SEMI ITERATIVE METHODS FOR THE DRAZIN INVERSE SOLUTION OF LINEAR EQUATIONS IN BANACH SPACES

N. Castro González
Departamento de Matemática Aplicada
Facultad de Informática
Universidad Politécnica de Madrid
Spain

J. J. Koliha
Department of Mathematics and Statistics
University of Melbourne
Australia

ABSTRACT

We study semi-iterative methods for an approximate solution of a linear equation $x = Tx + c$ with a bounded linear operator T acting on a Banach space X, considering separately the cases when $\lambda = 1$ is an isolated and accumulation spectral point for T. We give necessary and sufficient conditions for the convergence of a semi-iterative method, utilizing the generalized Drazin inverse introduced by the second author.

AMS subject classification: primary 47A50, 47A10, secondary 47A35, 65J10

Keywords and phrases: Linear equations, semi-iterative methods, Drazin inverse.

1 INTRODUCTION AND PRELIMINARIES

In [1], Castro extended results of Eiermann, Marek and Niethammer [2] from the matrix setting to a class of bounded linear operators on Banach spaces. The purpose of this paper is to analyze the convergence of semi-iterative methods applied to linear operator equation $x = Tx + c$ with as few restrictions on the operator T as possible. We lay a foundation that can serve as a theoretical background for numerical applications to infinite systems of linear equations, abstract Cauchy problems, partial differential equations and Markov processes. In the process we obtain approximations of the generalized Drazin inverse introduced by Koliha [6], and an extension of previous results to the case when the resolvent of T has an essential singularity at 1. By using a different approach, we achieve a simplification of proofs compared with [1] and [2].

The results of this paper can be applied to approximate solutions of more general equations $Ax = b$ if we consider a splitting $A = M - N$ such that M^{-1} exists. We then reduce the equation to an equivalent equation $x = Tx + M^{-1}b$ with $T = M^{-1}N$. The theory developed here provides us with the possibility of constructing effective methods for computing approximate solutions combining a semi-iterative method applied to the nondiscretized equation and discretizing them at each certain number of iterations, and applying another semi-iterative method to the discretized problem.

Semi-iterative methods are used to solve the equation

$$x = Tx + c \qquad (c \text{ given}), \tag{1.1}$$

where T is a bounded linear operator on X, with $I - T$ not necessarily invertible. Many of the results of this paper will require the concept of the (generalized) Drazin inverse of the operator $I - T$.

By $\mathcal{B}(X)$ we denote the Banach algebra of all bounded linear operators on a complex Banach space X with the norm $\|T\| = \sup\{\|Tx\| : \|x\| \leq 1\}$. The nullspace and the range of $T \in \mathcal{B}(X)$ will be denoted by $\mathcal{N}(T)$ and $\mathcal{R}(T)$, respectively. Further, $\sigma(T)$ will denote the spectrum of T, $r(T)$ the spectral radius of T, $\rho(T)$ the resolvent set of T, and $R(\lambda; T) = (\lambda I - T)^{-1}$ the resolvent of T. The *peripheral spectrum* of T is the set of all $\lambda \in \sigma(T)$ such that $|\lambda| = r(T)$. The set of all isolated points of $\sigma(T)$ will be written as $\operatorname{iso}\sigma(T)$, and the set of all accumulation points of $\sigma(T)$ as $\operatorname{acc}\sigma(T)$. An operator $T \in \mathcal{B}(X)$ is a *projection* if $T^2 = T$. A *quasinilpotent* operator T is algebraically characterized by the condition that $I - TU$ is invertible for every $U \in \mathcal{B}(X)$ commuting with T; topologically, this is equivalent to $r(T) = 0$.

We call a direct sum $X = X_1 \oplus X_2$ of subspaces of X a *topological direct sum* if both X_1 and X_2 are closed; otherwise we say that the direct sum is *algebraic*.

2 DRAZIN INVERSE

In practice we often encounter a situation when the operator $T \in \mathcal{B}(X)$ is not invertible, but the resolvent $R(\lambda; T)$ exists in some punctured neighborhood of 0; in this case 0 is an isolated singularity of the holomorphic function $R(\lambda; T)$, and an isolated point of the set $\sigma(T)$. It is advantageous to regard such operator as 'nearly invertible', and to assign T

a 'near inverse'. Such program has been carried out in [6], and is based on a generalized Drazin inverse; a good account of the conventional Drazin inverse is given in Nashed [8, pp. 317–320].

First we need a characterization of the isolated spectral points of an operator $T \in \mathcal{B}(X)$. We observe that $\mathbb{C} = \rho(T) \cup \operatorname{iso}\sigma(T) \cup \operatorname{acc}\sigma(T)$. With T we associate the sets

$$\mathcal{H}(T) = \{x \in X : \lim_{n\to\infty} \|T^n x\|^{1/n} = 0\}, \tag{2.1}$$

$$\mathcal{K}(T) = \{x \in X : (\exists x_n \in X)\ Tx_n = x_{n-1},\ x_0 = x,\ \sup_n \|x_n\|^{1/n} < \infty\}. \tag{2.2}$$

$\mathcal{H}(T)$ and $\mathcal{K}(T)$ are subspaces of X, not necessarily closed. The subspaces $\mathcal{H}(T)$ and $\mathcal{K}(T)$ are hyperinvariant under T, and satisfy the inclusions

$$\begin{aligned} &\mathcal{N}(T^n) \subset \mathcal{N}(T^{n+1}) \subset \mathcal{N}(T^\infty) \subset \mathcal{H}(T), \\ &\mathcal{R}(T^n) \supset \mathcal{R}(T^{n+1}) \supset \mathcal{R}(T^\infty) \supset \mathcal{K}(T), \quad n = 0, 1, 2, \dots, \end{aligned}$$

where $\mathcal{N}(T^\infty) = \bigcup_n \mathcal{N}(T^n)$ is the hyperkernel of T and $\mathcal{R}(T^\infty) = \bigcap_n \mathcal{R}(T^n)$ the hyperrange of T.

