Control
and Cybernetics
VOL. 10 (1981) No. 1–2

Convergent approximation methods for ill-posed problems. Part I — General Theory

by

THOMAS I. SEIDMAN*

University of Maryland,
Baltimore County, USA and
Université de Nice, France

A comprehensive study of approximation algorithms for solving various ill-posed problems is presented in the paper. Results concerning convergence of the proposed methods are given with proofs.

Part I offers an exposition of the general framework. The main approximation results and a discussion of approaches most frequently used for solving the problems under consideration (penalty functions and minimization, regularization) are presented. In particular algorithms for ill-posed problems with only discrete knowledge of data resulting from measurements, using so called generalized interpolation, are discussed.

Part II is devoted to some applications. Approximation algorithms for equations involving Nemytsky operators are discussed. Results concerning constructive approaches to solving inverse problems for parabolic equations (recovery of a diffused signal) as well as to solving more general identification problems for distributed systems are presented.

1. Introduction

In his discussion of partial differential equations [27], Hadamard called a problem

$$f(x)=b \quad where \quad f\colon X\supset D\to Y,\ b\in Y \tag{1.1}$$

well-posed if, for each b in Y, there would exist a unique solution x_* in D for (1.1) depending continuously on b in a suitable domain $\hat{D}\subset Y$ — i.e., if f were to have a continuous inverse $g\colon Y\supset\hat{D}\to X$. This may be generalized (this, essentially, is the context of the theory of pseudoinverses and normal solvability; see, e.g., [49], [53]) to let $g\colon R(f)\to X$ be a suitable continuous right inverse of f, permitting consideration of possible nonsurjectivity of f (but with the range of f closed in Y) and nonuniqueness (e.g., adjoining an auxiliary criterion to select a unique solution: say, selecting the one of minimum X—norm). Conversely, we now wish to consider *ill-posed* problems, meaning problems for which there can be *no continuous right inverse* g; the usual symptom of this is that the range of f is not closed in Y.

* This research was partly supported by the U.S.Army Research Office undei ARO Grant DAAG-29-77-G-0061

In practice, problems of the form (1.1) arise in circumstances for which the data (b and the map f) are available only as *approximations* — through calculation, theoretical analysis or measurement. If one considers the approximating problems

$$f_k(x)=b_k, \tag{1.2}$$

then the "solutions" $x_k:=g_k(b_k)$ provide a sequence of approximants to $x_*=g(b)$. It is easily seen that, for well-posed problems, the properties of *stability* (g_k equicontinuous) and *consistency* $\big(f(g_k(b))\to b\big)$ imply *convergence* $\big(b_k\to b$ implies $x_k=$ $=g_k(b_k)\to x_*\big)$. Conversely, one expects (e.g., note the Lax Equivalence Theorem [43]) that, for consistent approximation schemes $\{g_k\}$, convergence is *equivalent* to stability and implies continuity of g (well-posedness).

The prevailing attitude following [27], was that ill-posed problems did not occur for properly modeled questions of physical interest — fortunately, as the inevitable imprecisions of formulation, measurement and finitary computation would preclude any possibility of useful inference otherwise. It happens, however, that numerous ill-posed problems are, indeed, of practical interest and, fortunately, that useful computation is, in fact, possible in that case — through requiring more than the usual care in formulation and execution: (1.2) cannot be safely used as an approximation scheme nor (see Theorem 7.1 below) can the standard projection method of minimizing $\|f_k(x)-b_k\|$ over a sequence of subspaces $\{X_k\}$. The most widely (and most successfully) used computational approach to the solution of ill-posed problems has been the method of *regularization*, due originally to Tikhonov ([69], etc.). There is a considerable and growing body of research (largely in the Russian literature; in particular see [47]) into the theoretical properties and various applications of (variants of) this method.

A number of classes of ill-posed problems have been extensively studied, including:

(1.3)

(a) determination of Sturm-Liouville operators from spectral data; see, e.g., [35] and the deep results of [23],

(b) inverse problems (determination of functions, such as coefficients, defining a differential equation from various observations of solutions); see, e.g., Section 7, below, and further references there,

(c) nonstandard Cauchy problems (as, for example, the backwards heat equation); see, e.g., [33], [41], [38] and Section 6 below,

(d) integral geometry (determining a function from integral averages over a suitable family of sets); see, e.g., [34] [42],

(e) integral equations of the first kind (including numerical inversion of various integral transforms); see, e.g., [56],

(f) continuation (analytical continuation and "field problems", estimating, say, a pressure field in hydrology or an electrostatic field globally from local observations),

etc.

For most of these problems the principal thrusts of the body of theoretical analysis have been toward establishing uniqueness and toward demonstrating a form of well-posedness in a restricted context, showing that the solution depended continuously on b if it were *a priori* known to lie in a sufficiently restricted set (say, by the assumption of bounds for some suitable norm — see, e.g., [33], [55] and note also, [5], Theorem 1 of [61]; compare Theorem 2.5, below). Of these classes (1.3), the class of inverse problems (1.3.b), in particular, has been the subject of extensive investigation (we specifically cite the considerable work by J. R. Cannon and by G. Chavent). It is worth noting that much of the investigation in this area, especially that which has appeared in the system theoretic and engineering literature, suffers from a profound defect (compare Theorem 7.1, below): it investigates the theoretical properties and the feasibility and convergence of computational schemes as applied to the problem *after* expressing the unknown function in terms of (an approximating) parametric representation without considering the sensitivity of the computation to the effect of approximation at *this* stage of the total formulation.

There are two principal viewpoints from which to consider the computational aspects of (1.1). First one may assume that one has given a specific (perhaps approximate) form $\tilde{f}$ of the map $f: X \supset D \to Y$ and a specific observation $\tilde{b}$ approximating b together with *a priori* information as to a subset $\tilde{D} \subset D$ within which to seek the solution and estimates of the "goodness of approximation" of $\tilde{f}$ to f and of $\tilde{b}$ to b (these estimates may be explicit bounds or may be parameters of a statistical error distribution). One then seeks to obtain (by some feasible computational procedure) the best possible (or maximum likelihood) "estimate" $\tilde{x}$ on the basis of the available data and also a bound (or confidence regions) for the "goodness of approximation" of the computed $\tilde{x}$ to the desired true solution x_*. This may also be reversed to determine the accuracy and "complexity" of measurement and computation required in order to attain a specified goodness of approximation of $\tilde{x}$ to x_*. Second, one may take a formulation such as the above and embed it as one of a *sequence* of increasingly more accurate (and more complex) approximations defining a *computation scheme*. One then asks whether the sequence of computed approximations $\{x_k\}$ converges to x_* and seeks to obtain also a *rate of convergence* (asymptotic error bounds). The second of these (i.e., the asymptotic analysis) is the viewpoint adopted in this paper. The focus of our present concern is with the convergence to x_* of the computed sequence $\{x_k\}$, without regard for any optimality properties of individual approximants (but note Remark 4.7). (Note that, in the case of deterministic error bounds, the "best estimate", in a min-max sense, for the problem with approximate data will just be the "center" of the set $\tilde{S}$ of x in X consistent with the available information. In most cases $\tilde{S}$ will be unbounded so no "best estimate" exists; even if S_k is bounded for the sequentially embedded viewpoint, the sequence of approximants $\{\tilde{x}_k\}$ computed by letting each $\tilde{x}_k$ be the "best estimate" for the k-th approximating problem, need not converge to x).

DEFINITIONS: An *approximation scheme* for (1.1) consists of a sequence of maps $f_k: X \supset D_k \to Y$ with $\{f_k\}$ converging to f in some appropriate sense and a sequence of sets $B_k \subset Y$ with $B_k \downarrow \{b\}$. Such an approximation scheme is *regular* if $S_k := \{x \in D_k \subset X : f_k(x) \in B_k\}$ is nonempty for each (large) k and if there exists some convergent sequence $\{x_k\}$ with each x_k in S_k. *A solution algorithm* consists of a sequence of (computationally feasible=finitarily realizable) algorithms

$$\underset{\sim}{\sigma}_k : (\tilde{f}, \tilde{B}) \mapsto \tilde{x}, \tag{1.4}$$

i.e., each $\underset{\sim}{\sigma}_k$ is defined for suitable classes of computationally definable maps $\tilde{f}$ and subsets $\tilde{B} \subset Y$ and produces an element $\tilde{x}$ in X. A problem of the form (1.1) is *approximation solvable* if there exists a solution algorithm $\{\underset{\sim}{\sigma}_k\}$ such that, for each f of a specified class and each b in the range of f, the computed sequence $x_k := \underset{\sim}{\sigma}_k(f_k, B_k)$ converges to the desired solution x_* of (1.1) for any regular approximation scheme $\{(f_k, B_k)\}$.

