



**Michigan  
Technological  
University**

Michigan Technological University  
**Digital Commons @ Michigan Tech**

---

Michigan Tech Publications

---

8-1-2020

## Article the singular value expansion for arbitrary bounded linear operators

Daniel K. Crane

*Michigan Technological University, dkcrane@mtu.edu*

Mark S. Gockenbach

*Michigan Technological University*

Follow this and additional works at: <https://digitalcommons.mtu.edu/michigantech-p>



Part of the [Mathematics Commons](#)

---

### Recommended Citation

Crane, D., & Gockenbach, M. (2020). Article the singular value expansion for arbitrary bounded linear operators. *Mathematics*, 8(8). <http://doi.org/10.3390/MATH8081346>

Retrieved from: <https://digitalcommons.mtu.edu/michigantech-p/2664>

Follow this and additional works at: <https://digitalcommons.mtu.edu/michigantech-p>



Part of the [Mathematics Commons](#)

Article

# The Singular Value Expansion for Arbitrary Bounded Linear Operators

Daniel K. Crane and Mark S. Gockenbach <sup>\*,†</sup>

Department of Mathematical Sciences, Michigan Technological University, Houghton, MI 49931, USA; dkcrane@mtu.edu

\* Correspondence: msgocken@udel.edu

† Current address: Department of Mathematical Sciences, University of Delaware, Newark, DE 19716, USA.

Received: 9 July 2020; Accepted: 7 August 2020; Published: 12 August 2020



**Abstract:** The singular value decomposition (SVD) is a basic tool for analyzing matrices. Regarding a general matrix as defining a linear operator and choosing appropriate orthonormal bases for the domain and co-domain allows the operator to be represented as multiplication by a diagonal matrix. It is well known that the SVD extends naturally to a compact linear operator mapping one Hilbert space to another; the resulting representation is known as the singular value expansion (SVE). It is less well known that a general bounded linear operator defined on Hilbert spaces also has a singular value expansion. This SVE allows a simple analysis of a variety of questions about the operator, such as whether it defines a well-posed linear operator equation and how to regularize the equation when it is not well posed.

**Keywords:** singular value expansion; inverse problems; regularization

## 1. Introduction

One of the most powerful ideas in linear algebra is diagonalization, which renders many problems completely transparent. For example, if  $A \in \mathbb{R}^{n \times n}$  is a symmetric matrix, the spectral theorem implies that there exists an orthogonal matrix  $V \in \mathbb{R}^{n \times n}$  and a diagonal matrix  $D \in \mathbb{R}^{n \times n}$  such that

$$A = VDV^T \quad (1)$$

(see, for instance, ([1], Section 7.1)). To say that  $V$  is orthogonal means that  $V^T V = I$ , the  $n \times n$  identity matrix, which implies that the columns  $v_1, v_2, \dots, v_n$  of  $V$  form an orthonormal basis for  $\mathbb{R}^n$ . If  $D$  has diagonal entries  $\lambda_1, \lambda_2, \dots, \lambda_n$ , then we can also write

$$A = VDV^T = \sum_{i=1}^n \lambda_i v_i v_i^T.$$

If  $A \in \mathbb{R}^{m \times n}$  is a general matrix (not assumed to be symmetric or even square), the singular value decomposition (SVD) allows us to diagonalize the matrix, at the cost of using two different orthonormal bases. There exist orthogonal matrices  $U \in \mathbb{R}^{m \times m}$  and  $V \in \mathbb{R}^{n \times n}$  and a diagonal matrix  $S \in \mathbb{R}^{m \times n}$  such that

$$A = USV^T$$

and  $S = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_n)$ , with  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$  (see ([2], Lectures 4–5) or ([1], Chapter 8)). This decomposition follows immediately from the spectral theorem for symmetric matrices; in fact,  $A^T A = VS^T S V^T$  is the spectral decomposition of the symmetric matrix  $A^T A$ .

Among its other virtues, the SVD reveals the rank of  $A$  and bases for the four fundamental subspaces associated with  $A$ . If  $A$  has rank  $r$ , then exactly  $r$  of the singular values  $\sigma_1, \sigma_2, \dots, \sigma_n$  are positive, and we can write the SVD in the reduced form

$$A = \hat{U}\hat{S}\hat{V}^T = \sum_{i=1}^r \sigma_i u_i v_i^T,$$

where  $u_1, u_2, \dots, u_m \in \mathbb{R}^m$ ,  $v_1, v_2, \dots, v_n \in \mathbb{R}^n$  are the columns of  $U$ ,  $V$ , respectively,  $\hat{U} = [u_1|u_2|\dots|u_r]$ ,  $\hat{V} = [v_1|v_2|\dots|v_r]$ , and  $\hat{S} = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r)$ . The four fundamental subspaces are represented by orthonormal bases as follows:

$$\begin{aligned} \text{col}(A) &= \text{sp}\{u_1, u_2, \dots, u_r\}, \\ \mathcal{N}(A^T) &= \text{col}(A)^\perp = \text{sp}\{u_{r+1}, u_{r+2}, \dots, u_m\}, \\ \mathcal{N}(A) &= \text{sp}\{v_{r+1}, v_{r+2}, \dots, v_n\}, \\ \text{col}(A^T) &= \mathcal{N}(A)^\perp = \text{sp}\{v_1, v_2, \dots, v_r\}. \end{aligned}$$

It is well known that spectral theory is considerably more complicated for linear operators defined on infinite-dimensional spaces. However, for compact operators, the finite-dimensional theory carries through almost unchanged. Throughout the rest of this paper,  $X$  and  $Y$  will denote real Hilbert spaces. If  $T : X \rightarrow Y$  is linear, then  $T$  is called compact if and only if  $\{Tx_n\}$  has a convergent subsequence in  $Y$  for every bounded sequence  $\{x_n\}$  in  $X$  (this is equivalent to  $T$ 's being continuous when the weak topology is imposed on  $X$  and the norm topology on  $Y$ ).