Proposition 2.1 (Mbekhta [7, Théorème 1.6]) *Let $T \in \mathcal{B}(X)$. Then the following are equivalent.*

(i) *$0 \notin \operatorname{acc}\sigma(T)$.*

(ii) *X is a topological direct sum $X = \mathcal{H}(T) \oplus \mathcal{K}(T)$.*

(iii) *If T is neither invertible nor quasinilpotent, $T = T_1 \oplus T_2$ (with respect to a topological direct sum $X = X_1 \oplus X_2$), where T_1 is quasinilpotent and T_2 is invertible.*

If (iii) *holds, then $X_1 = \mathcal{H}(T)$ and $X_2 = \mathcal{K}(T)$.*

If was proved recently by the second author and Pak Wai Poon that the preceding proposition is true if the direct sum in (ii) is algebraic with one of the spaces closed. We mention that 0 is a pole of $R(\lambda; T)$ if and only if X is a topological direct sum $X = \mathcal{H}(T) \oplus \mathcal{K}(T)$ with $\mathcal{H}(T) = \mathcal{N}(T^\nu)$ and $\mathcal{K}(T) = \mathcal{R}(T^\nu)$ for some ν, and $\mathcal{H}(T) \neq 0$.

In Section 5 we need to consider the spaces $\mathcal{K}_s(T)$ defined by

$$\mathcal{K}_s(T) = T^{-s}\mathcal{K}(T) = \{x \in X : T^s x \in \mathcal{K}(T)\}, \quad s = 0, 1, 2, \dots \tag{2.3}$$

Then $(\mathcal{K}_s(T))_0^\infty$ is an increasing sequence of subspaces of X. We have the following characterization.

Lemma 2.2 *Let $T \in \mathcal{B}(X)$ and $0 \in \operatorname{iso}\sigma(T)$. Then:*

(i) *For each integer $s \geq 0$,*

$$\mathcal{K}_s(T) = \mathcal{N}(T^s) + \mathcal{K}(T).$$

(ii) *If 0 is a pole of $R(\lambda; T)$ of order ν, then*

$$\mathcal{K}_s(T) = \begin{cases} \mathcal{N}(T^s) + \mathcal{R}(T^{\nu-s}) & \text{if } 0 \leq s < \nu \\ X & \text{if } s \geq \nu. \end{cases}$$

(iii) *If 0 is an essential singularity of $R(\lambda; T)$, then $\mathcal{K}_s(T) \neq X$ for all s.*

Proof (i) If $s \geq 0$, then $T^s(\mathcal{N}(T^s) + \mathcal{K}(T)) = T^s\mathcal{K}(T) \subset \mathcal{K}(T)$; hence $\mathcal{N}(T^s) + \mathcal{K}(T) \subset T^{-s}\mathcal{K}(T)$. Conversely, if $T^s x \in \mathcal{K}(T)$, write $x = u + v$ with $u \in \mathcal{H}(T)$ and $v \in \mathcal{K}(T)$. Then $T^s u = T^s x - T^s v \in \mathcal{H}(T) \cap \mathcal{K}(T)$. Hence $T^s u = 0$, and $u \in \mathcal{N}(T^s)$. This proves $\mathcal{K}_s(T) \subset \mathcal{N}(T^s) + \mathcal{K}(T)$.

(ii) If 0 is a pole of $R(\lambda; T)$ of order ν, then $\mathcal{H}(T) = \mathcal{N}(T^\nu)$ and $\mathcal{K}(T) = \mathcal{R}(T^\nu)$. Let $0 \leq s < \nu$. We have $\mathcal{K}_s(T) = \mathcal{N}(T^s) + \mathcal{K}(T) \subset \mathcal{N}(T^s) + \mathcal{R}(T^{\nu-s})$. Conversely, let $x \in \mathcal{N}(T^s) + \mathcal{R}(T^{\nu-s})$; then $x = u + T^{\nu-s}w$ with $u \in \mathcal{N}(T^s)$ and $w \in X$. Therefore $T^s x = T^s u + T^s T^{\nu-s} w = T^\nu w \in \mathcal{R}(T^\nu) = \mathcal{K}(T)$, so that $x \in T^{-s}\mathcal{K}(T) = \mathcal{K}_s(T)$. Consequently, $\mathcal{N}(T^s) + \mathcal{R}(T^{\nu-s}) \subset \mathcal{K}_s(T)$.

Let $s \geq \nu$. Then $\mathcal{K}_s(T) = \mathcal{N}(T^s) + \mathcal{K}(T) = \mathcal{H}(T) + \mathcal{K}(T) = X$.

(iii) follows from (i) and the fact that $\mathcal{N}(T^s) \neq \mathcal{H}(T)$ for all s if 0 is an essential singularity of $R(\lambda; T)$.

We give a functional definition of the (generalized) Drazin inverse which seems to be best suited to the methods of this paper.

