.

Theorem 4.1. *Let X be a Banach space over the field $\mathbb{R}$, K be a proper wedge in X, and $A_1 : X \to X$, $A_2 : X \to X$ be completely continuous linear operators leaving the wedge K invariant and satisfying the condition*

$$K^{\diamond} \subseteq \ker(A_1 - A_2)\,. \tag{4.1}$$

In addition, assume that relation (3.2) *is satisfied with some constant $\alpha \in [0, +\infty)$ and element $f \in X$ possessing the property*

$$f \mathrel{\varepsilon\!\!-}_{K;\,\mathrm{im}\,(A_1-A_2)} 0. \tag{4.2}$$

Then the spectral radius of the operator $A_1 - A_2$ admits the estimate

$$r\,(A_1 - A_2) \le \alpha \tag{4.3}$$

Proof. It follows from the Riesz–Schauder theory (see, e. g., [2]) that, due to the complete continuity of the operator $A_1 - A_2$, its spectrum consists of countably many eigenvalues.

It is easy to see that condition (4.2) implies the existence of a linear manifold $H \subseteq X$ satisfying inclusion (3.4) and such that relation (3.3) holds for the element f (one can put $H := \mathrm{im}\,(A_1 - A_2)$). Assumption (4.1), by virtue of Lemma 2.48, implies that every non-zero eigenvalue of the operator $A_1 - A_2$ is K-substantial in the sense of Definition 2.44. Therefore, under the conditions assumed, Theorem 3.1 can be used.

Application of Theorem 3.1 guarantees that an arbitrary non-zero eigenvalue λ of the operator $A_1 - A_2$ admits estimate (3.5). Considering the least upper bound of $|\lambda|$ in the left-hand side of relation (3.5) with respect to all the non-zero eigenvalues λ of $A_1 - A_2$, we arrive immediately at inequality (4.3). □

Recall that the binary relation "$\mathrel{\varepsilon\!\!-}_{K;\,H}$" appearing in condition (4.2) has been introduced by Definition 2.16.

Remark 4.2. The complete continuity of $A_1 - A_2$ in the proof of Theorem 4.1 is used only to guarantee that this operator has discrete spectrum (in fact, an upper bound for the discrete spectrum is established under the conditions specified). Note also that, in the case where K is a solid and normal cone, the assertion of Theorem 4.3 can also be proved by using Theorems 5.3 and 5.5 of [7].

Corollary 4.3. *Let $K \subset X$ be a wedge and $A_1 : X \to X$, $A_2 : X \to X$ be completely continuous linear operators leaving the wedge K invariant, satisfying condition* (4.1) *and such that relation* (3.2) *is true with some $\alpha \in [0, +\infty)$ and element $f \in X$ possessing the property*

$$f \mathrel{\varepsilon\!\!-}_{K} 0. \tag{4.4}$$

Then the spectral radius of the operator $A_1 - A_2$ admits estimate (4.3)*.*

Proof. In view of Lemma 2.13, assumption (4.4) ensures that f possesses property (4.2). Therefore, in order to obtain the required statement, it is sufficient to apply Theorem 4.1. □

For the sake of completeness we note that, by Lemma 2.19, assumption (4.4) is satisfied, in particular, in the classical case where K is solid and the element f belongs its interior.

Remark 4.4. Condition (4.1) for the operator $A : X \to X$ in Theorem 4.1 and Corollary 4.3 is satisfied automatically if K is a cone.

Corollary 4.5. *If $A : X \to X$ is a completely continuous linear operator leaving invariant a wedge K and, moreover, satisfying condition* (2.65) *and the relation*

$$Af \leqq_K \alpha f \tag{4.5}$$

with some constant $\alpha \in [0, +\infty)$ and element $f \in X$ possessing property (4.4)*, then the estimate*

$$r(A) \le \alpha \tag{4.6}$$

is true.

Proof. Corollary 4.5 is an immediate consequence of Corollary 4.3 in the case where $A_1 = A$ and $A_2 = 0$. □

Remark 4.6. A number of theorems that, for a positive operator A preserving a cone $K \subset X$, allow one to deduce estimate (4.6) from relation (4.5) are well-known, e. g., from [6, 7, 15]. Apart from those relying on the f-positivity of the operator, in the results mentioned, the element f appearing in (4.5) is supposed either to lie in the interior of K (Theorem 16.2 (b) of [6]) or to be its quasi-interior element (Theorem 16.2 (a) from [6]). It is important to note that condition (4.2), which also expresses a certain kind of the strong positivity* of the element f and, under these circumstances, replaces the two conditions mentioned, depends on the image space of the operator $A_1 - A_2$.

Corollary 4.3 also implies an analogue of Corollary 4.5 for the "negative" operators.

Corollary 4.7. *Let $A : X \to X$ be a completely continuous linear operator such that $A(-K) \subseteq K$ and, moreover, the relation*

$$Af \geqq_K -\alpha f$$

be satisfied with some $\alpha \in [0, +\infty)$ and $f \in X$ possessing property (4.4)*. Then the spectral radius of A admits estimate* (4.6).

Proof. It suffices to put $A_1 = 0$ and $A_2 = -A$ in Corollary 4.3. □

Corollary 4.10 established below is an example of application of Theorem 4.1 with H different from X. Prior to its formulation, we introduce a definition.

*See Section 2.3.

Definition 4.8. Let K be a wedge in X and f be an element from X. An operator $A : X \to X$ is said to be *f-bounded* with respect to K if, for every $x \in X$, there exists a constant $\beta \in (-\infty, +\infty)$ such that

$$-\beta f \leqq_K Ax \leqq_K \beta f.$$

In other words, A is f-bounded if the element Ax is f-measurable for all x. It follows from Proposition 2.11 that, in the pathological cases where $0 \leqq_K f \leqq_K 0$ or f is incomparable with zero, every operator A which is f-bounded with respect to K has the property $\operatorname{im} A \subseteq K^\diamond$.

Remark 4.9. An operator f-bounded with respect to K is, in particular, f-bounded from above in the sense of the definition from [5], Chapt. 2, §1. The converse statement is not true.

Corollary 4.10. *Let $f \geqq_K 0$ be a given element and $A_k : X \to X$, $k = 1, 2$, be completely continuous linear operators leavfing the wedge K invariant, satisfying condition* (4.1)*, and f-bounded with respect to K.*

Then the existence of a non-negative constant α for which relation (3.2) *is satisfied implies estimate* (4.3) *for the spectral radius of the operator $A_1 - A_2$.*

Proof. Setting

$$H := X_K(f), \tag{4.7}$$

we see that condition (3.4) is satisfied due to the f-boundedness of A_1 and A_2 with respect to K. By Lemma 2.14, the element f satisfies relation (2.8) and, hence,

$$f \,\underset{K;\,X_K(f)}{\multimap}\, 0.$$

Therefore, condition (3.3) holds with H given by (4.7), and it remains to apply Theorem 4.1. □

5. Importance of condition (4.1)

As is seen from the proof of Theorem 4.1, the applicability of statements on K-substantial eigenvalues is guaranteed by condition (4.1). It is natural to expect that estimating the spectrum of an operator on the base of assumptions of type (3.2) is not possible any more if one admits the existence of a non-zero eigenvalue which is not K-substantial, and imposes no additional conditions on A.

The following example [11] shows that the assumption on the fulfilment of condition (4.1)in Theorem 4.1 is essential and, generally speaking, cannot be omitted.

Example 5.1. Let us consider the set

$$K = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : x_1 \geq 0,\ x_2 \in \mathbb{R} \right\}. \tag{5.1}$$

Obviously, K is a solid wedge in $X := \mathbb{R}^2$, and the blade of K has the form

$$K^\diamond = \left\{ \begin{pmatrix} 0 \\ c \end{pmatrix} : c \in \mathbb{R} \right\}.$$

It is not difficult to verify that the linear operator A given by the matrix $A = \left(\begin{smallmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{smallmatrix}\right)$ leaves invariant the set K given by (5.1) if, and only if

$$a_{11} \geq 0, \qquad a_{12} = 0. \tag{5.2}$$

Furthermore, one can show that, under condition (5.2), A vanishes on $K^{\diamond}$ if, and only if

$$a_{22} = 0. \tag{5.3}$$

A vector $f = \binom{f_1}{f_2}$ belongs to the interior of K if, and only if

$$f_1 > 0. \tag{5.4}$$

whereas the corresponding condition (4.5) is equivalent to the inequality $\alpha f_1 \geq a_{11} f_1$, which, in view of (5.4), means that

$$\alpha \geq a_{11}. \tag{5.5}$$

If condition (2.65) or, which is the same in our case, equality (5.3) is violated, then assumption (4.5), generally speaking, cannot guarantee the validity of the estimate $r(A) \leq \alpha$ for the spectral radius of A. Indeed, it is clear from (5.2) that $r(A) = \max\{a_{11}, |a_{22}|\}$ and, hence,

$$r(A) \geq |a_{22}|. \tag{5.6}$$

However, if the inequalities

$$|a_{22}| > \alpha \geq a_{11} \geq 0 \tag{5.7}$$

hold, then the assertion of Corollary 4.5 in the case considered would have the form $r(A) \leq \alpha$, which is impossible in view of (5.6) and (5.7).

Thus, condition (2.65) Corollary 4.5 (and, therefore, condition (4.1) in Theorem 4.1), generally speaking, cannot be dropped.

6. Unique solvability of linear equations

Theorem 4.1 allows one to obtain efficient conditions under which the linear equation

$$x = A_1 x - A_2 x + q, \tag{6.1}$$

where A_1 and A_2 are linear operators, possesses a unique solution for an arbitrary element q from X.

Corollary 6.1. *Let X be a real Banach space, $K \subset X$ be a wedge, and $A_i : X \to X$, $i = 1, 2$, be completely continuous linear operators leaving K invariant and satisfying condition* (4.1)*. In addition, assume that relation* (3.2) *is satisfied with some constant $\alpha \in [0, 1)$ and element $f \in X$ possessing property* (4.2).

Then equation (6.1) *is uniquely solvable for arbitrary $q \in X$, and its solution x is represented by the convergent Neumann series*

$$x = \sum_{k=0}^{+\infty} (A_1 - A_2)^k q. \tag{6.2}$$

Proof. It suffices to notice that, by virtue of Theorem 4.1, the conditions assumed guarantee that the spectrum of the operator $A_1 - A_2$ is contained in the closure of the unit disk in $\mathbb{C}$. $\square$

In the cases where the fact of convergence of series (6.2) is less important, one may prefer to use the following

Corollary 6.2. *Let K be a proper wedge in a real Banach space X and $A_i : X \to X$, $i = 1, 2$, be bounded linear operators leaving K invariant and satisfying condition* (4.1)*. Assume that relation* (3.2) *is true, where $\alpha \in [0, 1)$ and $f \in X$ is an element relation* (4.2) *is satisfied.*

Then the homogeneous equation

$$x = A_1 x - A_2 x \tag{6.3}$$

has no non-trivial solutions. If, moreover, the operators A_1 and A_2 are such that $A_1 - A_2$ is a Fredholm operator of index 0*, then equation* (6.1) *is uniquely solvable for an arbitrary $q \in X$.*

Proof. In view of assumption (4.1) and Lemma 2.48, it follows from Theorem 3.1 that operator (3.6) has no eigenvalues outside the open interval $(-1, 1)$ and, in particular, the number 1 is not an eigenvalue for the operator mentioned. Therefore, zero is the unique solution of the homogeneous equation (6.3). The unique solvability of equation (6.1) for any q is guaranteed by the assumption that $A_1 - A_2$ is a Fredholm operator of index 0. $\square$

Corollary 6.2 allows one to obtain the following statement.

