

The Effective Choice of the Smoothing Norm in Regularization*

By Jane Cullum**

Abstract. We consider ill-posed problems of the form

$$(1) \quad g(t) = \int_0^1 K(t, s)f(s) ds, \quad 0 \leq t \leq 1,$$

where g and K are given, and we must compute f . The Tikhonov regularization procedure replaces (1) by a one-parameter family of minimization problems—Minimize $(\|Kf - g\|^2 + \alpha\Omega(f))$ —where Ω is a smoothing norm chosen by the user. We demonstrate by example that the choice of Ω is not simply a matter of convenience. We then show how this choice affects the convergence rate, and the condition of the problems generated by the regularization. An appropriate choice for Ω depends upon the character of the compactness of K and upon the smoothness of the desired solution.

1. Introduction. We consider ill-posed problems of the following type. Given a kernel K and a function g , find a function f such that

$$(1) \quad \int_0^1 K(t, s)f(s) ds = g(t), \quad 0 \leq t \leq 1.$$

In (1) we assume K is square-integrable and its null space is $\{0\}$. For example, if

$$(2) \quad \begin{aligned} K(t, s) &= 1, & t > s, \\ &= 0, & t < s, \end{aligned}$$

then the solution of (1) is the derivative of g . We will also write (1) as $Kf = g$.

From (1), using quadrature rules, we can generate—in general rectangular but typically square—linear algebraic systems

$$(3) \quad Ax = b,$$

where A has full rank and whose solutions approximate the solution of (1).

A problem is ill-posed in the Hadamard sense, if its solution does not depend continuously upon the data. Equations of type 1 are ill-posed. Note, however, that

Received May 16, 1977; revised April 17, 1978.

AMS (MOS) subject classifications (1970). Primary 45L05, 65D25, 65F35.

* This research was sponsored by the Air Force Office of Scientific Research (AFSC), United States Air Force, under contract F44620-76-C-0022.

** The material in this paper was presented at the 1976 Fall SIAM meeting, Atlanta, Georgia, October 16–18, 1976.

© 1979 American Mathematical Society
0025-5718/79/0000-0010/\$06.50

the associated algebraic systems (3) are well-posed, since A has full rank. However, for numerical computations problems must not only be well-posed, but also well-conditioned. Small variations in the problem data must not yield large variations in the solution. If we obtain the problem in (3) by applying standard quadrature rules to (1), it is easy to demonstrate that we obtain ill-conditioned algebraic systems. This happens because the operator in (1) is compact; i.e., its singular values have a limit point at 0. Recall that each such operator has a singular value decomposition Smithies [1],

$$(4) \quad K = \sum_{n=1}^{\infty} \sigma_n v_n u_n^T.$$

The functions u_n and v_n , $n = 1, 2, \dots$, in (4) are, respectively, eigenvector systems for the symmetric operators K^*K and KK^* . The σ_n^2 are the eigenvalues of K^*K and the σ_n are called the singular values of K . Equality in (4) is in the L_2 sense.

Since the matrices A approximate K and $\sigma_n \downarrow 0$, the singular values of A will also converge to zero as the approximation to K is refined. Thus, A will become ill-conditioned. We note that for crude approximations, A may be fairly well-conditioned.

Our discussion focuses on the Tikhonov regularization procedure, Tikhonov [2].

Definition 1. We say a one-parameter family of problems, denoted by $P(\alpha)$ for $0 \leq \alpha \leq 1$, is a regularizing family for a problem P if

- (a) For $0 \leq \alpha \leq 1$, each $P(\alpha)$ is well-posed; and
- (b) The solution f_α of $P(\alpha)$ converges to the solution f_0 of the original problem as $\alpha \downarrow 0$.

Convergence of the solutions is measured in some relevant norm. For numerical work we want at least pointwise convergence.

For an equation of type (1), Tikhonov [2] introduces the following regularizing family

$$(5) \quad P(\alpha) : \text{Minimize } \|Kf - g\|^2 + \alpha\Omega(f).$$

In (5), $\|\cdot\|$ is typically an L_2 -norm, measuring the residual error in solving (1), and $\Omega(f)$ is a smoothing norm which, in practice, is often set equal to

$$(6) \quad \Omega_1(f) = \int_0^1 f^2 + \int_0^1 (f^{(1)})^2.$$

We use $f^{(j)}$ to denote the j th derivative of f .

Regularization requires the assumption that the desired solution is smooth (has several derivatives), and that any errors in the data supplied are random and not smooth. We want the regularization procedure to generate an approximation to our given problem that (a) is better conditioned, (b) mollifies the effects of any random noise in our data, and (c) gives us a physically meaningful solution that approximates the true solution in some reasonable sense.

We cannot achieve these objectives by simply minimizing the residual error. The residual provides no measure of the lack of smoothness of a function f , since the compactness of K may smooth oscillations in f . The residual controls only the low

frequency components of f . Control over the high frequency components is exercised by the smoothing norm Ω in (5). $\Omega(f)$ will be small, only if f is smooth. Thus, the minimization in (5) forces the solutions f_α of the P_α to be smooth. The selection of α in (5) determines the tradeoff between these two forms of control. We refer the reader to Wahba [19] for a discussion of the choice of α .

Much of the existing literature on regularization focuses on constructing regularizing families and proving that the required convergence occurs. In fact, the implication in some papers (see Chechkin [3] for example) is that the choice of Ω is a matter of convenience, and does not critically affect the behavior of the approximations. We will argue, however, that the choice of Ω is important even for smooth kernels because numerical work demands more than simple convergence.

Regularization is normally applied to problems where the data is determined experimentally and, thus, contains noise or errors. We have, for any $\alpha \geq 0$, the inequality

$$(7) \quad \|f(\alpha, \epsilon) - f_0\| \leq \|f(\alpha, \epsilon) - f(\alpha, 0)\| + \|f(\alpha, 0) - f_0\|,$$

where f_0 denotes the desired solution, $f(\alpha, 0)$ denotes the solution of (5) with no error in g , and $f(\alpha, \epsilon)$ denotes the solution of (5) with error ϵ in g . We note that we will use ϵ to denote both an error function and the norm of this error function. Clearly, the size of the first term on the right-hand side of (7) is controlled by the condition of $P(\alpha)$ and the size of the error ϵ . If $\epsilon \neq 0$, as $\alpha \downarrow 0$, this first term will get arbitrarily large. The size of the second term depends upon the rate of convergence of the approximation $P(\alpha)$ to the original problem P , as $\alpha \downarrow 0$. Inequality (7) clearly demonstrates that there are two forces at work. To obtain a reasonable approximation, each $P(\alpha)$ must be well-conditioned, and we must achieve a reasonable rate of convergence.

We will demonstrate, for certain classes of operators and solutions, how the choice of Ω affects both the condition and the rate of convergence. Its effect upon the condition is usually, however, of primary importance.

In Section 2 we extend some results of Franklin [4]. He considered three problems: (a) the analytic continuation of a harmonic function that is known on a circle of radius $R < 1$ to the unit disk; (b) the backward heat equation in one dimension; and (c) the differentiation of a smooth function. He showed that the regularizing families obtained for problems (a) and (b) using Ω_1 in (6) are poor approximations to the original problem when there is noise in the data. However, the corresponding family obtained for problem (c), differentiation, has reasonable convergence properties. He did not elaborate. In Section 4 we will explain these results for problems (a) and (b) in terms of the lack of sensitivity of the condition of the members of the associated regularizing families to simple order changes in the smoothing norm Ω . This lack of sensitivity is due to the exponential decay rate of the singular values of the associated operators. First, however, in Section 2 we introduce two norms that yield regularizing families that are 'good' approximations to problems (a) and (b).

In Section 3, we consider in detail the problem of differentiating a function and derive relationships between the rate of convergence, the order of the smoothing norm Ω and the smoothness of the desired solution. Furthermore, we demonstrate that the condition of the associated regularizing approximations improves as the order of Ω is increased, whereas the corresponding rate of convergence deteriorates. However, the overall approximation given by (7) can improve. The results derived in Section 3 apply to a family of operators that includes differentiation.