The spectral theorem for self-adjoint compact operators (see ([3], Section 4.3) or ([4], Section 8.2)) says that if  $T : X \rightarrow X$  is compact and self-adjoint, then there exists an orthonormal sequence  $\{\phi_k\}$  in  $X$  and a sequence  $\{\lambda_k\}$  of nonzero real numbers such that

$$T = \sum_{k=1}^{\infty} \lambda_k \phi_k \otimes \phi_k, \tag{2}$$

where the outer product  $\phi_k \otimes \phi_k : X \rightarrow X$  is the bounded linear operator defined by

$$(\phi_k \otimes \phi_k)x = \langle \phi_k, x \rangle_X \phi_k \text{ for all } x \in X.$$

For ease of exposition, we will assume that the sequences  $\{\phi_k\}$  and  $\{\lambda_k\}$  are infinite sequences; in the contrary case,  $T$  is a finite-rank operator and the infinite series becomes a finite sum.

For a general (that is, not necessarily self-adjoint) compact operator  $T : X \rightarrow Y$ , we can derive the singular value expansion (SVE) of  $T$  by applying the spectral theorem for self-adjoint compact operators to  $T^*T$  (see [3], Section 4.4). The result is

$$T = \sum_{k=1}^{\infty} \sigma_k \psi_k \otimes \phi_k,$$

where  $\{\phi_k\}$  is an orthonormal sequence in  $X$ ,  $\{\psi_k\}$  is an orthonormal sequence in  $Y$ , and  $\sigma_1 \geq \sigma_2 \geq \dots$  are positive numbers converging to zero. Moreover,  $\{\phi_k\}$  is a complete orthonormal set for  $\mathcal{N}(T)^\perp$  and  $\{\psi_k\}$  is a complete orthonormal set for  $\overline{\mathcal{R}(T)}$ .

The SVE of a compact operator has many applications, particularly in analyzing linear operator equations of the form  $Tx = y$  (that is, given  $y \in Y$ , find or estimate  $x \in X$  satisfying  $Tx = y$ ). When  $T$  is compact and not of finite rank, this equation is ill-posed in the sense that the solution  $x$  (if it exists) does not depend continuously on the data  $y$ ; in this case, the equation is often referred to as an inverse problem. The singular value expansion of  $T$  is useful in analyzing approaches to

regularizing  $Tx = y$ , that is, to computing a stable (approximate) solution to the equation in the presence of noisy data.

When  $T : X \rightarrow Y$  is not necessarily compact, there still exists a singular value expansion in the following form.

**Theorem 1.** *Let  $X$  and  $Y$  be real Hilbert spaces and let  $T : X \rightarrow Y$  be a bounded linear operator. Then there exist a Borel space  $(M, \mathcal{A}, \mu)$  with a second-countable topology  $\mathcal{T}$ , isometries  $V : L^2(\mu) \rightarrow X, U : L^2(\mu) \rightarrow Y$ , and an essentially bounded measurable function  $\sigma : M \rightarrow \mathbb{R}$  such that*

$$T = Um_{\sigma}V^{\dagger}, \tag{3}$$

where  $V^{\dagger}$  is the generalized inverse of  $V$  and  $m_{\sigma}$  is the multiplication operator defined by  $\sigma$ :

$$\begin{aligned} m_{\sigma} : L^2(\mu) &\rightarrow L^2(\mu), \\ m_{\sigma}f &= \sigma f \text{ for all } f \in L^2(\mu). \end{aligned}$$

Moreover,  $\sigma > 0$  a.e.

By Borel space, we mean a measure space  $(M, \mathcal{A}, \mu)$  such that there is a topology  $\mathcal{T}$  defined on the set  $M$  and  $\mathcal{A}$  (the collection of measurable subsets of  $M$ ) is the  $\sigma$ -algebra of Borel sets of  $\mathcal{T}$ . As noted in the theorem, the topology is guaranteed to be second-countable, that is, to have a countable base. Furthermore, note that for an isometry  $V, V^{\dagger} = V^*$ .

We will call the representation of Theorem 1 the SVE of  $T$ . Pietsch ([5], Section D.3) outlines a short proof of Theorem 1 based on the polar decomposition. We give a direct proof in the next section that is analogous to the derivation of the SVD of a matrix  $A$  from the spectral decomposition of  $A^T A$ . In Section 3, we derive some basic results about this form of the SVE, including its relationship to the classical SVE of a compact operator and how to recognize from the SVE when  $\mathcal{R}(T)$  fails to be closed. We also include a brief discussion of the relationship of the SVE to notions of  $s$ -numbers (the generalization of singular values) that have appeared in the literature. In Section 4, we analyze the inverse problem  $Tx = y$ , including methods for regularizing the equation, using Theorem 1. The results in Section 4 are not new, but we hope to convince the reader that the analysis based on Theorem 1 is particularly convenient. We conclude with a brief discussion in Section 5.

## 2. The SVE of a Bounded Linear Operator

As noted above, the SVD of a matrix  $A \in \mathbb{R}^{m \times n}$  can be derived from the spectral decomposition of the symmetric matrix  $A^T A \in \mathbb{R}^{n \times n}$ . In the same way, the SVE described by Theorem 1 can be derived from the following spectral theorem for a bounded, self-adjoint linear operator.

**Theorem 2.** *Let  $T$  be a bounded and self-adjoint linear operator mapping a real Hilbert space  $X$  into itself. Then there exists a Borel space  $(M, \mathcal{A}, \mu)$  with a second-countable topology  $\mathcal{T}$ , a unitary operator  $V : L^2(\mu) \rightarrow X$ , and an essentially bounded measurable function  $\theta : M \rightarrow \mathbb{R}$  such that*

$$T = Vm_{\theta}V^{-1},$$

where  $m_{\theta}$  is the multiplication operator defined by  $\theta$ .

