NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

MULTILEVEL APPROACHES TO ILL-POSED EQUATIONS

by

Barry Lee
B.Sc., Simon Fraser University, Burnaby, 1989

THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
in the Department
of
Mathematics and Statistics

SIMON FRASER UNIVERSITY
February, 1992

ISBN 0-315-83690-3

Approval

Name : Barry Lee

Degree : Master of Science

Title of Thesis : Multilevel Approaches to Ill-posed Equations

Examining committee :

Chairman: *Dr. G.A.C. Graham*

Dr. M.R. Trummer
Senior Supervisor

Dr. R.D. Russell

Dr. E. Van Vleck

Dr. Matt Yedlin
External Examiner
Departments of Geophysics and Electrical Engineering
University of British Columbia

Date Approved : April 6, 1992

PARTIAL COPYRIGHT LICENSE

I hereby grant to Simon Fraser University the right to lend my thesis, project or extended essay (the title of which is shown below) to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this work for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this work for financial gain shall not be allowed without my written permission.

Title of Thesis/Project/Extended Essay

Multilevel Approaches to Ill-posed Equations

Author:

(signature)

Barry Lee

(name)

April 2, 1993

(date)

Abstract

The computational cost for solving ill-posed problems can be extremely large: $\frac{n^3}{2}+\frac{jn^3}{6}$ for the Tikhonov method, where j is the number of Newton iterations, and in^2 for the Landweber iteration, where i is the number of Landweber iterations. Several fast techniques exist such as the QR method of Eldén and Voevodin, and the generalized Landweber iteration of Strand. But the former method first involves an expensive transformation process applied to the operator, and then a search for the regularization parameter. This expensive transformation can induce rounding errors, which would affect the parameter search. Also, the latter method involves polynomials of the operator, and hence may not be practically applicable. In this thesis, we present several multilevel approaches which can reduce the computational cost dramatically. We present a multilevel Landweber iteration that resembles the Twomey-Tikhonov method, and a multilevel Tikhonov method with zero'th-order and first-order stabilizers. Using the multilevel Tikhonov method with a first-order stabilizer, we also present a multilevel technique for a class of nonlinear first kind equations.

Central to these multilevel methods are the effects of the interpolation error– how does the interpolation error belonging in the null space of the operator K affect the approximation?; the damping effect of the multilevel process– can we interpret the methods in terms of "selective" dampening of frequencies of the error on different grid levels?; and the parameter selection– how can we choose a good parameter and how is the parameter of one grid related to the parameter of the next grid? It will be shown that the effects of the interpolation error can be controlled by the choice of interpolation operator. We also will give an intuitive view of the damping effect of the multigrid schemes, and we will describe some generalized discrepancy principles and the quasi-optimal parameter selection procedures. Showing a relationship between the parameter of one grid to the parameter of the next grid definitely requires more investigation.

Numerical experiments conducted demonstrate that the multilevel methods can be very effective. We describe some of these experiments in the last chapter.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The view of ill-posed problems has changed since the time of Hadamard. Indeed, ill-posed problems arise in many of the natural sciences and will continue to arise in more problems as technology advances. Seismology, mineral and oil exploration, computer tomography, plasma diagnostics, radioastronomy, and image processing are some of the present diverse fields which constantly encounter this branch of mathematics. And ironically, Hadamard's classical example of the Cauchy problem for the Laplace equation itself describes many physical models appearing in several of these fields. We give examples of this and of other ill-posed problems below.

1.1 Inverse Problems in Gravimetry

Consider the problem of determining the topography of the earth's base. This problem is confronted in mineral and oil exploration since mineral and oil deposits are often related to specific characteristic forms of the base and intrusions. One approach for determining the topography is to extract information from gravitational observations at the surface of the earth. Since the intensity of the gravitational field at the earth's surface is affected by the inhomogenities of the terrain and of the density of material within it, and since the base and the intrusions are of higher densities than that of the sediment, the topography of the base and intrusions can be constructed using these gravitational observations.

1.1.1 Continuation of Static Fields

It is known that away from the source of the gravitational field, the potential of the field and the intensity components satisfy Laplace's equation. The problem is to determine how the intensity continues towards its source.

To see the significance of this problem, consider figure 1.1. If the distance between two adjacent intrusions is greater than the distance between the intrusions and the surface, then

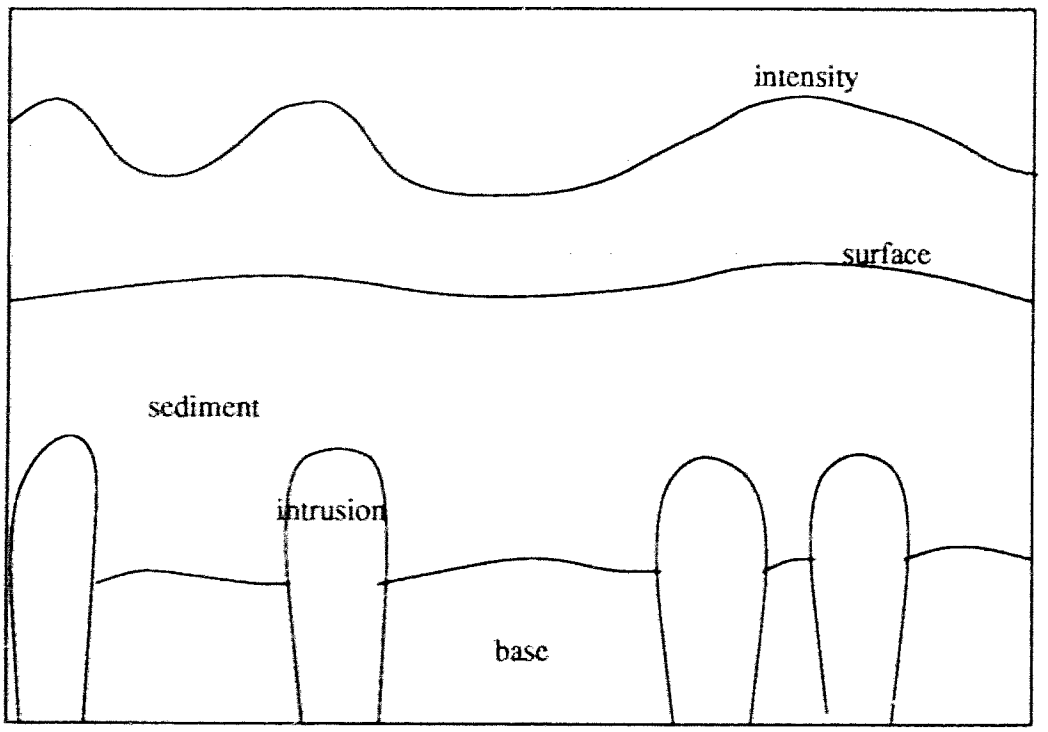

Figure 1.1: Continuation of static fields.

two local maxima will appear on the graph of the gravitational field intensity. Otherwise, if the distance between two adjacent intrusions is less than their distance from the surface, then only one local maximum will appear. The latter case illustrates the uncertain number of intrusions corresponding to a single local extremum. This uncertainty is resolved by determining the intensity at a depth below the surface. If the new graph maintains a single local maximum at the interval of interest, then the reliability of the former graph is strengthened. But if an additional local maximum appears, then it is possible that there are two intrusions beneath the interval. In mineral and oil drillings, it is desirable to drill

between the bodies.

Now let the earth's surface coincide with the x-axis, and let $u(x, y)$ be the vertical intensity component of the gravitational field generated by several bodies located at a depth of $y < -H$. Then

$$\begin{aligned} \Delta u &= 0 \\ u(x, 0) &= g(x), \end{aligned} \tag{1.1}$$

where $g(x)$ is known. The unknown is the intensity function

$$f(x) = u(x, -h), \qquad 0 < h < H.$$

Since equation (1.1) is equivalent to the first kind integral equation

$$\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{h f(\xi)}{h^2 + (x - \xi)^2} d\xi = g(x), \tag{1.2}$$

which is the integral equation formulation of the Cauchy problem for the Laplace equation, the continuation problem reduces to Hadamard's problem. Equation (1.2) is an ill-posed linear integral equation.

1.1.2 Contact Surfaces and the Prediction of Gas- and Oil-bearing Structures

Another inverse problem in gravimetry involves the study of multilayered media. Only the single-layered case will be discussed.

Gas and oil deposits are often characterized by dome-shaped uplifts in the base's topography. These structural problems are simulated using the concept of contact surfaces at which the medium's density abruptly changes. Let $z = -h$ be a horizontal plane, and assume that the density of a bounded region abutting this plane deviates from the density of the plane. Denote the bounded region by D. Also, let $\Delta\rho(x, y, z)$ denote the excessive density of D, let $u(x, y, z)$ be the potential of the anomalous gravitational field, and let γ denote the gravitational constant. Then

$$u(x, y, z) = \frac{\gamma}{4\pi} \iiint_D \frac{\Delta\rho}{[(x - \xi)^2 + (y - \eta)^2 + (z - \zeta)^2]^{\frac{1}{2}}} \, d\xi \, d\eta \, d\zeta. \tag{1.3}$$

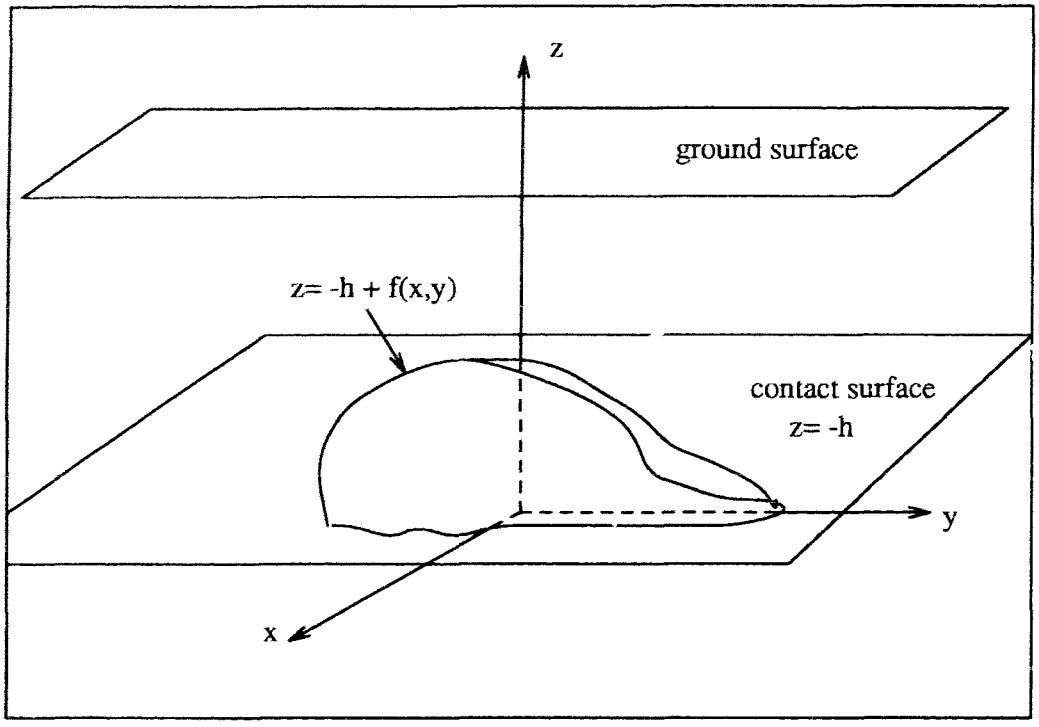

Figure 1.2: Contact surface.

Moreover, the vertical component of the intensity of the anomalous field $H(x,y)$ is given by

$$\begin{aligned} H(x,y) &= -\frac{\partial}{\partial z}\, u(x,y,z)_{|z=0} \\ &= -\frac{\gamma}{4\pi}\iiint_D \frac{\zeta\Delta\rho}{[(x-\xi)^2+(y-\eta)^2+\zeta^2]^{\frac{3}{2}}}\, d\xi\, d\eta\, d\zeta. \end{aligned} \tag{1.4}$$

(1.4) does not have a unique solution.

Now assume that $\Delta\rho$ is a known constant, that D is bounded above by $-h+f(x)$, and for simplicity, consider only the 2-dimensional case. We have

$$\begin{aligned} H(x) &= -\frac{\Delta\rho\gamma}{2\pi}\frac{\partial}{\partial y}\left\{\int_a^b\int_{-h}^{-h+f(\xi)} \ln\left[\frac{1}{[(x-\xi)^2+(y-\eta)^2]^{\frac{1}{2}}}\right] d\xi\, d\eta\right\}_{|y=0} \\ &= -\frac{\Delta\rho\gamma}{2\pi}\int_a^b\int_{-h}^{-h+f(\xi)} \frac{\eta}{(x-\xi)^2+\eta^2}\, d\xi\, d\eta \\ &= -\frac{\Delta\rho\gamma}{4\pi}\int_a^b \ln\left[\frac{(x-\xi)^2+(h-f(\xi))^2}{(x-\xi)^2+h^2}\right] d\xi. \end{aligned} \tag{1.5}$$

This is an ill-posed nonlinear first kind integral equation in $f(x)$.

1.2 Medical Tomography

The procedures of this last example are not only applicable to medical tomography, but they also are applied in other fields such as plasma physics, which frequently studies plasmas that reach several tens of millions of degrees– no measuring device can withstand this temperature!

To solve these inverse problems, the following stages are performed :

- remove the instrumental disturbances, i.e. solve

$$Bg = v,$$

  where B is an operator defined by the device;

- determine the attributes of the unknown function f from the data g, i.e. solve

$$Kf = g,$$

  where K is the operator of the problem;

- display the attributes of f in some visual form, i.e. solve

$$Dw = f,$$

  where D is an operator defined by the visual apparatus and w is the "brightness function."

Figure 1.3 illustrates these stages and the scanning process in medical tomography. Here, the rays are sent through the object (e.g. the head) at different paths by rotating the detector and x-ray source together. Variables l and θ are parameters of the path L on the path Q, and

$$l = x\cos\theta + y\sin\theta.$$

The three stages are given by

$$\int_{-\infty}^{\infty} b(l-l')g(l',\theta)\,dl' \;=\; v(l,\theta), \tag{1.6}$$

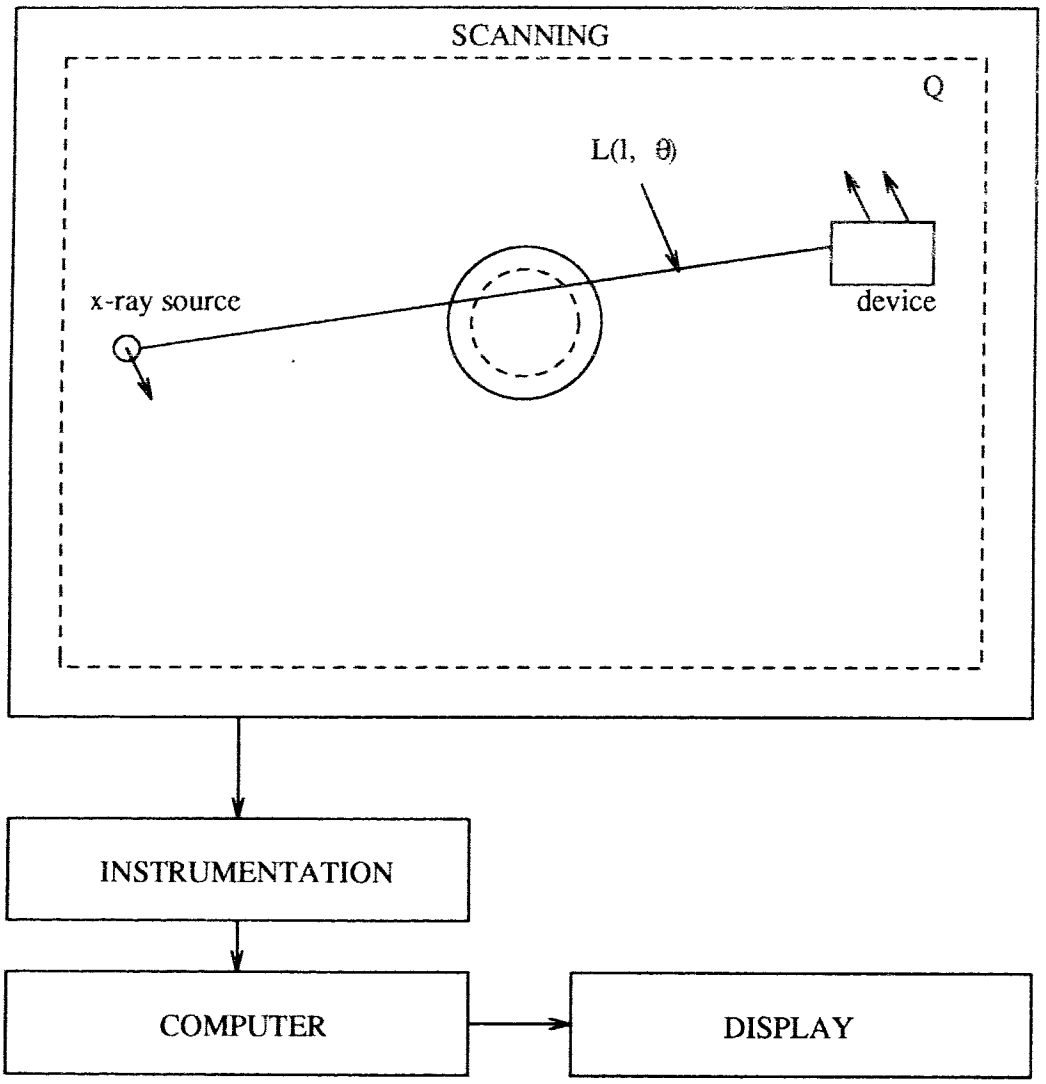

Figure 1.3: Medical tomography.

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{f(\xi,\eta)}{[(x-\xi)^2+(y-\eta)^2]^{\frac{1}{2}}}\, d\xi\, d\eta \;=\; S(x,y) := \frac{1}{\pi}\int_0^{\pi} \hat{g}(l,\theta)\, d\theta, \tag{1.7}$$

$$\int\int_G \hat{d}(x-\xi, y-\eta) w(\xi,\eta)\, d\xi\, d\eta \;=\; f(x,y), \tag{1.8}$$

where G is the domain of support for $w(x,y)$. Again ill-posed integral equations must be solved.

As an example of how an algorithm will inherit the instability of any one of these

equations, consider solving the 1-dimensional convolution equation

$$\begin{aligned} Kf &= \int_{-\infty}^{\infty} k(s-t)f(t)\,dt \\ &= g(s), \qquad -\infty < s < \infty, \end{aligned} \tag{1.9}$$

using Fourier transforms. Assume that

$$\begin{aligned} g(s) &\in \mathcal{L}_2(-\infty,\infty), \\ f(t) &\in \mathcal{L}_1(-\infty,\infty), \\ k(s) &\in \mathcal{L}_2(-\infty,\infty), \end{aligned}$$

and that

$$g(s) = g_T(s) + v(s),$$

where $g_T(s)$ is the exact right-hand side and $v(s)$ is an interference. Applying Fourier transforms to (1.9) we have

$$\hat{K}(\omega) * \hat{f}(\omega) = \hat{g}_T(\omega) + \hat{v}(\omega),$$

which implies that

$$\begin{aligned} \hat{f}(\omega) &= \frac{\hat{g}_T(\omega)}{\hat{K}(\omega)} + \frac{\hat{v}(\omega)}{\hat{K}(\omega)} \\ &= \hat{f}_T(\omega) + \frac{\hat{v}(\omega)}{\hat{K}(\omega)}. \end{aligned}$$

Hence,

$$\begin{aligned} f(t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(\omega)e^{i\omega t}\,d\omega \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}_T(\omega)e^{i\omega t}\,d\omega + \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\hat{v}(\omega)}{\hat{K}(\omega)}e^{i\omega t}\,d\omega \\ &= f_T(t) + \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\hat{v}(\omega)}{\hat{K}(\omega)}e^{i\omega t}\,d\omega. \end{aligned}$$

But the last integral can be arbitrarily large since in general, $K(\omega)$ and $v(\omega)$ approach zero as $|\omega| \to \infty$ at different rates. In particular, if the random function $v(\omega)$ contains high frequencies, the integral may even diverge.

Clearly there is a need for solving ill-posed problems, and as illustrated in the last example, there is a need for stable algorithms for solving them. Fortunately, stable algorithms for solving ill-posed problems exist. In this thesis, we examine some of these "regularization algorithms" for solving first kind operator equations with emphasis on solving linear first kind Fredholm integral equations. We then examine the possibility of applying multilevel processing to these algorithms, in order to reduce the computational cost yet obtain the same convergence rates.

Chapter 2

Theoretical Foundations

This section presents some of the fundamental theory and concepts used throughout this thesis. Section 2.1 reviews some basic elements of functional analysis, beginning with Hilbert spaces and ending with some spectral theory on Hilbert spaces. Most of the theorems in this section are merely stated. Section 2.2 elaborates several of these concepts for Fredholm integral equations. The emphasis here is on spectral theory, especially the Picard conditions for solvability of Fredholm integral equations of the first kind. The final section presents the rudiments of multigrid for problems with a continuous base. Multigrid for purely discrete models will not be presented.

2.1 Hilbert Spaces and Linear Operators

2.1.1 Hilbert Spaces

Definition 2.1 *A nonempty set $\mathcal{H}$ together with a real-valued (complex-valued) function $(\cdot,\cdot)$ from $\mathcal{H}\otimes\mathcal{H}$ into $\Re$ (C) is called a* Hilbert space *if $\mathcal{H}$ is a real (complex) vector space and $(\cdot,\cdot)$ satisfies the following properties:*

1. $(x,x)\geq 0$, *and* $(x,x)=0$ *if and only if* $x=0$;
2. $(x+y,z)=(x,z)+(y,z)$ *for all* $x,y,z\in\mathcal{H}$;
3. $(\lambda x,y)=\lambda(x,y)$ *for all* $x,y\in\mathcal{H},\lambda\in\Re$;

4. $(x, y) = (y, x)$ $((x, y) = \overline{(y, x)})$ *for all* $x, y \in \mathcal{H}$;

5. *if* $\{x_n\} \subset \mathcal{H}$,

$$\lim_{n,m\to\infty} (x_n - x_m, x_n - x_m) \longrightarrow 0,$$

then there exists an element $x \in \mathcal{H}$ *such that*

$$\lim_{n\to\infty} (x_n - x, x_n - x) = 0.$$

The function $(\cdot, \cdot)$ is called an inner product, and without property (5), $\mathcal{H}$ is called an inner product space.

If the norm $\|x\| = (x, x)^{\frac{1}{2}}$ is defined on $\mathcal{H}$ then the following theorem holds:

Theorem 2.1 (Schwarz Inequality) *In a Hilbert space* $\mathcal{H}$,

$$|(x, y)| \leq \|x\| \, \|y\| .$$

Using this theorem, it can be shown that a Hilbert space is a Banach space with the previously defined norm. Property (5) states the completeness property of $\mathcal{H}$.

The familiar examples of Hilbert spaces are

$$l_2 := \left\{ \{a_i\} : \sum_{i=1}^{\infty} |a_i|^2 < \infty \right\},$$

and

$$\mathcal{L}_2 := \left\{ f(t) : \int_a^b |f(t)|^2 \, dt < \infty \right\}$$

with the respective inner products

$$(a, b) = \sum_{i=1}^{\infty} a_i b_i , \qquad a = \{a_i\} \text{ and } b = \{b_i\} ,$$

and

$$(f, g) = \int_a^b f(t) g(t) \, dt.$$

The latter example will appear frequently in this thesis.

Hilbert spaces possess two properties that are fundamental in the theory of ill-posed problems. They concern the concepts of convexity and orthogonality:

Definition 2.2 *A subset C of a normed linear space is said to be convex if $[tx+(1-t)y] \in C$ for all $x, y \in C$ and all $t \in [0,1]$.*

Definition 2.3 *Let x, y be two points in an inner product space E. If $(x,y) = 0$, then x and y are said to be orthogonal. This relation is written as $x \perp y$. Let S be any subset of E. If x is orthogonal to each element of S, then x is said to be orthogonal to S, and this is written as $x \perp S$. The set of elements of E that are orthogonal to S is called the orthogonal complement of S, and is denoted by $S^{\perp}$ –*

$$S^{\perp} = \{ y \in E : x \perp y \text{ for all } x \in S \}. \tag{2.1}$$

The two properties are given in theorems 2.2 and 2.3 .

Theorem 2.2 *Let S be a closed convex set in a Hilbert space $\mathcal{H}$. For every point $x_0 \in \mathcal{H}$, there exists a unique point $y_0 \in S$ such that*

$$\|x_0 - y_0\| = \inf_{y \in S} \|x_0 - y\|.$$

Theorem 2.3 (Projection Theorem) *Let S be a closed linear subspace of a Hilbert space $\mathcal{H}$. Any $x_0 \in \mathcal{H}$ can be written in the form $x_0 = y_0 + z_0$, where $y_0 \in S$, $z_0 \in S^{\perp}$. The elements y_0, z_0 are uniquely determined by x_0.*

Proofs of these two theorems as well as most of the other theorems in this section can be found in any book of modern analysis. We note simply that using $x_0 = 0$ in theorem 2.2, S contains a unique element of minimal norm.

2.1.2 Linear Operators and Functionals

Let X and Y be two linear spaces, and let T be a function from a subset D_T of X into Y. T is called an operator, and D_T and $T(D_T) \subset Y$ are called respectively the domain and range of T. Furthermore, if D_T is a linear subspace of X, and if

$$T(\gamma_1 x_1 + \gamma_2 x_2) = \gamma_1 T x_1 + \gamma_2 T x_2, \qquad \gamma_1, \gamma_2 \text{ scalars},$$

then T is called a linear operator. Now assuming that $D_T = X$, the set of linear operators from X into Y can be made into a linear space by defining the operations

$$\begin{aligned}(T+S)x &= Tx + Sx \\ (\gamma T)x &= \gamma Tx.\end{aligned}$$

This space will be denoted by $\mathcal{L}(X,Y)$.

When X and Y are normed linear spaces, and $D_T = X$, the norm of T also can be defined. First, T is said to be bounded if there exists a constant $\tilde{C}$ such that

$$\|Tx\|_Y \leq \tilde{C}\,\|x\|_X$$

for all $x \in X$. T is said to be unbounded otherwise, and the subspace of all bounded operators in $\mathcal{L}(X,Y)$ is denoted by $\mathcal{B}(X,Y)$. The norm of T is

$$\|T\| := \sup_{x\neq 0} \frac{\|Tx\|_Y}{\|x\|_X} = \sup_{\|x\|=1} \|Tx\|_Y .$$

Trivially then, for $T \in \mathcal{B}(X,Y)$,

$$\|Tx\|_Y \leq \|T\|\,\|x\|_X .$$

Of course, bounded linear operators play a larger role in applications. One reason for this is that a bounded linear operator implies a continuous linear operator– i.e., a linear operator that satisfies the condition that if $x_n \to x$, then $Tx_n \to Tx$. The converse also holds. Another reason is that for an arbitrary $T \in \mathcal{B}(X,Y)$, a sequence of T_n's $\in \mathcal{B}(X,Y)$ that "approximate" T may exists. The following definition clarifies this:

Definition 2.4 *A sequence $\{T_n\}$ of bounded linear operators from a normed linear space X into a normed linear space Y is said to be uniformly convergent if there exists a bounded linear operator T from X into Y such that*

$$\|T_n - T\| \longrightarrow 0 \quad as\ n \to \infty.$$

$\{T_n\}$ is said to be uniformly convergent to T.

$\{T_n\}$ is said to converge strongly or pointwise to T if

$$\|T_n x - Tx\|_Y \longrightarrow 0 \quad as\ n \to \infty\ for\ any\ x \in X.$$

Now consider the case when X is a normed linear space and $Y = \Re$. Linear operators defined on X into $\Re$ (C) are called real (complex) linear functionals. These operators will be denoted with lower-case Roman letters superscripted with asterisks (e.g. x^*), and $\mathcal{B}(X, \Re)$ will be denoted by X^*. X^*, which is called the dual or conjugate space of X, is also a Banach space. Thus, the dual of X^* is defined. Moreover, to each element $x \in X$, a corresponding element $\hat{x}$ in the dual of X^* is given by $\hat{x}(x^*) = x^*(x)$ for all $x^* \in X^*$. This correspondence between X and X^{**} is called a natural imbedding of X into X^{**}.

Definition 2.5 *Let X be a Banach space. If the natural imbedding of X into X^{**} is equal to X^{**}, then X is called reflexive.*

Linear functionals introduce another form of convergence in a normed linear space X. Let $\{x_n\}$ be a sequence in X. $\{x_n\}$ is said to be weakly convergent if there exists an element $x \in X$ such that

$$\lim_{n\to\infty} x^*(x_n) = x^*(x)$$

for any $x^* \in X^*$. The element x is called the weak limit of $\{x_n\}$, and $\{x_n\}$ is said to converge weakly to x. Furthermore, a sequence $\{x_n\} \subset X$ is called a weak Cauchy sequence if $\{x^*(x_n)\}$ is a Cauchy sequence for any $x^* \in X^*$; the space X is said to be weakly complete if every weak Cauchy sequence has a weak limit; and a set $S \subset X$ is said to be weakly closed if the weak limit of every weakly convergent sequence in S is in S.

There is also a connection between bounded linear functionals defined on a Hilbert space $\mathcal{H}_1$ and the space $\mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$, where $\mathcal{H}_2$ is another Hilbert space. The connection is a result of the Riesz reprentation theorem:

Theorem 2.4 (Riesz Representation Theorem) *For every continuous linear functional x^* on $\mathcal{H}_1$, there exists a unique element $z \in \mathcal{H}_1$, called the representer of x^*, such that*

$$x^*(x) = (x, z)_1 \qquad \textit{for every } x \in \mathcal{H}_1$$

and

$$\|x^*\| = \|z\|.$$

Let $T \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ and $y \in \mathcal{H}_2$, and define the linear functional x^* by

$$x^*(x) = (Tx, y)_2.$$

Clearly x^* is continuous because

$$|x^*(x)| \leq (\|T\| \, \|y\|_2) \, \|x\|_1 .$$

Hence, by Riesz theorem, there exists a $z_y \in \mathcal{H}_1$ such that

$$x^*(x) = (x, z_y)_1 \; ;$$

hence, there exists a mapping $y \longrightarrow z_y$. Moreover, this mapping is not only linear, but it is also continuous since by taking $x = z_{y_1} - z_{y_2}$,

$$\begin{aligned} \|z_{y_1} - z_{y_2}\|_1^{\,2} &= (x, z_{y_1} - z_{y_2})_1 \\ &= (Tx, y_1 - y_2)_2 \\ &\leq \|T\| \, \|x\|_1 \|y_1 - y_2\|_2 , \end{aligned}$$

or

$$\|z_{y_1} - z_{y_2}\|_1 \leq \|T\| \, \|y_1 - y_2\|_2 .$$

If $y_1 \rightarrow y_2$, then $z_{y_1} \rightarrow z_{y_2}$.

Definition 2.6 *The mapping $y \longrightarrow z_{y_1}$ is called the adjoint operator of T, and is denoted T^*. It is the unique operator in $\mathcal{B}(\mathcal{H}_2, \mathcal{H}_1)$ satisfying*

$$(Tx, y)_2 = (x, T^* y)_1$$

for all $x \in \mathcal{H}_1$ and $y \in \mathcal{H}_2$.

When $\mathcal{H}_1 = \mathcal{H}_2$, T is said to be self-adjoint if $T = T^*$.

Example 1: Let K be the integral operator defined on the real $\mathcal{L}_2[a, b]$ space given by

$$Kf(s) = \int_a^b k(s,t) f(t) \, dt ,$$

where $k(s,t) \in \mathcal{L}_2([a,b] \otimes [a,b])$. K is a bounded linear operator for if $\|f\|_2 \leq 1$, then

$$\begin{aligned} |Kf(s)|^2 &= \left| \int_a^b k(s,t) f(t) \, dt \right|^2 \\ &\leq \int_a^b |k(s,t)|^2 \, dt \int_a^b |f(t)|^2 \, dt \\ &\leq \int_a^b |k(s,t)|^2 \, dt , \end{aligned}$$

and so

$$\|Kf\|_2 \leq \|k(\cdot,\cdot)\|_2 = \left[\int_a^b \int_a^b |k(s,t)|^2 \, ds \, dt\right]^{\frac{1}{2}} .$$

Now since the adjoint of K must satisfy

$$(Kf, g)_2 = (f, K^* g)_2 ,$$

$$\begin{aligned} \int_a^b \left[\int_a^b k(s,t) f(t) \, dt\right] g(s) \, ds &= \int_a^b \left[\int_a^b k(s,t) g(s) \, ds\right] f(t) \, dt \\ &= \int_a^b f(t) \left[K^* g(s)\right] dt \end{aligned}$$

for all $f(t) \in \mathcal{L}_2[a,b]$. Thus

$$K^* g(s) = \int_a^b k(s,t) g(s) \, ds .$$

Example 2: Let S be a closed subspace of a Hilbert space $\mathcal{H}$. By the projection theorem, any $x \in \mathcal{H}$ can be written as

$$x = x_1 + x_2 \in S \oplus S^{\perp} .$$

Now consider the operator $P_S \in \mathcal{B}(\mathcal{H}, S)$ defined by

$$P_S x = x_1 .$$

P_S is called the orthogonal projection operator of $\mathcal{H}$ onto S. It is self-adjoint and satisfies the properties

$$\begin{aligned} P_S &= {P_S}^2 , \\ \|P_S\| &= 1 , \\ \|x - P_S x\| &= \inf_{y \in S} \{\|x - y\|\} . \end{aligned}$$

Several useful properties of the adjoint operator are

$$\begin{aligned} (T_1 T_2)^* &= {T_2}^* {T_1}^* \\ (T^*)^* &= T \\ (T_1 + T_2)^* &= {T_1}^* + {T_2}^* \\ (\gamma T)^* &= \gamma T^* , \qquad \gamma \in \Re \\ \|T^*\| &= \|T\| \\ \|T T^*\| &= \|T^* T\| = \|T\|^2 . \end{aligned}$$

A crucial theorem involving T^* is theorem 2.5.