In Section 4 we extend the results obtained for differentiation to more general convolution equations. The results in Section 4 are split into five theorems. The first theorem yields majorant estimates on the rate of convergence of Tikhonov regularizing families of order p for a large class of convolution operators. The second theorem yields minorant estimates on these rates for a subclass of the operators considered in Theorem 1. Together these two theorems yield the exact order of the rate of convergence of the regularizing families for this subclass. The third and fourth theorems in Section 4, again for subclasses, relate the order of Ω to the condition of the approximating problems generated by the regularization. Together these theorems yield, for a subclass of the operators considered, an overall estimate of the error in the approximations. We conclude Section 4 with a theorem that tells us that if our operator K is symmetric and positive definite, then there is a better-conditioned regularizing family which can be used in place of (5).

Sections 3 and 4 use Fourier transform analysis. The arguments are similar to those given in Aref'eva [5]. We will discuss this paper as well as the earlier papers Arsenin and Ivanov [6], [7], Arsenin and Savelova [8], Goncharkii, Leonov, and Yagola [9] in more detail in Section 4.

In Section 5 we make some remarks about several ways that a regularization can be implemented numerically using the associated Euler equation,

$$(8) \quad (K^*K + \alpha B)f = K^*g,$$

where B is the differential operator obtained from Ω by variational arguments.

The discussion is not chronological since some of the Russian work predates Franklin [4]. The results described give, for convolution equations, heuristics that can be used in estimating the appropriateness of a proposed regularization.

2. **Franklin [4].** In this section we examine the two classical ill-posed problems considered by Franklin [4]:

- (a) The analytic continuation of a harmonic function in the unit circle (see Eq. (17)) and
- (b) The solution of the backwards heat equation in 1-dimension.

THEOREM 1 (FRANKLIN [4]). *The regularizing families for problems (a) and (b) above, obtained using the norm in (6), and assuming $d_1\epsilon^2 \leq \alpha \leq d_2\epsilon^2$, satisfy the following estimates of the error in the approximation of the desired solution f_0 by the solution $f(\alpha, \epsilon)$ of $P(\alpha)$ with $g_\epsilon = g_0 + \epsilon$,*

$$(9) \quad (a) \quad \|f(\alpha, \epsilon) - f_0\| \leq [C_1/(-\ln \alpha)],$$

$$(10) \quad (b) \quad \|f(\alpha, \epsilon) - f_0\| \leq C_2 [T/(-\ln \alpha)]^{p/2}.$$

In (10), $[0, T]$ is the interval of interest for the heat equation, and this result was obtained with

$$(11) \quad \Omega_p(f) = \sum_{j=0}^p m_j \|f^{(j)}\|^2, \quad p = 1, 2, \dots \text{ and } m_j \geq 0.$$

$\|\cdot\|$ denotes the L_2 -norm.

We note that the setting in Franklin [4] is different from what we will use in later sections. Franklin considered varying α and ϵ simultaneously. He asked the question: If $\|Ku\| \leq \epsilon$ and $\Omega(u) \leq 1$, then how does the $\|u\|$ behave as $\epsilon \downarrow 0$? He used this to define a modulus of regularization. He proved that the ordinary modulus of convergence can be bounded by multiples of the modulus of regularization. Thus, estimating the modulus of regularization yields a corresponding estimate on the modulus of convergence as ϵ and α decrease simultaneously to zero. The estimates in (9) and (10) were obtained using the modulus of regularization. Franklin [4] used Fourier series expansions of the solutions of problems (a) and (b). In this section we will use similar arguments to derive error estimates for other norms. The class of problems considered in Section 4 will include these two problems. However, we also consider them separately here, because through Franklin's arguments [4] we can see explicitly the control exercised by the residual term in (5) on the low frequency components of the solution, and the control exercised by the smoothing norm Ω in (5) on the high frequency components of the solution. In Sections 3 and 4 we will see how the smoothing norm mollifies the condition of the regularizing approximations.

If we know for problem (a) that for the desired solution the quantity

$$(12) \quad \Omega_\infty^c(f) = \sum_{j=0}^{\infty} \|f^{(j)}\|^2 c^{2j}/(2j)!$$

exists and is finite for some $c > 0$, where the L_2 -norm is over the interval of interest, then we can use Ω_∞^c to define a regularizing family for problem (a). We will demonstrate that for $c \geq -2 \ln R$, the regularizing family generated by Ω_∞^c yields 'good' approximations to problem (a).

LEMMA 1. *Let f be C^∞ on $(-\infty, \infty)$ and assume $f^{(j)}(t) \rightarrow 0$ as $|t| \rightarrow \infty$ for $j = 0, 1, 2, \dots$. Assume also that the Fourier transform of f exists and satisfies*

$$(13) \quad |\hat{f}(\omega)| \leq M \exp(-\mu |\omega|)$$

for all $|\omega| > S$ and for some $M > 0$. Then, if $2\mu > c$, $\Omega_\infty^c(f)$ is finite.

Proof. By (13) the Fourier transform of each derivative of f exists and is in L_1 . By Parseval's equality,

$$(14) \quad \|f^{(j)}\|^2 \leq \|\omega^j \hat{f}\|^2,$$

where the ω -norm is on $(-\infty, \infty)$. Therefore, each term in (12) is majorized by

$$(15) \quad (\bar{M}/(2j)!) \int_0^\infty (c\omega)^{2j} \exp(-2\mu |\omega|) \leq \bar{M} (c/2\mu)^{2j}.$$

Combining (15) and (12), we obtain

$$\Omega_{\infty}^c(f) \leq \bar{M} \sum_{j=0}^{\infty} (c/2\mu)^{2j}.$$

But, this series is finite, whenever $2\mu > c$. Q.E.D.

THEOREM 2. *If for the solution f_0 of the analytic continuation problem $\Omega_{\infty}^c(f_0) < \infty$, then for the regularizing family defined in (5) using Ω_{∞}^c in (12), we obtain the following majorant estimate of the error as $\alpha \downarrow 0$,*

$$(16) \quad \|h_{\alpha}\|^2 \equiv \|f(\alpha, \epsilon) - f_0\|^2 \leq Q^* \alpha^{c/(c-2 \ln R)}.$$

We have assumed that $d_1 \epsilon^2 \leq \alpha \leq d_2 \epsilon^2$.

Proof. We repeat Franklin's analysis, using instead the Ω_{∞}^c -norm. We are given a harmonic function $u(r, \theta)$ in the unit disk with known values for some $R < 1$, $g_0 = u(R, \theta)$, and asked to determine its values $f_0 = u(1, \theta)$ on the unit circle. We must solve the integral equation of the first kind,

$$(17) \quad g_0 = (1/2\pi) \int_0^{2\pi} [(1 - R^2)/(1 - 2R \cos(\theta - \theta_1) + R^2)] f_0(\theta_1) d\theta_1.$$

We know the solution is periodic, so we write the error $h_{\alpha} = f(\alpha, \epsilon) - f_0$ as

$$h_{\alpha}(\theta) = \sum_{n=0}^{\infty} (A_n \cos n\theta + B_n \sin n\theta).$$

If $\|Kh_{\alpha}\|^2 \leq \alpha \bar{Q}$ and $\Omega_{\infty}^c(f_0) < Q$, we obtain, assuming that the interchanges of summation, integration, and differentiation are correct,

$$(18) \quad \Omega_{\infty}^c(h_{\alpha}) = 2\pi A_0^2 + \pi \sum_{n=1}^{\infty} \cosh(cn)(A_n^2 + B_n^2) \leq 2Q$$

and

$$(19) \quad \|Kh_{\alpha}\|^2 = 2\pi A_0^2 + \pi \sum_{n=1}^{\infty} R^{2n}(A_n^2 + B_n^2) \leq \alpha \bar{Q}.$$

Following Franklin, we balance the contribution from the low frequency terms in (19) against the contribution from the high frequency terms in (18). Let $N = N(\alpha)$ be the solution of the equation

$$(20) \quad \alpha \cosh cN = 2R^{2N}.$$

Since for $n \leq N$, $R^{2n} \geq R^{2N}$; and for $n > N$, $\cosh cn > \cosh cN$, we obtain

$$(21) \quad \|h_{\alpha}\|^2 \leq \|Kh_{\alpha}\|^2 R^{-2N} + \Omega_{\infty}^c(h_{\alpha}) (\cosh cN)^{-1} \leq 2\bar{Q} (\cosh cN)^{-1}.$$

Now consider the asymptotic behavior of N as a function of α . Clearly, as $\alpha \downarrow 0$, $N \uparrow \infty$, and

$$(22) \quad N \sim \ln(4/\alpha)/(c - 2 \ln R).$$

Substituting (22) into (21), we obtain (16). Q.E.D.