This notion of approximation solvability is weaker than that of [52] but this seems necessitated by the applicability to ill-posed problems.

A preliminary version of the material of Sections 2 and 3 has been made available (1975 in preprint form [59]). Section 2 is concerned with the correct notion of convergence of $\{f_k\}$ to f in defining an approximation scheme, as above, and exhibiting a general form of solution algorithm with which one can demonstrate approximation solvability for a wide class of ill-posed problems. The argument for convergence is essentially that initially developed in [13], which provided the author's original introduction to ill-posed problems. The particular scheme employed in [13] and discussed more abstractly (with a far simpler convergence proof in the linear, Hilbert space case) in [58] is the method of "generalized interpolation" developed in Section 4, below; a preliminary form of portions of that section was presented in [50]. The general form of the solution algorithm in Section 2 is that of a sequence of constrained optimization problems so that the method of regularization does not appear directly in this form. Computationally, these approximating problems are reformulated in Section 3, for reduction to algorithmic form, by the use of Lagrange multipliers or penalty functions (although directly applicable to well-posed problems, it is worth noting [54] in this context) and this reformation provides the connection with the regularization approach.

As a complement to the rather abstract treatment of these first sections, Part II further develops and explicates these ideas in the context of application to some more concrete problems. A more detailed and further developed presentation of the material of Section 6 appeared elsewhere [65]. A preliminary version of Examples 7.2, 7.3 (along with other material) was presented in [62] and a more extensive development of this approach to problems of system identification for distributed parameter systems is in preparation for appearance elsewhere.

I would like to thank the late William Chewning for initiating the work which led to this investigation. That problem [13] is rather unusual in that a well-posed but computationally inaccessible problem was replaced by a computationally feasible ill-posed problem: abstractly, A, B are compact linear operators with $R(B) \subset R(A)$, making the map: $z \mapsto$(minimum norm solution x of

$$Ax + Bz = 0, \tag{1.5}$$

i.e., solution of (1.5) in $N(A)^{\perp}$) continuous, but it is convenient to solve, instead, the ill-posed problem

$$Ax=b \quad (b:=-Bz), \quad x\in N(A)^{\perp}. \tag{1.6}$$

Thanks are also due to the ARO for its support and to numerous individuals — most significantly to J. R. Cannon, G. Chavent, W. C. Chewning, C. Groetsch, M. Z. Nashed and G. Wahba — for the stimulus provided by their work, their correspondence and numerous conversations.

2. Approximation Theorems

We assume henceforth that X is a Efimov-Stečkin space [19], i.e., a reflexive Banach space for which one has

$$x_k \rightharpoonup x, \ \|x_k\|\to\|x\| \ \textit{imply} \ x_k\to x. \tag{2.1}$$

This holds, for example, in any uniformly convex (=uniformly rotund) Banach space [16] such as any Hilbert space or, e.g., $W^{1,p}$ with $1<p<\infty$. We use X_w to denote the space X taken with its weak topology.

Definitions: Given a sequence of sets $\{S_k\}$, an *associated sequence* is any $\{s_k\}$ with $s_k\in S_k$ for each k. A sequence of sets $\{S_k\}$ will be called *subconvergent to S* if, for every convergent associated subsequence, the limit is in S, i.e., if

$$s_{k(j)}\in S_{k(j)}, \ s_{k(j)}\to s \quad \textit{imply} \ s\in S. \tag{2.2}$$

A sequence of mappings $\{f_k: X\supset D_k\to Y\}$ will be called *graph subconvergent* to $f: X\supset D\to Y$ in $X_w\times Y$ if the sequence of graphs $F_k:=\{(x,f_k(x)): x\in D_k\}$ is subconvergent in $X_w\times Y$ to $F:=\{(x,f(x)): x\in D\}$, i.e., if

$$x_{k(j)}\in D_{k(j)}, \ x_{k(j)}\rightharpoonup x, \ f_{k(j)}(x_{k(j)})=:y_{k(j)}\to y \ \textit{imply} \ x\in D \ \textit{and} \ y=f(x). \tag{2.3}$$

In connection with the definition of *approximation scheme* given in Section 1, we take graph subconvergence in $X_w\times Y$ as the appropriate notion of convergence of $\{f_k\}$ to f.

Remarks 2.1: If $\{f_k\}$ is a constant sequence ($f_k=f$ for each k), then (2.3) corresponds to the assertion that f has a closed graph in $X_w\times Y$. If f_k, f are continuous linear maps (Y also a Banach space), then (2.3) is implied by the strong convergence of $\{f_k^*\}$ to f^*. If $\{D_k\}$ is subconvergent to D and $\{f_k\}$ converges to f uniformly on bounded sets, then $\{f_k\}$ is graph subconvergent to f. If f_k, f may be multivalued (i.e., relations rather than mappings), then the notion of graph subconvergence clearly still applies. ∎

Suppose, now, we consider a (possibly ill-posed) problem

$$f(x)=b \tag{2.4}$$

or, more generally (permitting F to be a relation as in the case of evaluating an unbounded operation: $F: =\{(x, y): x=Ly\}$ with L unbounded and possibly not injective),

$$(x, b) \in F. \tag{2.4'}$$

If (2.4′) does not uniquely determine x, then we impose the further condition that x is to have minimum norm among all solutions of (2.4′). We make the assumption that:

The equation (2.4) (*or* (2.4′)) *has a unique solution* x_* *of minimum norm.* (2.5)

REMARKS 2.2: The condition (2.5) includes the assumption that the right hand side of the problem is known to be in the range of f so that existence is not at issue. Note that (2.5) involves only the "true" limit problem (2.4) or (2.4′) and does *not* involve the approximating maps $\{f_k\}$. In considering (2.4), some sufficient conditions for the uniqueness of x_* are: uniqueness for (2.4) itself; f is continuous and linear (so the solution set of (2.4) is a closed subspace of X); $f: X \to Y = X^*$ is demicontinuous and maximal monotone (so the solution set of (2.4) is closed and convex). Note that if the desired solution of (2.4) is known to be isolated, then uniqueness may be obtained by redefining f, replacing it by its restriction to a suitable neighborhood of the solution.

In terms of the above, a *regular approximation scheme* for, say, (2.4) with, $f: X \supset D \to Y$ and b in Y satisfying (2.5) will be a sequence $\{(f_k, B_k)\}$ such that $f_k: X \supset D_k \to Y$ with (2.3) holding, $B_k \supset Y$ with each associated sequence converging to b, the sets

$$S_k: =\{x \in D_k: f_k(x) \in B_k\} \tag{2.6}$$

(or

$$S_k: =\{x \in X: (x, y) \in F_k \text{ for some } y \in B_k\} \tag{2.6'}$$

if one considers (2.4′)) are nonempty for large k and some associated sequence $\{\tilde{x}_k\}$ for $\{S_k\}$ converges to x_*.

Often this last condition is verified by showing that there is a sequence $\{\tilde{x}_k\}$ associated with $\{D_k\}$ converging (rapidly enough) to x_* and taking $\{B_k\}$ shrinking slowly enough to $\{b\}$ to have $f_k(\tilde{x}_k)$ in B_k. Note that for purposes of computation it is often important that the domain of each approximating map be (essentially) finite dimensional. Suppose $\{(f_k, B_k)\}$ is a given regular approximation sequence for which this may not be the case but suppose that there exists $\hat{D}_k \subset D_k$ such that each $\tilde{x}_k$ in D_k is (asymptotically) approximable by $\hat{x}_k$ in $\hat{D}_k$ (i.e., $\|\tilde{x}_k - \hat{x}_k\| \to 0$ so $\tilde{x}_k \to x_*$ implies $\hat{x}_k \to x_*$) and such that $\hat{D}_k$ is in the range of a parametrizing map $\pi_k: R^{m(k)} \supset R_k \to X$ so that the computation on $\hat{D}_k$ can be replaced, after composing f_k with π_k, by computation constricted to the finite-dimensional domain R_k. Let $\hat{f}_k$ be the restriction of f_k to $\hat{D}_k$ and note that if the regularity condition ($\tilde{x}_k \in S_k$, $\tilde{x}_k \to x_*$) was verified as suggested, the corresponding condition (perhaps after replacing B_k by a slightly enlarged $\hat{B}_k$) can similarly be verified to show that $\{(\hat{f}_k, \hat{B}_k)\}$ is also a regular approximation scheme.

