

# Optimal numerical approximation of a linear operator

H.F. Weinberger

Institute for Mathematics and its Applications  
University of Minnesota

(To appear in Linear Algebra and its Applications)

## Abstract

Let  $S$  be a bounded linear transformation from a Hilbert space  $B$  to a Hilbert space  $\Sigma$ .  $Su$  can be thought of as the solution of a linear differential equation with right-hand side, initial data, or boundary data  $u$ . Given the incomplete information  $Nu = v$ ,  $\|u\|_B < 1$  about the data where  $N$  is a linear operator from  $B$  to a Euclidean space  $E_n$  and a linear interpolation operator  $M$  from  $E_m$  to  $\Sigma$ , one defines the optimal approximation to  $Su$  to be the point  $Ma(v)$  in the range of  $M$  which is the center of the smallest ball containing all points of the form  $Su$  with  $Nu = v$  and  $\|u\|_B < 1$  and centered in  $M$ .

A characterization is given for the optimal approximation  $Ma(v)$ . It is shown to be unique and, in general, nonlinear.

Simpler approximations and relations with other concepts of optimality are investigated.

---

This research was supported by the National Science Foundation through grant  
NSF MCS 81-02609

## 1. Introduction

Many linear problems of numerical analysis can be formulated in the following way: One is given a set of  $n$  linear data  $Nu = v$  and a bound for the norm  $\|u\|_B$  of an otherwise unknown element  $u$  of a Hilbert space  $B$ . One wishes to find a best approximation to the element  $Su$ , where  $S$  is a bounded linear operator from  $B$  to another Hilbert space  $\Sigma$ . For example,  $Su$  may be the solution of an ordinary or partial differential equation with right-hand side, initial data, or boundary data  $u$ .

The approximation is to involve finitely many computations.

When  $\Sigma$  is one-dimensional, it was observed by Golomb and the author [7] that the set of  $u$  in  $B$  which are compatible with the data is a hypercircle, which is symmetric about its center  $q(v)$ . Hence the set of possible values of  $Su$  is an interval centered at  $Sq(v)$ . Therefore the best approximation to  $Su$  is  $Sq(v)$  in the sense that the maximum possible error  $|Su - c|$  among  $u$  which satisfy  $Nu = v$ ,  $\|u\|_B < 1^{(*)}$  is minimized by  $c = Sq(v)$ . This best approximation  $Sq(v)$  depends linearly upon the data  $v$ .

This reasoning works just as well when  $\Sigma$  is higher-dimensional. The mapping from  $v$  to the optimal approximation is always the linear map  $Sq(v)$ . (See [10, Theo. 3].)

---

(\*) The apparently more general inequality  $\|u\|_B < k$  is easily reduced to this case, either by replacing  $u$  by  $u/k$  and  $v$  by  $v/k$  or by redefining the norm on  $B$ .

When  $\Sigma$  is infinite-dimensional, one wishes to restrict oneself to elements which can be defined by finitely many computations. This can be done in two different ways.

Aubin [2], Babuska, Prager, and Vitasek [3], and Miranker [11] made the computation finite by replacing the operator  $S$  by  $\psi S$  where  $\psi$  is a linear map from  $\Sigma$  to the Euclidean space  $E_m$ , (that is, a set of  $m$  linear functionals) and found the above optimal approximation  $\psi S q(v)$ , which is linear. The connection with the method of [7] was pointed out by Micchelli and Miranker [9].

It appears more natural to me to define an interpolation operator  $M$  from  $E_m$  to  $\Sigma$  and to call an element  $\hat{M}a(v)$  of the range of  $M$  an optimal approximation to  $Su$  if the maximum error

$$d(a;v) \equiv \sup \{ \|Su - Ma\|_{\Sigma} : Nu = v, \|u\|_B < 1 \} .$$

is minimized when  $a = \hat{a}$ . We shall characterize the optimal approximation  $\hat{M}a$  in Section 3 and show that it is uniquely defined.

It is interesting to know whether the optimal approximation is again linear. It is shown in Section 4 that in most cases  $Ma(v)$  coincides with a linear transformation on  $v$  when  $v$  is small, but not when  $v$  is large. Thus  $\hat{M}a(v)$  is, in general, not linear in  $v$ .

A somewhat coarser error bound for an approximation scheme  $Ma(v)$ , where  $a(v)$  is any mapping from  $E_n$  to  $E_m$ , is given by the functional

$$\begin{aligned} D[a(\cdot)] &= \sup \{ \|Su - Ma(Nu)\|_{\Sigma} : \|u\|_B < 1 \} \\ &= \sup \{ d(a(v);v) : \exists u \ni Nu = v, \|u\|_B < 1 \} \end{aligned}$$

Micchelli and Rivlin [10] call a mapping  $M_a$  for which  $D[a(\cdot)]$  takes on its minimum value an optimal recovery by a restricted algorithm.

It is clear that the optimal approximation  $M_{\hat{a}}(v)$  is such a restricted optimal recovery and that  $M_{\tilde{a}}(v)$  is a restricted optimal recovery if and only if

$$d(\hat{a}(v); v) < d(\tilde{a}(v); v) < D[\hat{a}]$$

for all  $v$ . Thus any restricted optimal recovery  $\tilde{a}$  must agree with  $\hat{a}$  at those  $v$  at which  $d(\hat{a}(v), v)$  attains its maximum, but there is leeway in  $\tilde{a}$  at the other values of  $v$ . In general,  $M_{\hat{a}}(v)$  gives a smaller error bound than  $M_{\tilde{a}}(v)$  for most values of  $v$ .