Corollary 6.3. *Let K be a proper wedge in a real Banach space X and $A_i : X \to X$, $i = 1, 2$, be bounded linear operators leaving K invariant, satisfying the condition*

$$K^{\diamond} \subseteq \ker A_1 \cap \ker A_2. \tag{6.4}$$

Assume also that relation (3.2) *is true, where $\alpha \in [0, 1)$ is a constant and $f \in X$ is an element possessing property* (4.2)*. Then the homogeneous equation*

$$x = \sigma_1 A_1 x + \sigma_2 A_2 x \tag{6.5}$$

has no non-trivial solutions for any values $\{\sigma_1, \sigma_2\} \subset \{-1, 1\}$.

If, in addition, for some $\{\sigma_1, \sigma_2\} \subset \{-1, 1\}$, the mapping $\sigma_1 A_1 + \sigma_2 A_2$ is a Fredholm operator of index 0*, then the equation*

$$x = \sigma_1 A_1 x + \sigma_2 A_2 x + q, \tag{6.6}$$

is uniquely solvable for an arbitrary $q \in X$.

Proof. Let us define the operators $\tilde{A}_i : X \to X$, $i = 1, 2$, by putting

$$\tilde{A}_1 := \frac{1 + \sigma_1}{2} A_1 + \frac{1 + \sigma_2}{2} A_2 \tag{6.7}$$

and

$$\tilde{A}_2 := \frac{1-\sigma_1}{2} A_1 + \frac{1-\sigma_2}{2} A_2. \tag{6.8}$$

One can verify that the relations

$$\sigma_1 A_1 + \sigma_2 A_2 = \tilde{A}_1 - \tilde{A}_2$$

and

$$A_1 + A_2 = \tilde{A}_1 + \tilde{A}_2 \tag{6.9}$$

are true. It is also easy to see from (6.7) and (6.8) that both operators $\tilde{A}_1$ and $\tilde{A}_2$ leave invariant the wedge K.

Assumption (6.4) guarantees the fulfilment of the inclusion

$$K^{\diamond} \subseteq \ker(\tilde{A}_1 - \tilde{A}_2).$$

Moreover, by virtue of (6.9), condition (3.2) can be rewritten as

$$\tilde{A}_1 f + \tilde{A}_2 f \leqq_K \alpha f.$$

We have thus shown that Corollary 6.2 can be applied with A_1 and A_2 replaced by $\tilde{A}_1$ and $\tilde{A}_2$, respectively. $\square$

7. An example of a linear integral equation

We illustrate the idea of the results above on an example. Let us consider the problem on the continuous solutions of the equation

$$x(t) = \int_0^1 h(t,s)\, x(\omega(s))\, ds + q(t), \quad t \in [0,1], \tag{7.1}$$

where $q : [0,1] \to \mathbb{R}$ is continuous, the function $\omega : [0,1] \to [0,1]$ is measurable, $h(t,\cdot) : [0,1] \to \mathbb{R}$ is Lebesgue integrable for all $t \in [0,1]$, and $h(\cdot, s) : [0,1] \to \mathbb{R}$ is continuous for almost every s from $[0,1]$.

7.1. The general argument deviation. The techniques of Section 4 allow one toobtain the following theorem.

Theorem 7.1. *Let there exist a non-negative continuous function $\psi : [0,1] \to \mathbb{R}$ such that*

$$\operatorname*{vrai\,max}_{t\in[0,1]\setminus\Gamma_{\psi,\omega}} \frac{1}{\psi(\omega(t))} \int_0^1 |h(\omega(t),s)|\, ds < +\infty \tag{7.2}$$

and

$$h(\omega(t),s) = 0 \quad \textit{for a. e. } t \in \Gamma_{\psi,\omega} \textit{ and a. e. } s \in [0,1], \tag{7.3}$$

where

$$\Gamma_{\psi,\omega} := \{t \in [0,1] \mid \psi(\omega(t)) = 0\}. \tag{7.4}$$

Moreover, assume that the inequality

$$\operatorname*{vrai\,max}_{t\in[0,1]\setminus\Gamma_{\psi,\omega}} \frac{1}{\psi(\omega(t))}\int_0^1 |h(\omega(t),s)|\,\psi(\omega(s))\,ds < 1 \tag{7.5}$$

is satisfied.

Then equation (7.1) *possesses a unique solution for any continuous function* $q : [0,1] \to \mathbb{R}$.

Proof. Equation (7.1) can obviously be rewritten in form (6.1), where the operators A_i, $i = 1,2$, in the Banach space $X := C([0,1],\mathbb{R})$ of all the continuous scalar functions on $[0,1]$ are introduced by the formulae

$$(A_i x)(t) := \int_0^1 \max\{(-1)^{i+1}h(t,s),0\}\,x(\omega(s))\,ds, \quad t\in[0,1],\ i=1,2, \tag{7.6}$$

for any continuous $x : [0,1] \to \mathbb{R}$. Clearly, each of these operators leaves invariant the wedge

$$\mathscr{K}_\omega := \{u \in C([0,1],\mathbb{R}) \mid u(\omega(t)) \geq 0 \text{ for a. e. } t \in [0,1]\}.$$

It is easy to show that mappings (7.6) are completely continuous linear operators possessing property (4.1) with respect to the wedge $K = \mathscr{K}_\omega$.

Let $\mathscr{C}_{\psi,\omega}$ be the set of all the continuous functions $x : [0,1] \to \mathbb{R}$ for which there exists a non-negative constant β_x such that the estimate

$$|x(\omega(t))| \leq \beta_x \psi(\omega(t))$$

is true for a. e. $t \in [0,1]$.

It is easy to see that $\mathscr{C}_{\psi,\omega}$ is a linear manifold in $C([0,1],\mathbb{R})$. Assumptions (7.2) and (7.3) and formula (7.6) imply that, for any x from $C([0,1],\mathbb{R})$, the estimate

$$|\,(A_1x)\,(\omega(t)) - (A_2x)\,(\omega(t))| \leq \psi(\omega(t))\,\Delta \max_{\xi\in[0,1]} |x(\xi)| \tag{7.7}$$

is true at a. e. t from $[0,1] \setminus \Gamma_{\psi,\omega}$, where

$$\Delta := \operatorname*{vrai\,max}_{t\in[0,1]\setminus\Gamma_{\psi,\omega}} \frac{1}{\psi(\omega(t))}\int_0^1 |h(\omega(t),s)|\,ds. \tag{7.8}$$

Indeed, let $x \in C([0,1],\mathbb{R})$ be arbitrary. According to (7.6), we obtain

$$\begin{aligned} |(A_1x)\,(\omega(t)) - (A_2x)\,(\omega(t))| &= \left|\int_0^1 h(\omega(t),s)\,x(\omega(s))\,ds\right| \\ &\leq \max_{\xi\in[0,1]} |x(\xi)| \int_0^1 |h(\omega(t),s)\,x(\omega(s))|\,ds \end{aligned} \tag{7.9}$$

for a. e. $t \in [0,1]$. For almost all $t \notin \Gamma_{\psi,\omega}$, we have $\psi(\omega(t)) \neq 0$ and, thus, by (7.8),

$$\int_0^1 |h(\omega(t),s)\,x(\omega(s))|\,ds \leq \Delta\,\psi(\omega(t)),$$

which, together with (7.9), yields (7.7). Moreover, assumption (7.3) guarantees that

$$(A_1 x)(\omega(t)) = (A_2 x)(\omega(t)) \tag{7.10}$$

for a. e. $t \in \Gamma_{\psi,\omega}$. Therefore, estimate (7.7) is true almost for all t from the entire interval $[0,1]$, not only those lying outside the set $\Gamma_{\psi,\omega}$.

This means that condition (3.4) is satisfied with the above definitions of A_1 and A_2 and $H := \mathscr{C}_{\psi,\omega}$.

Furthermore, the function ψ satisfies the condition

$$\psi \mathrel{\unicode{x22F2}}_{\mathscr{K}_\omega;\mathscr{C}_{\psi,\omega}} 0, \tag{7.11}$$

which means that (3.3) holds with $f := \psi$ and $K := \mathscr{K}_\omega$. Indeed, according to Definition 2.16, relation (7.11) means that, for any x from $\mathscr{C}_{\psi,\omega}$, there exists a constant $\beta \geq 0$ such that

$$|x(\omega(t))| \leq \beta\psi(\omega(t))$$

at almost every point t from the interval $[0,1]$. However, the above property is an immediate consequence of the definition of the set $\mathscr{C}_{\psi,\omega}$.

Finally, condition (7.3) and inequality (7.5) guarantee that

$$\int_0^1 |h(\omega(t),s)|\,\psi(\omega(s))\,ds \leq \alpha\psi(\omega(t)) \quad \text{for a. e. } t \in [0,1], \tag{7.12}$$

where

$$\alpha := \operatorname*{vrai\,max}_{t\in[0,1]\setminus\Gamma_{\psi,\omega}} \frac{1}{\psi(\omega(t))}\int_0^1 |h(\omega(t),s)|\,\psi(\omega(s))\,ds$$

and, by (7.5), $\alpha < 1$. In view of (7.6), relation (7.12) can be rewritten in form (3.2) for $K := \mathscr{K}_\omega$. Applying Theorem 4.1, we conclude that 1 is a regular value for the operator (3.6) corresponding to the given problem. □

The above theorem implies, for example, the following statement.

Corollary 7.2. *Assume that, for certain $\tau \in [0,1]$ and $\gamma \geq 0$, the functions $h : [0,1]^2 \to \mathbb{R}$ and $\omega : [0,1] \to [0,1]$ satisfy the conditions*

$$\operatorname*{vrai\,max}_{t\in[0,1]\setminus\omega^{-1}(\tau)} \frac{1}{|\omega(t)-\tau|^\gamma}\int_0^1 |\omega(s)-\tau|^\gamma\,|h(\omega(t),s)|\,ds < 1, \tag{7.13}$$

and

$$\operatorname*{vrai\,max}_{t\in[0,1]\setminus\omega^{-1}(\tau)} \frac{1}{|\omega(t)-\tau|^\gamma}\int_0^1 |h(\omega(t),s)|\,ds < +\infty. \tag{7.14}$$

In the case where

$$\operatorname{mes}\omega^{-1}(\tau) > 0, \tag{7.15}$$

assume, in addition, that

$$h(\tau,s) = 0 \quad \textit{for a. e. } s \in [0,1]. \tag{7.16}$$

Then equation (7.1) *has a unique solution for any continuous function* $q : [0,1] \to \mathbb{R}$.