Theorem 2.5 *If $T \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$, then*

$$\begin{aligned} N(T)^\perp &= \overline{R(T^*)} \qquad R(T)^\perp = N(T^*) \\ N(T^*)^\perp &= \overline{R(T)} \qquad R(T^*)^\perp = N(T). \end{aligned}$$

Here, $N(T)$ is the null space of T defined by

$$N(T) = \{x : Tx = 0\},$$

and $R(T)$ is the range of T.

2.1.3 Compact Operators

The importance of this class of operators warrants a separate subsection. Compact operators play a central role in the theory of linear integral equations and the theory of ill-posed problems. Before examining these operators, however, we define a compact set.

Definition 2.7 *Let (X, ρ) be a metric space. X is said to be compact if every sequence in X contains a convergent subsequence. Equivalently, X is compact if X is complete, and for every $\epsilon > 0$, X can be covered with a finite number of open spheres of radius ϵ.*

Definition 2.8 *Let $T \in \mathcal{B}(X, Y)$ where X and Y are normed linear spaces. If $\overline{T(B)}$ is compact in Y for each bounded set $B \subset X$, then T is called a compact operator.*

Next, we give a collection of important facts concerning compact operators.

Definition 2.9 *An operator is said to be finite rank if its range is finite-dimensional.*

Trivially, a finite rank operator is compact.

Theorem 2.6 *A compact linear operator maps weakly convergent sequences into convergent sequences.*

Theorem 2.7 *Let X be a normed linear space. If $T_1, T_2 \in \mathcal{B}(X, X)$ and T_1 is compact, then T_1T_2 and T_2T_1 are compact.*

Theorem 2.8 *Let X be a normed linear space, and let Y be a Banach space. An operator $T \in \mathcal{B}(X,Y)$ is compact if and only if its adjoint is compact.*

Theorem 2.9 *Let X be a normed linear space, let Y be a Banach space, and let $T \in \mathcal{B}(X,Y)$. If $\|T - T_n\| \to 0$ as $n \to \infty$, where each $T_n \in \mathcal{B}(X,Y)$ is compact, then T is compact.*

Using the fact that every Hilbert space contains a complete orthonormal system, and using theorem 2.9, it can be shown that the integral operator of example (1) is compact.

2.1.4 Spectral Theory

Two of the most common operator equations are

$$\lambda x - Kx = y \qquad \lambda \text{ a nonzero scalar,} \tag{2.2}$$

$$Kx = y, \tag{2.3}$$

of which the Fredholm integral equations of the second and first kinds are respective examples of. The setting is $x, y \in \mathcal{H}$, a Hilbert space, and $K \in \mathcal{B}(\mathcal{H},\mathcal{H})$ and is compact. The problem is to find x given y and K. Equivalently, assuming that $(\lambda I - K)^{-1}$ and K^{-1} exist, the problem is to find expressible forms of these inverses.

Another problem is deriving conditions for the existence of a solution. Necessary and sufficient conditions for the existence of a solution for arbitrary right-hand terms are a crucial part of the theory of first and second kind operator equations. As a preparation for the study of this and the previous problem for the more difficult first kind equation, we consider these problems only for equation (2.2) in this subsection.

Define the sets N_λ, $N_\lambda{}^*$, R_λ, and $R_\lambda{}^*$ to be

$$N_\lambda = \{x : \lambda x - Kx = 0\}, \qquad N_\lambda{}^* = \{z : \overline{\lambda} z - K^* z = 0\},$$

$$R_\lambda = \{y : y = \lambda x - Kx, x \in \mathcal{H}\}, \qquad R_\lambda{}^* = \{w : w = \overline{\lambda} z - K^* z, z \in \mathcal{H}\}.$$

It can be shown that N_λ and $N_\lambda{}^*$ are finite-dimensional, and that R_λ and $R_\lambda{}^*$ are closed. A simple application of theorem 2.5 then verifies that

Theorem 2.10 *(a) $R_\lambda = {N_\lambda}^{*\perp}$ - i.e., the equation $\lambda x - Kx = y$ has a solution if and only if y lies in the orthogonal complement of ${N_\lambda}^*$.*
(b) ${R_\lambda}^* = N_\lambda{}^\perp$ - i.e., the equation $\overline{\lambda} z - K^* z = w$ has a solution if and only if w lies in the orthogonal complement of N_λ.

In fact, the following stronger theorem holds:

Theorem 2.11 *For any $\lambda \neq 0$,*

$$\begin{aligned} \dim N_{\lambda I - K} &= \dim N_{\overline{\lambda} I - K^*} < \infty , \\ R_{\lambda I - K} &= {N_{\overline{\lambda} I - K^*}}^{\perp} , \qquad R_{\overline{\lambda} I - K^*} = {N_{\lambda I - K}}^{\perp} . \end{aligned}$$

Here, the notation of the subspaces has been changed to stress the variability of λ.

Theorem 2.11 is essentially the Fredholm alternative. It states that either

1. for any $y \in \mathcal{H}$, there exists a unique solution of $(\lambda I - K)x = y$; or
2. there exists a nonzero solution of $(\overline{\lambda} I - K^*)z = 0$.

Condition (1) holds if dim $N_{\lambda - K} = \dim N_{\overline{\lambda} - K^*} = 0$. If (2) holds, then dim $N_{\lambda - K} = \dim N_{\overline{\lambda} - K^*}$ is finite, and the equation $(\lambda I - K)x = y$ has a "general solution" if and only if y is in the orthogonal complement of $N_{\overline{\lambda} - K^*}$. Similar statements hold for the adjoint equation $(\overline{\lambda} I - K^*)z = w$.

Actually the Fredholm alternative emphasizes more on the parameter λ.

Definition 2.10 *The resolvent set $\rho(K)$ of K consists of all complex numbers λ for which $(\lambda I - K)^{-1}$ is a bounded operator. For these values, the operator*

$$R(\lambda; K) = (\lambda I - K)^{-1}$$

is called the resolvent of K, and λ is called a regular value of K.

One way to represent the resolvent of K is to take a geometric series expansion for $(\lambda I - K)^{-1}$,

$$(\lambda I - K)^{-1} = \sum_{n=0}^{\infty} \frac{1}{\lambda^{n+1}} K^n .$$

However, this is valid only for

$$|\lambda| > |\sigma(K)| := \sup_{\lambda \in \sigma(K)} \{|\lambda|\} ,$$

where $\sigma(K)$ is defined below.

Definition 2.11 *The spectrum $\sigma(K)$ of K consists of the complement of $\rho(K)$ in $\mathcal{C}$. These values fall into three sets:*

1. *continuous spectrum- the range of $(\lambda I - K)$ is dense in $\mathcal{H}$ and $(\lambda I - K)^{-1}$ exists but is unbounded;*
2. *residual spectrum- $(\lambda I - K)^{-1}$ exists, but its domain is not dense in $\mathcal{H}$; and*
3. *eigenvalues- $(\lambda I - K)$ does not have an inverse; i.e., there exists a nonzero x such that $(\lambda I - K)x = 0$.*

The Fredholm alternative states that for nonzero λ, λ is either an eigenvalue or a regular value of K. Note that in an infinite-dimensional space, for $\lambda = 0$, $(\lambda I - K)^{-1} = -K^{-1}$ cannot be bounded since otherwise $KK^{-1} = I$ would be a compact operator. Hence, $0 \in \sigma(K)$.

Theorem 2.12 *Let K be a compact linear operator in an infinite-dimensional Banach space. Then $\sigma(K)$ consists of 0 and either a finite number of eigenvalues or an infinite sequence of eigenvalues that converges to 0.*

2.1.5 Spectral Family of Self-Adjoint Operators

Now let K be a self-adjoint operator belonging to $\mathcal{B}(\mathcal{H}, \mathcal{H})$ which need not be compact. The spectrum of K simplifies, and an elegant spectral theory can be developed.

Definition 2.12 *A self-adjoint operator $K \in \mathcal{B}(\mathcal{H}, \mathcal{H})$, where $\mathcal{H}$ is a Hilbert space is said to be positive if*

$$(Kx, x) \geq 0 \qquad \text{for all } x \in \mathcal{H}.$$

If K is positive, we write $K \geq 0$.

For two self-adjoint operators K and T, the notation $K \geq T$ will be used if $(K - T) \geq 0$.

Theorem 2.13 *The spectrum of a self-adjoint operator K is contained in the real line, and if a and b are real numbers satisfying $aI \leq K \leq bI$, then $\sigma(K) \subset [a,b]$. In fact, with*

$$\mathrm{m} := \inf_{\|x\|=1} (Kx,x) \quad \text{and} \quad \mathrm{M} := \sup_{\|x\|=1} (Kx,x),$$

$\sigma(K) \subset [\mathrm{m},\mathrm{M}]$.

Theorem 2.14 *If $K \in \mathcal{B}(\mathcal{H},\mathcal{H})$ is self-adjoint, then*

$$\|K\| = |\sigma(K)| .$$

$|\sigma(K)|$ *is called the spectral radius of K.*

Now consider the linear operator T defined on $\Re^n$. Suppose that the matrix representation of T is real and symmetric, and has n distinct eigenvalues λ_i which are indexed according to increasing magnitude: $\lambda_i < \lambda_{i+1}$. Let $\{u_i\}$ be the orthonormal basis of eigenvectors corresponding to $\{\lambda_i\}$. The spectral familty of T is the family of projection-valued functions $\{E_\lambda\}$ that satisfies

$$\begin{aligned} E_\lambda &= 0 \quad \lambda \leq \lambda_1 \\ E_\lambda &= P_{S_i} \quad \lambda_i < \lambda \leq \lambda_{i+1} \\ E_\lambda &= I \quad \lambda_n < \lambda \end{aligned}$$

where S_i is the subspace spanned by $\{u_j\}_{j=1}^{i}$, and P_{S_i} is the projection onto S_i. Defining the projection $\Delta_i E$ by

$$\begin{aligned} \Delta_1 E &= P_{S_1} \\ \Delta_i E x &= (E_{\lambda_i} - E_{\lambda_{i-1}})x = (x,u_i)u_i \quad i = 2,\ldots,n, \end{aligned}$$

then

$$Tx = \sum_{i=1}^{n} \lambda_i (x,u_i)u_i = \sum_{i=1}^{n} \lambda_i \Delta_i E x .$$

This concept can be generalized to include self-adjoint operators defined on infinite-dimensional Hilbert spaces. A spectral family generated by a self-adjoint operator K is a family of projections $\{E_\lambda\}$ for $\lambda \in \Re$ satisfying

1. $E_\lambda \leq E_\mu$ if $\lambda < \mu$;

2. E_λ is strongly continuous from the left, i.e.

$$\lim_{\lambda \nearrow \mu} E_\lambda x = E_\mu x \quad \text{for any } x \in \mathcal{H};$$

3. $E_\lambda = 0$ if $\lambda <$ m, $E_\lambda = I$ if $\lambda >$ M.

Let $f(\lambda)$ be a continuous complex-valued function for m $\leq \lambda \leq$ M. Assume that $f(\lambda)$ can be extended to M $< \lambda <$ M + 1 so that the extended function, which we denote by $f(\lambda)$ again, is continuous on [m,M+1). The integral

$$\int_m^{M+\epsilon} f(\lambda)\, dE_\lambda \qquad (0 < \epsilon < 1)$$

can be defined:

Let Π be a partition of [m,M+ ϵ]:

$$\mathrm{m} = \lambda_0 < \lambda_1 < \ldots < \lambda_n = \mathrm{M} + \epsilon.$$

That is, it is a partition of an interval containing $\sigma(K)$. Let

$$\Delta_i E = E_{\lambda_i} - E_{\lambda_{i-1}} \quad \text{and} \quad \Delta_i = [\lambda_{i-1}, \lambda_i].$$

The integral is the limit of the sum

$$\sum_{i=1}^{n} f(\mu_i)\Delta_i E$$

as $|\pi| := \max_i(\lambda_i - \lambda_{i-1})$ approaches zero, where $\mu_i \in \Delta_i$.

Theorem 2.15 *To every self-adjoint operator K there corresponds a unique family $\{E_\lambda\}$ of projections, $-\infty < \lambda < \infty$, satisfying*

1. *any bounded linear operator that commutes with K commutes with E_λ;*
2. *$E_\lambda \leq E_\mu$ if $\lambda < \mu$;*
3. *E_λ is strongly continuous form the left;*
4. *$E_\lambda = 0$ if $\lambda < m$, $E_\lambda = I$ if $\lambda > M$;*

5.

$$K = \int_m^{M+\epsilon} \lambda \, dE_\lambda \qquad (0 < \epsilon < 1) .$$

Next,

Theorem 2.16 *Let K be a self-adjoint operator, and let $\{E_\lambda\}$ satisfy (1)-(5) of theorem (2.15). Then for any polynomial $p(\lambda)$,*

$$p(K) = \int_m^{M+\epsilon} p(\lambda) \, dE_\lambda .$$

Theorem 2.17 *For any polynomial $p(\lambda)$ and any $x \in \mathcal{H}$,*

$$(p(K)x, x) = \int_m^{M+\epsilon} p(\lambda) \, d(E_\lambda x, x) .$$

Moreover, if $f(\lambda)$ is any continuous function for m $\leq \lambda \leq$ M, $f(K)$ is defined to be

$$f(K) = \int_m^{M+\epsilon} f(\lambda) \, dE_\lambda .$$

Theorems 2.16 and 2.17 also hold:

$$\begin{aligned} f(K)x &= \int_m^{M+\epsilon} f(\lambda) \, dE_\lambda x \\ (f(K)x, x) &= \int_m^{M+\epsilon} f(\lambda) \, d(E_\lambda x, x) . \end{aligned}$$

The significance of this functional representation is illustrated in the analysis and construction of linear iterative methods. For example, the inverse of an invertible positive operator K can be approximated by a sequence of operators $\{f_n(K)\}$, where each f_n is continuous in an interval containing $\sigma(K)$, and where $\{f_n\}$ converges uniformly to λ^{-1} on this interval.

Finally, assume that K is a compact self-adjoint operator. From a previous discussion, the spectrum of K consists of zero and a finite number of eigenvalues or an infinite sequence of eigenvalues that converges to zero. Also, the eigenspace corresponding to each nonzero eigenvalue is finite-dimensional. We have

Theorem 2.18 *Let $\lambda_1, \lambda_2, \ldots$ be the eigenvalues of K which are repeated according to the dimension of the associated eigenspace, and let $u_1, u_2, \ldots$ be the corresponding orthonormal eigenvectors. Then for any $x \in \mathcal{H}$,*

$$Kx = \sum_n \lambda_n (x, u_n) u_n \,.$$

Furthermore, given any real-valued function f defined on $\sigma(K)$, $f(K)$ is given by

$$f(K) = \sum_n f(\lambda_n)(x, u_n) u_n \,,$$

and the following theorems hold:

Theorem 2.19 *Let K be a compact self-adjoint operator, and let f be a continuous real-valued function defined on $\sigma(K)$. Then*

$$\sigma(f(K)) = f(\sigma(K)) \,.$$

Theorem 2.20 *Let K be a compact self-adjoint operator, let f be a continuous real-valued function defined on $\sigma(K)$, and let $\{f_n\}$ be a sequence of continuous real-valued functions which converges to f uniformly on $\sigma(K)$. Then*

$$\|f_n(K) - f(K)\| \to 0 \,.$$

2.2 Integral Equations

Several examples of integral equations, and the crux of the theory of linear Fredholm integral equations of the second kind have been given already. The objective of this section is to develop the theory of Fredholm integral equations of the first kind.

Consider the following generic Fredholm integral equation of the first kind:

$$\int_a^b k(s,t) f(t)\, dt = g(s), \qquad a \le s \le b \tag{2.4}$$

or

$$Kf = g \,, \tag{2.5}$$

where K is the integral operator defined on $\mathcal{L}_2[a,b]$ generated by the kernel $k(s,t) \in \mathcal{L}_2([a,b] \otimes [a,b])$. The difficulty of this problem stems from the fact that $N(K)$ may be non-trivial, and more importantly, from the fact that K^{-1} may not be continuous even though K^{-1} exists. The latter difficulty implies that small changes in the data $g(s)$ may cause dramatic changes in the solution. For example, with $a = 0$ and $b = \pi$, using the Riemann–Lebesgue lemma, for sufficiently large n, the slight perturbation of

$$A \int_0^{\pi} k(s,t) sin(nt)\, dt$$

to $g(s)$ will cause a change of $A\, sin(nt)$ to the solution of the original problem. This change to the solution can be quite large.

More insight into this instability problem is revealed through the singular value decomposition of the operator. Since K is compact, K^*K is compact, and since K^*K is positive, all the eigenvalues of K^*K are nonnegative:

$$0 \leq (K^*Kx, x) = (ex, x) = e\,,$$

where e is an eigenvalue and x is the corresponding orthonormal eigenvector. Thus the eigenvalues can be represented as

$$\lambda_1{}^2 \leq \lambda_2{}^2 \leq \ldots .$$

Now if the associated orthonormal eigenvector system is denoted by $\{v_n\}$, and if μ_n is defined by

$$\mu_n = \lambda_n{}^{-1}$$

and u_n is defined by

$$u_n = \mu_n K v_n\,,$$

then the set $\{u_n, v_n; \mu_n\}$ is called the singular system of K. The μ_n's[1] are called the singular values of K, and the v_n's and u_n's, which form bases for $\overline{R(K^*)}$ and $\overline{R(K)}$ respectively, are called the singular functions of K.

(Observe that the relation

$$\mu_n K^* u_n = v_n \tag{2.6}$$

holds.)

[1]Some authors denote the singular values with λ_n.

Theorem 2.21 *Equation (2.5) has a solution if and only if*

1. $g \in N(K^*)^\perp$, *and*
2. $$\sum_{n=1}^{\infty} \lambda_n{}^{-2} |(g, u_n)|^2 = \sum_{n=1}^{\infty} \mu_n{}^2 |(g, u_n)|^2 < \infty .$$

Under these conditions, the solution is

$$f = \sum_{n=1}^{\infty} \mu_n (g, u_n) v_n .$$

Theorem 2.21 states the Picard conditions for the solvability of first kind integral equations. Note that since $\mathcal{H} = N(K^*) \oplus N(K^*)^\perp = N(K^*) \oplus \overline{R(K)}$, the second condition "transfers" g from $\overline{R(K)}$ to $R(K)$. Also note that if g is perturbed by δg, then (2.) for $\tilde{g} = g + \delta g$ may not hold, or even if it does, the series

$$\sum_{n=1}^{\infty} \mu_n (\delta g, u_n) v_n$$

may be noticeable. The first misfortune is elaborated in the next theorem.

Theorem 2.22 *Suppose $\mu_n \to \infty$ as $n \to \infty$. Let $g \in R(K)$ and $\epsilon > 0$ be arbitrary. Then there exists a function $g' \in \mathcal{L}_2$ such that*

$$\begin{aligned} g' &\notin R(K), \\ g' &\in \overline{R(K)}, \\ \|g - g'\| &< \epsilon . \end{aligned}$$

Proof: Since $\mu_n \to \infty$, a sequence of real numbers $\{a_n\}$ can be constructed such that $\sum_{n=1}^{\infty} |a_n|^2 < \epsilon^2$ but $\sum_{n=1}^{\infty} \mu_n{}^2 a_n{}^2$ diverges. The Riesz–Fischer theorem [2] shows that

$$g'' := \sum_{n=1}^{\infty} a_n u_n$$

[2] Riesz–Fischer Theorem: Given an enumberable orthonormal system $\{e_n\}$ in $\mathcal{L}_2$, and a sequence of complex numbers $\{\hat{a}_n\}$, a necessary and sufficient condition for the series $\sum_{n=1}^{\infty} \hat{a}_n e_n$ to be convergent in the mean is that $\sum_{n=1}^{\infty} |\hat{a}_n|^2 < \infty$.

belongs in $\mathcal{L}_2$. Clearly g'' does not satisfy condition (*2.*) and so by the Picard conditions, $g'' \in \overline{R(K)}$, $g'' \notin R(K)$. The $\mathcal{L}_2$-function

$$g' = g + g''$$

also belongs in $\overline{R(K)}$ but not in $R(K)$. It satisfies $\|g - g'\| < \epsilon$ by Parseval's equality.

It is obvious that the existence of g' is guaranteed by the divergence of μ_n. Hence, if $\mu_n \to \infty$, then $R(K)$ is not closed. Moreover, suppose $R(K) \neq \overline{R(K)}$. Then since the range of a compact operator is closed if and only the operator is of finite rank, $R(K)$ must be infinite-dimensional. Using relation (2.6), the set $\{v_n\}$ must be infinite, and so the set $\{{\lambda_n}^2\}$ is infinite (recall that the eigenspace of each eigenvalue of a compact self-adjoint operator is finite-dimensional). Since ${\lambda_n}^2 \to 0$, $\mu_n \to \infty$.

But whether $R(K)$ is closed or not, solving first kind Fredholm integral equations numerically is usually difficult. Even in the absence of data error, the unavoidable error created by the discretization of the equation implies that a perturbed equation is always solved. Thus, only the presence of some large singular values may cause a noticeable difference between the solutions of the perturbed equation and the unperturbed equation. Hence, solving first kind Fredholm integral equations using standard methods for solving second kind Fredholm equations is not advised.

2.3 Multigrid

Multigrid is a fast iterative technique which has not only been extremely successful in solving partial differential equations, but is also being used to solve second kind integral equations. Presently multigrid– algebraic multigrid – is being applied to first kind equations derived from tomography problems (Brandt (1989), Szeliski and Terzopoulus (1989)). The difficulties in applying multigrid to first kind equations are however severe. One cannot blindly apply multigrid to a standard iterative method for solving ill-posed equations since the errors introduced by the multigrid scheme may not be in the range of the operator. There is also the basic question of whether multilevel processing is more efficient than a standard procedure in solving ill-posed equations. These are some of the severe difficulties which shall be examined lightly later. Right now, we present the elements of multigrid.

The backbone of standard multigrid methods is the continuous base of the problem. Working on several scales of discretization and exploiting the continuous base of the problem, these methods damp out the error components of an approximate solution much faster than a basic iterative scheme does. To illustrate this, consider the 1-dimensional boundary-valued problem

$$\begin{aligned} \mathcal{A}u(x) &= f \quad \text{in } \Omega = \{x : 0 < x < 1\} \\ u(0) = u(1) &= 0\,, \end{aligned} \tag{2.7}$$

where $\mathcal{A} = -\frac{d^2}{dx^2}$.

Let the interval Ω be divided into $(n^h + 1)$ subintervals each of length $h = \frac{1}{(n^h+1)}$, and let $x_i^h = ih$, $i = 1, 2, \ldots, n^h$, be the interior grid points. The set of grid points will be denoted by Ω^h. Approximating the derivative using a centered difference scheme gives

$$\frac{-u(x_{i+1}^h) + 2u(x_i^h) - u(x_{i-1}^h)}{h^2} + \tau_i = f(x_i^h)\,, \tag{2.8}$$

where $\tau_i = \frac{h^2}{12}u^{(4)}(\overline{x}_i^h)$, $x_{i-1}^h < \overline{x}_i^h < x_{i+1}^h$, is the local truncation error. Dropping the truncation error, (2.8) converts into a set of difference equations

$$\frac{-u_{i+1} + 2u_i - u_{i-1}}{h^2} = f_i\,, \qquad 1 \le i \le n^h\,, \tag{2.9}$$

where the u_i's and f_i's are components of the grid functions $\underline{u}^h, \underline{f}^h \in \Re^{n^h}$. (2.9) in matrix notation is

$$A_h \underline{u}^h = \underline{f}\,. \tag{2.10}$$

This system can be solved using, for example, the damped Jacobi or the Gauss-Seidel iteration.

The solution of (2.10) will not satisfy (2.8) exactly. Hence, we define the global error to be

$$\underline{e}_g^h = \underline{u}^h - \underline{w}^h\,,$$

where $\underline{w}^h$ is the vector with components $u(x_i^h)$.

Another quantity is the algebraic error. Note that an iterative process is performed to "solve" (2.10), and hence all that is available is an approximation $\underline{v}^h$ to $\underline{u}^h$. The algebraic error is

$$\underline{e}_a^h := \underline{v}^h - \underline{u}^h\,.$$

A simple application of the triangle inequality confirms the intuition that the magnitude of $\left\|\underline{v}^h - \underline{w}^h\right\|$ depends on both the the number of iterations of the iterative procedure (algebraic error) and the size of h (global error).

Now the quantity that is of interest to us is the algebraic error. To see the effect a linear iterative method has on this error, let the iteration be written as

$$(\underline{v}^h)^i = M_h(\underline{v}^h)^{i-1} + N_h\underline{f}^h\,,$$

where M_h is called the iteration matrix. The i'th iterate can be expressed also as

$$(\underline{v}^h)^i = (M_h)^i(\underline{v}^h)^{i-1} + N_h^{(i)}\underline{f}^{(h)}\,,$$

where $N_h^{(i)} = \sum_{k=0}^{i-1} M_h^k N_h$. Obviously the solution of (2.10) must be a fixed point of the iteration, and consequently

$$(\underline{e}_a^h)^i = M_h(\underline{e}_a^h)^{i-1} \quad \text{or} \quad (\underline{e}_a^h)^i = (M_h)^i(\underline{e}_a^h)^0\,.$$

Hence, the iteration will converge for all $(\underline{v}^h)^0$ if and only if the spectral radius of M_h is less than one.

But even when the spectral radius of M_h is less than one, the convergence may be slow due to the slow attentuation of some error components in the initial approximation. For example, consider the damped Jacobi iteration. Decomposing A_h to $(D_h - B_h)$ where $D_h := 2h^2 I$ and $B_h := D_h - A_h$, the damped Jacobi iteration is

$$(\underline{v}^h)^i = (\underline{v}^h)^{i-1} - \omega h^2[A_h(\underline{v}^h)^{i-1} - \underline{f}^h] \quad \text{with } \omega \in \left(0, \frac{1}{2}\right]\,. \tag{2.11}$$

It has the iteration matrix

$$M_h = I - \omega h^2 A_h$$

with eigenvalues

$$\lambda_j(\omega) = 1 - 4\omega\sin^2\left(\frac{j\pi h}{2}\right) \qquad 1 \le j \le n^h$$

and corresponding eigenvectors

$$\underline{ew}^h{}_j = \sqrt{2h}\,[\sin(kj\pi h)]_{k=1}^{n^h} \qquad 1 \le j \le n^h\,.$$

Next, classify the subspaces $span\left\{\underline{ew^h}_j : \frac{1}{2} \leq jh < 1\right\}$ and $span\left\{\underline{ew^h}_j : 0 < jh < \frac{1}{2}\right\}$ as the high frequency set and low frequency set respectively. Then fixing ω to $\frac{1}{2}$, since $\lambda_1\left(\frac{1}{2}\right) \approx 1 - \frac{\pi^2 h^2}{2}$ and $\lambda_{n_h}\left(\frac{1}{2}\right) \approx -1 + \frac{\pi^2 h^2}{2}$, given the initial error

$$\left(\underline{e}_a^h\right)^0 = \sum_{j=1}^{n^h} \alpha_j \underline{ew^h}_j \,, \tag{2.12}$$

the error after ν iterations is

$$\left(\underline{e}_a^h\right)^\nu = \sum_{j=1}^{n^h} \alpha_j \left(\lambda_k\left(\frac{1}{2}\right)\right)^\nu \underline{ew^h}_j \,, \tag{2.13}$$

implying that the highest and lowest frequency components of the initial error have not been damped down dramatically. Fixing ω to $\frac{1}{4}$, although the spectral radius of M_h now is larger than the spectral radius of M_h when $\omega = \frac{1}{2}$, the high frequencies are damped down strongly after ν iterations. Thus, the new error is smooth.

2.3.1 Two Grid Method

The last statement hints on another iterative approach to solving (2.10). From (2.11), it is clear that the correction of $(\underline{v}_h)^{i-1}$ is obtained from the defect $\underline{d}_h := [A_h(\underline{v}_h)^{i-1} - \underline{f}_h]$. A rearrangement of the defect together with (2.10) gives the defect equation

$$A_h \underline{e}_a^h = \underline{d}^h \,, \tag{2.14}$$

which has a smooth solution. (2.14) is of the same form as (2.10). (2.14) also can be considered as a discretization of the differential equation

$$\mathcal{A}e_a = d \,, \qquad e_a(0) = e_a(1) = 0 \qquad \text{in } \Omega = \{x : 0 < x < 1\} \tag{2.15}$$

where e_a and d are the functions $\underline{e}_a^h$ and $\underline{d}^h$ approximate. This can be seen by considering $\underline{v}_h^i$ as the approximation to $v(x)$, the solution of the boundary-valued problem

$$\begin{aligned} \mathcal{A}v &= f + d \\ v(0) = v(1) &= 0 \,. \end{aligned} \tag{2.16}$$

Subtracting (2.7) from (2.16) gives

$$\begin{aligned} \mathcal{A}(v-u) &= \mathcal{A}e_a = d \\ e_a(0) = e_a(1) &= 0\,. \end{aligned} \tag{2.17}$$

The important fact is that e_a is smooth. Thus (2.15) can be discretized on a coarser grid, which we conveniently take to be Ω^H with $H = 2h$:

$$A_H \underline{e}_a^H = \underline{d}^H\,. \tag{2.18}$$

But since only $\underline{d}^h$ is available, the right-hand side of (2.18) poses a problem. To define $\underline{d}^H$, a restriction operator $I_h{}^H$ which maps grid functions defined on Ω^h onto grid functions defined on Ω^H is needed. Assuming that $I_h{}^H$ has been defined and (2.18) has been solved, a new approximation is

$$(\underline{v}_h)^i \longleftarrow (\underline{v}_h)^i - I_H{}^h \underline{e}_a^H\,, \tag{2.19}$$

where $I_H{}^h$ is an interpolation operator from $\Re^{n^H}$ into $\Re^{n^h}$. (2.19) is called a coarse-grid correction. It may also be written as

$$(\underline{v}^h)^i \longleftarrow (\underline{v}^h)^i - I_H{}^h A_H{}^{-1} I_h{}^H \left[A_h(\underline{v}^h)^i - \underline{f}^h\right]\,. \tag{2.20}$$

Turning to the operators $I_h{}^H$ and $I_H{}^h$, note that since $(\underline{e}_a^h)^\nu$ is smooth, $\underline{d}^h$ is also smooth:

$$\begin{aligned} \underline{d}^h &= A_h(\underline{v}^h)^\nu - \underline{f}^h \\ &= A_h\left[(\underline{e}_a^h)^\nu + \underline{u}^h\right] - \underline{f}^h \\ &= A_h\left[\sum_{j=1}^{n^h} \alpha_j \left(\lambda_j\left(\frac{1}{4}\right)\right)^\nu \underline{ew}^h{}_j\right] \\ &= \frac{4}{h^2}\sum_{j=1}^{n^h} \alpha_j \left(\lambda_j\left(\frac{1}{4}\right)\right)^\nu \sin^2\left(\frac{j\pi h}{2}\right) \underline{ew}^h{}_j\,, \end{aligned} \tag{2.21}$$

which is smooth because for high frequencies, $\left(\lambda_j\left(\frac{1}{4}\right)\right)^\nu \ll 1$ for ν sufficiently large. An appropriate restriction operator is then

$$\left[I_h{}^H \underline{d}^h\right](x) = \frac{1}{4}\left[\underline{d}^h(x-h) + 2\underline{d}^h(x) + \underline{d}^h(x+h)\right] \qquad \text{for } x \in \Omega^h\,. \tag{2.22}$$

The trivial injection

$$\left[I_h{}^H \underline{d}^h\right](x) = \underline{d}(x) \qquad \text{for all } x \in \Omega^H \subset \Omega^h \tag{2.23}$$

is another good choice. As for the operator $I_H{}^h$, note that the smoothness of $\underline{e}_a^h$ implies that

$$\underline{e}_a^h \approx I_H{}^h \underline{e}_a^h$$

if $I_H{}^h$ is taken to be the interpolant

$$\left[I_H{}^h \underline{e}_a^h\right](x) = \begin{cases} \underline{e}_a^H(x) & \text{if } x \in \Omega^H \\ \left[\underline{e}_a^H(x-h) + \underline{e}_a^H(x+h)\right]/2 & \text{otherwise.} \end{cases} \tag{2.24}$$

But even with the correct choices of operators $I_h{}^H$ and $I_H{}^h$, it is still unclear how the coarse-grid correction helps in obtaining an accurate solution. Iteration (2.20) by itself is in fact non-convergent; for take a nonzero element $\underline{g}^h \in N(I_h{}^H)$ and let

$$(\underline{v}^h)^0 = A_h{}^{-1}\left(\underline{f}^h + \underline{g}^h\right)$$

so that $\underline{d}^h = A_h(\underline{v}^h)^0 - \underline{f}^h = \underline{g}^h$, $\underline{d}^H = I_h{}^H \underline{g}^h = 0$, and

$$(\underline{v}^h)^\nu = (\underline{v}^h)^0 \not\to A_h{}^{-1}\underline{f}^h.$$

It is the combination of smoothing iteration and coarse-grid correction that results in a rapidly convergent scheme. Its steps are

1. pre-smooth ν_1 times on the fine grid;
2. restrict the defect onto the coarse grid;
3. solve the defect equation;
4. interpolate the solution of the defect equation;
5. post-smooth ν_2 times with the new approximation.

Here, $\nu_1, \nu_2 \geq 0$ but both cannot be zero.

Theorem 2.23 *The iteration matrix of a two-grid method using a smoother S_h and pre-smoothing ν times is*

$$M_h = M_h(\nu) = \left(I - I_H{}^h A_H{}^{-1} I_h{}^H A_h\right) S_h{}^\nu.$$

Proof: In the linear iteration

$$(\underline{v}^h)^i = (M_h)^i(\underline{v}^h)^0 + N_h^{(i)}\underline{f}^h,$$

choose $\underline{f}^h = 0$, replace M_h with S_h and i with ν, and apply the coarse-grid correction (2.20).