It can be shown [14] that if one defines

$$\kappa = \sup \left\{ \frac{\|Su\|}{\|u\|_B} : Nu = 0 \right\} \quad (1.1)$$

and

$$\kappa^* = \sup \left\{ \frac{\|S^* \phi\|_B}{\|\phi\|_{\Sigma}} : M^* \phi = 0 \right\}, \quad (1.2)$$

then for a linear map  $Av$

$$D[A] \geq \max(\kappa, \kappa^*),$$

and that there is an  $\tilde{A}$  for which equality holds. Micchelli and Rivlin [10, Theo. 1 and Theo. H] have proved the inequality

$$D[a(v)] \geq \max(\kappa, \kappa^*)$$

for any map. Consequently, a linear map  $\tilde{A}$  for which

$$D[\tilde{A}] = \max(\kappa, \kappa^*)$$

provides a restricted optimal recovery  $\tilde{M}\tilde{A}$ , and

$$D[\tilde{A}] = D[\hat{a}] = \max(\kappa, \kappa^*) .$$

The set of all linear maps  $\tilde{A}$  which provide restricted optimal recoveries (there are, in general, many of them) was recently characterized by Davis, Kahan, and the author [6]. Related results have been found by M.G. Krein [8], Parrott [12], Arsene and Gheonda [1], Smul'jan and Janovskaya [13], and David [4,5].

Section 5 investigates some further properties of the optimal approximation and derives some simpler approximations which are not quite optimal. The application of the results to the numerical solution of boundary value problems is indicated in Section 6.

## 2. FORMULATION OF THE PROBLEM

We wish to approximate a bounded linear operator  $S$  from a Hilbert space  $B$  to a Hilbert space  $\Sigma$  in the following sense. We are given a bound

$$\|u\|_B \leq 1$$

and the set of linear equations  $Nu = v$ , where  $N$  is a bounded linear operator from  $B$  onto the Euclidean space  $E_n$ , but no other information. We fix a linear injection  $M$  (an interpolation) from the Euclidean space  $E_m$  to  $\Sigma$ . We wish to find an element  $a$  of  $E_m$  such that the maximum possible error

$$d(a; v) \equiv \sup\{\|Sv - Ma\|_\Sigma \mid Nv = v, \|v\|_B \leq 1\} \quad (2.1)$$

which can be made when  $Su$  is approximated by  $Ma$  is minimized.

In other words, we seek the ball of smallest radius which contains the set  $\{Sv \mid Nv = v, \|v\|_B \leq 1\}$  and whose center lies in the range of  $M$ .

We must of course, require that the set of admissible  $v$  is not empty. That is, the linear data  $Nu = v$  and the bound  $\|u\|_B \leq 1$  must be compatible. This compatibility condition is most easily stated by noting that if

$$\pi v = v - N^*(NN^*)^{-1}Nv \quad (2.2)$$

is the orthogonal projection of  $v$  onto the null space of  $N$ , then

$$q(v) \equiv v - \pi v = N^*(NN^*)^{-1}Nv \quad (2.3)$$

is the element of smallest norm in the plane  $\{v \mid Nv = v\}$ . Consequently, the set of  $v$  in (2.1) is not empty if and only if

$$\|q(v)\|^2 \leq 1 \quad (2.4)$$

This is a quadratic inequality to be satisfied by  $v$ .

If  $\|q(v)\| = 1$ , then the set of admissible  $v$  consists of the single point  $q(v)$ , and the optimal  $a$  is immediately found to be

$$a = (M^* M)^{-1} M^* S q(v) \quad (2.5)$$

If we define the orthogonal projection

$$Pw = w - M(M^* M)^{-1} M^* w \quad (2.6)$$

onto the orthogonal complement of the range of  $M$ , we see that the  $a$  defined by (2.5) can also be defined by the equation

$$Ma = Sq(v) - PSq(v) \quad (2.7)$$

The identity

$$\|Sv - \frac{1}{2} M(a_1 + a_2)\|^2 = \frac{1}{2} [\|Sv - Ma_1\|^2 + \|Sv - Ma_2\|^2] - \|\frac{1}{2} M(a_1 - a_2)\|^2$$

shows that the function  $[d(a,v)]^2$  is strictly convex in  $a$ . Since it approaches infinity as  $\|a\| \rightarrow \infty$ ,  $d(a,v)$  attains its minimum at a unique point  $\hat{a}$ , and we shall characterize this point.

Suppose for the moment that when  $a$  is optimal, the supremum in (2.1) is attained at a unique value  $r$  of  $v$ . Then the Euler equation which states that  $\|Sr - Ma\|$  is a maximum among admissible  $v$  and a minimum among admissible  $a$  gives the symmetric linear system

$$\begin{aligned} (\lambda - S^* S)r + S^* Ma + N^* b &= 0 \\ M^* Sr - M^* Ma &= 0 \\ Nr &= v \end{aligned} \quad (2.8)$$

and, if  $\lambda \neq 0$ , the condition

$$\|r\| = 1 \quad (2.9)$$

which states that the constraint  $\|v\| \leq 1$  is the active.

The system (2.8) can be written in the form

$$\begin{aligned} (\lambda I - \Pi S^* P S \Pi) \Pi r &= \Pi S^* P S q(v) \\ (I - \Pi) r &= q(v) . \end{aligned} \quad (2.10)$$

Since  $\|S \Pi\| = \kappa$ ,  $\|\Pi S^* P S \Pi\| \leq \kappa^2$ .

Therefore when  $\lambda > \kappa^2$  the operator in (2.10) is invertible, so that the system (2.8) is uniquely solvable. We denote the solution by  $\{r(\lambda, v), a(\lambda, v), b(\lambda, v)\}$ . This solution is clearly linear in  $v$  and analytic in  $\lambda$  for  $\lambda > \kappa^2$ .

The following lemma describes the behavior of (2.8) when  $\lambda \geq \kappa^2$ .

**LEMMA 2.1** The norm  $\|r(\lambda, v)\|$  of the solutions of (2.8) for  $\lambda > \kappa^2$  is nonincreasing in  $\lambda$ . The system (2.8) for  $\lambda = \kappa^2$  has a solution if and only if  $\|r(\lambda, v)\|$  is uniformly bounded for  $\lambda > \kappa^2$ . If this is the case, then as  $\lambda$  decreases to  $\kappa^2$ , the family  $\{r(\lambda, v), a(\lambda, v), b(\lambda, v)\}$  converges strongly to that solution  $\{r(\kappa^2, v), a(\kappa^2, v), b(\kappa^2, v)\}$  for which  $\|r\|$  has the smallest value.