Definition 2.3 Let $T \in \mathcal{B}(X)$ be such that $0 \notin \operatorname{acc}\sigma(T)$, and let X be decomposed as $X = \mathcal{H}(T) \oplus \mathcal{K}(T)$. The *Drazin inverse* T^D of T is defined by $T^D = T^{-1}$ if $\mathcal{H}(T) = 0$ and by $T^D = 0$ if $\mathcal{K}(T) = 0$; if $\mathcal{H}(T) \neq 0$ and $\mathcal{K}(T) \neq 0$, T^D is defined by

$$T^D = (T_1 \oplus T_2)^D = 0 \oplus T_2^{-1}, \tag{2.4}$$

where $T = T_1 \oplus T_2$ is the decomposition of T from Proposition 2.1. This definition reduces to the conventional Drazin inverse when T_1 is nilpotent of index q. In this case $\lambda = 0$ is a pole of order q of the resolvent $R(\lambda; T)$. (A pole of order 0 is by definition a resolvent point of T.)

If $\mu \notin \operatorname{acc}\sigma(T)$, then $T - \mu I$ is Drazin invertible, and the operator

$$P = P(\mu; T) = I - (T - \mu I)(T - \mu I)^D$$

is called the *spectral projection* of T at μ. We observe that there are three possibilities:

(i) If $P = 0$, then μ is a resolvent point of T and $(T - \mu I)^D = (T - \mu I)^{-1}$.

(ii) If $P = I$, then $\sigma(T) = \{\mu\}$, and $(T - \mu I)^D = 0$.

(iii) If $P \neq 0$ and $P \neq I$, then μ is an isolated singularity for the resolvent $R(\lambda; T)$, and $(T - \mu I)^D \neq 0$.

Suppose that μ is in isolated spectral point of T, that B is the Drazin inverse of $T - \mu I$ and that P is the spectral projection of T at μ. Then there is $r > 0$ such that the following Laurent expansion converges whenever $0 < |\lambda - \mu| < r$ (see [6]):

$$R(\lambda; T) = \sum_{n=1}^{\infty} (\lambda - \mu)^{-n} (T - \mu I)^{n-1} P - \sum_{n=0}^{\infty} (\lambda - \mu)^n B^{n+1}. \tag{2.5}$$

For completeness we give an algebraic definition of the (generalized) Drazin inverse which is close to the original Drazin definition. The equivalence of the two definitions, expressed in the following proposition, follows from [6, Theorem 7.1]. First we recall that $A \in \mathcal{B}(X)$ is quasinilpotent if $I - AU$ is invertible for every $U \in \mathcal{B}(X)$ commuting with A. Nilpotent operators are quasinilpotent.

Proposition 2.4 *An operator $S \in \mathcal{B}(X)$ is the Drazin inverse of $T \in \mathcal{B}(X)$ if and only if the following conditions are satisfied.*

$$TS = ST, \quad TS^2 = S, \quad T - T^2 S \text{ is quasinilpotent.} \tag{2.6}$$

The definition of the Drazin inverse given by (2.6) reduces to the conventional Drazin inverse if and only if $T - T^2 S = T(I - TS)$ is nilpotent. In this case $(T(I - TS))^q = 0$ for some nonnegative integer q. Observing that $I - TS$ is a projection, we get $T^q(I - TS) = 0$, that is, $T^{q+1} S = T^q$. If q is the smallest such integer, 0 is a pole of the resolvent $R(\lambda; T)$ of order q.

3 SEMI ITERATIVE METHODS

Given an operator $T \in \mathcal{B}(X)$ and an element $c \in X$, we first define the *iteration*

$$x_m = Tx_{m-1} + c \qquad (m \geq 1), \tag{3.1}$$

where $x_0 \in X$ is an initial approximation. We observe that

$$x_m = T^m x_0 + \sum_{j=0}^{m-1} T^j c \qquad (m \geq 1).$$

A *semi-iterative method* (SIM) with respect to the iteration (3.1) is given by

$$y_0 = x_0, \quad y_m = \sum_{i=0}^{m} \pi_{m,i} x_i \qquad (m \geq 1), \tag{3.2}$$

where $\mathcal{P} = [\pi_{m,i}]$ $(m, i \geq 0)$ is an infinite lower triangular matrix with complex entries satisfying

$$\sum_{i=0}^{m} \pi_{m,i} = 1 \text{ for all } m.$$

Let us remark that the condition $\sum_{i=0}^{m} \pi_{m,i} = 1$ together with $0 \leq \pi_{m,i} \leq 1$ ensures that the semi-iterates y_m belong to the convex hull of the iterates x_m; however, methods dependent on convexity are not considered in the sequel.

Following [1, 2, 3] we define polynomials p_m and q_{m-1} by

$$p_m(\lambda) = \sum_{i=0}^{m} \pi_{m,i} \lambda^i = (\lambda - 1) q_{m-1}(\lambda) + 1 \quad (m \geq 1); \tag{3.3}$$

we note that $p_m(1) = 1$. The semi-iterates y_m defined by (3.2) can be expressed as

$$y_m = p_m(T) y_0 + q_{m-1}(T) c \qquad (m \geq 1); \tag{3.4}$$

this follows from [3, Lemma 1].

Formula (3.4), while convenient for theoretical considerations, is unsuitable for the computational determination of the semi-iterates y_m; the construction described by (3.2), where x_m are determined recursively by (3.1) makes more efficient use of the computer memory and is less succeptible to the round-off error.

