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

Convergent approximation methods for ill-posed problems.
Part II — Applications

by

THOMAS I. SEIDMAN
University of Maryland,
Baltimore County, USA and
Université de Nice, France

5. Nemytsky Operators

Let Ω be a bounded domain in R^m, let $X:=L^p(\Omega)$ with $1<p<\infty$ and $Y:=L^{p'}(\Omega)$. Let $\varphi: \Omega\times R\to R$ satisfy Carathéodory conditions and a growth condition of the form

$$|\varphi(t,r)|\leqslant c(t)\,[a(t)+|r|]^{\theta} \quad \text{for} \quad t\in\Omega,\ r\in R$$
$$\text{with } a\in L^p(\Omega) \text{ and } c\in L^{\gamma}(\Omega) \text{ where } \frac{1}{\gamma}+\frac{\theta}{p}=\frac{1}{p'}. \tag{5.1}$$

Then the *Nemytsky operator* f defined by φ:

$$[f(x)](t):=\varphi(t,x(t)) \tag{5.2}$$

is known [39] to be well-defined and continuous from X to Y.

Suppose, for the time being, we assume that,

$$\text{for a.e. } t\in\Omega,\ \varphi(t,\cdot) \text{ is strictly increasing}. \tag{5.3}$$

Then the inverse function $\psi(t,\cdot)$ is well-defined; ψ will also satisfy Carathéodory conditions and we may define a Nemytsky operator g by $[g(y)](t):=\psi(t,y(t))$ for suitable y. In this case f will be injective but, unless ψ happens to satisfy appropriate growth conditions, meaning that φ grows rapidly enough, the range of f will be dense but not closed, the inverse map $f^{-1}=g$ will not be continuous and the "inverse substitution" problem

$$f(x)=b \tag{5.4}$$

will be ill-posed. (An explicit example might be: $[f(x)](t)=[x(t)]^{1/3}$ with $p=p'=2$. The range of f is clearly dense and the inverse operator: $y\mapsto y^3$ is not continuous.)

The inversion (5.4) is, of course, trivial if φ, b are exactly available but otherwise our earlier considerations apply.

REMARK 5.1: Observe, first, that Y cannot be appropriately topologized as Y_w if the problem is actually nonlinear (i.e., unless φ has the affine form: $\varphi(t, r)= =a+rb$ with $a \in L^{p'}$ and $b \in L^s$; $s=pp'/(p-p')$, $p \geqslant p' > 1$) since
The graph of, given by (5.2), *is not closed in* $X_w \times Y_w$ *if* φ *is nonlinear.*

Proof: If φ were not of the form $a+rb$, then there would be a set $\Omega_0 \subset \Omega$ of positive measure and (rational) constants r, d, ε ($d, \varepsilon > 0$) such that either one has

$$[\varphi(t, r+d)+\varphi(t, r-d)-2\varphi(t, r)] > 5\varepsilon \quad \text{for} \quad t \in \Omega_0$$

or one has

$$[\varphi(t, r+d)+\varphi(t, r-d)-2\varphi(t, r)] < -5\varepsilon \quad \text{for} \quad t \in \Omega_0$$

(for definiteness, take: $>5\varepsilon$). There would then be a subset $\Omega_* \subset \Omega_0$ of positive measure $2\mu^*$ and (rational) constants α, β, γ such that for each $t \in \Omega_*$

$$\alpha \leqslant \varphi(t, r+d) < \alpha+\varepsilon, \ \beta \leqslant \varphi(t, r-d) < \beta+\varepsilon, \ \gamma-\varepsilon \leqslant \varphi(t, r) < \gamma.$$

By the definition of $\Omega_0 \supset \Omega_*$, one has $(\alpha+\beta-2\gamma) > \varepsilon$. Now there exists a sequence $\{u_j: j=0, 1, ...\}$, orthogonal in $L^2(\Omega)$, with each $u_j=0$ on $\Omega \setminus \Omega_*$ and taking only the values ± 1 on Ω_* ($u_0 \equiv 1$ on Ω_*). For $j=1, 2, ...$, let $x_j := r+du_j$. Observe that $x_j \to r$ in $X := L^p(\Omega)$ (regardless of $1 < p < \infty$) and that $\{y_j := f(x_j)\}$ is bounded in Y so, passing to a subsequence if necessary, $y_j \to \hat{y}$ in Y. For $j=1, 2, ...$, let $\Omega_j^{\pm} := \{t \in \Omega_*: \ u_j(t) = \pm 1\}$ and note that each $\Omega_j^{\pm}$ has measure μ_* since $\langle u_0, u_j \rangle = 0$. Thus,

$$\langle 1, y_j \rangle = \int_{\Omega \setminus \Omega^*} \varphi(\cdot, r) + \int_{\Omega_j^+} \varphi(\cdot, r+d) + \int_{\Omega_j^-} \varphi(\cdot, r-d) \geqslant \int_{\Omega} \varphi(\cdot, r) +$$

$$+ \int_{\Omega_j^+} [\alpha-\gamma] + \int_{\Omega_j^-} [\beta-\gamma] = \langle 1, f(r) \rangle + \mu_* (\alpha+\beta-2\gamma).$$

Thus, $\hat{y} := w-\lim f(x_j) \neq f(r) = f(w-\lim x_j)$ and the graph of f is not closed in $X_w \times Y_w$. ∎

