Zeitschrift für Analysis und ihre Anwendungen
Journal for Analysis and its Applications
Volume 13 (1994), No. 4, 667–681

Chebyshev Polynomial Iterations and Approximate Solutions of Linear Operator Equations

A. P. Zabrejko and P. P. Zabrejko

Abstract. An iteration scheme for the approximate solution of a linear operator equation in a Banach space is discussed from the viewpoint of Chebyshev polynomials. The optimal rate of convergence is described by numerical characteristics which are similar to (but different from) the classical Chebyshev constants. The abstract results are illustrated by some examples which frequently arise in applications.

Keywords: *Polynomial iterations, Chebyshev polynomials, Chebyshev characteristics*

AMS subject classification: Primary 41A10, secondary 47A50, 65J10

0. Introduction

Consider the equation

$$Ax = f \tag{1}$$

where A is a continuous linear operator in a Hilbert or Banach space $\boldsymbol{X}$, and $f \in \boldsymbol{X}$ is given. One of the most useful methods for finding approximate solutions of equation (1) is based on the successive iterations

$$x_{j+1} = P(A)x_j + \tilde{f} \qquad (j = 0, 1, 2, \ldots) \tag{2}$$

with some initial value x_0; here $P(A)$ is the value of some polynomial $P = P(\lambda)$ at A, and $\tilde{f} \in \boldsymbol{X}$ is an appropriate element depending on f and P. If the iterated sequence (2) converges, its limit x_* solves the equation

$$x = P(A)x + \tilde{f}. \tag{3}$$

Consequently, if the equations (1) and (3) are equivalent, the iterations (2) provide approximate solutions of the problem (1). It is clear that the polynomial P should then

A. P. Zabrejko: Universitet Druzhby Narodov, Matematicheskij Fakultet, Ul. Miklukho-Maklaja 16, R-117923 Moscow, Russia
P. P. Zabrejko: Belgosuniversitet, Matematicheskij Fakultet, Pl. Nezavisimosti, BR-220050 Minsk, Belarus

ISSN 0232-2064 / $ 2.50

be chosen in such a way that the convergence of the iterations (2) to x_* is as fast as possible.

This problem has been studied by many authors. For instance, in case of a positive definite self-adjoint operator A in a Hilbert space $\boldsymbol{X}$ several results have been obtained in [1 - 3, 5, 16, 22, 25]. For such operators it is rather obvious that, among all polynomials of degree n, the Chebyshev polynomial $T(\lambda)$ of A (i.e. the polynomial T satisfying the condition $T(0) = 1$ and having the smallest maximum on the spectrum of A) give the fastest rate of convergence of the iterations (2).

Later it became clear that the basic ideas which had been employed for studying the iterations (2) in the setting of positive definite self-adjoint operators, carry over as well to equations with arbitrary continuous linear operators in Banach spaces (see, e.g., [12, 13, 15]). In this connection, however, the idea of using Chebyshev polynomials receded somewhat into the background.

The purpose of the present paper is to discuss the iteration scheme (2) in general Banach spaces from the viewpoint of Chebyshev polynomials. It turns out that this leads to several new problems which are related to the study of Chebyshev polynomials for various subsets of the complex plane $\boldsymbol{C}$, and which bear a certain resemblance to the classical problems studied in [6, 7, 9, 10]. In particular, we shall show that the optimal rate of convergence of the successive iterations (2) is given by some numerical characteristic χ_n depending on the spectrum of A (which is different but analogous to the classical Chebyshev constants); moreover, the limit

$$\chi = \lim_{n\to\infty} \chi_n^{1/n}$$

exists and this new characteristic χ is analogous to the classical Fekete transfinite diameter of the spectrum of A (see [9, 10]).

The calculation of the corresponding Chebyshev polynomials and characteristics χ_n for different sets is a difficult problem even for $n = 1$. Below we shall give three examples to illustrate how to carry out explicit calculations in case $n = 1$; this leads to some new results on the convergence of the approximations (2) (cf. [12, 13]). Moreover, we provide a "recipe" for such computations in a fairly general setting. Unfortunately, in case $n > 1$ the problem of finding the precise values of χ_n and χ is still unsolved; an exceptional case is when the set under consideration is a real interval (see [7] and Lemma 6 below). Some results about Chebyshev characteristics in fact may be found in [17 - 20, 23, 24] but they concern only the case when the set is a union of finitely many intervals and points. It is worth mentioning that one may find good approximations of Chebyshev polynomials by using tools from the modern theory of extremal problems.

1. The Chebyshev polynomial methods

Let $\boldsymbol{X}$ be a complex Banach space and A a continuous linear operator in $\boldsymbol{X}$. Consider the equation (1) in the space $\boldsymbol{X}$, where $x \in \boldsymbol{X}$ is unknown and $f \in \boldsymbol{X}$ is given. The main purpose of this paper is to examine iteration methods for approximate solutions to equation (1) which are realized through successive calculations of powers of A and involve the right-hand side f and some initial approximation x_0. To describe this idea more precisely, let

$$P(\lambda) = c_0 + c_1\lambda + \ldots + c_n\lambda^n$$

be an arbitrary complex polynomial of degree n, and let

$$P(A) = c_0 I + c_1 A + \dots + c_n A^n$$

as usual. It is obvious that $P(A)$ defined by this relation is a continuous linear operator in $\boldsymbol{X}$.

Consider now the successive iterations (2), where $x_0 \in \boldsymbol{X}$ is some initial approximation. If the sequence $(x_j)_j$ given in (2) converges, its limit x_* is a solution of the equation (3). Thus, as already stated, the sequence $(x_j)_j$ converges to the solution of equation (1) if the equations (1) and (3) are equivalent.