Comments. If $c = -2 \ln R$, then the exponent in (16) equals $\frac{1}{2}$. For fixed R , as we let $c \uparrow \infty$ this exponent approaches 1. As we decrease c , it decreases to 0. Note that the eigenvalues of K^*K decay like $\exp(2n \ln R)$ as $n \uparrow \infty$. Choosing a norm with a growth rate that matches or exceeds this decay rate, that is, $c \geq -2 \ln R$, yields a reasonable rate of convergence. To use Ω_∞^c we must know that for the desired solution $\Omega_\infty^c(f_0) < \infty$.

We can achieve a better rate of convergence if we use an even stronger norm.

$$(23) \quad \bar{\Omega}_\infty^c \equiv \sum_{j=0}^{\infty} c^j \|f^{(j)}\|^2 / j! \quad \text{for } c > 0.$$

Then we have the following: the proof is similar to that of Lemma 1.

LEMMA 2. *Let $f \in C^\infty$ on $(-\infty, \infty)$ and assume $f^{(j)}(t) \rightarrow 0$ as $|t| \rightarrow \infty$ for all $j = 0, 1, 2, \dots$. Assume also that the Fourier transform of f exists and satisfies*

$$(24) \quad |\hat{f}(\omega)| \leq M \exp(-\mu\omega^2)$$

for all $|\omega| > S$, for some $M > 0$. If $\mu > c/2$, then $\bar{\Omega}_\infty^c(f)$ is finite.

THEOREM 3. *If for the solution f_0 of the analytic continuation problem $\bar{\Omega}_\infty^c(f_0) < \infty$ for some $c > 0$, then for the regularizing family defined in (5) using $\bar{\Omega}_\infty^c$ in (23), we obtain the following majorant estimate of the error in the approximation as $\alpha, \epsilon \downarrow 0$ with $d_1 \epsilon^2 \leq \alpha \leq d_2 \epsilon^2$,*

$$(25) \quad \|h_\alpha\|^2 = \|f(\alpha, \epsilon) - f_0\|^2 \leq Q^* \alpha.$$

Proof. The proof proceeds as before with small changes. Equation (20) becomes

$$(26) \quad \alpha \exp(cN^2) = 2R^{2N}.$$

The asymptotic solution of this satisfies

$$(27) \quad N \sim (-\ln(\alpha/2)/c)^{1/2},$$

for a given R and sufficiently small α , or for a given α and sufficiently large R . Then using (27) and the analog of (21), we obtain (25). Q.E.D.

Comments. We note that to use Ω_∞^c or $\bar{\Omega}_\infty^c$ numerically we would have to use Fourier transforms, and we would have to know that our desired solution had sufficient smoothness to guarantee the existence of Ω_∞^c or $\bar{\Omega}_\infty^c$. We note that as $R \downarrow 0$, the eigenvalues R^n , $n = 1, 2, \dots$, of the operator K in (17) decrease more rapidly, and the error in the approximation increases. Thus, by picking a norm whose growth rate matches or exceeds the decay rate of the given operator, we can achieve 'good' orders of approximation.

If we apply the same analysis to the backwards heat equation, we obtain the following.

THEOREM 4. *If for the solution f_0 of the backwards heat equation in 1-dimension we have $\bar{\Omega}_\infty^c(f_0) < \infty$ for some $c > 0$, then for the regularizing family defined in (5)*

using $\bar{\Omega}_\infty^c$ in (23), we obtain the following majorant estimate of the error in the approximation,

$$(28) \quad \|h_\alpha\|^2 = \|f(\alpha, \epsilon) - f_0\|^2 \leq Q^* \alpha^{c/(c+2T)}.$$

The heat equation is considered on the interval $(0, T)$, and again we assume $d_1 \epsilon^2 \leq \alpha \leq d_2 \epsilon^2$.

Proof. The proof proceeds as before. In this case we obtain the equation

$$(29) \quad \alpha \exp(N^2(c + 2T)) = 2.$$

Taking logarithms of both sides, we obtain

$$N^2(c + 2T) = -\ln(\alpha/2).$$

Therefore,

$$N^2 = -\ln(\alpha/2)/(c + 2T)$$

and

$$\|h_\alpha\|^2 \leq \bar{Q} \exp(-cN^2) \leq \bar{Q}^* \alpha^{c/(c+2T)}. \quad \text{Q.E.D.}$$

THEOREM 5. *If we use Ω_∞^c , we obtain an estimate of the form*

$$(30) \quad \|h_\alpha\|^2 \leq Q^* \exp(-\bar{c}(-\ln \alpha)^{1/2}),$$

where \bar{c} involves c and T .

The difference in the convergence of the regularizing families for problems (a) and (b) can be explained by the difference in the decay rates of the eigenvalues of the two operators. For problem (a) the eigenvalues decay like R^n ; for problem (b) like $\exp(-n^2 T)$. Thus, the norm Ω_∞^c with a growth rate of $\exp(cn)$ cannot counteract the rapid decay of $\exp(-n^2 T)$. For problem (b) we must go to $\bar{\Omega}_\infty^c$, whose growth rate is $\exp(cn^2)$, to achieve an overall approximation that is some power of α .

The preceding analysis points out the need to have an estimate of the decay rate of the singular values of the given operator K as an aid in choosing an appropriate Ω . We are of course constrained in what order Ω we can use by the smoothness of the desired solution. In order for the convergence arguments to be valid, $\Omega(f_0)$ must be defined.

Arguments of the preceding type assumed α and ϵ varied simultaneously and, in fact, that $d_1 \epsilon^2 \leq \alpha \leq d_2 \epsilon^2$. Consequently, the effect of the choice of Ω on the rates of convergence and on the condition of the approximating problems cannot be obtained from such arguments. In Sections 3 and 4 we attempt to estimate directly rates of convergence, and the condition of members of the Tikhonov regularizing families for certain classes of convolution operators.

In the next section, we examine the problem of differentiation in detail. We will be using Fourier transforms so instead of Eq. (1) we use

$$(31) \quad \int_{-\infty}^t f(s) ds = g(t), \quad -\infty < t < \infty.$$

We will show that increasing the order of Ω decreases the rate of convergence but improves the condition of the approximating problems generated.

3. Differentiation. Now let us use differentiation to understand the effect of various choices of the norm Ω on the rate of convergence and on the condition of the regularizing approximations. We want to demonstrate the interaction between the choice of Ω , the rate of convergence, the condition, the smoothness of the desired solution, and the compactness of the kernel. We will work on the interval $(-\infty, \infty)$ using Eq. (31), and use Fourier transform analysis. In Sections 3 and 4 it is assumed that each kernel K and each solution f has a Fourier transform. The smoothness of the solution is measured in terms of the decay rate of its Fourier transform for large ω . The compactness of the kernel is measured in terms of the decay rate of its Fourier transform. We further assume that the transforms of $f(\alpha, \epsilon)$, f_0 , g , and g_ϵ are in $L_2(-\infty, \infty)$.

The theorems in this section are valid for any operator K with Fourier transform

$$(32) \quad |\hat{K}(\omega)| = D\omega^{-r}, \quad r > 0.$$

This includes, in particular, the derivative operators. We use arguments similar to those in Aref'eva [5].

THEOREM 6. *We assume that the desired solution f_0 has $m \geq 1$ derivatives, that each $f_0^{(j)} \in L_1(-\infty, \infty)$, $0 \leq j \leq m$, and its Fourier transform \hat{f}_0 satisfies $|\omega^{j+1}\hat{f}_0|^2 < R_{j+1}$ for all ω and $|\hat{f}_0| > D_1\omega^{-m-1}$ for large ω . Then for any Eq. (1) whose kernel K satisfies (32), we have the following. For sufficiently small α , the rate of convergence of the Tikhonov regularizing families obtained using Ω_p in (11), varies as*

$$(33) \quad \|f(\alpha, 0) - f_0\| \leq Q\alpha^a, \quad \text{where } a = (2m + 1)/(4p + 4r), \text{ for } m \leq 2p + 2r - 1.$$

For $m > 2p + 2r - 1$, a becomes

$$(34) \quad (4p + 4r - 1)/(4p + 4r).$$

The dynamic condition of the members of these families varies as

$$(35) \quad \alpha^b, \quad \text{where } b = -r/(p + r).$$

Moreover, for any m , $p \geq 0$, and $r > 0$, an overall estimate of the error in using the solution of $P(\alpha)$ with $g_\epsilon = g_0 + \epsilon$ to approximate f_0 is of the form

$$(36) \quad \|f(\alpha, \epsilon) - f_0\| \leq C_1\alpha^q + \epsilon C_2\alpha^b,$$

where $b = -r/(2p + 2r)$ and $q = (2m + 1)/(4p + 4r)$ if $m \leq 2p + 2r - 1$, and $q = (4p + 4r - 1)/(4p + 4r)$, otherwise.

