Integral equations of the first kind, inverse problems and regularization: a crash course

C. W. Groetsch
School of Science and Mathematics, The Citadel, Charleston, SC 29409 U.S.A.
E-mail: charles.groetsch@citadel.edu

Abstract. This paper is an expository survey of the basic theory of regularization for Fredholm integral equations of the first kind and related background material on inverse problems. We begin with an historical introduction to the field of integral equations of the first kind, with special emphasis on model inverse problems that lead to such equations. The basic theory of linear Fredholm equations of the first kind, paying particular attention to E. Schmidt’s singular function analysis, Picard’s existence criterion, and the Moore-Penrose theory of generalized inverses is outlined. The fundamentals of the theory of Tikhonov regularization are then treated and a collection of exercises and a bibliography are provided.

1. Historical Background
Most of history is guessing, and the rest is prejudice.
W. & A. Durant

The recasting of Ivar Fredholm’s theory of linear integral equations of the second kind by David Hilbert and Erhardt Schmidt in the first decade of the last century has had an enormous influence on modern mathematics. Indeed, the Hilbert-Schmidt geometrization of analysis, which is an outgrowth of Fredholm’s theory of integral equations of the second kind, is one of the great triumphs of twentieth century mathematics. On the other hand, it seems that integral equations of the first kind are much less familiar, even exotic. But, as the terminology suggests, integral equations of the first kind came first (the term “integral equation” seems to have been coined by du Bois-Reymond in 1888 [15]). As remarked by one of the early chroniclers of the theory of integral equations, Maxime Bôcher [14]:

The theory of integral equations may be regarded as dating back at least as far as the discovery by Fourier of the theorem concerning integrals which bears his name; for, though this was not the point of view of Fourier, this theorem may be regarded as a statement of the solution of a certain integral equation of the first kind.

Here Bôcher is referring to Fourier’s “inversion” formula for the integral equation of the first kind

\[ g(x) = \sqrt{\frac{2}{\pi}} \int_0^\infty \cos (x\xi)f(\xi) d\xi \]

namely,

\[ f(x) = \sqrt{\frac{2}{\pi}} \int_0^\infty \cos (x\xi)g(\xi) d\xi. \]
A more generally accepted view is that the theory of integral equations began with an inverse problem – Abel’s analysis of the mechanical problem of finding the curve of descent, given the time of descent as a function of the vertical distance of fall \[1\]. The problem is to find the unknown path in the plane along which a particle will fall, under the influence of gravity alone, so that at each instant the time of fall is a known function of the distance fallen. Suppose that the particle falls from height \(z\) and that the path of descent is parameterized by arclength \(s\), that is, at time \(t\) the length of arc traversed is \(s(t)\) (\(s(0) = 0\)). Assuming that the particle starts from rest, we find by equating the gain in kinetic energy to the loss in potential energy that:

\[
\frac{1}{2} \left( \frac{ds}{dt} \right)^2 = a(z - y),
\]

where \(a\) is the gravitational acceleration. Integrating this, we find that the time of descent from \(z\) to the base line \(y = 0\), \(g(z)\) is given by

\[
g(z) = \int_{y=0}^{y=z} ds \sqrt{2a(z - y)}.
\]

Setting \(\frac{ds}{dy} = -f(y)\), this gives

\[
g(z) = \int_{0}^{z} f(y) \sqrt{2a(z - y)} dy. \tag{1}
\]

This equation, in which the descent time \(g\) is a given function and \(f\) is the unknown function, from which a parameterization of the desired path may be obtained, is an example of an integral equation of the first kind.

In the two decades following Abel’s publication of the solution to his mechanical problem, the most discussed problem in planetary astronomy concerned the strange behavior of the outer most known planet: Uranus. Its orbit did not conform to that which Newton’s law of gravity, applied to the then known solar system, would dictate. This led some, including the Astronomer Royal, George Airy, to question the validity of Newton’s law of gravitational force. Others felt that Newton’s theory could account for the irregularities in Uranus’s orbit by hypothesizing the existence of a still more distant but as yet undiscovered planet. This, of course, turned out to be the case (see for example \[87\], \[42\] and \[86\] for popular accounts of this mathematical hunt for the new planet, Neptune). Inspired by Abel’s result, and perhaps motivated by the earlier controversy over the true law of gravitational attraction, Ferdinand Joachimstahl \[53\], published in 1861 a solution of a simple inverse problem leading to an integral equation of the first kind.

Suppose an infinitely long wire of unit mass density, identified with the vertical \(y\)-axis, exerts a total horizontal attraction of \(g(h)\) on a unit mass that is \(h\) units distant from the wire. Suppose the force of attraction between two masses is the product of the masses times an unknown function \(f(r)\) of the distance \(r\) between the masses (e.g., in Newton’s law, \(f(r) \propto r^{-2}\)). Using \(\theta\) to denote the acute angle between the horizontal and the segment connecting the point mass to an infinitesimal mass element at position \(y\) on the vertical attracting line, one finds that

\[
g(h) = 2 \int_{0}^{\infty} f(r) \cos \theta dy = 2 \int_{0}^{\infty} f(r) \frac{h}{r} dy = 2h \int_{h}^{\infty} \frac{f(r)}{\sqrt{r^2 - h^2}} dr.
\]

Therefore, if the attraction \(g\) is given, then the unknown law of attraction \(f\) is the solution of the integral equation

\[
g(h) / 2h = \int_{h}^{\infty} \frac{f(r)}{\sqrt{r^2 - h^2}} dr.
\]
Joachimstahl's posing of a model inverse problem as an integral equation of the first kind was of little consequence, but at the turn of the twentieth century a general theory of integral equations of the second kind was developed, first by Volterra and Fredholm, and then by Hilbert and Schmidt. In large part the theory was developed to address existential mathematical questions arising out of specific problems in mechanics and mathematical physics (see [14] for an excellent source on the early development of integral equations; see also [10]). Hilbert and his student Schmidt constructed an entirely new geometrical context, now called Hilbert space, for the analysis of linear integral equations of the second kind. At first the theory was restricted to integral equations with symmetric kernels, but Schmidt's invention of systems of what are now known as "singular functions" (adjungierter Eigenfunktionen, see [85]) allowed extension of the theory to non-symmetric problems.

Schmidt's theory of singular function expansions gave E. Picard [78] the tools he needed to prove his necessary and sufficient condition for the existence of solutions of a Fredholm integral equation of the first kind (now called the Picard's Criterion). Picard was concerned only with existence of solutions and did not address the problem of approximation of solutions.

As will be seen below, Picard's analysis of Fredholm integral equations of the first kind shows that these equations model ill-posed problems. The notion of an ill-posed problem, or more correctly that of a well-posed problem, was introduced by Hadamard [43] more than a century ago (see also [44]). Within the context of the theory of partial differential equations Hadamard termed a problem well-posed if it has a solution (existence), it does not have more than one solution (uniqueness), and this solution depends continuously on the data of the problem (stability). Problems that are not well-posed have come to be called ill-posed, and Fredholm integral equations of the first kind are prime exemplars of ill-posed problems (see the exercises). Hadamard was initially of the opinion that only well-posed problems were appropriate object of study in mathematical physics. But, as will be seen in the many examples and exercises given below, numerous ill-posed problems arise naturally in modern science and technology. The regularization theory introduced by Phillips [77] and Tikhonov [92], and developed to a high degree by Tikhonov's school, is a methodology whose goal is to stand up to the most serious numerical challenge that ill-posed problems present, namely instability.

In 1943 Tikhonov [91] made an important theoretical advance concerning stability of solutions of inverse problems. He showed that the stability of the solution of the inverse problem of determining the spatial distribution of mass situated below a surface that produces a given gravitational potential on the surface, can be assured if the allowable mass distributions are restricted to lie in a compact subset of a certain function space. The stability of the solution of the inverse problem is then a consequence of what has now become a familiar abstract topological theorem, namely, that a continuous injective function defined on a compact set has a continuous inverse.

The new technologies arising out of World War II necessitated practical numerical solutions of inverse problems and the fledgling digital computers of the day provided for the first time the means of realizing these solutions. In particular, in the 1950's attempts to construct numerical solutions of integral equations of the first kind brought an awareness of the peculiar instabilities that are inherent in such solutions (see e.g., [12]: “the effects of very different climatic histories will be indistinguishable ...”; [33]: “These solutions ... are much less accurate, bearing little resemblance to the true solution.”; [29]: “The major difficulty being observed ... is the instability of the solution of the integral equation ...”). The need to tame, or at least mitigate, these instabilities led Phillips and Tikhonov to independently develop the first regularization methods for approximation of solutions in 1962 and 1963, respectively (see [77], [92]). While Phillips' work was based largely on intuitive ideas, Tikhonov's approach led to a rich mathematical structure that has been developed to an extraordinary extent over the past four decades (see also [41] for more on the relationship between the work of Phillips and Tikhonov).
2. Some Model Inverse Problems

The cause is hidden, but the result is known.

Ovid

In this section we introduce a half-dozen model inverse problems that may be formulated as integral equations of the first kind. We deal exclusively with inverse problems of causation type. That is, we assume a known model in which some cause evolves continuously into a unique effect. The problem is to determine the cause given sufficient knowledge of the effect. Such inverse problems are typically modeled by integral equations of the first kind. We present a few examples in this section; further examples may be found in the exercises and in the works listed in the bibliography.

2.1. A Gravitation Problem

Suppose mass is distributed on a circular ring of radius 1/2 centered on the origin with density \( f = f(\theta) \), where \( \theta \) is the polar angle. At points on the concentric circle of radius 1 in the same plane, the centrally directed component of gravitational force \( g(\varphi) \), at polar angle \( \varphi \) is measured. Then a relationship of the form

\[
g(\varphi) = \int_0^{2\pi} k(\phi, \theta) f(\theta) d\theta
\]

relating the density and centrally directed force is easy to derive. The distance \( r \) between a mass element at position \( \theta \) on the inner ring and a point on the outer ring located at polar angle \( \varphi \) satisfies, according to the law of cosines:

\[
r^2 = \frac{5}{4} - \cos (\varphi - \theta).
\]

Similarly, by the law of cosines, the angle \( \psi \) between the centrally directed vector emanating from the attracted point on the unit circle and the vector from the attracted point and the gravitating element \( f(\theta)d\theta \) on the inner circle satisfies

\[
\cos \psi = \frac{1 - \frac{1}{4} \cos (\varphi - \theta)}{r}
\]

and hence the total centrally directed force on a point at polar angle \( \varphi \) on the outer circle is

\[
g(\varphi) = \gamma \int_0^{2\pi} \frac{2 - \cos (\varphi - \theta)}{(5 - 4 \cos (\varphi - \theta))^{3/2}} f(\theta) d\theta
\]

where \( \gamma \) is the universal gravity constant. The inverse problem of determining the interior mass distribution \( f \) from observation of the force \( g \) on the outer ring is thus formulated as an integral equation of the first kind.