Lemma 1. *Let $\boldsymbol{X}$ be a complex Banach space and A a continuous linear operator in $\boldsymbol{X}$. If*

$$P(0) = 1, \tag{4}$$

then every solution of equation (1) is a solution of equation (3), where

$$\tilde{f} = Q(A)f, \tag{5}$$

and the polynomial Q is defined by

$$P(\lambda) = 1 - \lambda Q(\lambda). \tag{6}$$

Conversely, if condition (4) holds and 1 is not an eigenvalue of the operator $P(A)$, then any solution of equation (3), with $\tilde{f}$ given by (5), is a solution of equation (1).

Proof. Let x_* be a solution of equation (1). Then $A^j x_* = A^{j-1} f \quad (j = 1, 2, \dots)$ and hence, by (4),

$$P(A)x_* = x_* + c_1 f + \dots + c_n A^{j-1} f.$$

By the definition of Q we have

$$x_* = P(A)x_* + Q(A)f = P(A)x_* + \tilde{f},$$

i.e. x_* satisfies equation (3). Conversely, suppose that condition (4) holds, 1 is not an eigenvalue of $P(A)$, and x_* is a solution of equation (3). We have then

$$x_* = P(A)x_* + \tilde{f} = P(A)x_* + Q(A)f.$$

Assume now that x_* is not a solution of equation (1). Then the element $z_* = Ax_* - f$ is not zero and hence, by (5) and (6), we obtain

$$\begin{aligned} z_* - P(A)z_* &= (I - P(A))(Ax_* - f) \\ &= A(I - P(A))x_* - (I - P(A))f \\ &= AQ(A)f - AQ(A)f \\ &= 0. \end{aligned}$$

We conclude that $z_* = P(A)z_*$, contradicting the assumption that 1 is not an eigenvalue of the operator $P(A)$ ∎

Lemma 1 allows us to consider different iteration processes based on appropriately chosen polynomials P, which will converge to the solution of equation (1). Here one should try to choose a polynomial which gives the fastest convergence of the iterations (2); this emphasizes the need of studying the rate of convergence of the sequence $(x_j)_j$ given in (2).

In what follows, we shall call the iteration process (2) involving a polynomial P with condition $P(0) = 1$ a *polynomial iteration method of degree n*. In order to study such methods, we recall some basic facts on iterations of the operator equation

$$x = Bx + g \tag{7}$$

in a Banach space $\boldsymbol{X}$ (cf. [11, 15, 16, 26, 27]). According to the classical Gel'fand formula, the spectral radius $\rho(B) = \sup\{|\lambda| : \lambda \in \operatorname{sp} B\}$ of a continuous linear operator B in $\boldsymbol{X}$ may be calculated either by

$$\rho(B) = \lim_{j\to\infty} ||B^j||^{1/j} \qquad \text{or by} \qquad \rho(B) = \inf_j ||B^j||^{1/j}.$$

It is worth mentioning that the spectral radius $\rho(B)$ is also equal to the infimum of all possible operator norms $||B||$ of B, running over all norms on $\boldsymbol{X}$ which are equivalent to the original norm $||\cdot||$.

The following result is essential for what follows.

Lemma 2 (see [27]). *Suppose that $\rho(B) < 1$. Then the successive iterations*

$$x_{j+1} = Bx_j + g \qquad (j = 0, 1, 2, \ldots) \tag{8}$$

converge to the solution x_* of equation (7) for any $g \in \boldsymbol{X}$ and any initial value $x_0 \in \boldsymbol{X}$. Moreover, the estimate

$$||x_j - x_*|| \le \Delta_j(B)\,||g|| \qquad (j = 0, 1, 2, \ldots) \tag{9}$$

holds, where

$$\Delta_j(B) = ||B^j(I - B)^{-1}|| \qquad (j = 0, 1, 2, \ldots). \tag{10}$$

The analysis of the sequence (10) has not found much attention yet for general operators B. Usually, the rate of convergence of a positive sequence $(\delta_j)_j$ to zero is characterized by the corresponding *Ljapunov exponent*

$$\lambda = \limsup_{j\to\infty} \delta_j^{1/j}.$$

In fact, if $\lambda < 1$, the sequence $(\delta_j)_j$ converges to zero faster than any geometrical progression with ratio $\rho \in (\lambda, 1)$, but not faster than any geometrical progression with ratio $\rho \in (0, \lambda)$. Since

$$\lim_{j\to\infty} \Delta_j(B)^{1/j} = \rho(B),$$

the error $\delta_j = ||x_j - x_*||$ for the iterations (8) behaves, loosely speaking, like a geometrical progression with ratio $\rho = \rho(B)$. More precisely, one may show that an estimate

$$\Delta_j(B) \le H_B(\rho)\rho^j \qquad (\rho(B) < \rho < 1;\ j = 0, 1, 2, \ldots)$$

holds, where $H_B(\rho)$ is an *error-characteristic function* of the operator B. If $\boldsymbol{X}$ is finite-dimensional, this function has the asymptotic behavior

$$H_B(\rho) \asymp (\rho - \rho(B))^{1-d},$$

where d is the maximal order of the poles of the resolvent $R(\lambda, B) = (\lambda I - B)^{-1}$ on the spectral circumference $|\lambda| = \rho(B)$. If $\boldsymbol{X}$ is infinite dimensional, a similar result holds if the peripherical spectrum of B is Fredholm. If B is a self-adjoint or, more generally, a normal operator, the function $H_B(\rho)$ is constant and may be calculated by means of the spectral representation of B.