When the convergence of SIMs is investigated for operator equations, it is convenient to consider several distinct cases. The case when 1 is the resolvent point of T leads to the unique solution of $x = Tx + c$ for each $x \in X$; this has been extensively analyzed in the literature. We look first at the case when $1 \in \operatorname{iso}\sigma(T)$. It is then possible to obtain new results for the convergence of SIMs as analogues of [2, Theorems 1 and 2] and [1, Theorems 3.3 and 3.8] by separating the cases $c \in \mathcal{K}(I-T)$ and $c \notin \mathcal{K}(I-T)$. We then consider the case when $1 \in \operatorname{acc}\sigma(T)$; this situation does not arise for matrix equations as eigenvalues of matrices are isolated in $\sigma(T)$. In [1], only the case that 1 is a pole of $R(\lambda; T)$ was considered.

4 THE CASE $1 \in \operatorname{iso}\sigma(T),\ c \in \mathcal{K}(I-T)$

In this section we analyze the convergence of SIMs for an operator $T \in \mathcal{B}(X)$ for which $\lambda = 1$ is an isolated point of the spectrum $\sigma(T)$. The results previously derived in [1] required the peripheral spectrum of T to consist of poles of the resolvent, and required $\lambda = 1$ to be at most a pole of $R(\lambda; T)$.

If $1 \in \operatorname{iso}\sigma(T)$, we assume throughout this section that the operator T has a decomposition $T = T_1 \oplus T_2$ with respect to the direct sum $X = \mathcal{H}(I-T) \oplus \mathcal{K}(I-T)$, where $I_1 - T_1$ is quasinilpotent and $I_2 - T_2$ is invertible. By P we will denote the spectral projection $P = I - (I-T)(I-T)^D$ of T at $\lambda = 1$. We note that P is the projection with the range $\mathcal{H}(I-T)$, and $I - P$ the projection with the range $\mathcal{K}(I-T)$.

The space $\mathcal{H}(I-T)$ contains the union of all nullspaces $\mathcal{N}((I-T)^n)$, and in the case that 1 is a pole of $R(\lambda; T)$, $\mathcal{H}(I-T)$ coincides with one of the nullspaces, $\mathcal{N}((I-T)^\nu)$.

The way to view the space $\mathcal{K}(I-T)$ is to observe that it is contained in the intersection of all ranges $\mathcal{R}((I-T)^n)$, and that for every $x \in \mathcal{K}(I-T)$ there is a recursive string of solutions to $(I-T)x_n = x_{n-1}$, starting with $x_0 = x$; all the x_n are contained in $\mathcal{K}(I-T)$. Also, from the definition it follows that there is a constant $c > 0$ such that $\|x_n\| \le c^n \|x\|$. Another important feature of $\mathcal{K}(I-T)$ is that $(I-T)\mathcal{K}(I-T) = \mathcal{K}(I-T)$. If 1 is a pole of the resolvent of T, $\mathcal{K}(I-T)$ reduces to one of the ranges $\mathcal{R}((I-T)^\nu)$.

If $c \in \mathcal{K}(I-T)$, we set $\widetilde{x} = (I_2 - T_2)^{-1}c = (I-T)^D c$; then $\widetilde{x}$ is a solution to $x = Tx + c$ which lies in the space $\mathcal{K}(I-T)$.

The following result generalizes [1, Theorem 3.3] and [2, Theorem 1]. Example 4.5 below will show that our result is strictly stronger than the quoted theorems.

Theorem 4.1 *Let $T \in \mathcal{B}(X)$ be such that $1 \in \operatorname{iso}\sigma(T)$, and let $c \in \mathcal{K}(I-T)$. Then the following statements are equivalent:*

(i) *The* SIM *converges for any initial approximation* $y_0 \in \mathcal{N}(I-T) + \mathcal{K}(I-T)$ *to a solution of* $x = Tx + c$.

(ii) $\lim_{m\to\infty} p_m(T)v = 0$ *for every* $v \in \mathcal{K}(I-T)$.

(iii) $\lim_{m\to\infty} q_{m-1}(T)v = (I-T)^D v$ *for every* $v \in \mathcal{K}(I-T)$.

If one of the above conditions holds, then

$$\lim_{m\to\infty} y_m = Py_0 + (I_2 - T_2)^{-1}c = (I - (I-T)(I-T)^D)y_0 + (I-T)^D c. \tag{4.1}$$

Proof (ii) $\iff$ (iii). In view of the identity

$$q_{m-1}(T)v = (I_2 - T_2)^{-1}(I - p_m(T))v$$

valid for every $v \in \mathcal{K}(I-T)$, and in view of Definition 2.3, (ii) and (iii) are equivalent.

(i) $\implies$ (ii). Let $v \in \mathcal{K}(I-T)$. For the SIM (y_m), choose an initial approximation $y_0 = \widetilde{x} + v$. Then $y_0 \in \mathcal{K}(I-T) \subset \mathcal{N}(I-T) + \mathcal{K}(I-T)$, and

$$\begin{aligned} y_m &= p_m(T)y_0 + q_{m-1}(T)c = p_m(T_2)(\widetilde{x} + v) + (I_2 - T_2)^{-1}(I - p_m(T_2))c \\ &= p_m(T_2)v + \widetilde{x}. \end{aligned}$$

Hence

$$p_m(T)v = p_m(T_2)v = y_m - \widetilde{x} = (I_2 - T_2)^{-1}((I_2 - T_2)y_m - c) \to 0$$

as $(I_2 - T_2)y_m \to c$ by hypothesis.