2.2. Extrapolation of Band-limited Signals

The problem of extrapolation of band-limited signals provides another instance of an inverse problem that gives rise to a Fredholm integral equation of the first kind. In communications engineering a (square integrable) function of time \( f(t), \quad (-\infty < t < \infty) \) is regarded as a “signal” which is a superposition over a continuous spectrum of frequency components. The basic tool in the theory is the time-to-frequency domain map given by the Fourier transform:

\[
\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt.
\]
The Fourier inversion theorem than allows one to analyze a signal in terms of all its frequency components. However, any real signal detector can discern signals only over a limited range of frequencies. A signal \( f(t) \) is called \textit{band-limited} if there is a \( \Omega > 0 \) such that \( \hat{f}(\omega) = 0 \) if \( |\omega| > \Omega \), that is, if the signal contains no frequencies of magnitude greater than \( \Omega \). Band limiting is then accomplished mathematically simply by multiplying \( \hat{f} \) by \( 1_{[-\Omega, \Omega]} \), the characteristic function of the frequency interval \([-\Omega, \Omega] \):

\[
1_{[-\Omega, \Omega]}(\omega) = \begin{cases} 
1, & \omega \in [-\Omega, \Omega] \\
0, & \text{otherwise}
\end{cases}
\]

Let us call the corresponding filtered signal \( g(t) \). Then

\[
g = \mathcal{F}^{-1}\{1_{[-\Omega, \Omega]}\hat{f}\}
\]

and hence, by the convolution theorem,

\[
g = \mathcal{F}^{-1}\{1_{[-\Omega, \Omega]}\} * f.
\]

But, by the Fourier inversion formula,

\[
\mathcal{F}^{-1}\{1_{[-\Omega, \Omega]}\}(t) = \frac{1}{2\pi} \int_{-\Omega}^{\Omega} e^{it\omega} d\omega = \frac{\sin \Omega t}{\pi t}.
\]

Therefore, given the detected band-limited signal \( g(t) \), the full signal \( f \) is reconstituted (or \( g \) is “extrapolated”) by solving the integral equation of the first kind

\[
g(t) = \int_{-\infty}^{\infty} \frac{\sin (\Omega(t - \tau))}{\pi(t - \tau)} f(\tau) d\tau
\]

for \( f(t) \).

2.3. Palaeoclimatology

A well-studied problem in palaeoclimatology concerns the reconstruction of the temperature history, extending to very early times, of the earth’s surface from temperature measurements taken in a borehole at the present time (see e.g., [12], [5]). In this model problem we assume that the borehole is in a semi-infinite (positive \( z \) direction extending downward from the surface) homogeneous earth with constant thermal diffusivity (which for simplicity we take to be 1). We are interested in determining the deviation of temperature from a fixed reference level, a quantity that we will denote by \( u(z, t) \) for \( z, t \geq 0 \). Given temperature measurements \( u(z, T) = g(z) \) taken down the borehole at the present time \( T \), we wish to find the surface temperature history \( f(t) = u(0, t) \) at prior time \( t \leq T \). The simplified governing equations are then

\[
\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial z^2}
\]

\[u(z, 0) = 0, \quad u(0, t) = f(t)\]

and to match the data at time \( T \) we require in addition that

\[u(z, T) = g(z)\]

If we denote by \( s > 0 \) a Laplace transform variable and take Laplace transforms with respect to \( t \), using, as is customary, the corresponding upper case letters for the Laplace transformed variable, the we have

\[
U'' = sU, \quad U(0) = F
\]
where the primes indicate differentiation with respect to $z$. The bounded solution of this ordinary differential equation has the form
\[ U = F(s) \exp(-\sqrt{s}z) \]
and hence, by the convolution theorem,
\[ u(z,T) = \int_0^T w(T) f(T - \tau) d\tau \]
where
\[ w(t) = \mathcal{L}^{-1}\{\exp(-z\sqrt{s})\} = \frac{z}{\sqrt{4\pi t^3}} \exp\left(-\frac{z^2}{4t}\right). \]
and hence if we denote by $\varphi(\tau) = f(T - \tau)$ the surface temperature at a time $\tau \leq T$ in the past, then this surface temperature history satisfies the Fredholm integral equation of the first kind:
\[ g(z) = \frac{z}{\sqrt{4\pi}} \int_0^T T^{-3/2} \exp\left(-\frac{z^2}{4T}\right) \varphi(\tau) d\tau, \]
where $g(z)$ is the borehole temperature profile at the present time.

2.4. An Immunology Problem
We now develop a simple model in immunology relating to the reaction of antigens with antibodies in an equilibrium state (see [47]). Our aim is to derive an integral equation of the first kind for the probability density of the equilibrium constant of the antigen-antibody reaction, which we take to be a random variable.

Consider first the simplest case in which an antigen $AG$ combines with an antibody $AB$ to form a bound antigen-antibody complex $AGAB$:
\[ AG + AB \rightleftharpoons_{k_+}^{k_-} AGAB. \]
The dynamics of the reaction are governed by rate constants, an association rate $k_+$ and a disassociation rate $k_-$. The rate of association is taken to be proportional to the product of the concentrations of antigen and antibody, that is, the association rate is

\[ k_+ [AG][AB] \]
where the brackets indicate concentrations. Similarly, the disassociation rate is

\[ k_- [AGAB]. \]
At equilibrium, we have
\[ k_+ [AG][AB] = k_- [AGAB] \]
and hence, if we define the equilibrium constant, $x$, by $x = k_+/k_-$, then
\[ x[AG][AB] = [AGAB]. \quad (2) \]
The total number of antibodies, $AB_t$, consists of free antibodies $AB$ and bound antibodies $AGAB$ and the concentrations satisfy
\[ [AB_t] = [AB] + [AGAB]. \]
Substituting for $[AB]$ in (2), we obtain

$$x[AG]([AB] - [AGAB]) = [AGAB]$$

and hence

$$\frac{[AGAB]}{[AB]} = \frac{x[AG]}{1 + x[AG]}.$$  \hspace{1cm} (3)

The left hand side of this equation is the fraction of antibody molecules in the bound state. To simplify notation, we will denote the concentration of free antigen by $h$, that is, $h = [AG]$. If we denote the number of antigen molecules bound per molecule of antibody by $\nu(h)$, then, assuming the antibody molecules are $n$-valent (i.e., that each antibody molecule has $n$ receptor sites at which antigen molecules attach), we have

$$n[AGAB] = \nu(h)[AB]$$

and hence by (3)

$$\frac{xh}{1 + xh} = \frac{\nu(h)}{n}.$$  \hspace{1cm} (4)

Finally, we suppose that the equilibrium constant $x$ is actually a random variable with probability density $p(x)$, then, interpreting $\nu(h)$ as the average number of bound antigen molecules per antibody molecule, we have

$$\int_0^\infty \frac{xh}{1 + xh} p(x)dx = \frac{\nu(h)}{n}.$$  \hspace{1cm} (4)

This Fredholm integral equation of the first kind for the probability density $p(x)$ is called the antigen binding equation. The quantity $\nu(h)$ can be determined experimentally for various concentrations $h$ and the goal is to find the density $p(x)$.

2.5. Steady State Heat Distributions

We now treat a simple model inverse problem in the theory of two-dimensional steady state heat conduction. Consider the problem of determining the temperature flux (cause) on the left edge of a semi-infinite strip from observation of the temperature on that face (effect) when the temperature in the strip is at steady state. The problem may be stated mathematically as follows. Let

$$\Omega = \{(x, y) : 0 < x, \quad 0 < y < \pi\}$$

and suppose $u = u(x, y)$ is a function defined on the closure of $\Omega$ and satisfying

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \quad \text{in} \quad \Omega$$

and

$$u(x, 0) = u(x, \pi) = 0 \quad \text{for} \quad x > 0.$$  \hspace{1cm} (5)

Suppose we wish to determine the temperature flux

$$f(y) = \frac{\partial u}{\partial x}(0, y), \quad 0 < y < \pi$$

given the temperature distribution $g(y) = u(0, y)$.  \hspace{1cm} (5)
Elementary separation of variables techniques lead to a representation of the form

\[ u(x, y) = \sum_{n=1}^{\infty} a_n e^{-nx} \sin ny. \]

Proceeding formally, we find that

\[ f(y) = \sum_{n=1}^{\infty} (-na_n) \sin ny \]

and hence

\[ a_n = -\frac{2}{n\pi} \int_{0}^{\pi} f(\xi) \sin n\xi d\xi, \]

while

\[ g(y) = \sum_{n=1}^{\infty} a_n \sin ny \]

\[ = -\sum_{n=1}^{\infty} \frac{2}{n\pi} \int_{0}^{\pi} f(\xi) \sin n\xi d\xi \sin ny \]

\[ = \int_{0}^{\pi} k(y, \xi) f(\xi) d\xi \]

where

\[ k(y, \xi) = -\frac{2}{\pi} \sum_{n=1}^{\infty} \frac{1}{n} \sin ny \sin n\xi. \]

Again the inverse problem is modeled by an integral equation of the first kind.

2.6. Polymer Sedimentation in a Centrifuge

High speed centrifuges can be used to determine the distribution of molecular weights in the chemical constituents of polymers. In this laboratory procedure, a polymer solute is centrifuged until sedimentation equilibrium is achieved. The distribution of molecular weights will then determine a concentration distribution along the radial direction of the centrifuge rotor from the meniscus surface to the sample bottom. This concentration profile may be measured by optical means and the inverse problem consists of determining the molecular weight distribution from the measured concentration profile. Thermodynamical and rheological considerations ([36], [62]) lead to the integral equation of the first kind (known as Fujita’s equation)

\[ g(r) = g_0 \int_{0}^{\infty} \frac{\lambda w \exp \left(-\lambda wr\right)}{1 - \exp (-\lambda w)} f(w) dw \]

where \( w \) is a variable representing the molecular weight (in practice confined to a bounded interval of positive numbers), \( f(w) \) is the molecular weight distribution (i.e., \( f(w)\Delta w \) is the fraction of molecular weights in the sample between \( w \) and \( w+\Delta w \)), \( g_0 \) is the initial concentration in the mixed solute, \( \lambda \) is a parameter determined by the rotor speed, and \( r \) is a radial measure related to the distance from the rotor center (\( r = 1 \) at the meniscus and \( r = 0 \) at the sample bottom). Of course \( f(w) = 0 \) for \( w \) sufficiently large, so the upper limit of integration is actually finite.
3. Integral Equations of the First Kind

Some mathematicians still have a kind of fear whenever they encounter a Fredholm integral equation of the first kind.

Francesco Tricomi

All of the inverse problems of the preceding section gave rise to integral equations of the first kind. For the purposes of this lecture we will consider one-dimensional integral equations on a finite interval, that is, equations of the form

\[ g(x) = \int_{a}^{b} k(x,t)f(t)dt, \quad c \leq x \leq d \]  (5)

where \( k(\cdot, \cdot) \) is a given kernel. Volterra equations of the first kind, namely equations of the form

\[ g(x) = \int_{a}^{x} k(x,t)f(t)dt \]

form a special class of first kind equations in which the kernel satisfies \( k(x,t) = 0 \) for \( t > x \). However, special theory and techniques are applicable to Volterra integral equations of the first kind, as they are intimately related to ordinary differential equations.

In the very special case when the kernel \( k(\cdot, \cdot) \) is degenerate, that is, when it has the form

\[ k(x,t) = \sum_{j=1}^{n} X_j(x)T_j(t) \]

where \( X_j \) and \( T_j \) are functions of a single variable (we may take the \( X_j \) as well as the \( T_j \) to be linearly independent), the equation (5) evidently has a solution if and only if

\[ g \in \text{span}\{X_1, \ldots, X_n\}. \]

In this case solving (5) reduces to the problem of solving a finite system of linear equations. Note that none of the kernels of the integral equations of the first kind appearing in the previous section is degenerate.

A common abstract framework for inverse problems can be constructed in terms of operator equations of the first kind, that is, equations of the form

\[ Kf = g \]  (6)

where \( K : D(K) \subseteq X \rightarrow Y \) is an operator defined on a subset \( D(K) \) of a normed linear space \( X \) and taking values in a normed linear space \( Y \). The equation (6) is well-posed if it has a unique solution \( f \in D(K) \) for each \( g \in Y \) (that is, the inverse operator \( K^{-1} : Y \rightarrow D(K) \) exists) and this unique solution depends continuously on \( g \) (that is, \( K^{-1} \) is continuous). If (6) is well-posed, then \( f \) is stable with respect to small changes in \( g \). It is the stability issue that is of primary concern when attempting to solve (6) because, as can be seen in many of the examples from the previous section, in practical circumstances \( g \) is often a measured quantity and therefore is subject to observational errors. Stability then simply means that small errors in \( g \) will lead to small errors in the solution \( f \).