Lemma 2 gives us a rather simple device for studying the rate of convergence of the iterations (2). To make this precise, we need a useful, though trivial, version of the spectral mapping theorem (see [4]). Given a non-empty bounded subset $\mathcal{M} \subset \boldsymbol{C}$ and a polynomial P, we write

$$\nu(P, \mathcal{M}) = \sup\left\{|\mathcal{P}(\lambda)| : \lambda \in \mathcal{M}\right\}. \tag{11}$$

Lemma 3. *Let $\boldsymbol{X}$ be a complex Banach space and A a continuous linear operator in $\boldsymbol{X}$. Then $\nu(P, \mathrm{sp}A) = \rho(P(A))$.*

By Lemma 3, the rate of convergence of the iterations (2) essentially depends on the size of the Chebyshev norm $\nu(P, \mathcal{M})$ of the polynomial P on the compact set $\mathcal{M} = \mathrm{sp}A$. In particular, it is important for us to describe those polynomials P for which $\nu(P, \mathrm{sp}A) < 1$, and to make this description as explicit as possible.

Let us introduce some notation. Given a non-empty compact subset $\mathcal{M} \subset \boldsymbol{C}$ we write

$$\chi_n(\mathcal{M}) = \inf \nu(P, \mathcal{M}), \tag{12}$$

where the infimum is taken over all polynomials P of degree n with $P(0) = 1$. The number $\chi_n(\mathcal{M})$ $(n = 1, 2, \ldots)$ is called the *Chebyshev characteristic of order n* in what follows. A *Chebyshev polynomial* of degree n for $\mathcal{M}$ is a polynomial $T_{\mathcal{M},n}$ with $T_{\mathcal{M},n}(0) = 1$ at which the infimum in (12) is achieved, i.e.

$$\chi_n(\mathcal{M}) = \sup\left\{|T_{\mathcal{M},n}(\lambda)| : \lambda \in \mathcal{M}\right\}.$$

One can show (as in the classical situation, see [9, 10]) that, given $\mathcal{M} \subseteq \boldsymbol{C}$, there exists a Chebyshev polynomial of degree n for $\mathcal{M}$; this Chebyshev polynomial is unique if $\mathcal{M}$ contains infinitely many points and $0 \notin \mathcal{M}$.

We point out that the Chebyshev characteristics (12) have the properties

$$\chi_n(\mathcal{M}_1) \le \chi_n(\mathcal{M}_2) \qquad (\mathcal{M}_1 \subseteq \mathcal{M}_2 \subset \boldsymbol{C}) \tag{13}$$

$$\chi_{m+n}(\mathcal{M}) \le \chi_m(\mathcal{M})\chi_n(\mathcal{M}) \qquad (\mathcal{M} \subset \boldsymbol{C}) \tag{14}$$

$$\chi_n(\lambda\mathcal{M}) = \chi_n(\mathcal{M}) \qquad (\mathcal{M} \subset \boldsymbol{C}, \lambda \in \boldsymbol{C}). \tag{15}$$

The property (13) follows easily from the properties of the upper bound of nested sets. To see that inequality (14) is true, fix two Chebyshev polynomials $T_{\mathcal{M},m}$ and $T_{\mathcal{M},n}$ for

$\mathcal{M}$ of degree m and n, respectively. The product $P = T_{\mathcal{M},m}T_{\mathcal{M},n}$ is then a polynomial (not necessarily Chebyshev) of degree $m+n$ such that $P(0) = 1$ and

$$\sup\{|P(\lambda)|: \lambda \in \mathcal{M}\} \leq \chi_m(\mathcal{M})\chi_n(\mathcal{M}).$$

Passing to the infimum in this inequality, we conclude that $\chi_{m+n}(\mathcal{M}) \leq \chi_m(\mathcal{M})\chi_n(\mathcal{M})$. To prove the last property (15) it is sufficient to notice that the class of polynomials P satisfying $P(0) = 1$ is invariant under change of argument from z to λz.

From (14) it follows that thc limit

$$\chi(\mathcal{M}) = \lim_{n\to\infty} \chi_n(\mathcal{M})^{1/n} \tag{16}$$

exists; we call $\chi(\mathcal{M})$ the *limit Chebyshev characteristic* of the set $\mathcal{M}$. Evidently,

$$\chi(\mathcal{M}) = \inf_n \chi_n(\mathcal{M})^{1/n}. \tag{17}$$

We point out that the Chebyshcv characteristics $\chi_n(\mathcal{M})$ $(n = 1, 2, \dots)$ and $\chi(\mathcal{M})$ are different from classical Chebyshev characteristics considered, for example, in [9, 10]. For instance, while the new Chebyshev characteristics are homogeneous of degree 0 (see (15)) the classical Chebyshev characteristics $\xi_n(\mathcal{M})$ $(n = 1, 2, \dots)$ and $\xi(\mathcal{M})$ are homogeneous of degree 1 (i.e. $\xi_n(\lambda\mathcal{M}) = |\lambda|\xi_n(\mathcal{M})$ $(n = 1, 2, \dots)$ and $\xi(\lambda\mathcal{M}) = |\lambda|\xi(\mathcal{M})$ $(\mathcal{M} \subset \boldsymbol{C})$). This means, intuitively speaking, that the numbers $\chi_n(\mathcal{M})$ and $\chi(\mathcal{M})$ introduced above give information on the shape and the position of the set $\mathcal{M}$, rather than on its size.

We return to the iterations (2). For the reader's ease, we summarize our previous discussion with the following

Theorem 1. *Let $\boldsymbol{X}$ be a complex Banach space and A a continuous linear operator in $\boldsymbol{X}$. Then the following statements hold:*

a) *The iterations (2) with some polynomial P satisfying $P(0) = 1$ converge if and only if $\nu(P, \mathrm{sp}A) < 1$; moreover, the rate of convergence in this case may be estimated by*

$$\|x_j - x_*\| \leq H_B(\rho)\rho^j \qquad (\nu(P, \mathrm{sp}A) < \rho < 1;\ j = 0, 1, 2, \dots).$$

In particular, $H_{P(A)}(\rho) \asymp 1$ if $\mathrm{sp}\,A$ consists only of simple Fredholm points on the set $\{\lambda: |P(\lambda)| = \nu(P, \mathrm{sp}A)\}$.