Proof. It suffices to apply Theorem 7.1 with

$$\psi(t) := |t-\tau|^{\gamma}, \qquad t \in [0,1],$$

in which case set (7.4) is given by the formula

$$\Gamma_{\psi,\omega} = \{t \in [0,1] \mid \omega(t) = \tau\}.$$

Assumption (7.3) of the theorem mentioned is satisfied in this case. Indeed, if the set $\omega^{-1}(\tau)$ has zero measure, then property (7.3) is obvious, whereas in the case where (7.15) is true, condition (7.3) is satisfied due to assumption (7.16). □

Remark 7.3. If the function $\omega : [0,1] \to [0,1]$ possesses the property

$$\operatorname*{vrai\,min}_{t\in[0,1]} |\omega(t)-\tau| > 0, \tag{7.17}$$

then condition (7.14) of Corollary 7.2 is a consequence of its assumption (7.13). Indeed, it follows from (7.13) and (7.17) that

$$\begin{aligned}\operatorname*{vrai\,max}_{t\in[0,1]\setminus\omega^{-1}(\tau)} &\frac{1}{|\omega(t)-\tau|^{\gamma}} \int_0^1 |h(\omega(t),s)|\,ds \\ &\le \frac{1}{\varepsilon^{\gamma}} \operatorname*{vrai\,max}_{t\in[0,1]\setminus\omega^{-1}(\tau)} \frac{1}{|\omega(t)-\tau|^{\gamma}} \int_0^1 |\omega(s)-\tau|^{\gamma}\,|h(\omega(t),s)|\,ds < \frac{1}{\varepsilon^{\gamma}},\end{aligned}$$

where $\varepsilon := \operatorname{vrai\,min}_{t\in[0,1]} |\omega(t)-\tau|$. Therefore, relation (7.14) is true.

Remark 7.4. Condition (7.13) of Corollary 7.2 is unimprovable in the sense that the corresponding non-strict inequality

$$\operatorname*{vrai\,max}_{t\in[0,1]\setminus\omega^{-1}(\tau)} \frac{1}{|\omega(t)-\tau|^{\gamma}} \int_0^1 |\omega(s)-\tau|^{\gamma}\,|h(\omega(t),s)|\,ds \le 1. \tag{7.18}$$

does not guarantee the unique solvability of equation (7.20) for all continuous q. In order to show this, it is sufficient to consider the simplest functional equation

$$x(t) = x(\vartheta) + q(t), \qquad t \in [0,1], \tag{7.19}$$

where ϑ is a given point from $[0,1]$ and $q : [0,1] \to \mathbb{R}$ is a continuous function. Obviously, equation (7.19) can be rewritten as (7.1) with $\omega(s) := \vartheta$ and $h(t,s) := 1$ for all t and almost every s from $[0,1]$. Equation (7.19) has no continuous solutions for any continuous function $q : [0,1] \to \mathbb{R}$ satisfying the inequality $q(\vartheta) \neq 0$. Nevertheless, the corresponding condition (7.18) is true in the form of an equality with arbitrary $\vartheta \neq \tau$ and non-negative γ. Note that, for $\gamma = 1$, one can also refer to the example of equation (7.28) from Remark 7.7.

7.2. The case of a power transformation of argument. For instance, in the case of the equation

$$x(t) = \int_0^1 h(t, s)\, x(s^\alpha)\, ds + q(t), \quad t \in [0, 1], \tag{7.20}$$

where $\alpha \in (0, +\infty)$ is a constant and $q : [0, 1] \to \mathbb{R}$ is a continuous function, the following statement is true.

Corollary 7.5. *Assume that there exists some $\tau \in [0, 1]$ for which the conditions*

$$\sup_{t\in[0,1]\setminus\{\tau\}} \frac{1}{|t - \tau|^\gamma} \int_0^1 |s^\alpha - \tau|^\gamma\, |h(t,\ s)|\, ds < 1 \tag{7.21}$$

and

$$\sup_{t\in[0,1]\setminus\{\tau\}} \frac{1}{|t - \tau|^\gamma} \int_0^1 |h(t,\ s)|\, ds < +\infty \tag{7.22}$$

are satisfied with a certain $\gamma \in [0, +\infty)$. Then equation (7.20) *is uniquely solvable for any continuous function $q : [0, 1] \to \mathbb{R}$.*

Proof. Obviously, assumption (7.21) implies that

$$\sup_{t\in[0,1]\setminus\{\sqrt[\alpha]{\tau}\}} \frac{1}{|t^\alpha - \tau|^\gamma} \int_0^1 |s^\alpha - \tau|^\gamma\, |h(t^\alpha,\ s)|\, ds < 1,$$

which means that condition (7.13) is satisfied with

$$\omega(t) := t^\alpha, \qquad t \in [0, 1]. \tag{7.23}$$

In view of (7.22), condition (7.14) holds with ω given by (7.23). Thus, Corollary 7.2 can be applied. □

The following statement gives somewhat simpler but more restrictive conditions sufficient for the solvability of equation (7.20).

Corollary 7.6. *Equation* (7.20) *has a unique continuous solution for any continuous q, provided that the inequality*

$$\operatorname*{vrai\,max}_{s\in[0,1]} \sup_{t\in[0,1]\setminus\{\tau\}} \left|\frac{h(t, s)}{t - \tau}\right| < \frac{\alpha + 1}{2\alpha\tau^{1+\frac{1}{\alpha}} - (\alpha + 1)\,\tau + 1} \tag{7.24}$$

is satisfied with some $\tau \in [0, 1]$.

Proof. In view of (7.24), there exists a $\delta \in [0, 1)$ such that

$$\left|\frac{h(t, s)}{t - \tau}\right| \le \frac{\delta\,(\alpha + 1)}{2\alpha\tau^{1+\frac{1}{\alpha}} - (\alpha + 1)\,\tau + 1} \tag{7.25}$$

for almost every $s \in [0,1]$ and all $t \in [0,1] \setminus \{\tau\}$. Using estimate (7.25) and taking into account the identity

$$\int_0^1 |\xi^\alpha - \tau|\, d\xi = \frac{2\alpha\tau^{1+\frac{1}{\alpha}} + 1}{\alpha + 1} - \tau, \tag{7.26}$$

we conclude that

$$\frac{1}{|t-\tau|}\int_0^1 |\xi^\alpha - \tau|\,|h(t,\xi)|\, d\xi \le \int_0^1 |\xi^\alpha - \tau|\, d\xi \operatorname*{vrai\,max}_{s\in[0,1]} \left|\frac{h(t,s)}{t-\tau}\right| \le \delta < 1$$

for every t different from τ, i. e., condition (7.21) is satisfied with $\gamma := 1$. Moreover, by virtue of (7.25), condition (7.22) is satisfied with this value of γ. Applying Corollary 7.5, we obtain the required assertion. □

Corollary 7.6 implies, in particular, that equation (7.20) is uniquely solvable if

$$\operatorname*{vrai\,max}_{s\in[0,1]} \sup_{t\in(0,1]} t^{-1}\,|h(t,s)| < \alpha + 1. \tag{7.27}$$

It should be noted that (7.27) is weaker than the condition

$$\operatorname*{vrai\,max}_{s\in[0,1]} \sup_{t\in(0,1]} t^{-1}\,|h(t,s)| < 1,$$

which is obtained by using the standard techniques (e. g., Theorem 5.5 of [7] with $E = C([0,1],\mathbb{R})$, K defined as the cone of non-negative functions, $y_0 \equiv 1$, and the operator A given by the expression in the right-hand side of (7.20)).

Remark 7.7. None of conditions (7.24) and (7.27) can be weakened. Indeed, consider the equation

$$x(t) = \frac{(\alpha+1)\,|t-\tau|}{2\alpha\tau^{1+\frac{1}{\alpha}} - \tau\,(\alpha+1) + 1}\int_0^1 x(s^\alpha)\, ds, \qquad t \in [0,1], \tag{7.28}$$

where $\alpha \in (0,+\infty)$ and $\tau \in [0,1]$ are arbitrary constants. Obviously, equation (7.28) has form (7.20) with

$$h(t,s) := \frac{(\alpha+1)\,|t-\tau|}{2\alpha\tau^{1+\frac{1}{\alpha}} - \tau\,(\alpha+1) + 1} \tag{7.29}$$

for all t and almost every s from $[0,1]$. Moreover, due to formulae (7.26) and (7.29), we have

$$\operatorname*{vrai\,max}_{s\in[0,1]} \sup_{t\in[0,1]\setminus\{\tau\}} \left|\frac{h(t,s)}{t-\tau}\right| = \frac{\alpha+1}{2\alpha\tau^{1+\frac{1}{\alpha}} - (\alpha+1)\,\tau + 1}.$$

However, the homogeneous equation (7.28) has the non-trivial solution

$$x(t) = \frac{(\alpha+1)\,|t-\tau|}{2\alpha\tau^{1+\frac{1}{\alpha}} - \tau\,(\alpha+1) + 1}, \qquad t \in [0,1].$$

Thus, we see that condition (7.24) is unimprovable. In order to show the optimality of condition (7.27), it is sufficient to put $\tau := 0$ in (7.28).

References

[1] Birkhoff, G.: *Lattice Theory*, Amer. Math. Soc., Providence, R. I., 1948.

[2] Dunford, N. and Schwartz, J. T.: *Linear Operators. Part I*, John Wiley & Sons Inc., New York, 1988.

[3] Jentzsch, P.: *Über Integralgleichungen mit positivem Kern*, J. Reine Angew. Math., **141** (1912), 235–244.

[4] Kantorovich, L. V. and Akilov, G. P.: *Functional Analysis*, 2nd edn., Pergamon Press, Oxford, 1982.

[5] Krasnoselskii, M. A.: *Positive Solutions of Operator Equations*, Noordhoff, Groningen, 1964.

[6] Krasnoselskii, M. A., Lifshits, E. A., and Sobolev, A. V.: *Positive Linear Systems. Method of Positive Operators*, Heldermann, Berlin, 1989.

[7] Krasnoselskii, M. A., Vainikko, G. M., Zabreiko, P. P., Rutitskii, Y. B., and Stetsenko, V. Y.: *Approximate Solution of Operator Equations*, Noordhoff, Groningen, 1972.

[8] Krein, M.: *Propriétés fondamentales des ensembles coniques normaux dans l'espace de Banach*, C. R. (Doklady) Acad. Sci. URSS (N.S.), **28** (1940), 13–17.

[9] Krein, M. and Krein, S.: *On an inner characteristic of the set of all continuous functions defined on a bicompact Hausdorff space*, C. R. (Doklady) Acad. Sci. URSS (N.S.), **27** (1940), 427–431.

[10] Krein, M. G. and Rutman, M. A.: *Linear operators leaving invariant a cone in a Banach space*, No. 26 in Amer. Math. Soc. Transl., Amer. Math. Soc., R. I., 1950.

[11] Rontó, A.: *On substantial eigenvalues of linear operators leaving invariant a closed wedge*, in: *Folia FSN Universitatis Masarykianae, Mathematica*, vol. 13, Masaryk University, 2003, pp. 235–246.