We shall treat linear Fredholm integral equations of the first kind in a Hilbert space setting. Such equations may be phrased abstractly in the form

\[ Kf = g \]  (7)
where $K : H_1 \rightarrow H_2$ is a bounded linear operator on a real Hilbert space $H_1$, taking values in a real Hilbert space $H_2$. Typically, these Hilbert spaces will be spaces of square integrable functions and the kernel will be a square integrable function of two variables, giving rise to a compact operator.

The inner product on any Hilbert space (we will restrict our attention to real spaces) shall be denoted by $\langle \cdot, \cdot \rangle$ and the associated norm shall be symbolized by $\| \cdot \|$, that is

$$\|f\|^2 = \langle f, f \rangle.$$

The canonical example of a Hilbert space is $L^2[a, b]$, the space of Lebesgue measurable square integrable functions with inner product

$$\langle \varphi, \psi \rangle = \int_a^b \varphi(t)\psi(t)dt.$$

A given kernel $k(\cdot, \cdot) \in L^2([c, d] \times [a, b])$ induces the bounded linear operator $K : L^2[a, b] \rightarrow L^2[c, d]$ defined by

$$g(x) = (Kf)(x) = \int_a^b k(x, t)f(t)dt.$$  \hspace{1cm} (8)

The adjoint of a bounded linear operator $K : H_1 \rightarrow H_2$ is the bounded linear operator $K^* : H_2 \rightarrow H_1$ satisfying

$$\langle K\varphi, \psi \rangle = \langle \varphi, K^*\psi \rangle$$

for all $\varphi \in H_1$ and all $\psi \in H_2$. The adjoint of the linear integral operator $K$ defined by (8) is the integral operator $K^*$ given by

$$(K^*g)(t) = \int_c^d k(x, t)g(x)dx.$$

A sequence $\{f_n\}$ is said to converge weakly to a vector $f$ in the Hilbert space, denoted, $f_n \rightharpoonup f$, if

$$\langle f_n, \varphi \rangle \rightarrow \langle f, \varphi \rangle$$

for all $\varphi$ in the Hilbert space. Note that if the operator $K$ is a finite rank operator, e.g., an integral operator induced by a degenerate kernel, and $f_n \rightarrow f$, then $Kf_n \rightarrow Kf$. An operator with this property is called compact. Square integrable kernels generate compact linear operators on the corresponding $L^2$-spaces (see, e.g., [59]).

E. Schmidt’s [85] theory of singular functions is a foundational tool for the analysis of compact linear operators on Hilbert space. In a nutshell, Schmidt’s theory establishes the existence of a singular system $\{v_j, u_j; \mu_j\}_{j=1}^\infty$ for a compact linear operator with infinite rank. Here $\{v_j\}_{j=1}^\infty$ is a complete orthonormal system for $N(K)^\perp$, the orthogonal complement of the nullspace of $K$, $\{u_j\}_{j=1}^\infty$ is a complete orthonormal system for the closure of the range of $K$, that is, for $\overline{R(K)} = N(K^*)^\perp$. The positive numbers $\{\mu_j\}_{j=1}^\infty$ are called the singular values of $K$. The sequence of singular values converges to zero and the singular values provide a link between the two orthonormal systems of singular vectors:

$$Kv_j = \mu_j u_j \quad \text{and} \quad K^* u_j = \mu_j v_j.$$
the equation (6), with a compact linear operator $K$, has a solution for a given $g \in \overline{R(K)}$ if and only if the singular components of $g$ decay sufficiently rapidly for the condition

$$ \sum_{j=1}^{\infty} \mu_j^{-2} |\langle g, u_j \rangle|^2 < \infty $$

to hold. This condition has come to be known as Picard’s Criterion.

For example, in the one-dimensional heat equation (see Exercise 11), the singular system is

$$ v_j(s) = u_j(s) = \sqrt{2/\pi} \sin js; \quad \mu_j = e^{j^2/2}. $$

In this case a temperature distribution $g$ lies in the range of the appropriate compact operator $K$ if and only if

$$ \sum_{j=1}^{\infty} e^{j^2} |\langle g, u_j \rangle|^2 < \infty $$

that is, the Fourier coefficients of $g$ must decay very rapidly indeed.

A solution $f$ of (7) exists if and only if $g \in R(K)$, the range of $K$. Since $K$ is linear, $R(K)$ is a subspace of $H_2$, however, it generally does not exhaust $H_2$, as can be seen in many of the examples of the previous section. Therefore, a traditional solution of (7) will exist only for a restricted class of functions $g$ (see the exercises). If we are willing to broaden our notion of solution, we may enlarge the class of functions $g$ for which a type of generalized solution exists to a dense subspace of functions in $H_2$. This is accomplished by introducing the idea of a least squares solution. A function $f \in H_1$ is called a least squares solution of (7) if

$$ \|Kf - g\| = \inf \{\|Ku - g\| : u \in H_1\}. $$

This is equivalent to saying that $Pg \in R(K)$, where $P$ is the orthogonal projector of $H_2$ onto $\overline{R(K)}$, the closure of the range of $K$. Now, $Pg \in R(K)$ if and only if

$$ g = Pg + (I - P)g \in R(K) + R(K)^\perp. $$

Therefore, a least squares solution exists if and only if $g$ lies in the dense subspace $R(K) + R(K)^\perp$ of $H_2$. By extending the notion of solution to the idea of least squares solution, we have guaranteed the existence of a generalized, i.e., least squares, solution of (7) for all $g$ in a dense subspace of $H_2$.

In taking up the issue of uniqueness, we note that (10) is equivalent to the condition

$$ Kf - g \in R(K)^\perp = N(K^*) $$

that is,

$$ K^*Kf = K^*g, $$

where $K^*$ is the adjoint of $K$. From (12) we see that there is a unique least squares solution if and only if

$$ \{0\} = N(K^*K) = N(K), $$

and that the set of all least squares solutions is closed and convex. Therefore, there is a unique least squares solution of smallest norm, and it is this solution that we will adopt as our generalized solution of (7). The mapping $K^\dagger$ that associates with a given

$$ g \in \mathcal{D}(K^\dagger) = R(K) + R(K)^\perp $$
the unique least squares solution having smallest norm, $K^\dagger g$, is called the Moore-Penrose generalized inverse of $K$.

In this scheme $K^\dagger$ is then the mechanism which provides a unique (least squares) solution of (7) for any $g \in \mathcal{D}(K^\dagger)$. In this sense, $K^\dagger$ settles the issues of existence and uniqueness for generalized solutions of (7). However, the big issue remains. Namely, in order for (7) to be well-posed in the sense of Hadamard for generalized solutions it is necessary that $K^\dagger$ be continuous. The next result, which summarizes the basic properties of $K^\dagger$, shows precisely when this is the case. Recall that an operator is called closed if its graph is closed in the product Hilbert space $H_1 \times H_2$.

$$K^\dagger : \mathcal{D}(K^\dagger) \to H_1 \text{ is a closed densely defined linear operator which is bounded if and only if } R(K) \text{ is closed.}$$

To see this, note first that $\mathcal{D}(K^\dagger) = R(K) + R(K)^\bot$ is evidently dense in $H_2$. The linearity of $K^\dagger$ follows easily from the definition. To see that $K^\dagger$ is closed, note that if $\{g_n\} \subseteq \mathcal{D}(K^\dagger)$, $g_n \to g$ and $K^\dagger g_n \to f$, then $f \in N(K)^\bot$, since $K^\dagger g_n \in N(K)^\bot$, and that

$$K^*g_n = K^*K K^\dagger g_n \to K^*Kf.$$

But $K^*g_n \to K^*g$ and hence $K^*Kf = K^*g$, that is, $f$ is a least squares solution lying in $N(K)^\bot$. Therefore we find $g \in \mathcal{D}(K^\dagger)$ and $K^\dagger g = f$, i.e., $K^\dagger$ is closed. Suppose now that $R(K)$ is closed, then $\mathcal{D}(K^\dagger) = H_2$ and $K^\dagger$ is a closed everywhere defined linear operator, therefore $K^\dagger$ is bounded. On the other hand, if $K^\dagger$ is bounded and $Kf_n \to g$, where $\{f_n\} \subseteq N(K)^\bot$, then

$$f_n = K^\dagger Kf_n \to K^\dagger g \text{ and } Kf_n \to KK^\dagger g.$$

Therefore $g = KK^\dagger g \in R(K)$, and $R(K)$ is closed.

The great majority of integral equations of the first kind encountered in applications have square integrable kernels and hence generate operators on $L^2$ which are compact. Solving such equations, in the generalized sense above, then involves evaluating the operator $K^\dagger$ and the solution process is stable if and only if $R(K)$ is closed. Now, it is easy to see that if $K$ is compact, then $R(K)$ is closed if and only if it is finite dimensional (see Exercise 36) and hence the only compact operators $K$ for which $K^\dagger$ is bounded are those with finite dimensional range. In the context of integral equations this says that the only Fredholm integral equations of the first kind giving rise to well-posed problems on $L^2$ are those whose kernels are degenerate.

For compact linear operators $K$ we can give a convenient representation for $K^\dagger$ in terms of the singular system $\{v_j, u_j; \mu_j\}$ discussed above. Indeed, if $g \in \mathcal{D}(K^\dagger)$, then

$$g = g_1 + g_2, \quad g_1 \in R(K) \text{ and } g_2 \in R(K)^\bot.$$  

Since $u_j \in R(K)$, we then have $\langle g, u_j \rangle = \langle g_1, u_j \rangle$ for all $j$ and hence the vector

$$f = \sum_{j=1}^{\infty} \frac{\langle g_1, u_j \rangle}{\mu_j} v_j = \sum_{j=1}^{\infty} \frac{\langle g, u_j \rangle}{\mu_j} v_j$$

exists by Picard’s criterion and satisfies $Kf = g_1$ and $f \in N(K)^\bot$. Thus $f$ is a least squares solution lying in $N(K)^\bot$, that is,

$$K^\dagger g = \sum_{j=1}^{\infty} \frac{\langle g, u_j \rangle}{\mu_j} v_j.$$  \hfill (13)
This representation of $K^\dagger g$ shows very clearly that $K^\dagger$ is unbounded if $R(K)$ is infinite dimensional. Indeed, a perturbation in $g$ of the form $\epsilon u_n$ gives a new right hand side of the form

$$g' = g + \epsilon u_n$$

satisfying $\|g' - g\| = \epsilon$. Yet the generalized solutions satisfy

$$\|K^\dagger g - K^\dagger g'\| = \epsilon \frac{\mu_n}{\mu_n} \to \infty \text{ as } n \to \infty.$$ 

This type of instability of the solution process is the hallmark of continuous inverse problems posed in infinite dimensional function spaces. It was a desire to develop methods that are capable of mitigating this instability that led Phillips and Tikhonov to launch the theory of regularization for Fredholm integral equations of the first kind in the early sixties of the twentieth century (see [41]).

4. Regularization Theory

Every restriction corresponds to a law of nature,
a regularization of the universe
Carl Sagan

The idea of the method of regularization is to replace an ill-posed Fredholm integral equation of the first kind

$$\int_0^1 k(x,t)f(t)dt = g(x)$$

by a nearby well-posed Fredholm integral equation of the second kind. We will express this equation abstractly, as we did in the previous section, as an equation of the form

$$Kf = g$$

where $K$ is a compact linear operator from a Hilbert space $H_1$ into a Hilbert space $H_2$. We have seen that generally this equation does not have a unique solution, therefore we seek a particular generalized solution, namely the least squares solution of minimum norm. That is, we assume that $g \in D(K^\dagger)$ and our aim is to approximate $K^\dagger g$. We know that, ignoring the trivial case in which the kernel $k(\cdot, \cdot)$ is degenerate, the generalized solution $K^\dagger g$ depends discontinuously on $g$, but we would like our approximations to depend continuously on $g$ and to be defined for all $g \in H_2$, as the approximations may be based on “rough” data which do not necessarily give rise to a vector in $D(K^\dagger)$. That is, our scheme involves exchanging the solution of the ill-posed problem for the solution for an approximating well-posed problem.