Alternatively, note that the solution algorithms proposed below are variational in nature so that, even if D_k is not finite dimensional in any sense, one may know in advance (from properties of the minimization problem) that the approximant to be determined in D_k actually must lie in a subset $\hat{D}_k \subset D_k$ (e.g., a spline space; compare [44]). If this $\hat{D}_k$ is as above, then it is clear that, even though f_k remains defined on the larger domain D_k, computation can be done on $\hat{D}_k$ and so on $\boldsymbol{R}^{m(k)}$. ∎

We now propose the following *solution algorithm* $\{\underset{\sim}{\sigma}_k\}$. Suppose a sequence $\delta_k \to 0+$ is given. For any $\hat{f}: X \supset \hat{D} \to Y$ and $\hat{B} \subset Y$ set

$$S(\hat{f}, \hat{B}) := \{x \in \hat{D} : \hat{f}(x) \in \hat{B}\}, \ \nu(\hat{f}, \hat{B}) := \inf\{\|x\| : x \in S(\hat{f}, \hat{B})\}$$

(with $\nu = \infty$ if $S(\hat{f}, \hat{B})$ is empty). Then $\underset{\sim}{\sigma}_k(\hat{f}, \hat{B})$ is defined by taking

$$\underset{\sim}{\sigma}_k(\hat{f}, \hat{B}) \in \{x \in S(\hat{f}, \hat{B}) : \|x\| \leqslant \nu(\hat{f}, \hat{B}) + \delta_k\}; \tag{2.7}$$

the definition of $\underset{\sim}{\sigma}_k(\hat{f}, \hat{B})$ involves making an arbitrary selection from the set on the right of (2.7). Thus, $x_k := \underset{\sim}{\sigma}_k(f_k, B_k)$ is defined as being a solution of the approximating problem for (2.4):

$$f_k(x) \in B_k \tag{2.8}$$

having close to minimum norm.

THEOREM 2.3: *Let X satisfy* (2.1), *let Y be a Hausdorff space and let $f: X \supset D \to Y$ and b satisfy* (2.5). *Then* (2.4) (*more generally*, (2.4')) *is approximation solvable through the application of the solution algorithm $\{\underset{\sim}{\sigma}_k\}$ defined by* (2.7) *to an arbitrary regular approximation scheme $\{(f_k, B_k)\}$. I.e., if we set $x_k := \underset{\sim}{\sigma}_k(f_k, B_k)$ so $x_k \in D_k$ is a solution of* (2.8) *having close to minimum norm, then*

$$x_k \to x_* . \tag{2.9}$$

Proof: By assumption, there is some (fixed) sequence $\tilde{x}_k \to x_*$ associated with $\{S_k\}$. Thus, setting $\nu_k := \nu(f_k, B_k) := \inf\{\|x\| : x \in S_k\}$, we have $\nu_k \leqslant \|\tilde{x}_k\| \to \|x_*\|$ so $\{\|x_k\| \leqslant \nu_k + \delta_k\}$ is bounded. Any subsequence of $\{x_k\}$ contains a weakly convergent subsubsequence $x_{k(j)} \rightharpoonup \hat{x}$. Now $f_{k(j)}(x_{k(j)}) =: y_{k(j)} \in B_{k(j)}$ since $x_{k(j)} \in S_{k(j)}$. By assumption, $y_{k(j)} \to b$ so the graph subconvergence of $\{f_k\}$ to f implies that $\hat{x} \in D$ and $f(\hat{x}) = b$, i.e., that $\hat{x}$ is a solution of (2.4). On the other hand, lim sup $\|x_k\| \leqslant$ lim sup $(\nu_k + \delta_k) \leqslant$ lim $(\|\tilde{x}_k\| + \delta_k) = \|x_*\|$ so $\|\hat{x}\| \leqslant$ lim sup $\|x_{k(j)}\| \leqslant \|x_*\|$. By (2.5) this means $\hat{x} = x_*$ so $x_{k(j)} \rightharpoonup x_*$. Since this holds for a subsubsequence of every subsequence of $\{x_k\}$, we have $x_k \rightharpoonup x_*$. Since $\|x_*\| \leqslant$ lim inf $\|x_k\| \leqslant$ lim sup $\|x_k\| \leqslant \|x_*\|$, we have $\|x_k\| \to \|x_*\|$ so, by (2.1), we also have $x_k \to x_*$. ∎

REMARK 2.4: If it were not *known* that the right hand side b of (2.4) were in the range of f, then the solution algorithm $\{\underset{\sim}{\sigma}_k\}$ — solving (2.8) for a solution of (approximately) minimum norm for each k — can still be applied provided (2.8) does indeed have solutions for large k. If the sequence $\{x_k\}$ is bounded then, as is easily seen from the proof above, a subsequence converges to a solution of (2.4). Thus, if (2.4) were to have *no* solutions (b not in $R(f)$), the sequence $\{x_k\}$ obtained from the algorithm would have $\|x_k\| \to \infty$. This could be turned around to give an existence theorem: if one could somehow show that (2.8) has a bounded sequence of solutions $\{\hat{x}_k\}$ with $\|\hat{x}_k\| \leqslant M$, then it would follow that b is in the range of f and (2.4) has a solution x with $\|\hat{x}\| \leqslant M$. ∎

The theorem above gives convergence of the approximating sequence $\{x_k\}$ but no indication of the *rate of convergence* nor any estimate on the approximation error for any given k. It should be clear that this lack of an error bound is inevitable. Suppose, for simplicity, we were to have $f_k = f$ linear for each k. If the B_k were to contain balls of radius $\varepsilon_k \to 0$ and if one could obtain bounds $\beta_k \to 0$ on $\|x_k - x_*\|$, then this would imply continuity of a right inverse of f and thus contradict the assumed ill-posedness of the problem. On the other hand we shall see that a rate of convergence=asymptotic error bound may exist if some additional structure (compactness) is introduced. Compare the considerations of [5].

To consider the notion of a rate of convergence for the algorithm, we must view it as applying to the problem (2.4) determined for each b in the range $R(f)$. Thus, we assume we are given X satisfying (2.1) and $f: X \supset D \to Y$ satisfying:

for each b in the range of f there is a unique minimum norm solution $x_* = x_*(b)$ of (2.4). (2.5′)

Let $\{f_k: X \supset D_k \to Y\}$ satisfy (2.3) and set

$$S_k(b) := \{x \in D_k : f_k(x) \in B_k(b)\} \tag{2.10}$$

where $B_k(b) \subset [b + B_k^0]$. (This notation assumes, for simplicity, that Y is a linear space) with $B_k^0 \to \{0\}$ in the sense that $\tilde{y}_k \in B_k^0$ for each k implies $\tilde{y}_k \to 0$. We assume that for each $b \in R(f)$ and each k we have $S_k(b)$ non-empty and set $\nu_k(b) := \inf\{\|x\| : x \in S_k(b)\}$. Given a sequence $\delta_k \to 0+$, set

$$\begin{aligned} S_k'(b) &:= \{x \in S_k(b) : \|x\| \leqslant \nu_k(b) + \delta_k\}, \\ \psi_k(b) &:= \{\|x - x_*(b)\| : x \in S_k'(b)\}. \end{aligned} \tag{2.11}$$

The content of Theorem 2.3, then, is that $\psi_k(b) \to 0$. Having a (uniform) rate of convergence would mean that $\psi_k \to 0$ where

$$\psi_k := \sup\{\psi_k(b) : b \in R(f)\}. \tag{2.12}$$

THEOREM 2.5: *Let X, Y, f, $\{f_k\}$, $\{B_k(b)\}$ be as above so* (2.1), (2.5′) *are satisfied and* (2.10) *defines nonempty $S_k(b)$ for $b \in R(f)$. Given $\delta_k \to 0+$, define $\psi_k(b)$ by* (2.11). *Suppose $D_* := \{x_*(b) : b \in R(f)\}$ is compact and f is continuous on D_*. Then $\psi_k \to 0$, so* (2.8) *gives an approximation x_k to $x_*(b)$ with the error bound*

$$\|x_k - x_*(b)\| \leqslant \psi_k \to 0. \tag{2.13}$$

Proof: Suppose $\psi_k \not\to 0$. Then there would be a $\mu > 0$ and a sequence b_k in $R(f)$ such that (selecting a subsequence if necessary): $\psi_k(b_k) \geqslant \mu$ and $x_k^* := x_*(b_k) \to x_* =$