Comments. The smoothness of the solution helps the rate of convergence, but of course does nothing for the condition. Using (36), we see that increasing the order of Ω adversely affects the rate of convergence, but improves the condition. As we increase the decay rate r of \hat{K} , the condition and the rate of convergence deteriorate. As $r \uparrow \infty$, the estimate of the condition approaches $1/\alpha$ and of the rate approaches $O(1)$.

By rate of convergence, we will always mean the size of the L_2 -norm of the discrepancy caused by the regularization, $h_\alpha = f(\alpha, 0) - f_0$.

By condition we mean the dynamic condition. To briefly explain dynamic condition, consider the following. Given a symmetric matrix A , one measure of the ill-conditioning of A is the ratio, $\max |\lambda_j| / \min |\lambda_j|$, where λ_j are the eigenvalues of A .

For a circulant matrix C defined by the vector $c = (c_0, c_1, \dots, c_{n-1})$, we have a direct relationship between the eigenvalues of C and the finite Fourier transform of c , Gray [10]. Namely, $\lambda_j(C)$ equals the j th Fourier coefficient of c . Therefore, we can express the condition of a circulant matrix in terms of the maximum and minimum Fourier coefficients of c , or the eigenvalues of C .

We have a generalization of this result to symmetric Toeplitz matrices, see Widom [11], for example. A matrix T_n is Toeplitz if its entries, $T_n(i, j)$, are a function only of the difference $i - j$. Thus, T_n is constant on each of its diagonals and defined by a vector $c = (c_0, c_1, \dots, c_{n-1})$. The relevant theorem states roughly that if we consider the entries of $c = (c_0, c_1, \dots, c_{n-1})$ as the Fourier coefficients of a L_2 -function ϕ , then all the eigenvalues of T_n satisfy $m \leq \lambda_i(T_n) \leq M$, where $m = \text{ess inf } \phi$ and $M = \text{ess sup } \phi$ on the real line.

Since we have convolution operators, each can be approximated by a Toeplitz matrix (although this may not be the approximation one would use in practice). The Toeplitz matrix obtained is essentially a discretized form of the kernel, and the kernel is the Fourier transform of its transform. Thus, for this particular discretization we can use the maximum and minimum of $|\hat{K}|$ to estimate the maximum and minimum eigenvalues and thus, the condition of the matrix obtained.

It is common practice in the engineering literature, see Makhoul [12], for example, to estimate the condition of a convolution operator by the ratio of the maximum of the modulus of its Fourier transform to the minimum. The number obtained is called the dynamic condition of the operator.

Comments. In the proofs we assume p, m, r and j are integers. The proofs could be modified to accept fractional quantities. Also, we use only the p th derivative term in Ω_p . This is sufficient to give us some understanding of the relationships between the different quantities.

Proof of Theorem 6. First consider the rate of convergence of the associated regularizing family as a function of p, m , and r . Consider Ω_p given by (11) for $p = 1, 2, \dots$. We use Fourier transforms and the Euler equation (8) to compute f_α . Then the discrepancy caused by the regularization $h_\alpha = f(\alpha, 0) - f_0$, using Parseval's equality, satisfies

$$(37) \quad \|h_\alpha\|^2 = \alpha^2 \int_0^\infty \omega^{4p+4r} |\hat{f}|^2 / (\bar{c} + \alpha \omega^{2p+2r})^2 \pi.$$

Part a: Rate of Convergence—Majorant Estimate. We split the integral in (37) into the sum of integrals over $(0, M)$ and (M, ∞) where M will be chosen subsequently to optimize the error estimate. If we have $|\omega^{j+1} \hat{f}|^2 < R_{j+1}$, then we have for any M , and $0 \leq j \leq m$,

$$(38) \quad \int_M^\infty |\hat{h}_\alpha|^2 \leq R_j \int_M^\infty \omega^{-2j-2} = \bar{R}_j M^{-2j-1}.$$

Moreover, for $0 \leq j \leq J = \min(2p + 2r - 1, m)$,

$$(39) \quad \int_0^M |\hat{h}_\alpha|^2 \leq R'_j \alpha^2 \int_0^M \omega^{4p+4r-2j-2} = R'_j \alpha^2 M^{4p+4r-2j-1}.$$

Combining (38) and (39), we obtain the majorant estimate for any $j \leq J$ of the form,

$$(40) \quad \|h_\alpha\|^2 \leq E_1 \alpha^2 M^{4p+4r-2j-1} + E_2 M^{-1-2j}.$$

If we choose M to minimize the expression in (40), we obtain for each such j ,

$$(41) \quad M_j = C \alpha^{-1/(2p+2r)}.$$

Thus, the optimal choice of M is independent of $j \leq J$, and the optimal majorant estimate obtainable from (40) is

$$(42) \quad \|h_\alpha\|^2 \leq Q_j \alpha^a \quad \text{with } a = (2j + 1)/(2p + 2r),$$

where Q_j depends on j but not on α . In particular, for a given r and p , if the smoothness of f satisfies $m \leq 2p + 2r - 1$, then the best order of α in (42) corresponds to $j = m$ for which we obtain (33). If $m > 2p + 2r - 1$, then the best order is obtained with $j = 2p + 2r - 1$, when we get

$$(43) \quad \|h_\alpha\|^2 \leq \bar{Q} \alpha^{\bar{a}} \quad \text{with } \bar{a} = (4p + 4r - 1)/(2p + 2r).$$

Note. The constant terms in the various equations do not involve α and do not significantly affect the arguments so no attempts have been made to be precise with them. We note, moreover, that the estimates in (42) and (43) hold for all $\alpha > 0$.

Part b: Rate of Convergence—Minorant Estimate. We must obtain a lower bound on $\|h_\alpha\|^2$. Since $|\hat{f}| > D_1 \omega^{-m-1}$ for large ω and $|\hat{K}| = D \omega^{-r}$ everywhere, we have for sufficiently small α that

$$(44) \quad \|h_\alpha\|^2 \geq \bar{C} \int_M^\infty \alpha^2 \omega^{4p+4r-2m-2} / (1 + \alpha \omega^{2p+2r})^2,$$

where M is defined in (41). But for $\omega > M$ we have

$$(45) \quad \alpha \omega^{2p+2r} > C > 0.$$

Therefore, we have with $Q = (C + 1)/C$, on (M, ∞) the denominator in (44) satisfies

$$(46) \quad (1 + \alpha \omega^{2p+2r}) \leq Q \alpha \omega^{2p+2r}.$$

From (46) we obtain

$$\|h_\alpha\|^2 \geq \bar{Q} \int_M^\infty \omega^{-2m-2} = Q^* M^{-2m-1}.$$

Using (41), we obtain for sufficiently small α

$$(47) \quad \|h_\alpha\|^2 \geq Q' \alpha^a \quad \text{with } a = (2m + 1)/(2p + 2r).$$

The coefficient Q' in (47) depends on p , m , and r but not α .

Combining (33) and (47) for $m \leq 2p + 2r - 1$, we obtain an exact order estimate of the rate of convergence

$$(48) \quad \|h_\alpha\|^2 \leq Q_1 \alpha^a \quad \text{with } a = (2m + 1)/(2p + 2r).$$

For $m > 2p + 2r - 1$, we have only that the convergence is at least as fast as

$$(49) \quad \|h_\alpha\|^2 \leq Q_2 \alpha^{\bar{a}} \quad \text{with } \bar{a} = (4p + 4r - 1)/(2p + 2r).$$

Part c: Condition. The condition of a minimization problem typically refers to the condition of the second derivative operator of the given functional. In our case this is the operator in the Euler equation (8)

$$(50) \quad \phi = K * K + \alpha B,$$

where B is generated from Ω using variational arguments.