The success of the two-grid method is accredited to the good approximation of the low frequencies on the coarse scale. First, observe how the smooth frequencies are transformed onto the coarser grid. The following diagram illustrate that smooth modes on the fine grid appear more oscillatory on the coarse grid. This also can be shown mathematically by

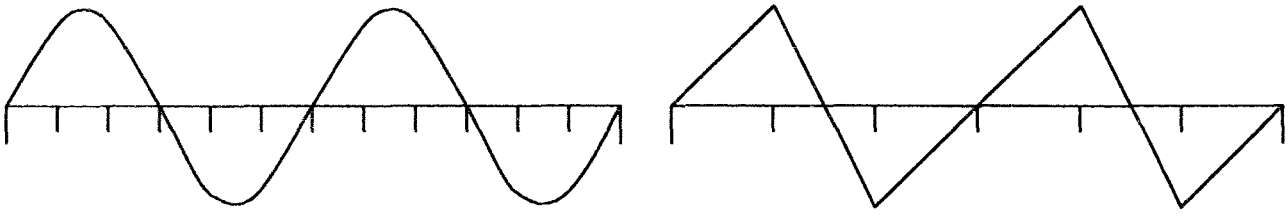

Figure 2.1: Frequency representation on two grids. The left diagram shows a low frequency on a fine grid, while the right diagram shows a high frequency on a coarse grid.

comparing the components of the eigenvectors of S_h and S_H at common points:

$$\begin{aligned} \underline{ew}^H{}_j &= \sqrt{2H}\,[\sin(pj\pi H)]_{p=1}^{(n^h-1)/2} \\ &= \sqrt{2}\sqrt{2h}\,[\sin(2pj\pi h)]_{p=1}^{(n^h-1)/2} \qquad j = 1, 2, \ldots, \frac{n^h-1}{2} \end{aligned}$$

and

$$\underline{ew}^h{}_j = \sqrt{2h}\,[\sin(kj\pi h)]_{k=1}^{n^h} \qquad j = 1, 2, \ldots, n^h,$$

so that at the common points of Ω^h and Ω^H (i.e. $k = 2p$),

$$\underline{ew}^H{}_j = \sqrt{2}\,\left[\underline{ew}^h{}_j\right].$$

For $1 \le j \le \frac{n^h-1}{2}$, the j'th mode on the fine grid transforms into the j'th mode of the coarse grid, or since there are only $\frac{n^h-1}{2}$ coarse-grid modes, half being oscillatory, some of the smooth fine grid modes transform into oscillatory coarse-grid modes. Now solving the defect equation exactly and interpolating the result, the smooth error components on the fine grid are effectively reduced, although some of the oscillatory fine-scale modes are excited by the interpolation. Additional two-grid iterations will resolve this latter problem, and will further diminish the smooth error components.

2.3.2 Multigrid Method

Another approach to solving (2.10) is created by solving the defect equation iteratively: why not apply a two-grid scheme to the defect equation, or why not apply the two-grid recursively until some lower level is reached where the defect equation there is solved exactly? Let l denote the grid level with step-size $h_l = 2^{-1-l}$. Solving the original defect equation with a smoother which effectively damps down the high frequency components, since only some of the smooth modes on the finest level transform to oscillatory modes on level $(n-1)$, only some of the smooth fine-grid modes will be effectively damped down. However, some of these other smooth modes can be approximated by solving the new defect equation on level $(n-2)$ and interpolating the result into Ω^{n-1}. The strategy should be obvious now. Modes that cannot be approximated well on one level will be approximated via solving the new defect equations on descending coarser levels.

Chapter 3

Regularization Algorithms

With the preliminary concepts developed, we now set up the stage for a discussion of numerical procedures for solving ill-posed problems. This chapter presents regularization methods, and develops some of the vital theory which will be used in the next chapter. Section 3.1 introduces the concept of generalized inverses for both operators with closed range and operators with non-closed range. Section 3.2 introduces the general regularization method and develops some theory for the case when the forcing term is known exactly, while section 3.3 develops some theory for the case when the forcing term is perturbed. Lastly, section 3.4 gives some important examples of regularization methods.

3.1 Generalized Inverses

As noted earlier, the problem $Kf = g$ is ill-posed if K^{-1} does not exists or if it does, it is not continuous. A generalization of the inverse operator is then needed.

The first step in defining such an operator is to modify the concept of a solution. Rather than having a solution satisfy the equation exactly, a solution is required only to satisfy the equation in the least square sense:

Definition 3.1 *An element $f \in \mathcal{H}_1$ is called a least squares solution of $Kf = g$ if*

$$\|Kf - g\| = \inf_{x \in \mathcal{H}_1} \|Kx - g\| .$$

If $\hat{f}$ is a least squares solution and $\left\|\hat{f}\right\| \leq \|f\|$ for all other least squares solution f, then $\hat{f}$ is called a least squares solution of minimal norm.

Now a fundamental theorem characterizing least squares solutions is

Theorem 3.1 *Let K be a bounded linear operator from a Hilbert space $\mathcal{H}_1$ into a Hilbert space $\mathcal{H}_2$. The following conditions are equivalent:*

1. *f is a least squares solution of $Kf = g$;*
2. *$K^*Kf = K^*g$;*
3. *$Kf = Pg$, where P is the orthogonal projection of $\mathcal{H}_2$ onto $\overline{R(K)}$.*

From (*3.*), it is clear that $Kf = g$ may not have a least squares solution if K does not have a closed range. But if there does exist a least squares solution f, and if $N(K) \neq \{0\}$, then the least squares solution is not unique since $f + f_1$, where f_1 is a nonzero element of $N(K)$, is another least squares solution. Also, from (*2.*) and the continuity and linearity of K, the set of least squares solution forms a closed convex set:

Let f_1, f_2 be least squares solutions. Then

$$\begin{aligned} K^*K[tf_1 + (1-t)f_2] &= tK^*Kf_1 + (1-t)K^*Kf_2 \\ &= tK^*g + (1-t)K^*g \\ &= K^*g\,. \end{aligned}$$

Let $\{f_n\}$ be sequence of least squares solutions which converges to f. Then $Pg = Kf_n \longrightarrow Kf$.

Hence, if $R(K)$ is closed, there exists a unique least squares solution of minimal norm.

For the case $R(K) \neq \overline{R(K)}$, if $g \in R(K) \oplus R(K)^{\perp}$, then there exist a set of least squares solutions, and using the same reasoning as above, there exists a unique least squares solution of minimal norm.

We now can define the generalized inverse of K.

Definition 3.2 *Let $K \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$, and let $D(K^\dagger) = R(K) \oplus R(K)^{\perp}$. The operator $K^\dagger : D(K^\dagger) \rightarrow \mathcal{H}_1$ which assigns to each $g \in D(K^\dagger)$ the unique least squares solution of minimal norm is called the generalized inverse of K.*

Theorem 3.2 *$K^\dagger$ is a linear operator.*

Theorem 3.3 *$K^\dagger$ is bounded if and only if $R(K)$ is closed.*

In particular, because a compact operator has a closed range if and only if it is of finite rank, $K^\dagger$ for a compact operator K is bounded if and only if $R(K)$ is finite-dimensional. This fact forecasts possible stability problems in solving the first kind equation $Kf = g$ even in the least squares sense.

We conclude this section with the following useful theorem:

Theorem 3.4 *Let $K : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ be a compact operator with singular system $\{u_n, v_n; \mu_n\}$, and let $g \in D(K^\dagger)$. Then*

$$\begin{aligned} K^\dagger g &= \sum_{n=1}^{\infty} \mu_n (Pg, u_n) v_n \\ &= \sum_{n=1}^{\infty} \mu_n (g, u_n) v_n \,, \end{aligned}$$

where P is the orthogonal projector onto $\overline{R(K)}$.

3.2 Regularization Methods

Most of the discussion so far has created the view that ill-posed equations cannot be solved effectively. But now we present an effective method which has been investigated thoroughly in Soviet literature. The basic idea behind this method is simple: rather than solving the equation using the original operator, an approximating operator having a bounded inverse is used. To justify this method however is not simple.

Let X and Y be two metric spaces, and let $G : X \longrightarrow Y$ be a mapping defined on $D(G) \subset X$. The task is to find the image of $G(x)$ given x. In terms of the first kind equation

$$Kf = g \,, \qquad f \in \mathcal{H}_1,\ g \in \mathcal{H}_2 \,,$$

$G = K^{-1}$, $X = \mathcal{H}_2$, $Y = \mathcal{H}_1$, $D(G) = K\mathcal{H}_1$, and the problem is to find $K^{-1}g$. Now if only x_δ satisfying $\rho_X(x_\delta, x) \leq \delta$ is available, and if the problem is well-posed, then $G(x_\delta)$ can be taken as an approximation to $G(x)$, and $G(x_\delta) \rightarrow G(x)$ as $\delta \rightarrow 0$. But if the equation is

ill-posed, x_δ may not belong to $D(G)$, and even if does, $G(x_\delta)$ need not converge to $G(x)$ as $\delta \to 0$. To circumvent this difficulty, regularizing algorithms are employed.

Definition 3.3 (Goncharsky, 1987) *Let R_δ be an operator defined on the pair (x_δ, δ), $x_\delta \in X$ and $0 < \delta \leq \delta_0$, and with range in Y. Consider the error*

$$\Delta(R_\delta, \delta, x) := \sup_{\{x_\delta \in X:\, \rho_X(x_\delta, x) \leq \delta\}} \rho_Y(R_\delta(x_\delta), G(x)) .$$

If $\lim_{\delta \to 0} \Delta(R_\delta, \delta, x) = 0$ for any $x \in D(G)$, then G is said to be regularizable on $D(G)$, and R_δ is called a regularizing operator.

In other words, if G is regularizable, then there exists a family operators $\{R_\delta\}$ such that

$$R_\delta(x_\delta) \to G(x) \qquad \text{as } \delta \to 0$$

for any $x \in D(G)$.

Of course, the regularizing operators should be chosen so that they are more tractable to computation than G is. Referring back to the initial discussion, the approximating operators having bounded inverses are regularizing operators.

Note that the error $\Delta(R_\delta, \delta, x)$ requires the knowledge of the exact value x, which is not known. This obstacle raises the question of how much information is needed to obtain an approximate solution of an ill-posed problem. Specifically, can an approximate solution be constructed using only x_δ under the condition that $\rho_X(x_\delta, x) \to 0$ as $\delta \to 0$? The next theorem shows that the minimum amount of information needed to construct an approximate solution of a first kind equation involving a compact operator is the pair (x_δ, δ).

Theorem 3.5 (Goncharsky, 1987) *G is regularizable on $D(G)$ by the family $R_\delta = R(\cdot)$ (i.e., R is independent of δ) if and only if G can be continued to all of X, the continuation being continuous on $D(G)$ in X.*

Now although (x_δ, δ) is the minimal amount of information required to regularize a first kind equation, still not all first kind equations can be regularized using only (x_δ, δ). The next two theorems illustrate some necessary and sufficient conditions for the regularizability of G.

Theorem 3.6 (Goncharsky, 1987) *Let $G : X \longrightarrow Y$ be the limit point on $D(G)$ of the functions G_n which are continuous all over X. Then G is regularizable on $D(G)$.*

Theorem 3.7 (Goncharsky, 1987) *Suppose a separable space Y is a convex subset of a linear normed space, and G is regularizable on $D(G)$. Then G is a limit point on $D(G)$ of a sequence of functions G_n which are continuous on X.*

Theorems 3.6 and 3.7 suggest choosing G_α where $G_\alpha \in \{G_n\}$ and where α is a weight function compromising between the fidelity of the approximate solution and the stability of the regularized equation.

Henceforth, we consider only the regularization of the equation

$$Kf = g, \qquad K \text{ compact.}$$

Recall that for a self-adjoint operator A, A may be expressed as

$$A = \int_{\sigma(A)} \lambda \, dE_\lambda ,$$

E_λ being its spectral family. Also, if h is a continuous function for $\lambda \in \sigma(A)$, then

$$h(A) = \int_{\sigma(A)} h(\lambda) \, dE_\lambda .$$

Formally then,

$$\begin{aligned} f &= K^\dagger g \\ &= h(K)g \\ &= \int_{\sigma(K)} \frac{dE_\lambda g}{\lambda} , \end{aligned}$$

where $h(\lambda) = \frac{1}{\lambda}$, assuming that K is self-adjoint and a solution exists.

The obvious problem is that K has a limiting spectral value of zero, and so any regularization algorithm must eliminate this difficulty. Such an algorithm can be constructed in the following way (Bakushinskii (1967), Groetsch (1984)):

i) Create a uniformly bounded real-valued function $\varphi(\lambda, \alpha)$ defined on $\lambda \in \sigma(K)$ and $\alpha > 0$, and satisfying

a. $\sup_{\lambda\in\sigma(K)} \frac{|\varphi(\lambda,\alpha)|}{\lambda} = K_\alpha < \infty$,

b. $\varphi(0,\alpha) = 0, \quad \alpha \neq 0$,

c. $\lim_{\alpha\to 0} \varphi(\lambda,\alpha) = 1$, uniformly for all $\lambda \in (c,\infty) \cap \sigma(K)$, $0 < c \ll 1$.

$\varphi(\lambda,\alpha)$ in a sense approximates the constant function 1.

ii) Define $h(\lambda,\alpha)$ to be

$$h(\lambda,\alpha) = \begin{cases} \frac{\varphi(\lambda,\alpha)}{\lambda}, & \lambda \neq 0 \\ L & |L| \leq K_\alpha,\ \lambda = 0. \end{cases}$$

Then $h(K,\alpha)$ is a regularizing operator of the equation provided that α as a continuous nonnegative function of δ satisfies $\alpha(\delta) \to 0$, and provided that $K_\alpha\delta \to 0$ as $\delta \to 0$.

Proof: Note that $\|h(K,\alpha)\| = K_\alpha$ since K is self-adjoint and compact. Hence,

$$\begin{aligned} \left\|f_\alpha^\delta - f\right\| &\leq \|f_\alpha - f\| + \left\|f_\alpha - f_\alpha^\delta\right\| \\ &\leq \|f_\alpha - f\| + \|h(K,\alpha)\|\,\|g - g_\delta\| \\ &\leq \|f_\alpha - f\| + K_\alpha\delta, \end{aligned}$$

where $f_\alpha = h(K,\alpha)\,g$ and $f_\alpha^\delta = h(K,\alpha)\,g_\delta$. Using the singular function expansions of f_α and f, it can be shown that $\|f_\alpha - f\| \longrightarrow 0$ as $\alpha \to 0$, or as $\delta \to 0$ (eg. Bakushinskii (1965), Groetsche (1984)). It follows that

$$\lim_{\delta\to 0} \left\|f_\alpha^\delta - f\right\| = \lim_{\delta\to 0} \left\|h(K,\alpha)\,g_\delta - K^\dagger g\right\| = 0$$

if $\lim_{\delta\to 0} K_\alpha\delta = 0$.

A similar procedure can be used to construct an r.a. (regularization algorithm) when K is an arbitrary compact operator, except that the equation to be regularized is now $K^*Kf = K^*g$, and except that $\varphi(\lambda,\alpha)$ must now satisfy

$$\sup_{\lambda\in\sigma(K^*K)} \frac{|\varphi(\lambda,\alpha)|}{\sqrt{\lambda}} = \hat{K}_\alpha < \infty, \qquad \alpha \neq 0.$$

Defining

$$h(\lambda,\alpha) = \begin{cases} \frac{\varphi(\lambda,\alpha)}{\lambda}, & \lambda \neq 0 \\ L & |L| \leq \infty,\ \lambda = 0, \end{cases}$$

the regularization operator is $h(K^*K,\alpha)K^*$.

Before making any examination of this r.a., some modifications will be made. Instead of the function $\varphi(\lambda,\alpha)$, the function $R_\alpha(t)$ which is continuous on $\sigma(K^*K)$ and which approximates $h(t)=\frac{1}{t}$ will be used. $R_\alpha(t)$ will be continuous for each $\alpha>0$, and will converge to $\frac{1}{t}$ as $\alpha\to 0$, and $|tR_\alpha(t)|$ will be uniformly bounded.

Given these changes, since $K^*Kf=K^*g$ is the equation to be regularized, $R_\alpha(K^*K)K^*g$ must approximate $K^\dagger g$. This is indeed the case.

Theorem 3.8 *Suppose that $\{R_\alpha\}_{\alpha>0}$ is a family of continuous real-valued functions on $[0,\|K\|^2]\supseteq\sigma(K^*K)$ satisfying*

$$R_\alpha(t) = \frac{1}{t} \quad \text{as } \alpha\to 0, \tag{3.1}$$

$$|tR_\alpha(t)| \le M \quad \text{for all } \alpha>0,\ t\in[0,\|K\|^2], \text{ and } M \text{ finite}. \tag{3.2}$$

Then $R_\alpha(K^*K)K^*g\longrightarrow K^\dagger g$ as $\alpha\to 0$ for each $g\in D(K^\dagger)$.

Proof: The identity $p(K^*K)K^*=K^*p(KK^*)$ for any polynomial p, and the Weierstrass approximation theorem imply that $R_\alpha(K^*K)K^*=K^*R_\alpha(KK^*)$. Hence, $R_\alpha(K^*K)K^*g\in R(K^*)$. Next, using the singular system $\{u_n,v_n;\mu_n\}$ of K,

$$\begin{aligned}
R_\alpha(K^*K)K^*g &= \sum_{n=1}^{\infty} R_\alpha({\mu_n}^{-2})(K^*g,v_n)v_n \\
&= \sum_{n=1}^{\infty} R_\alpha({\mu_n}^{-2})(g,Kv_n)v_n \\
&= \sum_{n=1}^{\infty} R_\alpha({\mu_n}^{-2}){\mu_n}^{-1}(g,u_n)v_n \\
&= \sum_{n=1}^{\infty} \mu_n{\mu_n}^{-2}R_\alpha({\mu_n}^{-2})(g,u_n)v_n .
\end{aligned}$$

Considering each term of the sum as a function of α and using (3.2) and theorem 3.4, the partial sums $S_p(\alpha)$ form a Cauchy sequence independent of α:

$$\begin{aligned}
\|S_p(\alpha)-S_q(\alpha)\|^2 &= \left\| \sum_{n=p+1}^{q} \mu_n{\mu_n}^{-2}R_\alpha({\mu_n}^{-2})(g,u_n)v_n \right\|^2 \\
&= \sum_{n=p+1}^{q} {\mu_n}^2\left|{\mu_n}^{-2}R_\alpha({\mu_n}^{-2})\right|^2|(g,u_n)|^2
\end{aligned}$$

$$\leq \quad M^2 \sum_{n=p+1}^{q} {\mu_n}^2 |(g, u_n)|^2 \longrightarrow 0 \qquad \text{for sufficiently large } p \text{ and } q\,.$$

Thus, the series converges uniformly in α. The result follows trivially.

The requirement that $g \in D(K^\dagger)$ is crucial to the convergence of $R_\alpha(K^*K)K^*g$ for it can be shown that if $g \notin D(K^\dagger)$, then for any sequence $\alpha_n \to 0$, $\{R_{\alpha_n}(K^*K)K^*g\}$ is not weakly convergent (Groetsch (1984)). Moreover, to ensure good convergence rates, further requirements are needed. For example, consider the case when g is known exactly. Now because $R(K) = R(KP_{N(K)^\perp})$ where P_S denotes the projection operator onto the subspace S, and because $P_{N(K)^\perp}x = P_{N(K^*K)^\perp}x = \lim_{\nu\to 0+}(K^*K)^\nu x$ [1], a reasonable requirement is $Pg \in R(K(K^*K)^\nu)$. This change enforces the new condition

$$t^\nu |1 - tR_\alpha(t)| \leq \omega(\alpha, \nu) \qquad t \in [0, \|K\|^2]\,, \tag{3.3}$$

where $\omega(\alpha, \nu)$ satisfies $\omega(\alpha, \nu) \to 0$ as $\alpha \to 0$ for each $\nu > 0$.

3.3 The Perturbed Equation

For the perturbed data case, the regularization parameter must be described more specifically. It is a continuous nonnegative function $\alpha : [0, \infty) \longrightarrow [0, \infty)$ with $\alpha(0) = 0$, and satisfying some condition which ensures the convergence of $R_{\alpha(\delta)}(K^*K)K^*g$ to $K^\dagger g$ as $\delta \to 0$. Letting $r(\alpha) := \max_{t\in[0,\|K\|^2]} |R_\alpha(t)|$, we delve into finding this condition.

Lemma 3.9 $\left\|K(f_\alpha - f_\alpha^\delta)\right\| \leq \delta M$.

Proof: A simple application of the bound (3.2) gives the above fact:

$$\begin{aligned}
\left\|K(f_\alpha - f_\alpha^\delta)\right\|^2 &= \left(K^*K(f_\alpha - f_\alpha^\delta), f_\alpha - f_\alpha^\delta\right) \\
&= \left(K^*K\left[R_\alpha(K^*K)K^*g - R_\alpha(K^*K)K^*g_\delta\right], (f_\alpha - f_\alpha^\delta)\right) \\
&= \left(K\left[R_\alpha(K^*K)K^*g - R_\alpha(K^*K)K^*g_\delta\right], K(f_\alpha - f_\alpha^\delta)\right)
\end{aligned}$$

[1] Use the spectral theorem for compact self-adjoint operators:

$$(K^*K)^\nu x = \sum {\lambda_n}^{2\nu}(x, v_n)v_n\,,$$

where ${\lambda_n}^2$ and v_n is an eigenvalue–eigenvector pair of K^*K.

$$\begin{aligned} &= \left([KK^*R_\alpha(KK^*)](g-g_\delta), K(f_\alpha - f_\alpha^\delta)\right) \\ &\leq M\delta \left\|K(f_\alpha - f_\alpha^\delta)\right\| . \end{aligned}$$

Lemma 3.10 $\left\|f_\alpha - f_\alpha^\delta\right\| \leq \delta\sqrt{Mr(\alpha)}$.

Proof: Use lemma 3.9 and the definition of $r(\alpha)$:

$$\begin{aligned} \left\|f_\alpha - f_\alpha^\delta\right\|^2 &= \left(f_\alpha - f_\alpha^\delta, R_\alpha(K^*K)K^*(g-g_\delta)\right) \\ &= \left(f_\alpha - f_\alpha^\delta, K^*R_\alpha(KK^*)(g-g_\delta)\right) \\ &= \left(K(f_\alpha - f_\alpha^\delta), R_\alpha(KK^*)(g-g_\delta)\right) \\ &\leq \delta^2 Mr(\alpha) . \end{aligned}$$

Theorem 3.11 *Let $g \in D(K^\dagger)$, $\alpha(\delta) \to 0$ and $\delta^2 r(\alpha(\delta)) \to 0$ as $\delta \to 0$. Then*

$$R_\alpha(K^*K)K^*g_\delta \longrightarrow K^\dagger g \qquad \text{as } \delta \to 0 .$$

Proof:

$$\begin{aligned} \left\|K^\dagger g - f_{\alpha(\delta)}^\delta\right\| &\leq \left\|K^\dagger g - f_{\alpha(\delta)}\right\| + \left\|f_{\alpha(\delta)} - f_{\alpha(\delta)}^\delta\right\| \\ &\leq \left\|K^\dagger g - f_{\alpha(\delta)}\right\| + \delta\sqrt{Mr(\alpha)} . \end{aligned}$$

By theorem 3.8 and the hypothesis that $\delta^2 r(\alpha(\delta)) \to 0$ as $\delta \to 0$,

$$f_{\alpha(\delta)}^\delta = R_\alpha(K^*K)K^*g_\delta \longrightarrow K^\dagger g .$$

Hence, the condition is, $\alpha(\delta) \to 0$ and $\delta^2 r(\alpha(\delta)) \to 0$ as $\delta \to 0$.

But this condition is not sufficient for practical purposes since usually all that is available is an operator K_η satisfying $\|K - K_\eta\| \leq \eta$. We leave the investigation of this problem until the discussion of each particular regularization algorithm.

3.4 Examples

Example 1. Landweber–Fridman Iteration

$$f_{n+1}^\delta = f_n^\delta + \mu(K^*g_\delta - K^*Kf_n^\delta) , \qquad f_0^\delta = 0,\ 0 < \mu < \frac{2}{\|K^*K\|} = \frac{2}{\|K\|^2} .$$

The generating function of this iteration is

$$R_n(t) = \mu \sum_{k=1}^{n} (1 - \mu t)^{k-1}, \qquad t \in \left[0, \|K\|^2\right] .$$

Defining the regularization parameter to be $\alpha(\delta) = \frac{1}{\beta(\delta)}$, $\lfloor \beta(\delta) \rfloor = n(\delta)$, there is a natural correspondence between the generating functions $\hat{R}_{\alpha(\delta)}(t)$ and $R_n(t)$. Now since

$$R_n(t)\left[(1 - \mu t) - 1\right] = \mu\left[(1 - \mu t)^n - 1\right] ,$$

or

$$-tR_n(t) = \left[(1 - \mu t)^n - 1\right] ,$$

$M = 1$. Also,

$$\begin{aligned} r(n) &= \max_{t \in \sigma(K^*K)} |R_n(t)| \\ &= \max_{t \in \sigma(K^*K)} \mu \left| \sum_{k=1}^{n} (1 - \mu t)^{k-1} \right| \\ &= \mu n . \end{aligned}$$

Thus choosing $n(\delta)$ such that $\delta^2 \mu n(\delta) \to 0$ as $\delta \to 0$, $f^{\delta}_{n(\delta)} \longrightarrow K^{\dagger} g$.

Example 2: Tikhonov Regularization

$$f^{\delta}_{\alpha(\delta)} = \left[K^*K + \alpha(\delta)\right]^{-1} K^* g_{\delta}$$

This is the classical regularization algorithm. Its generating function is $R_{\alpha}(t) = (t + \alpha)^{-1}$. Hence, $M = 1$, $r(\alpha) = \frac{1}{\alpha}$, and $f^{\delta}_{\alpha(\delta)} \longrightarrow K^{\dagger} g$ if $\frac{\delta^2}{\alpha(\delta)} \to 0$ as $\delta \to 0$.

These two examples are also part of another class of regularization methods described by the equation

$$f^{\delta}_{\alpha(\delta)} = \left[I - K^*K \hat{R}_{\alpha(\delta)}(K^*K)\right] u_0 + \hat{R}_{\alpha(\delta)}(K^*K) K^* g_{\delta} ,$$

where u_0 is an initial approximation, and $\hat{R}_{\alpha(\delta)}(t)$ satisfies

$$\begin{aligned} \sup_{t \in [0, \|K\|^2]} \left| \hat{R}_{\alpha(\delta)}(t) \right| &\leq \frac{\gamma}{\alpha(\delta)} , \\ \sup_{t \in [0, \|K\|^2]} t^p \left| 1 - t\hat{R}_{\alpha(\delta)}(t) \right| &\leq \gamma_p \, \alpha(\delta)^p, \qquad 0 \leq p \leq p_0 , \end{aligned}$$

γ, γ_p, p_0 constants (Vainikko (1982 and 1983)).

Example 3: Landweber–Fridman Iteration

Again $\alpha(\delta) = \frac{1}{\beta(\delta)}$, $\lfloor \beta(\delta) \rfloor = n(\delta)$. Or more simply, $\alpha(\delta) = \frac{1}{n(\delta)}$. Hence,

$$f^{\delta}_{\alpha(\delta)} = (I - \mu K^* K)^{n(\delta)} f_0 + \mu \sum_{k=1}^{n(\delta)} (I - \mu K^* K)^{k-1} K^* g_\delta$$

where f_0 need not be zero. Now using the identity $(I - \mu K^* K)^n = I - \mu \sum_{k=1}^{n} K^* K (I - \mu K^* K)^{k-1}$,

$$\hat{R}_{\alpha(\delta)}(t) = \mu \sum_{k=1}^{n} (1 - \mu t)^{k-1} .$$

Example 4: Iterative Tikhonov Method

$$f^{\delta}_{n,\alpha(\delta)} = [\alpha(\delta) + K^* K]^{-1} \left(\alpha(\delta) f^{\delta}_{n-1,\alpha(\delta)} + K^* g_\delta \right) \qquad n = 1, 2, \ldots, m .$$

Let $f^{\delta}_{0,\alpha} = f_0$ be an initial approximation, and let $f^{\delta}_{\alpha} = f^{\delta}_{m,\alpha}$. Then the above iteration generates

$$\begin{aligned} f^{\delta}_{\alpha(\delta)} &= [\alpha(\delta)]^m [\alpha(\delta) I + K^* K]^{-m} f_0 + \sum_{j=1}^{m} [\alpha(\delta)]^{j-1} [\alpha(\delta) I + K^* K]^{-j} K^* g_\delta \\ &= [\alpha(\delta)]^m [\alpha(\delta) I + K^* K]^{-m} f_0 + \sum_{j=1}^{m} \frac{1}{\alpha(\delta)} \left[I + \frac{1}{\alpha(\delta)} K^* K \right]^{-j} K^* g_\delta , \end{aligned}$$

implying that

$$\begin{aligned} \hat{R}_{\alpha(\delta)}(t) &= \frac{1}{\alpha(\delta)} \sum_{j=1}^{m} \left(1 + \frac{t}{\alpha(\delta)} \right)^{-j} \\ &= \frac{1}{t} \left[1 - \left(1 + \frac{1}{\alpha(\delta)} t \right)^{-m} \right] . \end{aligned}$$

Choosing $m = 1$ and $f_0 = 0$, the classical Tikhonov method reappears.

Example 5: Implicit Iterative Method

$$f^{\delta}_{n} = [\tau I + K^* K]^{-1} \left(\tau f^{\delta}_{n-1} + K^* g_\delta \right) , \qquad n = 1, 2, \ldots \text{ and } \tau > 0 .$$

The regularization parameter here is also $\frac{1}{n(\delta)}$. This gives

$$f^{\delta}_{n(\delta)} = \tau^{n(\delta)} [\tau I + K^* K]^{-n(\delta)} f_0 + \sum_{j=1}^{n(\delta)} \tau^{j-1} [\tau I + K^* K]^{-j} K^* g_\delta ,$$

and

$$\hat{R}_{\alpha(\delta)}(t) = \sum_{j=1}^{m} \tau^{j-1} [\tau + t]^{-j} .$$

Notice that for the last two examples, no condition on $\alpha(\delta)$ to guarantee the convergence of $f^{\delta}_{\alpha(\delta)}$ to $K^{\dagger}g$ has been specified in this section.

Chapter 4

Discrepancy Principles and Multilevel Procedures

Several questions concerning r.a.'s remain to be discussed. They concern the numerical realization of these algorithms: how should the regularization parameter be chosen systematically without considerably impairing the convergence rate, and how do the the discretization errors and the unavoidable random errors affect the effectiveness of the r.a., and what modifications must be made to retain its effectiveness? Some of these questions will be answered for several methods in this chapter. Section 4.1 will investigate the discrepancy principle, which is an indispensable component in the choosing of the regularization parameter. This principle is then applied to the Landweber iteration in section 4.2, where we make a thorough investigation of the realization of this iteration and describe a multilevel adaption to it. In section 4.3, the discrepancy principle will be applied to the classical Tikhonov method, and a multilevel process will be described. Also in this section, we will introduce the concept of stabilizers and look at the quasi-optimal parameter choice, and with these two, we will propose a multilevel Tikhonov procedure for solving nonlinear first kind equations.

4.1 Discrepancy Principle

A fitting way to start this section is to quote Morozov: "the magnitude of the residual $\left[\left\|K_\eta f_\alpha^\delta - g_\delta\right\|\right]$ must be commensurate with the inconsistency measure $[\inf_{x\in\mathcal{H}_1}\|Kx-g\|]$ and the accuracy of specification of input information of the problem."[1]

Recall that the regularization parameter is a weight function compromising between the fidelity of the approximate solution and the stability of the regularized equation. One would like an α that results in good fidelty and good stability. A systematic way of choosing such an α is to apply the discrepancy principle which chooses the parameter satisfying the criterion that the magnitude of the residual formed by the resulting solution be commensurate with the inconsistency measure and the data error. We establish the validity of this principle for the classical Tikhonov method first.

Assume that

$$\|g_\delta\|^2 > \delta^2 + {\mu_\eta}^2 , \tag{4.1}$$

where ${\mu_\eta}^2$ is the inconsistency measure

$$ {\mu_\eta}^2 := \inf_{x\in\mathcal{H}_1}\|K_\eta x - g_\delta\|^2 \qquad \text{with } K_\eta \text{ satisfying } \|K-K_\eta\|\le\eta .$$

The decomposition

$$g_\delta = \nu_\eta + \omega_\eta , \qquad \nu_\eta \in N(K_\eta^*),\ \omega_\eta \in \overline{R(K_\eta)} ,$$

and the inequality

$$\frac{\|g_\delta\|^2}{\delta^2} \ge \frac{\|g_\delta\|^2}{\delta^2+{\mu_\eta}^2}$$

show that (4.1) is sort of a natural condition for one would desire both that g_δ has some component in $\overline{R(K_\eta)}$[2], and that the signal-to-noise ratio is greater than one. Indeed, if (4.1) is not satisfied then $f_{\alpha(\eta,\delta)}^{(\eta,\delta)} = 0$ may be taken as an approximate solution, and the data may be too corrupted to permit any useful mathematical analysis.

Also assume that the operator is known exactly, and take the view that $g \in R(K)$, absorbing any $N(K^*)$ component of g into the error in g_δ. Hence, (4.1) modifies into the form

$$\|g_\delta\|^2 > c\delta^2 \qquad c \ge 1, \tag{4.2}$$

[1] Morozov (1984), p.32.

[2] Note that $\|\nu_\eta\|^2 = {\mu_\eta}^2$ and $\|g_\delta\|^2 = \|\nu_\eta\|^2 + \|\omega_\eta\|^2$. Hence, (4.1) implies that $\|\omega_\eta\|^2 > \delta^2$.