(ii) $\implies$ (i) Let $y_0 = u + v$ with $u \in \mathcal{N}(I-T)$ and $v \in \mathcal{K}(I-T)$; note that $Py_0 = u$. We have $T^j u = u$ for all $j \geq 1$, and $p_m(T)u = u$. Then

$$\begin{aligned} y_m &= p_m(T)y_0 + q_{m-1}(T)c = p_m(T)u + p_m(T)v + (I_2 - T_2)^{-1}(I - p_m(T))c \\ &= u + \widetilde{x} + p_m(T)(v - \widetilde{x}) \to u + \widetilde{x} = Py_0 + (I-T)^D c, \end{aligned}$$

as $v - \widetilde{x} \in \mathcal{K}(I-T)$. The element $u + \widetilde{x}$ is a solution to $x = Tx + c$.

A special case of the semi-iterative method (3.2) is the iteration (3.1) which is obtained for the choice of $\mathcal{P}$ as the infinite unit matrix. In this case we obtain the equivalence of the following conditions.

Corollary 4.2 *Let $1 \in \operatorname{iso}\sigma(T)$ and let $c \in \mathcal{K}(I-T)$. The following statements are equivalent.*

(i) *The iterative method* (3.1) *converges for all $x_0 \in \mathcal{N}(I-T) + \mathcal{K}(I-T)$.*

(ii) *$T^n v \to 0$ for all $v \in \mathcal{K}(I-T)$.*

(iii) *$\sum_{n=0}^{\infty} T^n v = (I-T)^D v$ for all $v \in \mathcal{K}(I-T)$.*

If any of the preceding conditions holds, then $r(T) \le 1$.

Proof The only fact needed to be proved is the statement about the spectral radius. The operator T is quasinilpotent or $T = T_1 \oplus T_2$ where $I_1 - T_1$ is quasinilpotent and $I_2 - T_2$ invertible by Proposition 2.1. We have $\sigma(T_1) = \{1\}$. The sequence (T_2^m) is pointwise bounded, and, by the principle of uniform boundedness, uniformly bounded, say $\|T_2^n\| \le M$. We observe that

$$r(T_2) = \lim_{m\to\infty} \|T_2^m\|^{1/m} \le \lim_{m\to\infty} M^{1/m} = 1.$$

Hence $r(T) = \max\{r(T_1), r(T_2)\} = 1$.

Example 4.3 If the peripheral spectrum of T consists only of poles of $R(\lambda; T)$, then conditions (i)–(iii) of the preceding corollary hold if and only if $r(T) \le 1$ and $\sigma(T)\backslash\{1\}$ does not meet the unit circle. This result is due to Castro [1, Corollary 3.4].

Example 4.4 Let $T \in \mathcal{B}(H)$ be a normal operator in a Hilbert space H such that $R(\lambda; T)$ has a simple pole at 1. Then $\mathcal{H}(I-T) = \mathcal{N}(I-T)$ and $\mathcal{K}(I-T) = \mathcal{R}(I-T)$. In this case the Drazin inverse $(I-T)^D$ coincides with the Moore–Penrose inverse $(I-T)^\dagger$. We can improve on [1, Corollary 3.6] by applying Theorem 4.1 together with [1, Remark 2.2]: If $c \in \mathcal{R}(I-T)$, then the SIM (y_m) defined by (3.2) converges to the minimal norm solution of $x = Tx + c$ if and only if $y_0 \in \mathcal{R}(I-T)$.

In the next example we construct an operator T to which we can apply the results of this paper but which is not covered by the theory developed in [1] or [2]. The operator T, defined on the bilateral ℓ^1-space, has an essential singularity at 1. We identify the spaces $\mathcal{H}(I-T)$, $\mathcal{K}(I-T)$ and $\mathcal{N}(I-T) + \mathcal{K}(I-T)$, and apply the Cesàro summability SIM to T (with a special choice of the coefficients β_n defined below).

Example 4.5 Let $X = \ell^1$ be the set of all bilateral sequences $x = (\xi_n)_{-\infty}^{\infty}$ of complex numbers with $\|x\| = \sum_{n=-\infty}^{\infty} |\xi_n| < \infty$; $\|x\|$ is the norm of $x \in X$. Let $(\beta_n)_1^{\infty}$ be a bounded sequence of complex numbers which does not have 0 in its closure. Let S be the operator defined by

$$S(\ldots, \xi_{-2}, \xi_{-1}, \widehat{\xi_0}, \xi_1, \xi_2, \xi_3, \ldots,) = (\ldots, \beta_2\xi_{-2}, \beta_1\xi_{-1}, \widehat{0}, 0, 0, \xi_1, 0, \tfrac{1}{3}\xi_3, 0, \tfrac{1}{5}\xi_5, \ldots)$$

for any $x = (\xi_n) \in X$. In terms of the standard Schauder basis $(e_n)_{-\infty}^{\infty}$ for X,

$$Se_{2k} = 0, \qquad Se_{2k+1} = \frac{1}{2k+1} e_{2k+3} \qquad \text{if } k = 0, 1, 2, \ldots,$$
$$Se_{-n} = \beta_n e_{-n} \qquad \text{if } n = 1, 2, \ldots$$

Then S maps X into itself, and $S \in \mathcal{B}(X)$. If σ is the closure of the set $\{\beta_n\}$ in $\mathbb{C}$, it follows that $\sigma(S)$ is the disjoint union of closed sets σ and $\{0\}$. Consequently, 0 is an isolated spectral point of S. Let

$$\begin{aligned} X_1 &= \{(\xi_n) \in X : \xi_n = 0 \text{ if } n < 0\}, \\ X_2 &= \{(\xi_n) \in X : \xi_n = 0 \text{ if } n \geq 0\}, \\ X_3 &= \{(\xi_n) \in X : \xi_n = 0 \text{ if } n = 2k+1,\ k \geq 0\}. \end{aligned}$$