For differentiation,

$$(51) \quad \phi(f) = \int_t^1 \int_0^s f + \alpha f - \alpha f^{(2)}.$$

If we approximate this operator and f by discretizing, then we obtain a matrix. However, the type of matrix obtained, Toeplitz, symmetric, etc., depends upon the quadrature rule used for $K * K$ and the approximation of the differential operator chosen. The algebraic condition may vary depending upon the particular approximation chosen. Therefore, we choose to look at this operator directly and to use Fourier transform analysis. We use the dynamic condition, computing the minimum value of the modulus of the Fourier transform of our given operator. This notion of condition is only applicable to convolution operators.

Therefore, we form the Fourier transform of ϕ and consider

$$(52) \quad h(\omega) = |\hat{\phi}(\omega)| = \omega^{-2r} + \alpha \omega^{2p}.$$

In (52) we have again ignored the lower order derivative terms in Ω_p , because they have only secondary effects upon the conclusions. The minimum of h occurs at $\omega = (r/p\alpha)^{1/(2p+2r)}$ and is of the order

$$(53) \quad C\alpha^a, \quad a = r/(r + p).$$

(53) is a lower bound on the smallest eigenvalue of the operator ϕ . Observe that ϕ has no maximum. However, in numerical work, there is a limit on the sampling size, and the maximal singular value of any discretization of ϕ depends primarily on this sample size. To see this consider the following: We know, see, e.g. Wilkinson [13], that for any matrix A that approximates ϕ , all the eigenvalues of A satisfy $|\lambda_i(A)| \leq \|A\|$ for any consistent matrix norm. In particular, the maximal row sum norm is a bound on the eigenvalues.

So we can conclude that as the order of Ω is increased, the condition of the approximating problems improves, and the critical order in α is given in (35).

Part d. Overall Estimate of Error in the Approximation. Now let us consider the overall error in (7). Using Fourier transforms, we find that the

$$(54) \quad \|f(\alpha, \epsilon) - f(\alpha, 0)\| = \|\hat{f}(\alpha, \epsilon) - \hat{f}(\alpha, 0)\| \leq C_2 \epsilon \alpha^b \quad \text{with } b = -r/(2r + 2p).$$

Combining (54) and (48), we obtain (36). Q.E.D.

Theorem 6 tells us that for small decay rates r we can achieve fairly well-conditioned, approximating problems with a small order norm Ω .

We have the following simple theorem.

THEOREM 7. *Under the hypotheses of Theorem 6, for*

$$(55) \quad \alpha < (C_2 \epsilon / C_1)^2,$$

$m \leq 2p + 2r - 1$, and $2r > 2m + 1$, the choice Ω_m yields the regularizing family that minimizes the overall error in (3.6).

Proof of Theorem 7. Since $2r > 2m + 1$, the minimizer of (36) is larger than the solution p^* of the equation $\alpha^b = \epsilon C_2 / C_1$, where $b = (m + 1/2 + r)/(2r + 2p)$. Therefore, if p^* is in the interval $p \geq m$, then the minimum value corresponds to $p = m$. Taking logarithms of both sides of this equation and using (55), we get

$$2p^* + 2r > 2m + 2r + 1.$$

Therefore, the optimal choice is $p = m$. Q.E.D.

In the next section we consider more general convolution operators, again on the interval $(-\infty, \infty)$.

4. General Convolution Equations. In this section we extend the results obtained in the previous section for operators of the form $|\hat{K}(\omega)| = C\omega^{-r}$, $r > 0$, to more general convolution operators. Arsenin and Ivanov [6], Arsenin and Savelova [8], Goncharkii, Leonov and Yagola [9] and Aref'eva [5] have each derived estimates on the asymptotic rates of convergence (as $\alpha \downarrow 0$), and on the overall error bounds attainable when the data is inaccurate, for certain classes of convolution equations on the interval $(-\infty, \infty)$. Each paper used Fourier transform analysis. The families considered in Aref'eva [5] will be described. These include those considered in the other papers. The papers are listed chronologically. The Arsenin and Ivanov [6] and Arsenin and Savelova [8] papers obtained majorant estimates that relate the decay of the transform of the operator K , the order of the norm Ω , and the smoothness of the desired solution. They used complex variable theory to evaluate the error integrals and obtained pointwise estimates. Goncharkii, Leonov and Yagola [9] enlarged the class of kernels under consideration and used the L_2 -norm of the Fourier transform of the error so that they could work with real integrals. However, they did not consider general norms Ω or the effects of the smoothness of the desired solution. Aref'eva [5] considered the same class of kernels as in [9]. She answered a slightly different question. Generalizing the class of admissible regularizers, she obtained the following theorem. For a precise statement see Aref'eva [5].

THEOREM 8 (AREF'EVA [5]). *For a given level of noise, σ , the optimal approximation f_σ to the desired solution f_0 of (1) is obtained as the solution of (8) with*

$\alpha\hat{B}$ replaced by

$$(56) \quad \sigma^2/|\hat{f}_0|^2.$$

Thus, in the presence of noise, the optimal choice of regularizer depends only on the decay rate of the Fourier transform of the desired solution. We note that the expression in (56) may or may not have a function analytic interpretation as the norm of some identifiable function space. All Tikhonov regularizers are, of course, defined initially in terms of such norms. Using (56), Aref'eva [5] derives estimates on the optimal overall error attainable for a given σ with various assumptions on the decay rates of \hat{f}_0 and \hat{K} . Since (56) is not a priori computable, Aref'eva then returns to Tikhonov regularizers and shows that for certain classes of problems, she can achieve to the same order, the error that she gets using the optimal regularizer in (56).

In this paper we consider only Tikhonov regularizers. We will demonstrate that the character of the compactness of K determines whether or not the approximation obtained using a particular family of regularizers is sufficient to make sense computationally. As we saw by examples in Section 2, with exponentially decaying operators, in order to obtain an approximation of the order α^a for some $a > 0$, we had to use exponentially increasing smoothing norms. Obviously, the use of such norms is limited by the smoothness of the desired solution, since the convergence arguments are valid only when $\Omega(f_0)$ is defined.

Numerically, we have to worry about the condition of the approximating problems generated. As in Section 3, we will use the dynamic condition.

The discussion has three parts. First, for a large class of operators, we obtain in Theorem 9 majorant estimates of the rates of convergence of approximating regularizing families obtained from Ω_p . In Theorem 10, for a subclass of these operators, we obtain minorant estimates of these rates. Together Theorems 9 and 10 give us the order of the rates of convergence for this subclass. In Theorems 11 and 12, for a further subclass, we estimate the dynamic condition of the approximating problems generated using Ω_p (Theorem 11) and using Ω_∞^c and $\bar{\Omega}_\infty^c$ (Theorem 12). Together these theorems give us an estimate on the overall behavior of the error in approximating our given solution f_0 by the solution of $P(\alpha)$ for $g_\epsilon = g_0 + \epsilon$.

Using these estimates, one can argue that typically one would obtain the best error by maximizing the order of Ω . (For a specific problem one would have to be more precise in such arguments and take into account the dependence of the coefficients that appear in the estimates upon p .) For a function f_0 with m derivatives this corresponds to taking $p = m$. We note, however, that in practice we often do not know m , and we are likely to choose a low order Ω because it is easier to work with. In such a situation it is important to have some estimate of the decay rate of our operator. Since we are nominally working with symmetric, Toeplitz operators, the associated Fourier transforms give us such an estimate. Using this estimated decay rate, we can use inequalities similar to those obtained in these theorems to indicate whether the proposed regularization makes sense numerically.

The following are slight modifications of definitions in Aref'eva [5].

Definition 2 (Aref'eva [5]). A kernel K is of *type 1* if: (a) Its transform, $|\hat{K}(\omega)|$ has a denumerable number of zeros on the real axis of ω . (b) There are no limit points for the zeros of $\hat{K}(\omega)$ with the exception of $\omega = +\infty$, and if $N(M)$ is the number of zeros in $(0, M)$, then $N(M) \leq CM^\nu$ for some $\nu \geq 0$, $C \geq 0$. (c) There exists a system of nonoverlapping ϵ -neighborhoods of the zeros of $\hat{K}(\omega)$ on $\omega \geq 0$ such that $|\hat{K}(\omega)| \geq a_i |\omega - \omega_i|^s$ for $\omega \in [\omega_i - \epsilon_i, \omega_i + \epsilon_i]$, where the $\omega_i > 0$ are the finite nonzero, zeros of \hat{K} and $a_i \geq a > 0$, $s > 0$. (d) $\hat{K}(0) \neq 0$. (e) Outside the ϵ -neighborhoods, $|\hat{K}(\omega)| \geq D\omega^{-r}$, $r \geq 0$, $D > 0$.