This version of the spectral theorem is usually stated in terms of a complex Hilbert space  $X$  and complex  $L^2(\mu)$  (see, for instance, [6] for an accessible exposition), but it can be verified that the same proof yields this representation when  $X$  is a real Hilbert space and  $L^2(\mu)$  denotes the space of real-valued square-integrable functions.

To derive Theorem 1 from Theorem 2, we require the following preliminary result.

**Lemma 1.** ([7], Lemma 2.1) Let  $(M, \mathcal{A}, \mu)$  be a measure space and let  $\theta : M \rightarrow [0, \infty)$  be a measurable function that is positive a.e. Define

$$S = \left\{ f \in L^2(\mu) : \theta^{-1}f \in L^2(\mu) \right\}.$$

Then  $S$  is dense in  $L^2(\mu)$ .

We can now prove a special case of Theorem 1.

**Theorem 3.** Let  $X$  and  $Y$  be real Hilbert spaces and let  $T : X \rightarrow Y$  be a bounded linear operator with  $\mathcal{N}(T) = \{0\}$ . Then there exist a Borel space  $(M, \mathcal{A}, \mu)$  with a second-countable topology  $\mathcal{T}$ , a unitary operator  $V : L^2(\mu) \rightarrow X$ , an isometry  $U : L^2(\mu) \rightarrow Y$ , and an essentially bounded measurable function  $\sigma : M \rightarrow [0, \infty)$  such that

$$T = Um_\sigma V^{-1}.$$

Moreover,  $\sigma > 0$  a.e. and  $\mathcal{R}(U) = \overline{\mathcal{R}(T)}$ .

**Proof.** By Theorem 2, there exist a measure space  $(M, \mathcal{A}, \mu)$ , a unitary operator  $V : L^2(\mu) \rightarrow X$ , and a bounded measurable function  $\theta : M \rightarrow \mathbb{R}$  such that

$$T^*T = Vm_\theta V^{-1}.$$

We first show that  $\theta \geq 0$  a.e., which obviously follows if we prove that

$$\langle m_\theta f, f \rangle_{L^2(\mu)} \geq 0 \text{ for all } f \in L^2(\mu).$$

However,  $m_\theta = V^{-1}T^*TV$  and hence, for any  $f \in L^2(\mu)$ ,

$$\langle m_\theta f, f \rangle_{L^2(\mu)} = \left\langle V^{-1}T^*TVf, f \right\rangle_{L^2(\mu)} = \langle TVf, TVf \rangle_Y \geq 0$$

(using the fact that  $V^{-1} = V^*$  because  $V$  is unitary). Therefore,  $\theta \geq 0$  a.e., as desired, and we can assume that  $\theta \geq 0$  everywhere.

Now define  $E = \{x \in M : \theta(x) = 0\}$ . If  $\mu(E) > 0$ , then  $\chi_E \neq 0$  in  $L^2(\mu)$ , where  $\chi_E$  is the characteristic function of the set  $E$ ; this implies that  $V\chi_E \neq 0$  in  $X$  and hence that  $T^*TV\chi_E \neq 0$  (since  $\mathcal{N}(T^*T) = \mathcal{N}(T)$  is trivial). However,

$$T^*TV\chi_E = Vm_\theta\chi_E = V(\theta\chi_E) = 0$$

because  $\theta = 0$  on  $E$  and  $\chi_E = 0$  on  $M \setminus E$ . This contradiction shows that  $\mu(E)$  must be zero, that is,  $\theta > 0$  a.e. in  $M$ .

Therefore, if we define  $\sigma = \sqrt{\theta}$  and

$$S = \left\{ f \in L^2(\mu) : \sigma^{-1}f \in L^2(\mu) \right\},$$

then Lemma 1 applies and we see that  $S$  is dense in  $L^2(\mu)$ . We define  $U : S \rightarrow Y$  by  $U = TVm_{\sigma^{-1}}$ . Since  $\sigma^{-1}f \in L^2(\mu)$  for all  $f \in S$ ,  $U$  is well-defined, and we see that it is linear and densely defined. We now show that  $\|Uf\|_Y = \|f\|_{L^2(\mu)}$  for all  $f \in S$ . We have

$$\begin{aligned} \|Uf\|_Y^2 &= \langle Uf, Uf \rangle_Y = \langle TVm_{\sigma^{-1}}f, TVm_{\sigma^{-1}}f \rangle_Y \\ &= \left\langle f, m_{\sigma^{-1}}V^{-1}T^*TVm_{\sigma^{-1}}f \right\rangle_{L^2(\mu)} \\ &= \langle f, m_{\sigma^{-1}}m_{\sigma^2}m_{\sigma^{-1}}f \rangle_{L^2(\mu)} \\ &= \langle f, f \rangle_{L^2(\mu)} = \|f\|_{L^2(\mu)}^2. \end{aligned}$$

However, now we see that  $U$  is bounded and densely defined and hence it extends to a bounded operator defined on all of  $L^2(\mu)$ . We will use  $U$  to denote this extension as well (therefore  $U : L^2(\mu) \rightarrow Y$  satisfies  $U|_S = TVm_{\sigma^{-1}}$ ). By continuity, we have  $\|Uf\|_Y = \|f\|_{L^2(\mu)}$  for all  $f \in L^2(\mu)$  and hence  $U$  is an isometry.

Next, we show that  $T = Um_{\sigma}V^{-1}$ . For each  $x \in X$ ,  $m_{\sigma}V^{-1}x \in S$  because  $m_{\sigma^{-1}}m_{\sigma}V^{-1}x = V^{-1}x \in L^2(\mu)$ . However, then we see that, for each  $x \in X$ ,

$$Um_{\sigma}V^{-1}x = TVm_{\sigma^{-1}}m_{\sigma}V^{-1}x = TVV^{-1}x = Tx.$$

Therefore,  $Um_{\sigma}V^{-1} = T$ , as desired.