b) *The polynomial iteration methods of degree n (or of arbitrary degree, respectively) apply to solving the equation (1) if and only if $\chi_n(\mathrm{sp}A) < 1$ $(\chi(\mathrm{sp}A) < 1$, respectively).*

c) *If this inequality holds, the fastest rate of convergence (in the sense of the Lyapunov exponent) of the iterations (2) is achieved for the Chebyshev polynomial $P = T_{\mathrm{sp}A,n}$; the corresponding Ljapunov exponent in this case is precisely $\chi_n(\mathrm{sp}A)$.*

When applying Theorem 1 to a specific problem, one encounters some natural questions. First of all, the question arises how to calculate $\chi_n(\mathrm{sp}A)$ at least for some values of n for the iteration (2). Next, the question arises how to calculate the norm bounds $\Delta_n(P(A))$ (see (10)) and the error characteristic function $H_{P(A)}(\rho)$ for the corresponding Chebyshev polynomials. These questions have not been answered yet in the literature.

It should be noted that in applications one usually does not know the spectrum sp$A$ of A explicitely, but one may *localize* sp$A$, i.e. find a compact set $\mathcal{M} \supseteq \mathrm{sp}A$. By the monotonicity property (13), for applying Theorem 1 it suffices then to show that $\chi_n(\mathcal{M}) < 1$ ($\chi(\mathcal{M}) < 1$, respectively). A necessary, but not sufficient condition for these inequalities to be true is the relation $0 \notin \mathcal{M}$.

2. Iteration methods of first order

In this section we examine the simplest case, namely Chebyshev characteristics and Chebyshev polynomials of first order for arbitrary sets in the complex plane $\boldsymbol{C}$. Every first degree polynomial P satisfying (4) may be written in the form

$$P(\lambda) = 1 - z^{-1}\lambda \tag{18}$$

and coincides with the Chebyshev polynomial $T_{\mathcal{M},1}$ if and only if z in (18) satisfies

$$\sup\left\{|1 - z^{-1}\lambda| : \lambda \in \mathcal{M}\right\} = \chi_1(\mathcal{M}). \tag{19}$$

Therefore, the problem of applying the iteration method of first order reduces to the problem of finding the number z in (19); this number z is called the *Chebyshev center* of the set $\mathcal{M}$. Given $\mathcal{M} \subset \boldsymbol{C}$ and $z \in \boldsymbol{C}$, by $\theta(z, \mathcal{M})$ we denote the radius of the minimal disc centered at z which contains the set $\mathcal{M}$.

It is known (see [12, 13]) that the inequality $\chi_1(\mathcal{M}) < 1$ is true if and only if $\mathcal{M}$ is situated in some halfplane $\mathrm{Re}\, c\lambda > 0$ ($c \in \boldsymbol{C}$ fixed).

Lemma 4. *For compact sets $\mathcal{M} \subset \boldsymbol{C}$, the equality*

$$\chi_1(\mathcal{M}) = \min_z |z|^{-1}\theta(z, \mathcal{M}) \tag{20}$$

holds.

Proof. From the definition of $\chi_1(\mathcal{M})$ it follows that

$$\begin{aligned}
\chi_1(\mathcal{M}) &= \min_z \sup\left\{|1 - z^{-1}\lambda| : \lambda \in \mathcal{M}\right\} \\
&= \min_z \sup\left\{|z|^{-1}|z - \lambda| : \lambda \in \mathcal{M}\right\} \\
&= \min_z |z|^{-1} \sup\left\{|z - \lambda| : \lambda \in \mathcal{M}\right\} \\
&= \min_z |z|^{-1}\theta(z, \mathcal{M})
\end{aligned}$$

as claimed ∎

Lemma 4 provides a rather simple method for calculating $\chi_1(\mathcal{M})$ and $T_{\mathcal{M},1}$ for compact sets in $\boldsymbol{C}$. To make this more transparent, some further notation is in order. Given $\mathcal{M} \subset \boldsymbol{C}$, by extr$\mathcal{M}$ we denote as usual the set of all extemal points of $\mathcal{M}$, and by round$\mathcal{M}$ the set of all rounding points of $\mathcal{M}$ (i.e. points z with the property that there exists a disc D such that $\mathcal{M} \subseteq D$ and $z \in \partial D$). Obviously, round$\mathcal{M} \subseteq$ extr$\mathcal{M}$; equality holds, for example, if $\mathcal{M}$ is a polygon. We remark that the set round$\mathcal{M}$ is not necessarily closed. Consider the (multivalued) map

$$D_{\mathcal{M}}(\lambda) = \left\{ z \in \mathcal{M} : \ |z - \lambda| = \theta(\lambda, \mathcal{M}) \right\}. \tag{21}$$

It is not hard to see that $D_{\mathcal{M}}(\lambda) \neq \emptyset$ and $D_{\mathcal{M}}(\lambda) \subseteq \mathrm{round}\mathcal{M}$ for all $\lambda \in \boldsymbol{C}$; moreover,

$$\bigcup \left\{ D_{\mathcal{M}}(\lambda) : \ \lambda \in \boldsymbol{C} \right\} = \mathrm{round}\mathcal{M}.$$

We use the multivalued function (21) for rewriting (20) in a form which is more suitable for our computations below. From the definition (21) it follows indeed that

$$\theta(z, \mathcal{M}) = |z - \lambda| \quad \text{for all} \quad \lambda \in D_{\mathcal{M}}(z).$$

Now, for $z \in \mathrm{round}\mathcal{M}$ let

$$D_{\mathcal{M}}^{(-1)}(z) = \left\{ \lambda \in \boldsymbol{C} : \ z \in D_{\mathcal{M}}(\lambda) \right\} \tag{22}$$

denote the large pre-image of z under the multivalued function (21). By the upper semi-continuity of the map (21), the set $D_{\mathcal{M}}^{(-1)}(z)$ is always closed; moreover,

$$\bigcup \left\{ D_{\mathcal{M}}^{(-1)}(z) : \ z \in \mathrm{round}\mathcal{M} \right\} = \boldsymbol{C}.$$