We note that K can have infinitely many zeros satisfying (c) only if $\epsilon_i \rightarrow 0$ as $i \rightarrow \infty$.

Definition 3 (Aref'eva [5]). A kernel is of *type 2* if its transform has no finite zeros and for large ω , $|\hat{K}(\omega)| \geq C \exp(-a\omega)$ for some $a > 0$, $C > 0$.

Within type 1 Aref'eva [5] determines a subclass for which she derives a lower bound estimate of the same order. Aref'eva [5] also classifies solutions f_0 in terms of the behavior of their Fourier transforms. She further subdivides type 1 into $0 \leq s < 1/2$, integrable singularities, and $s > 1/2$, nonintegrable singularities. We consider only $s < 1/4$. We use Aref'eva type arguments to derive our estimates in Theorems 9 and 10. We note that the arguments in Section 3 also mimic Aref'eva [5] arguments.

THEOREM 9. Consider the norms Ω_p , $p \geq 1$, and Eq. (1). Let the solution f_0 of (1) have m derivatives with each $f_0^{(j)} \in L_1(-\infty, \infty)$, $0 < j \leq m$ and $|\omega^{j+1} \hat{f}(\omega)|^2 \leq R_{j+1}$ for all ω .

(a) Let K be of type 1 with $0 \leq s < 1/4$ and $\nu \leq 4s$. Then the discrepancy $h_\alpha \equiv f(\alpha, 0) - f_0$ satisfies the following as $\alpha \downarrow 0$. For $m \leq 2p + 2r - 1$

$$(57) \quad \|h_\alpha\|^2 \leq C_1 \alpha^b \quad \text{with } b = (2m + 1)/(2p + 2r).$$

For $m > 2p + 2r - 1$,

$$(58) \quad b = (4p + 4r - 1)/(2p + r).$$

(b) Let K be of type 2 with $|\hat{K}| \geq D \exp(-a\omega)$ for some $a > 0$ and for all ω . Then

$$(59) \quad \|h_\alpha\|^2 \leq C_3 (-\ln(b\alpha^2)/2a)^{-2(m+1)}.$$

The proofs in this section will also treat p , r , m and j as integers. The proofs could be modified to handle fractional values. As in Section 3, we have used only the p th derivative term in Ω_p .

Proof of Theorem 9. (The argument parallels that in Theorem 2 of Aref'eva [5].)

Part a. Type 1 Operators. The proof is analogous to that given for Theorem 6 using an expression like (37) for $\|h_\alpha\|^2$ except that it is now necessary to split the integral over $(0, M)$ into integrals over neighborhoods of the zeros of \hat{K} , and integrals over the complements of these neighborhoods. For each neighborhood $N_i = (\omega_i - \epsilon_i, \omega_i + \epsilon_i)$ of a zero ω_i of $|\hat{K}|$ we have

$$\int_{N_i} |\hat{h}_\alpha|^2 \leq \alpha^2 \int_{N_i} \omega^{4p} |\hat{f}|^2 / (\alpha_i^4 |\omega - \omega_i|^{4s}) \leq \bar{C} \alpha^2 \epsilon_i^{1-4s+d_j},$$

where $d_j = \max(0, 4p - 2j - 2)$.

There are at most M^ν such intervals in $(0, M)$ and clearly each $\epsilon_i < M$, so we obtain

$$\sum_i \int_{N_i} |\hat{h}_\alpha|^2 \leq \bar{C} \alpha^2 M^{1-4s+\nu+d_j}.$$

Now consider the integrals over the complements \bar{N}_i of the N_i . On the complements $|\hat{K}| > D\omega^{-r}$

$$\sum_i \int_{\bar{N}_i} |\hat{h}_\alpha|^2 \leq \bar{D} \alpha^2 \int_0^M \omega^{4p+4r} |\hat{f}|^2 \leq \alpha^2 \bar{D} M^{4p+4r-2j-1}$$

for any $j \leq J = \min(m, 2p + 2r - 1)$.

The integral from (M, ∞) , as in Theorem 6, satisfies (38). Therefore,

$$(60) \quad \|h_\alpha\|^2 < \alpha^2 \bar{C} M^\mu + \bar{D} M^{-1-2j},$$

where $\mu = \max(1 + \nu - 4s + d_j, 4p + 4r - 2j - 1)$. Since $\nu \leq 4s$ and $j \leq 2p + 2r - 1$, we have for each j , $\mu = 4p + 4r - 2j - 1$. Minimizing (60) over M , we obtain

$$(61) \quad M_{\text{opt}} = C_0 \alpha^{-1/(2p+2r)}.$$

Therefore, for $m \leq 2p + 2r - 1$ we obtain (57). For $m > 2p + 2r - 1$ we obtain (58).

Part b. Type 2 Operators. We have assumed that \hat{K} has no real zeros. We again split the error estimate (37) into integrals over $(0, M)$ and (M, ∞) .

$$(62) \quad \int_0^M |\hat{h}_\alpha|^2 \leq \int_0^M \alpha^2 \omega^{4p} e^{4a\omega} |\hat{f}|^2 / (D^4 + \alpha^2 \omega^{4p} e^{4a\omega}) \leq \alpha^2 C_1 e^{4aM} M^d,$$

where $d = \max(0, 4p - 2m - 2)$. Moreover, we obtain (38) again.

Combining (38) and (62) and then minimizing over M , we obtain the transcendental equation for the optimal M

$$4aC_1 \alpha^2 e^{4aM} = C_2 M^{-2m-2-d}.$$

Taking logarithms, we find that the solution

$$(63) \quad M^* \sim [-\ln(b\alpha)/2a] [1 - \ln(-\ln(b\alpha)/2a)(2(m+1)/(2(m+1) - 2\ln b\alpha))],$$

where $b = 4aC_1/C_2$. Therefore, a majorant estimate is of the form

$$(64) \quad \|h_\alpha\|^2 \leq C(-\ln(b\alpha^2)/(2a))^{-2d(m+1)}.$$

For α small, d is approximately 1.

Comments. Theorem 9 states that we can obtain the same majorant of the rate of convergence that was obtained in Theorem 6 when the growth rate ν of the number of zeros in the interval $(0, M)$ is not greater than $4s$. The exponent s in Definition 1 is

a measure of the rate of decrease of \hat{K} in a neighborhood of any zero. Theorem 9 also states that for type 2 operators the rate of convergence depends primarily on the smoothness of the desired solution and is inversely proportional to $-\ln \alpha$.

THEOREM 10. *Consider the norms Ω_p , $p \geq 1$, and Eq. (1). Let the desired solution f_0 have m derivatives with each $f_0^{(j)} \in L_1(-\infty, \infty)$, $0 \leq j \leq m$. Let its Fourier transform \hat{f}_0 satisfy $|\omega^{j+1}\hat{f}_0| \leq R^{j+1}$ for all ω , and $D_2\omega^{-m-1} \geq |f_0| \geq D_1\omega^{-m-1}$ for large ω .*

Let K be of type 1 with $0 \leq s < 1/4$ and $\nu = 0$. ($\nu = 0$ corresponds to a finite number of zeros.) Moreover, assume $|\hat{K}(\omega)| \leq \bar{D}\omega^{-r}$ as $\omega \uparrow \infty$. Then for sufficiently small α , we obtain the following lower bound:

$$(65) \quad \|h_\alpha\|^2 \geq Q\alpha^a \quad \text{with } a = (2m + 1)/(2p + 2r).$$

Proof of Theorem 10. As $\alpha \downarrow 0$, $M(\alpha)$ in (60) becomes arbitrarily large. Therefore, for sufficiently small α , we have

$$\|h_\alpha\|^2 \geq \int_M^\infty (\alpha^2 \omega^{4p+4r} |\hat{f}|^2) / (\bar{D} + \alpha^2 \omega^{4p+4r}).$$

Furthermore, we have for $\omega > M$, $\alpha^2 \omega^{4p+4r} > C_0$. Therefore, for some Q ,

$$\|h_\alpha\|^2 \geq Q \int_M^\infty |\hat{f}|^2 = \bar{Q}M^{-2m-1} = \bar{Q}\alpha^a$$

with $a = (2m + 1)/(2p + 2r)$. Q.E.D.