[12] Rontó, A.: *Upper bounds for the eigenvalues of compact linear operators in a preordered Banach space*, Preprint No. 159, Academy of Sciences of the Czech Republic, Mathematical Institute, Praha, 2005.

[13] Stetsenko, V. Y.: *On bounds of the spectra of certain classes of linear operators*, Sov. Math. Dokl., **5** (1964), 1090–1093.

[14] Vulikh, B. Z.: *Introduction to the theory of partially ordered spaces*, Wolters-Noordhoff Scientific Publications Ltd., Groningen, 1967.

[15] Zabreiko, P. P., Krasnoselskii, M. A., and Stetsenko, V. Y.: *Bounds for the spectral radius of positive operators*, Math. Notes, **1** (1967), 306–310.

Author's address

Andrei Rontó:

Mathematical Institute, Academy of Sciences of the Czech Republic, Žižkova 22, CZ-61662 Brno, Czech Republic

E-mail address: `ronto@ipm.cz`

Doi: 10.18514/MMN.2006.128

Miskolc Mathematical Notes
Vol. 7 (2006), No. 1, pp. 91–99

HU ISSN 1787-2413
electronic version

ON THE ERGODIC AND SPECTRAL PROPERTIES OF GENERALIZED BOOLE TRANSFORMATIONS. I

ANATOLIY K. PRYKARPATSKY AND JACEK FELDMAN

[Received: May 23, 2005]

Abstract. The invariant ergodic measures for generalized Boole type transformatins are studied making use of the invariant quasi-measure approach, based on some special solutions to the Frobenius-Perron operator.

Mathematics Subject Classification: Primary 34A30, 34B05 Secondary 34B15

Keywords: generalized Boole transformations, ergodic dynamical systems, invariant quasi-measures, Frobenius-Perron operator

1. Introduction

We will consider the following generalized Boole transformation

$$\mathbb{R} \ni y \to \varphi(y) := \alpha y + a - \sum_{j=1}^{N} \frac{\beta_j}{y - b_j} \in \mathbb{R}, \tag{1.1}$$

where a and $b_j \in \mathbb{R}$, $j = \overline{1,N}$, are some real and $\alpha, \beta_j \in \mathbb{R}_+$, $j = \overline{1,N}$, are positive parameters. It generalizes that classical Boole transformation [1] $\mathbb{R} \ni y \mapsto \varphi(y) := y - 1/y \in \mathbb{R}$, which appeared to be ergodic [2] with respect to the invariant standard infinite Lebesgue measure on $\mathbb{R}$. In the case where $\alpha = 1, a = 0$, the similar ergodicity result was proved in [3–5] making use of the specially despised inner function notion. The related spectral properties were in part studied in [5]. In spite of these results the case $\alpha \neq 1$ still persists to be challenging as the only related result [6] concerns the following special case of (1.1): $\mathbb{R} \ni y \mapsto \varphi(y) := \alpha y - \beta/y \in \mathbb{R}$ for $0 < \alpha < 1$ and arbitrary $\beta \in \mathbb{R}_+$. The corresponding invariant measure appeared to be finite absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$ and equal to

$$d\mu(x) := \frac{\sqrt{\beta(1-\alpha)}dx}{\pi[x^2(1-\alpha)+\beta]}, \tag{1.2}$$

The first author was supported in part by a local AGH grant.

where $x \in \mathbb{R}$. The ergodicity for the invariant measure (1.2) now can be easily proved. Recall here, concerning a general nonsingular mapping $\varphi : \mathbb{R} \to \mathbb{R}$, the problem of constructing the measure preserving ergodic measures is analized [6, 7] by means of studying the spectral properties of the adjoint Frobenius-Perron operator $\hat{T}_\varphi \varrho : L_2(\mathbb{R};\mathbb{R}) \to L_2(\mathbb{R};\mathbb{R})$, where, by definition,

$$\hat{T}_\varphi \varrho(x) := \sum_{y \in \{\overline{\varphi'(x)}\}} \varphi(y) J_\varphi^{-1}(y) \tag{1.3}$$

for any $\varrho \in L_2(\mathbb{R};\mathbb{R}_+)$ and $J_\varphi^{-1}(y) := \left|\frac{d\varphi(y)}{dy}\right|$, $y \in \mathbb{R}$. Then if $\hat{T}_\varphi \varrho = \varrho$, $\varrho \in L_2(\mathbb{R};\mathbb{R}_+)$, then the expression $d\mu(x) := \varrho(x)dx$, $x \in \mathbb{R}$, will be an invariant, in general infinite, measure with respect to the mapping $\varphi : \mathbb{R} \to \mathbb{R}$.

Another way to finding a general algorithm for finding such an invariant measure was devised in [8, 9], making use of the generating measure function method.

Below we will study another special cases of the generalized Boole transformation (1.1), for which we deliver the corresponding invariant measures and prove the related ergodicity and spectral properties.

2. Invariant measures and ergodic transformations

We will start by analyzing the following Boole type surjective transformation

$$\mathbb{R} \ni y \to \varphi(y) := \alpha y + a - \frac{\beta}{y-b} \in \mathbb{R} \tag{2.1}$$

for any $a, b \in \mathbb{R}$ and $2\beta := \gamma^2 \in \mathbb{R}_+$. The transformation (2.1) at $\alpha = 1/2$ and $b = 2a \in \mathbb{R}$ will be measure preserving with respect to a measure like (1.2). Namely, the following lemma holds.

Lemma 2.1. *The Boole type mapping* (2.1) *is measure preserving with respect to the measure*

$$d\mu(x) := \frac{|\gamma| dx}{\pi[(x-2a)^2 + \gamma^2]}, \tag{2.2}$$

where $x \in \mathbb{R}$ and $\gamma^2 := 2\beta \in \mathbb{R}_+$.

Proof. A proof follows easily from the fact that the function

$$\varrho(x) := \frac{\gamma}{\pi[(x-2a)^2 + \gamma^2]} \tag{2.3}$$

satisfies for all $x \in \mathbb{R} \setminus \{2a\}$ the determining condition (1.3):

$$\hat{T}_\varphi \varrho(x) := \sum_I \varrho(y_\pm) y'_\pm(x), \tag{2.4}$$

where, by definition, $\varphi(y_\pm(x)) := x$ for any $x \in \mathbb{R}$. The relation (2.4) is, evidently, equivalent to the next invariance condition

$$\sum_\pm d\mu(y_\pm(x)) = d\mu(x) := \mu(dx) \tag{2.5}$$

for any infinitesimal subset $dx \subset \mathbb{R}$. □

The question about the ergodicity of mapping (2.1) is solved here easily by the following theorem.

Theorem 2.2. *Measure* (2.3) *is ergodic with respect to transformation* (2.1) *at* $\alpha = 1/2$ *and* $b = 2a \in \mathbb{R}$ *as such one is equivalent to the canonical ergodic mapping* $\mathbb{R}/\pi\mathbb{Z} \ni s \mapsto \psi(s) := 2s \bmod (\pi Z) \in \mathbb{R}/\pi\mathbb{Z}$ *with respect to the standard Lebesgue measure on* $\mathbb{R}/\pi\mathbb{Z}$.

Proof. Put by definition $\mathbb{R}/\pi\mathbb{Z} \ni s \mapsto \xi(s) = y \in \mathbb{R}$, where

$$\xi(s) := \gamma \operatorname{ctg} s + 2a, \tag{2.6}$$

Then transformation (2.1) at $\alpha = 1/2$, $b = 2a \in \mathbb{R}$ and $\gamma^2 := 2\beta \in \mathbb{R}_+$ yields under mapping (2.6)

$$\begin{aligned} \varphi(y) = \varphi(\xi(s)) &= \frac{\gamma}{2} \operatorname{ctg} s + 2a - \frac{\gamma}{2} \operatorname{tg} s = \frac{\gamma(\cos^2 s - \sin^2 s)}{2 \sin s \cos s} + 2a \\ &= \gamma \frac{\cos 2s}{\sin 2s} + 2a = \gamma \operatorname{tg} 2s + 2a := \xi(2s) \end{aligned} \tag{2.7}$$

for any $s \in \mathbb{R}/\pi\mathbb{Z}$. Relation (2.7) means that transformation (2.1) is conjugated [5, 7] with the transformation

$$\mathbb{R}/\pi\mathbb{Z} \ni s \longmapsto \psi(s) = 2s \bmod (\pi\mathbb{Z}) \in \mathbb{R}/\pi\mathbb{Z}, \tag{2.8}$$

that is the following diagram is commutative:

$$\begin{tikzcd} \mathbb{R}/\pi\mathbb{Z} \arrow[r, "\xi"] \arrow[dr, "\psi"'] & \mathbb{R} \arrow[d, "\varphi"] \\ & \mathbb{R}/\pi\mathbb{Z} \end{tikzcd} \tag{2.9}$$

It is easy now to check that the measure (2.2) under the conjugation (2.9) transforms into the standard normalized Lebesgue measure on $\mathbb{R}/\pi\mathbb{Z}$:

$$d\mu(x)|_{x=\gamma \operatorname{ctg} s+2a} = \frac{ds\gamma^2 \left| d(\operatorname{ctg} s)/ds \right|}{\pi(\gamma^2 \operatorname{ctg}^2 s + \gamma^2)} = \frac{1}{\pi} \frac{\sin^2 s \cdot \sin^{-2} s \, ds}{\cos^2 s + \sin^2 s} = \pi^{-1} ds, \tag{2.10}$$

where $s \in \mathbb{R}/\pi\mathbb{Z}$. The infinitesimal measure $\pi^{-1}ds$ on $\mathbb{R}/\pi\mathbb{Z} \ni s$ as well as the infinitesimal measure (2.2) on $\mathbb{R}$ are normalized, being thus probabilistic. Now it is enough to make use of the fact that the measure $\pi^{-1}ds$ on $\mathbb{R}/\pi\mathbb{Z} \ni s$ on the interval $[0, \pi] \simeq \mathbb{R}/\pi\mathbb{Z}$ is ergodic [6, 7]. □