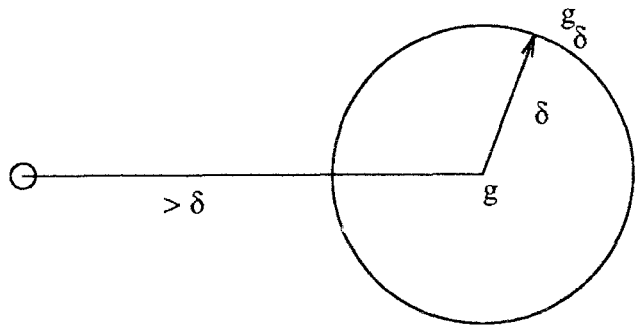

Figure 4.1: Limiting size of $\delta : \delta < \frac{\|g\|}{2}$.

which we shall take as

$$\|g_\delta\|^2 > \delta^2 \tag{4.3}$$

(Note that since $g \in R(K)$, (4.3) does not have the problem that $f^\delta_{\alpha(\delta)} = 0$ may be taken as an approximate solution.)

Now the discrepancy principle chooses the unique $\alpha(\delta)$ that satisfies

$$\rho(\alpha) := \left\| K f^\delta_{\alpha(\delta)} - g_\delta \right\| = \delta \,. \tag{4.4}$$

The uniqueness follows from the next theorem.

Theorem 4.1 *Suppose that $g \in R(K)$, and that g and g_δ satisfy*

$$\|g - g_\delta\|^2 \leq \delta^2 < \|g_\delta\|^2 \,.$$

Then $\rho(\alpha)$ is continuous, increasing, and contains δ in its range.

Proof: Because the approximate solution produced by Tikhonov's method is

$$f^\delta_{\alpha(\delta)} = \left[K^* K + \alpha(\delta)\right]^{-1} K^* g_\delta \,,$$

$$\begin{aligned}
K f^\delta_{\alpha(\delta)} - g_\delta &= K \left[K^* K + \alpha(\delta) I\right]^{-1} K^* g_\delta - g_\delta \\
&= K K^* \left[K K^* + \alpha(\delta) I\right]^{-1} g_\delta - \left(I - P_{R(K)^\perp}\right) g_\delta - P_{R(K)^\perp} g_\delta \\
&= \sum_n \frac{{\lambda_n}^2}{{\lambda_n}^2 + \alpha} (g_\delta, u_n) u_n - \sum_n (g_\delta, u_n) u_n - P_{R(K)^\perp} g_\delta \\
&= \sum_n -\frac{\alpha}{{\lambda_n}^2 + \alpha} (g_\delta, u_n) u_n - P_{R(K)^\perp} g_\delta \\
&= \sum_n \frac{-\alpha {\mu_n}^2}{1 + \alpha {\mu_n}^2} (g_\delta, u_n) u_n - P_{R(K)^\perp} g_\delta
\end{aligned}$$

where $\{u_n, v_n; \mu_n\}$ is the singular system of K. Hence,

$$\rho^2(\alpha) = \sum_n \left(\frac{\alpha \mu_n{}^2}{1 + \alpha \mu_n{}^2} \right)^2 |(g_\delta, u_n)|^2 + \left\| P_{R(K)^\perp} g_\delta \right\|^2 ,$$

which is continuous and increasing. The continuity and monotonicity of $\rho(\alpha)$ clearly follows.

Next, taking the limit of $\rho^2(\alpha)$ as $\alpha \to \infty$ and $\alpha \to 0$, we have

$$\lim_{\alpha \to \infty} \rho^2(\alpha) = \left\| \left(I - P_{R(K)^\perp} \right) g_\delta \right\|^2 + \left\| P_{R(K)^\perp} g_\delta \right\|^2 = \|g_\delta\|^2 > \delta^2$$

and

$$\lim_{\alpha \to 0} \rho^2(\alpha) = \left\| P_{R(K)^\perp} g_\delta \right\|^2 \leq \|g - g_\delta\|^2 \leq \delta^2 .^3$$

Thus, the range of $\rho(\alpha)$ contains δ.

Further insight into the discrepancy principle is revealed through an examination of a particular error bound on the approximate solution. Let $r(\alpha; g_\delta) := g_\delta - K f^\delta_{\alpha(\delta)}$. Then

$$\begin{aligned} K^* r(\alpha; g_\delta) &= K^* g_\delta - K^* K f^\delta_\alpha \\ &= K^* g_\delta - K^* K \left[K^* K + \alpha I \right]^{-1} K^* g_\delta \\ &= \alpha \left[K^* K + \alpha I \right]^{-1} K^* g_\delta \\ &= \alpha f^\delta_\alpha , \end{aligned} \tag{4.5}$$

and so

$$\begin{aligned} \left\| f - f^\delta_\alpha \right\|^2 &= \|f\|^2 - 2 \left(f, f^\delta_\alpha \right) + \left\| f^\delta_\alpha \right\|^2 \\ &= \|f\|^2 - \frac{2}{\alpha} \left(K^* r(\alpha; g_\delta), f \right) + \left\| f^\delta_\alpha \right\|^2 \\ &= \|f\|^2 - \frac{2}{\alpha} \left(r(\alpha; g_\delta), g_\delta \right) + \frac{2}{\alpha} \left(r(\alpha; g_\delta), g_\delta - g \right) + \left\| f^\delta_\alpha \right\|^2 \\ &\leq E(\alpha; g_\delta) , \end{aligned}$$

where $E(\alpha; g_\delta) := \|f\|^2 - \frac{2}{\alpha} \left(r(\alpha; g_\delta), g_\delta \right) + \frac{2\delta}{\alpha} \rho(\alpha) + \left\| f^\delta_\alpha \right\|^2$. This error bound is minimized when α satisfies (4.4).

---

[3] $P_{R(K)^\perp} g_\delta \in N(K^*)$, $(g - g_\delta) \in N(K^*) \oplus N(K^*)^\perp$; i.e., $g - g_\delta = g - (g + \delta g) = -\delta g = -\delta g_{N(K^*)^\perp} - \delta g_{N(K^*)}$.

Theorem 4.2 (Groetsch, 1984) *Suppose that $g \in R(K)$, and that g and g_δ satisfy*

$$\|g - g_\delta\|^2 \leq \delta^2 < \|g_\delta\|^2 .$$

Then $E(\alpha; g_\delta)$ is minimized if and only if $\rho(\delta) = \delta$.

The proof of this theorem involves taking the Fréchet derivative of an operator. Also, it uses the fact that a positive operator has a square root.

Definition 4.1 *Let E_1, E_2 be normed linear spaces, and suppose that Z is an open subset of E_1. A function $F : Z \longrightarrow E_2$ is said to be Fréchet differentable at $x_0 \in Z$ if there is an operator $F'(x_0) \in \mathcal{B}(E_1, E_2)$ such that*

$$\lim_{\|h\| \to 0} \frac{\|F(x_0 + h) - F(x_0) - F'(x_0)h\|}{\|h\|} = 0 .$$

$F'(x_0)$ is called the Fréchet derivative of F at x_0.

Proof: First, if

$$[K^*K + \alpha(\delta)I]\, f^\delta_{\alpha(\delta)} = K^* g_\delta ,$$

then

$$[K^*K + \alpha(\delta)I] \frac{df^\delta_\alpha}{d\alpha} = -f^\delta_{\alpha(\delta)}$$

or

$$\frac{df^\delta_\alpha}{d\alpha} = -\left[K^*K + \alpha(\delta)I\right]^{-1} f^\delta_{\alpha(\delta)} = -\frac{1}{\alpha(\delta)} \left[K^*K + \alpha(\delta)I\right]^{-1} K^* r(\alpha; g_\delta) .$$

Next, $\rho(\alpha)$ is positive for $\alpha > 0$. This follows from the fact that if $\rho(\alpha) = 0$ for some positive α, then $r(\alpha; g_\delta) = 0$, or equivalently, using (4.5), $f^\delta_{\alpha(\delta)} = [K^*K + \alpha(\delta)I]^{-1} K^* g_\delta = 0$. But then $g_\delta \in N(K^*) = R(K)^\perp$, and consequently the contradiction $\delta^2 \geq \|g - g_\delta\|^2 = \|g\|^2 + \|g_\delta\|^2 > \|g\|^2 + \delta^2$ arises.

Now taking the derivative of $E(\alpha; g_\delta)$, we have

$$\begin{aligned} \frac{d}{d\alpha} E(\alpha; g_\delta) &= 2\left(\frac{df^\delta_\alpha}{d\alpha}, f^\delta_\alpha\right) + \frac{2}{\alpha^2}\left(r(\alpha; g_\delta), g_\delta\right) - \frac{2}{\alpha}\left(\frac{d}{d\alpha}(g_\delta - K f^\delta_\alpha), g_\delta\right) - \frac{2\delta}{\alpha^2}\rho(\alpha) \\ &\quad + \frac{2\delta}{\alpha\rho(\alpha)}\left(\frac{d}{d\alpha}(K f^\delta_\alpha - g_\delta), K f^\delta_\alpha - g_\delta\right) \end{aligned}$$

$$
\begin{aligned}
&= \frac{2}{\alpha}\left(\frac{df_\alpha^\delta}{d\alpha}, K^*r(\alpha;g_\delta)\right) + \frac{2}{\alpha^2}\left(r(\alpha;g_\delta), g_\delta\right) + \frac{2}{\alpha}\left(K\frac{df_\alpha^\delta}{d\alpha}, g_\delta\right) \\
&\quad - \frac{2\delta}{\alpha^2\rho(\alpha)}\left[\rho^2(\alpha) - \alpha\left(K\frac{df_\alpha^\delta}{d\alpha}, -r(\alpha;g_\delta)\right)\right] \\
&= -\frac{2}{\alpha^2}\left([K^*K+\alpha I]^{-1}K^*r(\alpha;g_\delta), K^*r(\alpha;g_\delta)\right) + \frac{2}{\alpha^2}\left(r(\alpha;g_\delta), g_\delta\right) \\
&\quad - \frac{2}{\alpha^2}\left(K[K^*K+\alpha I]^{-1}K^*r(\alpha;g_\delta), g_\delta\right) \\
&\quad - \frac{2\delta}{\alpha^2\rho(\alpha)}\left[\left(r(\alpha;g_\delta), r(\alpha;g_\delta)\right) + \left(K[K^*K+\alpha I]^{-1}K^*r(\alpha;g_\delta), -r(\alpha;g_\delta)\right)\right] \\
&= -\frac{2}{\alpha^2}\left\{\left(KK^*[KK^*+\alpha I]^{-1}r(\alpha;g_\delta), r(\alpha;g_\delta)\right) - \left(r(\alpha;g_\delta), g_\delta\right)\right. \\
&\quad \left. + \left(KK^*[KK^*+\alpha I]^{-1}r(\alpha;g_\delta), g_\delta\right)\right\} \\
&\quad - \frac{2\delta}{\alpha^2\rho(\alpha)}\left(r(\alpha;g_\delta) - KK^*[KK^*+\alpha I]^{-1}r(\alpha;g_\delta), r(\alpha;g_\delta)\right) \\
&= -\frac{2}{\alpha^2}\left\{\left(KK^*[KK^*+\alpha I]^{-1}r(\alpha;g_\delta), r(\alpha,g_\delta)\right)\right. \\
&\quad \left. + \left(KK^*[KK^*+\alpha I]^{-1}r(\alpha;g_\delta) - r(\alpha;g_\delta), g_\delta\right)\right\} \\
&\quad - \frac{2\delta}{\alpha^2\rho(\alpha)}\left((KK^*+\alpha I - KK^*)[KK^*+\alpha I]^{-1}r(\alpha;g_\delta), r(\alpha;g_\delta)\right) \\
&= -\frac{2}{\alpha^2}\left\{\left(KK^*[KK^*+\alpha I]^{-1}r(\alpha;g_\delta), r(\alpha,g_\delta)\right)\right. \\
&\quad \left. + \left((KK^*-KK^*-\alpha I)[KK^*+\alpha I]^{-1}r(\alpha;g_\delta), g_\delta\right)\right\} \\
&\quad - \frac{2\delta}{\alpha\rho(\alpha)}\left([KK^*+\alpha I]^{-1}r(\alpha;g_\delta), r(\alpha;g_\delta)\right) \\
&= -\frac{2}{\alpha^2}\left\{\left(KK^*[KK^*+\alpha I]^{-1}r(\alpha;g_\delta) - \alpha[KK^*+\alpha I]^{-1}r(\alpha;g_\delta), g_\delta\right)\right. \\
&\quad \left. - \left(KK^*[KK^*+\alpha I]^{-1}r(\alpha;g_\delta), Kf_\alpha^\delta\right)\right\} \\
&\quad - \frac{2\delta}{\alpha\rho(\alpha)}\left([KK^*+\alpha I]^{-1}r(\alpha;g_\delta), r(\alpha,g_\delta)\right) \\
&= -\frac{2}{\alpha^2}\left\{\left([KK^*-\alpha I][KK^*+\alpha I]^{-1}r(\alpha;g_\delta), g_\delta\right)\right. \\
&\quad \left. - \left(KK^*[KK^*+\alpha I]^{-1}r(\alpha;g_\delta), Kf_\alpha^\delta\right)\right\}
\end{aligned}
$$

$$
\begin{aligned}
& -\frac{2\delta}{\alpha\rho(\alpha)}\left([KK^*+\alpha I]^{-1}r(\alpha;g_\delta),r(\alpha,g_\delta)\right)\\
= & -\frac{2}{\alpha^2}\left([KK^*+\alpha I]^{-1}r(\alpha,g_\delta),[KK^*-\alpha I]g_\delta-KK^*Kf_\alpha^\delta\right)\\
& -\frac{2\delta}{\alpha\rho(\alpha)}\left([KK^*+\alpha I]^{-1}r(\alpha;g_\delta),r(\alpha,g_\delta)\right)\\
= & -\frac{2}{\alpha^2}\left([KK^*+\alpha I]^{-1}r(\alpha,g_\delta),K\left(K^*g_\delta-K^*Kf_\alpha^\delta\right)-\alpha g_\delta\right)\\
& -\frac{2\delta}{\alpha\rho(\alpha)}\left([KK^*+\alpha I]^{-1}r(\alpha;g_\delta),r(\alpha,g_\delta)\right)\\
= & -\frac{2}{\alpha^2}\left([KK^*+\alpha I]^{-1}r(\alpha;g_\delta),\alpha Kf_\alpha^\delta-\alpha g_\delta\right)\\
& -\frac{2\delta}{\alpha\rho(\alpha)}\left([KK^*+\alpha I]^{-1}r(\alpha;g_\delta),r(\alpha,g_\delta)\right)\\
= & \frac{2}{\alpha}\left([KK^*+\alpha I]^{-1}r(\alpha;g_\delta),r(\alpha;g_\delta)\right)-\frac{2\delta}{\alpha\rho(\alpha)}\left([KK^*+\alpha I]^{-1}r(\alpha;g_\delta),r(\alpha,g_\delta)\right)\\
= & \frac{2}{\alpha}\left(1-\frac{\delta}{\rho(\alpha)}\right)\left([KK^*+\alpha I]^{-1}r(\alpha;g_\delta),r(\alpha;g_\delta)\right)\\
= & \frac{2}{\alpha}\left(1-\frac{\delta}{\rho(\alpha)}\right)\left\|[KK^*+\alpha I]^{-\frac{1}{2}}r(\alpha;g_\delta)\right\|^2,
\end{aligned}
$$

where the factor $[KK^*+\alpha I]^{-\frac{1}{2}}$ follows from the positivity of $[KK^*+\alpha I]^{-1}$. Since $r(\alpha;g_\delta)\neq 0$, $\left\|[KK^*+\alpha I]^{-\frac{1}{2}}r(\alpha;g_\delta)\right\|^2>0$. Hence, from the monotonicity of $\rho(\alpha)$,

$$\frac{d}{d\alpha}\left(E(\alpha;g_\delta)\right)>0 \qquad \text{when } \rho(\alpha)>\delta,$$

and

$$\frac{d}{d\alpha}\left(E(\alpha;g_\delta)\right)<0 \qquad \text{when } \rho(\alpha)<\delta.$$

Hence, $\frac{d}{d\alpha}E(\alpha;g_\delta)=0$ if and only if $\rho(\alpha)=\delta$.

But the minimization of this bound is relevant only if the minimum is not too large. Consider $\lim_{\delta\to 0}E(\alpha;g_\delta)$. A consequence of applying the discrepancy principle is that $\alpha(\delta)\to 0$ as $\delta\to 0$. This follows from the contradiction that if $\alpha\to\tilde{C}>0$ as $\delta_n\to 0$, then $g=0$ and therefore,

$$\|g_{\delta_n}\|>\delta_n\geq\|g-g_{\delta_n}\|=\|g_{\delta_n}\|:$$

$$
\begin{aligned}
0 = \lim_{n\to\infty} \delta_n = \lim_{n\to\infty} \rho\left(\alpha(\delta_n)\right) &= \lim_{n\to\infty} \left\| K f^{\delta_n}_{\alpha(\delta_n)} - g_{\delta n} \right\| \\
&= \left\| K[K^*K + \tilde{C}I]^{-1} K^* g - g \right\| \\
&= \left\| [KK^* + \tilde{C}I]^{-1} KK^* g - g \right\| ,
\end{aligned}
$$

giving the equation $KK^*g + \tilde{C}g = KK^*g$. The terms $-\frac{2}{\alpha}\left(r(\alpha; g_\delta), g_\delta\right)$ and $\frac{2\delta}{\alpha}\rho(\alpha) = \frac{2\delta^2}{\alpha}$ thus will pose problems on the limiting upper bound unless $\frac{\delta^2}{\alpha}$ is bounded as $\delta \to 0$.[4] What is required is the regularity of Tikhonov's method with the discrepancy principle.[5]

Theorem 4.3 (Groetsch, 1984) *Let g and g_δ satisfy the hypotheses of theorem 4.2. Then Tikhonov's method with the discrepancy principle is a regularizing algorithm; i.e. $f^\delta_{\alpha(\delta)} \to K^\dagger g$ as $\delta \to 0$.*

Three results of functional analysis will be used in this proof. They are

1. the weak compactness property of a Hilbert space– i.e., every bounded sequence in a Hilbert space contains a weakly convergent subsequence;

---

[4] Note that

$$|(r(\alpha; g_\delta), g_\delta)| \le \|r(\alpha; g_\delta)\| \, \|g_\delta\| \le \delta\rho(\alpha) = \delta^2 .$$

[5] Recall that $\frac{\delta^2}{\alpha(\alpha)} \to 0$ as $\delta \to 0$ is only a sufficient condition for $f^\delta_{\alpha(\delta)} \to K^\dagger g$ as $\delta \to 0$. But if $\frac{\delta^2}{\alpha(\delta)} \to c > 0$, then it can be shown that $f^{\delta_n}_{\alpha(\delta_n)} \to K^\dagger g$ weakly for any sequence $\{\delta_n\}$ with $\delta_n \to 0$. In this case, $\left\{ f^{\delta_n}_{\alpha(\delta_n)} \right\}$ is said to be weakly regular.

Other possibilities of the limiting process are

1. $\lim_{\delta\to 0} \frac{\delta^2}{\alpha(\delta)} = \infty$;
2. $\lim_{\delta\to 0} \inf \frac{\delta^2}{\alpha(\delta)} = b > 0$ and $\lim_{\delta\to 0} \sup \frac{\delta^2}{\alpha(\delta)} = a < \infty$, $a \neq b$;
3. $\lim_{\delta\to 0} \inf \frac{\delta^2}{\alpha(\delta)} = 0$ and $\lim_{\delta\to 0} \sup \frac{\delta^2}{\alpha(\delta)} = a < \infty$, and
4. $\lim_{\delta\to 0} \inf \frac{\delta^2}{\alpha(\delta)} = d \ge 0$ and $\lim_{\delta\to 0} \sup \frac{\delta^2}{\alpha(\delta)} = \infty$.

The divergence of $\frac{\delta^2}{\alpha(\delta)}$ is a sufficient condition for the existence of a sequence of $\{\delta_n\}$ with $\delta_n \to 0$ and elements g_{δ_n} with $\|g - g_{\delta_n}\| \le \delta_n$ such that $\{f^{\delta_n}_{\alpha(\delta_n)}\}$ is not weakly convergent (Groetsch, 1984, p.25); and although the finite superior limit is a sufficient condition for the weak regularity of $\{f^{\delta_n}_{\alpha(\delta_n)}\}$ (Groetsch, 1984, pp. 23-24), the nonzero inferior limit is a sufficient condition for the existence of a sequence $\{\delta_n\}$ with $\delta_n \to 0$ and $\{g_{\delta_n}\}$ with $\|g - g_{\delta_n}\| \le \delta_n$ such that $\{f^{\delta_n}_{\alpha(\delta_n)}\}$ is not strongly convergent to $K^\dagger g$ (Groetsch, 1984, p.26). The fourth condition is sufficient for $\{f^{\delta_n}_{\alpha(\delta_n)}\}$ not to be strongly convergent to $K^\dagger g$. We can conclude that if Tikhonov's method with the discrepancy principle is an r.a., then $\frac{\delta^2}{\alpha(\delta)} \to 0$ as $\delta \to 0$.

Note that the weak regularity of $\{f^\delta_{\alpha(\delta)}\}$ will also guarantee that $E(\delta; \alpha)$ will be bounded.

2. the H-property of a Hilbert space– i.e., if $z_n \to z$ weakly and $\|z_n\| \to \|z\|$ in a Hilbert space, then $z_n \to z$ strongly; and
3. the weak lower semi-continuity of a convex functional– i.e., if z^* is a continuous convex functional, and if $z_n \to z$ weakly, then $\lim_{n\to\infty} \inf z^*(z_n) \geq z^*(z)$.

Proof: Introduce the functional

$$F_{\alpha(\delta)}(z; g_\delta) := \left\| K z^\delta_{\alpha(\delta)} - g_\delta \right\|^2_{\mathcal{H}_2} + \alpha(\delta) \left\| z^\delta_{\alpha(\delta)} \right\|^2_{\mathcal{H}_1},$$

which has a minimum if and only if

$$[K^*K + \alpha(\delta) I] z^\delta_{\alpha(\delta)} = K^* g_\delta .$$

Clearly then,

$$\begin{aligned} \|r(\alpha(\delta); g_\delta)\|^2_{\mathcal{H}_2} + \alpha(\delta) \left\| f^\delta_{\alpha(\delta)} \right\|^2_{\mathcal{H}_1} &= F_{\alpha(\delta)} \left(f^\delta_{\alpha(\delta)}; g_\delta \right) \\ &\leq F_{\alpha(\delta)} \left(f; g_\delta \right) \\ &= \|g - g_\delta\|^2_{\mathcal{H}_2} + \alpha(\delta) \|f\|^2_{\mathcal{H}_1} \\ &\leq \delta^2 + \alpha(\delta) \|f\|^2_{\mathcal{H}_1} . \end{aligned}$$

Now because the discrepancy principle is applied, $\|r(\alpha(\delta); g_\delta)\|^2 = \delta^2$ and consequently,

$$\left\| f^\delta_{\alpha(\delta)} \right\|^2 \leq \|f\|^2 \tag{4.6}$$

for any $\delta > 0$. This last inequality together with the weak compactness property imply that for any sequence $\{\delta_n\}$ with $\delta_n \to 0$, there exists a subsequence $\{\delta_{n_k}\}$ such that

$$f^{\delta_{n_k}}_{\alpha(\delta_{n_k})} \longrightarrow z \in \mathcal{H}_1 \text{ weakly.}$$

Moreover, $Kz = g$ since

$$\begin{aligned} g_{\delta_{n_k}} &\to g, \\ K f^{\delta_{n_k}}_{\alpha(\delta_{n_k})} &\to Kz, \qquad \text{(compactness of } K\text{)}, \\ \left\| K^* f^{\delta_{n_k}}_{\alpha(\delta_{n_k})} - g_{\delta_{n_k}} \right\| &= \delta_{n_k} \to 0 \qquad \text{(discrepancy principle)}. \end{aligned}$$

Using the fact that $f^{\delta_{n_k}}_{\alpha(\delta_{n_k})} \in R(K^*)$, $z \in \overline{R(K^*)} = N(K)^{\perp}$,[6] and hence

$$z = f = K^{\dagger} g.$$

But every subsequence of $\left\{f^{\delta_n}_{\alpha(\delta_n)}\right\}$ must itself satisfy (4.6) elementwise. Hence, every subsequence itself contains a subsequence which converges weakly to f. Hence, every subsequence $\left\{f^{\delta_n}_{\alpha(\delta_n)}\right\}$ must converge weakly to f for otherwise for some subsequence $\left\{f^{\delta_{n_k}}_{\alpha(\delta_{n_k})}\right\}$, for some functional f^*,

$$\lim_{n_k \to \infty} f^* \left(f^{\delta_{n_k}}_{\alpha(\delta_{n_k})}\right) \neq f^*(f).$$

However, letting

$$b_{n_k} := f^* \left(f^{\delta_{n_k}}_{\alpha(\delta_{n_k})}\right) \in \Re \qquad \text{and} \qquad b := f^*(f) \in \Re,$$

the last limit is equivalent to the existence of an open sphere about b containing no points of $\{b_{n_k}\}$– i.e., $\{b_{n_k}\}$ cannot contain a subsequence which converges weakly to b, or equivalently, $\left\{f^{\delta_{n_k}}_{\alpha(\delta_{n_k})}\right\}$ does not contain a subsequence which weakly converges to f.

We now have $f^{\delta_n}_{\alpha(\delta_n)} \to f$ weakly. We also have by the weak lower semi-continuity of $\|\cdot\|$ and by (4.6) that

$$\begin{aligned} \|f\| &\leq \lim_{n \to \infty} \inf \left\|f^{\delta_n}_{\alpha(\delta_n)}\right\| \\ &\leq \lim_{n \to \infty} \sup \left\|f^{\delta_n}_{\alpha(\delta_n)}\right\| \\ &\leq \|f\|. \end{aligned}$$

Thus, $\left\|f^{\delta_n}_{\alpha(\delta_n)}\right\| \to \|f\|$, and resorting to the H-property of $\mathcal{H}_1$,

$$f^{\delta}_{\alpha(\delta)} \to f \qquad \text{as } \delta \to 0.$$

For the case when $f \in R(K^*)$, a convergence rate for this strategy exists:

Theorem 4.4 (Groetsch, 1984) *If $f \in R(K^*)$, then $\left\|f - f^{\delta}_{\alpha(\delta)}\right\| = O\left(\sqrt{\delta}\right)$.*

[6] Take the functional $(\cdot, y)$ where $y \in N(K)$. Then

$$0 = \lim_{n_k \to \infty} \left(f^{\delta_{n_k}}_{\alpha(\delta_{n_k})}, y\right) = (z, y).$$

In fact, this convergence rate cannot be improved if K is of infinite rank, although it can be improved to $O(\delta)$ if K is of finite rank (Groetsch, 1984).

The major difficulty with the discrepancy principle is that the exact operator is usually not accessible for computation. To treat this deficiency, the principle is modified to the choosing of the parameter that satisfies

$$\rho^2_{(\eta,\delta)}(\alpha(\eta,\delta)) = \left(\delta + \eta \left\| f^{(\eta,\delta)}_{\alpha(\eta,\delta)} \right\|\right)^2 + \nu_\eta^2$$

(Goncharsky *et al* (1973), Goncharsky (1987)). Several aspects of this generalized discrepancy principle are given below.

Theorem 4.5 (Goncharsky *et al*, 1973) *For all η and δ such that $0 < \eta \le \eta_0$, $0 < \delta \le \delta_0$ for some constants η_0, δ_0, let (4.1) be satisfy. Then for every pair (η, δ), there exists a unique number $\alpha(\eta, \delta)$ which satisfies*

$$\rho^2_{(\eta,\delta)}(\alpha(\eta,\delta)) = \left(\delta + \eta \left\| f^{(\eta,\delta)}_{\alpha(\eta,\delta)} \right\|\right)^2 + \nu_\eta^2$$

Theorem 4.6 (Goncharsky *et al*, 1973) *Let $\{(\eta_n, \delta_n)\}$, $0 < \eta_n \le \eta_0$, $0 < \delta_n \le \delta_0$, be any sequence convergent to zero such that (4.1) is satisfied. Then Tikhonov's method with the generalized discrepancy principle is an r.a..*

Moreover, this modified scheme is "order-optimal" on some subset of $\mathcal{H}_1$.

Definition 4.2 (Goncharsky, 1987) *A mapping R_{δ^*} is optimal on E at a fixed $\delta = \delta^*$ if*

$$\sup_{x\in E} \Delta(R_{\delta^*}, \delta^*, x) = \inf_{R:X\to Y} \sup_{x\in E} \Delta(R, \delta^*, x).$$

The algorithm R_δ is called the optimal one on E if it is optimal for any $\delta > 0$.

Definition 4.3 *An algorithm R_δ is order-optimal in E if for any $\delta > 0$,*

$$\sup_{x\in E} \Delta(R_\delta, \delta, x) = \hat{C} \inf_{R:X\to Y} \sup_{x\in E} \Delta(R, \delta, x),$$

$\hat{C} \ge 1$ and independent of δ.

Theorem 4.7 (Goncharskii, 1987) *For linear first kind equations, Tikhonov's method with the generalized discrepancy principle is order-optimal on sets of the type $BS(0,R)$. Here, B is a compact operator from a reflexive space Z into $\mathcal{H}_1$, and $S(0,r) \subset Z$ is a sphere of sufficiently large radius r.*

Of course, defining a suitable operator B and a sphere $S(0,r)$ is quite a task in itself information about the solution need to be known. But it is not necessary to know the set $BS(0,r)$ to obtain this order-optimal convergence rate.

Turning to the iterative methods, analogues of the discrepancy principle will be considered. These principles will be applied to iterative methods of the class

$$f_{\alpha(\eta,\delta)}^{(\eta,\delta)} = \left[I - K_\eta^* K_\eta \hat{R}_\alpha(K_\eta^* K_\eta)\right] f_0 + \hat{R}_\alpha(K_\eta^* K_\eta) K_\eta^* g_\delta \,, \tag{4.7}$$

where $\hat{R}_\eta(t)$ satisfies

$$\sup_{0\le t\le \|K_\eta\|^2} \left|\hat{R}_\alpha(t)\right| \;\le\; \frac{\gamma}{\alpha}\,, \tag{4.8}$$

$$\sup_{0\le t\le \|K_\eta\|^2} t^p \left|1 - t\hat{R}_\alpha(t)\right| \;\le\; \gamma_p \alpha^p, \qquad 0 \le p \le p_0, \tag{4.9}$$

and where γ, γ_p, p_0 are positive constants, and f_0 is an initial approximation.

Using (4.7), we have

$$\begin{aligned} K_\eta f_{\alpha(\eta,\delta)}^{(\eta,\delta)} - g_\delta &= K_\eta \left[I - K_\eta^* K_\eta \hat{R}_\alpha(K_\eta^* K_\eta)\right] f_0 - \left[I - K_\eta \hat{R}_\alpha(K_\eta^* K_\eta) K_\eta^*\right] g_\delta \\ &= \left[I - K_\eta K_\eta^* \hat{R}_\alpha(K_\eta K_\eta^*)\right] K_\eta f_0 - \left[I - K_\eta K_\eta^* \hat{R}_\alpha(K_\eta K_\eta^*)\right] g_\delta \\ &= \left[I - K_\eta K_\eta^* \hat{R}_\alpha(K_\eta K_\eta^*)\right] (K_\eta f_0 - g_\delta)\,. \end{aligned} \tag{4.10}$$

Also, using (4.9) with the assumption that $p_0 > \frac{1}{2}$ (see theorem 4.10), for any $K_\eta w$,

$$\begin{aligned} \left\| \left[I - K_\eta K_\eta^* \hat{R}_\alpha(K_\eta K_\eta^*)\right] K_\eta w \right\| &= \left\| K_\eta \left[I - K_\eta^* K_\eta \hat{R}_\alpha(K_\eta^* K_\eta)\right] w \right\| \\ &\le \left[\sup_{0\le t\le \|K_\eta\|^2} t^{\frac{1}{2}} \left|1 - t\hat{R}_\alpha(t)\right| \right] \|w\| \\ &\le \gamma_{\frac{1}{2}} \alpha^{\frac{1}{2}} \|w\| \longrightarrow 0 \qquad \text{as } \alpha \to 0. \end{aligned}$$

An application of the Banach-Steinhaus theorem[7] then shows that

$$\left[I - K_\eta K_\eta^* \hat{R}_\eta(K_\eta K_\eta^*)\right] z \longrightarrow 0 \qquad \text{as } \alpha \to 0 \text{ for all } z \in \overline{R(K_\eta)}. \tag{4.11}$$

Now with results (4.10) and (4.11) and the decomposition of g_δ, we have

$$\begin{aligned}
\left\| K_\eta f_{\alpha(\eta,\delta)}^{(\eta,\delta)} - g_\delta \right\| &= \left\| \left[I - K_\eta K_\eta^* \hat{R}_\alpha(K_\eta K_\eta^*)\right](K_\eta f_0 - g_\delta) \right\| \\
&\le \left\| \left[I - K_\eta K_\eta^* \hat{R}_\alpha(K_\eta K_\eta^*)\right](K_\eta f_0 - \omega_\eta) \right\| + \left\| \left[I - K_\eta K_\eta^* \hat{R}_\alpha(K_\eta K_\eta^*)\right] \nu_\eta \right\| \\
&= \left\| \left[I - K_\eta K_\eta^* \hat{R}_\alpha(K_\eta K_\eta^*)\right](K_\eta f_0 - \omega_\eta) \right\| + \left\| \nu_\eta - K_\eta \hat{R}_\alpha(K_\eta^* K_\eta) K_\eta^* \nu_\eta \right\| \\
&= \left\| \left[I - K_\eta K_\eta^* \hat{R}_\alpha(K_\eta K_\eta^*)\right](K_\eta f_0 - \omega_\eta) \right\| + \|\nu_\eta\| \\
&\longrightarrow \inf_{z \in \mathcal{H}_1} \|K_\eta z - g_\delta\| \qquad \text{as } \alpha \to 0.
\end{aligned} \tag{4.12}$$

Two other central facts are

$$\left\| K_\eta^* \left(K_\eta f_{\alpha(\eta,\delta)}^{(\eta,\delta)} - g_\delta \right) \right\| \to 0 \qquad \text{as } \alpha \to 0, \tag{4.13}$$

and

$$\inf_{z \in \mathcal{H}_1} \|K_\eta z - g_\delta\| \le \delta + \|f\| \eta \tag{4.14}$$

These two results and (4.12) are the bases for the following stopping criteria for (4.7) (Vainikko (1982,1983)):

1. Choose numbers b_1, $b_2 \ge b_1$ and $\theta \in (0,1)$. If for $\alpha^*(\eta,\delta) \gg 1$,

$$\left\| K_\eta f_{\alpha^*(\eta,\delta)}^{(\eta,\delta)} - g_\delta \right\| \le b_2 \left(\delta + \|f_*\| \eta\right), \tag{4.15}$$

where f_* is the exact least squares solution closest to f_0, then take $\alpha = \alpha^*$. Else choose an $\alpha(\eta,\delta)$ such that both (4.15) is satisfied, and for some $\tau \in [\alpha, \frac{\alpha}{\theta}]$,

$$\left\| K_\eta f_{\tau(\eta,\delta)}^{(\eta,\delta)} - g_\delta \right\| \ge b_1 \left(\delta + \|f_*\| \eta\right) \tag{4.16}$$

is satisfied.