The generalized solution $f = K^\dagger g$ is a least squares solution and therefore it satisfies the normal equations

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

where $K^*$ is the adjoint of $K$. Now, the self-adjoint compact operator $K^*K$ has nonnegative eigenvalues and therefore, for any positive number $\alpha$, the operator $K^*K + \alpha I$, where $I$ is the identity operator on $H_1$, has strictly positive eigenvalues. In particular, the operator $K^*K + \alpha I$ has a bounded inverse, that is, the problem of solving the equation

$$(K^*K + \alpha I)f_\alpha = K^*g$$  \hspace{1cm} (14)$$

is well-posed. This second kind equation is called a regularized form of the normal equations and its unique solution

$$f_\alpha = (K^*K + \alpha I)^{-1}K^*g$$
is called the Tikhonov approximation to $K^\dagger g$, the minimum norm solution of the normal equations.

The first order of business in studying these Tikhonov approximations is to show that they converge to $K^\dagger g$ as $\alpha \to 0$. This can be accomplished conveniently in terms of a singular system $\{v_j, u_j; \mu_j\}$ for $K$. Recall that $\{v_j\}$ is a complete orthonormal set for $N(K)\perp$, $\{u_j\}$ is a complete orthonormal set for $\overline{R(K)}$, $\mu_j \to 0$, and

$$Kv_j = \mu_j u_j, \quad K^* u_j = \mu_j v_j.$$  \hspace{1cm} (15)

From this we see that $\alpha f_\alpha = K^* g - K^* K f_\alpha$ and hence

$$f_\alpha \in R(K^*) \subseteq N(K)^\perp.$$  

Therefore, we may expand $f_\alpha$ in terms of the singular vectors $\{v_j\}$:

$$f_\alpha = \sum_{j=1}^{\infty} \langle f_\alpha, v_j \rangle v_j.$$  

Similarly,

$$K^* g = \sum_{j=1}^{\infty} \langle K^* g, v_j \rangle v_j = \sum_{j=1}^{\infty} \langle g, Kv_j \rangle v_j$$

$$= \sum_{j=1}^{\infty} \mu_j \langle g, u_j \rangle v_j$$

and substituting these results into (14) we find (using (15))

$$\sum_{j=1}^{\infty} (\mu_j^2 + \alpha) \langle f_\alpha, v_j \rangle v_j = \sum_{j=1}^{\infty} \mu_j \langle g, u_j \rangle v_j.$$  

Therefore,

$$\langle f_\alpha, v_j \rangle = \frac{\mu_j}{\mu_j^2 + \alpha} \langle g, u_j \rangle$$

and hence

$$f_\alpha = \sum_{j=1}^{\infty} \frac{\mu_j}{\mu_j^2 + \alpha} \langle g, u_j \rangle v_j.$$  

The true minimum norm least squares solution is,

$$K^\dagger g = \sum_{j=1}^{\infty} \frac{1}{\mu_j} \langle g, u_j \rangle v_j.$$  

Therefore,

$$\|f_\alpha - K^\dagger g\|^2 = \sum_{j=1}^{\infty} \left( \frac{\alpha}{\mu_j (\mu_j^2 + \alpha)} \right)^2 |\langle g, u_j \rangle|^2.$$  \hspace{1cm} (16)

Now, since

$$\left( \frac{\alpha}{\mu_j (\mu_j^2 + \alpha)} \right)^2 |\langle g, u_j \rangle|^2 \leq \frac{1}{\mu_j^2} |\langle g, u_j \rangle|^2$$
and
\[
\sum_{j=1}^{\infty} \frac{1}{\mu_j^2}|(g, u_j)|^2 = \|K^\dagger g\|^2 < \infty,
\]
we may, in passing to the limit as \(\alpha \to 0\) in (16), interchange the limit and summation, giving
\[
\lim_{\alpha \to 0} \|f_\alpha - K^\dagger g\|^2 = 0.
\]
The vectors \(\{f_\alpha\}\) are therefore genuine approximations to \(K^\dagger g\) in the sense that
\[
f_\alpha \to K^\dagger g \quad \text{as} \quad \alpha \to 0.
\]
Moreover, since for each fixed \(\alpha > 0\), the operator \((K^*K + \alpha I)^{-1}K^*\) is bounded, we see that the Tikhonov approximation \(f_\alpha\) depends continuously on \(g\), for each fixed \(\alpha > 0\).

To summarize, in Tikhonov regularization, we approximate the minimum norm least squares solution \(K^\dagger g\), which depends discontinuously on \(g\), by a vector \(f_\alpha\), which is a continuous function of \(g\) for all \(g \in H_2\) and depends on a regularization parameter \(\alpha > 0\). In short, the Tikhonov regularization method approximates an ill-posed problem by a family of nearby well-posed problems.

In our models in the first section we saw that the function \(g\) in (7) is typically a measured or observed quantity and hence in practice the true \(g\) is not available to us. The best we can hope for is some estimate \(g^\delta\) of \(g\) satisfying
\[
\|g - g^\delta\| \leq \delta
\]
where \(\delta\) is a known bound on the measurement error. Instead of forming the regularized approximation with the true \(g\), we must make do with the available data \(g^\delta\) and form the regularized approximations
\[
f_\alpha^\delta = (K^*K + \alpha I)^{-1}K^*g^\delta.
\]
Now, we know that the approximations \(f_\alpha\) using “clean” data \(g\) converge to the minimum norm least squares solution \(K^\dagger g\); it is therefore reasonable to compare \(f_\alpha^\delta\) to \(f_\alpha\):
\[
f_\alpha^\delta - f_\alpha = (K^*K + \alpha I)^{-1}K^*(g^\delta - g).
\]
From this we find (since \((K^*K + \alpha I)^{-1}K^* = K^*(KK^* + \alpha I)^{-1}\))
\[
\|f_\alpha^\delta - f_\alpha\|^2 = \langle K^*(KK^* + \alpha I)^{-1}(g^\delta - g), K^*(KK^* + \alpha I)^{-1}(g^\delta - g) \rangle = \langle (KK^* + \alpha I)^{-1}(g^\delta - g), \ K^*(KK^* + \alpha I)^{-1}(g^\delta - g) \rangle.
\]
But, by the spectral mapping theorem,
\[
\|KK^*(KK^* + \alpha I)^{-1}\| \leq 1 \quad \text{and} \quad \|(KK^* + \alpha I)^{-1}\| \leq 1/\alpha,
\]
and hence
\[
\|f_\alpha^\delta - f_\alpha\| \leq \delta/\sqrt{\alpha}.
\]
This inequality represents a stability bound for the approximation \(f_\alpha^\delta\). It illustrates the classic dilemma in the analysis of ill-posed problems: for fixed \(\delta > 0\), the bound blows up as \(\alpha \to 0\), mirroring the instability in the underlying problem.

These considerations show that, for a fixed error level \(\delta\), letting the regularization parameter \(\alpha\) approach zero generally results in an unstable process. Choosing a suitable regularization parameter, based on the error in the data, then becomes the heart of the matter. Following
Tikhonov, we say that a choice $\alpha = \alpha(\delta)$ leads to a regular algorithm for the ill-posed problem (7) if
\[ \alpha(\delta) \to 0 \quad \text{and} \quad f_{\alpha(\delta)}^\delta \to K^\dagger g \quad \text{as} \quad \delta \to 0. \]

Since
\[ \| f_{\alpha(\delta)}^\delta - K^\dagger g \| \leq \| f_{\alpha(\delta)}^\delta - f_{\alpha(\delta)} \| + \| f_{\alpha(\delta)} - K^\dagger g \| \leq \delta / \sqrt{\alpha(\delta)} + \| f_{\alpha(\delta)} - K^\dagger g \| \]
and since we have shown that $f_{\alpha(\delta)} \to K^\dagger g$ as $\alpha(\delta) \to 0$, we see that the condition
\[ \delta^2 / \alpha(\delta) \to 0 \quad \text{as} \quad \delta \to 0 \]
is sufficient to ensure that Tikhonov’s method gives a regular algorithm for (7).

Tikhonov’s method also has a very important variational interpretation. Remember that the idea of the method is to approximate the generalized solution in a stable way. A reasonable way to attempt to do this is to minimize an augmented least squares functional
\[ F_\alpha(x) = \| Kx - g^\delta \|^2 + \alpha \| x \|^2. \] (18)
In this functional the first term, when small, guarantees that $x$ is “nearly” a least squares solution, while the second term tends to damp out instabilities in $x$. Now, the functional $F_\alpha$ in (18) actually achieves a minimum on $H_1$. The easiest way to see this is to note that if we define a norm $| \cdot |$ on the product Hilbert space $H_1 \times H_2$ by
\[ |\{u, v\}|^2 = \|v\|^2 + \alpha \|u\|^2, \]
then (18) measures the (squared) distance of the vector $\{0, g^\delta\} \in H_1 \times H_2$ from the graph of $K$, which is a closed convex set in $H_1 \times H_2$. Therefore, there is a vector $x \in H_1$ minimizing (18).

Any minimizer $z$ of (18) must satisfy
\[ \frac{d}{dt} \{ \| K(z + tw) - g^\delta \|^2 + \alpha \| z + tw \|^2 \} |_{t=0} = 0 \] (19)
for all $w \in H_1$. Expressing the squared norms in terms of the inner product and expanding the quadratic forms we find that this is equivalent to
\[ \langle Kz - g^\delta, Kw \rangle + \alpha \langle z, w \rangle = 0 \]
or
\[ \langle (K^*K + \alpha I)z - K^*g^\delta, w \rangle = 0 \]
for all $w \in H_1$. That is
\[ (K^*K + \alpha I)z = K^*g^\delta. \]
We therefore see that the unique minimizer of the augmented least squares functional (18) is
\[ f_{\alpha}^\delta = (K^*K + \alpha I)^{-1} K^*g^\delta \]
which is the Tikhonov approximation discussed previously. The variational characterization of the Tikhonov approximation as the minimizer of the Tikhonov functional (18) has important analytical and computational implications for Tikhonov’s method.
According to condition (17), an a priori choice \( \alpha = \alpha(\delta) \) of the regularization parameter satisfying \( \delta^2/\alpha(\delta) \to 0 \) as \( \delta \to 0 \) leads to a regular algorithm for the solution of \( Kf = g \). Although this asymptotic result may be theoretically satisfying, it would seem that a choice of the regularization parameter that is based on the actual computations performed, that is, an a posteriori choice of the regularization parameter, would be more effective in practice. One such a posteriori strategy is the discrepancy principle of Morozov. The idea of the strategy is to choose the regularization parameter so that the size of the residual \( \|KF^\delta - g^\delta\| \) is the same as the error level in the data:

\[
\|KF^\delta - g^\delta\| = \delta. \tag{20}
\]

Exercise 40 provides a motivation for this choice. According to this exercise, the vector \( f \) of minimum norm satisfying the requirement

\[
\|Kf - g^\delta\| \leq \delta
\]

also satisfies the active constraint

\[
\|Kf - g^\delta\| = \delta
\]

and hence it seems reasonable to ask the approximate solution \( f^\delta_\alpha \) to also satisfy this condition.

Assuming that the signal-to-noise ratio is larger than one, that is, \( \|g^\delta\| > \delta \), and that \( g \in R(K) \), then it is not hard to see that there is a unique positive parameter \( \alpha \) satisfying (20). To do this, we use the singular value decomposition:

\[
\|KF^\delta_\alpha - g^\delta\|^2 = \sum_{j=1}^{\infty} \left( \frac{\alpha}{\mu_j^2 + \alpha} \right)^2 |\langle g^\delta, u_j \rangle|^2 + \|Pg^\delta\|^2 \tag{21}
\]

where \( P \) is the orthogonal projector of \( H_2 \) onto \( R(K)^\perp \). From (21) we see that the real function

\[
\varphi(\alpha) = \|KF^\delta_\alpha - g^\delta\|
\]

is a continuous, increasing function of \( \alpha \) satisfying (since \( Pg = 0 \))

\[
\lim_{\alpha \to -0^+} \varphi(\alpha) = \|Pg^\delta\| = \|Pg^\delta - Pg\| \leq \|g^\delta - g\| \leq \delta
\]

and

\[
\lim_{\alpha \to \infty} \varphi(\alpha) = \|g^\delta\| > \delta.
\]

Therefore, by the intermediate value theorem, there is a unique \( \alpha = \alpha(\delta) \) satisfying (20). This choice of the regularization parameter is called the choice by the discrepancy method.

