1-1-2021

Commuting Perturbations of Operator Equations

Xue Xu
Harbin University

Jiu Ding
University of Southern Mississippi, jiu.ding@usm.edu

Journal of Applied Analysis and Computation
Volume 11, Number 4, August 2021, 1691–1698
Website:http://www.jaac-online.com
DOI:10.11948/20190382

COMMUTING PERTURBATIONS OF OPERATOR EQUATIONS*

Xue Xu[1] and Jiu Ding[2,†]

Abstract Let X be a Banach space and let $T : X \to X$ be a bounded linear operator with closed range. We study a class of commuting perturbations of the corresponding operator equation, using the concept of the spectral radius of a bounded linear operator. Our results extend the classic perturbation theorem for invertible operators and its generalization for arbitrary operators under the commutability assumption.

Keywords Operator equation, generalized inverse, commuting perturbation, spectral radius, projection, least squares.

MSC(2010) 15A18.

1. Introduction

Let X be a Banach space and let $B(X)$ denote the Banach space of all bounded linear operators $T : X \to X$ with the operator norm $\|T\| = \sup\{\|Tx\| : \|x\| = 1\}$. In this paper we study a class of perturbations for the operator equation $Tx = b$, where $T \in B(X)$ with closed range and b is a given vector in X.

In the literature, for example [2, 3, 8, 9], of the perturbation theory for operator equations and related generalized inverses of bounded linear operators from a Banach space to a Banach space, a common assumption for various perturbation results is that the perturbed operator $T + \delta T$ satisfies the inequality $\|\delta T\| < 1/\|T^+\|$, or more generally either $\|\delta T T^+\| < 1$ or $\|T^+\delta T\| < 1$, so that the classic Banach lemma can be used, where T^+ is a generalized inverse associated with two given projections to be defined in the next section. The fundamental lemma for the perturbation theory of linear operators says that, if $E \in B(X)$ satisfies $\|E\| < 1$, then the bounded linear operator $I - E$ is one-to-one and onto, and the power series $\sum_{k=0}^{\infty} E^k$ converges to the bounded linear operator $(I - E)^{-1}$ absolutely. Moreover, $\|(I - E)^{-1}\| \le 1/(1 - \|E\|)$. Here I denotes the identity operator. When the above inequality is applied to the perturbation analysis of bounded linear operators, E is taken to be either $\delta T T^+$ or $T^+\delta T$ in different situations.

However, the perturbation δT of T may not be small enough in norm so that the perturbation condition $\|\delta T\| < 1/\|T^\dagger\|$ is not satisfied. The purpose of the present paper is to weaken the common assumption on the size of δT but still guarantee

† The corresponding author. Email: jiuding@gmail.com(J. Ding)
[1] Department of Mathematics, Harbin University, Harbin, Heilongjiang, 150001, China
[2] School of Mathematics and Natural Sciences, The University of Southern Mississippi, Hattiesburg, MS 39406, USA
* The research of Xue Xu was partially supported by NSFH(No. LH2020A002) and HUDF(No. 2019101).

the applicability of Banach's lemma. The special perturbation that we consider is the one that computes with the unperturbed invertible operator or a generalized inverse of the unperturbed general operator. That is, the perturbation δT of the original operator T satisfies the commutability condition $\delta TT = T\delta T$ when T is one-to-one and onto or $\delta TT^+ = T^+\delta T$ in general. This kind of perturbations is called commuting perturbations. With commuting perturbations, we are able to use the concept of spectral radius to achieve our goal. We shall weaken the assumption on the size of the perturbation in the classic perturbation theorem [7] in operator theory and the generalizations [4, 6] of the classic perturbation result. The classic result states that, if T is one-to-one and onto, and if $\|\delta T\| < 1/\|T^{-1}\|$, then the solution x^* of the operator equation $Tx = b$ and the solution y^* of the perturbed operator equation $(T + \delta T)y = b + \delta b$ satisfy the inequality

$$\frac{\|y^* - x^*\|}{\|x^*\|} \le \frac{\|T\|\|T^{-1}\|}{1 - \|T^{-1}\|\|\delta T\|}\left(\frac{\|\delta b\|}{\|b\|} + \frac{\|\delta T\|}{\|T\|}\right), \tag{1.1}$$

which has been generalized to non-invertible operators in [4, 6].