---

[7] Banach-Steinhaus Theorem: The following conditions are together necessary and sufficient for a sequence of continuous linear operators $\{T_n\}$, mapping a Banach space X into a Banach space Y, to converge on X to a linear operator:

1. $$\|T_n\| \le M \qquad n = 1, 2, \ldots;$$

2. $\{T_n x\}$ is a Cauchy sequence for each x in some dense subset E of X.

2. Similar to (1.) above, but with (4.15) and (4.16) replaced by

$$\left\| K_\eta f^{(\eta,\delta)}_{\alpha^*(\eta,\delta)} - g_\delta \right\| \le b_2 \left(\delta + \left\| f^{(\eta,\delta)}_{\alpha^*(\eta,\delta)} \right\| \eta \right), \tag{4.17}$$

and

$$\left\| K_\eta f^{(\eta,\delta)}_{\tau(\eta,\delta)} - g_\delta \right\| \ge b_1 \left(\delta + \left\| f^{(\eta,\delta)}_{\tau(\eta,\delta)} \right\| \eta \right) \tag{4.18}$$

respectively.

3. Also similar to (1.), but now with the replacements being

$$\left\| K_\eta f^{(\eta,\delta)}_{\alpha^*(\eta,\delta)} - g_\delta \right\| \le b_2 \left[\delta + \left(\left\| f^{(\eta,\delta)}_{\alpha^*(\eta,\delta)} - f_0 \right\| + \|f_0\| \right) \eta \right] \tag{4.19}$$

and

$$\left\| K_\eta f^{(\eta,\delta)}_{\tau(\eta,\delta)} - g_\delta \right\| \ge b_1 \left[\delta + \left(\left\| f^{(\eta,\delta)}_{\tau(\eta,\delta)} - f_0 \right\| + \|f_0\| \right) \eta \right]. \tag{4.20}$$

4. Choose numbers $b_1 > 0$, $b_2 \ge b_1$ and $\theta \in (0,1)$. If for $\alpha^*(\eta,\delta) \gg 1$,

$$\left\| K_\eta^* \left(K_\eta f^{(\eta,\delta)}_{\alpha^*(\eta,\delta)} - g_\delta \right) \right\| \le b_2 (\delta + \eta), \tag{4.21}$$

then take $\alpha = \alpha^*$. Otherwise, choose any α that satisfies both (4.21) and

$$\left\| K_\eta^* \left(K_\eta f^{(\eta,\delta)}_{\tau(\eta,\delta)} - g_\delta \right) \right\| \ge b_1 (\delta + \eta), \tag{4.22}$$

for some $\tau \in \left[\alpha, \frac{\alpha}{\theta}\right]$.

Criteria (2.) and (3.) are further stipulated with the condition that

$$\alpha \ge \frac{1}{\beta} (\delta + \eta)^2,$$

and if these α's do not satisfy the appropriate stopping condition, then α is choosen to be $\frac{1}{\beta}(\delta+\eta)^2$. Here, $\beta > 0$ is a fixed constant.

Indeed, iteration (4.7) together with any of these stopping criteria is an r.a.. This is elaborated in the concluding theorems of this section.

Theorem 4.8 (Vainikko, 1982) *Let $g \in R(K)$, and let (4.8)–(4.9) hold. Then iteration (4.7) with $\hat{R}_\alpha(t) : \left[0, \|K_\eta\|^2\right] \to \mathcal{C}$, and with $\alpha(\eta,\delta)$ chosen such that*

$$\alpha(\eta,\delta) \to 0, \quad \frac{(\delta+\eta)^2}{\alpha(\eta,\delta)} \to 0 \qquad \text{as } \delta,\ \eta \to 0,$$

is an r.a..

Theorem 4.9 (Vainikko, 1982) *Let $P_{\overline{R(K)}}g \in R(K)$, and let (4.8)-(4.9) hold. Then iteration (4.7) with $\hat{R}_\alpha(t) : \left[0, \|K_\eta\|^2\right] \to C$, and with $\alpha(\eta,\delta)$ chosen such that*

$$\alpha(\eta,\delta) \to 0, \quad \frac{(\delta^2+\eta)}{\alpha(\eta,\delta)} \to 0 \qquad \text{as } \delta,\ \eta \to 0,$$

is an r.a..

Theorem 4.10 (Vainikko, 1982) *Let $g \in R(K)$, and let (4.8)-(4.9) hold with $p_0 > \frac{1}{2}$, $\gamma_0 = 1$. Let the parameter $\alpha(\eta,\delta)$ be chosen according to any of the criteria (1.)-(3.) with the assumption that $N(K) = 0$ or else*

$$b_1 \left[1 - (\gamma\gamma'\beta)^{\frac{1}{2}} \max\left\{\frac{\|K\|}{\|g\|}, 1\right\}\right] > 1$$

where

$$\gamma' := \sup_{0 \le t \le \|K\|^2} t\left|\hat{R}_\alpha(t)\right| \le 1 + \gamma_0$$

for criterion (2.). Then

$$\frac{(\delta+\eta)^2}{\alpha(\eta,\delta)} \to 0, \quad \left\|f^{(\eta,\delta)}_{\alpha(\eta,\delta)} - K^\dagger g\right\| \to 0 \qquad \text{as } \delta,\ \eta \to 0.$$

Theorem 4.11 (Vainikko, 1982) *Let $P_{\overline{R(K)}}g \in R(K)$, and let (4.8)-(4.9) hold with $p_0 > 1$. Let the parameter $\alpha(\eta,\delta)$ be chosen according to criterion (4.) with $b_1 > \gamma_0 \left\|g - P_{\overline{R(K)}}g\right\|$. Then*

$$\frac{(\delta+\eta)}{\alpha(\eta,\delta)} \to 0, \quad \left\|f^{(\eta,\delta)}_{\alpha(\eta,\delta)} - K^\dagger g\right\| \to 0 \qquad \text{as } \delta,\ \eta \to 0.$$

4.2 Landweber Iteration

We now take a closer look at the Landweber iteration and propose a multilevel adaption to it.

This iteration is essentially the successive approximation procedure applied to the least squares equation

$$K^*Kf = K^*g.$$

It is known that this iteration produces approximations which converge to $K^\dagger g + f_{0_{N(K)}}$ if $g \in D(K^\dagger)$ (Strand (1972 and 1973)). In particular, the approximations converge to $K^\dagger g$ if $f_0 \in R(K^*)$. However, it is also true that $\|f_k\| \to \infty$ when $g \notin D(K^\dagger)$. For ill-posed problems then, the Landweber iteration cannot be applied without employing some sophisticated stopping criterion which logically should depend on the magnitude of the data error. What really is required is a stopping criterion which together with the Landweber iteration forms an r.a..

Having employed such a stopping condition, this r.a. filters out the error-source high frequency components; that is, n is automatically chosen to allow only a compromising amount of high frequency components to enter into the approximation. To see this for the exact operator case, recall that

$$K^\dagger g = \sum_{j=1}^{\infty} \mu_j (g, u_j) v_j \,,$$

and

$$K^* g_\delta = \sum_{j=1}^{\infty} {\mu_j}^{-1} (g_\delta, u_j) v_j .$$

Now if $f_0^\delta \in R(K^*)$, say

$$f_0^\delta = K^* w = \sum_{j=1}^{\infty} {\mu_j}^{-1} (w, u_j) v_j ,$$

then

$$\begin{aligned}
f_n^\delta &= f_{n-1}^\delta + K^* \left(g_\delta - K f_{n-1}^\delta \right) \\
&= (I - K^*K)^n f_0^\delta + \sum_{i=0}^{n-1} (I - K^*K)^i K^* g_\delta \\
&= (I - K^*K)^n \sum_{j=1}^{\infty} {\mu_j}^{-1} (w, u_j) v_j + \sum_{i=0}^{n-1} (I - K^*K)^i \sum_{j=1}^{\infty} {\mu_j}^{-1} (g_\delta, u_j) v_j \\
&= \sum_{j=1}^{\infty} {\mu_j}^{-1} (w, u_j) \left(1 - {\lambda_j}^2 \right)^n v_j + \sum_{i=0}^{n-1} \sum_{j=1}^{\infty} {\mu_j}^{-1} \left(1 - {\lambda_j}^2 \right)^i (g_\delta, u_j) v_j \\
&= \sum_{j=1}^{\infty} \lambda_j \left(1 - {\lambda_j}^2 \right)^n (w, u_j) v_j + \sum_{j=1}^{\infty} \sum_{i=0}^{n-1} \lambda_j \left(1 - {\lambda_j}^2 \right)^i (g_\delta, u_j) v_j .
\end{aligned}$$

Taking $f_0^\delta = 0$,

$$f_n^\delta = \sum_{j=1}^{\infty} \sum_{i=0}^{n-1} \lambda_j \left(1 - {\lambda_j}^2\right)^i (g_\delta, u_j) v_j ,$$

and a further iteration adds the term

$$\sum_{j=1}^{\infty} \lambda_j \left(1 - {\lambda_j}^2\right)^n (g_\delta, u_j) v_j$$

to f_n^δ. Now since

$$\begin{aligned} \lambda_j \left(1 - {\lambda_j}^2\right)^n &= \lambda_j \left[1 - n{\lambda_j}^2 + \frac{n(n-1)}{2}{\lambda_j}^4 - \frac{n(n-1)(n-2)}{6}{\lambda_j}^6 + \cdots + (-1)^n {\lambda_j}^{2n}\right] \\ &= \lambda_j - n{\lambda_j}^3 + O\left({\lambda_j}^5\right) \\ &\approx \lambda_j \end{aligned} \tag{4.23}$$

for large j's, f_{n+1}^δ contains an additional multiple of about $\lambda_j(g_\delta, u_j)$ of each high frequency v_j. Furthermore, because $\sum_{i=0}^{n-1} \lambda_j \left(1 - {\lambda_j}^2\right)^i$ is the product of λ_j and the $(n-1)$'st Taylor polynomial of $\frac{1}{x}$ expanded about 1 and evaluated at ${\lambda_j}^2$,

$$f_n^\delta \longrightarrow \sum_{j=1}^{\infty} \mu_j(g, u_j) v_j .$$

An r.a. producing stopping criterion hence would produce a good approximation by limiting the growth of high frequency components.

For first kind Fredholm integral equations, the iteration is

$$f_n^\delta(s) = f_{n-1}^\delta(s) + \int_a^b k(v, s) \left[g_\delta(t) - \int_a^b k(v, t) f_{n-1}^\delta(t)\, dt \right] dv . \tag{4.24}$$

Equation (4.24) is solved numerically after being discretized into the matrix-vector iteration

$$\underline{f}_{n,h}^\delta = \underline{f}_{n-1,h}^\delta + hK_{hh}^t \left[\underline{g}_{\delta h} - hK_{hh} \underline{f}_{n-1,h}^\delta \right], \tag{4.25}$$

where h is the stepsize of the discretization, and K_{hh} is the discrete version of the kernel with stepsize h in both variables. The question that naturally arises is how well the solutions of (4.24) and (4.25) agree. Equivalently, the question is how well the singular system of the integral operator K is approximated. One noticeable problem is that if K is of infinite rank, then its infinite singular system must be approximated by the finite singular system of K_{hh}.

We will not delve into this or any other problems arising from the discretization of (4.24). But we will require the discretization to be dependent on the quadrature error which will be arranged to be $O(\delta)$.

In passing, $\underline{f}^{\delta}_{n,h}$ and $f^{\delta}_{n}(s)$ do agree to an extent determined by the quadrature error. Since the generating function of the Landweber iteration is

$$\hat{R}_n(t) = \frac{1}{t}\left[1-(1-t)^n\right],$$

and assuming that $\|K\|^2 \leq 1$ and $\|K_{hh}\|^2 \leq 1$ (conditions which always can be fulfilled by some scaling),

$$\begin{aligned}\sup_{0\leq t\leq 1} t^p \left|1-t\hat{R}_n(t)\right| &= \sup_{0\leq t\leq 1} t^p \left|(1-t)^n\right| \\ &= \begin{cases} 1=\gamma_0 & \text{for } p=0 \\ \left(\frac{p}{p+n}\right)^p \left(\frac{n}{p+n}\right)^n \leq p^p n^{-p} = \gamma_p n^{-p} & \text{for } p>0. \end{cases}\end{aligned}$$

Hence, theorem 4.10 is applicable, and so both

$$\left\| f^{\delta}_{n}(s) - K^{\dagger}g \right\| \longrightarrow 0 \qquad \text{as } \delta \to 0\,,$$

and

$$\left\| \underline{f}^{\delta}_{n,h} - K^{\dagger}g \right\| \longrightarrow 0 \qquad \text{as the quadrature error and } \delta \to 0\,.$$

Returning to equations (4.24) and (4.25) with the assumption that $g \in D(K^{\dagger})$, note that the errors after k iterations are respectively

$$\begin{aligned} e^{\delta}_{k}(s) &:= K^{\dagger}g(s) - f^{\delta}_{k}(s) \\ &= \sum_{j=1}^{\infty} \mu_j (g,u_j) v_j - \sum_{j=1}^{\infty} \left\{ \sum_{i=0}^{k-1} \lambda_j \left(1-{\lambda_j}^2\right)^i \right\} (g_\delta, u_j) v_j \\ &= \sum_{j=1}^{\infty} \mu_j (g,u_j) v_j - \sum_{j=1}^{\infty} \left\{ \sum_{i=0}^{k-1} \lambda_j \left(1-{\lambda_j}^2\right)^i \right\} (g+\delta g, u_j) v_j \\ &= \sum_{j=1}^{\infty} \mu_j (g,u_j) v_j - \sum_{j=1}^{\infty} \mu_j \left[1-\left(1-{\lambda_j}^2\right)^k\right] (g+\delta g, u_j) v_j \\ &= \sum_{j=1}^{\infty} \mu_j \left(1-{\lambda_j}^2\right)^k (g,u_j) v_j - \sum_{j=1}^{\infty} \mu_j \left[1-\left(1-{\lambda_j}^2\right)^k\right] (\delta g, u_j) v_j\,, \end{aligned}$$

and

$$\begin{aligned}\underline{e}_{kh}^{\delta} &= K_{hh}^{\dagger}\underline{g}_h - \underline{f}_{kh}^{\delta} \\ &= \sum_{j=1}\mu_{jh}\left(1-{\lambda_{jh}}^2\right)^k(\underline{g}_h,\underline{u}_{jh})\underline{v}_{jh} - \sum_{j=1}\mu_{jh}\left[1-\left(1-{\lambda_{jh}}^2\right)^k\right](\underline{\delta g}_h,\underline{u}_{jh})\underline{v}_{jh}\,.\end{aligned}$$

Regularization enforces some weighting between the two sums of each of the above expressions. With this view in mind, another approach to the iteration can be derived.

Suppose the Landweber iteration is applied to the first kind equation $K\hat{e}_k^{\delta} = \hat{d}_{\delta} := g - Kf_k^{\delta}$. After m iterations, we have the approximation

$$\begin{aligned}\hat{e}_{km}^{\delta} &= \sum_{j=1}^{\infty}\left\{\sum_{i=0}^{m-1}\lambda_j\left(1-{\lambda_j}^2\right)^i\right\}(\hat{d}_{\delta},u_j)v_j \\ &= \sum_{j=1}^{\infty}\mu_j\left[1-\left(1-{\lambda_j}^2\right)^m\right](g,u_j)v_j - \sum_{j=1}^{\infty}\mu_j\left[1-\left(1-{\lambda_j}^2\right)^m\right](Kf_k^{\delta},u_j)v_j \\ &= \sum_{j=1}^{\infty}\mu_j\left[1-\left(1-{\lambda_j}^2\right)^m\right](g,u_j)v_j \\ &\quad -\sum_{j=1}^{\infty}\left[1-\left(1-{\lambda_j}^2\right)^m\right]\left\{\sum_{p=1}^{\infty}\mu_p\left[1-\left(1-{\lambda_p}^2\right)^k\right](g_{\delta},u_p)(v_p,v_j)\right\}v_j \\ &= \sum_{j=1}^{\infty}\mu_j\left[1-\left(1-{\lambda_j}^2\right)^m\right](g,u_j)v_j \\ &\quad -\sum_{j=1}^{\infty}\left[1-\left(1-{\lambda_j}^2\right)^m\right]\left\{\mu_j\left[1-\left(1-{\lambda_j}^2\right)^k\right](g_{\delta},u_j)\right\}v_j \\ &= \sum_{j=1}^{\infty}\mu_j\left[1-\left(1-{\lambda_j}^2\right)^m\right](g,u_j)v_j\left[1-1+\left(1-{\lambda_j}^2\right)^k\right] \\ &\quad -\sum_{j=1}^{\infty}\mu_j\left[1-\left(1-{\lambda_j}^2\right)^m\right]\left[1-\left(1-{\lambda_j}^2\right)^k\right](\delta g,u_j)v_j \\ &= \sum_{j=1}^{\infty}\mu_j\left(1-{\lambda_j}^2\right)^k\left[1-\left(1-{\lambda_j}^2\right)^m\right](g,u_j)v_j \\ &\quad -\sum_{j=1}^{\infty}\mu_j\left[1-\left(1-{\lambda_j}^2\right)^m\right]\left[1-\left(1-{\lambda_j}^2\right)^k\right](\delta g,u_j)v_j\end{aligned}$$

As $m \to \infty$, this expression tends to e_k^δ, and if regularization is performed, then $f_k^\delta + \hat{e}_{km}^\delta$ should be a good approximation to the regularized solution of $Kf = g_\delta$.

However, $\hat{d}_\delta$ is usually not available, but rather

$$d_\delta := g_\delta - Kf_k^\delta$$

is. Applying the Landweber iteration with this forcing term produces

$$\begin{aligned} e_{km}^\delta &= \sum_{j=1}^{\infty} \mu_j \left(1-{\lambda_j}^2\right)^k \left[1-\left(1-{\lambda_j}^2\right)^m\right](g,u_j)v_j \\ &\quad + \sum_{j=1}^{\infty} \mu_j \left[1-\left(1-{\lambda_j}^2\right)^m\right]\left(1-{\lambda_j}^2\right)^k (\delta g,u_j)v_j \\ &= \hat{e}_{km}^\delta + \sum_{j=1}^{\infty} \mu_j \left[1-\left(1-{\lambda_j}^2\right)^m\right](\delta g,u_j)v_j\,, \end{aligned}$$

which does not approach e_k^δ, and in fact, may diverge because the sum may diverge. This infers a quicker halting of the r.a. for the present equation than the previous defect equation. Nevertheless, for regularized values of m, $f_k^\delta + e_{km}^\delta$ should not be too poor of an approximation to the regularized solution of $Kf_\delta = g_\delta$.

With the difference of the regularizations of the two defect equations stated, it is questionable whether this approach to the Landweber iteration is practically helpful. But it is. First, two ideas that may alleviate the difference are to apply a different r.a. to the defect equation, and to have a different f_k^δ – i.e., f_k^δ is not the k'th iteration of the Landweber procedure with starting approximation $f_0^\delta = 0$. The first idea is actually required in a multigrid procedure, while the second idea will always be the case if an original equation–defect equation cycle is performed. Second, this approach opens up the realm of multilevel processing to some ill-posed problems, which we investigate next.

The ideas of multilevel processing start with the construction of the defect equation. Let f_k be the k'th iteration of the Landweber procedure applied to the exact equation

$$Kf = g, \qquad g \in R(K^\dagger). \tag{4.26}$$

f_k may be considered to be the least squares solution of

$$Kf = g - d, \qquad d \in R(K^\dagger). \tag{4.27}$$

Also let f_k^δ be the k'th iteration corresponding to

$$Kf = g_\delta \,. \tag{4.28}$$

We have

$$\begin{aligned} \|d - d_\delta\| &= \left\| g - Kf_k - g_\delta + Kf_k^\delta \right\| \\ &\leq \|g - g_\delta\| + \left\| K\left(f_k - f_k^\delta\right) \right\| \\ &\leq \delta + \delta M \qquad \text{(lemma 3.9)} \\ &= 2\delta \,, \end{aligned}$$

which enables the r.a. to stop sooner.

The same analysis and results hold for the discretized versions of (4.26)–(4.28). But now the defect equation can be solved on a coarser grid where the singular system of K can be approximated almost as well as on the fine grid. However, if we denote the elements of the fine grid equation with subscript h and those of the coarse grid equation with subscript H, then a major problem with this multilevel procedure is that any component of the interpolated coarse grid solution belonging in $N(K_{hh})$ will always remain in the regularized solution of $K_{hh}\underline{f}_h = \underline{g}_{\delta h}$. The interpolated solution therefore must be in $R(K_{hh}{}^*)$, a requirement that cannot be easily satisfied if the Landweber iteration is performed on the defect equation.

Suppose that instead of solving the least squares equation

$$K_{HH}{}^* K_{HH} \underline{e}_{kH}^\delta = K_{HH}{}^* \underline{d}_{\delta H} \,,$$

the equation

$$\left[K_{HH}{}^* K_{HH} + \alpha I_{HH}\right] \underline{e}_{kH}^\delta = K_{HH}{}^* \underline{d}_{\delta H}$$

with suitably chosen parameter α is solved. Then

$$\underline{e}_{kH}^\delta = \frac{1}{\alpha}\left[K_{HH}{}^* \underline{d}_{\delta H} - K_{HH}{}^* K_{HH} \underline{e}_{kH}^\delta\right] \,,$$

which suggests the interpolation

$$I_H{}^h \underline{e}_{kH}^\delta = \frac{1}{\alpha}\left[K_{hh}{}^* \underline{d}_{\delta h} - K_{hh}{}^* K_{hH} \underline{e}_{kH}^\delta\right] \,. \tag{4.29}$$

Now using (4.29), a two-grid iteration matrix describing the new procedure is

$$M_h := \left\{ I_{hh} - {I_H}^h \left[{K_{HH}}^* K_{HH} + \alpha I_{HH} \right]^{-1} {K_{HH}}^* {I_h}^H K_{hh} \right\} {S_{hh}}^m , \tag{4.30}$$

where I_{aa} is the identity matrix on grid $a = h, H$, ${I_H}^h$ is the projection operator between grids H and h, and ${S_{hh}}^m := [I_{hh} - {K_{hh}}^* K_{hh}]^m$ is the m'th iterate of the Landweber iteration matrix. The full iteration is

$$\underline{f}^{\delta}_{kh} = {S_{hh}}^m \underline{f}^{\delta}_{k-1,h} + {I_H}^h \left[{K_{HH}}^* K_{HH} + \alpha I_{HH} \right]^{-1} {K_{HH}}^* {I_h}^H \left(\underline{g}_{\delta h} - K_{hh} {S_{hh}}^m \underline{f}^{\delta}_{k-1,h} \right) . \tag{4.31}$$

The braced part of (4.30) can be viewed as an approximation to the iteration matrix

$$T_h := \left\{ I_{hh} - \left[{K_{hh}}^* K_{hh} + \alpha I_{hh} \right]^{-1} {K_{hh}}^* K_{hh} \right\} , \tag{4.32}$$

which corresponds to the iterated Twomey–Tikhonov scheme

$$\underline{f}^{\delta}_{kh} = \underline{f}^{\delta}_{k-1,h} + \left[{K_{hh}}^* K_{hh} + \alpha I_{hh} \right]^{-1} {K_{hh}}^* \left(\underline{g}_{\delta h} - K_{hh} \underline{f}^{\delta}_{k-1,h} \right) . \tag{4.33}$$

(4.33) captures the high frequency components more rapidly than the Landweber iteration does as can be seen in its k'th iterate with $\underline{f}^{\delta}_{0h} = 0$:

$$\underline{f}^{\delta}_{kh} = \sum_{j=1} \mu_{jh} \left[1 - \left(1 - \frac{1}{1 + \alpha {\mu_j}^2} \right)^k \right] \left(\underline{g}_{\delta h}, \underline{u}_{jh} \right) \underline{v}_{jh} .$$

Hence, if an appropriate α is chosen and if an r.a. producing stopping criterion is used, then the Twomey–Tikhonov scheme should require only a few iterations.

Of course, this quicker capturing of relevant components is what we want (4.31) to achieve. A look at the Landweber residual shows that this is quite possible:

$$\left\| K \underline{f}^{\delta}_{kh} - \underline{g}_{\delta h} \right\| = \left\| \sum_{j=1} \left(1 - {\lambda_j}^2 \right)^k \left(\underline{g}_{\delta h}, \underline{u}_{jh} \right) \underline{u}_{jh} \right\| .$$

After each iteration, very little is reduced for λ_j's≈ 0, but a moderate reduction is possible for moderately sized λ_j's. So, to attain the stopping criterion bound, it is the slow capturing of medium frequency components that creates the problem– frequencies that may be approximated well on the coarse grid since

$$\underline{e}^{\delta}_{kH} = \sum_{j=1} \frac{\mu_{jH}}{1 + \alpha {\mu_{jH}}^2} \left({I_h}^H \underline{d}_{\delta h}, \underline{u}_{jH} \right) \underline{v}_{jH} .$$

But the two-grid scheme is ineffective if the solution is composed mainly of high frequency components. This is true even for an ideal multigrid which consists of a restriction operator that projects the defect closely onto the space spanned by the low–medium frequencies of K_{HH}, and an interpolation operator that interpolates the regularized error closely into the corresponding space of K_{hh}. Little is achieved because the coarse grid procedure cannot capture the high frequencies well. In this case, the coarse grid step should be stopped, and further Landweber iterations should be performed on the fine grid. To recognize such a switch, the residual can be monitored: if the residual does not change much after a coarse-grid correction, then only additional Landweber iterations on the fine grid should be performed.

Turning to the parameter choice and the computational costs, α should not be too small to permit magnification of rounding errors. A good choice can be obtained cheaply on a grid coarser than H. Letting this grid be $4h$ and letting $H = 2h$, the cost breakdown is

$2n^2$ per Landweber iteration,
$\frac{n^3}{12}$ for the formation and Cholesky decomposition of $[K_{HH}{}^t K_{HH} + \alpha I_{HH}]$
$\frac{n^2}{4}$ for the adjustment $K_{HH}{}^* \underline{d}_{\delta H}$,
$\frac{5n^2}{2}$ for the interpolation, and
$\approx \frac{n^3}{64}$ for the refinement of α.

(The cost of the restriction has been omitted since only the trivial injection has been applied in the numerical examples.)

4.3 Tikhonov Regularization

The realization of the Tikhonov procedure with the discrepancy principle is slightly more involved than the Landweber iteration. It may consists of solving two linear equations which are nested in a Newton iteration:

given $\alpha_0(\eta, \delta)$, both

$$[K_\eta{}^* K_\eta + \alpha_0(\eta, \delta) I] f^{(\eta,\delta)}_{\alpha_0(\eta,\delta)} = K_\eta{}^* g_\delta \tag{4.34}$$

and

$$[K_\eta{}^* K_\eta + \alpha_0(\eta,\delta) I] \frac{df^{(\eta,\delta)}_{\alpha(\eta,\delta)}}{d\alpha}_{|\alpha_0(\eta,\delta)} = -f^{(\eta,\delta)}_{\alpha_0(\eta,\delta)} \tag{4.35}$$

are solved, and then $\alpha_0(\eta,\delta)$ is updated according to

$$\begin{aligned}
\alpha_k(\eta,\delta) &= \alpha_{k-1}(\eta,\delta) - \frac{\rho^2_{(\eta,\delta)}(\alpha_{k-1}(\eta,\delta)) - \left(\delta + \eta \left\| f^{(\eta,\delta)}_{\alpha_{k-1}(\eta,\delta)} \right\|\right)^2 - \nu_\eta{}^2}{\frac{d}{d\alpha}\left[\rho^2_{(\eta,\delta)}(\alpha(\eta,\delta)) - \left(\delta + \eta \left\| f^{(\eta,\delta)}_{\alpha(\eta,\delta)} \right\|\right)^2\right]_{|\alpha_{k-1}(\eta,\delta)}} \\
&= \alpha_{k-1}(\eta,\delta) \\
&\quad + \frac{\rho^2_{(\eta,\delta)}(\alpha_{k-1}(\eta,\delta)) - \left(\delta + \eta \left\| f^{(\eta,\delta)}_{\alpha_{k-1}(\eta,\delta)} \right\|\right)^2 - \nu_\eta{}^2}{2\left(\alpha_{k-1}(\eta,\delta) + \eta^2 + \frac{\delta\eta}{\left\| f^{(\eta,\delta)}_{\alpha_{k-1}(\eta,\delta)} \right\|}\right)\left(\frac{df^{(\eta,\delta)}_{\alpha(\eta,\delta)}}{d\alpha}_{|\alpha_{k-1}(\eta,\delta)}, f^{(\eta,\delta)}_{\alpha_{k-1}(\eta,\delta)}\right)}
\end{aligned} \tag{4.36}$$

Note that equations (4.34) and (4.35) must be solved for each Newton iteration. Note also that some simplification is achieved if

$$\nu_\eta{}^2 = 0\,, \tag{4.37}$$

for then the recurrence reduces to

$$\begin{aligned}
\alpha_k(\eta,\delta) &= \alpha_{k-1}(\eta,\delta) - \frac{\rho_{(\eta,\delta)}(\alpha_{k-1}(\eta,\delta)) - \left(\delta + \eta \left\| f^{(\eta,\delta)}_{\alpha_{k-1}(\eta,\delta)} \right\|\right)}{\frac{d}{d\alpha}\left[\rho_{(\eta,\delta)}(\alpha(\eta,\delta)) - \eta \left\| f^{(\eta,\delta)}_{\alpha(\eta,\delta)} \right\|\right]_{|\alpha_{k-1}(\eta,\delta)}} \\
&= \alpha_{k-1}(\eta,\delta) \\
&\quad + \frac{\rho_{(\eta,\delta)}(\alpha_{k-1}(\eta,\delta)) - \left(\delta + \eta \left\| f^{(\eta,\delta)}_{\alpha_{k-1}(\eta,\delta)} \right\|\right)}{\left(\frac{\alpha_{k-1}(\eta,\delta)}{\rho_{(\eta,\delta)}(\alpha_{k-1}(\eta,\delta))} + \frac{\eta}{\left\| f^{(\eta,\delta)}_{\alpha_{k-1}(\eta,\delta)} \right\|}\right)\left(\frac{df^{(\eta,\delta)}_{\alpha(\eta,\delta)}}{d\alpha}_{|\alpha_{k-1}(\eta,\delta)}, f^{(\eta,\delta)}_{\alpha_{k-1}(\eta,\delta)}\right)}\,.
\end{aligned} \tag{4.38}$$

We henceforth assume that (4.37) holds.

One of the major problems of (4.38) is that the function

$$h_1(\alpha(\eta,\delta)) := \rho_{(\eta,\delta)}(\alpha(\eta,\delta)) - \delta - \eta \left\| f^{(\eta,\delta)}_{\alpha(\eta,\delta)} \right\|$$

is not necessarily convex. Thus, Newton's method may be sensitive to the initial approximation. However, for the exact operator case, the modified equation

$$h_2\left(\alpha(\delta)\right) := h_1\left(\frac{1}{\alpha(\delta)}\right) = \rho_\delta\left(\frac{1}{\alpha(\delta)}\right) - \delta$$

is convex (Gordonova and Morovoz(1973), Morovoz(1984)), and so Newton's method will converge for any starting $\alpha > 0$. The recurrence becomes

$$\alpha_k(\delta) = \alpha_{k-1}(\delta) - \frac{\left[\rho_\delta\left(\frac{1}{\alpha_{k-1}(\delta)}\right) - \delta\right] \rho_\delta\left(\frac{1}{\alpha_{k-1}(\delta)}\right) \alpha_{k-1}^3(\delta)}{\left(\frac{df^\delta_{\alpha(\delta)}}{d\alpha}|_{\frac{1}{\alpha_{k-1}(\delta)}}, f^\delta_{\frac{1}{\alpha_{k-1}(\delta)}}\right)}. \tag{4.39}$$

Turning to operation counts, the count for this algorithm can be very large. For example, for first kind Fredholm integral equations, it is

$\frac{n^3}{2}$ – formation of $K_{hh}{}^tK_{hh}$, and
$j\frac{n^3}{6}$ – Cholesky factorizations after $(j-1)$ Newton iterations.

This large value reveals the problem of the standard approach: a direct solver is used for each iteration even though the solution of the previous iteration may be of sufficient accuracy for use. Applying a standard iterative method which does use the previous solution may not resolve this problem either, because of the ill-conditioned systems. A suggestion is to apply a multilevel process.

Let $N = \frac{n}{2}$, and assume that Tikhonov's method with the discrepancy principle obtains a solution after k Newton iterations. This coarse level procedure does not only derive a good approximate α, but it also derives a solution that may be interpolated into the fine grid to be used as an initial approximation for some iterative linear system solver. Now since the Newton scheme requires two systems to be solved and since a good α is available, a switch over to a generalized discrepancy principle which requires solving only one system should be performed. The created system can be solved iteratively, and as a new parameter is obtained using

$$\alpha_k = \mu\alpha_{k-1}, \qquad 0 < \mu < 1,$$

the solution of this system can be used as an initial approximation for the next system. Moreover, because the parameter changes only slightly, this initial approximation is accurate enough to prevent slow convergence of the linear system solver.