We observe that X_1 and X_2 are closed complementary subspaces of X invariant under S. Write S_1, S_2 for the restriction of S to X_1, X_2, respectively. Then $\sigma(S_1) = \{0\}$, but S_1 is not nilpotent. Further, $\sigma(S_2) = \sigma$, and S_2 is invertible on X_2 since $0 \notin \sigma$. By Proposition 2.1, $\mathcal{H}(S) = X_1$, $\mathcal{K}(S) = X_2$. The null space $\mathcal{N}(S)$ is the infinite dimensional space $X_3 \cap X_1$; we observe that $\mathcal{N}(S^n) = \mathcal{N}(S)$ for all n; by Lemma 2.2, 0 is an essential singularity of $R(\lambda; S)$.

Define $T \in \mathcal{B}(X)$ by $T = I + S$. Then 1 is an essential singularity of T, and

$$\mathcal{H}(I - T) = X_1, \quad \mathcal{K}(I - T) = X_2, \quad \mathcal{N}(I - T) + \mathcal{K}(I - T) = X_3.$$

Let $\beta_1 = \beta_2 = \cdots = -1 + i$ in the definition formula of the operator S above. For the resulting operator $T = I + S$ we consider the SIM based on the Cesàro summability with the polynomials $p_m(\lambda) = (m+1)^{-1} \sum_{k=0}^{m} \lambda^k \qquad (m \geq 0)$, so that, for any $c \in X$,

$$p_m(T)y_0 + q_{m-1}(T)c = \frac{1}{m+1} \sum_{k=0}^{m} x_k. \tag{4.2}$$

If $v = (\dots, \xi_{-2}, \xi_{-1}, \widehat{0}, 0, \dots) \in X_2$, then $T^k v = (\dots, i^k\xi_{-2}, i^k\xi_{-1}, \widehat{0}, 0, \dots) = i^k v$, and $p_m(T)v = p_m(i)v$. Then $p_m(T)v \to 0$ for all $v \in X_2$ as $|p_m(i)| \le \sqrt{2}/(m+1)$.

Let $c \in X_1$. If $v \in X_2 \backslash \{0\}$, $T^m v = i^m v \not\to 0$. Hence by Corollary 4.2 the iterative method $x_m = Tx_{m-1} + c$ does not converge for all $x_0 \in X_3$. However, according to Theorem 4.1, the SIM (4.2) converges for any initial approximation $y_0 \in X_3$ to a solution of $x = Tx + c$. From the definition of T we can calculate $(I-T)^D$: If $x = u + v$ with $u \in X_1$ and $v \in X_2$, then

$$(I-T)^D x = (I-T)^D(u+v) = (I-T)^D v = \tfrac{1}{2}(1+i)v.$$

5 THE CASE $1 \in \operatorname{iso}\sigma(T)$, $c \notin \mathcal{K}(I-T)$

As in the preceding section we assume that $T = T_1 \oplus T_2$, where $I_1 - T_1$ is quasinilpotent and $I_2 - T_2$ invertible. By P we denote the spectral projection of T at 1.

If $c \notin \mathcal{K}(I-T)$, we set $\widetilde{x} = (I_2 - T_2)^{-1}(I-P)c = (I-T)^D(I-P)c$; then $\widetilde{x}$ satisfies $(I-T)\widetilde{x} = (I-P)c$ and $\widetilde{x} \in \mathcal{K}(I-T)$. This unique solution of $x = Tx + (I-P)c$ contained in $\mathcal{K}(I-T)$ is called the *Drazin inverse solution* to $x = Tx + c$.

The following result generalizes [1, Theorem 3.8] as well as [2, Theorem 2]. In its proof we will use the characterization of the spaces $\mathcal{K}_s(I-T)$ derived in Lemma 2.2. Example 5.3 below shows that our result is strictly stronger that the corresponding results of [1, 2].

Theorem 5.1 *Let* $1 \in \operatorname{iso}\sigma(T)$, *and let there be an integer* $s \ge 1$ *such that one of the following equivalent conditions holds.*

(i) $\lim_{m\to\infty} p_m(T)x = Px$ *for any* $x \in \mathcal{K}_{s+1}(I-T)$.

(ii) $\lim_{m\to\infty} q_{m-1}(T)x = (I-T)^D x$ *for any* $x \in \mathcal{K}_s(I-T)$.

Then, for each $c \in \mathcal{K}_s(I-T)$, *the* SIM (y_m) *converges for any* $y_0 \in \mathcal{K}_{s+1}(I-T)$ *to the limit*

$$\widetilde{x}(y_0) = Py_0 + (I-T)^D c. \tag{5.1}$$

Proof (i) $\Longrightarrow$ (ii). Under condition (i) we have

$$p_m(T)u = u + \sum_{j=1}^{s} \frac{p_m^{(j)}(1)}{j!}\,(I_1 - T_1)^j u \to u$$

for each $u \in \mathcal{N}((I-T)^{s+1})$; then it follows that $p_m^{(j)}(1) \to 0$ as $m \to \infty$ for $1 \le j \le s$. (Note that, in general, 1 is not a pole of the resolvent of T. However, 1 is a pole of the resolvent of T_{s+1}, the restriction of T to $\mathcal{N}((I-T)^{s+1})$.) Also, $p_m(T)v \to 0$ for each $v \in \mathcal{K}(I-T)$.