Our main results, which improve the above classic inequality, will be presented in Section 3 after proving a preliminary result in the next section. We summarize the results and give a concluding remark in Section 4.

2. Spectral Radius

For $T \in B(X)$ let $N(T)$ and $R(T)$ be the null space, which is closed since T is continuous, and the range of T, respectively. We assume that $R(T)$ is closed, and both $N(T)$ and $R(T)$ are complemented by closed subspaces $N(T)^c$ and $R(T)^c$ respectively, so that $X = N(T) \oplus N(T)^c = R(T) \oplus R(T)^c$. Let P and Q be the projections from X onto $N(T)$ along $N(T)^c$ and onto $R(T)$ along $R(T)^c$, respectively.

The operator T is one-to-one and onto from $N(T)^c$ to $R(T)$. The *generalized inverse* $T^+ \in B(X)$ of T with respect to P and Q is defined by letting $T^+y = x$ for all $y \in R(T)$, where x is the unique element of $N(T)^c$ such that $Tx = y$, letting $T^+y = 0$ for all $y \in R(T)^c$, and letting T^+ be extended to the whole space X by linearity. The generalized inverse T^+ is uniquely determined by the equalities

$$TT^+T = T,\ T^+TT^+ = T^+,\ T^+T = I - P,\ TT^+ = Q. \tag{2.1}$$

When X is a Hilbert space and one chooses the projections P and Q to be *orthogonal*, i.e., $N(T)^c = N(T)^\perp$ and $R(T)^c = R(T)^\perp$, where $M^\perp$ denotes the orthogonal complement of a subspace M of X, the corresponding generalized inverse is called the *Moore-Penrose generalized inverse* and is usually denoted by $T^\dagger$. In this special case (2.1) becomes

$$TT^\dagger T = T,\ T^\dagger TT^\dagger = T^\dagger,\ (T^\dagger T)^* = T^\dagger T,\ (TT^\dagger)^* = TT^\dagger,$$

where S^* is the adjoint of operator S. See [1] for more details on generalized inverses.

The perturbation theory of generalized inverses of bounded linear operators has been fruitful in the past four decades, starting with the pioneering work [8] by Nashed. All the perturbation theorems so far in the literature, however, have the assumption that $\|\delta TT^+\| < 1$ or $\|T^+\delta T\| < 1$, which is implied by the stronger inequality $\|\delta T\| < 1/\|T^+\|$ [2, 3, 5, 8]. But such assumptions are sometimes too strong and so not necessarily satisfied in many applications. In fact, the concept of the norm may not be the best tool for measuring the size of perturbations. It may not provide the intrinsic feature of an operator, which is related to the invertibility of the operator. As it turns out, the notion of the spectral radius of a linear operator plays an important role in the Banach lemma, which can be seen from Lemma 2.1 below. Let $\sigma(T)$ be the spectrum of T, that is, the collection of all complex numbers λ such that $T - \lambda I : X \to X$ is not one-to-one or onto.

Definition 2.1. The number

$$r(T) = \sup\{|\lambda| : \lambda \in \sigma(T)\}$$

is called the *spectral radius* of T.

Spectral radius may provide a better controlling number than norm in the assumption of perturbation analysis for operator equations. For example, let

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad \delta A = \begin{bmatrix} \epsilon & \delta \\ 0 & \epsilon \end{bmatrix}; \quad 0 < \epsilon \leq \delta. \tag{2.2}$$

Then the matrix 1-norm $\|\delta A A^{-1}\|_1 = \delta$. On the other hand, the spectral radius $r(\delta A A^{-1}) = \epsilon$. Note that $A\delta A = \delta A A$. From Theorem 3.1 in the next section, we can obtain a perturbation bound for any δ as long as ϵ is small enough.