$=x_*(b_*)$, using the compactness of D_*. Since $\psi_k(b_k)\geqslant\mu$, there also exists $\hat{x}_k$ in $S'_k(b_k)$ such that $\|x_k-x_k^*\|>\mu/2>0$. Since f is continuous at x_*, we have $b_k=$ $=f(x_k^*)\to f(x_*)=b_*$ so there are neighborhoods U_k such that $f(x_k^*)\in b_*+U_k$ with $U_k\to\{0\}$ in Y. Setting $B_k^*:=b_*+U_k+B_k^0$, we clearly have $B_k(b_k)\subset B_k^*$ so

$$S_k^*:=\{x\in D_k: f_k(x)\in B_k^*\}\supset S_k(b_k),$$

$$v_k^*:=\inf\{\|x\|: x\in S_k^*\}\leqslant v_k(b_k).$$

As in the proof of Theorem 2.3, we have $v_k^*\to\|x_*\|$. Now each $\hat{x}_k$ is in $S'_k(b_k)$ so

$$\|\hat{x}_k\|\leqslant v_k(b_k)+\delta_k\leqslant\|x_k^*\|+\delta_k\to\|x_*\|$$

so

$$0\leqslant\|\hat{x}_k\|-v_k^*=:\delta_k^*\to 0_+.$$

Applying Theorem 2.3 with b_*, B_k^*, δ_k^* for b, B_k, δ_k, we see that $\hat{x}_k\in S_k^{*'}:=\{x\in S_k^*:$ $:\|x\|\leqslant v_k^*+\delta_k^*\}$ implies $\hat{x}_k\to x_*(b_*)=x_*$. Since also $x_k^*\to x_*$, this contradicts the assumption that $\|x_k-x_k^*\|\geqslant\mu/2$. ■

REMARKS 2.6: The condition that D_* be compact is an automatic consequence of an assumed compactness of D when f is injective so $D_*=D$. In general, even with D compact D_* need not be closed and we note a simple finite-dimensional example ($X=\boldsymbol{R}^2$, $Y=\boldsymbol{R}^3$) for which all the other hypotheses of Theorem 2.4 are satisfied but D_* is not closed and $\psi_k\not\to 0$: let $f_k=f$ be given on $D:=[0,3]\times[0,1]$ by

$$f(\xi,\eta):=(\xi[2-\xi]_+,[\xi-2]\eta,[\xi-2]_+)$$

(α_+ means: α if $\alpha\geqslant 0$, 0 if $\alpha\leqslant 0$). One verifies easily that $D_*=([0,1]\cup[2,3])\times[0,1]$. Also, for any sequences $\delta_k\to 0_+$, $\varepsilon_k\to 0_+$, let $B_k(b):=\{y\in\boldsymbol{R}^3:\|y-b\|\leqslant\varepsilon_k\}$ and the hypotheses of Theorem 2.4 (other than closure of D_*) are satisfied. In this case one has ψ_k bounded below by 1 so $\psi_k\not\to 0$. ■

The notion of "graph subconvergence" considered here is entirely different from the notion of graph convergence for closed linear maps between Banach spaces (defined by the Hausdorff metric between the $X\times Y$ — norm unit spheres of the graphs) in [36], IV, 4. It is much closer to Stummel's notion (see, e.g., [66]) of a "discretely closed" sequence and is still more in the spirit of [67] although developed independently: cf. [13], [59]). Indeed, the similarity of "flavor" to Tanana's paper is even more marked if one notes that the compactness of D in Theorem 2.4 is plausibly the result of a "regularity condition" imposed in the form of assuming that the desired solution is actually not only in $D=D(f)\subset X$ but is in the range of a compact embedding $E: R\hookrightarrow X$. (We refer to R as the *regularity space*, e.g., $E: R=C^1\hookrightarrow X=L^2$. See, e.g., [1] for the use of this approach for similar error bounds in the context of well-posed problems.) Given an *a priori* bound M on the R-norm of x_*, we restrict f to the compact domain $cl\{x=Eu: \|u\|_R\leqslant M\}$, assuming it is already defined there. If f were homogeneous of some degree — e.g., if f were linear — one might factor M out of the estimate, relating the convergence rate to that for the case $M=1$. Results related to Theorem 2.4 appear in [30]; for a more detailed treatment of the linear case, see Lemma 4.4 below.

3. Penalty Functions and Minimization: Regularization

The algorithm of Theorem 2.3 requires specifying approximating sets $B_k \to \{b\}$ and finding a solution of (approximately) minimum norm for the problem:

$$f_k(x) \in B_k. \tag{3.1}$$

Equivalently, we have the minimization problem:

$$\textit{minimize } \|x\|, \ x \in S_k := \{x \in D_k : f_k(x) \in B_k\} \tag{3.1'}$$

and this viewpoint suggests the possible use of a penalty function approach to the constraint: $x \in S_k$.

Especially if the problem is presented in the approximate form

$$f_k(x) = b_k \tag{3.2}$$

(with $f_k \to f$, say in the sense of (2.3), and $b_k \to b$), then we are likely to have available an estimate of the magnitude of the error

$$\|f_k(x_*) - b_k\|_Y \leqslant \varepsilon_k \to 0 \tag{3.3}$$

where x_* is the true solution of the problem: $f(x) = b$ (the form of (3.3) assumes Y is also a Banach space but we will generalize that in Theorem 3.1, below). We may treat (3.3) as a definition of B_k so (3.3) takes the equivalent form

$$x_* \in B_k := \{x \in D_k : \|f_k(x) - b_k\| \leqslant \varepsilon_k\}. \tag{3.3'}$$

Comparing with (3.1'), this suggests including a penalty function $\lambda_k \| f_k(x) - b_k\|$ with a suitable choice of (large) λ_k in constructing the minimization problem:

$$\textit{minimize } J_k(x) := [\|x\| + \lambda_k \| f_k(x) - b_k\|], \qquad x \in D_k. \tag{3.4}$$

(Alternatively, one might arrive at this by observing that, if 0 is not in S_k, a minimum for (3.1') would occur on the boundary of S_k. Presumably, for B_k as in (3.3'), this is where

$$\|f_k(x) - b_k\| = \varepsilon_k \tag{3.5}$$

which may then be used as the constraint instead of the inequality (3.3). A Lagrange multiplier treatment gives the function $[\|x\| + \lambda_k (\| f_k(x) - b_k\| - \varepsilon_k)]$ for global minimization and, since $-\lambda_k \varepsilon_k$ is constant in minimizing with respect to x, this is equivalent to (3.4).)

Note that (3.4) can "equivalently" be replaced by minimization of, say,

$$\|x\|_X^p + \lambda_k \| f_k(x) - b_k\|_Y^{p'} \tag{3.6}$$

which, if the X—norm is modelled on L^p $(W^{r,p})$ for $1 < p < \infty$, etc., makes this differentiable provided f_k is differentiable. This is particularly advantageous in the case of Hilbert space norms (in which case we take $p, p' = 2$ so we have a quadratic cost criterion) and linear f_k. In that case minimization for (3.6) corresponds, after differentiation, to the *linear* problem:

$$[f_k^* f_k + \lambda_k^{-1} I]\, x = f_k^* b_k \tag{3.7}$$

which we recognize as (the regularized form of) the normal equation for (3.2). The relation between the appropriate choice of $\lambda=\lambda_k$ and the $\varepsilon=\varepsilon_k$ defining B_k (see, e.g., [45]) gives

$$\lambda' = -[\|x'\|^2+\lambda\| f_k x'\|^2]/\varepsilon<0$$

($'$ denotes $d/d\varepsilon$) so λ is strictly increasing as ε decreases and conversely. Clearly $\lambda\to\infty$ as $\varepsilon\to 0+$ (since, actually, $(\varepsilon\lambda)'<0$) and, if we assume b_n is in the closure of the range of f_n, it is easy to see that $\|f_n x-b_n\|=\varepsilon\to 0$ as $\lambda\to\infty$ in (3.4).