We close by showing that the choice \( \alpha(\delta) \) as given by the discrepancy method (20) leads to a regular scheme for approximating \( K^\dagger g \), that is

\[
f^\delta_{\alpha(\delta)} \to K^\dagger g \quad \text{as} \quad \delta \to 0.
\]

To do this it is sufficient to show that for any sequence \( \delta_n \to 0 \) there is a subsequence, which for notational convenience we will denote by \( \{\delta_k\} \), such that \( f^\delta_{\alpha(\delta_k)} \to K^\dagger g \). We are assuming that \( g \in R(K) \) and to simplify notation we set \( f = K^\dagger g \). Then \( f \) is the unique vector satisfying \( Kf = g \) and \( f \in N(K)^\perp \).

From the variational characterization of the Tikhonov approximation we have

\[
F_{\alpha(\delta)}(f^\delta_{\alpha(\delta)}) \leq F_{\alpha(\delta)}(f)
\]
that is, 
\[ \delta^2 + \alpha(\delta) \| f_{\alpha(\delta)}^\delta \|^2 = \| K f_{\alpha(\delta)}^\delta - g^\delta \|^2 + \alpha(\delta) \| f_{\alpha(\delta)}^\delta \|^2 \leq F_{\alpha(\delta)}(f) = \| g - g^\delta \|^2 + \alpha(\delta) \| f \|^2 \leq \delta^2 + \alpha(\delta) \| f \|^2 \]

and hence \( \| f_{\alpha(\delta)}^\delta \| \leq \| f \| \). Therefore, for any sequence \( \delta_n \to 0 \) there is a subsequence \( \delta_k \to 0 \)
with \( f_{\alpha(\delta_k)}^\delta \to z \), for some \( z \). Since
\[ f_{\alpha(\delta)}^\delta = K^* (KK^* + \alpha(\delta)I)^{-1} g^\delta \in R(K^*) \subseteq N(K)^\perp \]
and \( N(K)^\perp \) is weakly closed, we find that \( z \in N(K)^\perp \). Also, since
\[ \| K f_{\alpha(\delta_k)}^\delta - g^\delta_k \| \to 0 \]
we see that \( K f_{\alpha(\delta_k)}^\delta \to g \). But \( K \) is weakly continuous and therefore \( K f_{\alpha(\delta_k)}^\delta \to Kz \). It follows that \( Kz = g \) and \( z \in N(K)^\perp \), i.e., \( z = f \). Since \( \| f_{\alpha(\delta_k)}^\delta \| \leq \| f \| \), we then have
\[ \| f \|^2 = \lim_{k \to \infty} \langle f_{\alpha(\delta_k)}^\delta, f \rangle \leq \lim_{k \to \infty} \| f_{\alpha(\delta_k)}^\delta \| \cdot \| f \| \]
and therefore \( f_{\alpha(\delta_k)}^\delta \to f \) and \( \| f_{\alpha(\delta_k)}^\delta \| \to \| f \| \) and hence \( f_{\alpha(\delta_k)}^\delta \to f \), and the proof is complete.

5. Conclusion
The preceding discussion is the merest outline of the abstract theory of linear operator equations of the first kind and their relationship to integral equations, inverse problems and the method of regularization. We have considered Tikhonov regularization only, while there are many other methods of regularizing an ill-posed problem (see, e.g. [32]). The subject has a vast literature, a small, biased sample of which is contained in the list of references given below. Relatively elementary introductions to the field may be found in [99] and [40]; much more comprehensive accounts are available in the highly recommended references [59], [32] and [57].

6. Exercises

The wise . . . on exercise depend.
John Dryden

1. Abel’s Problem

The Abel transform is the integral transform \( A \) defined by
\[ (A\varphi)(z) = \frac{1}{\sqrt{\pi}} \int_0^z \frac{\varphi(t)}{\sqrt{z-t}} \, dt. \]
Therefore Abel’s integral equation for his mechanical problem reads: \( g = \sqrt{\frac{\pi}{2a}} Af \). Show that
\[ (A^2 \varphi)(x) = \int_0^x \varphi(t) dt. \]
It follows that if \( D \) is the differentiation operator, then \( DA^2 \varphi = \varphi \). Explain why the operator \( DA \) might then be regarded as “differentiation of one-half order.”
2. Huygens’s Tautochrone

Huygens was interested in an inverse synthesis problem in horology, namely, the design of an isochronous pendulum clock in which the period is independent of the amplitude. In terms of equation (1), he wanted to find an \( f \) for which \( g(z) = g \) is independent of \( z \). Show that if the time of descent \( g(z) \) is independent of \( z \), then \( (A^2 f)(z) = \alpha \sqrt{z} \), for some constant \( \alpha \) and hence

\[
\frac{ds}{dy} = -f(y) = a/\sqrt{y}
\]

for some constant \( a \). Show that this condition is satisfied by the cycloidal arc

\[
x = \frac{a^2}{2}(\varphi - \sin \varphi), \quad y = \frac{a^2}{2}(1 + \cos \varphi), \quad 0 \leq \varphi \leq \pi.
\]

3. Joachimstahl’s Inverse Problem

This inverse problem for the unknown law of attraction \( f(r) \) in terms of the total attractive force \( g(h) \) may be expressed as

\[
F(h) = \int_{\infty}^{h} \frac{f(r)}{\sqrt{r^2 - h^2}} dr, \quad \text{where} \quad F(h) = \frac{g(h)}{2h}.
\]

Show formally that a solution is given by

\[
f(r) = -\frac{2}{\pi} \int_{r}^{\infty} \frac{rF'(\rho)}{\sqrt{\rho^2 - r^2}} d\rho.
\]

The result

\[
\int_{h}^{\rho} \frac{r}{\sqrt{r^2 - h^2} \sqrt{\rho^2 - r^2}} dr = \frac{\pi}{2}
\]

may be useful.

4. A Transformation

Show that the integral equation in the previous problem may be reformulated as

\[
\Phi(\tau) = \int_{\tau}^{\infty} \frac{\varphi(z)}{\sqrt{z - \tau}} dz
\]

where \( \tau = h^2 \), \( \Phi(\tau) = F(\sqrt{\tau}) \) and \( \varphi(z) = \frac{1}{2} f(\sqrt{z})/\sqrt{z} \).

5. Axial Attraction

Given a positive continuous function \( r(x) \) defined for \(-1 \leq x \leq 1\), consider the gravitational attraction at position \( s > 1 \) on the \( x \)-axis induced by the surface of revolution resulting from revolving the graph of \( r \) about the \( x \)-axis. Assume the lineal mass density of the circular cross section at position \( x \) is a function \( \rho(x) \).

a) Show that the total axial component of attraction, \( g(s) \) at position \( s > 1 \) on the central axis is related to the functions \( r \) and \( \rho \) by the equations

\[
g(s) = \gamma \int_{-1}^{1} \frac{(s-x)\rho(x)r(x)}{|r(x)|^2 + (s-x)^2}^{3/2} dx
\]

for some constant \( \gamma \). Therefore, if the profile of the surface, specified by the function \( r \), is known, then the density distribution which induces a given axial attraction \( g(s) \) at position \( s \)
is the solution of a linear Fredholm integral equation of the first kind. On the other hand, if the density distribution \( \rho(x) \) is known, then the surface profile \( r \) is the solution of a nonlinear integral equation of the first kind.

b) Suppose \( r \) and \( \rho \) are both constant. Verify that \( g(s) \approx O(s^{-2}) \) as \( s \to \infty \).

c) Derive a relationship between \( g(s), r(x) \) and \( \rho(x) \) when \( s \in [-1, 1] \).

6. A Planar Gravitating Source
Consider a mass which is distributed on a plane which is parallel to the horizontal surface plane and \( z \) units below the surface plane. Let \( w(s, t) \) be the mass density (per unit area) at a point \((s, t)\) on the subsurface plane. Show that the vertical gravitational inhomogeneity at a point \((x, y)\) on the surface is

\[
\mu(x, y) = z \gamma \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} ((x-s)^2 + (y-t)^2 + z^2)^{-3/2} w(s, t) ds dt,
\]

where \( \gamma \) is a constant. Let \( h(x, y, r) \) be the average value of the mass density on a circle of radius \( r \) in the subsurface plane centered on the point \((x, y, -z)\), that is

\[
h(x, y, r) = \frac{1}{2\pi} \int_{0}^{2\pi} w(x - r \cos \theta, y - r \sin \theta) d\theta.
\]

Show that the relationship between this average \( h \) and the gravitational inhomogeneity \( \mu \) is given by the integral equation of the first kind

\[
\mu(x, y) = 2\pi \gamma z \int_{0}^{\infty} (z^2 + r^2)^{-3/2} h(x, y, r) r dr.
\]

7. A Star Shaped Gravitating Source
Consider a body with uniform density \( \rho \) which is contained within the unit circle and which is star shaped with respect to the origin. Suppose the boundary of the body is described in polar coordinates by \( r = f(\theta), 0 \leq \theta \leq 2\pi \). Show that if the centrally directed gravitational force at a point on the unit circle at polar angle \( \varphi \) is \( g(\varphi) \), then the unknown boundary profile \( f \) satisfies the nonlinear integral equation of the first kind

\[
g(\varphi) = \gamma \rho \int_{0}^{2\pi} k(\varphi, \theta, f(\theta)) d\theta
\]

where

\[
k(\varphi, \theta, z) = \int_{0}^{z} \frac{(1 - r \cos (\varphi - \theta)) r dr}{(1 + r^2 - 2r \cos (\varphi - \theta))^{3/2}}.
\]

8. Antigen Binding
Show that the change of variables \( h = e^{-s}, g(s) = \nu(e^{-s})/n, x = e^{t}, f(t) = e^{t}p(e^{t}) \) transforms the antigen binding equation into the integral equation

\[
\int_{-\infty}^{\infty} (1 + \exp(s-t))^{-1} f(t) dt = g(s).
\]

9. A General Heat Source
Consider the heat problem with sources

\[
\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f(x, t), \quad 0 < x < \pi, \quad 0 < t < \infty
\]
subject to the boundary and initial conditions
\[ u(0, t) = u(\pi, t) = 0, \quad u(x, 0) = 0. \]

We wish to reconstruct the source distribution \( f(x, t) \) from observation of the temperature history \( g(t) = u(a, t) \) at some interior point \( a \in (0, \pi) \). Show that formal separation of variables techniques lead, under suitable assumptions, to the representation
\[ g(t) = \int_0^t \int_0^\pi k(s, t-\tau)f(s, \tau)dsd\tau \]
where
\[ k(s, z) = \frac{2}{\pi} \sum_{n=1}^{\infty} e^{-n^2z} \sin na \sin ns. \]

10. IP in Steady State

Consider a bar of unit length, insulated at both ends, and radiating heat into an environment that is constantly at \( 0^\circ \) according to Newton’s law of cooling, with a distributed internal heat source \( f(x) \). The lateral surface temperature \( u \) may then be modeled by
\[ \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} - q^2 u + f(x) \quad 0 < x < 1, \]
where \( q^2 > 0 \) is a conductivity constant governing, via Newton’s law, the transfer of heat through the lateral surface. In steady state this becomes
\[ u'' - q^2 u = -f(x), \quad 0 < x < 1 \]
\[ u'(0) = u'(1) = 0. \]
Show that the inverse problem of determining the source distribution \( f(x) \) from observation of the steady state temperature distribution \( u(x) \) may be modeled by the Fredholm integral equation of the first kind
\[ u(x) = \int_0^1 k(x, s)f(s)ds \]
where
\[ k(x, s) = \begin{cases} \cosh qs \cosh q(1-x) \quad 0 \leq s \leq x \\ \frac{q \sinh q}{q \sinh q} \quad x \leq s \leq 1 \end{cases} \]