Let $x = u + v$ with $u \in \mathcal{N}((I-T)^s)$ and $v \in \mathcal{K}(I-T)$. Then we have

$$q_{m-1}(T)u = \sum_{j=0}^{s-1} \frac{q_{m-1}^{(j)}(1)}{j!}(I_1 - T_1)^j u = \sum_{j=1}^{s} \frac{p_m^{(j)}(1)}{j!}(I_1 - T_1)^{j-1} u \to 0,$$

and

$$q_{m-1}(T)v = (I_2 - T_2)^{-1}(I - p_m(T))v \to (I_2 - T_2)^{-1}v.$$

Then $q_{m-1}(T)x = q_{m-1}(T)u + q_{m-1}(T)v \to (I-T)^D x$.

(ii) $\implies$ (i). Let $x \in \mathcal{K}_{s+1}(I-T)$; then $(I-T)x \in \mathcal{K}_s(I-T)$, and

$$p_m(T)x = (I - q_{m-1}(T)(I-T))x \to (I - (I-T)^D(I-T))x = Px.$$

Suppose that $c \in \mathcal{K}_s(I-T)$. From (i) and (ii),

$$\lim_{m\to\infty} y_m = \lim_{m\to\infty} p_m(T)y_0 + \lim_{m\to\infty} q_{m-1}(T)c = Py_0 + (I-T)^D c$$

for any $y_0 \in \mathcal{K}_{s+1}(I-T)$.

Corollary 5.2 *Under the assumptions of the preceding theorem, the vector $\widetilde{x}(y_0)$ in (5.1) is a solution to the equation $x = Tx + (I-P)c$ for any $y_0 \in \mathcal{N}(I-T) + \mathcal{K}(I-T)$. Moreover, $\widetilde{x}(y_0)$ is the Drazin inverse solution to $x = Tx + c$ if and only if $y_0 \in \mathcal{K}(I-T)$.*

If we assume that 1 is a pole of $R(\lambda; T)$ of order ν, then we can apply Theorem 5.1 combined with Lemma 2.2 to recover [1, Theorem 3.8].

In the following example we discuss an operator T for which the previous theory developed in [1] or [2] does not apply, but which can be treated by methods of this paper. As in Example 4.5, T is defined on the bilateral ℓ^1-space and has an essential singularity at 1. We identify the spaces $\mathcal{H}(I-T)$, $\mathcal{K}(I-T)$, $\mathcal{N}(I-T) + \mathcal{K}(I-T)$, and $\mathcal{K}_s(I-T)$, and apply the modified Richardson's method investigated by Sidi [9] to T.

Example 5.3 Let X be the bilateral space ℓ^1 described in Example 4.5. Let $(\beta_n)_1^\infty$ be a bounded sequence of complex numbers with the closure σ such that $0 \notin \sigma$. Let U be the operator defined by

$$U(\ldots, \xi_{-2}, \xi_{-1}, \widehat{\xi_0}, \xi_1, \xi_2, \xi_3, \ldots,) = (\ldots, \beta_2\xi_{-2}, \beta_1\xi_{-1}, \widehat{\xi_1}, \tfrac{1}{2}\xi_2, \tfrac{1}{3}\xi_3, \ldots)$$

for any $x = (\xi_n) \in X$. Then U maps X into itself, and $U \in \mathcal{B}(X)$. The spectrum $\sigma(U)$ is the disjoint union of closed sets σ and $\{0\}$. Consequently, 0 is an isolated spectral point of U.

If s is a nonnegative integer, define

$$X_1 = \{(\xi_n) \in X : \xi_n = 0 \text{ if } n < 0\},$$
$$X_2 = \{(\xi_n) \in X : \xi_n = 0 \text{ if } n \geq 0\},$$
$$Y_s = \{(\xi_n) \in X : \xi_n = 0 \text{ if } n \geq s\}.$$

Then X_1 and X_2 are closed subspaces of X invariant under U, and $X = X_1 \oplus X_2$. If $U = U_1 \oplus U_2$ is the decomposition of U with respect to this direct sum of subspaces, then U_1 is quasinilpotent (but not nilpotent) and U_2 is invertible on X_2 ($0 \notin \sigma$). By Proposition 2.1, $\mathcal{H}(U) = X_1$, $\mathcal{K}(U) = X_2$. Further, $\mathcal{N}(U^s) = Y_s \cap X_1$ for every nonnegative integer s; Lemma 2.2 then shows that 0 is an essential singularity of U.

Define $T \in \mathcal{B}(X)$ by $T = I + U$. Then 1 is an essential singularity of T, and

$$\mathcal{H}(I-T) = X_1, \quad \mathcal{K}(I-T) = X_2, \quad \mathcal{K}_s(I-T) = Y_s,\ s = 1, 2, \ldots$$

We choose the β_n so that $|1 + \beta_n| \leq \alpha < 1$ for all n; then $r(T_2) < 1$. If $c \in \mathcal{K}(I-T) = X_1$, then the basic iterative method $x_m = Tx_{m-1} + c$ converges for any initial vector $x_0 \in \mathcal{N}(I-T) + \mathcal{K}(I-T) = Y_1$ to a solution of $x = Tx + c$.