We note that this is essentially the *method of regularization* proposed by Tikhonov ([69], etc.) and extensively studied since (note the monograph [47]). The original version took $f_k\equiv f$, $b_k\equiv b$ but variants have considered approximations as well. The considerable literature which has developed with respect to the method of regularization treats such topics as: approximation properties for various norms (see, e.g., [21]) and optimality, choice of λ_k (see, e.g., [45], [72], [15]), computational aspects (see, e.g., [26], [40], [25]), etc.

Observe that in many cases the approximating maps f_k will have finite rank. One then knows that x_k will be in the range of f_k^* so that (3.5), (3.7) can be considered in a purely finite dimensional setting (in (3.5) we replace b_k by its projection onto the range of f_k). The computational aspects of obtaining x_k (minimizing $\|x\|$ subject to (3.5)) and the relation to (3.7) are explored, e.g., in [22].

We generalize somewhat the approach leading to (3.7), etc., and consider X to be a reflexive Banach space with a specified coercive (i.e., bounded only on bounded sets), weakly lower semicontinuous (wlsc) functional $N: X\to[0,\infty]$ such that

$$x_k\rightharpoonup x,\ N(x_k)\to N(x)\ \textit{implies}\ x_k\to x \tag{3.8}$$

and consider Y to be a uniform Hausdorff space with its topology determined by a family $\{\rho_1,\rho_2,\ldots\}$ of semimetrics (with no loss of generality we assume $\rho_1\leqslant\rho_2\leqslant\ldots$).

THEOREM 3.1: *Let X, Y, N be as above; suppose $f: X\supset D\to Y$ and b in Y are such that* (2.4) *has a unique solution x_* minimizing N; let $N(x_*)<N_*<\infty$. Let $\{f_k\}$ satisfy* (2.3) *and $b_k\to b$. Assume there is a sequence $\{x_k\in D_k\}$ such that (replacing $\{(f_k,b_k)\}$ by a subsequence if necessary) one has*

$$\rho_k\,[f_k\,\tilde{x}_k),b_k]\leqslant\varepsilon_k \tag{3.9}$$

so that, taking $B_k:=\{y\in Y:\rho_k\,[y,b_k]\leqslant\varepsilon_k\}$, the sequence $\{(f_k,B_k)\}$ is a regular approximation scheme for (2.4). *Let $\varphi_*:[0,\infty]\to[0,\infty]$ be continuous, monotone and coercive ($r\to\infty$ implies $\varphi_*(r)\to\infty$) and, for each k, let $\varphi_k:[0,\infty]\times[0,\infty]\to[0,\infty]$ be increasing in each variable with*

$$\begin{aligned}
&\text{(a)}\quad r_k\to r\ \textit{implies}\ \varphi_k(r_k,\varepsilon_k)\to\varphi_*(r),\\
&\text{(b)}\quad \varphi_*(r)\leqslant\varphi_k(r,0)+\gamma_k,\\
&\text{(c)}\quad \{\varphi_k(r_k,\alpha_k)\}\ \textit{bounded implies}\ \{r_k\}\ \textit{bounded and}\ \alpha_k\to 0.
\end{aligned} \tag{3.10}$$

Set

$$J_k(x) := \varphi_k\big(N(x), \rho_k[f_k(x), b_k]\big) \quad \textit{for} \quad x \in D_k,$$
$$J_k^* := \inf\{J_k(x) : x \in D_k,\ N(x) \leqslant N_*\}. \tag{3.11}$$

If x_k is an approximate solution of

$$\textit{minimize } J_k(x) \textit{ over } D_k, \tag{3.12}$$

i.e., if x_k satisfies

$$J_k(x_k) \leqslant J_k^* + \delta_k \tag{3.12'}$$

for each k, then $x_k \to x_*$.

In the discussion, above, of regularization/penalty function method, one had $N(x) := \|x\|_X$ (so (3.8) is just (2.1)) and, for each k, $\rho_k(u, v) := \|u - v\|_Y$ and $\varphi_*(r) := r$, $\varphi_k(r, \alpha) := r + \lambda_k \alpha$ for (3.4) or $\varphi_*(r) := r^2$, $\varphi_k(r, \alpha) := r^2 + \lambda_k \alpha^2$ for (3.6)–(3.7). The inequality (3.12′) can be taken to define a solution algorithm $\{\underset{\sim}{\hat\sigma}_k\}$ in much the same way as (2.7) defined $\{\underset{\sim}{\sigma}_k\}$. The content of Theorem 3.1 is, then, that the sequence of unconstrained minimization problems (3.12) can be used to obtain approximation solvability for a wide class of ill-posed problems by applying this solution algorithm to the approximating sequence $\{(f_k, b_k)\}$, i.e., to the regular approximation scheme $\{(f_k, B_k)\}$ with B_k defined via (3.9) as above.

Proof: We have

$$J_k^* \leqslant J_k(\tilde x_k) := \varphi_k\big(N(\tilde x_k), \rho_k[f_k(\tilde x_k), b_k]\big) \leqslant \varphi_k\big(N(\tilde x_k), \varepsilon_k\big)$$

so

$$\limsup J_k^* \leqslant \varphi_*\big(N(x_*)\big) \tag{3.13}$$

by (3.10a), the convergence of $\{\tilde x_k\}$ to x_* and the assumed semicontinuity of $N(\cdot)$. Next, by (3.12), (3.10c) and the coercivity of $N(\cdot)$, we have $\{x_k\}$ bounded and $\rho_k[f_k(x_k), b_k] \to 0$. Since $\rho_1 \leqslant \rho_2 \leqslant \ldots$,

$$\rho_n[f_k(x_k), b_k] \to 0 \quad \textit{for each } n.$$

Thus, since $b_k \to b$, we have $f_k(x_k) \to b$ in Y. Further, each subsequence of $\{x_k\}$ contains a weakly convergent subsequence: $x_{k(j)} \rightharpoonup \hat x$. By the weak lower semicontinuity of $N(\cdot)$,

$$\begin{aligned}
\varphi_*\big(N(\hat x)\big) &\leqslant \varphi_*\big(\liminf N(x_k)\big) \leqslant \limsup \varphi_*\big(N(x_k)\big) \leqslant \\
&\leqslant \limsup\, [\varphi_k\big(N(x_k), 0\big) + \gamma_k] \leqslant \\
&\leqslant \limsup\, [J_k(x_k) + \gamma_k] \leqslant \\
&\leqslant \limsup\, [J_k^* + \delta_k + \gamma_k] \leqslant \varphi_*\big(N(x_*)\big)
\end{aligned} \tag{3.14}$$

(writing k for $k(j)$). Thus, $N(\hat x) \leqslant N(x_*)$. The graph subconvergence of $\{f_k\}$ to f in $X_w \times Y$ implies $f(\hat x) = b$ since $x_{k(j)} \rightharpoonup \hat x$, $f_{k(j)}(x_{k(j)}) \to b$; by the minimizing property of x_*, this means $\hat x = x_*$. This gives $x_k \rightharpoonup x_*$ for the full sequence $\{x_k\}$ since $\{x_{k(j)}\}$ was selected from an arbitrary subsequence. Since (3.14) shows $\lim \varphi_*\big(N(x_k)\big) = \varphi_*\big(N(x_*)\big)$ so $\lim N(x_k) = N(x_*)$, one has $x_k \to x_*$ by (3.8). ∎

REMARKS 3.2: Assuming suitable differentiability (for this Y should also be a linear space), the problem of minimizing the functional J_k over D might be approached by solving

$$J_k'(x)=0, \qquad x \in D \tag{3.15}$$

where J_k' is the Fréchet derivative dJ_k/dx. As with the special case (3.6)-(3.7) in which J_k' is affine, the problem (3.15) will typically be well-posed. Indeed, if N, ρ_k, φ_k are such that J_k is strictly convex, as with (3.6), then the map $J_k'\colon X \supset D \to X^*$ will be a monotone map and so, under standard hypotheses (cf., e.g., [50]), continuously invertible.

If the existence of a solution x_* and of $\{\tilde{x}_k\}$ were not known in advance, one might still attempt using (3.12) to obtain an approximating sequence. If the sequence $\{J_k\}$ were bounded, then one would still have that $f_k(x_k) \to b$ and that some subsubsequence of any subsequence converges weakly to a solution $\hat{x}$. (That this is then the best possible result is shown by the following examples: let X be a Hilbert space with orthonormal sequence $\{e_k\}$, let $Y=\boldsymbol{R}$ and let $f\colon X \supset D := \{0\} \mapsto 0$, $f_k\colon X \supset D_k := \{e_k\} \mapsto 0$ so $x_k = e_k \rightharpoonup x_* = 0$ but $x_k \not\to 0$; next, let $X=\boldsymbol{R}=Y$ and let $f\colon X \supset D := \{-1, 0, 1\} \mapsto 0$, $f_k\colon X \supset D_k := \{(-1)^k\} \mapsto 0$.)