It remains only to show that  $\mathcal{R}(U) = \overline{\mathcal{R}(T)}$ . Let  $y \in \overline{\mathcal{R}(T)}$ . Then there exists  $\{x_n\} \subset X$  such that  $Tx_n \rightarrow y$ , that is,  $Um_{\sigma}V^{-1}x_n \rightarrow y$ . It follows that  $\{Um_{\sigma}V^{-1}x_n\}$  is a Cauchy sequence and hence, because  $U$  is an isometry,  $\{m_{\sigma}V^{-1}x_n\}$  is a Cauchy sequence in  $L^2(\mu)$ . Suppose  $m_{\sigma}V^{-1}x_n \rightarrow f \in L^2(\mu)$ . Then

$$Uf = \lim_{n \rightarrow \infty} Um_{\sigma}V^{-1}x_n = y,$$

which shows that  $y \in \mathcal{R}(U)$ . Since  $\mathcal{R}(U) \subset \overline{\mathcal{R}(T)}$  by definition of  $U$ , it follows that  $\mathcal{R}(U) = \overline{\mathcal{R}(T)}$ . This completes the proof.  $\square$

Theorem 1 is an immediate corollary of Theorem 3.

**Proof of Theorem 1.** If we apply Theorem 3 to  $T|_{\mathcal{N}(T)^{\perp}}$ , we obtain

$$T|_{\mathcal{N}(T)^{\perp}} = Um_{\sigma}V_1^{-1},$$

where  $U : L^2(\mu) \rightarrow Y$  is an isometry and  $V_1 : L^2(\mu) \rightarrow \mathcal{N}(T)^{\perp}$  is unitary. We claim that

$$T = Um_{\sigma}V^{\dagger}, \tag{4}$$

where  $V : L^2(\mu) \rightarrow X$  is defined by  $Vf = V_1f$  for all  $f \in L^2(\mu)$ . Since  $V$  is obviously an isometry, proving that (4) holds will complete the proof. By definition,  $V^{\dagger}x$  is the minimum-norm least-squares solution of  $Vf = x$ . Moreover, since  $\mathcal{N}(V)$  is trivial,  $V^{\dagger}x$  is the unique least-squares solution of  $Vf = x$ , which is defined by

$$V^{\dagger}x = \begin{cases} V_1^{-1}x, & \text{if } x \in \mathcal{N}(T)^{\perp}, \\ 0, & \text{if } x \in \mathcal{N}(T). \end{cases}$$

However, then

$$x \in \mathcal{N}(T)^{\perp} \Rightarrow Um_{\sigma}V^{\dagger}x = Um_{\sigma}V_1^{-1}x = Tx$$

and

$$x \in \mathcal{N}(T) \Rightarrow Um_{\sigma}V^{\dagger}x = 0 = Tx.$$

This proves (4), which completes the proof.  $\square$

### 3. Some Properties of the SVE of a Bounded Linear Operator

#### 3.1. Relationship to the SVE of a Compact Operator

Suppose  $T : X \rightarrow Y$  is a compact operator with singular value expansion

$$T = \sum_{n=1}^{\infty} \sigma_n \psi_n \otimes \phi_n.$$

As noted above,  $\{\phi_n\}$  is a complete orthonormal sequence for  $\mathcal{N}(T)^\perp$  and  $\{\psi_n\}$  is a complete orthonormal sequence for  $\overline{\mathcal{R}(T)}$ .

Let us define  $M = \mathbb{Z}^+$ ,  $\mathcal{A} = \mathcal{P}(\mathbb{Z}^+)$  (the power set of  $\mathbb{Z}^+$ ), and  $\mu$  to be counting measure (that is, for an  $E \subset \mathbb{Z}^+$ ,  $\mu(E)$  is the cardinality of  $E$ ). Then  $L^2(\mu)$  is the space of square summable sequences of real numbers (usually denoted by  $\ell^2$ ) and, for  $\alpha = \{\alpha_k\} \in L^2(\mu)$ ,

$$\int \alpha^2 d\mu = \sum_{k=1}^{\infty} \alpha_k^2.$$

We define  $V : L^2(\mu) \rightarrow X$  by

$$V(\alpha_k) = \sum_{k=1}^{\infty} \alpha_k \phi_k.$$

Then it is straightforward to verify that  $V$  is an isometry and that  $V^\dagger = V^*$  is defined by  $(V^\dagger(x))_k = \langle \phi_k, x \rangle_X$  for all  $k \in \mathbb{Z}^+$ . The sequence  $\sigma = \{\sigma_k\}$  is bounded and measurable with respect to the measure space  $(M, \mathcal{A}, \mu)$  and

$$m_\sigma \alpha = \{\sigma_k \alpha_k\}.$$

Finally,  $U : L^2(\mu) \rightarrow Y$  is defined to be the extension to all of  $L^2(\mu)$  of  $TVm_{\sigma^{-1}}$ , which is given, for  $\alpha \in L^2(\mu)$  such that  $m_{\sigma^{-1}}\alpha$  also lies in  $L^2(\mu)$ , by

$$TVm_{\sigma^{-1}}\alpha = T \left( \sum_{k=1}^{\infty} \frac{\alpha_k}{\sigma_k} \phi_k \right) = \sum_{k=1}^{\infty} \sigma_k \frac{\alpha_k}{\sigma_k} \psi_k = \sum_{k=1}^{\infty} \alpha_k \psi_k.$$

Clearly this formula extends to every  $\alpha \in L^2(\mu)$ .

Therefore, for each  $x \in X$ , we have

$$\begin{aligned} Um_\sigma V^{-1}x &= \sum_{k=1}^{\infty} (m_\sigma V^{-1}x)_k \psi_k = \sum_{k=1}^{\infty} \sigma_k \langle \phi_k, x \rangle_X \psi_k \\ &= \left( \sum_{k=1}^{\infty} \sigma_k \psi_k \otimes \phi_k \right) x \\ &= Tx. \end{aligned}$$

This shows that  $T = Um_\sigma V^{-1}$  and also that  $Um_\sigma V^{-1}$  is just another way of writing the usual singular value expansion of  $T$ .