11. Thermal Time Travel

Consider a uniform bar of length \( \pi \) which is insulated on its lateral surface so that heat is constrained to flow in only one direction (the \( x \)-direction). With certain normalizations and scalings the temperature \( u(x, t) \) satisfies the partial differential equation
\[ \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad 0 < x < \pi. \]
We assume the ends of the bar are kept at temperature 0 and that the initial temperature distribution is a function \( f(x), 0 \leq x \leq \pi \), that is, the boundary and initial conditions
\[ u(0, t) = 0, \quad u(\pi, t) = 0, \quad u(x, 0) = f(x) \]
A standard direct problem in applied mathematics is to find some subsequent temperature distribution, say at $t = 1$, from this information. If we call this later temperature distribution $g(x) = u(x, 1)$, then the method of separation of variables leads to a representation of $g(x)$ in terms of the eigenfunctions $\sin nx$ of the form

$$g(x) = \sum_{n=1}^{\infty} a_n \sin nx$$

where the coefficients are given by

$$a_n = \frac{2}{\pi} \int_0^{\pi} f(s) \sin ns \ ds \ e^{-n^2}.$$  

Substituting this into the expression for $g$ and interchanging the summation and integration show that the initial temperature distribution $f$ is the solution of the integral equation of the first kind

$$g(x) = \int_0^{\pi} k(x, s)f(s)ds \quad (22)$$

where

$$k(x, s) = \frac{2}{\pi} \sum_{n=1}^{\infty} e^{-n^2} \sin nx \sin ns.$$  

Suppose $f$ and $g$ satisfy (22). Let $\epsilon > 0$ be an arbitrarily small number and and $M > 0$ be an arbitrarily large number; and let $f_M(s) = M \sin ms$. Show that the arbitrarily large perturbation $f_M$ in $f$ leads, for $m$ sufficiently large, to a perturbation of size (in $L^2$ norm) less than $\epsilon$ in $g$.

12. A Doubly Infinite Bar
Suppose a doubly infinite bar has temperature distribution $g(x) = u(x, 1)$ for $-\infty < x < \infty$ at time $t = 1$. We are interested in finding the initial temperature distribution $f(x) = u(x, 0)$. That is, we assume the model

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \quad -\infty < x < \infty, \quad t > 0.$$  

Show that the initial temperature distribution $f$ is a solution of the integral equation of the first kind

$$g(x) = \frac{1}{2\sqrt{\pi}} \int_{-\infty}^{\infty} f(s) \exp \left( -\frac{(x-s)^2}{4} \right) ds.$$  

13. Blackbody Radiation
A black body is an idealized physical object that absorbs all of the radiation falling upon it. When such a body is heated it emits thermal radiation from its surface at various frequencies. The power spectrum of the black body is the distribution of thermal power, per unit area of radiating surface, over the various frequencies. The power radiated by a unit area of surface at a given frequency $\nu$ depends on the absolute temperature $T$ of the surface and is given in appropriate units by Planck’s law:

$$P(\nu) = \frac{2\hbar^3}{c^2} \frac{1}{\exp(h\nu/kT) - 1}$$

where $c$ is the speed of light, $h$ is Planck’s constant and $k$ is Boltzmann’s constant.
Suppose that different patches of the surface of the radiating black body are at different temperatures. If $a(T)$ represents the area of the surface which is at temperature $T$, that is, $a(T)$ is the area-temperature distribution of the radiating surface, then the total radiated power at frequency $\nu$, $W(\nu)$, is given by

$$W(\nu) = \frac{2h\nu^3}{c^2} \int_0^\infty \left( \exp\left(\frac{h\nu}{kT}\right) - 1 \right)^{-1} a(T) dT.$$ 

The inverse problem of black body radiation (see [74]) is to find the area-temperature distribution $a(T)$ that can account for an observed power spectrum $W(\nu)$, that is, to solve the integral equation. Change variables by introducing $u = h/kT$ (the “coldness”) and let $w(\nu) = c^2W(\nu)/(2h\nu^3)$.

Show that the integral equation becomes

$$w(\nu) = \int_0^\infty \left( e^{nu} - 1 \right)^{-1} a(u) du.$$ 

Show formally that $w$ is the Laplace transform of

$$f(u) = \sum_{n=1}^\infty \frac{1}{n} a(u/n).$$

14. Irrigation

In traditional agriculture fields are often watered from elevated irrigation canals by removing a solid gate from a weir notch. We suppose that the depth of water in the canal is $h$ and that the notch is symmetric about a vertical center line so that the notch is represented by the region

$$\{(x, y) : -f(y) \leq x \leq f(y), 0 \leq y \leq H\}$$

where $H \geq h$ and $f$ is a given positive function specifying the shape of the symmetrical notch. By Torricelli’s law, the velocity of the effluent at height $y$ is $\sqrt{2a(h-y)}$, where $a$ is the gravitational acceleration, therefore the volume of flow per unit time through the notch is

$$2 \int_0^h \sqrt{2a(h-y)} f(y) dy$$

where $x = f(y)$ specifies the shape of the notch (see [17]). Suppose that one wishes to design a notch so that this quantity is a given function $g(h)$ of the water depth in the canal (or equivalently, suppose one wants to determine the shape $f$ from observations of the flow rate $g$). One then is led to solve the convolution equation

$$g(h) = \int_0^h 2\sqrt{2a(h-y)} f(y) dy.$$ 

(a) Show that the equation has at most one continuous solution.
(b) Solve the equation when $g(h) = h^{5/2}$.

15. Temperature Probes

Suppose a hostile environment is enclosed by a protective wall (the containment vessel of a nuclear reactor is a suitable mental image). It is desired to remotely monitor the internal temperature by passing a long (for our purposes we will assume infinitely long) bar through the wall and measuring the temperature $g$ at a point $x = a$ on the safe side of the wall.

If we denote the temperature at the point $x$ on the bar at time $t$ by $u(x, t)$, then the problem is to determine the internal temperature $f(t) = u(0, t)$ from measurements of $g(t) = u(a, t)$. 

23
Assume that the one-dimensional heat equation is satisfied, that the initial temperature of the bar is 0 and that the temperature is uniformly bounded. Then

\[ u_t = u_{xx}, \quad 0 < x < \infty. \]
\[ u(x,0) = 0 \]

Use Laplace transforms with respect to the variable \( t \) to show that \( u = \psi \ast f \) where \( \psi \) is the inverse Laplace transform of \( e^{-\sqrt{\pi}x} \), that is \[ \psi(x,t) = \frac{x}{2\sqrt{\pi}} t^{-3/2} e^{-x^2/4t}. \]

Therefore, at \( x = a \), we have an integral equation of the first kind

\[ g(t) = \frac{a}{2\sqrt{\pi}} \int_0^t \exp\left(-a^2/4(t-\tau)\right) \left(t-\tau\right)^{3/2} f(\tau) d\tau \]

relating the internal temperature \( f \) to the external temperature \( g \). Show that if \( u(x,t) = \sqrt{\frac{2}{n}} \sin \left(nt + \sqrt{\frac{2}{n}} x\right) \exp \sqrt{\frac{2}{n}} x \) then \( u_t = u_{xx}, u(0,t) = f(t), \) where \( f(t) = \sqrt{\frac{2}{n}} \sin nt \) and \( u_x(0,t) = h(t) \), where \( h(t) = \cos nt + \sin nt \). Show that if \( \epsilon > 0 \) (arbitrarily small) and \( M > 0 \) (arbitrarily large), then for any fixed \( a > 0 \) there are functions \( f \) and \( h \) satisfying the above conditions with \( \max |f(t)| < \epsilon \), \( \max |h(t)| \leq 2 \) and \( \max |u(a,t)| > M \).

16. Poisson’s Formula for the Half Plane

Poisson’s formula for the half plane states that if \( \phi(x,y) \) is a harmonic function in the upper half plane \( (y > 0) \), and if \( \phi(x,y) \to f(x) \) as \( y \to 0^+ \), then \( f \) is a solution of the integral equation of the first kind:

\[ \phi(x,y) = \frac{y}{\pi} \int_{-\infty}^{\infty} \frac{f(s)}{[(x-s)^2 + y^2]^{1/2}} ds. \]

Give an intuitive argument in favor of this result by showing that the function

\[ \Phi(s,x,y) = \frac{y}{\pi[(x-s)^2 + y^2]} \]

satisfies the following conditions:

\[ \frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} = 0, \]
\[ \int_{-\infty}^{\infty} \Phi(s,x,y) ds = 1, \]

and

\[ \lim_{y \to 0^+} \Phi(s,x,y) = \begin{cases} 0, & s \neq x \\ \infty, & s = x \end{cases} \]

17. Nonuniqueness in Potential Theory

Suppose \( \Omega \) is a closed convex subset of \( \mathbb{R}^3 \) with smooth boundary \( \partial \Omega \). The gravitational potential \( U(P) \) at a point \( P \notin \Omega \) is related to the distribution of mass density \( \rho(Q) \) for \( Q \in \Omega \) by

\[ U(P) = \gamma \int_{\Omega} \rho(Q) \frac{1}{|P-Q|} dV(Q) \]
where $\gamma$ is a positive constant. In other words, an interior mass distribution $\rho$ that gives rise to an exterior gravitational potential field $U$ is a solution of a Fredholm integral equation of the first kind. Show that if $R \in C^2(\Omega)$ is an arbitrary function satisfying

$$R|_{\partial\Omega} = 0 \quad \text{and} \quad \frac{\partial R}{\partial n}|_{\partial\Omega} = 0$$

where $\frac{\partial}{\partial n}$ is the normal derivative, then

$$\int_{\Omega} \Delta R(Q) \frac{1}{|P - Q|} dV(Q) = 0$$

(Hint: Use Green’s second identity). Therefore, the external potential does not uniquely determine the internal mass distribution.

18. Fujita’s Equation

a) Show that the function $g(r)$ specifying the concentration profile in Fujita’s equation is the Laplace transform of the function

$$\varphi(x) = \frac{g_0 x}{\lambda(1 - \exp(-x))} f\left(\frac{x}{\lambda}\right).$$

b) Conclude, by the Laplace transform inversion formula, that the function $f(w)$ has the form

$$f(w) = \frac{1 - \exp(-\lambda)}{2\pi wg_0 i} \int_{a-i\infty}^{a+i\infty} g(r) \exp(\lambda wr) dr,$$

for a suitable real number $a$. Is this of any practical use?

c) Suppose that $g_0 = 1$, and

$$g_1(r) = \int_0^\infty \frac{\lambda w}{1 - \exp(-\lambda w)} \exp(-\lambda rw)f_1(w) dw$$

where $f_1(w)$ is a function satisfying

$$f_1(w) \geq 0 \quad \text{and} \quad \int_0^\infty f_1(w) dw = 1.$$

Now let

$$f_n(w) = f_1(w)(1 + \sin n\pi \lambda w)/a_n, \quad n = 1, 2, \ldots$$

where

$$a_n = \int_0^\infty f_1(w)(1 + \sin n\pi \lambda w) dw.$$

Show that

$$g_n(r) = \int_0^\infty \frac{\lambda w}{1 - \exp(-\lambda w)} \exp(-\lambda rw)f_n(w) dw \to g_1(r) \quad \text{as} \quad n \to \infty$$

for each $r$, while

$$\frac{f_{4n+1}(\frac{1}{2\lambda})}{f_1(\frac{1}{2\lambda})} \to 2 \quad \text{as} \quad n \to \infty.$$
19. Instability
Consider the equation \( \int_0^s f(t)dt = g(s) \), \( s \in [0,1] \). Assuming that \( g \) is absolutely continuous and \( g(0) = 0 \), the unique solution is \( f(t) = g'(t) \). In particular, \( f = 0 \) if \( g = 0 \). For a given \( \epsilon > 0 \), let \( g_\epsilon(s) = \epsilon \sin(s/\epsilon^2) \). Then \( |g_\epsilon(s)| \leq \epsilon \). Show that the solution \( f_\epsilon(t) \) satisfies \( \max_t |f_\epsilon(t)| = 1/\epsilon \).