3. Ergodic measures: a new approach

Assume that there exists a holomorphic, in parameter $\omega \in \mathbb{C}_+$, function $\varrho_\omega \in H_2(\mathbb{C}_+;\mathbb{C})$, satisfying the following identity

$$\hat{T}_\varphi \varrho_\omega = \varrho_{\tilde{\varphi}(\omega)} \tag{3.1}$$

valid for any $\omega \in \mathbb{C}_+$ for some induced transformation $\mathbb{C}_+ \ni \omega \mapsto \tilde{\varphi}(\omega) \in \mathbb{C}_+$. If we now take $\omega := \bar{\omega} \in \mathbb{C}_+$ being a fixed point of the mapping $\tilde{\varphi} : \mathbb{C}_+ \to \mathbb{C}_+$, then an easy conclusion from 3.1 gives rise to the condition $\hat{T}_\varphi \varrho_{\bar{\omega}} = \varrho_{\bar{\omega}}$, meaning that the expression

$$d\mu(x) := \operatorname{Im} \varrho_{\bar{\omega}}(x) dx, \tag{3.2}$$

is a searched invariant measure for the transformation $\varphi : \mathbb{R} \to \mathbb{R}$. There exists still no general rule of constructing such functions $\varrho_\omega \in H_2(\mathbb{C}_+;\mathbb{C})$ analytic in $\omega \in \mathbb{C}_+$ and related induced mappings $\tilde{\varphi} : \mathbb{C}_+ \to \mathbb{C}_+$. Nevertheless, for solving this problem one can adapt some natural motivations related to the exact functional form of the determining Frobenius-Perron operator $\hat{T}_\varphi : L_2(\mathbb{R};\mathbb{R}) \to L_2(\mathbb{R};\mathbb{R})$. To explain this, let us consider the following Boole type transformation:

$$\mathbb{R} \ni \varphi(y) := \alpha y + a - \frac{\beta}{y-b} \in \mathbb{R}, \tag{3.3}$$

where $a, b \in \mathbb{R}$ and $\beta \in \mathbb{R}_+$. It is easy to observe that the Frobenius-Perron operator action on any $\varrho_\omega \in H_2(\mathbb{C}_+;\mathbb{C})$ can be represented as follows:

$$\begin{aligned} \hat{T}_\varphi \varrho_\omega &:= \varrho_\omega(y_+)y'_+ + \varrho_\omega(y_-)y'_- \\ &= \frac{(\omega - y_+)\varrho_\omega(y_+)(\omega - y_-)y'_-}{(\omega - y_+)(\omega - y_-)} + \frac{\varrho_\omega(y_-)(\omega - y_+)(\omega - y_-)y'_-}{(\omega - y_+)(\omega - y_-)} \\ &= \frac{k(\omega - y_-)y'_+ + k(\omega - y_+)y'_-}{(\omega - y_+)(\omega - y_-)} = \frac{-k[(\omega - y_+)(\omega - y_-)]'}{(\omega - y_+)(\omega - y_-)} \\ &= -k\frac{d}{dx}\ln[(\omega - y_+)(\omega - y_-)], \end{aligned} \tag{3.4}$$

where we put, by definition,

$$\varrho_\omega(x) = \frac{k}{\omega - x} \tag{3.5}$$

for all $\omega \in \mathbb{C}_+ \setminus \{x\}$, $x \in \mathbb{R}$, and some parameter $k \in \mathbb{R}$. As a result of (3.5) one can take

$$\varrho_\omega(y_+)(\omega - y_+) = k = \varrho_\omega(y_-)(\omega - y_-), \tag{3.6}$$

for all $x \in \mathbb{R}$ and $\omega \in \mathbb{C}_+$. Since the root functions $y_+, y_- : \mathbb{R} \to \mathbb{R}$ satisfy, by definition, the same equation

$$\omega(y_\pm(x)) = x, \tag{3.7}$$

for all $x \in \mathbb{R}$, the following identity for all $\omega \in \mathbb{C}_+$ is easily inferred from (3.7) owing to the general form of (3.3):

$$\alpha(\omega - y_+)(\omega - y_-) \equiv [\varphi(\omega) - x](\omega - b), \tag{3.8}$$

where

$$y_+(x) + y_-(x) = b + \frac{x-a}{2}, \qquad y_+(x)y_-(x) = \frac{bx - ab - \beta}{2}. \tag{3.9}$$

Whence, taking into account the expression (3.4), one gets that

$$\begin{aligned} \hat{T}_\varphi \varrho_\omega &= -k\frac{d}{dx}\ln([\varphi(\omega) - x](\omega - b)) \\ &= \frac{k(\omega - b)}{[\varphi(\omega) - x](\omega - b)} = \frac{k}{\varphi(\omega) - x} = \varrho_{\varphi(\omega)}, \end{aligned} \tag{3.10}$$

for all $x \in \mathbb{R}$ and $\omega \in \mathbb{C}_+$. Therefore, the induced mapping $\tilde{\varphi} : \mathbb{C}_+ \to \mathbb{C}_+$ is exactly the transformation $\varphi : \mathbb{C}_+ \to \mathbb{C}_+$, extended naturally from the axis $\mathbb{R}$ on the complex plane $\mathbb{C}_+$.

Let now $\tilde{\omega} \in \mathbb{C}_+$ be a fixed point of the induced mapping $\varphi : \mathbb{C}_+ \to \mathbb{C}_+$, that is $\varphi(\bar{\omega}) = \bar{\omega} \in \mathbb{C}_+$. Then from (3.10) one finds that $\hat{T}_\varphi \varrho_{\bar{\omega}} = \varrho_{\bar{\omega}}$, or the corresponding invariant quasi-measure on $\mathbb{R}$ has the form

$$d\mu(x) := \operatorname{Im} \frac{kdx}{\bar{\omega} - x} \tag{3.11}$$

for all $x \in \mathbb{R}$ and some suitable parameter $k \in \mathbb{C}$. As $\operatorname{Im} \varrho_{\bar{\omega}} \in L_2(\mathbb{R}; \mathbb{R}_+)$ at any $\bar{\omega} \in \mathbb{C}_+ \setminus \mathbb{R}$ and some $k \in \mathbb{C}$, then the invariant quasi-measure (3.11) becomes a true invariant measure. The obtained result can be formulated as the following theorem.

Theorem 3.1. *The quasi-measure* (3.11) *is invariant with respect to transformation* (3.3) *for any* $\alpha \in \mathbb{R}_+ \setminus \{1\}$; *for* $\alpha = 1$ *at the condition* $a \neq 0$, $\operatorname{Im} k \neq 0$, *it is reduced upon the set* $\mathbb{R}/\pi\mathbb{Z}$, *being equivalent to the standard Gauss measure.*

Proof. Indeed, the searched infinitesimal quasi-measure $d\mu(x)$ exist if there exists at least one fixed point of the equation $\varphi(\omega) = \omega$ for $\omega \in \mathbb{C}_+$. If $\alpha \neq 1$, this equation is equivalent to

$$(\alpha - 1)\omega^2 - \omega[(\alpha - 1)b - a] - (ab + \beta) = 0, \tag{3.12}$$

always possessing a solution $\bar{\omega} \in \mathbb{C}_+$, for which $\varphi(\bar{\omega}) = \bar{\omega}$. At $\alpha = 1$ the unique solution $\bar{\omega} = (ab + \beta)/a \in \mathbb{R}$ exists only at $a \neq 0$, $\operatorname{Im} k \neq 0$, at which the quasi-measure (3.11) becomes degenerate, thus reducing to the standard Gauss measure [3, 7] on $\mathbb{R}/\pi\mathbb{Z}$. □

The Theorem 3.1 states only that quasi-measure (3.11) is invariant with respect to the transformation (3.3), thereby the ergodicity if any still needs to be proved separately, on which we plan to do in detail in another place. Below we will proceed to study the general case of transformation (1.1), searching for a suitable invariant quasi-measure becoming a measure at some $\bar{\omega} \in \mathbb{C}_+/\mathbb{R}$, $k \in \mathbb{C}$.

4. Invariant measures: the general case

Consider the following equation

$$\varphi(y) = x, \tag{4.1}$$

where $x, y \in \mathbb{R}$ and the mapping $\varphi : \mathbb{C}_+ \to \mathbb{C}_+$ is given by expression (1.1) at some still fixed integer $N \in \mathbb{Z}_+ \setminus \{1\}$. The equation (4.1) can be rewritten as

$$\alpha \prod_{j=1}^{N+1} (y - y_j) = [\varphi(y) - x] \prod_{j=1}^{N} (y - b_j) \tag{4.2}$$

for all $x, y \in \mathbb{R}$ and some functions $y_j : \mathbb{R} \to \mathbb{R}$, $j = \overline{1, N+1}$. Relation (4.2) is naturally extended to the complex plane $\mathbb{C}_+$ as

$$\alpha \prod_{j=1}^{N+1} (\omega - y_j) = [\varphi(\omega) - x] \prod_{j=1}^{N} (\omega - b_j) \tag{4.3}$$

for any $\omega \in \mathbb{C}_+$.

Consider now relation (3.1), similarly to Section 3:

$$\begin{aligned}
\hat{T}_\varphi \varrho_\omega &= \sum_{j=1}^{N} \varphi(\omega)(y_j) y_j' = \sum_{j=1}^{N+1} \frac{\varphi(\omega)(y_j)(\omega - y_j \prod_{k \neq j}^{N+1} (\omega - y_k) y_j'}{\prod_{k=1}^{N+1} (\omega - y_k)} \\
&= \sum_{j=1}^{N+1} \frac{\varphi(\omega)(y_j)(\omega - y_j \prod_{k \neq j}^{N+1} (\omega - y_k) y_j'}{\prod_{k=1}^{N+1} (\omega - y_k)} = \sum_{j=1}^{N+1} \frac{k \prod_{k \neq j}^{N+1} (\omega - y_k) y_j'}{\prod_{k=1}^{N+1} (\omega - y_k)} \\
&= -k \frac{\frac{d}{dx} \prod_{k \neq j}^{N+1} (\omega - y_k)}{\prod_{k=1}^{N} (\omega - y_k)} = -k \frac{d}{dx} \ln \prod_{k=1}^{N+1} (\omega - y_k),
\end{aligned} \tag{4.4}$$

where we have put, as before,

$$\varrho_\omega(y_j)(\omega - y_j) = k, \tag{4.5}$$

for all $j = \overline{1, N+1}$, $\omega \in \mathbb{C}_+$, and some parameters $k \in \mathbb{C}$. This evidently means that

$$\varrho_\omega(y) = \frac{k}{\omega - y} \tag{4.6}$$

for any $y \in \mathbb{R}$ and $\omega \in \mathbb{C}_+$.

Having substituted now expression (4.3) into (4.4), one gets easily that

$$\hat{T}_\varphi \varrho_\omega(x) = \frac{k}{\varphi(\omega) - x} = \varrho_{\varphi(\omega)}(x), \tag{4.7}$$

for all $x \in \mathbb{R}$ and any $\omega \in \mathbb{C}_+$. Thus, the invariant quasi-measure for the discrete dynamical system (1.1) will be given by the same expression (3.11), where $\bar{\omega} \in \mathbb{C}_+$

is some fixed point of the mapping $\varphi : \mathbb{C}_+ \to \mathbb{C}_+$. This means that

$$\alpha\bar{\omega} + a - \sum_{j=1}^{N} \frac{\beta_j}{\bar{\omega} - b_j} = \bar{\omega}, \tag{4.8}$$

or, equivalently,

$$\alpha\bar{\omega}\prod_{j=1}^{N}(\bar{\omega} - b_j) + a\prod_{j=1}^{N}(\bar{\omega} - b_j) - \sum_{j=1}^{N}\beta_j\prod_{k\neq j}^{N}(\bar{\omega} - b_k) = \bar{\omega}\prod_{j=1}^{N}(\bar{\omega} - b_j), \tag{4.9}$$

for some $\bar{\omega} \in \mathbb{C}_+$. Assume now that $\alpha \neq 1$; then it is easy to find that the algebraic equation (4.9) possesses exactly $N+1 \in \mathbb{Z}_+$ roots, which can be used for constructing the invariant quasi-measure (3.11). The case $\alpha = 1$ gives rise to the condition

$$a\prod_{j=1}^{N}(\bar{\omega} - b_j) = \sum_{j=1}^{N}\beta_j\prod_{k\neq j}^{N}(\bar{\omega} - b_k), \tag{4.10}$$

which always possesses roots if $N \geq 2$ and $a \in \mathbb{R}$ is arbitrary. Thereby, one can formulate the following theorem for the case $N \geq 2$.