Concerning the interpolation operator, the expensive operator used in the Landweber scheme need not be used. The systems created here are non-singular, and so any suitable polynomial interpolation can be used; i.e., interpolation errors belonging in $N(K_{hh})$ pose no problem here. One fact that should be exploited is that the regularized solution should be smooth. Hence, higher-order polynomial interpolation should be used.

Choosing a smoother requires more investigation. Recall that multigrid applied to differential equations has the effect of damping out certain error components at certain levels. The damping effect is controlled by the smoother. But this effect may not occur when multigrid is applied to a system derived from an integral equation. For example, Gauss-Seidel can be applied to

$$\left[h^2 {K_{hh}}^t K_{hh} + \alpha I\right] \underline{f}_\alpha^\delta = {K_{hh}}^t \underline{g}_\delta \,,$$

but what does this smoothing do to the error components? Selective components need not be damped out rapidly. To see this, let

$$B = \left[h^2 {K_{hh}}^t K_{hh} + \alpha I\right] ,$$

and

$$B = D + L + L^t \,,$$

where D and L are respectively the diagonal and strictly lower triangular parts of B. After k iterations, the error is

$$\begin{aligned} \underline{e}^k &= \left[-(D+L)^{-1} L^t\right]^k \underline{e}^0 \\ &= \left[-(D+L)^{-1}(B - L - D)\right]^k \underline{e}^0 \\ &= \left[I - (D+L)^{-1} B\right]^k \underline{e}^0 . \end{aligned}$$

Now consider the effect of the relaxation on an eigenvalue–eigenvector pair of B, which we denote by $\left\{\hat{\lambda}_i, \underline{\hat{v}}_i\right\}$. Also let $\left\{\tilde{\lambda}_i, \underline{\tilde{v}}_i\right\}$ denote the eigenvalue–eigenvector pair of $[D+L]^{-1}$. Then if

$$\underline{\hat{v}}_i = \sum_{j=1}^{n} c_j \underline{\tilde{v}}_j \,,$$

$$\left[I - (D+L)^{-1} B\right]^k \underline{\hat{v}}_i = \sum_{j=1}^{n} c_j \underline{\tilde{v}}_j - k\hat{\lambda}_i \sum_{j=1}^{n} c_i \tilde{\lambda}_j \underline{\tilde{v}}_j + \frac{k(k-1)}{2} \hat{\lambda}_i \sum_{j=1}^{n} c_j \tilde{\lambda}_j \left[(D+L)^{-1} B\right] \underline{\tilde{v}}_j$$

$$+\cdots+(-1)^k\hat{\lambda}_i\sum_{j=1}^{n}c_j\tilde{\lambda}_j\left[(D+L)^{-1}B\right]^{(k-1)}\tilde{v}_j\,.$$

Nothing about the damping effect of the $\underline{\hat{v}}_i$ frequency can be determined from the last expression except that the frequencies can mix. Hence, in general, the damping effect of the Gauss–Seidel relaxation is unknown. This is rather unfortunate because the spectral properties of the systems derived from the regularized first kind equations are roughly known.

One smoother that is suitable for the regularized equations is the (preconditioned) conjugate gradient iteration. The eigenvalue–eigenvector systems of interest are those of the original systems. But the problem now is that there is a cluster of eigenvalues near α, and since the polynomial curve fit of the conjugate gradient method must equal one at zero, the convergence may be slow (Jennings (1977)). Preconditioning may relieve some of this problem. But then this problem may not be too serious for the high frequency components do not contribute too much to the solution anyways. Elaborating, if the residual vector $\underline{r}$ is expressed as

$$\underline{r}^{(0)}=\sum_{j=1}^{n}s_i^{(0)}\underline{\hat{v}}_i\,,$$

then after j conjugate gradient iterations,

$$s_i^{(j)}=p^j(\lambda_i)s_i^{(0)}\,,$$

where $p^j(\lambda)$ is the polynomial fit of degree j. Assuming that the interpolation introduces only a small amount of high frequency error, the $s_i^{(0)}$'s corresponding to high frequencies would be small already. Hence, although the $s_i^{(j)}$'s for large i's may not decrease too dramatically, they are small enough to allow only a small contribution of high frequency components to enter into $\underline{r}^{(j)}$.

Another way to see a possible rapid convergence of the conjugate gradient iteration is to view the large eigenvalues of B as outliers. This view is reasonable for regularized first kind equations since the eigenvalues usually converge rapidly to zero. Now using the analysis of Jennings, a smaller estimate of the number of conjugate gradient iterations required can be obtained.

Two other things concerning the conjugate gradient must be emphasized. First, the

above problem of the smoother occurs only in the first system created on the fine grid. Subsequent systems should not pose much problem because good initial approximations are available. Second, the conjugate gradient iteration must have a stopping condition that relates to the regularization procedure. This is given in the following theorem:

Theorem 4.12 *If the magnitude of the residual of the iterative solver is $O[(\delta+\eta)^2]$, where η is the quadrature error in approximating K with a semi-discrete operator K_η, then*

$$\hat{f}^{(\eta,\delta)}_{\alpha(\eta,\delta)} \longrightarrow K^\dagger g \qquad \text{as } \delta,\ \eta \to 0$$

provided that one of the generalized principles (1.)-(3.) is used. Here, $\hat{f}^{(\eta,\delta)}_{\alpha(\eta,\delta)}$ is the solution obtained by the iterative solver.

Proof: By theorem 4.10, using any one of the three rules, we have $\frac{(\delta+\eta)^2}{\alpha(\eta,\delta)} \to 0$ as $\delta, \eta \to 0$ and

$$\left\| f^{(\eta,\delta)}_{\alpha(\eta,\delta)} - K^\dagger g \right\| \longrightarrow 0 \qquad \text{as } \delta,\ \eta \to 0\,.$$

Thus,

$$\begin{aligned}
\left\| \hat{f}^{(\eta,\delta)}_{\alpha(\eta,\delta)} - K^\dagger g \right\| &\leq \left\| \hat{f}^{(\eta,\delta)}_{\alpha(\eta,\delta)} - f^{(\eta,\delta)}_{\alpha(\eta,\delta)} \right\| + \left\| f^{(\eta,\delta)}_{\alpha(\eta,\delta)} - K^\dagger g \right\| \\
&= \left\| [K_\eta{}^* K_\eta + \alpha(\eta,\delta) I]^{-1} \underline{r} \right\| + \left\| f^{(\eta,\delta)}_{\alpha(\eta,\delta)} - K^\dagger g \right\| \\
&\leq \frac{\|\underline{r}\|}{\alpha(\eta,\delta)} + \left\| f^{(\eta,\delta)}_{\alpha(\eta,\delta)} - K^\dagger g \right\| \longrightarrow 0 \qquad \text{as } \delta,\ \eta \to 0\,.
\end{aligned}$$

The result of the theorem is a bit discouraging since $(\delta+\eta)^2$ may be rather small. In practice, what can be done is to stop the iterative solver earlier, but to progress to this value as $\alpha(\eta,\delta)$ is processed. For example, the iteration may be stopped at $(\delta+\eta)^{\frac{3}{2}}$ for the first system, at $\mu(\delta+\eta)^{\frac{3}{2}}$ for the next system, etc..

Summarizing, the two-grid method obtains good estimates of α and the regularized solution on the coarse grid using the discrepancy principle. Interpolating this solution to the fine grid and using this α, by switching over to a generalized discrepancy principle which requires solving only one system, an iterative system solver can be used on the fine grid.

Now this procedure can be applied to more than two grid levels. Using p levels, the coarsest level is solved using the discrepancy principle with Cholesky decomposition, and

then the higher levels are solved using a generalized discrepancy principle with an iterative system solver. This may reduce the operation count substantially. For example, for $p = 3$ with grid sizes $h, 2h, 4h$, the breakdown is

$\frac{n^3}{2}$ – formation of $K_{hh}{}^t K_{hh}$,
$j\frac{n^3}{384}$ – Cholesky factorization at the coarsest grid,
$O\left(\frac{n^2}{4}\right)$ – iterative solver for all the systems on the middle grid, and
$O\left(n^2\right)$ – iterative solver for all the systems on the fine grid.

This multilevel method also can be applied to some nonlinear first kind integral equations after a change in the parameter selection procedure. To see why such a modification is required, we take an excursion to examine the stabilization of the least squares problem.

So far the exact operator case of (4.34) has been viewed as a perturbation of the least squares equation. A better way to justify the use of (4.34) is to view it as the Euler equation of

$$M^{\alpha(\delta)}[z, g_\delta] := \rho^2\left(\alpha(\delta)\right) + \alpha(\delta)\Omega(z) \tag{4.40}$$

where Ω is a nonnegative stabilization functional defined on a dense subset F_1 of $D(K)$ and satisfying

1. the exact solution belongs in F_1, and
2. for arbitrary positive M,

$$F_M := \{z : \Omega(z) \leq M\}$$

is compact in F_1.

The solution of the Euler equation of (4.40) minimizes $M^\alpha[z, g_\delta]$, and if the parameter is chosen such that $\rho\left(\alpha(\delta)\right) = \delta$, then with $F_\delta := \{z : \|Kz - g_\delta\| \leq \delta\}$, the solution also satisfies

$$\min_{F_1 \cap F_\delta} \Omega(z).^8$$

[8] Here, we are assuming that $M_0 \cap F_\delta = \emptyset$, where

$$M_0 := \left\{\hat{z} : \Omega(\hat{z}) = \inf_{z \in F_1} \Omega(z)\right\}$$

(Tikhonov and Arsenin (1977)).

For

$$[K^*K + \alpha(\delta)I]\, f^\delta_{\alpha(\delta)} = K^* g_\delta\,,$$

the corresponding stabilizer is $\Omega(z) = \|z\|^2$. A general form is the n'th-order stabilizer defined by

$$\Omega(z) = \int_a^b \sum_{i=0}^{n} a_i(s) \left[\frac{d^i z}{ds^i}\right]^2 ds\,,$$

where the $a_i(s)$'s are nonnegative functions. Applying this stabilizer to a first kind Fredholm equation produces a corresponding integro–differential equation with a set of boundary conditions (the Euler equation). For example, using a first-order stabilizer, the integro-differential equation is

$$\int_a^b \int_a^b k(v,s)k(v,t) f^\delta_{\alpha(\delta)}(t)\, dv\, dt - \alpha(\delta)\left\{\frac{d}{ds}\left[a_1(s)\frac{df^\delta_{\alpha(\delta)}}{ds}\right] - a_0(s) f^\delta_{\alpha(\delta)}\right\} = \int_a^b k(v,s) g_\delta(v)\, dv \qquad (4.41)$$

with boundary conditions $f^\delta_{\alpha(\delta)}(a) = \overline{f_1}$, $f^\delta_{\alpha(\delta)}(b) = \overline{f_2}$.

Returning to the previous discussion, a modification is required because the stabilizer corresponding to the usual regularization equation will be changed. With this change, the regularization equation converts into the form

$$[K^*K + \alpha(\delta)L]\, f^\delta_{\alpha_\delta} = K^* g_\delta\,. \qquad (4.42)$$

Clearly the generalized discrepancy rules (1.)–(3.) cannot be applied now.

One remedy to this drawback is the quasi-optimal selection method. Consider the case $L = I$. A "smooth" regularized solution is

$$\hat{f}^\delta_{\alpha(\delta)} := f^\delta_{\alpha(\delta)} - \alpha(\delta)\frac{df^\delta_{\alpha(\delta)}}{d\alpha}\,, \qquad (4.43)$$

which is obtained by solving

$$[K^*K + \alpha(\delta)I]\, \hat{f}^\delta_{\alpha_\delta} = K^* g_\delta + \alpha(\delta) f^\delta_{\alpha(\delta)}\,.$$

Naturally a desired property of a good parameter is that

$$\hat{f}^\delta_{\alpha(\delta)} \approx f^\delta_{\alpha(\delta)}$$

i.e., the solution is already smooth. Thus, choosing the parameter that minimizes

$$\left\| \alpha(\delta) \frac{df^{\delta}_{\alpha(\delta)}}{d\alpha} \right\|$$

provides us a solution with this property. In practice, if the parameter is processed according to the formula

$$\alpha_k = \mu \alpha_{k-1}\,, \qquad 0 < \mu < 1\,,$$

then since

$$\begin{aligned} \alpha(\delta) \frac{df^{\delta}_{\alpha(\delta)}}{d\alpha} &= \alpha(\delta) \lim_{\mu \to 1-0} \frac{f^{\delta}_{\alpha(\delta)} - f^{\delta}_{\mu\alpha(\delta)}}{\alpha(\delta)(1-\mu)} \\ &= \lim_{\mu \to 1-0} \frac{f^{\delta}_{\alpha(\delta)} - f^{\delta}_{\mu\alpha(\delta)}}{1-\mu}, \end{aligned}$$

a good parameter is one that minimizes $\left\| f^{\delta}_{\alpha_n(\delta)} - f^{\delta}_{\alpha_{n-1}(\delta)} \right\|$. Such an α is called the quasi-optimal parameter.

Now the good news is that the transition of the quasi-optimal method for $L \neq I$ poses no problem. In relation to multigrid, for first kind equations with non-trivial stabilizers, the parameter selection method on the fine grids of the previous multigrid procedure can be replaced with this selection method. Using the discrepancy principle on the coarsest grid to obtain a good initial parameter, the quasi-optimal procedure may not be too burdensome.

Turning to nonlinear Fredholm equations, consider

$$K[s, f(t)] := \int_a^b k(s,t,f(t))\,dt = g(s)\,, \qquad c \leq s \leq d. \tag{4.44}$$

We will assume that

1. $K[s, f_1(t)] \neq K[s, f_2(t)]$ if $f_1(t) \neq f_2(t)$,
2. $K[s, f(t)]$ is a continuous operator from $[c,d] \otimes C[a,b]$ into $\mathcal{L}_2[c,d]$,
3. $k_f(s,t,f(t))$ and $k_{ff}(s,t,f(t))$ are continuous in f in a neighborhood of the exact solution, and
4. $\int_a^b k_f(s,t,f(t))w(t)\,dt = 0$, where $f(t)$ is the exact solution, has in the Sobolev space W_2^1 only the trivial solution

(Tikhonov (1964)). Linearizing (4.44) about some function f_0^δ we have by Newton's method

$$K_f[s, f_0^\delta] f_1^\delta = K_f[s, f_0^\delta] f_0^\delta + g_\delta - K[s, f_0^\delta], \tag{4.45}$$

where K_f is the Fréchet derivative of K. Then with the functional

$$M^\alpha \left[f_1^\delta, K_f[s, f_0^\delta] f_0^\delta + g_\delta - K[s, f_0^\delta] \right] = \int_c^d \left\{ K_f[s, f_0^\delta] f_{1,\alpha(\delta)}^\delta - K_f[s, f_0^\delta] f_0^\delta - g_\delta + K[s, f_0^\delta] \right\}^2 ds + \alpha(\delta) \int_a^b \left(\frac{df_{1,\alpha(\delta)}^\delta}{ds} \right)^2 dt,$$

we obtain the Euler equation

$$\int_c^d \int_a^b k_f \left(v, s, f_0^\delta(s) \right) k_f \left(v, t, f_0^\delta(t) \right) f_{1,\alpha(\delta)}^\delta(t) \, dv \, dt - \alpha(\delta) \frac{d^2 f_{1,\alpha(\delta)}^\delta}{ds^2} = \int_c^d k_f \left(v, s, f_0^\delta(s) \right) h_\delta(v) \, dv, \tag{4.46}$$

where

$$h_\delta(s) := g_\delta(s) + \int_a^b k_f \left(s, t, f_0^\delta(t) \right) f_0^\delta(t) \, dt - \int_a^b k \left(s, t, f_0^\delta(t) \right) dt.$$

Equation (4.46) is linear, and hence can be solved easily. However, the cost of obtaining an accurate approximation is prohibitively large since for each α, several Newton iterations are performed and each Newton iteration requires a matrix formation and a matrix–matrix product. Some savings are made if the matrix is kept fixed once a good approximation is available, but the cost is still high. In contrast, a multilevel procedure will reduce the cost noticeably. This reduction is not only achieved through the use of an iterative system solver, but it is mainly achieved by a reduction in the number of matrix formations and matrix products. Solving (4.46) on the coarse grid via the Cholesky–quasi-optimal parameter method, a good α and approximate solution are obtained. On this grid, matrix formations and matrix products are performed per Newton iteration. Interpolating into a finer grid, if the solution is accurate enough, only the initial Newton iteration requires matrix formations and matrix products. Subsequent Newton iterations on this grid will not involve matrix formations and products, only system solving using the conjugate gradient iteration.

Chapter 5

Numerical Results

Numerical experiments will be reported in this chapter. All the examples involve integral equations which have been discretized using the composite trapezoid rule. The multilevel methods of course are not restricted to this equally-spaced quadrature, since the major obstacle is to construct ${K_{hh}}^t K_{hh}$ and its coarse grid versions in $\frac{n^3}{2}$ operations, or if K_{hh} is decomposed as $A_{hh}\underline{w}_h$, where A_{hh} is the matrix of kernel evaluations at the nodes and $\underline{w}_h$ is the vector of quadrature weights, the task is to construct ${A_{hh}}^t A_{hh}$ and its coarse grid versions in $\frac{n^3}{2}$ operations. This obstacle does not even arise in the multilevel Landweber iteration since ${K_{hh}}^t K_{hh}$ need not be constructed. In the Tikhonov schemes, the task can be achieved by building the finer matrices using the previously formed coarser ones.

5.1 Multilevel Landweber Iteration

Several standard procedures shall be performed for all the examples of this scheme. First, the stepsize of the quadrature rule will be set to

$$h = \sqrt{\frac{12\delta}{b_3}},$$

where b_3 is some upper bound on $\|K_{tt}\|$. Second, the discrete equations

$$hK_{hh}\underline{f}_h^\delta = \underline{g}_{h,\delta}$$

will be "normalized" by dividing each through with the trace of $h^2 K_{hh}{}^t K_{hh}$. Third, the parameter α is chosen using the Morozov discrepancy principle on the coarsest grid (stepsize= 4h) with the restriction that $\alpha \geq 0.0005$ in order to prevent large magnification of rounding errors in the interpolation procedure. And fourth, the parameters b_1, b_2 in stopping criterion (2.) will be set to 1.1.

Example 1. Antenna design theory (Bakushinskii and Sizikov (1982))

$$\int_{-\pi}^{\pi} \cos(st) f(t)\, dt = 2\pi \left\{ Si\left[(1+s)\pi\right] + Si\left[(1-s)\pi\right] \right\} \qquad -\pi \leq s, t \leq \pi\,.$$

Here, $Si[s]$ is the sine integral

$$Si[s] = \int_0^s \frac{\sin(v)}{v}\, dv\,,$$

and the solution is

$$f(t) = 2\pi \frac{\sin(t)}{t}\,.$$

We transform this equation so that the limits of integration are zero and one:

$$\int_0^1 \cos\left[\pi^2 (2w-1)(2v-1)\right] f(2\pi v - \pi)\, dv = Si[2\pi w - \pi + 1] + Si[1 - 2\pi w + \pi]\,, \qquad 0 \leq w, v \leq 1\,.$$

After discretizing this latter equation and normalizing the result, the right-hand side is perturbed by $trace\left[h^2 K_{hh}{}^t K_{hh}\right] \delta$. Using the upper bound $\|K\| \leq 1$ and $\|K_{vv}\| \leq \pi^4$, we obtain the results given in tables 5.1 and 5.2.

δ/h	time (real)	iterations	$\|error\|_2$
0.001/0.0078125	3.44	18	0.0324145
0.0005/0.0078125	4.01	21	0.0313566
0.0001/0.0019531	1079.81	348	0.0103711

Table 5.1: Antenna design theory using a standard Landweber method.

For the multigrid method, ten Landweber iterations were performed before each restriction. It is clear that this method is more efficient once a sufficient number of Landweber iterations are required for the standard method.

δ/h	time (real)	alpha	$\|error\|_2$
0.001/0.0078125	5.34	0.02300	0.0256249
0.0005/0.0078125	5.33	0.02293	0.0254669
0.0001/0.0019531	201.21	0.00250	0.00739896

Table 5.2: Antenna design theory using a multigrid Landweber scheme.

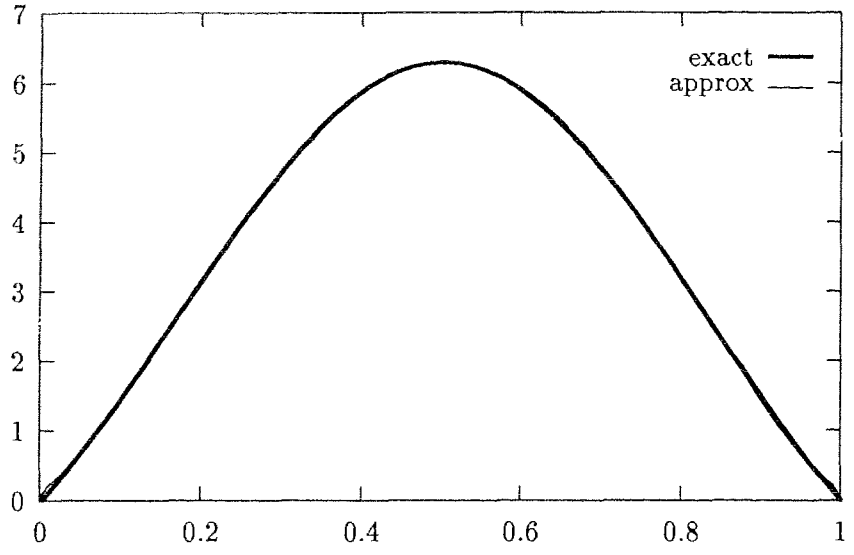

Figure 5.1: Antenna design theory for $\delta = 0.0005$.

Example 2. (Delves and Mohamed (1985))

$$\int_0^1 e^{s^{2.75}t} f(t)\,dt = \frac{e^{s^{2.75}+1} - 1}{s^{2.75}+1}, \qquad 0 \le s,t \le 1\,.$$

The exact solution is

$$f(t) = e^t\,.$$

Perturbing the forcing term as in example 1, we obtain the results summarized in tables 5.3 and 5.4.

Ten Landweber iterations also were performed before each restriction.

This example illustrates a problem associated with the Landweber iteration and the Tikhonov method with a zero'th-order stabilizer. From figure 5.3, we see that the approximation is good away from the endpoints, although it is not very good near the endpoints.

δ/h	time (real)	iterations	$\|error\|_2$
0.001/0.125	0.25	235	0.457333
0.0005/0.0625	1.09	325	0.344821
0.0001/0.03125	6.76	560	0.251526
0.00005/0.015625	31.84	687	0.180891

Table 5.3: Example 2 using a standard Landweber method.

δ/h	time (real)	alpha	$\|error\|_2$
0.001/0.125	0.39	0.630119	0.457326
0.0005/0.0625	1.28	0.282816	0.344790
0.0001/0.03125	5.85	0.127290	0.251518
0.00005/0.015625	19.00	0.0623390	0.180885

Table 5.4: Example 2 using a multigrid Landweber scheme.

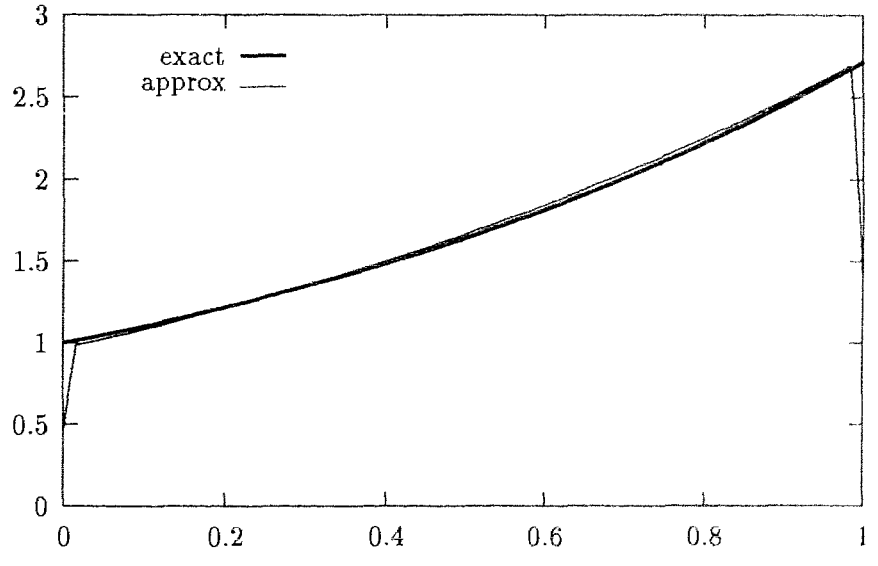

Figure 5.2: Example 2 for $\delta = 0.00005$.

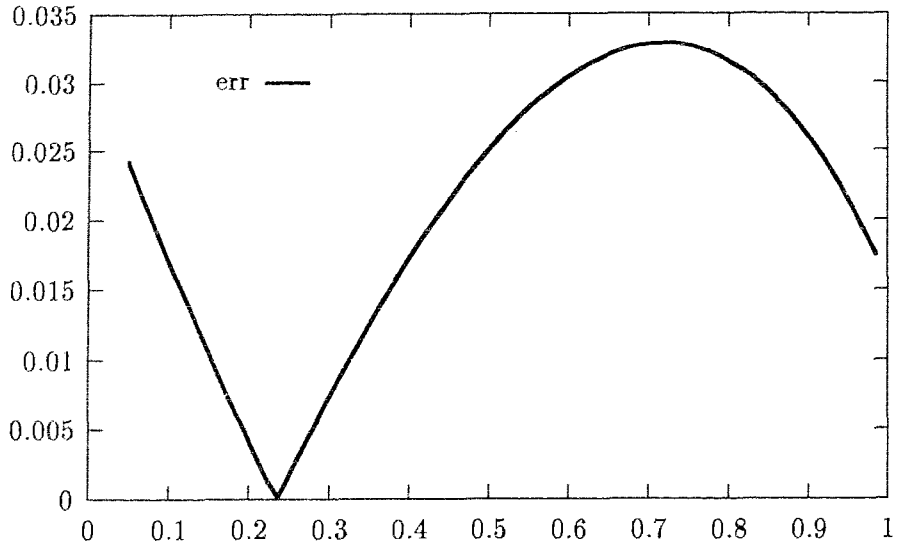

Figure 5.3: Example 2 error for $\delta = 0.0005$.

Providing boundary values may improve the accuracy near these points.

Example 3. Green's function for a vibrating string with fixed endpoints (Strand (1972,1974))

$$\int_0^1 k(s,t)f(t)\,dt = \frac{s}{30}\left(3 - 5s^2 + 3s^4 - s^5\right), \qquad 0 \le s,t \le 1$$

where

$$k(s,t) = \begin{cases} (1-s)t & 0 \le t \le s \le 1 \\ s(1-t) & 0 \le s \le t \le 1. \end{cases}$$

The forcing term was chosen such that the exact solution

$$f(t) = t - 2t^3 + t^4$$

is approximated well by the first singular function of K.

Again we perturb the forcing term by $trace[h^2 K_{hh}{}^t K_{hh}]\delta$, and relax ten times before restricting the defect onto grid $2h$. Results are summarized in tables 5.5 and 5.6.

The results should not be too surprising. Because the solution is composed mainly of a low harmonic, very few Landweber iterations are required to obtain a good approximation.

δ/h	time (real)	iterations	$\|error\|_2$
0.001/0.015625	0.12	2	0.000915895
0.0005/0.0078125	0.66	3	0.000906966
0.0001/0.00390625	3.37	4	0.000864410

Table 5.5: Green's function using a standard Landweber method.

δ/h	time (real)	alpha	$\|error\|_2$
0.001/0.015625	0.56	0.0561587	0.000670274
0.0005/0.0078125	3.17	0.0200675	0.000658018
0.0001/0.00390625	22.55	0.00740150	0.000210933

Table 5.6: Green's function using a multigrid Landweber scheme.

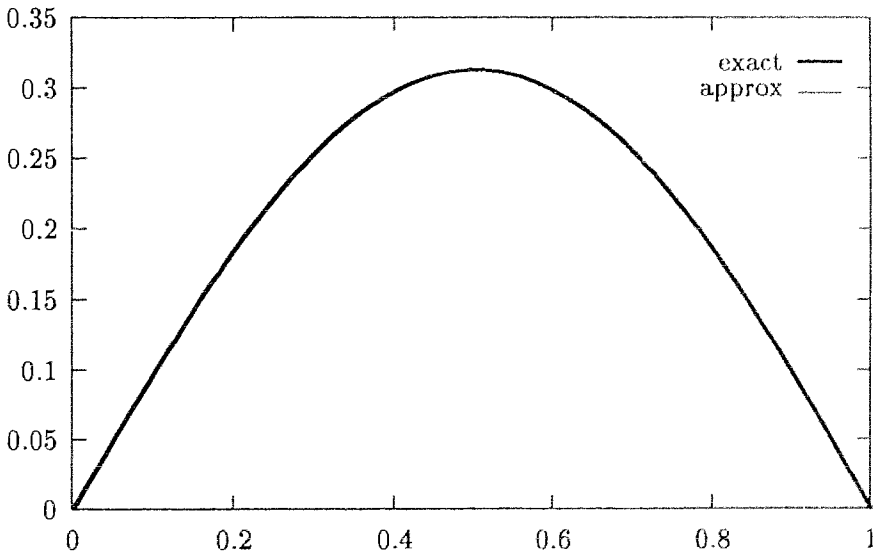

Figure 5.4: Green's function for $\delta = 0.0005$.

Hence, the multilevel scheme should be less efficient.

Example 4.

$$\frac{1}{\pi}\int_{-1}^{1}\frac{f(t)}{(s-t)^2+1}\,dt = -\frac{1}{\pi}\left\{2+(s^2-s-1)\left[\arctan(s+1)-\arctan(s-1)\right]\right.$$
$$\left.+\left(\frac{1}{2}-s\right)\ln\left[\frac{(s+1)^2+1}{(s+1)^2-1}\right]\right\}, \qquad -1\le s,t\le 1$$

with the exact solution

$$f(t)=t(1-t)\,.$$

The usual perturbation of $trace[h^2K_{hh}{}^tK_{hh}]\delta$ and the usual ten relaxation cycle produce the next two tables.

δ/h	time (real)	iterations	$\|error\|_2$
0.005/0.03125	0.99	82	0.312120
0.001/0.015625	32.18	700	0.179916
0.0005/0.0015625	106.14	2310	0.162019

Table 5.7: Example 4 using a standard Landweber method.

δ/h	time (real)	alpha	$\|error\|_2$
0.005/0.03125	1.04	0.0486373	0.266286
0.001/0.015625	4.98	0.0105314	0.177603
0.0005/0.0015625	14.49	0.0102397	0.161406
0.0001/0.0015625	229.20	0.000786000	0.0788235

Table 5.8: Example 4 using a multigrid Landweber scheme.

5.2 Tikhonov Method with Zero'th-order Stabilizer

The examples for this scheme are of more practical interest than those examined in the Landweber method. One of them is a first kind Volterra integral equation which will be converted into a Fredholm equation. Another involves discrete data which will be interpolated using cubic splines. And for some of them, it is known that higher-order stabilizers

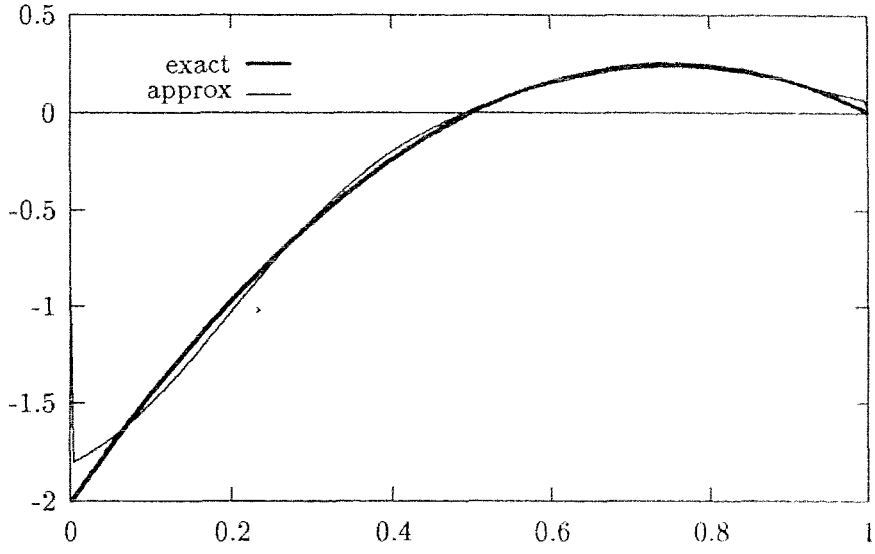

Figure 5.5: Example 4 for $\delta = 0.0001$.

are more appropriate. We examine these examples again using first-order stabilizers with known boundary values in the next section.

Common procedures in all of these examples include an initial parameter choice of one, a preconditioning using the diagonal of $[h^2 K_{hh}{}^t K_{hh} + \alpha I]$, a piecewise linear interpolation operator between two consecutive grids, and a generalized discrepancy principle with the function

$$h_2(\alpha) = \rho_{(\eta,\delta)}\left(\frac{1}{\alpha}\right) - \left(\delta + \eta \left\| f^{(\eta,\delta)}_{\frac{1}{\alpha(\eta,\delta)}} \right\|\right) ,$$

where

$$\eta = \frac{h\sqrt{(b-a)(d-c)}}{3}\left(\|K\| + \frac{h}{4}\max_{s,t}\left|\frac{\partial^2 K}{\partial t^2}\right|\right)$$

(Goncharskii, Leonov, and Yagola (1973)), and where a, b are the limits of integration, and c, d are the bounds on s. Only approximations of $\|K\|$ and $\max_{s,t}\left|\frac{\partial^2 K}{\partial t^2}\right|$ are used.