Proof. We choose  $\lambda_1 > \lambda_2 > \kappa^2$ , subtract the first equation of (2.8) for  $\lambda = \lambda_2$  from that for  $\lambda_1$  and take the scalar product with  $r(\lambda_1, v) - r(\lambda_2, v)$  to find that



$$\begin{aligned}
& \frac{1}{2}(\lambda_1 - \lambda_2) \{ \|r(\lambda_1, \nu)\|^2 - \|r(\lambda_2, \nu)\|^2 \} + \\
& + \frac{1}{2}(\lambda_1 + \lambda_2) \|r(\lambda_1, \nu) - r(\lambda_2, \nu)\|^2 - \\
& - \|S[r(\lambda_1, \nu) - r(\lambda_2, \nu)]\|^2 + \|M[a(\lambda_1, \nu) - a(\lambda_2, \nu)]\|^2 = 0 .
\end{aligned} \tag{2.11}$$

Since  $\frac{1}{2}(\lambda_1 + \lambda_2) > \kappa^2$ , we conclude that

$$\|r(\lambda_1, \nu)\|^2 \leq \|r(\lambda_2, \nu)\|^2 \quad \text{for } \lambda_1 > \lambda_2 > \kappa^2 . \tag{2.12}$$

Therefore as  $\lambda \searrow \kappa^2$ ,  $\|r(\lambda, \nu)\|$  either goes to  $+\infty$  or remains bounded.

In the latter case, there is a sequence  $\lambda_p$  decreasing to  $\kappa^2$  such that the sequence  $\{r(\lambda_p, \nu), a(\lambda_p, \nu), b(\lambda_p, \nu)\}$  converges weakly to a solution of (2.8) with  $\lambda = \kappa^2$ , which we denote by  $\{r(\kappa^2, \nu), a(\kappa^2, \nu), b(\kappa^2, \nu)\}$ .

The derivation of (2.12) is still valid when  $\lambda_2 = \kappa^2$ . Thus  $\|r(\kappa^2, \nu)\| \geq \liminf \|r(\lambda, \nu)\|$ , which implies that the convergence is strong, and that

$$\|r(\kappa^2, \nu)\| = \lim_{\lambda \searrow \kappa^2} \|r(\lambda, \nu)\| . \tag{2.13}$$

The inequality (2.12) with  $\lambda_2 = \kappa^2$  still applies when  $r(\kappa^2, \nu)$  is replaced by any solution of (2.8) with  $\lambda = \kappa^2$ . Consequently the limit solution  $\{r(\kappa^2, \nu), a(\kappa^2, \nu), b(\kappa^2, \nu)\}$  is the solution of (2.8) with  $\lambda = \kappa^2$  which has the smallest value of  $\|r\|$ . This solution is unique even if (2.8) has other solutions. Hence the whole family  $\{r(\lambda, \nu), a(\lambda, \nu), b(\lambda, \nu)\}$  converges to  $\{r(\kappa^2, \nu), a(\kappa^2, \nu), b(\kappa^2, \nu)\}$  as  $\lambda$  decreases to  $\kappa^2$ .

The inequality (2.12) with  $\lambda_2 = \kappa^2$  and  $r(\lambda_2, \nu)$  replaced by any solution of (2.8) also shows that if there is a solution of (2.8), then  $\|r(\lambda, \nu)\|$  for  $\lambda > \kappa^2$  is bounded. Thus the Lemma is proved.

## 3. The main theorem.

We are now ready to state and prove our principal result.

**THEOREM 3.1** Suppose that  $\|q(v)\| < 1$  and let  $\kappa^2$  be the norm of the restriction of  $S$  to the null space of  $N$ .

If the system

$$\begin{aligned} \kappa^2 r - S^* S r + S^* M a + N^* b &= 0 \\ M^* S r - M^* M a &= 0 \\ N r &= v \end{aligned} \tag{3.1}$$

has a solution with  $\|r\| \leq 1$ , define  $\lambda(v) = \kappa^2$ . If this system has no solution with  $\|r\| \leq 1$ , then there is a unique value  $\lambda(v)$  of  $\lambda > \kappa^2$  such that the solution of the system (2.8) satisfies  $\|r(\lambda(v), v)\| = k$ .

In either case the optimal approximation is given by  $\hat{a}(v) = a(\lambda(v), v)$  where  $\{r(\lambda, v), a(\lambda, v), b(\lambda, v)\}$  is the solution of (2.8), and the corresponding optimal error bound is

$$d(a(\lambda(v), v); v) = \lambda(v) + v \cdot b(\lambda(v)). \tag{3.2}$$

Proof: Suppose that for some  $\lambda \geq \kappa^2$  the solution  $\{r, a, b\}$  of the first and third equations in (2.8) has  $\|r\| = 1$ . Then if  $Nv = v$  we have

$$\begin{aligned} \|Sv - Ma\|^2 &= \|S(v - r)\|^2 + 2(S(v - r), Sr - Ma) + \|Sr - Ma\|^2 \\ &\leq \kappa^2 \|v - r\|^2 + 2(v - r, \lambda r + N^* b) + \|Sr - Ma\|^2 \\ &\leq \lambda \|v - r\|^2 + 2\lambda(v - r, r) + \|Sr - Ma\|^2 \\ &= \lambda(\|v\|^2 - 1) + \|Sr - Ma\|^2. \end{aligned}$$