Consequently,

$$\begin{aligned}
\min_z |z|^{-1} \ & \sup \left\{ |z - \lambda| : \ \lambda \in \mathcal{M} \right\} \\
&= \min_z \left\{ |z|^{-1} \ |z - \lambda| : \ \lambda \in D_{\mathcal{M}}(\lambda) \right\} \\
&= \min_{z,\lambda} \left\{ |1 - z^{-1}\lambda| : \ \lambda \in D_{\mathcal{M}}(\lambda) \right\} \\
&= \min_{z,\lambda} \left\{ |1 - z^{-1}\lambda| : \ z \in D_{\mathcal{M}}^{(-1)}(\lambda) \right\} \\
&= \min_{\lambda \in \mathrm{round}\mathcal{M}} \ \min \left\{ |1 - z^{-1}\lambda| : \ z \in \ D_{\mathcal{M}}^{(-1)}(\lambda) \right\} \\
&= \min_{\lambda \in \mathrm{round}\mathcal{M}} \ \min \left\{ |1 - z| : \ z \in \ \lambda \left(D_{\mathcal{M}}^{(-1)}(\lambda) \right)^{-1} \right\} \\
&= \min_{\lambda \in \mathrm{round}\mathcal{M}} \ \mathrm{dist} \left(1, \lambda \left(D_{\mathcal{M}}^{(-1)}(\lambda) \right)^{-1} \right)
\end{aligned}$$

or, in other words,

$$\min_z |z|^{-1} \, \theta(z, \mathcal{M}) = \min_{\lambda \in \mathrm{round}\mathcal{M}} \ \mathrm{dist} \left(1, \lambda \left(D_{\mathcal{M}}^{(-1)}(\lambda) \right)^{-1} \right). \tag{23}$$

Formula (23) is of particular interest, since the set round$\mathcal{M}$ usually consists only of few points (e.g., if $\mathcal{M}$ is a polygon), and hence the right-hand side in (23) may be calculated rather easily. We summarize with the following

Theorem 2. *Let $\mathcal{M} \in \boldsymbol{C}$ be non-empty and compact. Then*

$$\chi_1(\mathcal{M}) = \min_{\lambda \in \mathrm{round}\mathcal{M}} \mathrm{dist}\left(1, \lambda\left(D_{\mathcal{M}}^{(-1)}(\lambda)\right)^{-1}\right).$$

Moreover, the Chebyshev center z for which (20) holds belongs to the set

$$\gamma(\mathcal{M}) = \bigcap \left\{D_{\mathcal{M}}^{(-1)}(\lambda): \ \lambda \in \delta(\mathcal{M})\right\}$$

where $\delta(\mathcal{M})$ consist of all points $\lambda \in \mathrm{round}\mathcal{M}$ for which the minimum in (23) is attained.

Let us now give three important examples which show how to apply Lemma 4 and Theorem 2 in practice.

Example 1. Let

$$\mathcal{M} = B(\mu, R) = \left\{\lambda \in \boldsymbol{C} : |\lambda - \mu| \leq R\right\} \tag{24}$$

where $\mu \in \boldsymbol{C}\backslash\{0\}$ and $0 < R < |\mu|$. We claim that then

$$\chi_1(\mathcal{M}) = |\mu|^{-1}R. \tag{25}$$

In fact, since the characteristic $\chi_1(\mathcal{M})$ is homogeneous of degree 0, as already observed, we have

$$\chi_1(\mathcal{M}) = \chi_1\big(B(\mu, R)\big) = \chi_1\big(\mu B(1, |\mu|^{-1}R)\big) = \chi_1\big(B(1, |\mu|^{-1}R)\big) = |\mu|^{-1}R$$

as claimed. Observe, in particular, that $\chi_1(\mathcal{M}) < 1$.

In the following two examples we calculate $\chi_1(\mathcal{M})$ for sets $\mathcal{M}$ which are symmetric with respect to the real axis. For such sets $\mathcal{M}$, the function

$$\phi(c) = \sup \left\{|1 - c\lambda| : \ \lambda \in \boldsymbol{C}\right\} \tag{26}$$

assumes the same value at pairs $(c, \bar{c})$ of conjugate complex numbers. Moreover, since (26) is a convex function, its minimum is necessarily attained on the real axis.

Example 2. Let

$$\mathcal{M} = S(m, R) = \left\{\lambda \in \boldsymbol{C} : \ \mathrm{Re}\,\lambda \geq m, \ |\lambda| \leq R\right\} \tag{27}$$

where $0 < m < R$. We claim that

$$\chi_1(\mathcal{M}) = \left(1 - R^{-2}m^2\right)^{1/2} \tag{28}$$

and that the Chebyshev center z is given by $z = m^{-1}R^2$. In fact, by the symmetry of $\mathcal{M}$ and the above geometric reasoning we conclude that

$$\theta(z, \mathcal{M})^2 = (m - z)^2 + R^2 - m^2,$$

hence ($c = z^{-1}$)

$$\phi(c)^2 = 1 - 2mc + R^2c^2.$$

This function is minimized precisely for $c = mR^{-2}$. Putting this into $\phi(c)$ yields (28).

Observe, in particular, that $\chi_1(\mathcal{M}) < 1$ also in this example. The following example is somewhat more complicated.