Theorem 4.1. *Expression* (3.11) *at some $k \in \mathbb{C}$ determines, in general, the infinitesimal invariant quasi-measure for the generalized Boole transformation (1.1) at all $N \geq 2$ with arbitrary parameters $a, b_j \in \mathbb{R}$ and $\alpha, \beta_j \in \mathbb{R}_+$, $j = \overline{1, N+1}$.*

It is an important task now to separate from the obtained set of invariant quasi-measures (3.11) those which are positively defined and ergodic with respect to transformation (1.1) at $N \geq 2$. The positivity condition simply means that the determining equation (4.9) must possess at least one pair of complex conjugated roots with a nontrivial imaginary part. Having analyzed, concerning this criteria, roots of equation (4.9) one can find that the following statement analogous to that proved in [6] holds.

Theorem 4.2. *The generalized Boole transformation* (1.1) *for any $N \geq 1$ is necessarily ergodic with respect to the measure* (3.11) *at some $\bar{\omega} \in \mathbb{C}_+ \setminus \mathbb{R}$ and $k \in \mathbb{C}$ iff $\alpha = 1$ and $a = 0$. If $\alpha = 1$ and $a \neq 0$, the transformation* (1.1) *is not ergodic being totally dissipative, that is the wandering set $\mathscr{D}(\varphi) := \bigcup \mathscr{W}(\varphi) = \mathbb{R}$, where $\mathscr{W}(\varphi) \subset \mathbb{R}$ are subsets such that all sets $\varphi^{-n}(\mathscr{W})$, $n \in \mathbb{Z}$, are disjoint.*

Sketch of proof. It is easy to see that for $N \geq 2$ at $\alpha = 1$ and $a = 0$ the determining algebraic equation (4.9) always possesses exactly $N - 1 \in \mathbb{Z}_+$ real roots $\bar{\omega}_j \in \mathbb{R}$, $j = \overline{1, N-1}$. Therefore, the invariant quasi-measure expression (3.11) is degenerate for $\bar{\omega}_j \in \mathbb{R}$, $j = \overline{1, N-1}$, giving rise to the conclusion that the corresponding invariant measure $d\mu(x) = dx$, $x \in \mathbb{R}$, is the standard Lebesgue measure on $\mathbb{R}$. Its ergodicity with respect to transformation (1.1) then follows from the fact that the corresponding dissipative set $\mathscr{D}(\varphi) = \varnothing$ and the unique invariant set subalgebra $I(\varphi) = \mathbb{R}$. □

The statements similar to above can be formulated for the most generalized Boole type transformation

$$\mathbb{R} \ni y \to \varphi(y) := \alpha y + a + \int_{\mathbb{R}} \frac{d\nu(s)}{s-y} \in \mathbb{R}, \tag{4.11}$$

where $a \in \mathbb{R}$, $\alpha \in \mathbb{R}_+$ and a measure ν on $\mathbb{R}$ has the compact support $\operatorname{supp} \nu \subset \mathbb{R}$, being such that the following natural conditions [3]

$$\int_{\mathbb{R}} \frac{d\nu(s)}{1+s^2} = a, \quad \int_{\mathbb{R}} d\nu(s) < \infty, \tag{4.12}$$

hold. Concerning the extension of the transformation (4.11) on the upper part $\mathbb{C}_+$ of the complex plane $\mathbb{C}$ in such a way that $\operatorname{Im} \varphi(\omega) \geq 0$ for all $\omega \in \mathbb{C}_+$, the following representation

$$\varphi(\omega) = \alpha\omega + a + \int_{\mathbb{R}} \frac{1+s\omega}{s-\omega} d\sigma(s), \tag{4.13}$$

holds [3, 11], where a measure $d\sigma$ on $\mathbb{R}$ is closely related to the measure $d\nu$.

The general properties of mapping (4.13) were in part studied in [6] in the framework of the theory of inner functions. The invariant measures corresponding to (4.11) and their ergodic properties can be also treated effectively making use of the analytical and spectral properties of the associated Frobenius–Perron transfer operator (1.3). These aspects of the problem being are as important as interesting, and are planned to be studied elsewhere.

5. Acknowledgements

The authors are cordially grateful to professors D. Blackmore (NJ, USA), F. Pszytycki (Warsów) and T. Downarowicz (Wroclaw) for valuable discussions of the ergodic measure properties related to the generalized Boole transformations.

References

[1] Boole, G. *On the comparison of transcendents with certain applications to the theory of definite integrals*, Philos Trans. Royal Soc. London, **147** (1957), 745–803.

[2] Adler, R. and Weiss, B.: *The ergodic, infinite measure preserving transformation of Bool*, Israel Journal of Math., **16** (1973), 263–278.

[3] Aaronson, J.: *Ergodic theory for inner functions of the upper half plane*, Ann. Inst. H. Poincaré, BXIV (1978), 233–253.

[4] Aaronson J.: *A remark on this existence of inner functions*, J. London Math. Soc, **23** (1981), 469–474.

[5] Aaronson, J.: *The eigenvalues of nonsingular transformations*, Israel J. Math., **45** (1983), 297–312.

[6] Aaronson, J.: *An Introduction to Infinite Ergodic Theory*, Amer. Math. Soc, 1997.

[7] Katok, A. and Hasselblatt, B.: *Introduction to the Modern Theory of Dynamical Systems*, Cambridge University Press, Cambridge, 1997.

[8] Prykarpatsky A. K.: *On invariant measure structure of a class of ergodic discrete dynamical systems*, Nonlinear Oscillations, **3** (2000), No. 1, 78–83.

[9] Prykarpatsky, A. K. and Brzychczy, S.: *On invariant measure structure of a class of ergodic discrete dynamical systems*, Proceedings of the International Conference SCAN 2000/Interval 2000, September 19–22, Karlsruhe, Germany.

[10] Pollycott, M. and Yari, M.: *Dynamical Systems and Ergodic Theory*, London Math. Society, Cambridge University Press, 1998.

[11] Ablowitz, M. J. and Fokas, A. S.: *Complex Variables: Introduction and Applications*, Cambridge University Press, 1997.

Authors' addresses

Anatoliy K. Prykarpatsky:

The AGH University of Science and Technology, Department of Applied Mathematics, Krakow 30059, Poland

Jacek Feldman:

The AGH University of Science and Technology, Department of Applied Mathematics, Krakow 30059, Poland

Doi: 10.18514/MMN.2006.127

Miskolc Mathematical Notes
Vol. 7 (2006), No. 1, pp. 101–108

HU ISSN 1787-2413
electronic version

CONVERGENCE FIELDS OF REGULAR MATRIX TRANSFORMATIONS OF SEQUENCES OF ELEMENTS OF BANACH SPACES

T. VISNYAI

[Received: May 23, 2005]

ABSTRACT. In [4, 7] some properties of convergence fields of regular matrix transformations of bounded sequences of real numbers are presented. We shall prove a generalization of Steinhaus' theorem for sequences of a Banach space and show that results of [4] cen be generalize for a space of sequences of elements of a Banach space $(X, \|\cdot\|)$.

Mathematics Subject Classification: 46B45,40D09

Keywords: convergence field, Baire category, porosity

1. INTRODUCTION

Let $(X, \|\cdot\|)$ be a Banach space. The sequence $\alpha = (\alpha_k)$ converges to c if $\forall \varepsilon > 0$ $\exists k_0 \in \mathbb{N}$ $\forall k > k_0$: $\|\alpha_k - c\| < \varepsilon$. We write $\lim_{k\to\infty} \alpha_k = c$.

We define the following sets:

(a) $b = \{\alpha = (\alpha_k) : \alpha_k \in X, k = 1, 2, \ldots\}$;
(b) $B_\infty = \{\alpha = (\alpha_k) : \alpha_k \in X, k = 1, 2, \ldots : \exists K_\alpha > 0\ \forall k = 1, 2, \ldots \|\alpha_k\| \le K_\alpha\}$;
(c) $\Omega = \{\alpha = (\alpha_k) : \alpha_k \in X, k = 1, 2, \ldots : \forall k = 1, 2, \ldots \|\alpha_k\| = 0 \ \vee\ \|\alpha_k\| = 1\}$.

The set b contains all sequences of elements of X, the set B_∞ is the set of all bounded sequences of X and Ω is the set of all sequences of elements of X which have norm null or one. In what follows, for the set B, the following inclusions hold:

$$\Omega \subset B_\infty \subset B \subset b. \tag{1}$$

The notion of porosity has been introduced in [10]. It is a suitable tool to describe small sets in a metric space.

Let (Y, d) be a metric space, $Z \subset Y$. Let $y \in Y$, $\delta > 0$, and let $B(y, \delta)$ denote the set $\{x \in Y : d(x, y) < \delta\}$. We put

$$P(y, Z, \delta) = \sup\{t > 0 : \exists z \in B(y, \delta) \text{ such that } B(z, t) \subset B(y, \delta) \text{ and } B(z, t) \cap Z = \emptyset\}.$$

Supported by the Grant VEGA 1/3018/06.

If such $t > 0$ does not exist, we put $P(y, Z, \delta) = 0$. The numbers

$$\underline{p}(y, Z) = \liminf_{\delta\to 0^+} \frac{P(y, Z, \delta)}{\delta}$$

and

$$\overline{p}(y, Z) = \limsup_{\delta\to 0^+} \frac{P(y, Z, \delta)}{\delta}$$

are called the lower and upper porosity of the set Z at $y \in Y$, respectively.

A set $Z \subset Y$ for which $\overline{p}(y, Z) > 0$ for every $y \in Y$ is said to be porous in Y. Obviously every set porous in Y is nowhere dense in Y. If $\underline{p}(y, Z) = \overline{p}(y, Z) = p(y, Z)$, then the number $p(y, Z)$ is called porosity of Z at $y \in \overline{Y}$. If $p(y, Z) = 1$, then Z is said to be strongly porous at y. If for all $y \in Y$ we have $p(y, Z) = 1$ for $y \notin Z$ and $p(y, Z) = \frac{1}{2}$ for $y \in Z$, then Z is said to be strongly porous at Y. The set W is said to be σ-porous (σ-strongly porous) at Y if $W = \bigcup_{n=1}^{\infty} Z_n$ and each Z_n is porous (strongly porous) at Y.