We see that the maximum of this expression on the set  $\{v | Nv = v, \|v\| \leq 1\}$  is attained at  $v = r$ , so that if  $d(a)$  is defined by (2.1),

$$d(a; v) = \|Sr - Ma\| . \quad (3.3)$$

Suppose first that for some admissible  $v$  the system (3.1) has a solution  $\{r(\kappa^2, v), a(\kappa^2, v), b(\kappa^2, v)\}$  with  $\|r(\kappa^2, v)\| \leq 1$ . Also suppose that  $\kappa^2$  is the largest eigenvalue of the projection  $\Pi S^* S \Pi$  of  $S^* S$  on the null space of  $N$ . Then there is a solution  $\{\hat{r}, \hat{b}\}$  of the eigenvalue equation

$$\begin{aligned} \kappa^2 \hat{r} - S^* S \hat{r} + N^* \hat{b} &= 0 \\ N \hat{r} &= 0 , \end{aligned}$$

with  $r \neq 0$ . For any constant  $\alpha$ ,  $\{r(\kappa^2, v) + \alpha \hat{r}, a(\kappa^2, v), b(\kappa^2, v) + \alpha \hat{b}\}$  is a solution of the first and third equations of (2.8). If  $\|r(\kappa^2, v)\| < 1$ , the quadratic equation

$$\|r(\kappa^2, v) + \alpha \hat{r}\|^2 = 1$$

has two solutions  $\alpha_+$  and  $\alpha_-$  with  $\alpha_+ \alpha_- < 0$ . If  $\|r(\kappa^2, v)\| = 1$ , we set  $\alpha_+ = \alpha_- = 0$ .

The formula (3.3) now shows that

$$d(a(\kappa^2, v)) = \|S(r(\kappa^2, v) + \alpha_{\pm} \hat{r}) - Ma(\kappa^2, v)\| .$$

Moreover, for any  $m$ -vector  $c$

$$\begin{aligned} d(a(\kappa^2, v) + c)^2 &\geq \|S(r(\kappa^2, v) + \alpha_{\pm} \hat{r}) - M(a(\kappa^2, v) + c)\|^2 = \\ &= d(a(\kappa^2, v))^2 + 2(S(r + \alpha_{\pm} \hat{r}) - Ma, Mc) + \|Mc\|^2 \end{aligned}$$

$$= d(a(\kappa^2, \nu))^2 + 2\alpha_{\pm}(S\hat{r}, Mc) + \|Mc\|^2$$

because of the second equation in (2.8). Because  $\alpha_{+}\alpha_{-} \leq 0$ , at least one of the quantities  $\alpha_{\pm}(S\hat{r}, Mc)$  is nonnegative, so that  $d(a(\kappa^2, \nu) + c) > d(a(\kappa^2, \nu))$  for  $c \neq 0$ . We conclude that when  $\|r(\kappa^2, \nu)\| \leq 1$ , the vector  $a(\kappa^2, \nu)$  gives the optimal approximation, as the Theorem states.

We take the scalar product of  $r + 2\alpha_{\pm}r$  with the first equation of (3.1) to find the corresponding optimal error bound

$$\begin{aligned} d(a(\kappa^2, \nu))^2 &= \|S(r + \alpha_{\pm}r) - Ma\|^2 \\ &= \kappa^2 + \nu \cdot b(\kappa^2, \nu), \end{aligned}$$

which is (3.2) for this case.

If  $S^*S$  is not compact,  $\kappa^2$  may not be an eigenvalue. In this case for each  $\delta > 0$  there are elements  $\hat{r}_{\delta}$  and  $\hat{b}_{\delta}$  such that

$$\begin{aligned} \|\hat{r}_{\delta}\| &= 1, \\ N\hat{r}_{\delta} &= 0, \\ \|\kappa^2 \hat{r}_{\delta} - S^*S\hat{r}_{\delta} + N^*\hat{b}_{\delta}\| &< \delta. \end{aligned}$$

Computations like the above show that

$$d(a(\kappa^2, \nu))^2 \leq \|S(r + \alpha_{\pm}(\delta)r_{\delta}) - Ma\|^2 + 4\delta.$$

and

$$d(a+c)^2 \geq \|S(r + \alpha_{\pm}(\delta)r_{\delta}) - Ma\|^2 + \|Mc\|^2.$$

By letting  $\delta \rightarrow 0$ , we again find that  $a(\kappa^2, \nu)$  is optimal and that the error bound (3.2) is valid.

The above arguments with (3.1) replaced by (2.8),  $\kappa^2$  by  $\lambda > \kappa^2$ , and  $\alpha_+$  by 0 show that if  $\|r(\lambda, \nu)\| = 1$ , then the optimal approximation is given by  $a(\lambda, \nu)$  and the optimal error is given by (3.2).

It remains to be shown that if there is exactly one such  $\lambda$  if either (3.1) has no solution or if  $\|r(\kappa^2, \nu)\| > 1$ . By Lemma 2.1 this condition is equivalent to the condition  $\|r(\kappa^2 + 0, \nu)\| > \kappa$ .

In terms of the projections  $\Pi$  and  $P$  defined by (2.2) and (2.6), respectively, the system (2.8) can be written as

$$\lambda \Pi r - \Pi S^* P S \Pi r = \Pi S^* P S q(\nu). \quad (3.4)$$

Clearly  $\|\Pi r\|$  approaches zero as  $\lambda$  increases to infinity. Since  $\|r\|^2 = \|\Pi r\|^2 + \|q(\nu)\|^2$ , it follows that

$$\lim_{\lambda \rightarrow \infty} \|r(\lambda, \nu)\| = \|q(\nu)\| < 1.$$

Thus if  $\|r(\kappa^2 + 0, \nu)\| > 1$ , continuity shows that there is at least one value of  $\lambda < \kappa^2$  for which  $\|r(\lambda, \nu)\| = 1$  so that  $a(\lambda, \nu)$  is optimal.