If $c \notin \mathcal{K}(I-T)$ but $c \in \mathcal{K}_s(I-T) = Y_s$ for some $s \geq 1$, we consider the modified Richardson's method which was introduced by Sidi [9] as an extrapolation method for the Drazin inverse solution of singular linear systems. This method can be also interpreted as a semi-iterative method with the polynomials

$$p_{m+s}(\lambda) = \lambda^m \sum_{i=0}^{s} (-1)^i \binom{m+i-1}{i} (\lambda - 1)^i;$$

these polynomials satisfy $p_{m+s}(1) = 1$ and $p^{(j)}_{m+s}(1) = 0$, $1 \leq j \leq s$. Then

$$p_{m+s}(T)v \to Pv \text{ for all } v \in \mathcal{K}_{s+1}(I-T) = Y_{s+1},$$

where P is the spectral projection of T corresponding to 1. According to Theorem 5.1, we have

$$\lim_{m\to\infty} y_m = Py_0 + (I-T)^D c$$

for all $y_0 \in Y_{s+1}$.

6 THE CASE $1 \in \operatorname{acc} \sigma(T)$

If we restrict the semi-iterative methods to a certain class of infinite matrices $\mathcal{P}$, we are able to obtain necessary and sufficient conditions for the convergence of SIMs even in the case when 1 is an accumulation point of $\sigma(T)$. This requires the concept of $\mathcal{P}$-regularity, a weaker form of asymptotic $\mathcal{P}$-regularity studied in [4, 5].

Definition 6.1 Let $\mathcal{P}$ be an infinite matrix defining a SIM. We say that $T \in \mathcal{B}(X)$ is *$\mathcal{P}$-regular* if

$$(I-T)p_m(T)x \to 0 \quad \text{whenever} \quad p_m(T)x \text{ converges.}$$

For a $\mathcal{P}$-regular operator T we define

$$E_{\mathcal{P}}(T) = \{x \in X : p_m(T)x \text{ converges}\};$$

we note that $E_{\mathcal{P}}(T)$ is a subspace of X, not necessarily closed.

In the case that the equation (1.1) is consistent, the convergence of the SIM is described by the following theorem, which strenghtens [5, Proposition 1.4]. For any $c \in X$ we define $(I-T)^{-1}c = \{x \in X : x = Tx + c\}$.

Theorem 6.2 *Let $T \in \mathcal{B}(X)$ be a $\mathcal{P}$-regular operator, and let $c \in \mathcal{R}(I-T)$. Then the* SIM (3.2) *with an initial approximation $y_0 \in X$ converges to a solution x of the equation $x = Tx + c$ if and only if*

$$y_0 \in (I-T)^{-1}c + E_{\mathcal{P}}(T).$$

Proof Let $y_0 = u + v$, where $u \in (I-T)^{-1}c$ and $v \in E_{\mathcal{P}}(T)$. Then $(I-T)u = c$, and

$$\begin{aligned} p_m(T)y_0 + q_{m-1}(T)c &= p_m(T)(u+v) + q_{m-1}(T)(I-T)u \\ &= p_m(T)u + (I - p_m(T))u + p_m(T)v \\ &= u + p_m(T)v. \end{aligned}$$

Then

$$(I-T)(u + p_m(T)v) = (I-T)u + (I-T)p_m(T)v \to (I-T)u = c.$$

Conversely, assume that the SIM (3.2) converges to x for some initial approximation $y_0 \in X$. Choose $u \in (I-T)^{-1}c$, and set $v = y_0 - u$. Then

$$\begin{aligned} p_m(T)u + q_{m-1}(T)c &= p_m(T)u + q_{m-1}(T)(I-T)u \\ &= p_m(T)u + (I - p_m(T))u = u, \end{aligned}$$

and

$$\begin{aligned} y_m - u &= p_m(T)y_0 + q_{m-1}(T)c - \big(p_m(T)u + q_{m-1}(T)c\big) \\ &= p_m(T)(y_0 - u) = p_m(T)v, \end{aligned}$$

which shows that $v \in E_{\mathcal{P}}(T)$. Hence $y_0 = u + v \in (I-T)^{-1}c + E_{\mathcal{P}}(T)$.

REFERENCES

[1] N. Castro González, On the convergence of semiiterative methods to the Drazin inverse solution of linear equations in Banach spaces, *Collect. Math.* **46** (1995), 303–314.

[2] M. Eiermann, I. Marek and W. Niethammer, On the solution of singular linear systems of algebraic equations by semiiterative methods, *Numer. Math.* **53** (1988), 265–283.

[3] M. Eiermann and W. Niethammer, On the construction of semiiterative methods, *SIAM J. Numer. Anal.* **20** (1983), 1153–1160.

[4] J. J. Koliha, Ergodic theory and averaging iteration, *Canad. J. Math.* **25** (1973), 14–23.

[5] J. J. Koliha, The solution of linear equations in normed spaces by averaging iteration, *SIAM J. Math. Anal.* **5** (1974), 178–186.

[6] J. J. Koliha, A generalized Drazin inverse, *Glasgow Math. J.* **38** (1996), 367–381.

[7] M. Mbekhta, Généralisation de la décomposition de Kato aux opérateurs paranormaux et spectraux, *Glasgow Math. J.* **29** (1987), 159–175.

[8] M. Z. Nashed, Inner, outer and generalized inverses in Banach and Hilbert spaces, *Numer. Funct. Anal. Optim.* **9** (1987), 261–325.

[9] A. Sidi, Development of iterative techniques and extrapolation methods for Drazin inverse solution of consistent and inconsistent singular linear systems, *Linear Algebra Appl.* **167** (1992), 171–203.