#### 3.2. The SVE of Operators Related to $T$

Throughout the rest of the paper, we assume that  $T : X \rightarrow Y$  is a bounded linear operator from one real Hilbert space  $X$  to another such space  $Y$ , and that  $T = Um_\sigma V^\dagger$  is the SVE of Theorem 1. The associated Borel space is denoted by  $(M, \mathcal{A}, \mu)$  and the topology of  $M$  is denoted by  $\mathcal{T}$ .

Since  $V^* = V^\dagger$  and  $V^*V = I$  hold for any isometry  $V$ , we immediately have the following:

$$\begin{aligned} T^* &= Vm_\sigma U^\dagger, \\ T^*T &= Vm_{\sigma^2} V^\dagger, \\ TT^* &= Um_{\sigma^2} U^\dagger. \end{aligned}$$

### 3.3. Inverse Problems and the SVE

It is well known that the equation  $Tx = y$  represents a true inverse problem if and only if  $\mathcal{R}(T)$  fails to be closed (see ([3], Section 2.3)). In this case, the solution  $x$  (if it exists) does not depend continuously on the data vector  $y$ . One way to state this precisely is to note that the generalized inverse  $T^\dagger$  is unbounded if and only if  $\mathcal{R}(T)$  fails to be closed. We will study  $T^\dagger$  below in Section 4; for now, we prove the following necessary and sufficient condition for  $\mathcal{R}(T)$  to be closed.

**Theorem 4.** *The range of  $T$  is closed if and only if the function  $\sigma$  is bounded away from zero, that is, if and only if there exists  $\gamma > 0$  such that  $\sigma(t) \geq \gamma$  for almost all  $t \in M$ .*

**Proof.** It is a standard result that  $\mathcal{R}(T)$  is closed if and only if there exists  $\gamma > 0$  such that

$$\|Tx\|_Y \geq \gamma\|x\|_X \quad \forall x \in \mathcal{N}(T)^\perp \tag{5}$$

(see ([3], Theorem 2.20)). If there exists  $\gamma > 0$  such that  $\sigma(t) \geq \gamma$  for almost all  $t \in M$ , then

$$\|m_\sigma f\|_{L^2(\mu)} \geq \gamma\|f\|_{L^2(\mu)} \quad \forall f \in L^2(\mu).$$

We have  $\|Vf\|_X = \|f\|_{L^2(\mu)}$  for all  $f \in L^2(\mu)$ ; moreover,  $V$  defines an isomorphism from  $L^2(\mu)$  to  $\mathcal{N}(T)^\perp$ . Therefore,

$$\|Tx\|_Y = \|Um_\sigma V^\dagger x\|_Y = \|m_\sigma(V^\dagger x)\|_{L^2(\mu)} \geq \gamma\|V^\dagger x\|_{L^2(\mu)} = \gamma\|x\|_X \quad \forall x \in \mathcal{N}(T)^\perp.$$

Conversely, suppose  $\sigma$  is not bounded away from zero. It follows that

$$S_k = \{t \in M : \sigma(t) < 1/k\}$$

has positive measure for each  $k \in \mathbb{Z}^+$ . Therefore, with  $f_k = \chi_{S_k}$ , we have

$$\|m_\sigma f_k\|_{L^2(\mu)} < \frac{1}{k}\|f_k\|_{L^2(\mu)} \quad \forall k \in \mathbb{Z}^+$$

and hence

$$\|Tx_k\|_Y < \frac{1}{k}\|x_k\|_X \quad \forall k \in \mathbb{Z}^+,$$

where  $x_k = Vf_k$ . This shows that  $\mathcal{R}(T)$  fails to be closed in this case, and the proof is complete.  $\square$

### 3.4. s-Numbers

Given the utility of singular values for matrices and compact operators, it is natural to try to extend the concept to more general operators. This can be done in various ways. The Courant-Fischer characterization [8,9] of the singular values  $\sigma_1, \sigma_2, \dots$  of a compact operator  $T : X \rightarrow Y$  is the following:

$$\sigma_k = \min_{\substack{S \subset X \\ \dim(S) < k}} \max_{\substack{x \in S^\perp \\ x \neq 0}} \frac{\|Tx\|_Y}{\|x\|_X}, \quad k = 1, 2, \dots \tag{6}$$



Equation (6) can be taken as the definition of the  $s$ -numbers of an arbitrary bounded linear operator  $T : X \rightarrow Y$  by replacing “min” with “inf” and “max” with “sup.”

Alternatively, the singular values of a compact operator can be characterized [10] as

$$\sigma_k = \min_{\substack{\tau: X \rightarrow Y \\ \text{rank}(\tau) < k}} \|\tau - T\|, \tag{7}$$

which again extends to arbitrary bounded linear operators  $T$  by replacing “min” with “inf”. It can be shown that (6) and (7) are equivalent for operators on Hilbert space. In fact, Pietsch [11] formulated a list of five axioms that characterize  $s$ -numbers on Hilbert space (in the sense that two definitions, such as (6) or (7), that satisfy the axioms are equivalent).

The above definitions of  $s$ -numbers are limited, in that they may not give much information if the continuous spectrum of  $T^*T$  is nonempty. Fack and Kosaki [12] defined generalized  $s$ -numbers for certain operators in a von Neumann algebra, and their techniques allow for a nonempty continuous spectrum. In the context of the SVE presented in this paper, it would be natural to define the set of  $s$ -numbers of a bounded linear operator  $T : X \rightarrow Y$  as the essential range of  $\sigma$  (where  $T = Um_\sigma V^\dagger$ ); we refer the reader to the first author’s PhD dissertation [13] for a discussion. The relationship between these two approaches remains to be investigated.