Example 3. Let

$$\mathcal{M} = \Pi(m, M, H) = \Big\{\lambda \in \boldsymbol{C} : \ m \leq \operatorname{Re}\lambda \leq M, \ |\operatorname{Im}\lambda| \leq H\Big\} \tag{29}$$

where $0 < m \leq M < \infty$ and $H \geq 0$. We claim that

$$\chi_1(\mathcal{M}) = \begin{cases} \dfrac{(4H^2 + (M-m)^2)^{1/2}}{M+m} & \text{if } \dfrac{m}{m^2+H^2} \geq \dfrac{2}{M+m} \\ \dfrac{H}{(m^2+H^2)^{1/2}} & \text{if } \dfrac{m}{m^2+H^2} < \dfrac{2}{M+m} \end{cases} \tag{30}$$

and that the Chebyshev center z in (19) is given by

$$z = \begin{cases} \dfrac{M+m}{2} & \text{if } \dfrac{m}{m^2+H^2} \geq \dfrac{2}{M+m} \\ \dfrac{m^2+H^2}{m} & \text{if } \dfrac{m}{m^2+H^2} < \dfrac{2}{M+m}. \end{cases} \tag{31}$$

In fact, again by a simple geometric reasoning we see that

$$\theta(z, \mathcal{M})^2 = \begin{cases} (M-z)^2 + H^2 & \text{if } 0 < z < \dfrac{m+M}{2} \\ (m-z)^2 + H^2 & \text{if } \dfrac{m+M}{2} < z < \infty. \end{cases}$$

A straightforward computation shows then that ($c = z^{-1}$)

$$\phi(c)^2 = \begin{cases} (M^2 + H^2)c^2 - 2Mc + 1 & \text{if } \dfrac{2}{m+M} \leq c < \infty \\ (m^2 + H^2)c^2 - 2mc + 1 & \text{if } 0 < c \leq \dfrac{2}{m+M}. \end{cases} \tag{32}$$

The behaviour of this function is as follows: in the interval $\big[0, \frac{2}{(m+M)}\big)$ the function (32) attains its minimum

$$\frac{H^2}{m^2+H^2} \qquad \text{if } \frac{m}{m^2+H^2} < \frac{2}{m+H}$$

or

$$\frac{4H^2 + (M-m)^2}{(M+m)^2} \qquad \text{if } \frac{m}{m^2+H^2} \geq \frac{2}{m+H}.$$

Likewise, in the interval $\big[\frac{2}{m+M}, \infty\big)$ the function (32) attains its minimum

$$\frac{4H^2 + (M-m)^2}{(M+m)^2}$$

always at the left endpoint $c = \frac{2}{m+M}$. Thus,

$$\min_c \phi(c)^2 = \begin{cases} \dfrac{4H^2 + (M-m)^2}{(M+m)^2} & \text{if } \dfrac{m}{m^2+H^2} > \dfrac{2}{M+m} \\ \dfrac{H}{(m^2+H^2)^{1/2}} & \text{if } \dfrac{m}{m^2+H^2} \leq \dfrac{2}{M+m} \end{cases}$$

which implies (31). Observe that in case $\frac{m}{m^2+H^2} = \frac{2}{m+M}$ we simply get

$$\chi_1(\mathcal{M}) = \left(\frac{M-m}{M+m}\right)^{1/2}.$$

Example 3 becomes particularly simple in case $H = 0$, i.e. $\mathcal{M} = [m, M]$. Here we get

$$\chi_1(\mathcal{M}) = \frac{M-m}{M+m} \qquad \text{and} \qquad z = \frac{M+m}{2}.$$

We point out that we have again $\chi_1(\mathcal{M}) < 1$ in this example.

We are now going to exploit the above examples to derive rather subtle assertions on the convergence of first order polynomial iteration methods for (2). The following three theorems are immediate consequences of the calculations in the corresponding examples.

Theorem 3. *Let $\boldsymbol{X}$ be a complex Banach space and A a continuous linear operator in $\boldsymbol{X}$ such that* sp$A \subseteq B(\mu, R)$ *for some* $0 < R < |\mu|$. *Then the iterations*

$$x_{j+1} = (I - cA)x_j + cf \qquad (j = 0, 1, 2, \ldots) \tag{33}$$

with $c = \mu^{-1}$ converge, for each $f \in \boldsymbol{X}$, to the solution x_* of the equation (1). Moreover, the Ljapunov exponent of this convergence is precisely $|\mu|^{-1}R$.

Theorem 4. *Let $\boldsymbol{X}$ be a complex Banach space and A a continuous linear operator in $\boldsymbol{X}$ such that* sp$A \subseteq S(m, R)$ *for some* $0 < m < R$. *Then the iterations*

$$x_{j+1} = (I - cA)x_j + cf \qquad (j = 0, 1, 2, \ldots) \tag{34}$$

with $c = mR^{-2}$ converge, for each $f \in \boldsymbol{X}$, to the solution x_* of the equation (1). Moreover, the Ljapunov exponent of this convergence is precisely $(1 - m^2R^{-2})^{1/2}$.

Theorem 5. *Let $\boldsymbol{X}$ be a complex Banach space and A a continuous linear operator in $\boldsymbol{X}$ such that* sp$A \subseteq \Pi(m, M, H)$ *for some* $0 < m \leq M < \infty$ *and* $H \geq 0$. *Then the iterations*

$$x_{j+1} = (I - cA)x_j + cf \qquad (j = 0, 1, 2, \ldots) \tag{35}$$

with

$$c = \min\left\{\frac{2}{M+m}, \frac{m}{m^2+H^2}\right\}$$

converge, for each $f \in \boldsymbol{X}$, to the solution x_* of the equation (1). Moreover, the Ljapunov exponent of this convergence is precisely given by

$$\chi_1(\mathcal{M}) = \begin{cases} \dfrac{(4H^2 + (M-m)^2)^{1/2}}{M+m} & \text{if } \dfrac{m}{m^2+H^2} \geq \dfrac{2}{M+m} \\ \dfrac{H}{(m^2+H^2)^{1/2}} & \text{if } \dfrac{m}{m^2+H^2} < \dfrac{2}{M+m}. \end{cases}$$

In order to verify the hypotheses of these theorems, we have to find suitable localizations for the spectrum of A. If $\boldsymbol{X}$ is a Hilbert space, we may use, for instance, the classical fact [8] that $\mathrm{sp}A \subseteq W(A)$ where $W(A) = \overline{\{(Ax,x) : \|x\| = 1\}}$ denotes the *numerical range* of the operator A. The classical Toeplitz-Hausdorff theorem states that the numerical range of a continuous linear operator is always compact and convex; if A is normal, $W(A)$ coincides with the closed convex hull of $\mathrm{sp}A$.