The condition in Theorems 2.3 and 3.1 that a sequence such as $\{\tilde{x}_k\}$ exists is trivially necessary — the point of the result is that any sequence $\{x_k\}$ obtained by the algorithms (2.7), (3.12) must be like that — but is difficult to verify directly. A *sufficient set of conditions* implying both the *existence* of $\{\tilde{x}_k\}$ and the *graph subconvergence* of f_k to f is that:

Y is metrizable, the graph of f is closed in $X_w \times Y$, $D_k = D$ and $f_k \to f$ uniformly on bounded subsets. (3.15)

(Proof: Let $\tilde{x}_k = x_*$ for each k. Suppose $x_k \rightharpoonup x$, $y_k := f_k(x_k) \to y$. Then $\rho[f(x_k), y] \leqslant \rho[f_k(x_k), f(x_k)] + \rho[f_k(x_k), y] \to 0$. As the graph of f is closed, (x, y) is then in the graph so $y = f(x)$.) The condition that $D_k = D$ for each k can be relaxed if also f is continuous from X to Y. ∎

4. Approximation by Generalized Interpolation

In many cases the right hand side of the equation

$$f(x)=b \tag{4.1}$$

($f\colon X \supset D \to Y$) is known only through a discrete sequence of measurements. That is, one has (approximations to) the values $\{\beta^k := \eta_k(b)\}$ where $\{\eta_1, \eta_2, \ldots\}$ is a sequence of continuous functionals $\eta_k\colon Y \to \boldsymbol{R}$. If the sequence $\{\eta_k\}$ *separates points* (i.e., if $\eta_k(y) = \eta_k(\hat{y})$ for all k implies $y = \hat{y}$), then (4.1) is equivalent to the infinite system of equations

$$\eta_k(f(x)) = \beta^k \quad k=1, 2, \ldots . \tag{4.1'}$$

In many situations the measurement functionals η_k will correspond to point evaluations: if $b \in Y$ is a function $b(\cdot)$, then $\beta^j := \eta_j(b)$ might be $b(t_j)$. Consistently with such an interpretation, we might refer to this as a problem of *generalized interpolation* — interpolating the data $\beta_{(k)}$ by an element of the range of f: i.e., $b_{(k)} = f(x_k^*)$ where x_k^* is the minimum norm element of D consistent with the measurement values $\beta_{(k)}$. This approach is also related to the "collocation-projection" method of [71].

For purposes of computational approximation, we consider the sequence of problems:

$$\eta_j(f(x)) = \beta^j \quad j = 1, \ldots, k. \tag{4.2}$$

Introducing $H_k : y \mapsto (\eta_1(y), \ldots, \eta_k(y)) : Y \to \boldsymbol{R}^k$ and setting $f_{(k)} := H_k \circ f$ and $\beta_{(k)} := (\beta^1, \ldots, \beta^k)^*$, (4.2) takes the form

$$f_{(k)}(x) = \beta_{(k)}. \tag{4.2'}$$

Note that for fixed k the problem (4.2)=(4.2′) will not in general be ill-posed in the sense we have been considering since the codomain of $f_{(k)}$ is finite-dimensional; in particular, if f and each of the $\{\eta_j\}$ are linear then the range of $f_{(k)}$ will be closed (a subspace of $\boldsymbol{R}^k$). In the nonlinear case, if f is "smooth" (say, with a continuous Fréchet derivative), then one might expect $f_{(k)}$ to be *normally solvable* in the sense of Pogožaev [53], etc., so (4.2) would be treatable by established computational procedures.

In the linear case with X, Y Hilbert spaces, this approach to (4.1) was considered in [58]; compare, also, [46] (in which similar ideas are applied to evaluation of an unbounded linear operator) and [72]. Much of the remaining material of this section was presented in a preliminary form in [60].

THEOREM 4.1: *Let X be a reflexive Banach space, Y Hausdorff, let the graph of $f: X \supset D \to Y$ be closed in $X_w \times Y$ and let $\{\eta_j\}$ be a set of continuous functionals on Y which separates points of the range of f. Suppose N is a wlsc functional on X satisfying* (3.8) *and that there is a unique solution x_* of* (4.1) *minimizing N. Let $\delta_k \to 0+$ and let*

$$\begin{gathered} x_k \in S_k := \{x \in D : (4.2')\}, \\ N(x_k) \leqslant N_k^* + \delta_n \ \textit{with} \ N_k^* := \inf\{N(x) : x \in S_k\}. \end{gathered} \tag{4.3}$$

Then $x_k \to x_*$.

THEOREM 4.2: *Let X be a Hilbert space, Y a topological linear space, $f: X \supset D \to Y$ linear and with graph closed in $X \times Y$ and let $\eta_j \in Y^*$ be in the domain D^* of the adjoint $f^*: Y^* \supset D^* \to X^* = X$ with $\{\eta_j\}$ separating points of the range of f. Then for any b in the range of f there is a unique solution x_* of* (4.1) *having minimum norm* *and for each k there is a unique solution* x_k^* *of* (4.2′) *having minimum norm. This* x_k^* *is the orthogonal projection* $P_k x_*$ *of* x_* (*indeed*, $x_k^*=P_k x$ *for any solution* x *of* (4.1)) *onto the subspace* $X_k: =\mathrm{sp}\{f^* \eta_1, \ldots, f^* \eta_k\}$ *and can be computed by writing*

$$x_k^*=\sum_{j=1}^{k} \xi^j f^* \eta_j \tag{4.4}$$

and obtaining the coefficients $\xi_{(k)}: =(\xi^1, \ldots, \xi^k)^*$ *by solving*

$$\begin{gathered} A_k \xi_{(k)}=\beta_{(k)}, \\ A_k: =\left((a^{i,j})\right)_{i,j=1}^{k}, \quad a^{i,j}: =<f^* \eta_i, f^* \eta_j>. \end{gathered} \tag{4.5}$$

Proofs: Theorem 4.1 follows immediately from Theorem 2.3 with $f_k=f$ and $B_k: ==\{y \in Y: \eta_j(y)=\beta^j \text{ for } j=1, \ldots, k\}$ for $k=1, 2, \ldots$. It is easy to see that (modulo the replacement of (2.1) by (3.8)) the hypotheses of Theorem 2.3 are satisfied. Theorem 4.2 was proved in [58]; see also [13]. ∎

For the rest of this section we assume that, as in Theorem 4.2 we take X to be a Hilbert space, Y a topological linear space and $\{\eta_1, \eta_2, \ldots\}$ to be a total sequence in Y^*.

Remark 4.3: For $f: X \supset D \to Y$ nonlinear but Fréchet differentiable, suppose N_k^* in (4.3) is a minimum, taken on uniquely in S_k at x_k^*. Then, generalizing (4.4), (4.5), there is a representation

$$x_k^*=\sum_{j=1}^{k} \xi^j B^* \eta_j, \quad B: =f'(x_n), \tag{4.4′}$$

$$\begin{gathered} A_k \xi_{(k)}=\beta_{(k)}, \\ A_k: =\left((a_k^{i,j})\right)_{i,j=1}^{k}, \quad a_k^{i,j}=<B^* \eta_i, B^* \eta_j>. \end{gathered} \tag{4.5′}$$

To obtain this we apply the method of Lagrange multipliers to the problem of minimizing $\|x\|$ subject to the constraint (4.2). Thus, define F on $D \times \boldsymbol{R}^k$ by

$$F_k(x, \xi): =\frac{1}{2}\|x\|^2-\sum_{j=1}^{k} \xi^j\left[<\eta_j, f(x)>-\beta^j\right].$$

Differentiating with respect to x gives (4.4′) and substituting that in (4.2) then gives (4.5′). (The nature of the system (4.4′), (4.5′) suggests an iterative approach in which an approximant to x_n^* is used for evaluating f' to obtain B, A_k approximately and then the approximate system solved to obtain a new — possibly better — approximant to x_k^*.) ∎