To prove that this value of  $\lambda$  is unique we recall that  $\|r(\lambda, \nu)\|$  is nonincreasing by (2.12), which follows from (2.11). We see from (2.11) that equality for some  $\lambda_1 > \lambda_2$  holds if and only if  $r(\lambda_1, \nu) = r(\lambda_2, \nu)$  and  $a(\lambda_1, \nu) = a(\lambda_2, \nu)$ . The difference of the equations for  $\lambda = \lambda_1$ , and  $\lambda = \lambda_2$  then shows that  $r(\lambda_1, \nu) = r(\lambda_2, \nu) = q(\nu)$ . Then  $S^* P S q(\nu)$  lies in the range of  $N^*$  and it follows that  $r(\lambda, \nu) = q(\nu)$  for all  $\lambda$ . In

particular,  $\|r(\kappa^2, \nu)\| = \|q(\nu)\| < 1$ . We conclude that if  $\|r(\kappa^2 + 0, \nu)\| > 1$ ,  $\|r(\lambda, \nu)\|$  is strictly increasing so that the value of  $\lambda$  where  $\|r(\lambda, \nu)\| = 1$  is unique. We call this value  $\lambda(\nu)$ .

We have shown that when  $\|r(\kappa^2, \nu)\| > 1$ , the optimal approximation is again given by  $a(\lambda(\nu), \nu)$  and the optimal error bound by (3.2). Thus Theorem 3.1 is proved.

#### 4. Properties of the optimal approximation operator

In this section we examine the question of whether the optimal approximation  $\hat{a}(v)$  is given by a linear transformation on the data. When  $\|r(\kappa^2, v)\| \leq 1$ ,  $a$  is given by solving (3.1), so that it does depend linearly on  $v$ .

On the other hand, when  $\|r(\kappa^2, v)\| > 1$ ,  $\hat{a}(v) = a(\lambda(v), v)$ . Since  $\lambda(v)$  depends upon  $v$ , one cannot always expect this  $\hat{a}$  to be a linear function of  $v$ . However, the following example shows that the optimal approximation  $\hat{a}(v) = a(\lambda(v), v)$  may be linear in  $v$  for all  $v$  even though  $\lambda(v)$  varies. We present an extreme case where  $\lambda(v) > \kappa^2$  for all  $v \neq 0$ .

EXAMPLE 4.1 If

$$S = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}, M = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, N = (0 \ 1),$$
 we find that  $\kappa = 1$ , and the

solution of (2.8) is

$$r = \begin{pmatrix} 2v_1/(\lambda - 1) \\ v_1 \end{pmatrix}, \quad a = 3v_1, \quad b = \frac{-\lambda^2 + 5\lambda}{\lambda - 1} v_1,$$

so that

$$\lambda(v) = 1 + 2|v_1|(1 - v_1^2)^{-1/2}.$$

Note that  $a(\lambda, v)$  does not depend on  $\lambda$  in this example. Hence  $\hat{a} = 3v_1$ , which is linear in  $v_1$  for all admissible  $v_1$  (that is, for  $|v_1| \leq 1$ ), even though  $\lambda(v)$  varies with  $v_1$ . This states that  $Su$  is to be approximated by

$$3MNu = \begin{pmatrix} 0 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} .$$

The following theorem gives a necessary and sufficient condition for the optimal approximation  $\hat{a}(v)$  to be linear for all admissible  $v$ . We recall the definitions (2.2) and (2.6) of the projections  $\Pi$  onto the null space of  $N$  and  $P$  onto the orthogonal complement of the range of  $M$ .

**THEOREM 4.1** The mapping  $v \rightarrow \hat{a}(v)$  to the optimal approximation is linear if and only if there is a linear subspace  $J$  of the null space of  $N$  which is invariant under the operator  $S^*PS$  and which contains the range of  $\Pi S^*M$ .

Proof: Suppose that

$$\hat{a}(v) = Lv$$

where  $L$  is a linear transformation. If  $r(\lambda, v) \neq q(v)$  for some  $\lambda \geq \kappa^2$ , then  $\lambda = \lambda(v/k)$  if we define  $k = \|r(\lambda, v)\|$ . Since  $\hat{a}(v/k) = a(\lambda(v/k), v) = a(\lambda, v)/k$ , we have

$$a(\lambda, v) = Lv$$

To evaluate  $Lv$  we recall the second equation of (2.8) and the fact that  $r(\infty, v) = q(v) = (I - \Pi)r(\lambda, v)$ .

Thus we find that  $M^*Ma(\lambda, v) = M^*Sq(v)$ , so that

$$M^*S\Pi r(\lambda, v) = 0 . \tag{4.1}$$



If  $r(\lambda, \nu) = q(\nu)$ , this equation is still valid. Thus the linearity of  $\hat{a}(\nu)$  implies (4.1). Conversely, if (4.1) is valid,

$$\hat{a}(\nu) = (M^* M)^{-1} M^* S q(\nu) \quad (4.2)$$

which is linear in  $\nu$ .

We write the first two equations of (2.8) in the form  $\lambda r - S^* P S r = -N^* b$  to see that (4.1) is equivalent to

$$M^* S \Pi (\lambda I - S^* P S)^{-1} N^* b = 0.$$

If one prescribes  $b$  and solves the first two equations of (2.8), the third equation determines  $\nu$ . Therefore the mapping  $\nu \rightarrow b(\lambda, \nu)$  is onto  $E_n$ . Consequently,  $\hat{a}(\nu, k)$  is linear if and only if

$$M^* S \Pi (\lambda I - S^* P S)^{-1} N^* = 0 \quad \text{for } \lambda > \kappa^2. \quad (4.3)$$

We take the adjoint of this equation and expand the inverse in a power series in  $\lambda^{-1}$  to obtain the statement of the theorem.

We remark that the direct sum  $K$  of the range of  $M$  and  $SJ$  is an invariant space of  $S \Pi S^*$  which contains the range of  $M$  and lies in the null space of  $N S^* P$ , and that  $J = S^* \Pi K$ . Thus the existence of the invariant subspace  $K$  is also equivalent to the linearity of  $\hat{a}$ .

The infinite set of equations  $N(S^* P S)^{\ell} \Pi S^* M = 0$ ,  $\ell = 1, 2, \dots$  which is equivalent to the existence of  $J$ , is not likely to be satisfied.

There are some simpler sufficient conditions. For example, the condition  $N S^* P S \Pi = 0$ , which states that the null space of  $N$  is an invariant subspace of  $S^* P S$ , is equivalent to  $r(\lambda, \nu) = q(\nu)$  for all  $\nu$ .