Theorem 5, say, may be applied in this way if the operator A satisfies an estimate

$$m\|x\|^2 \leq \mathrm{Re}\,(Ax,x) \leq M\|x\|^2, \qquad |\mathrm{Im}\,(Ax,x)| \leq H\|x\|^2. \tag{36}$$

In fact, (36) immediately implies that $W(A) \subseteq \Pi(m,M,H)$. The corresponding partial case of Theorem 5 improves some previous results from [12, 13].

Similar results on the localization of $\mathrm{sp}A$ may be formulated through Lumer's numerical range $W(A)$ [21] of an operator A in a Banach space $\boldsymbol{X}$.

3. Iteration methods of higher order

It is natural to try to carry over the formulas we derived for $\chi_1(\mathcal{M})$ to the higher order Chebyshev characteristics $\chi_n(\mathcal{M})$ $(n > 1)$. As far as we know, no explicit information on $\chi_n(\mathcal{M})$ $(n > 1)$ is available in the literature. We confine ourselves here to some general remarks.

Any polynomial P of degree n which satisfies (4) can be written in the form

$$P(\lambda) = (1 - z_1^{-1}\lambda)...(1 - z_n^{-1}\lambda)$$

where $z_1, \dots, z_n \in \boldsymbol{C}$ are the roots of P. We use this to state a formula for $\chi_n(\mathcal{M})$ $(n > 1)$ which is parallel to that given in Lemma 4.

Lemma 5. *For compact sets $\mathcal{M} \subset \boldsymbol{C}$, the equality*

$$\chi_n(\mathcal{M}) = \min_{z_1,\dots,z_n} |z_1 \cdots z_n|^{-1}\, \Psi(z_1, \dots, z_n; \mathcal{M})$$

holds, where

$$\Psi(z_1, \dots, z_n; \mathcal{M}) = \max\Big\{|\lambda - z_1| \cdots |\lambda - z_n| :\ \lambda \in \mathcal{M}\Big\}.$$

Proof. From the definition of $\chi_n(\mathcal{M})$ it follows that

$$\begin{aligned}\chi_n(\mathcal{M}) &= \min_{z_1,\ldots,z_n} \max \left\{ |(1 - z_1^{-1}\lambda)\cdots(1 - z_n^{-1}\lambda)| : \lambda \in \mathcal{M} \right\} \\ &= \min_{z_1,\ldots,z_n} \max \left\{ |z_1 \cdots z_n|^{-1} \; |(z_1 - \lambda)\cdots(z_n - \lambda)| : \lambda \in \mathcal{M} \right\} \\ &= \min_{z_1,\ldots,z_n} |z_1 \cdots z_n|^{-1} \max \left\{ |(z_1 - \lambda)\cdots(z_n - \lambda)| : \lambda \in \mathcal{M} \right\} \\ &= \min_{z_1,\ldots,z_n} |z_1 \cdots z_n|^{-1} \; \Psi(z_1,\ldots,z_n;\mathcal{M})\end{aligned}$$

as claimed ∎

We remark that an analogue of Theorem 2 for $\chi_n(\mathcal{M})$ $(n > 1)$ leads to the study of *lemniscates* (*Cassini curves*) of the form

$$S(z_1,\ldots,z_n;r) = \left\{ \lambda \in \boldsymbol{C} : \left|(1 - z_1^{-1}\lambda)\cdots(1 - z_n^{-1}\lambda)\right| = r \right\}$$

and of domains of the form

$$B(z_1,\ldots,z_n;r) = \left\{ \lambda \in \boldsymbol{C} : \left|(1 - z_1^{-1}\lambda)\cdots(1 - z_n^{-1}\lambda)\right| \le r \right\}$$

which play here the same role as circumferences and discs in the case $n = 1$. The corresponding computations, however, are very cumbersome and do not lead to simple formulas for $\chi_n(\mathcal{M})$ $(n > 1)$. The only exception is the case of a real interval $\mathcal{M} = [m, M]$ $(0 < m \le M < \infty)$. Here the following is known (see [7, 11, 16]).

Lemma 6. *Let $\mathcal{M} = [m, M]$, where $0 < m \le M < \infty$. Then the Chebyshev polynomial $T_{\mathcal{M},n}$ of order n is given by*

$$T_{\mathcal{M},n}(\lambda) = \frac{\cos n \arccos\left((M-m)^{-1}(2\lambda - M - m)\right)}{\cos n \arccos\left((M-m)^{-1}(-M-m)\right)} \qquad (m \le \lambda \le M)$$

and the corresponding Chebyshev characteristic of $\mathcal{M}$ of order n is given by

$$\chi_n(\mathcal{M}) = |T_{\mathcal{M},n}(0)|. \tag{37}$$

Combining Lemma 6 with Theorem 1 we arrive at the following

Theorem 6. *Let $\boldsymbol{X}$ be a complex Banach space and A a continuous linear operator in $\boldsymbol{X}$ such that* sp$A \subseteq [m, M]$ *for some $0 < m \le M < \infty$. Then the iterations*

$$x_{j+1} = T_{\mathcal{M},n}(A)x_j + Q_n(A)f \qquad (j = 0, 1, 2, \ldots) \tag{38}$$

with the Chebyshev polynomial $T_{\mathcal{M},n}$ and polynomial $Q_n(\lambda) = \lambda^{-1}(1 - T_{\mathcal{M},n}(\lambda))$ converge, for each $f \in \boldsymbol{X}$, to the solution x_* of the equation (1). Moreover, the Ljapunov exponent of this convergence is precisely $|T_{\mathcal{M},n}(0)|$.