It is well known [7] that $r(T) \leq \|T\|$ for any operator norm $\| \; \|$, and

$$r(T) = \lim_{n\to\infty} \|T^n\|^{1/n} = \inf_{n\geq 1}\{\|T^n\|^{1/n}\}. \tag{2.3}$$

Lemma 2.1. *Let E be a bounded linear operator on a Banach space X such that $r(E) < 1$. Then $I - E$ is one-to-one and onto, so $(I - E)^{-1}$ exists as a bounded linear operator on X. Moreover,*

$$(I - E)^{-1} = \sum_{k=0}^{\infty} E^k,$$

where the convergence of the Neumann series is absolute.

Lemma 2.2. *Let T and S be bounded linear operators on a Banach space X such that $TS = ST$. Then*

$$r(TS) = r(ST) \leq r(T)r(S).$$

Proof. Since T and S commute, $(TS)^n = T^n S^n$ for all n. So by (2.3),

$$\begin{aligned} r(TS) &= \lim_{n\to\infty} \|(TS)^n\|^{1/n} = \lim_{n\to\infty} \|T^n S^n\|^{1/n} \\ &\leq \lim_{n\to\infty} \|T^n\|^{1/n} \cdot \lim_{n\to\infty} \|S^n\|^{1/n} \\ &= r(T)r(S). \end{aligned}$$

□

3. Commuting Perturbations of Operator Equations

Let $T \in B(X)$ and $b \in X$. Suppose that $\delta T \in B(X)$ and $\delta b \in X$. We shall give perturbation results for the operator equation

$$Tx = b \tag{3.1}$$

when it is perturbed to

$$(T + \delta T)y = b + \delta b, \tag{3.2}$$

under some additional assumptions on the perturbation.

We first consider the special case that T is one-to-one and onto. Let $\kappa = \|T\|\|T^{-1}\|$ be the condition number of T.

Theorem 3.1. *Let T be one-to-one and onto. If $T\delta T = \delta T T$ and $r(\delta T) < 1/r(T^{-1})$, then $T + \delta T$ is one-to-one and onto. Furthermore, the solution y^* of (3.2) and the solution x^* of (3.1) satisfy the inequality*

$$\frac{\|y^* - x^*\|}{\|x^*\|} \le \kappa \|(I + T^{-1}\delta T)^{-1}\| \left(\frac{\|\delta T\|}{\|T\|} + \frac{\|\delta b\|}{\|b\|} \right).$$

Proof. Since $T\delta T = \delta T T$, we have $\delta T T^{-1} = T^{-1}\delta T$. Then by Lemma 2.2 and the second assumption of the theorem, $r(\delta T T^{-1}) \le r(\delta T) r(T^{-1}) < 1$, so the operator $I + \delta T T^{-1}$ is one-to-one and onto from Lemma 2.1. Thus $T + \delta T = T(I + T^{-1}\delta T)$ is one-to-one and onto.

Subtracting $Tx^* = b$ from $(T + \delta T)y^* = b + \delta b$ gives

$$(T + \delta T)(y^* - x^*) = \delta b - \delta T x^*,$$

from which

$$y^* - x^* = (T + \delta T)^{-1}(\delta b - \delta T x^*) = (I + T^{-1}\delta T)^{-1} T^{-1} (\delta b - \delta T x^*).$$

It follows that

$$\begin{aligned} \frac{\|y^* - x^*\|}{\|x^*\|} &\le \|(I + T^{-1}\delta T)^{-1}\| \|T^{-1}\| \frac{\|\delta b - \delta T x^*\|}{\|x^*\|} \\ &\le \|T\|\|T^{-1}\| \|(I + T^{-1}\delta T)^{-1}\| \frac{\|\delta b\| + \|\delta T x^*\|}{\|T\|\|x^*\|} \\ &\le \kappa \|(I + T^{-1}\delta T)^{-1}\| \left(\frac{\|\delta b\|}{\|b\|} + \frac{\|\delta T\|}{\|T\|} \right). \end{aligned}$$

□

Remark 3.1. Since $r(T^{-1}) \le \|T^{-1}\|$ and $r(\delta T) \le \|\delta T\|$, if $\|\delta T\| < 1/\|T^{-1}\|$, then

$$r(\delta T) \le \|\delta T\| < \frac{1}{\|T^{-1}\|} \le \frac{1}{r(T^{-1})}.$$

Hence, Theorem 3.1 generalizes the classic perturbation result (1.1) when the perturbation is commuting.