The condition  $\text{PS}\Pi^* M = 0$ , which states that the range of  $M$  is an invariant set of  $\text{S}\Pi\text{S}^*$ , is satisfied when the interpolating functions  $m_i$  are eigenvectors of  $\text{S}\Pi\text{S}^*$ . This condition is satisfied in Example 4.1, in which the single interpolating vector lies in the null space of  $\text{S}\Pi\text{S}^*$ .

Theorem 4.1 shows that  $\hat{a}(k, v)$  is usually not linear in  $v$  for all admissible  $v$ .

On the other hand, as long as  $\|r(\kappa^2, v)\| \leq 1$  the mapping  $v \rightarrow \hat{a}(v)$  agrees with the linear transformation  $a(\kappa^2, v)$ .

Thus

$$\hat{a}(v) = a(\kappa^2, v) \quad \text{for } v \in Q \equiv \{v: \|r(\kappa^2, v)\| \leq 1\}.$$

We see from (3.4) that the set  $Q$  is a neighborhood of the origin if and only if the operator

$$(\lambda I - \Pi\text{S}^* \text{PS}\Pi)^{-1} \Pi\text{S}^* \text{PS}\Pi^* \tag{4.4}$$

remains bounded as  $\lambda$  decreases to  $\kappa^2$ . This is certainly true if  $\lambda$  is in the resolvent set of  $\Pi\text{S}^* \text{PS}\Pi$ . If  $S$  is compact, it is sufficient to assume that the null space of  $\kappa^2 I - \Pi\text{S}^* \text{PS}\Pi$  is contained in the null space of  $\kappa^2 I - \text{S}^* \text{PS}$ .

The set of vectors  $b$  on which the operator (4.4) remains bounded as  $\lambda \searrow \kappa^2$  is a linear space. Hence if the operator (4.4) is unbounded, the set of  $v$  such that  $\lambda(v/k) > \kappa^2$  is dense. Consequently, in this case  $\hat{a}$  is linear for  $|v|$  small only if the conditions of Theorem 4.1 are valid. We have proved the following theorem.

THEOREM 4.2. The optimal approximation operator  $v \rightarrow \hat{a}(v)$  coincides with a linear operator, namely (4.2), for all  $v$  in a neighborhood of the origin if and only if either the operator (4.4) is uniformly bounded for  $\lambda > \kappa^2$  or the conditions of Theorem 4.1 are satisfied.

## 5. Error bounds for non-optimal approximations.

In the derivation of (3.3) we have used the assumption that  $\|r\| = 1$ . If we drop this assumption, we obtain the inequality

$$d(a(\lambda, \nu))^2 \leq \|Sr - Ma\|^2 + \lambda(1 - \|r\|^2) \quad (5.1)$$

when  $\{r, a, b\}$  satisfies the first and third equations of (2.8).

Equality holds if and only if  $\|r\| = 1$ , but the error bound is valid more generally. If one assumes that the second equation of (2.8) is also satisfied, one obtains, as in the derivation of (3.2),

$$d(a(\lambda, \nu); \nu)^2 \leq \lambda + \nu \cdot b(\lambda, \nu) \quad (5.2)$$

for any  $\nu$  and  $\lambda \geq \kappa^2$ .

If we use the first equation of (2.8) and its derivative, we find that

$$\frac{d}{d\lambda} [\lambda + \nu \cdot b(\lambda, \nu)] = 1 - \|r(\lambda, \nu)\|^2.$$

Thus the right-hand side of (5.2) increases for  $\lambda > \lambda(\nu)$  and decreases for  $\lambda < \lambda(\nu)$ . Its minimum thus occurs, as it should, at the optimal value  $\lambda(\nu)$  where the inequality also becomes an equality.

This suggests a good strategy to use if  $\kappa$  is not known, but one has an upper bound  $\kappa'$  for it: If  $\|r(\kappa'^2, \nu)\| \leq 1$ , use the approximation  $a(\kappa'^2, \nu)$ . Otherwise find  $\lambda(\nu) > \kappa'^2$  and use the optimal approximation  $a(\lambda(\nu), \nu)$ .

Because  $M^*(Sr(\lambda, \nu) - Ma(\lambda, \nu)) = 0$ , we see that if  $\kappa^*$  is the norm (1.2) of the restriction of  $S^*$  to the null space of  $M^*$ ,

$$\begin{aligned}
\|Sr - Ma\|^2 &\geq \kappa^{*-2} \|S^*(Sr - Ma)\|^2 \\
&= \kappa^{*-2} \|\lambda r + N^*b\|^2 \\
&= \kappa^{*-2} [\lambda^2 \|r\|^2 + 2\lambda(v \cdot b) + \|N^*b\|^2] .
\end{aligned} \tag{5.3}$$

In deriving (5.2) we used the identity

$$\|Sr - Ma\|^2 = \lambda \|r\|^2 + v \cdot b .$$

Substituting this in (5.3), we see that

$$v \cdot b \leq - \frac{1}{2\lambda - \kappa^{*2}} [\lambda(\lambda - \kappa^{*2}) \|r\|^2 + \|N^*b\|^2] \tag{5.4}$$

provided  $2\lambda > \kappa^{*2}$ . We now see from (5.2) that

$$d(a(\lambda, v); v)^2 \leq \lambda \quad \text{for } \lambda \geq \mu^2$$

where

$$\mu = \max\{\kappa, \kappa^*\} . \tag{5.5}$$

In particular, we have the bound

$$d(a(\mu^2, v); v) \leq \mu . \tag{5.6}$$

This is the best bound which can be found by choosing an  $a$  which depends linearly on  $v$ . (Note that  $\mu$  is a fixed number.) The transformation  $v \rightarrow Ma(\mu^2, v)$  is, in fact, the center of the set of optimal operators of this kind obtained by Davis, Kahan, and the author [6].