Example 1. Harmonic continuation (Franklin (1974), Varah (1973))

$$\frac{1}{2\pi}\int_0^{2\pi} \frac{1-r^2}{1-2r\cos(s-t)+r^2} f(t)\,dt = g(s), \qquad 0 \le s, t \le 2\pi,\ r < 1\,.$$

The equation is the Poisson integral formula in the unit circle for fixed radius r. This formula gives the solution of the Dirichlet problem

$$\begin{aligned} &\tfrac{1}{r}\tfrac{\partial}{\partial r}(ru_r) + \tfrac{1}{r^2}u_{ss} = 0 && r < 1,\ 0 < s < 2\pi \\ &u(1,s) = f(s) && 0 < s < 2\pi \\ &u(r,0) = u(r,2\pi) && r < 1 \\ &u_s(r,0) = u_s(r,2\pi) && r < 1\,. \end{aligned}$$

But for our problem, $g(s) = u(r,s)$ is given approximately and $f(s) = u(1,s)$ must be found.

Another way to relate $f(s)$ to $g(s)$ is, if $f(s)$ has the Fourier series

$$f(s) = f_0 + \sum_{k=1}^{\infty} \left(f_k \cos(ks) + f_k' \sin(ks)\right)\,,$$

then $g(s)$ has the Fourier series

$$g(s) = f_0 + \sum_{k=1}^{\infty} r^k \left(f_k \cos(ks) + f_k' \sin(ks)\right)\,.$$

Clearly, small changes in high-order Fourier coefficients of $g(s)$ may cause large changes in $f(s)$.

In the numerical experiment, $r = \frac{1}{2}$,

$$g(s) = -1 + \frac{1}{2}\cos(s) + \frac{1}{8}\sin(3s)\,,$$

and $[a,b]$ and $[c,d]$ are transformed into $[0,1]$. The exact solution is

$$f(t) = -1 + \cos(t) + \sin(3t)\,.$$

Perturbing g to

$$g_\delta(s) = g(s)[1 + \theta\delta],$$

where θ is a random number chosen from a uniform distribution on $[-1,1]$, we obtained tables 5.9 and 5.10.

The cg iterations column of table 5.10 gives the number of preconditioned conjugate gradient iterations required for each updated α on grids 1 and 2. For example, for $\delta = 0.001$, three and seven conjugate gradient iterations were required on the systems of grid 2, and six and ten iterations on systems formed on grid 1.

δ/h	time (real)	alpha	$\|error\|_2$
0.001/0.03125	1.02	0.01505	0.344369
0.0002/0.015625	6.84	0.005145	0.174514
0.0001/0.0078125	50.51	0.002187	0.0868148
0.00002/0.00390625	411.54	0.001008	0.0428558

Table 5.9: Harmonic continuation using a standard Tikhonov method with zero'th-order stabilizer.

δ/tol	time (real)	alpha	cg iterations	$\|error\|_2$
$0.001/10^{-5}$	0.57	0.001995	3/7 6/10	0.0793932
$0.0002/10^{-6}$	2.45	0.0005496	4/8 7/11	0.0239888
$0.0001/10^{-6}$	13.77	0.0001465	4/8 7/10	0.00732991
$0.00002/10^{-8}$	93.88	0.0005098	6/10 7/10	0.00264081

Table 5.10: Harmonic continuation using a multigrid scheme with $\mu = 0.1$.

δ/tol	time (real)	alpha	cg iterations	$\|error\|_2$
$0.001/10^{-5}$	0.60	0.001247	3/3/4 4/4/4	0.311197
$0.0002/10^{-6}$	2.43	0.003435	4/4/4 4/6/4	0.126922
$0.0001/10^{-6}$	12.96	0.001816	4/4/4 4/4	0.0742827
$0.00002/10^{-8}$	93.10	0.0006372	6/7/4 7/7	0.0277327

Table 5.11: Harmonic continuation using a multigrid scheme with $\mu = 0.5$.

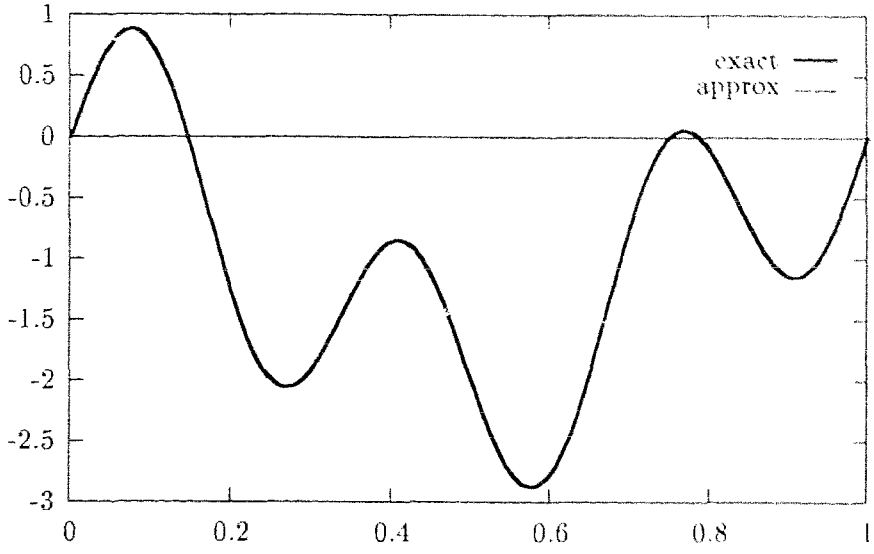

Figure 5.6: Harmonic continuation for $\mu = .1$, $\delta = 0.00002$.

Table 5.10 was obtained using $\mu = 0.1$. Different values of μ give different results as illustrated in table 5.11. Determining an optimal μ is a problem in itself.

Example 2. (Phillips (1962), Kryanev (1974))

$$\int_{-3}^{3} k(t-s)f(t)\,dt = g(s)\,, \qquad -6 \le s \le 6$$

where

$$k(z) = \begin{cases} 1+\cos(\frac{\pi z}{3}) & |z| \le 3 \\ 0 & |z| > 3 \end{cases}$$

and

$$g(s) = \begin{cases} (6-s)\left[1+\frac{1}{2}\cos(\frac{\pi s}{3})\right] + \frac{9}{2\pi}\sin(\frac{\pi s}{3}) & 0 \le s \le 6 \\ (6+s)\left[1+\frac{1}{2}\cos(\frac{\pi s}{3})\right] - \frac{9}{2\pi}\sin(\frac{\pi s}{3}) & -6 \le s \le 0\,. \end{cases}$$

The exact solution is

$$f(t) = \begin{cases} 1+\cos(\frac{\pi t}{3}) & |t| \le 3 \\ 0 & |t| \ge 3\,. \end{cases}$$

We also transform the intervals $[-3,3]$ and $[-6,6]$ into $[0,1]$, and perturb the data by $g(s)\theta\delta$.

δ/h	time (real)	alpha	$\|error\|_2$
0.001/0.03125	0.95	0.03832	0.181006
0.0002/0.015625	5.95	0.01785	0.104504
0.0001/0.0078125	44.43	0.008562	0.0627038
0.00002/0.00390625	364.13	0.004087	0.0404601

Table 5.12: "Phillip's" equation using a standard Tikhonov method with zero'th-order stabilizer.

δ/tol	time (real)	alpha	cg iterations	$\|error\|_2$
$0.001/10^{-5}$	0.48	0.02380	2/3/3 4/3	0.128577
$0.0002/10^{-6}$	2.14	0.01041	3/4/4 4/4	0.0709945
$0.0001/10^{-6}$	12.88	0.004729	4/3/4 5/5	0.0438212
$0.00002/10^{-8}$	89.95	0.002203	5/5/5 5/6	0.0299022

Table 5.13: "Phillip's" equation using a multigrid scheme with $\mu = 0.5$.

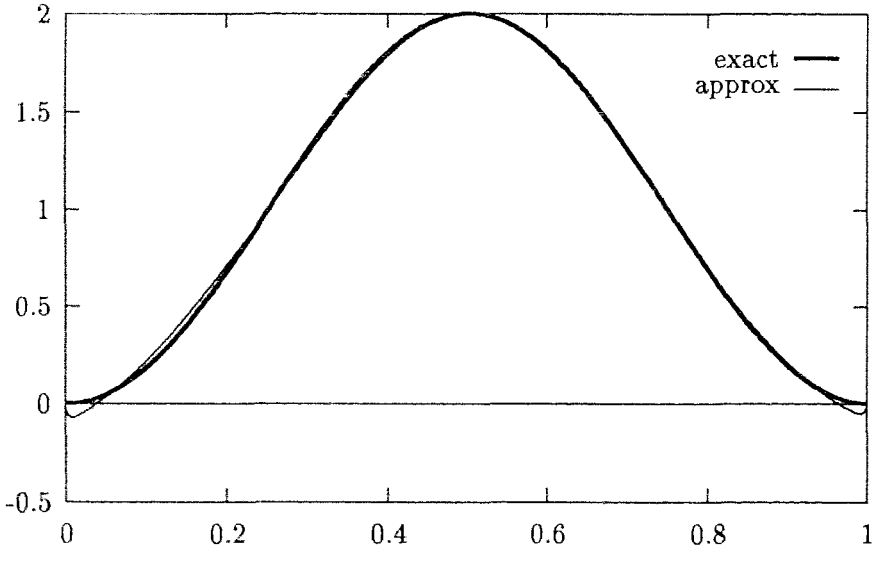

Figure 5.7: "Phillip's" equation for $\mu = 0.5$, $\delta = 0.00002$.

Example 3. Numerical differentiation (Tikhonov and Arsenin (1977), Marti (1980), Trummer (1984))

$$\int_0^s \frac{1}{(n-1)!}(s-t)^{n-1} f(t)\,dt = g(s)\,.$$

This equation first is converted into the Fredholm equation

$$\int_0^b k(s,t) f(t)\,dt = g(s)\,,$$

$$k(s,t) = \begin{cases} \frac{1}{(n-1)!}(s-t)^{n-1} & b \geq s > t \geq 0 \\ 0 & 0 \leq s \leq t \leq b\,. \end{cases}$$

Now choosing $n = 3$, $b = 4$, provided that $g \in W^3[0,4]$ and $g(0) = g'(0) = g''(0) = 0$, the solution of this equation is $g'''(t)$. In particular, with the choice

$$g(s) = \int_0^s e^{-(y-2)^4}\,dy - e^{-16}(s + 16s^2)\,,$$

and with the perturbation of $g(s)\theta\delta$, we have tables 5.14 and 5.15.

δ/h	time (real)	alpha	$\|error\|_2$
0.01/0.125	0.82	0.2183	1.858310
0.005/0.0625	4.95	0.09010	1.853227
0.001/0.03125	40.89	0.02372	1.810004
0.0002/0.015625	338.35	0.005029	1.711513

Table 5.14: Numerical differentiation using a standard Tikhonov method with zero'th-order stabilizer.

The unusually large number of conjugate gradient iterations is due to the inaccuracy of the approximations. In turn, the approximations are inaccurate due to the bad choice of stabilizer– the zero'th-order stabilizer simply minimizes the mean value of the approximations.

δ/tol	time (real)	alpha	cg iterations	$\|error\|_2$
$0.01/10^{-3}$	0.73	0.01209	7/14 18/28	1.771946
$0.005/10^{-4}$	3.47	0.004780	8/19 19/45	1.701079
$0.001/10^{-5}$	21.56	0.001996	12/32 22/76	1.624848
$0.0002/10^{-6}$	139.19	0.0007767	14/48 29/106	1.522646

Table 5.15: Numerical differentiation using a multigrid scheme with $\mu = 0.1$.

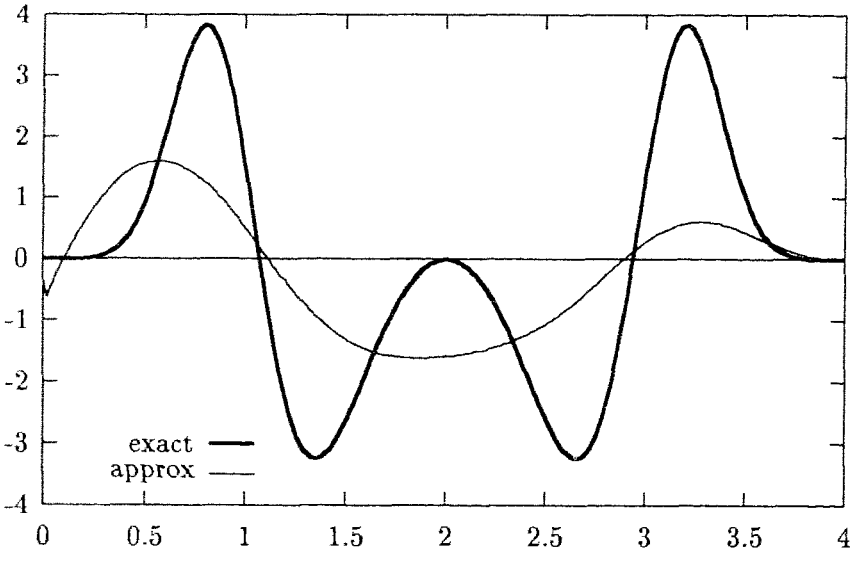

Figure 5.8: Numerical differentiation for $\mu = 0.1$, $\delta = 0.0002$.

Example 4. Spectral composition of radiation (Tikhonov and Arsenin (1977))

$$\int_1^{11} \left(1 - \frac{t}{s}\right) h(s-t) f(t)\, dt = g(s)\,, \qquad 1 \le s, t \le 11\,,$$

where $h(s-t)$ is a unit step function. Here, the unknown function characterizes the distribution of the particles, and the kernel, which may be obtained by passing a delta pulse through an experimental measuring device, describes such a device.

The exact solution was given at 41 points and then interpolated using cubic splines. Next, the experimental spectrum was obtained by numerically integrating this interpolated solution using a composite Simpson rule. After perturbing this spectrum by $g(s)\theta\delta$, the following results were obtained:

δ/h	time (real)	alpha	$\|error\|_2$
0.01/0.078125	41.90	0.05687	0.186767
0.001/0.0390625	339.92	0.02174	0.183865
0.0005 /0.01953125	2907.32	0.009660	0.181290

Table 5.16: Spectral composition of radiation using a standard Tikhonov method with zero'th-order stabilizer.

δ/tol	time (real)	alpha	cg iterations	$\|error\|_2$
$0.01/10^{-4}$	13.70	0.004141	3/7 5/15	0.0997999
$0.001/10^{-5}$	98.89	0.001233	3/12 9/29	0.0724816
$0.0005/10^{-5}$	729.57	0.0004835	5/15 10/25	0.0567522

Table 5.17: Spectral composition of radiation using a multigrid scheme with $\mu = 0.1$.

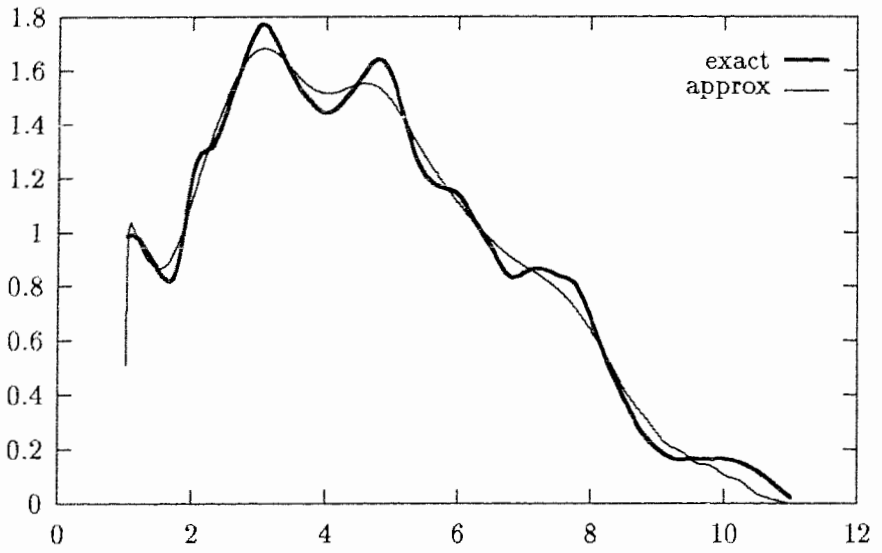

Figure 5.9: Spectral composition of radiation for $\mu = 0.1$, $\delta = 0.0005$.

Example 5. Backward heat equation (Tikhonov and Arsenin (1977), Varah (1973), King and Neubauer (1988))

$$\int_0^1 k(s,t)f(t)\,dt = g(s) = h\left(s-\frac{1}{4}\right) - h\left(s-\frac{3}{4}\right) \qquad 0 \le s,t \le 1$$

where

$$k(s,t) = \begin{cases} 0 & s \le t \\ \frac{e^{-\frac{1}{4(s-t)}}}{\sqrt{\pi(s-t)}} & s > t\,, \end{cases}$$

and

$$h(s) = \begin{cases} \frac{2}{\sqrt{\pi}}\sqrt{s}\,e^{-\frac{1}{4s}} - \frac{2}{\sqrt{\pi}}\int_{\frac{1}{2\sqrt{s}}}^{\infty} e^{-w^2}\,dw & s > 0 \\ 0 & s \le 0\,. \end{cases}$$

Solving this equation is equivalent to solving the ill-posed problem

$$\begin{cases} u_t = u_{xx} & x \in (0,\infty),\ t \in (0,1] \\ u(1,t) = g(t) & t \in (0,1] \\ |u(x,t)| < M & x \to \infty \\ u(x,0) = 0 & x \in (0,\infty) \end{cases}$$

for $-u_x(0,t)$.

The exact solution is

$$f(t) = \begin{cases} 1 & \frac{1}{4} \le t \le \frac{3}{4} \\ 0 & \text{elsewhere} \end{cases}$$

Using the perturbation $g(s)\delta\theta$ again, we have tables 5.18 and 5.19.

δ/h	time (real)	alpha	$\|error\|_2$
0.01/0.125	0.12	0.008677	0.360942
0.001/0.03125	1.21	0.0007986	0.214249
0.0001/0.0078125	57.27	0.00009595	0.166551
0.00001/0.00390625	461.74	0.00003047	0.146234

Table 5.18: Backward heat equation using a standard Tikhonov method with zero'th-order stabilizer.

δ/tol	time (real)	alpha	cg iterations	$\|error\|_2$
$0.01/10^{-3}$	0.11	0.005697	1 1/1/1/1/2	0.327268
$0.001/10^{-5}$	0.57	0.0006529	2/2/2 3/3	0.230515
$0.0001/10^{-6}$	13.04	0.00007873	3/3/3 4/4	0.170211
$0.00001/10^{-8}$	96.07	0.00001470	5/5/4 6/6/6	0.142153

Table 5.19: Backward heat equation using a multigrid scheme with $\mu = 0.5$.

5.3 Quasi-optimal Schemes

Examples 1, 3–5 of the previous section will be re-examined now with a first-order stabilizer using the quasi-optimal parameter choice.

For the standard method, the initial α will always be one, and the parameter μ will be 0.1. Thirteen iterates of the geometric progression for α, each requiring a Cholesky decomposition, will be performed. The solution $\underline{f}^{\delta}_{h,\alpha_i}$ that locally minimizes

$$\left\| \underline{f}^{\delta}_{h,\alpha_i} - \underline{f}^{\delta}_{h,\alpha_{i+1}} \right\|$$

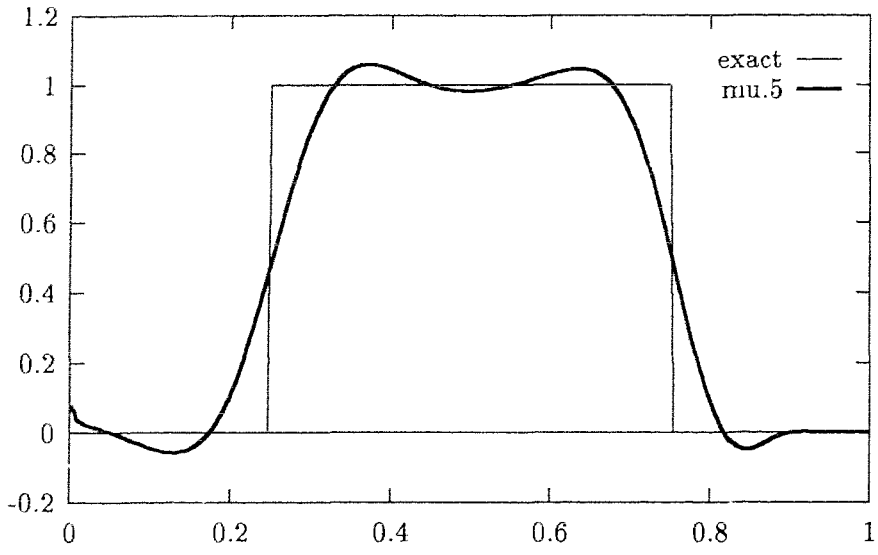

Figure 5.10: Backward heat equation for $\mu = 0.5$, $\delta = 0.00001$.

then is chosen to be the regularized solution. It must be emphasized that a local minimum is searched for rather than the extreme minimum because it is possible for two strongly "smoothed" solutions corresponding to a region of relatively large α's to differ minimally. The solution corresponding to the second minimum will always be chosen.

The parameter choosing procedure of the three-grid scheme is slightly better. Rather than using the same μ for all three grids, three different μ's with their values increasing as the grid is refined are used. We always will set $\mu^3 = 0.1$ on the coarsest grid, $\mu^2 = 0.25$ on the middle grid, and $\mu^1 = 0.75$ on the finest grid, unless otherwise specified. With $\mu^3 = 0.1$, ten iterates of the geometric refinement for α is performed, and then the regularized solution is chosen from the resulting set of solutions as in the standard procedure. Denoting the quasi-optimal parameter on the coarsest grid by α^3_{opt} , the initial α on the middle grid is taken to be $\frac{\alpha^3_{opt}}{\mu^3}$. Now on this refined grid, only five iterates of the geometric progression for α is performed, and the quasi-optimal solution is chosen from the resulting five approximations. Lastly, this process is repeated on the finest grid.

Turning to the linear system solver of the multilevel scheme, a preconditioned conjugate gradient method using the preconditioner $diag[h^2 {K_{hh}}^t K_{hh} + \alpha L_{hh} + \alpha I_{hh}]$ is applied on the

finer grids with stepsize h. The initial approximation of the first system on a finer grid is taken to be some interpolated version of the coarse grid quasi-optimal solution.

As for the effectiveness of the type of interpolation used, the piecewise linear and piecewise quadratic types appear to be the most effective. The interpolation scheme of the multilevel Landweber iteration was totally ineffective– the smaller α's of the first-order stabilizer permitted large rounding errors to creep into the interpolated solution, and hence, a large number of conjugate gradient iterations was required to solve the first system.

But no matter what interpolation was used, the number of conjugate gradient iterations was a bit larger than expected. This was true even for the "interpolation" which involved solving the initial system using a Cholesky decomposition. The reason for this problem is that the change to $[\alpha L_{hh} + \alpha I_{hh}]$ as α is refined is not minor. One way to correct this is to obtain a very good α on the coarse grid, which then would permit larger μ's to be used on the finer grids. But this approach was not fully investigated, one reason being that it may a bit unreasonable since the coarse grid parameter need not be an extremely good approximation of the fine grid parameter.

Another more basic problem with this multilevel approach is that it relies too heavily on the intuitive quasi-optimal parameter choice. And noting that $\frac{df_\alpha^\delta}{d\alpha}$ is approximated only numerically, we definitely should take some precautionary measures when applying this scheme.

Anyhow, all the examples were perturbed by $g(s)\delta\theta$. Some of the results are given in the tables below.

Example 1. Harmonic continuation

$$\frac{1}{2\pi}\int_0^{2\pi} \frac{0.75}{1-\cos(s-t)+0.25} f(t)\,dt = -1 + 0.5\cos(s) + 0.125\sin(3s), \qquad 0 \le s, t \le 2\pi.$$

Exact solution: $f(t) = -1 + \cos(t) + \sin(3t)$.

δ/h	time (real)	alpha	$\|error\|_2$
0.001/0.03125	1.07	10^{-8}	0.0122705
0.0002/0.015625	6.46	10^{-8}	0.00194001
0.0001/0.0078125	47.52	10^{-8}	0.00177101
0.00002/0.00390625	375.14	10^{-9}	0.000197044

Table 5.20: Harmonic continuation using a standard Tikhonov method with first-order stabilizer.

δ/tol	time (real)	alpha	cg iterations	$\|error\|_2$
$0.001/10^{-5}$	0.67	9.7656×10^{-12}	9/7/10/3/2 9/10/5/6/2	0.0655329
$0.0002/10^{-6}$	3.09	9.7656×10^{-12}	14/6/20/29/3 25/3/1/1/1	0.0182970
$0.0001/10^{-6}$	19.76	6.2500×10^{-9}	12/13/16/17/17 21/13/12/9/11	0.00105723
$0.00002/10^{-8}$	139.36	2.5000×10^{-9}	15/16/26/34/27 36/21/21/20/18	0.000180559

Table 5.21: Harmonic continuation using a multigrid scheme with piecewise linear interpolation.

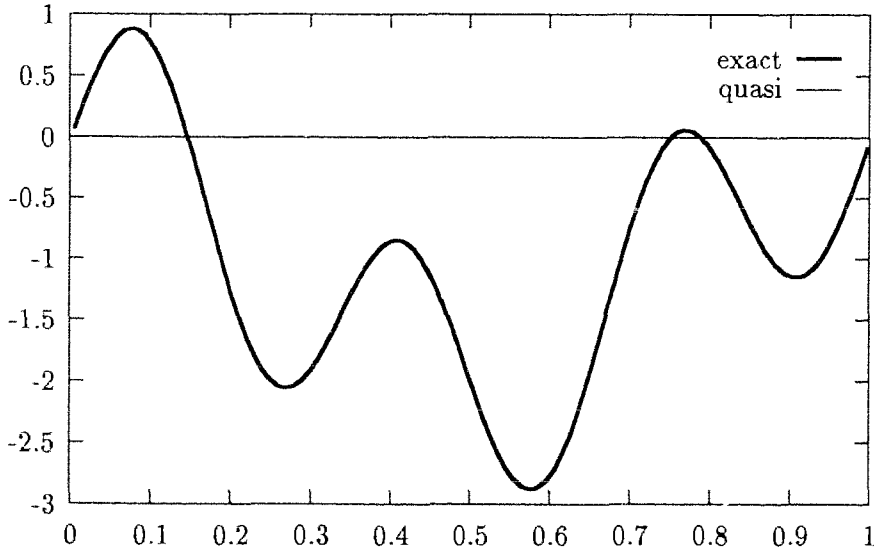

Figure 5.11: Harmonic continuation for $\delta = 0.00002$ using piecewise linear interpolation.

Example 2. Numerical differentiation

$$\int_0^4 k(s,t)f(t)\,dt = \int_0^s e^{-(y-2)^4}\,dy - e^{-16}(s+16s^2)\,,$$

$$k(s,t) = \begin{cases} \frac{1}{2!}(s-t)^2 & 4 \geq s > t \geq 0 \\ 0 & 0 \leq s \leq t \leq 4\,. \end{cases}$$

Exact solution: $f(t) = g'''(t)\,.$

δ/h	time (real)	alpha	$\|error\|_2$
0.01/0.125	0.98	10^{-3}	1.833796
0.005/0.0625	6.28	10^{-8}	0.659337
0.001/0.03125	46.66	10^{-9}	0.328415
0.0002/0.015625	372.38	10^{-10}	0.147929

Table 5.22: Numerical differentiation using a standard Tikhonov method with first-order stabilizer.

δ/tol	time (real)	alpha	cg iterations	$\|error\|_2$
$0.01/10^{-3}$	0.54	9.7656×10^{-12}	1/1/1/1 5/4/1/1	27.284139
$0.005/10^{-4}$	7.26	3.9063×10^{-5}	27/27/31/31 48/42/45/51	1.569780
$0.001/10^{-5}$	19.48	3.1250×10^{-9}	4/15/10/13 27/10/4/4	0.323278
$0.0002/10^{-6}$	184.24	1.5625×10^{-10}	5/8/33/28 115/20/40/10	0.148478

Table 5.23: Numerical differentiation using a multigrid scheme with "Cholesky" interpolation.

Example 3. Spectral composition of radiation

$$\int_1^{11} \left(1 - \frac{t}{s}\right) h(s-t)f(t)\,dt = g(s)\,, \qquad 1 \leq s,t \leq 11\,.$$

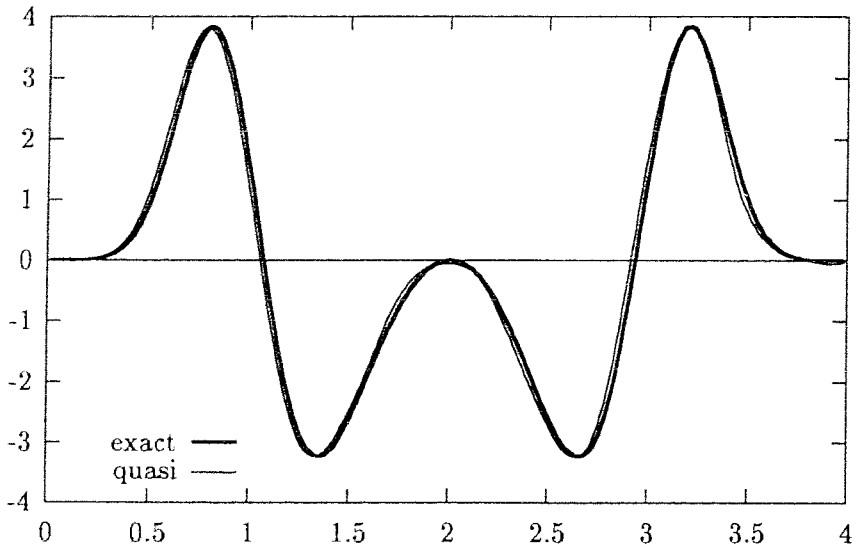

Figure 5.12: Numerical differentiation for $\delta = 0.00002$ using "Cholesky" interpolation.

δ/h	time (real)	alpha	$\|error\|_2$
0.01/0.078125	47.59	10^{-3}	0.100298
0.001/0.0390625	373.95	10^{-6}	0.0355995

Table 5.24: Spectral composition of radiation using a standard Tikhonov method with first-order stabilizer.

δ/tol	time (real)	alpha	cg iterations	$\|error\|_2$
$0.01/10^{-4}$	20.77	5×10^{-3}	15/11/7/6/6 20/20/19/14/11	0.100395
$0.001/10^{-6}$	160.99	3.125×10^{-6}	27/38/45/42/35 27/36/34/34/26	0.0315379

Table 5.25: Spectral composition of radiation using a multigrid scheme with piecewise quadratic interpolation.

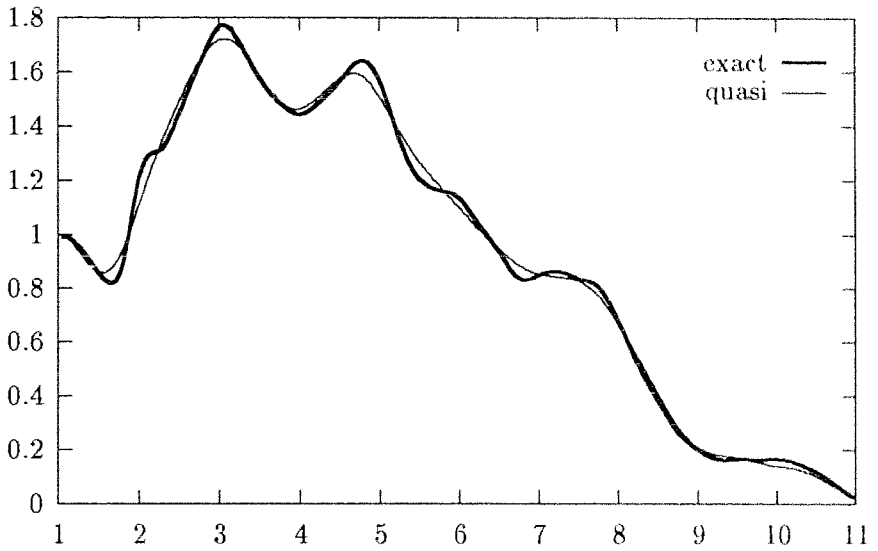

Figure 5.13: Spectral composition of radiation for $\delta = 0.001$ using piecewise quadratic interpolation.

Example 4. Backward heat equation

$$\int_0^1 k(s,t)f(t)\,dt = g(s) = h\left(s-\frac{1}{4}\right) - h\left(s-\frac{3}{4}\right) \qquad 0 \le s,t \le 1$$

where

$$k(s,t) = \begin{cases} 0 & s \le t \\ \frac{e^{-\frac{1}{4(s-t)}}}{\sqrt{\pi(s-t)}} & s > t\,, \end{cases}$$

and

$$h(s) = \begin{cases} \frac{2}{\sqrt{\pi}}\sqrt{s}\,e^{-\frac{1}{4s}} - \frac{2}{\sqrt{\pi}}\int_{\frac{1}{2\sqrt{s}}}^{\infty} e^{-w^2}\,dw & s > 0 \\ 0 & s \le 0\,. \end{cases}$$

Exact solution: $f(t) = \begin{cases} 1 & \frac{1}{4} \le t \le \frac{3}{4} \\ 0 & \text{elsewhere} \end{cases}$.

The parameters chosen for the multigrid method here are $\alpha_0 = 0.001$, $\mu^3 = 0.1$, $\mu^2 = 0.5$, $\mu^1 = 0.9$. These parameters were chosen so that the minimum of α is approximately 10^{-12}, the minimum value for the standard approach.

δ/h	time (real)	alpha	$\|error\|_2$
0.01/0.125	0.12	10^{-12}	0.350225
0.001/0.03125	1.04	10^{-10}	0.132364
0.0001/0.0078125	47.37	10^{-12}	0.0907106
0.00001/0.00390625	374.90	10^{-12}	0.08831142

Table 5.26: Backward heat equation using a standard Tikhonov method with first-order stabilizer.

δ/tol	time (real)	alpha	cg iterations	$\|error\|_2$
$0.01/10^{-3}$	0.11	8.2013×10^{-13}	1/1/1/1/1 3/1/1/1/1	0.275375
$0.001/10^{-6}$	0.57	8.2013×10^{-13}	9/6/1/2/1 6/1/3/3/2	0.230515
$0.0001/10^{-6}$	16.85	8.2013×10^{-13}	6/13/14/21/27 15/6/5/5/3	0.0965745
$0.00001/10^{-8}$	189.66	8.2013×10^{-13}	15/17/31/41/52 115/30/6/58/31	0.0798144

Table 5.27: Backward heat equation using a multigrid scheme with piecewise linear interpolation.