Further on, we recall the definition of a matrix transformation. Let $\mathscr{A} = (a_{nk})$ be an infinite matrix of real numbers. A sequence $\alpha = (\alpha_k)$, $\alpha_k \in X$, is said to be $\mathscr{A}$-limitable (limitable by the method $\mathscr{A}$) to the element $c \in X$ if, for $\beta = (\beta_n(\alpha))$, $\beta_n = \beta_n(\alpha) = \sum_{k=1}^{\infty} a_{nk}\alpha_k$, we have $\lim_{n\to\infty} \beta_n = c$.

If $\alpha = (\alpha_k)$ is $\mathscr{A}$-limitable to the element c, we write $\mathscr{A}\text{-}\lim_{k\to\infty} \alpha_k = c$. The method $\mathscr{A}$ defined by the matrix $\mathscr{A}$ is said to be regular if $\lim_{k\to\infty} \alpha_k = c$ implies that $\mathscr{A}\text{-}\lim_{k\to\infty} \alpha_k = c$.

Let $\mathscr{A}$ be a regular method. The symbol $F(\mathscr{A})$ denotes the set of all $\mathscr{A}$-limitable sequences of X. We put

$$F(\mathscr{A}) = \Big\{\alpha = (\alpha_k) \mid \alpha_k \in X, k = 1, 2, \dots : \text{ there exists } \lim_{n\to\infty} \beta_n, \\ \text{where } \beta_n = \sum_{k=1}^{\infty} a_{nk}\alpha_k\Big\}.$$

The set $F(\mathscr{A})$ is called the convergence field of the matrix transformation $\mathscr{A}$.

It is proved in [4] that $F(\mathscr{A})$ is a set of first Baire category in S, where S is a linear space of sequences of real numbers and $F(\mathscr{A})$ is strongly porous in l_∞ (l_∞ is the set of all bounded sequences of real numbers).

In this paper we show that these results can be generalized for the sequences of elements of Banach space $(X, \|\cdot\|)$.

2. Main results

Monograph [6] gives the Toeplitz theorem which provides necessary and sufficient condition for matrix $\mathscr{A}$ be a regular, i. e., when $\lim_{k\to\infty} x_k = t$ implies $\mathscr{A}\text{-}\lim_{k\to\infty} x_k = t$ for real sequences (x_k).

In [9], it is proved that Toeplitz theorem holds also for sequences of elements of a Banach space.

Theorem A. *Let* $(X, \|\cdot\|)$ *be a Banach space. Let* $\mathscr{A} = (a_{nk})$ *be an infinite matrix of real numbers. The necessary and sufficient condition for a sequence* $\beta_n(\alpha) = \sum_{k=1}^{\infty} a_{nk}\alpha_k$*, to converge to* $c \in X$ *as* $n \to \infty$ *where* $\lim_{k\to\infty} \alpha_k = c$, $\alpha_k \in X$ *is that matrix* $\mathscr{A}$ *satisfies the following three conditions:*

(a) $\exists M > 0,\ \forall n = 1, 2, \dots \ \sum_{k=1}^{\infty} |a_{nk}| \le M$;
(b) $\forall k = 1, 2, \dots \ \lim_{n\to\infty} a_{nk} = 0$;
(c) $\lim_{n\to\infty} \sum_{k=1}^{\infty} a_{nk} = 1$.

Further we will use the following theorem concerning the functions of Baire class one (see [8, p. 185]).

Theorem B. *Let* A *and* B *be metric spaces. Let* $\delta_n : A \to B$, $n = 1, 2, \dots$, *be a sequence of continuous operators, which converges pointwise to an operator* δ, *i. e.,*

$$\forall a \in A : \lim_{n\to\infty} \delta_n(a) = \delta(a).$$

Then the set of all discontinuity points of the operator $\delta : A \to B$ *is a set of the first Baire category.*

In [2], the Steinhaus theorem is proved under the condition that there does not exist a regular matrix which limits all sequences of 0's and 1's. Now we will show an analogue of the Steinhaus theorem for sequences of elements of X. First we prove the following:

Lemma 1. The metric space (Ω, d), where $d(\alpha,\beta) = \sum_{k=1}^{\infty} 2^{-k} \|\alpha_k - \beta_k\|$, $\alpha = (\alpha_k) \in \Omega$, $\beta = (\beta_k) \in \Omega$, is a complete metric space.

Proof. The function $d : \Omega \times \Omega \to \langle 0, \infty)$ is a metric.
Let $\alpha^{(n)} = (\alpha_k^{(n)}), n = 1, 2, \dots$ be a Cauchy sequence of elements of Ω. Thus,

$$\forall \varepsilon > 0\ \exists n_0 \in \mathbb{N} \quad \forall m, n > n_0 : \ d(\alpha^{(n)}, \alpha^{(m)}) < \varepsilon.$$

Choose $j \in \mathbb{N}$. Then

$$\varepsilon > d(\alpha^{(n)}, \alpha^{(m)}) = \sum_{k=1}^{\infty} \frac{\|\alpha_k^{(n)} - \alpha_k^{(m)}\|}{2^k} \ge \frac{1}{2^j} \|\alpha_j^{(n)} - \alpha_j^{(m)}\|.$$

The sequences $(\alpha_j^{(n)})_{n=1}^{\infty}$ of elements of X are Cauchy sequences. X is a complete metric space, therefore $(\alpha_j^{(n)})_{n=1}^{\infty}$ is convergent.

Let $\lim_{n\to\infty} \alpha_j^{(n)} = \alpha_j$. If $\alpha = (\alpha_j)_{j=1}^{\infty}$, then one can easily verify that that $\lim_{n\to\infty} \alpha^{(n)} = \alpha$ and, for each $j = 1, 2, \dots$, $\|\alpha_j\| = 0$ or $\|\alpha_j\| = 1$. □

Theorem C. *For any regular matrix* $\mathscr{A} = (a_{nk})$ *there exists a sequence in the set* Ω, *which is not limitable by the method* $\mathscr{A}$.

Proof. Let $\mathscr{A} = (a_{nk})$ be a regular matrix. Then there exists an $M > 0$ such that, for all $n = 1, 2, \ldots$,

$$\sum_{k=1}^{\infty} |a_{nk}| \leq M. \tag{2}$$

We prove that the operator $\delta_n : \Omega \to X$ is continuous at $\alpha = (\alpha_k) \in \Omega$ for each $n = 1, 2, \ldots$, where $\delta_n(\alpha) = \sum_{k=1}^{\infty} a_{nk}\alpha_k$, $\alpha = (\alpha_k) \in \Omega$. Let $\varepsilon > 0$ and $\alpha = (\alpha_k) \in \Omega$, $\beta = (\beta_k) \in \Omega$. Then

$$\|\delta_n(\alpha) - \delta_n(\beta)\| = \|\sum_{k=1}^{\infty} a_{nk}\alpha_k - \sum_{k=1}^{\infty} a_{nk}\beta_k\| \leq \sum_{k=1}^{\infty} |a_{nk}|\|\alpha_k - \beta_k\|.$$

We choose $n_0 \in \mathbb{N}$ so that $\sum_{k>n_0} |a_{nk}| < \frac{\varepsilon}{4}$. Then

$$\sum_{k>n_0} |a_{nk}|\|\alpha_k - \beta_k\| \leq \sum_{k>n_0} |a_{nk}|(\|\alpha_k\| + \|\beta_k\|) \leq 2 \cdot \frac{\varepsilon}{4} = \frac{\varepsilon}{2}. \tag{3}$$

If $d(\alpha, \beta) < \frac{\varepsilon}{2M2^{n_0}}$ for α and β, then we have

$$\frac{\|\alpha_k - \beta_k\|}{2^k} \leq d(\alpha, \beta) < \frac{\varepsilon}{2M2^{n_0}}$$

for all $k = 1, 2, \ldots n_0$. Therefore, $\|\alpha_k - \beta_k\| < \frac{\varepsilon}{2M}$ and

$$\sum_{k=1}^{n_0} |a_{nk}|\|\alpha_k - \beta_k\| < \frac{\varepsilon}{2M} \sum_{k=1}^{n_0} |a_{nk}| \leq \frac{\varepsilon}{2}. \tag{4}$$

According to (2)–(4), the inequality $d(\alpha, \beta) < \frac{\varepsilon}{2M2^{n_0}}$ implies that $\|\delta_n(\alpha) - \delta_n(\beta)\| < \varepsilon$. The operator δ_n, $n = 1, 2, \ldots$, is continuous at $\alpha = (\alpha_k)$.

Suppose that all sequences in Ω are limitable by the matrix $\mathscr{A} = (a_{nk})$. Hence, there exists $\lim_{n\to\infty} \delta_n(\alpha) = \delta(\alpha)$ for each $\alpha = (\alpha_k) \in \Omega$. Then, by Theorem B, the set of discontinuity points of the operator $\delta : \Omega \to X$ is a set of the first Baire category in Ω. On the other hand if $\alpha = (\alpha_k) \in \Omega$ and $\eta > 0$, then, in the open ball $S(\alpha, \eta)$, it is possible to find elements $\beta, \gamma \in \Omega$ such that

$$\|\delta(\beta) - \delta(\gamma)\| > \frac{1}{2}.$$

We choose $k_1 \in \mathbb{N}$ so that

$$\sum_{k>k_1} \frac{\|\alpha_k\|}{2^k} \leq \sum_{k>k_1} \frac{1}{2^k} < \frac{\eta}{4},$$

where, $\alpha = (\alpha_k) \in \Omega$. Put $\beta = (\beta_k)$, $\gamma = (\gamma_k)$, where $\beta_k = \gamma_k = \alpha_k$ for $k = 1, 2, \ldots k_1$ and $\beta_k = \Theta$, $\gamma_k = \xi$ for $k > k_1$. The element Θ is the null-element of X and $\xi \in X$ is an element with the property $\|\xi\| = 1$. Then

$$d(\alpha, \beta) = \sum_{k>k_1} \frac{\|\alpha_k - \Theta\|}{2^k} < \frac{\eta}{4}$$

and

$$d(\alpha,\gamma) = \sum_{k>k_1} \frac{\|\alpha_k - \xi\|}{2^k} < \frac{\eta}{2}.$$

Therefore, β and γ belong to $S(\alpha,\eta)$.

Due to the regularity the matrix $\mathscr{A}$ and Theorem A, it follows that

$$\delta(\beta) = \lim_{n\to\infty} \sum_{k=1}^{\infty} a_{nk}\beta_k = \lim_{n\to\infty} \sum_{k=1}^{k_1} a_{nk}\alpha_k$$

and

$$\delta(\gamma) = \lim_{n\to\infty} \sum_{k=1}^{\infty} a_{nk}\gamma_k = \lim_{n\to\infty} \sum_{k=1}^{k_1} a_{nk}\alpha_k + \lim_{n\to\infty} \sum_{k>k_1} a_{nk}\xi.$$

Since $\lim_{n\to\infty} \sum_{k>k_1} a_{nk} > \frac{1}{2}$, we have $\left\|\sum_{k>k_1} a_{nk}\xi\right\| = \left|\sum_{k>k_1} a_{nk}\right| \|\xi\| > \frac{1}{2}$ for every $n \in \mathbb{N}$ sufficiently large. Thus,

$$\|\delta(\gamma) - \delta(\beta)\| = \left\|\lim_{n\to\infty} \sum_{k>k_1} a_{nk}\xi\right\| > \frac{1}{2}.$$

Therefore the operator δ is discontinuous at each element α of Ω. Since Ω is a complete metric space, the set of all discontinuity points of δ is the set of the second Baire category in Ω — a contradiction. □

We introduce the notion of FK-space (see [1]).