We see from (5.4) that the error bound

$$d(a(\mu^2, \nu); \nu)^2 \leq \mu^2 + \nu \cdot b(\mu^2, \nu) \quad (5.7)$$

for approximation by the linear operator  $a(\mu^2, \nu)$  which comes from (5.2), is never worse than (5.6). It is, in fact, better unless  $\kappa^* \geq \kappa$  and  $\nu = \text{NS}^* w$  for some eigenvector  $w$  corresponding to the eigenvalue  $\kappa^{*2}$  of the orthogonal projection of  $SS^*$  onto the null space of  $M^*$ .

## 6. Examples.

In this section we shall indicate how our results can be used to approximate the solution of boundary value problems.

Consider the problem

$$\begin{aligned} Lw &\equiv -w'' + q(x)w = u && \text{in } (0,1) \\ w(0) &= w(1) = 0 \end{aligned} \tag{6.1}$$

where  $q$  is a given bounded nonnegative function, and define the solution operator

$$Su \equiv w . \tag{6.2}$$

For the sake of simplicity we assume that  $u(0) = u(1) = 0$ , and we suppose that we are given the values  $u(\xi_s)$  at  $0 < \xi_1 < \xi_2 < \dots < \xi_n < 1$ , and a bound

$$\int_0^1 u'^2 dx \leq 1 .$$

We define the Hilbert space  $B$  to be the closure of the set

$$\{v \in C^1 \mid v(0) = v(1) = 0, \int_0^1 v'^2 dx < \infty\}$$

in the norm

$$\|v\|_B^2 = \int_0^1 v'^2 dx .$$

Then

$$(Nv)_s = v(\xi_s)$$

and

$$N^* b = \sum_{s=1}^n G(x, \xi_s) b_s$$

where

$$G(x, \xi) = \frac{1}{2} (x + \xi - |x - \xi| - 2x\xi) .$$

We shall give two examples of approximation in different spaces  $\Sigma$  .

EXAMPLE 6.1. Let  $\Sigma$  be the closure of the set

$$\{w \in C^2, w(0) = w(1) = 0, \int_0^1 (Lw)^2 dx < \infty\}$$

in the norm

$$\|w\|_{\Sigma}^2 = \int_0^1 (Lw)^2 dx .$$

We choose  $m$  interpolating functions  $w_1, \dots, w_m$  which vanish at 0 and 1 and such that  $w_i''$  is square integrable. For example, we may choose splines which are  $C^1$  and piecewise  $C^2$  .

We observe that

$$(Su, h)_{\Sigma} = \int_0^1 u Lh dx = (u, S^* h)_B = - \int_0^1 u (S^* h)'' dx .$$

Thus  $S^* h$  is the solution of the problem

$$\begin{aligned} -(S^* h)'' &= Lh \\ S^* h(0) &= S^* h(1) = 0 . \end{aligned}$$

We note that when  $Nv = 0$  ,



$$\|Sv\|_{\Sigma}^2 = \int v^2 dx \leq \mu^2 \int v'^2 dx$$

where

$$\mu = \frac{1}{\pi} \max_{1 \leq s \leq n} |\xi_{s+1} - \xi_s|.$$

(We have put  $\xi_0 = 0$  and  $\xi_{n+1} = 1$ .) Equality holds when  $v$  is  $\sin \pi(x - \xi_s) / (\xi_{s+1} - \xi_s)$  on the longest interval and zero elsewhere, so that this  $\mu$  is the norm of  $S$  on the null space of  $N$ .

We differentiate the first equation of (2.8) twice to obtain the differential equation

$$\lambda r'' + r = LMa - (N^* b)'' \quad (6.3)$$

If we consider this equation on the whole interval  $(0,1)$ , we must treat  $(N^* b)''$  as a distribution. More importantly, the inverse of the differential operator on the left has poles at  $\lambda = 1/p^2 \pi^2$  for  $p=1,2,\dots$ , and some of these points lie above  $\mu^2$ . Consequently we prefer to use the equation  $Nr = v$  and to solve the equation (6.3) in each interval  $(\xi_s, \xi_{s+1})$ . Since  $(N^* b)'' = 0$  in these intervals, we find that

$$r(x) = \frac{v_s \sin(\xi_{s+1} - x) / \sqrt{\lambda} + v_{s+1} \sin(x - \xi_s) / \sqrt{\lambda}}{\sin(\xi_{s+1} - \xi_s) / \sqrt{\lambda}} \quad (6.4)$$

$$- \int_{\xi_s}^{\xi_{s+1}} \Gamma_s(x,y) \sum \text{Im}_j(y) a_j dy \quad \text{for } \xi_s \leq x \leq \xi_{s+1}$$

where

$$\Gamma_s(x,y) = \frac{\{\cos(x+y - \xi_s - \xi_{s+1})/\sqrt{\lambda} - \cos(|x-y| - \xi_{s+1} + \xi_s)/\sqrt{\lambda}\}}{2\sqrt{\lambda} \sin(\xi_{s+1} - \xi_s)/\sqrt{\lambda}} .$$

The second equation in (2.8) becomes

$$\int_0^1 [r(x) - \sum \text{Im}_i(x)a_i] \text{Im}_j(x) dx = 0 \quad \text{for } j=1, \dots, m .$$

We substitute (6.4) to obtain  $m$  equations for the  $m$ -vector  $a$ .

The solution (6.4) is well-defined for  $\lambda > \kappa^2$ . When  $\lambda = \kappa^2$  we must adjoin the orthogonality condition

$$v_s + v_{s+1} - \frac{1}{\pi} (\xi_{s+1} - \xi_s) \int_{\xi_s}^{\xi_{s+1}} \sin \pi(y - \xi_s) / (\xi_{s+1} - \xi_s) \sum \text{Im}_j(y) a_j dy = 0$$

on each interval of longest length. A linear combination of the eigenfunction corresponding to these longest intervals is added to the resulting limit as  $\lambda \rightarrow \kappa^2$  in (6.4) so that one still has a square matrix to invert.