#### 4. The SVE and Tikhonov Regularization

We believe that Theorem 1 will prove to be useful in a variety of applications. Here we show that it can be used to give transparent proofs of convergence theorems in the theory of Tikhonov regularization, the most popular method for addressing inverse problems.

We consider an equation of the form  $Tx = y$ . We are given  $y \in Y$  and wish to compute or estimate  $x \in X$  satisfying the equation. The problem is well-posed if there exists a unique solution  $x$  for each  $y \in Y$ , where  $x$  depends continuously on  $y$ . Existence fails, at least for some  $y \in Y$ , if  $\mathcal{R}(T)$  is a proper subspace of  $Y$ . However, in that case, it is common to settle for a least-squares solution of the equation, that is, an  $x \in X$  that minimizes  $\|Tx - y\|_Y^2$ . Uniqueness fails to hold if  $\mathcal{N}(T)$  is nontrivial, but we can select a unique (least-squares) solution by choosing the unique solution lying in  $\mathcal{N}(T)^\perp$ , which is equivalent to choosing the minimum-norm least-squares solution. The interesting case occurs when  $\mathcal{R}(T)$  fails to be closed. In that case,

1. Least-squares solutions exists only for  $y$  in the dense subspace  $\mathcal{R}(T) \oplus \mathcal{R}(T)^\perp$  of  $Y$ ;
2. For each  $y \in \mathcal{R}(T) \oplus \mathcal{R}(T)^\perp$ , there exists a unique minimum-norm least-squares solution  $x \in \mathcal{N}(T)^\perp$ , but  $x$  does not depend continuously on  $y$ .

The generalized inverse  $T^\dagger : D(T^\dagger) \rightarrow X$ , where  $D(T^\dagger) = \mathcal{R}(T) \oplus \mathcal{R}(T)^\perp$ , is defined by the condition that  $T^\dagger y$  is the minimum-norm least-squares solution of  $Tx = y$ . It follows from the above discussion that, when  $\mathcal{R}(T)$  fails to be closed, then  $T^\dagger$  is a densely defined unbounded linear operator. In this case, the problem  $Tx = y$ , even when interpreted as asking for the minimum-norm least-squares solution, is ill-posed in that the solution does not depend continuously on the data. In this case, we call  $Tx = y$  a (linear) inverse problem.

Many regularization techniques for solving  $Tx = y$  approximate  $T^\dagger$  by a family  $\{R_\lambda : \lambda > 0\}$  of bounded operators. Here  $\lambda$  is called the regularization parameter and it is required that  $R_\lambda \rightarrow T^\dagger$  pointwise as  $\lambda \rightarrow 0^+$ . The most popular regularization method is Tikhonov regularization, in which  $R_\lambda = (T^*T + \lambda I)^{-1}T^*$ . This operator arises from solving

$$\min_{x \in X} \|Tx - y\|_Y^2 + \lambda \|x\|_X^2.$$

We first show that  $R_\lambda y \rightarrow T^\dagger y$  for all  $y \in D(T^\dagger)$ . For convenience, we will write  $x_{\lambda,y} = R_\lambda y$ .

By definition,  $T^\dagger y \in \mathcal{N}(T)^\perp$ . Furthermore,  $x_{\lambda,y}$  is defined by the equation

$$(T^*T + \lambda I)x_{\lambda,y} = T^*y,$$

which implies that

$$x_{\lambda,y} = \lambda^{-1} (T^*y - T^*Tx_{\lambda,y}) \in \mathcal{R}(T^*) \subset \mathcal{N}(T)^\perp.$$

We will write  $\bar{y} = \text{proj}_{\overline{\mathcal{R}(T)}}y$ , and notice that  $T^*y = T^*\bar{y}$  since  $\bar{y} - y \in \mathcal{R}(T)^\perp = \mathcal{N}(T^*)$ . It follows that  $x_{\lambda,y} = (T^*T + \lambda I)^{-1}T^*y = (T^*T + \lambda I)^{-1}T^*\bar{y}$ . Moreover, since the least-squares solutions of  $Tx = y$  are precisely the solutions of  $T^*Tx = T^*y$ , it also follows that  $T^\dagger y = T^\dagger \bar{y}$ . These two facts (that  $T^\dagger y, x_{\lambda,y} \in \mathcal{N}(T)^\perp$  and that  $T^\dagger y, x_{\lambda,y}$  can be defined by  $\bar{y}$  in place of  $y$ ) make it convenient to use the singular value expansion as expressed in Theorem 3 (as opposed to the version of Theorem 1).

Suppose that, using the notation of Theorem 3,

$$T|_{\mathcal{N}(T)^\perp} = Um_\sigma V^{-1},$$

and recall that  $\sigma > 0$  a.e. in  $M$ . Since  $x = T^\dagger y$  satisfies  $Tx = \bar{y}$ ,

$$Um_\sigma V^{-1}T^\dagger y = \bar{y} \rightarrow T^\dagger y = Vm_{\sigma^{-1}}U^\dagger \bar{y}.$$

Furthermore,

$$T^*T + \lambda I = Vm_{\sigma^2}V^{-1} + \lambda VV^{-1} = Vm \left[ \sigma^2 + \lambda \right] V^{-1}$$

(where we write  $m_\theta$  as  $m[\theta]$  when it is convenient to do so), and hence

$$(T^*T + \lambda I)^{-1} = Vm \left[ (\sigma^2 + \lambda)^{-1} \right] V^{-1}.$$

This leads to

$$x_{\lambda,y} = (T^*T + \lambda I)^{-1} T^*\bar{y} = Vm \left[ \frac{\sigma}{\sigma^2 + \lambda} \right] U^\dagger \bar{y}.$$