Definition 1. A complete metric linear space (B,d) of sequences, i. e., $B \subset b$ (see (1)), is said to be an FK-space provided that the linear operators $\delta_k : B \to X$,

$$\delta_k(\alpha) = \alpha_k, \quad \alpha = (\alpha_k) \in B, \quad k = 1, 2\dots,$$

are continuous on B.

Proposition 1. *Let $(X, \|\cdot\|_1)$ be a Banach space. Let (B,d) be a complete metric linear space of sequences of elements of X. Then (B,d) is an FK-space if and only if the convergence in the sense of the metric d implies the pointwise convergence in the sense of the norm $\|\cdot\|_1$ for each $k = 1, 2, \dots$.*

Proof. Let (B,d) be an FK-space. The linear operator $\delta_k : B \to X$, $\delta_k(\alpha) = \alpha_k$, is continuous for each $k = 1, 2, \dots$. According to the Heine definition of continuity, the operator $\delta_k : B \to X$ is continuous if for each sequence $(\alpha^{(r)})_{r=1}^{\infty}$, $\alpha^{(r)} = (\alpha_s^{(r)})_{r=1}^{\infty}$, of elements of B such that $\alpha^{(r)} \to \alpha = (\alpha_s)$ as $r \to \infty$ by the metric d, we have $\delta_k(\alpha^{(r)}) \to \delta_k(\alpha)$ by the norm $\|\cdot\|_1$ as $r \to \infty$. In our case, $\delta_k(\alpha^{(r)}) = \alpha_k^{(r)}$ and $\delta_k(\alpha) = \alpha_k$ for each $r = 1, 2, \dots$. Consequently, $\alpha_k^{(r)} \to \alpha_k$ as $r \to \infty$. □

Now we denote by $F(\mathscr{A}) = \{\alpha = (\alpha_k) : \alpha_k \in B, k = 1, 2, \dots : \lim_{n\to\infty} \sum_{k=1}^{\infty} a_{nk}\alpha_k$ exists and equals c for $\alpha_k \to c\}$.

In the proof of Theorem 1 we shall use the following form of Banach's Subgroup Theorem (see [3]): *If G is any proper subgroup of a linear topological space E, then either G is of the first category in E, or G fails to satisfy the condition of Baire.*

Theorem 1. *Let (B, d) be an FK-space, $B_\infty \subset B \subset b$. Let $\mathscr{A} = (a_{nk})$ be a regular matrix of real numbers. Then the set $F(\mathscr{A})$ is of the first Baire category in B for every regular matrix $\mathscr{A}$.*

Proof. We set $\delta_{ns}(\alpha) = \sum_{k=1}^{s} a_{nk}\alpha_k,\ \alpha = (\alpha_k) \in B,\ n, s = 1, 2, \dots$. From the above assumption it follows that each of the operators $\delta_{ns} : B \to X$ is continuous on B. Thus, $\delta_n(\alpha) = \lim_{s\to\infty} \delta_{ns}(\alpha)$ is a linear operator of the first Baire class defined on the set $D_n = \{\alpha \in B :\ \lim_{s\to\infty} \delta_{ns}(\alpha)$ exists$\}$.

Obviously,

$$D_n = \left\{\alpha \in B : \forall p = 1, 2, \dots\ \exists s_0 \in \mathbb{N}\ \forall s, q \geq s_0 :\ \|\delta_{ns}(\alpha) - \delta_{nq}(\alpha)\|_1 \leq \frac{1}{p}\right\}$$
$$= \bigcap_{p=1}^{\infty} \bigcup_{s_0=1}^{\infty} \bigcap_{s\geq s_0} \bigcap_{q\geq s_0} \left\{\alpha \in B : \|\delta_{ns}(\alpha) - \delta_{nq}(\alpha)\|_1 \leq \frac{1}{p}\right\}.$$

Let $s > q$. Then every operator $\|\delta_{ns}(\alpha) - \delta_{nq}(\alpha)\|_1$ is continuous, every set D_n is an $F_{\sigma\delta}$ set. The common domain of all the operators δ_n, the set $\bigcap_{n=1}^{\infty} D_n$, is also an $F_{\sigma\delta}$ set.

Put $S(\mathscr{A}) = \bigcap_{n=1}^{\infty} D_n$. Then $F(\mathscr{A}) \subset S(\mathscr{A})$ and

$$F(\mathscr{A}) = \left\{\alpha \in S(\mathscr{A}) :\ \lim_{n\to\infty} \delta_n(\alpha) \text{ exists}\right\}$$
$$= \left\{\alpha \in\ S(\mathscr{A}) :\ \forall p = 1, 2, \dots\ \exists n_0 \in \mathbb{N}\ \forall m, n \geq n_0 :\ \|\delta_n(\alpha) - \delta_m(\alpha)\|_1 \leq \frac{1}{p}\right\}$$
$$= \bigcap_{p=1}^{\infty} \bigcup_{n_0=1}^{\infty} \bigcap_{n\geq n_0} \bigcap_{m\geq n_0} \left\{\alpha \in S(\mathscr{A}) :\ \|\delta_n(\alpha) - \delta_m(\alpha)\|_1 \leq \frac{1}{p}\right\}.$$

Each of the operators $\|\delta_n(\alpha) - \delta_m(\alpha)\|_1$ is a Baire one operator, hence the set $F(\mathscr{A})$ is a $G_{\delta\sigma\delta}$ set, therefore $F(\mathscr{A})$ satisfies the condition of Baire. Since $F(\mathscr{A})$ is a proper subgroup of B, Theorem 1 is a consequence of Banach's Subgroup Theorem. □

The following Lemma is proved in [5].

Lemma 2. Let Z be a convex nowhere dense set in a Banach space X. Then Z is strongly porous in X.

We shall show that in the set B_∞ endowed with the norm $\|\alpha\|_2 = \sup\{\|\alpha_k\|_1\}$, $\alpha = (\alpha_k) \in B_\infty$, the set $F(\mathscr{A})$ is strongly porous.

Theorem 2. *In the Banach space $(B_\infty, \|\cdot\|_2)$ the set $F(\mathscr{A})$ is strongly porous in B_∞ for any regular matrix $\mathscr{A}$.*

Proof. Each of the operators $\delta_n : B_\infty \to X$, $\delta_n(\alpha) = \sum_{k=1}^{\infty} a_{nk}\alpha_k$, $n = 1, 2, \ldots$, fulfils the Lipschitz condition with the same constant M.

Due to Theorem A, for $\alpha = (\alpha_k) \in B_\infty$, $\beta = (\beta_k) \in B_\infty$, we have

$$\|\delta_n(\alpha) - \delta_n(\beta)\|_1 = \left\| \sum_{k=1}^{\infty} a_{nk}(\alpha_k - \beta_k) \right\|_1 \le \|\alpha - \beta\|_2 \sum_{k=1}^{\infty} |a_{nk}| \le M \|\alpha - \beta\|_2 \,. \tag{5}$$

The set $F(\mathscr{A})$ is a subspace of B_∞ and it is convex. We show that $B_\infty \setminus F(\mathscr{A})$ is open in B_∞. Let $\alpha = (\alpha_k) \in B_\infty \setminus F(\mathscr{A})$. Then the sequence $(\delta_n(\alpha))$ does not satisfy Cauchy's conditions, i. e.,

$$\exists \varepsilon > 0 \ \forall n_0 \in \mathbb{N} \ \exists m, n > n_0 : \ \|\delta_n(\alpha) - \delta_m(\alpha)\|_1 \ge \varepsilon.$$

Let $\beta = (\beta_k) \in B_\infty$ such that

$$\|\alpha - \beta\|_2 < \frac{\varepsilon}{4M}. \tag{6}$$

Then, by (5), we have

$$\begin{aligned}\varepsilon &\le \|\delta_n(\alpha) - \delta_m(\alpha)\|_1 \\ &\le \|\delta_n(\alpha) - \delta_n(\beta)\|_1 + \|\delta_n(\beta) - \delta_m(\beta)\|_1 + \|\delta_m(\beta) - \delta_m(\alpha)\|_1 \\ &< 2M\|\alpha - \beta\|_2 + \|\delta_n(\beta) - \delta_m(\beta)\|_1.\end{aligned}$$

By (6), we have

$$\frac{\varepsilon}{2} < \|\delta_n(\beta) - \delta_m(\beta)\|_1.$$

Thus, $\beta = (\beta_k) \notin F(\mathscr{A})$ and

$$S\left(\alpha, \frac{\varepsilon}{4M}\right) \subset B_\infty \setminus F(\mathscr{A}),$$

which means that $B_\infty \setminus F(\mathscr{A})$ is open in B_∞.

Therefore, $F(\mathscr{A})$ is a closed nowhere dense subset of B_∞, and Theorem 2 follows from Lemma 2. □

References

[1] Bennet, G. and Kalton, N. J.: *Consistency theorems for almost convergence*, Trans. Amer. Math. Soc., **198** (1974), 23–43.

[2] Connor, J.: *A short proof of Steinhaus' theorem on summability*, Amer. Math. Monthly, **92** (1985), 420–421.

[3] Kelley, J. L. and Namioka, I.: *Linear Topological Spaces*, New York, 1963.

[4] Kostyrko, P.: *Convergence fields of regular matrix transformations*, Tatra Mount. Math. Publ., **28** (2004), 153–157.

[5] Olevskii, V.: *A note on the Banach-Steinhaus theorem*, Real Anal. Exchange, **17** (1991–1992), 399–401.

[6] Petersen, G.M.: *Regular matrix transformations*, London, 1966.

[7] Šalát, T.: *On convergence fields of regular matrix transformations*, Czechoslovak Math. J., **26** (1976), No. 101, 613–627.
[8] Sikorski, R.: *Funkcje rzeczywiste. I*, Warszawa, 1958.
[9] Visnyai, T.: *Limitation Methods of Sequences of Elements of a Banach Space* [in Slovak], Author's Abstract of PhD Thesis, Bratislava, 2002.
[10] Zajíček, L.: *Porosity and σ-porosity*, Real Anal. Exchange, **13** (1987-1988), 314–350.

Author's address

T. Visnyai:
Comenius University, Faculty of Mathematics, Physics and Informatics, Mlynská dolina, 842 48, Bratislava 4, Slovakia
E-mail address: visnyai@zoznam.sk

Miskolc Mathematical Notes

Abstracted in *Zentralblatt für Mathematik*, *Mathematical Reviews*, and Реферативный журнал „Математика”

Responsible for publication: Rector of the University of Miskolc
Published by the Miskolc University Press under the leadership of Dr. József Péter
Responsible for duplication: Works manager Mária Kovács
Number of copies printed: 200
Put to the Press on April 24, 2006
Number of permission: TU 2006-320-ME **HU e-ISSN 1787-2413**