20. Kantorovich’s Regularization
Show that the Volterra integral equations of the first kind
\[
\int_0^s k(s,t)f(t)dt = g(s)
\] (23)
may be converted to the Volterra integral equation of the second kind
\[
f(s) + \int_0^s \left( \frac{\partial k}{\partial t}(s,t)/k(s,s) \right) f(t)dt = g'(s)/k(s,s).
\] (24)
if \( k(s,t) \) and \( \frac{\partial k}{\partial t}(s,t) \) are continuous for \( 0 \leq t \leq s \leq 1 \), \( g'(s) \) is continuous for \( 0 \leq s \leq 1 \) and \( k(s,s) \neq 0 \) for \( 0 \leq s \leq 1 \).

21. Another Regularization
Let \( \varphi(s) = \int_0^s f(\tau)d\tau \). Apply integration by parts to (23) to obtain the Volterra integral equation of the second kind
\[
\varphi(s) - \int_0^s \left( \frac{\partial k}{\partial t}(s,t)/k(s,s) \right) \varphi(t)dt = g(s)/k(s,s).
\]
Does this circumvent the instability problem in (23)?

22. A Closed Form Solution
Show that the Fredholm integral equation of the first kind
\[
\int_0^1 |x-t|f(t)dt = g(x), \quad 0 \leq x \leq 1
\]
has a solution \( f \in C[0,1] \) if and only if \( g \in C^2[0,1] \) and \( g'(1) = -g'(0) = g(0) + g(1) \). Find the solution.

23. Another Transformation
Suppose \( \alpha > 0 \) and \( g \) is differentiable. Use the transformation \( \xi = xt \) to find a solution of the integral equation
\[
\int_0^1 t^\alpha f(xt)dt = g(x).
\]

24. No Solution
Show that
\[
\int_0^1 e^{st}f(t)dt = g(s), \quad 0 \leq s \leq 1
\]
has no bounded integrable solution if \( g(s) = |s - 1/2| \), \( 0 \leq s \leq 1 \).

25. Multiple Solutions
Suppose that for \( 0 \leq t \leq 1 \), \( k(s,t) = 0 \) for \( 0 \leq s < 1/2 \) and \( k(s,t) = 1 \) for \( 1/2 \leq s \leq 1 \). Show that \( f(t) = 0 \) and \( f(t) = t - 1/2 \) are both solutions of \( \int_0^1 k(s,t)f(t)dt = 0 \), \( 0 \leq s \leq 1 \).
26. Infinitely Many Solutions
Show that for every real number \( c \), \( f(t) = ct^2 \) is a solution of
\[
\int_0^s (3s - 4t) f(t) dt = 0.
\]

27. Unbounded Solutions
Let \( n \) be a positive integer and \( g_n(s) = \pi \epsilon \sin ns \). Note that \( \{g_n\} \) is uniformly bounded for all \( n \). Show that \( f_n(t) = \epsilon e^n \sin nt \) is a solution of
\[
\int_{-\infty}^{\infty} \frac{1}{1 + (s-t)^2} f_n(t) dt = g_n(s),
\]
yet the sequence of functions \( \{f_n\} \) is unbounded.

28. Tikhonov’s Lemma
Prove Tikhonov’s lemma [91]: If \( T \) is a continuous, one-to-one, surjective mapping \( T : X \to Y \), and \( X \) is compact, then \( T^{-1} \) is continuous.

29. Hanging Cables
Tikhonov’s lemma may be applied to the linearized hanging cable model (see [40]). In this model the vertical deflection of the cable \( g(x) \) is related to its lineal density distribution \( f(t) \) by
\[
\int_0^1 k(x, t) f(t) dt = g(x), \quad 0 \leq x \leq 1
\]
where
\[
k(x, t) = \begin{cases} 
  x(1-t)/T, & 0 \leq s \leq t \\
  t(1-x)/T, & t \leq s \leq 1
\end{cases}
\]
\( T \) is the constant tension in the cable, \( f(t) \) represents the unknown density distribution of the hanging cable and \( g(x) \) is the observed sag of the cable at position \( x \). Consider the operator as acting on the space \( C[0, 1] \), and let \( X \) be the class of densities \( f \) with bounded derivative satisfying
\[
\|f\|_2^2 := \|f\|_\infty^2 + \|f'\|_\infty^2 \leq C.
\]
Show that this is a compact set of densities in \( C[0, 1] \) and that the operator is injective. Conclude that the inverse of the operator \( K \) restricted to this class is continuous, that is, the inverse problem is stable for this class of densities.

30. Optimization
Prove that a closed convex subset of a Hilbert space contains a unique element of smallest norm.

31. Quasi-Solutions
Tikhonov’s idea of restricting allowable solutions of inverse problems to lie in a compact set, along with the fact that closed balls in Hilbert space are weakly compact, suggests the notion of a quasi-solution. Suppose that \( K \) is a compact linear operator with trivial nullspace acting on a Hilbert space. Given \( r > 0 \), a vector \( f_r \) is called an \( r \)-quasi-solution of the equation \( Kf = g \) if:
\[
\|f_r\| \leq r
\]
and
\[
\|Kf - g\| \leq \|K\varphi - g\| \quad \text{for all} \quad \varphi \quad \text{with} \quad \|\varphi\| \leq r.
\]
a) Prove that for each \( r > 0 \) there exits a unique \( r \)-quasi-solution. (Hint: Use the fact that closed convex sets in Hilbert space have the “unique closest point” property; see the previous exercise.)

b) Prove that \( r \)-quasi-solutions are weakly stable with respect to perturbations in \( g \), i.e., if \( \|g - g_n\| \to 0 \) and if \( f_r \) is the \( r \)-quasi-solution for \( Kf = g \) and \( f_{r,n} \) is the \( r \)-quasi-solution for \( Kf = g_n \), then \( f_{r,n} \to f_r \) (weak convergence) as \( n \to \infty \).

32. Least Squares Solutions
Suppose \( g \in D(K^\dagger) \). Show that \( K^\dagger g \) is the unique least squares solution in \( N(K)^\perp \) and that the set of all least squares solutions may be represented as \( K^\dagger g + N(K) \). Also show that if \( K \) represents the operator \( K \) restricted to \( N(K)^\perp \), then for any \( g \in D(K^\dagger) \), \( K^\dagger g = K^{-1}Pg \), where \( P \) is the orthogonal projector of \( H_2 \) onto \( R(K) \).

33. A Moore-Penrose Inverse
Let \( H_1 = H_2 = L^2[0, 1] \) and define \( K : H_1 \to H_2 \) by

\[
(Kf)(s) = \int_0^s f(t)dt.
\]

Show that

\[
R(K) = \{g \in L^2[0, 1] : g \text{ is abs. cont., } g' \in L^2[0, 1] \text{ and } g(0) = 0\},
\]

and that \( K^\dagger g = g' \) if \( g \in R(K) \).

34. Another Moore-Penrose Inverse
The definition of \( K^\dagger \) given above for a bounded linear operator \( K \) extends naturally to the case when \( K \) is a closed densely defined linear operator. Provide the details. Let \( H_1 = H_2 = L^2[0, 1] \) and let

\[
D(T) = \{f \in H_1 : f \text{ is absolutely cont. } f' \in H_1, f(0) = f(1) = 0\},
\]

and define \( T : D(T) \to H_2 \) by \( Tf = f' \). Show that \( D(T^\dagger) = H_2 \) and that

\[
(T^\dagger g)(t) = \int_0^t g(s)ds - t \int_0^1 g(s)ds.
\]

35. Generalized Picard Criterion
Suppose \( K \) is a compact linear operator with singular system \( \{v_j, u_j; \mu_j\} \) and \( \nu \geq 0 \). Prove that \( g \in R(K^*(K^*K)^\nu) \) if and only if \( g \in N(K^*K)^\perp \) and

\[
\sum_{j=1}^{\infty} \mu_j^{-\nu-2} |(g, u_j)|^2 < \infty.
\]

36. Compact Operators With Closed Range
If a bounded linear operator \( K \) has closed range, then \( \|Kx\| \geq m\|x\| \) for some \( m > 0 \) and all \( x \in N(K)^\perp \). From this conclude that a compact linear operator has closed range if and only if its range is finite-dimensional.

37. A Regularization
Let \( H_1 = H_2 = L^2[0, 1] \). Define \( K : H_1 \to H_2 \) by

\[
(Kf)(s) = \int_0^s f(t)dt.
\]
Suppose that \( g \) is absolutely continuous and \( g(0) = 0 \). Show that the Tikhonov approximation \( f_\alpha \) for the problem \( Kf = g \) is a solution of the boundary value problem
\[
\alpha f''_\alpha(t) - f'_\alpha(t) = -g'(t), \quad f_\alpha(1) = f'_\alpha(0) = 0.
\]
Solve this for \( f_\alpha \) and show that the solution depends \( L^2 \)--continuously on \( g \).

38. A Convergence Rate
Suppose that \( Kf = g \) where \( f = K^*Kw \), for some \( w \). Show that \( \|f_\alpha - f\| = O(\alpha) \), where \( f_\alpha \) is the Tikhonov approximation
\[
f_\alpha = (K^*K + \alpha I)^{-1}K^*g.
\]

39. Another Convergence Rate
Show that if \( g \in R(KK^*K) \) and \( \alpha = C\delta^{2/3} \) then \( \|f^*_\alpha - K^*g\| = O(\delta^{2/3}) \) for all \( g^\delta \) satisfying \( \|g - g^\delta\| \leq \delta \).

40. Optimization Again
Show that if \( \|f\| \) is a minimum subject to the constraint \( \|Kf - g^\delta\| \leq \delta \), then \( \|Kf - g^\delta\| = \delta \).

Acknowledgments
The author is grateful to two anonymous referees for several helpful comments, suggestions and corrections.