Comments. Theorems 9 and 10 yield exact order estimates for the rate of convergence for type 1 operators whenever $m \leq 2p + 2r - 1$, $0 \leq s < 1/4$, and $\nu = 0$.

Next we consider the condition of the approximating problems generated. We again use the dynamic condition, Makhoul [12].

THEOREM 11. *Consider the norms Ω_p , $p \geq 1$, and Eq. (1). Let \hat{K} be of type 1 or 2. Assume moreover that $|\hat{K}|$ is a monotone decreasing function of ω such that for type 1, $D\omega^{-r} < |\hat{K}| < \bar{D}\omega^{-r}$ outside an interval of 0, and for type 2, $C \exp(-a\omega) < |\hat{K}| < \bar{C} \exp(-a\omega)$ for all ω . Then for small α , the dynamic condition of the members of the approximating families generated is of order*

$$(66) \quad \bar{Q}\alpha^{-r/(p+r)}$$

for type 1 operators, and of order

$$(67) \quad \bar{Q}\alpha^{-1}((-\ln b\alpha)/2a)^{-2p}$$

for type 2 operators where $b = p/ac$.

Proof of Theorem 11. We consider the operator in the Euler equation (8),

$$\phi \equiv K^*K + \alpha B.$$

Part a. Type 1 Operators. Outside an interval of $\omega = 0$, $(0, \delta)$,

$$d\omega^{-2r} + \alpha\omega^{2p} \leq |\hat{\phi}(\omega)| \leq \bar{d}\omega^{-2r} + \alpha\omega^{2p}.$$

The minimum of $d\omega^{-2r} + \alpha\omega^{2p}$ occurs when

$$(68) \quad \omega^* = (rd/p\alpha)^{1/(2p+2r)},$$

and the corresponding minimum value is of the form,

$$(69) \quad Q\alpha^{r/(p+r)}.$$

For small α , ω in (68) is not in $(0, \delta)$. Therefore, the minimum value of $|\hat{\phi}(\omega)|$ lies between the two values of (69) obtained for d and \bar{d} . As we argued in Section 3, $|\hat{\phi}|$ has no maximum; however in practice the maximum is limited by the sample size. Therefore, we obtain (66).

Part b. Type 2 Operators. Using a similar argument for case (b), we obtain the equation

$$\alpha p \omega^{2p-1} = ac \exp(-2a\omega)$$

that we must solve. The solution is approximately

$$\omega^* \sim [-\ln(b\alpha)/2a] [1 - \ln(-\ln(b\alpha)/2a)((2p-1)/((2p-1) - \ln b\alpha))],$$

where $b = p/ac$ and for small α the minimum value of $\hat{\phi}$ is of order $Q\alpha(-\ln(b\alpha)/2a)^{2p}$. Q.E.D.

Comments. We see in Theorem 11 that for operators whose Fourier transforms decay like ω^{-r} , $r > 0$, increasing the order of Ω_p has a significant effect upon the condition of the approximating problems generated by the regularization. However, if the decay of $|\hat{K}|$ is exponential, then this effect is much less. As indicated by examples in Section 2 for operators whose transforms decay exponentially, to achieve a condition similar to (66), we must use a norm like Ω_∞^c in (12) or $\bar{\Omega}_\infty^c$ in (23).

THEOREM 12. *Use the norms Ω_∞^c and $\bar{\Omega}_\infty^c$ introduced in Section 2 in Eqs. (12) and (23). Let \hat{K} be of type 1 or type 2, satisfying the hypotheses of Theorem 11. Then the dynamic condition of the members of the approximating families generated is as follows. For K of type 1, the condition is majorized by*

$$(70) \quad Q((-\ln b\alpha)/c)^q,$$

where $q = 2r$ corresponds to Ω_∞^c , and $q = r$ to $\bar{\Omega}_\infty^c$. For type 2 operators we obtain for Ω_∞^c , the condition is of order

$$(71) \quad Q\alpha^{-d} \quad \text{with } d = 2a/(2a + c)$$

and for $\bar{\Omega}_\infty^c$ of order

$$(72) \quad Q((-\ln b\alpha)/c)^{1/2}.$$

Proof of Theorem 12. Part a. Type 1 Operators. As in the proof of Theorem 11, we compute the minimum of

$$(73) \quad |\hat{\phi}_1| = d\omega^{-2r} + \alpha \cosh c\omega$$

for Ω_∞^c and of

$$(74) \quad |\hat{\phi}_2| = d\omega^{-2r} + \alpha \exp(c\omega^2)$$

for $\bar{\Omega}_\infty^c$. Differentiating (73), we obtain

$$\alpha c \omega^{2r+1} = 2rd(\sinh c\omega)^{-1}.$$

But this equation has the same form as that obtained for case (b) in Theorem 11.

Therefore, its approximate solution is given by (63) with some renaming of constants and replacing α^2 by α .

$$\omega^* \sim (-\ln(b\alpha)/c)[1 - \ln(-\ln(b\alpha)/c)((2r+1)/(2r+1 - \ln b\alpha))].$$

Therefore, for small α ,

$$(75) \quad |\hat{\phi}_1(\omega)| \geq (c/(-\ln b\alpha))^{2r}.$$

For (74) we obtain

$$(76) \quad \alpha c \omega^{2r+2} = rd \exp(-c\omega^2).$$

The solution of (76) is approximately

$$c\omega^2 = -\ln(b\alpha)[1 - \ln(-\ln(b\alpha)/c)((r+1)/(r+1 - \ln(b\alpha)))],$$

where $b = c/rd$. Therefore, for small α ,

$$|\hat{\phi}_2(\omega)| \geq (c/(-\ln b\alpha))^r.$$

Part b. Type 2 Operators.

$$(77) \quad |\hat{\phi}_1| = d \exp(-2a\omega) + \alpha \exp(c\omega)$$

for Ω_∞^c ; and for $\bar{\Omega}_\infty^c$,

$$(78) \quad |\hat{\phi}_2| = d \exp(-2a\omega) + \alpha \exp(c\omega^2).$$

Differentiating (77) and solving for ω , we obtain

$$\omega \sim -\ln(b\alpha)/(2a+c),$$

where $b = c/2ad$. Therefore,

$$|\hat{\phi}_1| > Q\alpha^{\bar{d}} \quad \text{with } \bar{d} = 2a/(2a+c).$$

Differentiating (78) and solving for ω , we obtain

$$c\omega^2 \sim -\ln(b\alpha) - \ln((-\ln b\alpha)/c)^{1/2}.$$

Therefore,

$$|\hat{\phi}_2| \geq Q(c/(-\ln b\alpha))^{1/2}.$$

Combining these estimates, we can obtain overall estimates on the error, $\|f(\alpha, \epsilon) - f_0\|^2$.

We conclude this section with the following theorem which states that if K is symmetric and positive definite, then a regularizing family with improved convergence and condition can be used instead of Eq. (8).

THEOREM 13. *If the operator K in (1) is symmetric and positive definite, then*

for Ω_p , $p \geq 1$, the solutions of the problems $0 \leq \alpha \leq 1$,

$$(79) \quad (K + \alpha B)f = g$$

with the associated natural boundary conditions form a regularizing family for $Kf = g$ with respect to Ω_p .

We note that the decay rate for K is the square root of that for K^*K .

Proof of Theorem 13. Define operators

$$K_1 = \sum \sigma_n^{1/2} u_n u_n^T \quad \text{and} \quad K_2 = \sum u_n u_n^T / \sigma_n^{1/2}.$$

Then (79) is the Euler equation for the problem

$$(80) \quad \text{Minimize } \|K_1 f - K_2 g\|^2 + \alpha \Omega(f).$$

Equivalently,

$$(81) \quad \text{Minimize } \|Kf - g\|_{K_3}^2 + \alpha \Omega(f),$$

where $K_3 \equiv \sum u_n u_n^T / \sigma_n$.

We first must show that each of the families (80) and (81) is well defined. First note that the $\{u_n\}$ are complete since K is invertible. Moreover, we know that $Kf = g$ has a solution only if $\sum (g^T u_n)^2 / \sigma_n^2 < \infty$. Therefore, $K_1 f$, $K_2 g$ and $K_3 g$ are all well-defined.

f_α minimizes (80) if and only if it satisfies the associated Euler equation. Due to the symmetry, this equation is just (79). From Tikhonov [2] we know that the solution of (79) converges at least uniformly to the solution of

$$(82) \quad K_1 f = K_2 g$$

as $\alpha \downarrow 0$. But for the given g any solution of (82) is a solution of (1). Q.E.D.