Example 3.1. We take the matrix and its perturbation in (2.2) as an example. Let $b = (1,0)^T$ and $\delta b = (\epsilon, \epsilon)^T$. Then $x^* = (1,0)^T$ and $y^* = (1-\epsilon(1+\delta)/(1+\epsilon)^2, \epsilon/(1+\epsilon))^T$. The classic perturbation result works only when $\|\delta A\|_1 < 1/\|A^{-1}\|_1$, which means $\epsilon \le \delta < 1/4$. But since $A\delta A = \delta A A$, we can apply Theorem 3.1 , which only requires $r(\delta A) = \epsilon < 1/r(A^{-1}) = 1$ without any restriction to δ.

We can extend Theorem 3.1 to the general case that $T \in B(X)$ with closed range. The number $\kappa = \|T\|\|T^+\|$ is still called the condition number of T. The next theorem deals with the particular case that both (3.1) and (3.2) are consistent.

Theorem 3.2. *Let $T \in B(X)$ with closed range and let T^+ be a generalized inverse of T. Suppose that $T^+\delta T = \delta T T^+$ and $r(\delta T) < 1/r(T^+)$. If $b \in R(T)$ and $b + \delta b \in R(T + \delta T)$, then for any solution y of (3.2), there is a solution x of (3.1) such that*

$$\frac{\|y - x\|}{\|x\|} \le \kappa \|(I + T^+\delta T)^{-1}\| \left(\frac{\|\delta T\|}{\|T\|} + \frac{\|\delta b\|}{\|b\|} \right).$$

Proof. The assumption implies that $I + T^+\delta T$ is one-to-one and onto from Lemmas 2.1 and 2.2. Let $x = T^+b + (I - T^+T)y$. Then $y - x = T^+(Ty - b) \in N(T)^c$, so $T^+T(y - x) = y - x$. Subtracting $Tx = b$ from $(T + \delta T)y = b + \delta b$, we have $(T+\delta T)(y-x) = \delta b-\delta Tx$. Consequently, $(I+T^+\delta T)(y-x) = T^+(T+\delta T)(y-x) = T^+(\delta b - \delta Tx)$, from which

$$y - x = (I + T^+\delta T)^{-1}T^+(\delta b - \delta Tx).$$

The remaining proof is basically the same as the last part of that for Theorem 3.1. □

Remark 3.2. In fact, under the assumption of Theorem 3.2, $(T + \delta T)^+$ exists with $N((T + \delta T)^+) = N(T^+)$ and $R((T + \delta T)^+) = R(T^+)$. Moreover,

$$(T + \delta T)^+ = T^+(I + \delta T T^+)^{-1} = (I + T^+\delta T)^{-1}T^+.$$

Remark 3.3. If $\|\delta T\| < 1/\|T^+\|$, then

$$r(\delta T) \le \|\delta T\| < \frac{1}{\|T^+\|} \le \frac{1}{r(T^+)}.$$

Thus Theorem 3.2 generalizes the main result of [6] if the perturbation satisfies the commutability condition $T^+\delta T = \delta T T^+$.

Remark 3.4. The point x is actually the projection of y onto the solution set of (3.1) with respect to the decomposition $X = N(T) \oplus N(T)^c$. Therefore, in the case of Moore-Penrose generalized inverses, $\|y - x\|$ is the minimal distance of y to the solution set of (3.1), so the upper bound is the optimal one.

When the equation (3.1) is not consistent, any vector $T^+b + z$ with $z \in N(T)$ is called a projection solution of (3.1). With the help of the concept of residual for projection solutions, one can drop the consistency assumption for the original equation (3.1) and its perturbation (3.2), as the following theorem shows.

Theorem 3.3. *Let $T \in B(X)$ with closed range and let T^+ be a generalized inverse of T. Suppose that $T^+\delta T = \delta T T^+$ and $r(\delta T) < 1/r(T^+)$. Then for any projection solution y of (3.2), there is a projection solution x of (3.1) such that*

$$\frac{\|y - x\|}{\|x\|} \le \kappa \|(I + T^+\delta T)^{-1}\| \left(\frac{\|\hat{r}\|}{\|TT^+b\|} + \frac{\|\delta b\|}{\|TT^+b\|} + \frac{\|\delta T\|}{\|T\|} \right),$$

where $\hat{r} = (T + \delta T)y - (b + \delta b)$ *is the residual of* y.