We now assume further, for the rest of this section, that $f: X \to Y$ is linear so the setting is that of Theorem 4.2. In this context we can give more detail for the "rate of convergence" considerations of Theorem 2.5. As noted there, this is related in spirit to the results of [67], [48] which consider optimality with respect to certain assumed information. Again it is convenient to introduce a "regularity space" $\boldsymbol{R}$ which we now assume is also a Hilbert space densely embedded in X with compact embedding map $E: R \hookrightarrow X$ (thus, for x in the range of E, so $x=Er$ with $r \in R$, we identify x with r so $R \subset X$) and we now also introduce the eigenvectors $\{r_1, r_2, \dots\}$, taken orthonormal in R, and corresponding eigenvalues $\{\rho_1, \rho_2, \dots\}$, taken as ordered with $\rho_1 \geqslant \rho_2 \geqslant \dots > 0$, of the compact positive operator E^*E on R. Note that for $x = \sum_k \xi^k r_k$ in $R \subset X$,

$$\|x\|_R^2 = \sum_k |\xi^k|^2, \quad \|x\|_X^2 = \sum_k \rho_k |\xi^k|^2 .$$

Theorem 4.2 implies that the error of approximation of the desired solution x_* by x_n^* is given by

$$\|x_* - x_k^*\| = \|x_* - X_k\| :=\inf \{\|x_* - x\| : x \in X_k\}. \tag{4.6}$$

In Theorem 4.2 we took $X_k := \mathrm{sp}\,\{f^* \eta_1, \dots, f^* \eta_k\}$, which will generally be k-dimensional. As an indication of the rate of convergence (compare [21]) we introduce, for finite dimensional subspaces $S \subset X$,

$$\psi_M(S) := \sup \{\|x - S\| : x \in R, \quad \|x\|_R \leqslant M\} \tag{4.7}$$

and have the following.

LEMMA 4.4: *For $k := \dim S$ and any n*

$$M\sqrt{\rho_{k+1}} \leqslant \psi_M(S) = M\psi_1(S) \leqslant M\left[\rho_{n+1} + \sum_{j=1}^{n} \|r_j - S\|^2\right]^{1/2}. \tag{4.8}$$

Proof: Clearly $\psi_M = M\psi_1$ by homogeneity and the inequality $\psi_1(S)^2 \geqslant \rho_{n+1}$ follows a standard variational characterization of the eigenvalues (cf., e.g., [18], p. 908). For any $x \in R$ with $\|x\|_R \leqslant 1$, write $x = \sum_j \xi_j r_j$ (so $\sum_j |\xi_j|^2 \leqslant 1$), set $\tilde{x} := \sum_{j=1}^{n} \xi_j r_j$ and let $\hat{x}$ be the projection $P_* \tilde{x}$ of $\tilde{x}$ on S. Then

$$\|x - X_*\| \leqslant \|x - \tilde{x}\| + \|\tilde{x} - \hat{x}\| \leqslant \Big[\sum_{j=n+1}^{\infty} \rho_j |\xi^j|^2\Big]^{1/2} + \sum_{j=1}^{n} |\xi^j| \, \|r_j - P_* r_j\|$$

which gives $\|x - S\|^2 \leqslant [\rho_{n+1} + \sum_{j=1}^{n} \|r_j - S\|^2]$ by the Cauchy inequality. ∎

The *optimal convergence rate* $\{\hat{\psi}_k := \inf \{\psi_1(S) : \dim S = k\}\}$ is thus given by $\hat{\psi}_k = \sqrt{\rho_{k+1}}$, which is attainable if $\{r_1, r_2, \dots\}$ is in the range of f^* so $\{\eta_1, \dots\}$ can be taken to give $f^* \eta_j = r_j$ so $X_k = \mathrm{sp}\,\{r_1, \dots, r_k\}$. Note that this optimal rate depends only on the regularity condition and not at all on f — with the irrelevance of the ill-posedness of (4.1) due to the omission, thus far in this section, of any consideration of the effects of imprecision in the data. The number $\hat{\psi}_n$ is just the *n-width* of R as embedded in X by E; see [31], [32] for discussion and asymptotic estimates in some specific cases with $X = L^2(\Omega)$, $\Omega \subset IR^m$. We now consider the computation of (4.4), (4.5) when b, f may be only approximately available.

THEOREM 4.5: *Let X, Y be Hilbert spaces, let $f, \tilde{f}: X \to Y$ be continuous linear maps, let b be in the range of f with x_* the minimum norm solution of $fx=b$ and let $\{\eta_1, \ldots, \eta_n\}$ be in Y^* with $\{f^*\eta_1, \ldots, f^*\eta_k\}$ linearly independent. Define*

$$H: Y \to R^k: y \mapsto (\langle \eta_1, y\rangle, \ldots, \langle \eta_k, y\rangle)^*, \tag{4.9}$$

$$\begin{aligned} A: &= ((\alpha^{i,j}))_{i,j=1}^k, \qquad \tilde{A} = ((\tilde{\alpha}^{i,j}))_{i,j=1}^k, \\ \alpha^{i,j}: &= \langle f^*\eta_i, f^*\eta_j\rangle, \qquad \tilde{\alpha}^{i,j}: = \langle \tilde{f}^*\eta_i, \tilde{f}^*\eta_j\rangle, \end{aligned} \tag{4.10}$$

and assume $\|f-\tilde{f}\|$ small enough to ensure existence of $\tilde{A}^{-1}$. For any $\tilde{\beta}$ in R^k define $\tilde{x}$ in X by

$$\tilde{x}: = \sum_{j=1}^k \tilde{\xi}^j \tilde{f}^* \eta_j \text{ where } \tilde{\xi}: = (\tilde{\xi}^1, \ldots, \tilde{\xi}^k)^* = \tilde{A}^{-1}\tilde{\beta}, \tag{4.11}$$

i.e., by taking

$$\tilde{x} \text{ is the unique minimum norm solution of: } H\tilde{f}x = \tilde{\beta}. \tag{4.11'}$$

Then

$$\|x_* - \tilde{x}\|^2 \leqslant (\|x_* - X_0\| + \|P - \tilde{P}\|\,\|x_*\|)^2 + (\|\hat{f}^{-1}\|\,[\|Hb - \tilde{\beta}\| + \|H\|\,\|f - \tilde{f}\|\,\|x_*\|])^2 \tag{4.12}$$

where P and $\tilde{P}$ are the operators of orthogonal projection onto $X_0: = \mathrm{sp}\{f^*\eta_1, \ldots, f^*\eta_k\}$ and $\tilde{X}: = \mathrm{sp}\{\tilde{f}^*\eta_1, \ldots, \tilde{f}^*\eta_k\}$, respectively, and $\hat{f}$ is the restriction of $H\tilde{f}$ to $\tilde{X}$. One has estimates for $\|P-\tilde{P}\|$ and $\|\hat{f}^{-1}\|$ in terms of $\|H\|$, $\|f\|$, $\|f_0^{-1}\|$ and $\|\tilde{f}-f\|$ with $\|P-\tilde{P}\| \to 0$ and $\|\hat{f}^{-1}\| \to \|f_0^{-1}\|$ (f_0 is the restriction of Hf to X_0) as $\|\tilde{f}-f\| \to 0$.

Proof: The assumption that $\{f^*\eta_1, \ldots, f^*\eta_k\}$ is independent means that X_0 is n-dimensional and the Gramian A is invertible. An easy computation verifies that $A = f_0 f_0^* = Hf(Hf)^*$ and, similarly, that $\tilde{A} = \hat{f}\hat{f}^* = H\hat{f}(H\hat{f})^*$. Thus,

$$\begin{aligned} &\|A^{-1}\| = \|f_0^{-1}\|^2, \qquad \|\tilde{A}^{-1}\| = \|\hat{f}^{-1}\|^2, \\ &\|A - \tilde{A}\| = \|H(ff^* - \tilde{f}\tilde{f}^*)H^*\| \leqslant \|H\|^2 \|f - \tilde{f}\|(\|f\| + \|\tilde{f}\|) \end{aligned} \tag{4.13}$$

so f_0 is invertible and $\tilde{A}$ (hence also $\hat{f}$) will be invertible for $\|f-\tilde{f}\|$ small enough to make $\|A-\tilde{A}\|$, as estimated by (4.13), less than $\|A^{-1}\|^{-1}$. Henceforth we assume $\|\tilde{f}-f\|$ small enough that $\|\tilde{f}\| \leqslant 2\|f\|$ and $\|A-\tilde{A}\| \leqslant 1/2\|A^{-1}\|$, i.e., using (4.13), assume

$$\|\tilde{f} - f\| \leqslant \|f\| \min\{1, 1/6\|H\|^2\|f_0^{-1}\|\}. \tag{4.14}$$