5. Summary. In the preceding sections we have considered the question, how does our choice of Ω in (5) affect the condition and the rate of convergence of the approximating problems generated. For convolution equations, we have described the relationship of the condition and the rate of convergence to not only the choice of norm Ω , but also to the character of the compactness of the operator K , and the smoothness of the desired solution. The condition of the approximating problems generated is strongly dependent upon the order of Ω , and improves as we increase this order. For a given Ω , the condition deteriorates as we increase the decay rate of the singular values of K . The rate of convergence depends not only on the compactness of K and the order of Ω but also upon the smoothness of the desired solution. For a given operator K , the smoother the desired solution the better the rate of convergence. This rate deteriorates as either the order of Ω is increased or the decay rate of the operator is increased. With respect to the order of Ω , this deterioration is limited by the smoothness of the desired solution.

The estimates obtained in Sections 3 and 4 are crude. They are intended only to provide an appreciation for the relationship between the choice of the smoothing norm Ω , the compactness of K , and the smoothness of the solution. In particular, they imply that if the singular values of K decay exponentially, then regularization

with Ω_p in (11) cannot be expected to provide good approximations. It is interesting to note that the estimates indicate that for an operator whose singular values decay polynomially, nothing is gained by using an exponential norm.

These results give heuristics for determining the appropriateness of a proposed regularization for a convolution equation. In practice, we probably would not know the smoothness of the desired solution and would choose a lower order Ω because it is easier to work with. Moreover, we would probably be working on a finite interval, for example, $[0, 1]$ not $(-\infty, \infty)$.

In an implementation, we have at least three ways of solving the minimization problems in (5):

- (a) Direct discretization of the functionals in (5) or of the Euler equation in (8).
- (b) Solving the Euler equation, using Fourier transforms.
- (c) Inverting B in (8) with the natural boundary conditions and solving the

resulting integral equation of the second kind.

All three approaches have been used in practice, see Glasko, Kulik and Tikhonov [14] for approach (a), Anderssen and Bloomfield [15] for approach (b), and Cullum [16] for approach (c). Approach (a) is limited to lower order Ω , whereas approach (b) could be used with general regularizers as is done in Aref'eva [5]. In both approaches, (a) and (b), if we work with the Euler equation, we have the question of satisfying the natural boundary conditions associated with the Euler equation. The statement that (8) is equivalent to (5) includes these boundary conditions. With approach (c) we note that the solution obtained automatically satisfies the natural boundary conditions, because we must have

$$(83) \quad \alpha f = B^{-1}(K^*g - K^*Kf).$$

We also note, however, that if we use approach (c), then the resulting operator in (82) is only perturbed by αI from the rapidly decaying operator $B^{-1}K^*K$. Thus, this approach does not seem desirable. However, we can argue that Eq. (82) is not as bad as it seems. In contrast to what is often done in practice—namely, smoothing the given data and then solving $Kf = g_{\text{smooth}}$ —in approach (c) we simultaneously smooth the data $B^{-1}K^*g$ and alter the equation. Smoothing the data in this way removes the high frequency components of the noise in a way consistent with the original equation. To see this consider the following. For B^{-1} generated by (6) the eigenvectors have the oscillation property. That is, if we order the eigenvalues of B^{-1} , μ_n , such that $\mu_n \downarrow 0$ as $n \uparrow \infty$, then the number of zero crossings of the eigenvectors $w_n = \sin n\pi t$, $n = 1, 2, \dots$ (for $n = 0$, $w_0 = 1$), increases as $n \uparrow \infty$. Since the error in the modified data satisfies

$$B^{-1}K^*\epsilon = \sum_n (w_n^T K^*\epsilon) \mu_n w_n,$$

the high frequency components of $K^*\epsilon$ (which already are smoother than those of ϵ) are modified by μ_n .

We could obtain a symmetrized version of (82)

$$(84) \quad ((B_1 K^* K B_1) + \alpha I)h = B_1 K^* g$$

if we could compute $B_1 = B^{1/2}$. Then h will be derivatives of f . Once we have (84), we can relate it directly to the results obtained previously with $B = \text{identity}$, see Hilgers [17]. We note that if instead of using B with the natural boundary conditions, we use another operator with different boundary conditions, see for example Hilgers [17], then the approximations we generate will be good only if the desired solution satisfies these conditions. A comparison of the three approaches (a)–(c) is needed.

The preceding discussion demonstrates the importance of one's having some understanding of the given operator K . This can be used to estimate the appropriateness of any proposed regularization, and to give an indication of potential difficulties.

Mathematical Sciences Department
IBM Thomas J. Watson Research Center
Yorktown Heights, New York 10598

1. F. SMITHIES, *Integral Equations*, Cambridge Univ. Press, London, 1958.
2. A. N. TIKHONOV, "Solution of incorrectly formulated problems and the regularization method," *Soviet Math. Dokl.*, v. 4, 1963, pp. 1035–1038.
3. A. V. CHECHKIN, "A. N. Tikhonov's special regularizer for integral equations of the first kind," *U.S.S.R. Computational Math. and Math. Phys.*, v. 10, 1970, pp. 234–246.
4. JOEL N. FRANKLIN, "On Tikhonov's method for ill-posed problems," *Math. Comp.*, v. 28, 1974, pp. 889–907.
5. M. V. AREF'EVA, "Asymptotic estimates for the accuracy of optimal solutions of equations of the convolution type," *U.S.S.R. Computational Math. and Math. Phys.*, v. 14 (4), 1974, pp. 19–33.
6. V. YA. ARSEININ & V. V. IVANOV, "The solution of certain convolution type integral equations of the first kind by the regularization method," *U.S.S.R. Computational Math. and Math. Phys.*, v. 8, 1968, pp. 88–106.
7. V. YA. ARSEININ & V. V. IVANOV, "The effect of regularization of order p ," *U.S.S.R. Computational Math. and Math. Phys.*, v. 8, 1968, pp. 221–225.
8. V. YA. ARSEININ & T. I. SAVELOVA, "The application of the method of regularization to integral equations of the first kind of the convolution type," *U.S.S.R. Computational Math. and Math. Phys.*, v. 9, 1969, p. 1392–1396.
9. A. V. GONCHARSKIĬ, A. S. LEONOV & A. G. YAGOLA, "Some estimates of the rate of convergence of regularized approximations for equations of the convolution type," *U.S.S.R. Computational Math. and Math. Phys.*, v. 12 (3), 1972, pp. 243–254.
10. ROBERT M. GRAY, *Toeplitz and Circulant Matrices, A Review*, TR 6502-1, Stanford Electronics Lab., Stanford Univ., Stanford, Calif., 1971.
11. H. WIDOM, "Toeplitz matrices," *Studies in Real and Complex Analysis*, Math. Assoc. Amer. Studies in Math. (I. I. Hirschmann, Editor), Prentice-Hall, Englewood Cliffs, N. J., 1965, pp. 179–209.
12. JOHN MAKHOUL, "Linear prediction: A tutorial review," *Proc. IEEE*, v. 63 (4), 1975, pp. 561–580.
13. J. H. WILKINSON, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965.
14. V. B. GLASKO, N. I. KULIK & A. N. TIKHONOV, "Determination of the geoelectric cross-section by the regularization method," *U.S.S.R. Computational Math. and Math. Phys.*, v. 12 (1), 1972, pp. 174–186.
15. R. S. ANDERSSON & PETER BLOOMFIELD, "Numerical differentiation procedures for non-exact data," *Numer. Math.*, v. 22, 1974, pp. 157–182.
16. JANE CULLUM, "Numerical differentiation and regularization," *SIAM J. Numer. Anal.*, v. 8 (2), 1971, pp. 254–265.
17. JOHN W. HILGERS, "On the equivalence of regularization and certain reproducing kernel Hilbert space approaches for solving first kind problems," *SIAM J. Numer. Anal.*, v. 13, 1976, pp. 172–184.
18. JANE CULLUM, *Ill-Posed Problems, Regularization, and Singular Value Decompositions*, IBM Research Report RC 6465, Yorktown Heights, New York, 1977.
19. GRACE WAHBA, "Practical approximate solutions to linear operator equations when the data are noisy," *SIAM J. Numer. Anal.*, v. 14, 1977, pp. 651–667.