However, then

$$\begin{aligned} T^\dagger y - x_{\lambda,y} &= Vm_{\sigma^{-1}}U^\dagger \bar{y} - Vm \left[ \frac{\sigma}{\sigma^2 + \lambda} \right] U^\dagger \bar{y} \\ &= Vm \left[ \frac{1}{\sigma} - \frac{\sigma}{\sigma^2 + \lambda} \right] U^\dagger \bar{y} \\ &= Vm \left[ \frac{\lambda}{\sigma^2 + \lambda} \right] m_{\sigma^{-1}}U^\dagger \bar{y}. \end{aligned}$$

To show that  $T^\dagger y - x_{\lambda,y} \rightarrow 0$  as  $\lambda \rightarrow 0$ , it suffices to show that

$$m \left[ \frac{\lambda}{\sigma^2 + \lambda} \right] m_{\sigma^{-1}}U^\dagger \bar{y} \rightarrow 0 \text{ as } \lambda \rightarrow 0.$$

Moreover, since  $\bar{y} \in \mathcal{R}(T)$  (as opposed to merely belonging to  $\overline{\mathcal{R}(T)}$ —this follows from the fact that  $y \in D(T^\dagger) = \mathcal{R}(T) \oplus \mathcal{R}(T)^\perp$ ), it follows that  $\bar{f} = m_{\sigma^{-1}}U^\dagger \bar{y} \in L^2(\mu)$ . However, then, since  $\lambda/(\sigma^2 + \lambda)$  is bounded on  $M$  and goes to 0 pointwise as  $\lambda \rightarrow 0$ , it follows that

$$\left\| m \left[ \frac{\lambda}{\sigma^2 + \lambda} \right] m_{\sigma^{-1}}U^\dagger \bar{y} \right\|_{L^2(\mu)} = \int \frac{\lambda^2}{(\sigma^2 + \lambda)^2} \bar{f}^2 \rightarrow 0 \text{ as } \lambda \rightarrow 0$$

by the dominated convergence theorem. This shows that  $x_{\lambda,y} \rightarrow T^\dagger y$  as  $\lambda \rightarrow 0$ . Henceforth, we will write  $x_{0,y} = T^\dagger y$ .

We will prove two other results to demonstrate the usefulness of the singular value expansion. The result that we just defined shows that, for each  $y \in D(T^\dagger)$ ,  $x_{\lambda,y} \rightarrow x_{0,y}$  as  $\lambda \rightarrow 0$ . However, the result says nothing about the rate of convergence and, in fact, the convergence can be arbitrarily slow depending on the data  $y \in D(T^\dagger)$  (or, equivalently, on the solution  $x_{0,y}$ ). For certain  $x_{0,y}$ , though, we can bound the rate of convergence. We will not attempt to prove the most general theorem, but rather just consider what turns out to be the optimal rate of convergence. We will show that if  $x_{0,y} \in \mathcal{R}(T^*T)$ , then

$$\|x_{0,y} - x_{\lambda,y}\|_X = O(\lambda).$$

From above, we have

$$x_{0,y} - x_{\lambda,y} = Vm \left[ \frac{\lambda}{\sigma^2 + \lambda} \right] m_{\sigma^{-1}} U^\dagger \bar{y},$$

and  $\bar{y} = Tx_{0,y} = Um_\sigma V^{-1}x_{0,y}$ . Therefore,

$$x_{0,y} - x_{\lambda,y} = Vm \left[ \frac{\lambda}{\sigma^2 + \lambda} \right] V^{-1}x_{0,y}.$$

If we now assume that  $x_{0,y} \in \mathcal{R}(T^*T)$ , say  $x_{0,y} = T^*Tv_0 = Vm_{\sigma^2}V^{-1}v_0$  for some  $v_0 \in \mathcal{N}(T)^\perp$ , then we obtain

$$x_{0,y} - x_{\lambda,y} = Vm \left[ \frac{\lambda}{\sigma^2 + \lambda} \right] m_{\sigma^2}V^{-1}v_0 = \lambda Vm \left[ \frac{\sigma^2}{\sigma^2 + \lambda} \right] V^{-1}v_0.$$

Since  $V$  is an isometry and  $\left\| m \left[ \frac{\sigma^2}{\sigma^2 + \lambda} \right] \right\|$  is bounded by 1 (in general,  $\|m_\theta\| = \text{ess sup } \theta$ ), it follows that  $\|x_{0,y} - x_{\lambda,y}\|_X = O(\lambda)$ , as desired.

We can also prove the following converse result, namely, that if  $y \in D(T^\dagger)$  and  $\|x_{0,y} - x_{\lambda,y}\|_X = O(\lambda)$ , then  $x_{0,y} \in \mathcal{R}(T^*T)$ . We will use the fact, easily verified, that  $x \in \mathcal{R}(T^*T)$  if and only if  $m_{\sigma^{-2}}V^{-1}x \in L^2(\mu)$ , that is, if and only if  $V^{-1}x$  belongs to the domain of the densely defined operator  $m_{\sigma^{-2}}$ . Let us write  $f_0 = V^{-1}x_{0,y}$ ; then we must show that

$$\int \frac{f_0^2}{\sigma^4} d\mu < \infty.$$

We have

$$x_{0,y} - x_{\lambda,y} = Vm \left[ \frac{\lambda}{\sigma^2 + \lambda} \right] V^{-1}x_{0,y} = \lambda Vm \left[ \frac{1}{\sigma^2 + \lambda} \right] V^{-1}x_{0,y},$$

which implies that

$$\frac{\|x_{0,y} - x_{\lambda,y}\|_X^2}{\lambda^2} = \left\| Vm \left[ \frac{1}{\sigma^2 + \lambda} \right] V^{-1}x_{0,y} \right\|_X^2 = \left\| m \left[ \frac{1}{\sigma^2 + \lambda} \right] f_0 \right\|_{L^2(\mu)}^2.$$

Since  $\|x_{0,y} - x_{\lambda,y}\|_X = O(\lambda)$  by assumption, there exists a constant  $C > 0$  such that

$$\left\| m \left[ \frac{1}{\sigma^2 + \lambda} \right] f_0 \right\|_{L^2(\mu)}^2 \leq C \text{ for all } \lambda > 0,$$

that is,

$$\int \frac{f_0^2}{(\sigma^2 + \lambda)^2} d\mu \leq C \text{ for all } \lambda > 0.$$

Since  $f_0^2/(\sigma^2 + \lambda)^2$  converges monotonically to  $f_0^2/\sigma^4$  a.e. in  $M$  as  $\lambda \rightarrow 0$ , it follows from the monotone convergence theorem that

$$\int \frac{f_0^2}{\sigma^4} d\mu \leq C < \infty,$$

as desired.