5.4 Nonlinear First Kind Equations

The last example involves nonlinear Fredholm integral equations. We will investigate the contact surface problem

$$\frac{1}{4\pi}\int_a^b \ln\left[\frac{(s-t)^2+h^2}{(s-t)^2+(f(t)-h)^2}\right] dt = g(s), \qquad c \le s \le d,$$

for fixed bedrock level h using the quasi-optimal parameter choice. As mentioned earlier, the Euler equation for the linearized form of the perturbed equation is

$$\int_c^d \int_a^b k_f\left(v,s,f^\delta_{i,\alpha_j}(s)\right) k_f\left(v,t,f^\delta_{i,\alpha_j}(t)\right) f^\delta_{i+1,\alpha_j}(t)\,dv\,dt - \alpha_j \frac{df^\delta_{i+1,\alpha_j}(s)}{ds} =$$
$$\int_c^d k_f\left(v,s,f^\delta_{i,\alpha_j}(s)\right)\left[g_\delta(v) + \int_a^b k_f\left(v,t,f^\delta_{i,\alpha_j}(t)\right) f^\delta_{i,\alpha_j}(t)\,dt - \int_a^b k\left(v,t,f^\delta_{i,\alpha_j}(t)\right) dt\right] dv,$$

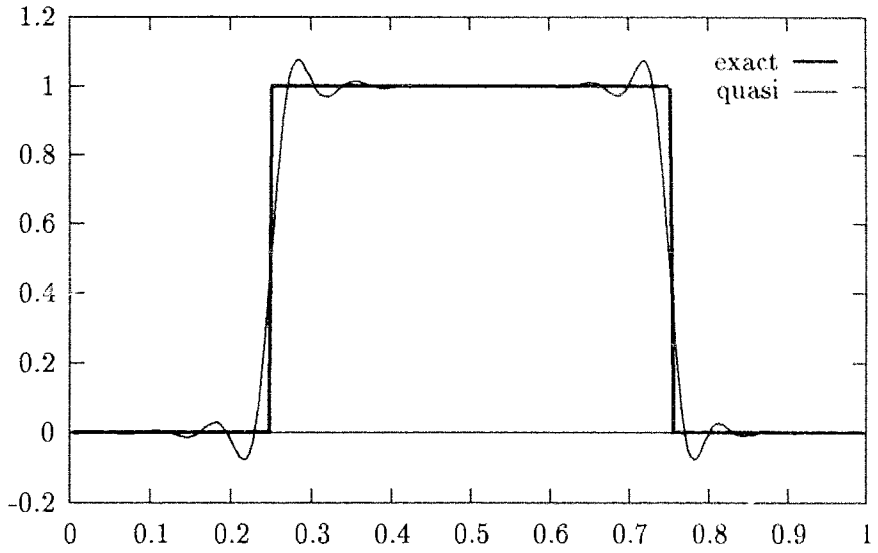

Figure 5.14: Backward heat equation for $\delta = 0.00001$ using piecewise linear interpolation.

where k and k_f respectively denote the nonlinear kernel and its partial derivative with respect to f. We will iterate this equation for a given initial f^{δ}_{0,α_j} until

$$\frac{\left\| f^{\delta}_{i,\alpha_j} - f^{\delta}_{i+1,\alpha_j} \right\|}{\left\| f^{\delta}_{i,\alpha_j} \right\|} < tol.$$

For the standard method, we again will perform thirteen iterates of the geometric progression for α using a particular μ. The initial approximation for the initial α will be $\underline{f}^{\delta}_{0,\alpha_0} = 0$, and the initial approximation for α_j will be the $\underline{f}^{\delta}_{i+1,\alpha_{j-1}}$ that satisfies the above tolerance condition. Now for each α_j, a sequence of $K_{f,hh}$'s and $K_{f,hh}{}^t K_{f,hh}$'s will be constructed, and a sequence of linear systems will be solved. After obtaining the thirteen solutions, the quasi-optimal solution will then be chosen as in the previous section but with the additional condition that the difference between the norm of this solution and a sample mean of the norms of the thirteen solutions be less than some sample deviation.

The three-grid method again will involve three different μ's: μ^3, μ^2, and μ^1. The coarse grid processing will be exactly the same as the standard procedure but with only ten iterates of the geometric progression. Its quasi-optimal approximation will be interpolated into the next grid using piecewise linear interpolation, and its quasi-optimal parameter

will be modified by a factor of $\frac{1}{\mu^3}$ and then used as the initial α on the next grid. On this refined grid, five iterates of the geometric progression for α will be performed with the initial iteration equation being processed as in the standard method; i.e., a sequence of $K_{f,hh}$'s and ${K_{f,hh}}^t K_{f,hh}$'s will be constructed, and a sequence of linear systems will be solved using Cholesky decomposition. The next four iteration equations however will not involve the construction of a sequence of $K_{f,hh}$'s and ${K_{f,hh}}^t K_{f,hh}$'s, although a sequence of linear systems will be solved using the preconditioned conjugate gradient scheme for each of these four iteration equations. These two matrices will be constructed only once using the solution of the initial iteration equation on this grid. Next, after obtaining the quasi-optimal solution, the whole procedure is repeated on the finest grid.

It may appear that the construction of a sequence of $K_{f,hh}$'s and ${K_{f,hh}}^t K_{f,hh}$'s for the standard method is a bit unfair. But this is precisely the disadvantage of the single-grid procedure. Linearization of the nonlinear operator permits an error of order $o\left(\left\|\underline{f} - \underline{f}^{\delta}_{i,\alpha_j}\right\|\right)$, which may be large if $\underline{f}^{\delta}_{i,\alpha_j}$ is a bad approximation to $\underline{f}$. Hence, $K_{f,hh}$ and ${K_{f,hh}}^t K_{f,hh}$ should be reconstructed for each iterate of the iteration equation. The finer grid of the multigrid scheme, on the other hand, has a good approximation after the first iteration equation has been solved. Elaborating, the quadrature and linearization errors are possibly balanced enough to omit the continual formation of $K_{f,hh}$ and ${K_{f,hh}}^t K_{f,hh}$.

Example 1. Contact surface problem

We use a set of discrete data and take $a = c = 14.8$; $b = d = 27.0$.

One-fifth of the components of the right-hand side were randomly chosen from a uniform distribution and perturbed by $(0.4 + 0.6\theta)\delta$, where θ is a random variable on $[-1, 1]$. The remaining components were perturbed by 0.4θ.

δ	time (real)	alpha	Newton iterations						$\|error\|_2$
0.01	327.07	3.9063×10^{-3}	2	2	$\cdots$	2	1	1	0.00512749
0.001	22701.70	6.1035×10^{-5}	3	2	$\cdots$	2	2	3	0.000807030
0.0005	20477.30	1.5259×10^{-5}	3	2	$\cdots$	2	2	2	0.000410676

Table 5.28: Contact surface equation using a standard Tikhonov method with first-order stabilizer.

δ	time (real)	alpha	Newton/cg iterations					$\|error\|_2$
0.01	74.64	1.7166×10^{-5}	3/-	1/1	1/1	1/1	1/1	0.00746543
0.001	3079.93	2.5749×10^{-5}	2/-	1/2	1/2	1/2	1/2	0.000834885
0.0005	3090.10	3.4332×10^{-5}	2/-	1/3	1/3	1/2	1/3	0.000415502

Table 5.29: Contact surface equation using a multigrid scheme with $\mu^1 = 0.75$, $\mu^2 = 0.5$, $\mu^3 = 0.25$.

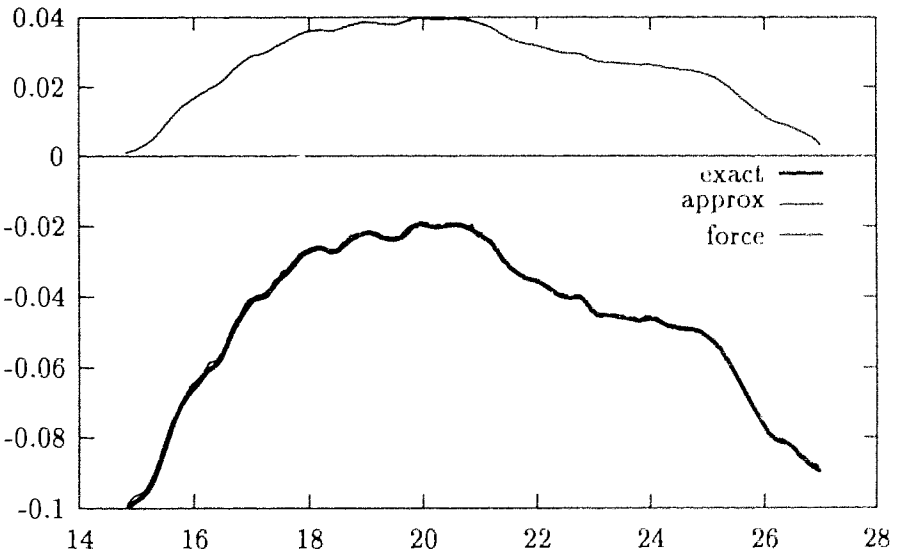

Figure 5.15: Contact surface for $\delta = 0.001$.

Chapter 6

Conclusion

We have seen through some analysis and numerical experiment that multilevel techniques can reduce the computational costs yet obtain the same convergence rates of several other standard techniques. Indeed, we have seen the complementary property that high frequencies, frequencies that are slowly captured by both the Landweber scheme and the c.g. iteration, are damped out by the r.a., and medium frequencies, frequencies that can be captured slowly on the fine grid, can be damped out on the coarser grids. This complementary property provides the necessary structure of a multilevel method; i.e., the property that components lying approximately in the null space of the smoother can be eliminated on the fine grid, and components lying near the orthogonal complement of the null space of the smoother can be eliminated by the coarse grid correction. Hence, we have seen that the smoother consists not only of an iterative method, but also of a discrepancy principle.

As for the discrepancy principles themselves, we have seen through a survey of the principle how indespensible they are. They are not only part of the smoothing process, but they also have an impact on the stopping criterion of the iterative linear system solver of a zero'th order Tikhonov method, as described in theorem 4.12. Moreover, taking advantage of the fact that several grids are used, by using several discrepancy principles, we have been able to obtain a parameter both accurately (Morozov discrepancy principle on the coarse grid) yet cheaply (generalized discrepancy principle on the finer grids) for the multilevel zero'th order Tikhonov scheme.

Many questions still need to be investigated though. For the multigrid Landweber scheme, it still must be shown whether this "pure" multigrid does form an r.a.. Although we have shown that this scheme approximates the Twomey-Tikhonov method, and although we have empirical evidence that the scheme does work it still must be shown that the approximations do converge to $K^{\dagger}g$ as $\delta \rightarrow 0$. This will most likely involve an investigation of the generating function of this multigrid scheme, and this may involve a generalization of the $R_{\alpha}(t)$ function since restriction, interpolation, and coarse grid operators are involved.

The multilevel zero'th order Tikhonov method requires even more future investigation. So far, all that has been created is a multilevel extrapolation technique. Through an analysis of the c.g. iteration and the decomposition of the regularized solution, only fine grid systems were solved after an extrapolation of a coarse grid solution and its corresponding parameter. Although this method proved very effective in most of the examples, it would be worthwhile to investigate a multilevel system solver. Such an implementation should not be too difficult for the zero'th order Tikhonov method since an underlying second kind integral equation exists. This modification also can be a starting point for creating a "pure" multigrid Tikhonov method, for each coarse grid defect equation can be viewed as a regularized first kind equation with a specific parameter. What should be examined cautiously, however, is the interpolation error introduced by the multilevel solver.

Of course, creating a "pure" multigrid Tikhonov method is a future goal. The major problem with such a scheme is the solution and parameter updating on the fine grid. If a first kind defect equation is regularized using Tikhonov's method, how should the interpolated regularized solution be added to the fine grid approximation, and how should the parameter be modified (remember, the regularized defect equation most likely will have a different parameter)? Note that this problem did not arise in the multilevel Landweber iteration since in there, the coarse grid "parameter" α indirectly modified the fine grid parameter n^{-1}. A possible remedy to this problem is to apply a weighted parameter updating scheme. The extrapolation analysis of King and Chillingworth (1979), and Groetsch and King (1979) may be helpful here.

Turning to the multilevel first order Tikhonov method, clearly a number of changes must be made to this scheme. Although it was more efficient than a standard Tikhonov method in the examples, its efficiency was not too impressive due to the large number of c.g. iterations

required. The first change that should be made is to apply the Morozov discrepancy principle on the coarsest grid, and then with larger μ's in the parameter progression formula

$$\alpha_k = \mu\alpha_{k-1} ,$$

we can apply the quasi-optimal parameter choice on the finer grids. Now with the larger μ's, the smaller changes in the regularization parameter may imply less c.g. iterations required to solve the updated regularized systems. Furthermore, multilevel system solvers can be applied to these systems.

And as for the nonlinear multigrid scheme, the field of investigation is enormous. One task is to try the multilevel scheme on other nonlinear first kind equations in order to determine any possible difficulties with the scheme. Two other tasks are to investigate other parameter choosing procedures, and to explore FAS-type methods, which would probably involve delicate nonlinear discrepancy principles. Now these goals may be ambitious, but the success of the multigrid method to the contact surface problem should be an impetus for such future research.

Bibliography

[1] Alifanov, O.M., Rumjancev, S.V., On the stability of iterative methods for the solution of linear ill-posed problems, Soviet Math. Dokl., Vol.20, 1133-1136 (1979).

[2] Arsenin, V.Y., Problems on computer diagnostics in medicine, in *Ill-posed Problem in the Natural Sciences*, A.N. Tikhonov, A.V. Goncharsky, eds., 206-219, MIR Publishing, Moscow (1987).

[3] Baker, C.T.H., *The Numerical Treatment of Integral Equations*, Oxford University Press, Oxford (1977).

[4] Bakushinskii, A.B., A numerical method for solving Fredholm integral equations of the 1'st kind, U.S.S.R. Comput. Maths. Math. Phys., Vol.5, No.4, 226-233 (1964).

[5] Bakushinskii, A.B., A general method of constructing regularizing algorithms for a linear ill-posed equation in Hilbert space, U.S.S.R. Comput. Maths. Math. Phys., Vol.7, No.3, 279-287 (1966).

[6] Bakushinskii, A.B., The discrepancy principle in the case of a disturbed operator for general regularizing algorithms, U.S.S.R. Comput. Maths. Math. Phys., Vol.22, No.4, 225-230 (1982).

[7] Bakushinskii, A.B., Sizikov, V.S., Some non-standard regularizing algorithms and their numerical realization, U.S.S.R. Comput. Maths. Math. Phys., Vol.22, No.3, 37-44 (1982).

[8] Baumeister, J., *Stable Solution of Inverse Problems*, Vieweg, Braunschweig(1987).

[9] Balakrishnan, A.V., *Applied Functional Analysis*, Springer Verlag, New York-Heidelberg-Berlin (1981).

[10] Bennell, R.P., Mason, J.C., Continuous approximation methods for the regularization and smoothing of integral transforms, Rocky Mountain Jour. of Math., Vol. 19, No.1, 51-66 (1989).

[11] Braess, D., On the combination of the multigrid method and conjugate gradients, in *Multigrid Method II*, W. Hackbusch, U. Trottenberg, eds., 52-64, Springer Verlag, Berlin-Heidelberg-New York-Tokyo (1986).

[12] Brandt, A., Multilevel computations: review and recent developments, in *Multigrid Method: Theory, Applications and Supercomputing*, S.F. McCormick, ed., 35-62, Marcel Dekker, New York (1988).

[13] Brandt, A., The Weizmann Institute research in multilevel computation: 1988 report, in *Proceedings of the Fourth Copper Mountain Conference on Multigrid Methods*, S. F. McCormick, J.E. Dendy, Jr., C. Farhat, G. Lonsdale, S.Y. Parter, J.W. Ruge, K. Stüben, eds, 13-53, SIAM, Philadelphia (1989).

[14] Brandt, A., Lubrecht, A.A., Multilevel multi-integration and fast solution of integral equations, The Weizmann Institute of Science Report (1988).

[15] Briggs, W.L., *A Multigrid Tutorial*, SIAM, Philadelphia (1987).

[16] Chen, W.H., Zhang, F.G., Multigrid strategy for effenciency improvement of GPST inversion algorithm, Appl. Num. Math., Vol.6, 431-446 (1989).

[17] Chatelin, F., Miranker,W.L., Acceleration by aggregation of succesive approximation methods, Linear Algebra Appl., Vol.43, 17-47 (1982).

[18] Delves, L.M., Mohamed, J.L., *Computational Methods for Integral Equations*, Cambridge University Press, Cambridge (1985).

[19] Diaz, J.B., Metcalf, F.T., On iteration procedures for equations of the first kind, Ax=y, and Picard's criterion for the existence of a solution, Math. Comp., Vol.24, No.112, 923-935 (1970).

[20] Dmitriev, V.I., Karus, E.V., Inverse problem in geophysical prospecting, in *Ill-posed Problems in the Natural Sciences*, A.N. Tikhonov, A.V. Goncharsky, eds., 70-76, MIR Publishers, Moscow (1987).

[21] Eldén, L., Algorithms for the regularization of ill-conditioned least squares problems, BIT, Vol.17, 134-145 (1977).

[22] Emelin, I.V., Krashosel'skii, M.A., A stopping rule in iteration procedures for solving ill-posed problems, Automat. Remote Control, Vol.39, 1783-1786 (1978).

[23] Emelin, I.V., Krasnosel'skii, M.A., On the theory of ill-posed problems, Soviet Math. Dokl., Vol.20, No.1, 105-109 (1979).

[24] Engl, H.W., Necessary and sufficient conditions for convergence of regularization methods for solving linear operator equations of the first kind, Numer. Funct. Anal. Optim., Vol.3, 201-222 (1981).

[25] Engl, H.W., Discrepancy principles for Tikhonov regularization of ill-posed problems leading to optimal convergence rates, J. Optimiz. Th. Appl., Vol.52, No.2, 209-215 (1987).

[26] Engl, H.W., On the choice of the regularization parameter for iterated Tikhonov regularization of ill-posed problems, J. Approx. Th., Vol.49, 55-63 (1987).

[27] Franklin, J.N., On Tikhonov's method for ill-posed problems, Math. Comp., Vol.28, No.128, 889-907 (1974).

[28] Friedman, A., *Foundations of Modern Analysis*, Dover, New York (1970).

[29] Gfrerer, H., An a posteriori parameter choice for ordinary and iterated Tikhonov regularization of ill-posed problems leading to optimal convergence rates, Math. Comp., Vol. 49, No.180, 507-522, S5-S12 (1987).

[30] Glasko, V. B., *Inverse problems of Mathematical Physics*, (English translation), American Institute of Physics, New York (1988).

[31] Glasko, V. B., Gushchin, G.V., Gushchina, L.G., Mudretsova, E.A., The use of boring data to reconstruct the shape of a contact by the regularization method, U.S.S.R. Comput. Maths. Math. Phys., Vol.14, No.5, 174-182 (1974).

[32] Glasko, V. B., Mudretsova, E.A., Strakhov, V. N., Inverse problems in gravimetry and magnetometry, in *Ill-posed Problems in the Natural Sciences*, A.N. Tikhonov, A.V. Goncharsky, eds., 115-124, MIR Publishers, Moscow (1987).

[33] Golub, G.H., Van Loan, C.F., *Matrix Computations*, John Hopkins University Press, Baltimore (1983).

[34] Golub, G.H., Overton, M.L., Convergence of a two-stage Richardson iterative procedure for solving systems of linear equations, in *Numerical Analysis* (912), G.A. Watson, ed., 125-139, Springer Verlag, Berlin-Heidelberg-New York (1982).

[35] Goncharsky, A.V., Ill-posed problems and their solution methods, in *Ill-posed Problems in the Natural Sciences*, A.N. Tikhonov, A.V. Goncharsky, eds., 21-52, MIR Publishers, Moscow (1987).

[36] Goncharsky, A.V., Leonov, A.S., Yagola, A.G., A generalized discrepancy principle, U.S.S.R. Comput. Maths. Math. Phys., Vol. 13, No.2, 25-37 (1973).

[37] Goncharsky, A.V., Leonov, A.S., Yagola, A.G., Finite-difference approximation of linear incorrectly posed problems, U.S.S.R. Comput. Maths. Math. Phys., Vol.14, No.1, 14-23 (1974).

[38] Gordonova, V.I., Morozov, V.A., Numerical parameter selection algorithms in the regularization method, U.S.S.R. Comput. Maths. Math. Phys., Vol. 13, No.3, 1-9 (1973).

[39] Greenbaum, A., Analysis of a multigrid method as an iterative technique for solving linear systems, SIAM J. Numer. Anal., Vol.21, No.3, 473-485 (1984).

[40] Groetsch, C.W., *Generalized Inverses of Linear Operators: Representation and Approximation*, Marcel Dekker, New York (1977).

[41] Groetsch, C.W., The parameter choice problem in linear regularization: a mathematical introduction, unpublished (1978).

[42] Groetsch, C.W., *Elements of Applicable Functional Analysis*, Marcel Dekker, New York (1980).

[43] Groetsch, C.W., On a regularization-Ritz method for Fredholm equation of the first kind, J. Integral Equations, Vol.4, 173-182 (1982).

[44] Groetsch, C.W., *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind*, Pitman Publishings, Boston (1984).

[45] Groetsch, C.W., Convergence analysis of a regularized kernel method for Fredholm integral equations of the first kind, Integral Equations and Operator Theory, Vol.13 (1990).

[46] Groetsch, C.W., King, J.T., Extrapolation and the moment of regularization for generalized inverses, J. Approx. Th., Vol.25, 233-247 (1979).

[47] Groetsch, C.W., Vogel, C.K., Asymptotic theory of filtering for linear operator equations with discrete noisy data, Math. Comput., Vol.49, No.180, 499-506 (1987).

[48] Guacaneme, J.E., An optimal parameter choice for regularized ill-posed problems, Integral Equations and Operator Theory, 610-613 (1988).

[49] Gushchin, G.V., Gushchina, L.G., Some regularizing algorithms and their use in the oil pool prospecting problem, U.S.S.R. Comput. Maths. Math. Phys., Vol. 16, No.4, 106-117 (1976).

[50] Hackbusch, W., *Multigrid Methods and Applications*, Springer Verlag, Berlin(1985).

[51] Hanke, M., Accelerated Landweber iterations for the solution of ill-posed equations, to appear.

[52] Hansen, P.C., The truncated svd as a method for regularization, BIT, Vol.27, 534-553 (1987).

[53] Hansen, P.C., Regularization, gvsd and truncated gsvd, BIT, Vol.29, 491-504.

[54] Hansen, P.C., Truncated singular value decomposition solutions to discrete ill-posed problems with ill-determined numerical rank, SIAM J. Sci. Stat. Comput., Vol.11, No.3, 503-518 (1990).

[55] Hanson, R.J., A numerical method for solving Fredholm integral equations of the first using singular values, SIAM J. Numer. Anal., Vol.8, No.3, 616-621 (1971).

[56] Hickey, K.K., Luecke, G.K., Remarks on Marti's method for solving first kind equations, SIAM J. Num. Anal., Vol.19, No.3, 623-628 (1982).

[57] Ivanov, V.K., Integral equations of the first kind and an approximation solution for the inverse problem of potential, Soviet Math. Dokl., Vol.3, No.1, 210-212 (1962).

[58] Ivanov, V.K., On linear problems which are not well-posed, Soviet Math. Dokl., Vol.3, No.4, 981-983 (1962).

[59] Jagola, A.G., On the choice of regularization parameter from the generalized residual principle, Soviet Math. Dokl., Vol.20, No.2, 275-277 (1979).

[60] Jennings, A., Influence of the eigenvalue spectrum on the convergence rate of the conjugate gradient method, J. Inst. Math. Applic., Vol.20, 61-72 (1977).

[61] Jesperson, D.C., Multigrid methods for partial differential equations, in *Studies in Numerical Analysis*, G.H. Golub, ed., Mathematical Assoc. Amer. (1984).

[62] Kammerer, W.J., Nashed, M.Z., Steepest descent for singular linear operators with nonclosed range, Appl. Anal., Vol.1, 143-159 (1971).

[63] Kammerer, W.J., Nashed, M.Z., Iterative methods for best approximate solutions of linear integral equations of the first and second kinds, J. Math. Anal. Appl., Vol.40, 547-573 (1972).

[64] Kantorovich, L.V., Akilov, G.P., *Functional Analysis*, (English translation), Pergamon Press, Oxford (1982).

[65] King, J.T., Serbin, S.M., Computational experiments and techniques for the penalty method with extrapolation, Math. Comp., Vol.32, No.141, 111-126 (1978).

[66] King, J.T., Chillingworth, D., Approximation of generalized inverses by iterated regularization, Numer. Funct. Anal. Optim., Vol.1, 499-513 (1979).

[67] King, J.T., Neubauer, A., A variant of finite-dimensional Tikhonov regularization with a posterior parameter choice, Computing, Vol.40, 91-109 (1988).

[68] Kryanev, A.V., An iterative method for solving incorrectly posed problems, U.S.S.R. Comput. Maths. Math. Phys., Vol.14, No.1, 24-35 (1974).

[69] Lavrent'ev, M.M., Romanov, V.G., Shishat-skiĭ, S.P., *Ill-posed Problems of Mathematical Physics and Analysis*, (English translation), American Mathematical Society, Providence (1986).

[70] Lee, J.W., Prenter, P.M., An analysis of the numerical solution of Fredholm integral equations of the first kind, Numer. Math., Vol.30, 1-23 (1978).

[71] Liskovets, O.A., Theory and the methods of solutions of ill-posed problems, J. Soviet Math., Vol.34, 1656-1696 (1986).

[72] Locker, J., Prenter, P.M, Regularization with differential operators. I. General theory, J. Math. Anal. Appl., Vol.74, 504-529 (1980).

[73] Mandel, J., On multilevel iterative methods for integral equations of the second kind and related problems, Numer. Math., Vol.46, No.46, 147-157 (1985).

[74] Mandel, J., Sekerka, B., A local convergence proof for the iterative aggregation method, Linear Algebra Appl., Vol.51, 163-172 (1983).

[75] Marti, J.T., An algorithm for recomputing minimum norm solutions of Fredholm integral equations of the first kind, SIAM J. Numer. Anal., Vol.15, 1071-1076 (1978).

[76] Marti, J.T., On the numerical stability in solving ill-posed problems, in *Numerical Treatment of Integral Equations*, J. Albrecht, L. Collatz, eds., 171-182, Birkhäuser Verlag, Basel (1980).

[77] Marti, J.T., On the convergence of an algorithm computing minimum-norm solutions of ill-posed problems, Math. Comput., Vol.34, No.150, 521-527 (1980).

[78] McCormick, S.F., An algebraic interpretation of multigrid methods, SIAM J. Numer. Anal., Vol.19, No.3, 548-560 (1982).

[79] Miranker, W.L., Pan, V.Y., Methods of aggregation, Linear Algebra Appl., Vol.29, 231-157 (1980).

[80] Morozov, V.A., Choice of parameter for the solution of functional equations by the regularization method, Soviet Math. Dokl., Vol.8, No.4, 1000-1003 (1967).

[81] Morozov, V.A., The error principle in the solution of operational equations by the regularization method, U.S.S.R. Comput. Maths. Math. Phys., Vol.8, No.2, 63-87 (1968).

[82] Morozov, V.A., On optimality of the discrepancy criterion in the problem of computing the values of unbounded operators, U.S.S.R. Comput. Maths. Math. Phys., Vol.11, No.4, 250-257 (1971).

[83] Morozov, V.A., Convergence of an approximate method of solving operator equations of the first kind, U.S.S.R. Comput. Maths. Math. Phys., Vol.13, No.1, 1-20 (1973).

[84] Morozov, V.A., The error principle in the solution of incompatible equations by Tikhonov regularization, U.S.S.R. Comput. Maths. Math. Phys., Vol.13, No.5, 1-16 (1973).

[85] Morozov, V.A., *Methods for Solving Incorrectly Posed Problems*, (English translation), Springer Verlag, New York (1984).

[86] Nashed, M.Z., Steepest descent for singular linear operator equations, SIAM J. Numer. Anal., Vol.7, No.3, 358-362 (1970).

[87] Nashed, M.Z., On moment-discretization and least-squares solutions of linear integral equations of the first kind, J. Math. Anal. Appl., Vol.53, 359-366 (1976).

[88] Natterer, F., The finite element method for ill-posed problems, R.A.I.R.O. Analyse Numérique, Vol.11, No.3, 271-278 (1977).

[89] Oldenburg, D.W., An introduction to linear inverse theory, IEEE Transactions on Geoscience and Remote Sensing, Vol. GE-22, No.6, 665-674 (1984).

[90] Parker, R.L., Understanding inverse theory, Ann. Rev. Earth Planet. Sci., Vol. 5, 36-64 (1977).

[91] Phillips, D.L., A technique for the numerical solution of certain integral equations of the first kind, J. Asso. Comput. Mach., Vol.9, 84-97 (1962).

[92] Plato, R., Vainikko, G., On the regularization of projection methods for solving ill-posed problems, Numer. Math., Vol.57, 63-79 (1990).

[93] Ruge, J.W., Stüben,K., Algebraic multigrid, in *Multigrid Methods*, S.F. McCormick, ed., 73-130, SIAM, Philadephia (1987).

[94] Saaty, T.L., *Modern Nonlinear Equations*, Dover, New York (1981).

[95] Saaty, T.L., Bram, J., *Nonlinear Mathematics*, Dover, New York (1981).

[96] Schock, E., Parameter choice by discrepacy principle for the approximate solution of ill-posed problems, Integral Equations and Operator Theory, Vol.7, 895-898 (1984).

[97] Smithies, F., *Integral Equations*, Cambridge Unversity Press, London (1958).

[98] Strand, O.N., *Theory and Methods for Operator Equation of the First Kind*, Doctorial thesis, Colorado State Univ., Fort Collins (1972).

[99] Strand, O.N., Theory amd methods related to the singular-function expansion and Landweber's iteration for integral equations of the first kind, SIAM J. Numer. Anal., Vol.11, No.4, 798-825 (1974).

[100] Stewart, G.W., The convergence of the method of conjugate gradients at isolated extreme points of the spectrum, Numer. Math., Vol.24, 85-93 (1975).

[101] Stüben,K., Algebraic multigrid(AMG): experiences and comparison, Appl. Math. Comp., Vol.13, 419-451 (1983).

[102] Szeliski, K., Terzopoulos, D., Parallel multigrid algorithms and computer vision applications, in *Proceedings of the Fourth Copper Mountain Conference on Multigrid Method*, S.F. McCormick, J.E. Dendy, Jr., C. Farhat,G. Lonsdale, S.Y. Parter, J.W. Ruge, K. Stüben, eds., 383-398, SIAM, Philadephia (1989).

[103] Tikhonov, A.N., Solution of nonlinear integral equations of the first kind, Soviet Math. Dokl., Vol.5, No.3, 835-838 (1964).

[104] Tikhonov, A.N., On the problems with approximately specified information, in *Ill-posed Problems in the Natural Sciences*, A.N. Tikhonov, A.V. Goncharsky, eds., 13-20, MIR Publishers, Moscow (1987).

[105] Tikhonov, A.N., Glasko, V.B., Use of the regularization method in non-linear problems, U.S.S.R. Comput. Maths. Math. Phys., Vol. 5, No.3, 93-107 (1965).

[106] Tikhonov, A.N., Arsenin, V.Y., *Solutions of Ill-posed Problems*, (English translation), John Wiley, New York (1977).

[107] Trummer, M.R., A method for solving ill-posed linear operator equations, SIAM J. Numer. Anal., Vol.21, No.4, 729-737 (1984).

[108] Twomey, S., On the numerical solution of Fredholm integral equations of the first kind by the inversion of the linear system produced by quadrature, J. Asso. Comput. Math., Vol.10, 97-101 (1963).

[109] Vainikko, G.M., Error estimates of the succesive approximation method for ill-posed problems, Automat. Remote Control, Vol.40, 356-363 (1980).

[110] Vainikko, G.M., The discrepancy principle for a class of regularization methods, U.S.S.R. Comput. Maths. Math. Phys., Vol.22, No.3, 1-19 (1982).

[111] Vainikko, G.M., The critical level of discrepancy in regularization methods, U.S.S.R., Comput. Maths. Math. Phys., Vol.23, No.6, 1-9 (1983).

[112] Varah, J.M., On the numerical solution of ill-conditioned linear systems with application to ill-posed problems, SIAM J. Numer. Anal., Vol.10, No.2, 257-267 (1973).

[113] Varah, J.M., A practical examination of some numerical methods for linear discrete ill-posed problems, SIAM Review, Vol.21, No.1, 1000-1111 (1974).

[114] Vasin, V.V., A general scheme for discretization of regularizing algorithms in Banach spaces, Soviet Math. Dokl., Vol.23, No.3, 494-498 (1981).

[115] Vasin, V.V., Iterative methods for solving ill-posed problems with a prior information in Hilbert spaces, U.S.S.R., Comput. Maths. Math. Phys., Vol.28, No.4, 6-13 (1988).

[116] Veretennikov, A. Yu., Krdsnosel'skiĭ, M.A., Regularizing stopping rules under conditions of random errors, Soviet Math. Dokl., Vol.27, No.1, 90-94 (1983).

[117] Voevodin, V.V., The method of regularization, U.S.S.R., Comput. Maths. Math. Phys., Vol.9, No.3, 228-232 (1969).

[118] Wahba, G., Convergence rates of certain approximate solutions to Fredholm integral equations of the first kind, J. Approx. Th., Vol.7, 167-185 (1973).

[119] Wahba, G., Practical approximate solutions to linear operator equation when the data are noisy, SIAM J. Numer. Anal., Vol.14, No.4, 651-667 (1977).

[120] Zhu, J., Using a hypercube to solve inverse problems in reservoir simulations, SIAM News, Vol.25, No.2 (1992).