Proof. First $(I + T^+\delta T)^{-1} \in B(X)$ exists. Let $x = T^+b + (I - T^+T)y$. Then

$$\begin{aligned} y - x &= T^+(Ty - b) = T^+(\hat{r} + \delta b - \delta T y) \\ &= T^+[\hat{r} + \delta b - \delta T(y - x) - \delta T x], \end{aligned}$$

from which $(I + T^+\delta T)(y - x) = T^+(\hat{r} + \delta b - \delta T x)$. Hence,

$$y - x = (I + T^+\delta T)^{-1} T^+(\hat{r} + \delta b - \delta T x). \tag{3.3}$$

It follows that

$$\begin{aligned} \frac{\|y - x\|}{\|x\|} &\le \|(I + T^+\delta T)^{-1}\| \|T\| \|T^+\| \left(\frac{\|\hat{r}\|}{\|T\|\|x\|} + \frac{\|\delta b\|}{\|T\|\|x\|} + \frac{\|\delta T\|}{\|T\|} \right) \\ &\le \kappa \|(I + T^+\delta T)^{-1}\| \left(\frac{\|\hat{r}\|}{\|Tx\|} + \frac{\|\delta b\|}{\|Tx\|} + \frac{\|\delta T\|}{\|T\|} \right) \\ &= \kappa \|(I + T^+\delta T)^{-1}\| \left(\frac{\|\hat{r}\|}{\|TT^+b\|} + \frac{\|\delta b\|}{\|TT^+b\|} + \frac{\|\delta T\|}{\|T\|} \right). \end{aligned}$$

The last equality is from the fact that $Tx = TT^+b$. □

In the case that X is a Hilbert space, the above perturbation bound can be further analyzed, using the decomposition technique for the proof of Theorem 3.1 in [4]. Projection solutions x in the Hilbert space are least squares solutions since they solve the minimization problem

$$\|Tx - b\| = \min\{\|Tz - b\| : z \in X\}.$$

Among all the least squares solutions, the one with the minimal norm is given by $x^* = T^\dagger b$.

Theorem 3.4. *Let* X *be a Hilbert space,* $T \in B(X)$ *with closed range, and* $T^\dagger$ *the Moore-Penrose generalized inverse of* T. *Suppose that* $T^\dagger \delta T = \delta T T^\dagger$ *and* $r(\delta T) < 1/r(T^\dagger)$. *Then for any least squares solution* y *of (3.2), there is a least squares solution* x *of (3.1) such that*

$$\begin{aligned} \frac{\|y - x\|}{\|x\|} \le \kappa \|(I + T^\dagger \delta T)^{-1}\| \Bigg[&\frac{\|(I - TT^\dagger)b\|}{\|TT^\dagger b\|} \|\delta T T^\dagger\| \\ &+ \left(\|\delta T T^\dagger\| + 1\right) \left(\frac{\|\delta b\|}{\|TT^\dagger b\|} + \frac{\|\delta T\|}{\|T\|} \right) \Bigg]. \end{aligned}$$

Proof. Let $x = T^\dagger b + (I - T^\dagger T)y$. Since $(T + \delta T)^\dagger \hat{r} = 0$ for $\hat{r} = (T + \delta T)y - (b + \delta b)$, from (3.3) in the proof of Theorem 3.3,

$$y - x = (I + T^\dagger \delta T)^{-1} \left\{ [T^\dagger - (T + \delta T)^\dagger]\hat{r} + T^\dagger(\delta b - \delta T x) \right\}. \tag{3.4}$$