Then $\|\tilde{A}^{-1}\| \leqslant 2\|A^{-1}\|$ and

$$\|\hat{f}^{-1}\| \leqslant 2\|f_0^{-1}\|, \qquad \|\tilde{A}^{-1} - A^{-1}\| \leqslant 2\|A^{-1}\|^2\|A - \tilde{A}\|. \tag{4.15}$$

Again an easy computation verifies that

$$P = (Hf)^* A^{-1}(Hf), \qquad P = (Hf)^* A^{-1}(Hf),$$

from which $\|P-\tilde{P}\|$ can be suitably estimated using (4.13), (4.15):

$$\|P - \tilde{P}\| \leqslant [3\|f_0^{-1}\| + 6\|f_0^{-1}\|^4\|H\|]\,\|H\|\,\|\tilde{f} - f\|. \tag{4.16}$$

One easily sees that (4.11) and (4.11′) are equivalent so $\tilde{x}=\tilde{f}^{-1}\tilde{\beta}$. Now set $\hat{\beta}:=H\tilde{f}x_*$ and $\hat{x}:=\tilde{P}x_*$. Note that

$$\tilde{f}\hat{x}=H\tilde{f}\tilde{P}x_*=H\tilde{f}x_*=:\hat{\beta} \quad so \quad \hat{x}=\tilde{f}^{-1}\hat{\beta}.$$

As $\tilde{P}$ is the orthogonal projection onto X and $\tilde{x},\hat{x}\in\tilde{X}$, one has

$$\|x_*-\tilde{x}\|^2=\|x_*-\hat{x}\|^2+\|\hat{x}-\tilde{x}\|^2\leqslant$$
$$\leqslant(\|x_*-Px_*\|+\|P-\tilde{P}\|\,\|x_*\|)^2+\|\tilde{f}^{-1}(\hat{\beta}-\tilde{\beta})\|^2$$

which just gives (4.12) since $Hb-\hat{\beta}=H(f-\tilde{f})x_*$. ∎

Now let $\{\eta_1,\eta_2,\dots\}$ be a total sequence in Y, let $f_k\to f$ (in operator norm) and let $b_k\to b$. For each k we can apply Theorem 4.5 with $\tilde{f}:=f_k$, $\tilde{\beta}:=H_k b_k$, X_k the range of $(H_k f)^*$, $\tilde{X}_k$ the range of $(H_k\tilde{f})^*$ and f_0 the restriction of f to X_k (let $\nu_k:=\|f_0^{-1}\|$; we assume f_k satisfies (4.14) for each k). For comparison with Theorem 2.5, let f, f_k be restricted to $D=D_M:=\{x\in R\subset X:\|x\|_R\leqslant M\}$ and define ψ_k by (2.12). Define $\psi_M(X_k)$ by (4.7), assume $\|x\|\leqslant\|x\|_R$ for $x\in R\subset X$ and substitute the estimates (4.15), (4.16) into (4.12). One then obtains

$$\begin{aligned}\psi_k\leqslant M\,[(\psi_1(X_k)+3\,(\nu_k+2\nu_k^4\|H_k\|)\|H_k\|\,\|f_k-f\|)^2+\\ +4\nu_k^2\|H_k\|^2(\|f_k-f\|+\|b_k-b\|/M)^2]^{1/2}.\end{aligned}\tag{4.17}$$

As k increases $\psi_1(X_k)\to 0$ and $\nu_k\to\infty$, reflecting the ill-posedness of the problem. Then (4.17) shows that $\psi_k\to 0$ if $b_k\to b$ and $f_k\to f$ rapidly enough. Note that one might be able to apply Theorem 2.5 in the context of Remark 4.3 but no suitable generalization is available for the more explicit estimate (4.17).

Clearly, just as Lemma 4.4 gives information about $\psi_1(X_k)$ one would like information about the best one one might hope for as to $\{\nu_k\}$.

LEMMA 4.6: *Let X, Y be Hilbert spaces and $f: X\to Y$ compact and linear. Let f_S be the restriction of f to the (finite-dimensional) subspace $S\subset X$. Let $\{u_j\}$ be the eigenvectors (taken orthonormal) and let $\{\omega_j\}$ be the corresponding eigenvalues of the positive operator f^*f, ordered so $\omega_1\geqslant\omega_2\geqslant\dots>0$. Then*

$$\min\{\|f_S^{-1}\|:\dim S=k\}=\omega_k^{-1/2}\tag{4.18}$$

with the minimum attained for $S:=\mathrm{sp}\{u_1,\dots,u_k\}$.

Proof: For dim $S=k$ there is always a nontrivial solution $x_S\in S$ of the system: $\langle u_j,x\rangle=0$ for $j=1,\dots,k-1$. Then $x_S=\sum_{j=k}^{\infty}\xi^j u_j$ and, setting $y:=fx_S=f_S x_S$,

$$\|y\|^2=\langle x_S,f^*fx_S\rangle=\sum_{j=k}^{\infty}\omega_j|\xi^j|^2\leqslant\omega_k\sum_{j=k}^{\infty}|\xi^j|^2=\omega_k\|x_S\|^2\leqslant\omega_k\|f_S^{-1}\|^2\|y\|^2$$

so $\|f_S^{-1}\|^2\geqslant 1/\omega_k$ for any k-dimensional S. Taking $S:=\mathrm{sp}\{u_1,\dots,u_k\}$, one has, for any $x=\sum_{j=1}^{k}\xi^j u_j\in S$, $y=fx=f_S x$,

$$\|y\|^2=\sum_{j=1}^{n}\omega_j|\xi^j|^2\geqslant\omega_k\sum_{j=1}^{k}|\xi^j|^2=\omega_k\|f_S^{-1}y\|^2$$

so $\|f_S^{-1}\|^2 \leqslant 1/\omega_k$ for this S (and *only* for this S among k-dimensional subspaces provided $\omega_{k+1} \neq \omega_k$). ■

REMARK 4.7: If one were just given a single approximant $\tilde{b}$ to b (for simplicity we assume $\tilde{f}=f$ available exactly) and the information

$$\|\tilde{b}-b\| \leqslant \varepsilon, \quad \|x_*\|_R \leqslant M, \tag{4.19}$$

then one might wish to select k and $\{\eta_1, ..., \eta_k\}$ so that (4.12) would give an *optimal estimate* for the error $\|x-x_*\|$. If possible, one would like $\psi_M(X_k)$ and ν_k to attain simultaneously the minima given by Lemmas 4.4 and 4.6 which would reduce (4.12) to

$$\|x_* - \tilde{x}\| \leqslant [M^2 \rho_k + \varepsilon^2/\omega_k]^{1/2} \tag{4.20}$$

(this assumes $\{\eta_1, ..., \eta_k\}$ taken orthonormal so $\|H_k\|=1$), permitting an optimal choice of k in terms of ε, M — although the information on $\{\rho_k, \omega_k\}$ typically available is likely, at best, to make this choice and the associated estimate (4.20) primarily asymptotic. For $\psi_M(X_k)$ and ν_k to be simultaneously minimizable one must have

$$\mathrm{sp}\,\{r_1, ..., r_k\} = \mathrm{sp}\,\{u_1, ..., u_k\} \subset \text{range of } f^* \tag{4.21}$$

so that one can take $\{\eta_1, ..., \eta_k\}$ (orthonormal) so that

$$X_k := \mathrm{sp}\,\{f^*\eta_1, ..., f^*\eta_k\} = \mathrm{sp}\,\{r_1, ..., r_k\} = \mathrm{sp}\,\{u_1, ..., u_k\}.$$

Even if (4.21) does not hold, so (4.20) does not hold exactly, it is often possible (especially on renorming the regularity space R with an equivalent Hilbert space norm which modifies the sequence $\{\rho_j\}$) to have an estimate of the form (4.20) holding asymptotically. Then the choice of k (in terms of the rate of convergence to b of a sequence of approximants) can again be made asymptotically optimal.

This use of (4.12) is complementary to the comment following (4.17). In each case, however, the use of (4.12) to bound the approximation error $\|x_* - \tilde{x}\|$ reflects the balance of the considerations involved in bounding the projection error $\|x_* + -\tilde{P}x_*\|$ and the sensitivity to the uncertainty in approximating b, f by $\tilde{b}$, $\tilde{f}$. This is especially clear in using (4.20) to make an optimal choice of k.