We mention the papers [17 - 20, 23, 24], where indeed $T_{\mathcal{M},n}$ and $\chi_n(\mathcal{M})$ are studied in the case when $\mathcal{M}$ is a finite union of intervals and points. Although no explicit formulas for $\chi_n(\mathcal{M})$ or $T_{\mathcal{M},n}(\lambda)$ are contained in these papers, it is shown that $\chi_n(\mathcal{M}) < \chi_n(\operatorname{co}\mathcal{M})$.

The authors express their gratitude to M. A. Krasnosel'skij and J. Appell for several helpful and interesting discussions, Mrs. H. Grieco for her help in preparing the manuscript of this article, and the nice girls from the Würzburg University Math Library for the opportunity to work with the rich mathematical literature.

References

[1] Abramov, A. A.: *On a method for accelerating iteration processes* (in Russian). Dokl. Akad. Nauk SSSR 74 (1950), 1051 - 1052.

[2] Birman, M. Sh.: *Some estimates for the method of steepest descent* (in Russian). Uspekhi Mat. Nauk 5 (1950), 152 - 155.

[3] Birman, M. Sh.: *On the computation of eigenvalues with the method of steepest descent* (in Russian). Uchen. Zap. Leningr. Gorn. Inst. 27 (1952), 209 - 216.

[4] Dunford, N. and J. T. Schwartz: *Linear Operators*. Part I: *General Theory*. Leyden: Int. Publ. 1963.

[5] Gavurin, M. K.: *The application of polynomials of best approximation for accelerating the convergence of iteration processes* (in Russian). Uspekhi Mat. Nauk 13 (1950), 156 - 160.

[6] Goluzin, G. M.: *Geometrical Theory of Functions of a Complex Variable* (in Russian). Moscow: Nauka 1966.

[7] Goncharov, V. L.: *Theory of Interpolation and Approximation of Functions* (in Russian). Moscow - Leningrad: Fizmatgiz 1954.

[8] Halmos, P.: *A Hilbert Space Problem Book*. Berlin: Springer - Verlag 1974.

[9] Hille, E.: *Analytic Function Theory* (Vol. 2). Boston: Ginn 1962.

[10] Hille, E.: *Some geometrical extremal problems*. J. Austral. Math. Soc. 6 (1966), 122 - 128.

[11] Kantorovich, L. V. and G. P. Akilov: *Functional Analysis* (in Russian). Moscow: Nauka 1984.

[12] Kozjakin, V. S. and M. A. Krasnosel'skij: *On some problems related to the method of minimal residuals* (in Russian). Zhurn. Vychisl. Mat. Mat. Fiz. 19 (1979), 508 - 510.

[13] Kozjakin, V. S and M. A. Krasnosel'skij: *Some remarks on the method of minimal residuals*. Numer. Funct. Anal. Optim. 4 (1981/82), 211 - 239.

[14] Krasnosel'skij, M. A.: *On the solutions of equations with self-adjoint operators by the method of successive approximations* (in Russian). Uspekhi Mat. Nauk 14 (1961), 161 - 165.

[15] Krasnosel'skij, M. A., Lifshits, E. A. and A. V. Sobolev: *Positive Linear Systems - the Method of Positive Operators*. Berlin: Heldermann Verlag 1989.

[16] Krasnosel'skij, M. A., Vajnikko, G. M., Zabrejko, P. P., Rutitskij, Ja. B. and V. Ja. Stetsenko: *Approximate Solutions of Operator Equations*. Groningen: Noordhoff 1972.

[17] Lebedev, V. I.: *Iteration methods for solving operator equations with a spectrum contained in several intervals*. J. Comp. Math. Math. Phys. 9 (1969), 1247 - 1252.

[18] Lebedev, V. I. and S. A. Finogenov: *Ordering of the iteration parameters in the cyclic Chebyshev iteration methods*. J. Comp. Math. Math. Phys. 11 (1971), 425 - 438.

[19] Lebedev, V. I. and S. A. Finogenov: *Solution of the parameter ordering problem in Chebyshev iteration methods*. J. Comp. Math. Math. Phys. 13 (1973), 21 - 41.

[20] Lebedev, V. I. and A. S. Finogenov: *Utilization of ordered Chebyshev parameters in iteration methods*. J. Comp. Math. Math. Phys. 16 (1976), 70 - 83.

[21] Lumer, G.: *Semi-inner product spaces*. Trans. Amer. Math. Soc. 10 (1961), 29 - 43.

[22] Natanson, I. P.: *On the theory of approximate solution of equations* (in Russian). Uchen. Zap. Leningr. Ped. Inst. 64 (1948), 93 - 104.

[23] Samokish, B. A.: *Investigation of the speed of convergence of the method of steepest descent* (in Russian). Uspekhi Mat. Nauk 12 (1957), 238 - 240.

[24] Samokish, B. A.: *The method of steepest descent in the problem of eigenvalues of semi-bounded operators* (in Russian). Izv. VUZov Mat. 5 (1958), 105 - 114.

[25] WIARDA, G.: *Integralgleichungen.* Leipzig - Berlin 1930.

[26] Zabrejko, P. P.: *Domains of convergence of the method of successive approximations for linear equations* (in Russian). Dokl. Akad. Nauk BSSR 29 (1985), 201 - 204.

[27] Zabrejko, P. P.: *Error estimates for successive approximations and spectral properties of linear operators.* Numer. Funct. Anal. Optim. 11 (1990), 823 - 838.

Received 29.04.1994