EXAMPLE 6.2 We now suppose that  $q \in C^1$  and let  $\Sigma$  be closure of

$$\{w | w \in C^3, w(0) = w(1) = Lw(0) = Lw(1) = 0, \int_0^1 (Lw)^2 dx < \infty\}$$

in the norm

$$\|w\|_{\Sigma}^2 = \int_0^1 (Lw)^2 dx .$$

Then  $S$  is an isometry from  $B$  to  $\Sigma$ , so that  $S^*S = I$ . Hence  $S^*h = Lh$  and  $\kappa = \kappa^* = 1$ .

We take interpolating functions  $m_i$  which are in  $C^2$  with  $m_i = m_i'' = 0$  at the ends. If splines are used, care must be taken to satisfy this boundary condition. The first equation in (2.8) becomes

$$(\lambda - 1)r + LMa + N^*b = 0 \quad (6.5)$$

If we take  $\lambda = \kappa^2 = 1$ , and assume that the functions  $l_{m_i}$  and  $n_j$  are linearly independent, the equation states that  $a = b = 0$ . The other two equations then are

$$M^*Sr = 0 ,$$

$$Nr = v .$$

The solution of this system with the smallest value of  $\|r\|$  is obtained by taking  $r$  of the form

$$r = S^*Ma' + N^*b' .$$

The first of the equations then states that  $Ma' = -(I - P)SN^*b'$  so that  $r = S^*PSN^*b'$ . The second equation becomes

$$NS^*PSN^*b' = v ,$$

$$\text{and } \|r(\kappa^2, v)\|^2 = b' \cdot NS^*PSN^*b' .$$

Whether this is larger or smaller than 1 depends on the interpolating functions  $m_i$ . If all the functions  $l_{m_i}$  vanish at the points  $\xi_s$ , then  $NS^*P = NS^*$ , so that  $N^*Nb' = v$ , and  $r = q(v)$ . In this case  $\|r(\kappa^2, v)\| = \|q(v)\| < 1$ , so that the optimal approximation is  $a = 0$  with the trivial optimal error bound  $\|Su\|^2 \leq 1$ .

On the other hand, if the functions  $w_i$  are chosen so that each component  $G(\xi_s, x)$  of  $N^*$  is so well approximated by a vector of the form  $Lmc$  that

$$\|PSN^*\| < |v| ,$$

then by Schwarz's inequality

$$\begin{aligned} |v|^2 &= v \cdot NS^* PSN^* b, \\ &\leq \|PSN^* v\| \|r\| \\ &< |v|^2 \|r\| \quad . \end{aligned}$$

That is,  $\|r(\kappa^2, v)\| > 1$ . In this case, then, one obtains a better bound by taking  $\lambda > 1$ .

When  $\lambda > 1$ , we solve (6.5) for  $r$  and substitute in the second equation to find that

$$r = -\frac{1}{\lambda} S^* (I - P) SN^* b - \frac{1}{\lambda - 1} S^* PSN^* b$$

so that the conditions  $Nr = v$  and  $\|r\| = k$  become

$$\begin{aligned} [(\lambda - 1) NS^* (I - P) SN^* + \lambda NS^* PSN^*] b &= -\lambda(\lambda - 1)v \\ b \cdot [(\lambda - 1)^2 NS^* (I - P) SN^* + \lambda^2 NS^* PSN^*] b &= \lambda^2 (\lambda - 1)^2 k^2 \quad . \end{aligned}$$

The above examples can be carried over to higher dimensions. In order to keep the pointwise evaluation operator  $N$  bounded, one must use a norm with higher derivatives in  $B$ .

If one wishes to use a norm with derivatives higher than one in  $\Sigma$  one must find elements  $m_i$  which satisfy  $Lw_i = 0$  as well as  $w_i = 0$  on the boundary, which is usually difficult.

## Bibliography

1. G. Arsene and A. Gheondea. Completing matrix contractions. *J. Operator Theory* 7(1982), pp. 179-189.
2. J.P. Aubin. Best approximation of linear operators. *S.I.A.M. J. on Numer. Anal.* 5(1968), pp. 518-521.
3. I. Babuska, M. Prager, and E. Vitasek. *Numerical Processes in Differential Equations*. Interscience, New York, 1966.
4. C. Davis. An extremal problem for extensions of a sesquilinear form. *Linear Alg. and Appl.* 13(1976), pp. 91-102.
5. C. Davis. Some dilation and representation theorems. Preprint.
6. C. Davis, W.M. Kahan, and H.F. Weinberger. Norm-preserving dilations and their applications to optimal error bounds. *S.I.A.M. J. on Numer. Anal.* 19(1982), pp. 445-469.
7. M. Golomb and H.F. Weinberger. Optimal approximations and error bounds. In *On Numerical Approximation*, ed. R. Langer, U. of Wisconsin Press, Madison, 1959, pp. 117-190.
8. M.G. Krein. The theory of self-adjoint extensions of semi-bounded Hermitian transformations and its applications. *Math. Sb.* 20(1947), pp. 431-495 and 21(1947), pp. 365-404.
9. C.A. Micchelli and W.L. Miranker. Optimal difference schemes for linear initial value problems. *SIAM J. on Num. Anal.* 10(1973), pp. 983-1009.
10. C.A. Micchelli and T.J. Rivlin. A survey of optimal recovery. In *Optimal Estimation in Approximation Theory*, ed. C.A. Micchelli and T.J. Rivlin, Plenum, New York, 1977.
11. W.L. Miranker. Galerkin approximations and the optimization of difference for boundary value problems. *S.I.A.M. J. on Numer. Anal.* 8(1971), pp. 486-496.
12. S. Parrott. On a quotient norm and the Sz.-Nagy - Foias lifting theorem. *J. Func. Anal.* 30(1978), pp. 311-328.
13. Ju. L. Smuljan and R.N. Janovskaja. Blocks of a contractive operator-matrix. *Izv. Vyss. Ucebn. Zaved. Mat.* (1981), #7, pp. 72-75.
14. H.F. Weinberger. On optimal numerical solution of partial differential equations. *S.I.A.M. J. on Numer. Anal.* 9(1972), pp. 182-198.