### 5. Discussion

The proofs of the last section are offered as an illustration of the power of the singular value expansion. The reader can compare these proofs to other treatments of the same results that can be found in the literature on inverse problems. In Groetsch’s monograph [14], the analysis is restricted to compact operators and Theorem 2.1.1, Corollary 3.1.2, and Theorem 3.2.2 correspond to our results; Groetsch’s proofs use the singular value expansion (2) for compact operators. The reader will see that our proofs are direct generalizations of the derivations given there, and also that there is no difficulty in extending his other conclusions to general bounded linear operators. Groetsch does present his theory in greater generality, with much of the analysis applying to a certain family of regularization operators  $R_\lambda$ , as opposed to just the Tikhonov approach. We restricted our presentation to Tikhonov regularization simply for convenience of exposition; there would be no difficulty in reproducing his results in the same level of generality.

To extend the results of [14] to general operators, the standard approach is to use the spectral representation of  $T^*T$  in the form

$$T = \int \alpha dE_\alpha,$$

where  $\{E_\alpha\}$  is the spectral resolution of  $T^*T$ , and apply the so-called functional calculus, which allows the representation of functions of  $T^*T$  via

$$f(T^*T) = \int f(\alpha) dE_\alpha.$$

It can be shown, for example, that

$$T^\dagger = \int \alpha^{-1} dE_\alpha.$$

A good reference for this approach is the book [15] by Engl, Hanke, and Neubauer, which (among other things) extends the results of [14] to general bounded linear operators. There is no intrinsic difficulty in doing so, but it may be argued that the arguments are less intuitive and therefore harder to follow. For instance, it is necessary to work with integrals of the following types:

$$\int (\dots) dE_\alpha, \int (\dots) d\|E_\alpha x\|_X^2, \int (\dots) d\langle E_\alpha x, y \rangle_X.$$

As Halmos stated in his popular expository article [6] on the spectral theorem (one of the most-downloaded articles from the American Mathematical Monthly),

The result (namely, the spectral theorem for Hermitian matrices, when expressed using a resolution  $\{E_\alpha\}$  of the identity) is not intuitive in any language; neither Stieltjes integrals with unorthodox multiplicative properties, nor bounded operator representations of function algebras, are in the daily toolbox of every working mathematician. In contrast, the formulation of the spectral theorem given below uses only the relatively elementary concepts of measure theory.

We believe that the singular value expansion for general bounded linear operators, as described above, offers a similarly intuitive tool that can replace the standard use of the functional calculus in many contexts.

**Author Contributions:** Conceptualization, M.S.G.; methodology, D.K.C. and M.S.G.; formal analysis, D.K.C. and M.S.G.; writing—original draft preparation, M.S.G.; writing—review and editing, D.K.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Gockenbach, M.S. *Finite-Dimensional Linear Algebra*; CRC Press: Boca Raton, FL, USA, 2010.
2. Trefethen, L.N.; Bau, D. *Numerical Linear Algebra*; Society for Industrial and Applied Mathematics: Philadelphia, PA, USA, 1997.
3. Gockenbach, M.S. *Linear Inverse Problems and Tikhonov Regularization*; Mathematical Association of America: Washington, DC, USA, 2016.
4. Young, N. *An Introduction to Hilbert Space*; Cambridge University Press: Cambridge, UK, 1988.
5. Pietsch, A. *Operator Ideals*; North-Holland Publishing Company: Amsterdam, The Netherlands, 1980.
6. Halmos, P.R. What does the spectral theorem say? *Am. Math. Mon.* **1963**, *70*, 241–247. [[CrossRef](#)]
7. Gockenbach, M.S. Generalizing the GSVD. *SIAM J. Numer. Anal.* **2016**, *54*, 2517–2540. [[CrossRef](#)]
8. Courant, R. Über die Eigenwerte bei den Differentialgleichungen der Mathematischen Physik. *Math. Z.* **1920**, *7*, 1–57. [[CrossRef](#)]
9. Fischer, E. Über quadratische Formen mit reellen Koeffizienten. *Monatshefte Math. Phys.* **1905**, *16*, 234–249. [[CrossRef](#)]
10. Fiedler, M.; Pták, V. Sur la meilleure approximation des transformations linéaires par des transformations de rang prescrit. *C. R. Acad. Sci. Paris* **1962**, *254*, 3805–3807.
11. Pietsch, A.  $s$ -Numbers of operators in Banach spaces. *Studia Math.* **1974**, *51*, 201–223. [[CrossRef](#)]
12. Fack, T.; Kosaki, H. Generalized  $s$ -numbers of  $\tau$ -measurable operators. *Pac. J. Math.* **1986**, *123*, 269–300. [[CrossRef](#)]
13. Crane, D.K. The Singular Value Expansion for Compact and Non-Compact Operators. Ph.D. Thesis, Michigan Technological University, Houghton, MI, USA, 2020.
14. Groetsch, C.W. *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind*; Pitman Publishing Inc.: Boston, MA, USA, 1984.
15. Engl, H.W.; Hanke, M.; Neubauer, A. *Regularization of Inverse Problems*; Mathematics and Its Applications; Kluwer Academic Publishers Group: Dordrecht, The Netherlands, 1996.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).