Using the decomposition (see formula (3.19) in Theorem 3.10 of [8])

$$\begin{aligned} [T^\dagger - (T + \delta T)^\dagger]\hat{r} = T^\dagger \delta T (T + \delta T)^\dagger - T^\dagger (\delta T T^\dagger)^* [I - (T + \delta T)(T + \delta T)^\dagger] \\ + (I - T^\dagger T)[(T + \delta T)^\dagger \delta T]^* (T + \delta T)^\dagger, \end{aligned}$$

we have

$$[T^\dagger - (T + \delta T)^\dagger]\hat{r} = -T^\dagger (\delta T T^\dagger)^* \hat{r}. \tag{3.5}$$

On the other hand, since y is a least squares solution of (3.2),

$$\begin{aligned}\|\hat{r}\| &= \|(T+\delta T)y-(b+\delta b)\| \le \|(T+\delta T)x-(b+\delta b)\| \\ &\le \|Tx-b\|+\|\delta b-\delta Tx\| = \|(I-TT^{\dagger})b\|+\|\delta b-\delta Tx\|. \end{aligned} \tag{3.6}$$

Therefore, denoting $\eta = \|(I+T^{\dagger}\delta T)^{-1}\|$, by (3.4), (3.5), and (3.6), we have

$$\begin{aligned}\frac{\|y-x\|}{\|x\|} &\le \eta\|T^{\dagger}\|\frac{\|\delta TT^{\dagger}\|\|\hat{r}\|+\|\delta b-\delta Tx\|}{\|x\|} \\ &\le \eta\|T^{\dagger}\|\frac{\|\delta TT^{\dagger}\|(\|(I-TT^{\dagger})b\|+\|\delta b-\delta Tx\|)+\|\delta b-\delta Tx\|}{\|x\|} \\ &\le \kappa\eta\left[\frac{\|\delta TT^{\dagger}\|\|(I-TT^{\dagger})b\|}{\|T\|\|x\|}+\frac{(\|\delta TT^{\dagger}\|+1)(\|\delta b\|+\|\delta Tx\|)}{\|T\|\|x\|}\right] \\ &\le \kappa\eta\left[\frac{\|(I-TT^{\dagger})b\|}{\|TT^{\dagger}b\|}\|\delta TT^{\dagger}\|+\left(\|\delta TT^{\dagger}\|+1\right)\left(\frac{\|\delta b\|}{\|TT^{\dagger}b\|}+\frac{\|\delta T\|}{\|T\|}\right)\right].\end{aligned}$$

□

4. Conclusions

In this paper, using the tool of the spectral radius instead of the norm, we have weakened the usual condition for the classic perturbation result of invertible operator equations and the extended perturbation result of general consistent operator equations, when the perturbation of the operator is commuting. We have also extended our results to the most general situation that the involved operator equations are inconsistent. Similar ideas may be applied to studying the perturbation bound when the perturbation satisfies the Nashed condition [8], or equivalently, when it is a stable perturbation [3].

References

[1] S. L. Campbell and C. D. Meyer, *Generalized inverses of linear transformations*, SIAM, 2009.

[2] G. Chen, M. Wei and Y. Xue, *Perturbation analysis of the least squares solution in hilbert spaces*, Linear algebra and its applications, 1996, 244, 69–80.

[3] G. Chen and Y. Xue, *Perturbation analysis for the operator equation $Tx = b$ in banach spaces*, Journal of Mathematical Analysis and Applications, 1997, 212(1), 107–125.

[4] J. Ding, *Perturbation of systems of linear algebraic equations*, Linear and Multilinear Algebra, 2000, 47(2), 119–127.

[5] J. Ding and L. Huang, *Perturbation of generalized inverses of linear operators in hilbert spaces*, Journal of mathematical analysis and applications, 1996, 198(2), 506–515.

[6] J. Ding and L. Huang, *A generalization of a classic theorem in the perturbation theory for linear operators*, Journal of mathematical analysis and applications, 1999, 239(1), 118–123.

[7] T. Kato, *Perturbation theory for linear operators*, 132, Springer Science & Business Media, 2013.

[8] M. Z. Nashed, *Perturbations and approximations for generalized inverses and linear operator equations*, in *Generalized inverses and applications*, Elsevier, 1976, 325–396.

[9] Y. Wang, *Theory and applications of generalized inverses of linear operators in banach spaces*, Journal of Mathematical Analysis and Applications, 2005, 83(2), 582–610.