References
[1] Abel N 1826 Auflö sung einer mechanische Aufgabe, Journal für die reine und angewandte Mathematik 1 153-157
[2] Abramowitz M and Stegun I 1964 Handbook of Mathematical Functions, National Bureau of Standards Applied Mathematics Series vol 55 (Washington, D.C.: U.S. Government Printing Office)
[3] Alifanov O 1994 Inverse Heat Transfer Problems (Berlin: Springer-Verlag)
[4] Allison H 1979 Inverse unstable problems and some of their applications, The Mathematical Scientist 4 9-30
[5] Anderssen R and Saull V 1973 Surface temperature history determination from borehole measurements, Mathematical Geology 5 269-283
[6] Anderssen R, deHoog F and Lukas M (Eds.) 1997 An Introduction to Inverse Scattering and Inverse Spectral Problems (Braunschweig: Friedrich Vieweg und Sohn)
[7] Bakushinsky A and Kokurin M 2004 Iterative Methods for Approximate Solution of Inverse Problems (New York: Springer)
[8] Baumeister J 1987 Stable Solution of Inverse Problems (Braunschweig: Friedrich Vieweg und Sohn)
[9] Bennett A 2002 Inverse Modeling of the Ocean and Atmosphere (Cambridge: Cambridge University Press)
[10] Bernkopf M 1966 The development of function spaces with particular reference to their origins in integral equation theory, Archive for History of Exact Sciences 3 1-96
[11] Bertero M 1989 Linear inverse and ill-posed problems, Advances in Electronics and Electron Physics 75 2-120
[12] Birch F 1948 The effect of Pleistocene climatic variation upon geothermal gradients, American Journal of Science 246 729-760
[13] Bitsadze A 1995 Integral Equations of the First Kind (Singapore: World Scientific)
[14] Böcher M 1908 An Introduction to the Theory of Inverse Equations, (Cambridge: Cambridge University Press)
[15] du Bois-Reymond P 1888 Bemerkungen über \( \Delta z = \frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} = 0 \), Journal für Mathematik 103 204-229
[16] Bölt B 1980 What can inverse theory do for applied mathematics and the sciences?, Australian Mathematical Society Gazette 7 69-78.
[17] Brenke W 1922 An application of Abel’s integral equation, American Mathematical Monthly 29 58-60
[18] Bukhemi A 2000 Introduction to the Theory of Inverse Problems (Utrecht: VSP)
[19] Chandan K, Colton D, et al (Eds.) 1997 An Introduction to Inverse Scattering and Inverse Spectral Problems (Philadelphia: SIAM)
[20] Cheney M 1994 An introduction to inverse problems, pp. 21-32 in “Mathematics of Computation 1943-1993: A Half-Century of Computational Mathematics”, W. Gautschi, Ed., Proceedings of Symposia in Applied Mathematics (Providence: American Mathematical Society)
[21] Cheney M 1995 *Inverse boundary-value problems*, American Scientist 85 448-455.
[22] Cheng J (Translator) 2006 *Inverse Problems: Activities for Undergraduates*, by C. Groetsch, Manadrin Chinese translation (Beijing: Tsinghua University Press and Springer).
[23] Coleman R 1989 *Inverse problems*, Journal of Microscopy 153 233-248.
[24] Colton D, Engl H et al (Eds.) 2000 *Surveys on Solution Methods for Inverse Problems* (Vienna: Springer-Verlag).
[25] Colton D and Kress R 1992 *Inverse Acoustic and Electromagnetic Scattering Theory* (New York: Springer-Verlag).
[26] Cakoni F and Colton D 2005 *Qualitative Methods in Inverse Scattering Theory* (New York: Springer).
[27] Craig I and Brown J 1986 *Inverse Problems in Astronomy*, (Bristol: Adam Hilger).
[28] Daubechies I, Defrise M and De Mol C 2004 *An iterative thresholding algorithm for linear inverse problems with a sparsity constraint*, Communications on Pure and Applied Mathematics 44 1413-1457.
[29] Douglas Jr. J 1960 *Mathematical programming and integral equations*, in Symposium on the Numerical Treatment of Ordinary Differential Equations, Integral and Integro-differential Equations, (Rome, 1960) (Basel: Birkhäuser) 209-274.
[30] Engl H 1995 *Inverse Problems* (Mexico City: Sociedad Matematica Mexicana).
[31] Engl H and Groetsch C (Eds.) 1987 *Inverse and Ill-posed Problems* (Orlando: Academic Press).
[32] Engl H, Hanke M and Neubauer A 1996 *Regularization of Inverse Problems* (Dordrecht: Kluwer).
[33] Fox L and Goodwin E 1953 *The numerical solution of non-singular linear integral equations*, Philosophical Transactions of the Royal Society of London (Series A) 245 501-534.
[34] French D et al 2006 *Numerical approximation of solutions of a nonlinear inverse problem arising in olfaction experimentation*, Mathematical and Computer Modelling 43 945-956.
[35] Fridman V 1953 *Method of successive approximations for Fredholm integral equations of the first kind*, Uspekhi Matematicheskikh Nauk 31 233-234 (In Russian).
[36] Fujita H 1962 *Mathematical Theory of Sedimentation Analysis* (New York: Academic Press).
[37] Gladwell G 1986 *Inverse Problems in Vibration* (Dordrecht: Martinus Nijhoff).
[38] Glasko V 1984 *Inverse Problems of Mathematical Physics* (New York: American Institute of Physics).
[39] Groetsch C 1984 *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind* (London: Pitman).
[40] Groetsch C 1993 *Inverse Problems in the Mathematical Sciences* (Braunschweig: Friedrich Vieweg und Sohn).
[41] Groetsch C 2003 *The delayed emergence of regularization theory*, Bollettino di Storia delle Scienze Matematiche 23 105-120.
[42] Grosser M 1962 *The Discovery of Neptune* (Cambridge, MA: Harvard University Press).
[43] Hadamard J 1902 *Sur les problèmes aux dérivées partielles et leur signification physique*, Princeton University Bulletin 13 49-52.
[44] Hadamard J 1948 *Sur le cas anormal du problème de Cauchy pour l’équation des ondes*, in Studies and Essays Presented to R. Courant on his 60th Birthday, January 8, 1948 (New York: Interscience Publishers).
[45] Hanke M and Hansen P-C 1993 *Regularization methods for large scale inverse problems*, Surveys of Mathematics in Industry 3 253-315.
[46] Hansen P-C 1997 *Rank Deficient and Discrete Ill-posed Problems* (Philadelphia: SIAM).
[47] Hanson R 1972 *Integral equations of immunology*, Communications of the Association for Computing Machinery 15 883-890.
[48] Hensel E 1991 *Inverse Theory and Applications for Engineers* (Englewood Cliffs, NJ: Prentice-Hall).
[49] Hofmann B 1986 *Regularization for Applied Inverse and Ill-posed Problems* (Leipzig: Teubner).
[50] Isakov V 1990 *Inverse Source Problems* (Providence: American Mathematical Society).
[51] Ivanov V 1962 *Integral equations of the first kind and an approximate solution for the inverse problem of potential*, Soviet Mathematics Doklady 3 210-212.
[52] Jaswon M and Symm G 1977 *Integral Equation Methods in Potential Theory and Electrostatics* (London: Academic Press).
[53] Joachimstahl J 1861 *Über ein Attractionsproblem*, Journal fuer die reine und angewandte Mathematik 58 135-137.
[54] Kaltenbacher B 2000 *Regularization by projection with a posteriori discretization level choice for linear and nonlinear ill-posed problems*, Inverse Problems 16 1523-1539.
[55] Kaneko A (Translator) 2002 *Inverse Problems: Activities for Undergraduates*, by C. Groetsch, (Japanese translation) (Tokyo: Saiensu-sha).
[56] Keller J 1976 *Inverse problems*, American Mathematical Monthly 83 107-118.
[57] Kirsch A 1996 *An Introduction to the Mathematical Theory of Inverse Problems* (New York: Springer).
[58] Kond J 1991 *Integral Equations* (Tokyo: Kodansha and Oxford: Clarendon Press).
[59] Kress R 1989 *Linear Integral Equations* (New York: Springer).
60. Kubo S (Ed.) 1992 Inverse Problems (Atlanta: Atlanta Technical Publications)
61. Lavrentiev M, Romanov V and Shishatskii S 1986 Ill-posed Problems of Mathematical Physics and Analysis, American Mathematical Society Translations of Mathematical Monographs, vol. 64, (Providence: American Mathematical Society)
62. Lee D 1970 On the determination of molecular weight distributions from sedimentation-diffusion equilibrium data at a single rotor speed, Journal of Polymer Science (A-2) 8 1039-1056
63. Louis A 1989 Inverse und schlecht gestellte Probleme (Stuttgart: Teubner)
64. McIver M 1991 An inverse problem in electromagnetic crack detection, IMA Journal on Applied Mathematics 47 127-145
65. Menke W 1984 Geophysical Data Analysis: Discrete Inverse Theory (Orlando: Academic Press)
66. Miyamoto S, Ikeda S, and Sawaragi Y 1978 Identification of distributed systems and the theory of regularization, Journal of Mathematical Analysis and Applications 63 77-95
67. Morozov V 1984 Methods for Solving Incorrectly Posed Problems (New York: Springer)
68. Nashed M (Ed.) 1976 Generalized Inverses and Applications (New York: Academic Press)
69. Nashed N and Scherzer O (Eds.) 2002 Inverse Problems, Image Analysis and Medical Imaging, Contemporary Mathematics (vol. 313) (Eds.: Providence: American Mathematical Society)
70. Natterer F 1986 The Mathematics of Computerized Tomography (New York: John Wiley and Sons)
71. Neittaanmäki P, Rudnicki M and A. Savini A 1996 Inverse Problems and Optimal Design in Electricity and Magnetism (Oxford: Clarendon Press)
72. Neumaier A 1998 Solving ill-conditioned and singular linear systems: a tutorial on regularization, SIAM Review 40 636-666
73. Onishi K, Tanuma K and Yamamoto M (Translators) 1996 Inverse Problems in the Mathematical Sciences, by C. Groetsch (Japanese translation) (Tokyo: Saiensusha)
74. Paley R and Wiener N 1934 The Fourier Transform in the Complex Domain (New York: American Mathematical Society)
75. Parker R 1994 Geophysical Inverse Theory (Princeton: Princeton University Press)
76. Petrov Yu and Sizikov V 2005 Well-Posed, Ill-Posed, and Intermediate Problems with Applications (Leiden, The Netherlands: Koninklijke Brill NV)
77. Phillips D 1962 A technique for the numerical solution of certain integral equations of the first kind, Journal of the Association for Computing Machinery 9 84-97
78. Picard E 1910 Sur un théorème général relatif aux équations intégrales de première espèce et sur quelques problèmes de physique mathématique, Rendiconti del Ciclolo Matematico di Palermo 29 79-97
79. Pöschel J and Trubowitz E 1986 Inverse Spectral Theory (New York: Academic Press)
80. Polyak R and Manzhirov A 1998 Handbook of Integral Equations (Boca Raton: CRC Press)
81. Ramm A 2005 Inverse Problems (New York: Springer)
82. Sabatier J 1988 Some Topics on Inverse Problems (Singapore: World Scientific)
83. Saitoh S 1997 Integral Transforms, Reproducing Kernels and Their Applications (London: Addison-Wesley, Longman, Ltd.)
84. Saitoh S 2005 Best approximation, Tikhonov regularization and reproducing kernels, Kodai Mathematics Journal 28 359-367
85. Schmidt E 1996 Zur Theorie der linearen und nichtlinearen Integralgleichungen I. Teil: Entwicklung willkürlicher Funktionen nach Systemen vorgeschriebener, Mathematische Annalen 63 443-476
86. Shechan W, Kollerstrom H and Waff C 2004 The case of the pilfered planet, Scientific American (December) 90-99
87. Standage T 2000 The Neptune File (New York: Walker)
88. Stein S 1996 Inverse problems for central forces, Mathematics Magazine 69 83-93
89. Tari G (Ed.) 1986 Inverse Problems, LNM 1225, (New York: Springer)
90. Tarantola A 1987 Inverse Problem Theory (Amsterdam: Elsevier)
91. Tikhonov A 1943 On the stability of inverse problems, Doklady Akademii Nauk SSSR 39 176-179
92. Tikhonov (Tikhonov) A 1963 Solution of incorrectly formulated problems and the regularization method, Soviet Mathematics Doklady 4 1035-1038
93. Tikhonov A and Arsenin V 1977 Solutions of Ill-posed Problems (Washington: Winston and Sons)
94. Wahlberg G (Ed.) 2003 Inside Out: Inverse Problems and Applications (New York: Cambridge University Press)
95. Vogel C 1987 An overview of numerical methods for nonlinear ill-posed problems, in [31], pp. 231-245.
96. Vogel C 2002 Computational Methods for Inverse Problems (Philadelphia: SIAM)
97. Wahba G 1977 Practical approximate solutions to linear operator equations when the data are noisy, SIAM Journal on Numerical Analysis 14 651-667
98. Wicksell S 1925 The corpuscle problem, Biometrika 17 84-99
[99] Wing G 1992 *A Primer on Integral Equations of the First Kind: The Problem of Deconvolution and Unfolding* (Philadelphia: SIAM)

[100] Wunsch C 1996 *The Ocean Circulation Inverse Problem* (Cambridge: Cambridge University Press)

[101] Yamaguti M, et al (Eds.) 1991 *Inverse Problems in Engineering Sciences* (Tokyo: Springer-Verlag)

[102] Zhdanov M 2002 *Geophysical Inverse Theory and Regularization* (Amsterdam: Elsevier)