Suppose X, Y, φ are as above (with φ satisfying (5.1), (5.3)) and define a Nemytsky operator f by (5.2). Approximate f by a sequence of Nemytsky operators $\{f_k\}$ defined as in (5.2) using $\{\varphi_k: \Omega \times \boldsymbol{R} \to \boldsymbol{R}\}$ satisfying Carathéodory conditions and

$$|\varphi_k(t, r)-\varphi(t, r)| \leqslant c_k(t)(a_k(t)+|r|)^{\theta_k} \quad \text{for} \quad t \in \Omega,$$

$$r \in \boldsymbol{R} \text{ with } a_k \in L^p(\Omega) \text{ and } c_k \in L^{\gamma_k}(\Omega) \text{ where } \frac{1}{\gamma_k}+\frac{\theta_k}{p}=\frac{1}{p'}. \tag{5.5}$$

It is easy to see that (5.1), (5.5) ensure that each f_k is well-defined and continuous from X to Y. Let b be in the range of f and set $x_*(t) := \psi(t, b(t))$. Clearly this x_* is the unique solution of (5.4). For some $M > \|x_*\|$, set

$$\varepsilon_k := \|c_k\| (\|a_k\|+M)^{\theta_k}.$$

THEOREM 5.2: *Let X, Y, φ, f, $\{\varphi_k\}$, $\{f_k\}$, b, x_*, $\{\varepsilon_k\}$ be as above and let $\{b_k\}$ be in Y with $\varepsilon_k' \geqslant \|b_k - b\|$. Then*

$$\|f_k(x) - f(x)\| \leqslant \varepsilon_k \text{ uniformly on } D_M := \{x \in X : \|x\| \leqslant M\}. \tag{5.6}$$

and, if $\varepsilon_k \to 0$, then $\{f_k\}$ is graph subconvergent to f (with f_k, f restricted to D_M). Let

$$\begin{aligned} x_k \in S_k &:= \{x \in X : \|f_k(x) - b_k\| \leqslant \varepsilon_k + \varepsilon_k'\}, \\ \|x_k\| &\leqslant \inf\{\|x\| : x \in S_k\} + \delta_k. \end{aligned} \tag{5.7}$$

Then, if $\varepsilon_k' \to 0$, $\varepsilon_k \to 0$ and $\delta_k \to 0+$, one has $x_k \to x_*$.

Proof: (5.6) follows immediately from (5.5). Suppose that $\tilde{x}_k \to \tilde{x}$ with $\|\tilde{x}_k\| \leqslant M$ and that $\tilde{y}_k := f_k(\tilde{x}_k) \to \tilde{y}$ in Y. Set $\hat{y}_k := f(\tilde{x}_k)$ and observe that (5.6) gives $\hat{y}_k \to \tilde{y}$. Thus, passing to a subsequence if necessary $\hat{y}_k(t) \to \tilde{y}(t)$ for a.e. $t \in \Omega$. By (5.3), $\psi(t, \cdot)$ is continuous so $\tilde{x}_k(t) = \psi(t, \hat{y}_k(t)) \to \psi(t, \tilde{y}(t)) =: \hat{x}(t)$; clearly $f(\hat{x}) = \tilde{y}$. Since $\{x_k\}$ converges weakly to $\tilde{x}$ and pointwise a.e. to $\hat{x}$, one has $\tilde{x} = \hat{x}$ so $f(\tilde{x}) = \tilde{y}$ which proves the graph subconvergence. If $\varepsilon_k \to 0$, $\varepsilon_k' \to 0$, then $y \in B_k := \{y \in Y : \|y - b_k\| \leqslant \varepsilon_k + \varepsilon_k'\}$ implies $\|y - b\| \leqslant \varepsilon_k + 2\varepsilon_k'$ so $B_k \downarrow \{b\}$. By (5.6), one has $x_* \in S_k$ for each k so Theorem 2.3 applies to give $x_k \to x_*$. ∎

REMARKS 5.3: We have so far considered φ defined on all of $\Omega \times \boldsymbol{R}$ so the domain of f is all of X. A frequently occurring variant of this is to have φ defined on $\Omega \times \boldsymbol{R}^+$ so the domain of f is

$$X^+ := \{x \in X : x(t) \geqslant 0 \text{ for a.e. } t \in \Omega\}$$

(more generally, $D(\varphi) := \{(t, r) \in \Omega \times \boldsymbol{R} : r_1(t) \leqslant t \leqslant r_2(t)\}$ and $D(f) := \{x \in X : (t, x(t)) \in D(\varphi) \text{ a.e.}\}$) and the proof above can easily be modified to treat this.

Suppose, next, that φ does not satisfy (5.3) and f is not monotone and not injective. Then an argument silmilar to the proof in Remark 5.1 shows that the graph of f will not be closed in $X_w \times Y$ and that $\{f_k\}$ is not graph subconvergent in $X_w \times Y$. Nevertheless, (5.3) can be replaced by the weaker condition

$$\begin{aligned} &\text{for a.e. } t \in \Omega \text{: for each } s \in \boldsymbol{R} \text{ either } \Psi(t, s) := \{r : \varphi(t, r) = s\} \text{ is empty or} \\ &\qquad \text{it contains a unique } r_* =: \psi(t, s) \text{ of minimum absolute value} \end{aligned} \tag{5.8}$$

and still have the final conclusion of Theorem 5.2: that $x_k \to x_*$.

First note that x_* given by $x_*(t) := \psi(t, b(t))$ will clearly be the unique minimum norm solution of (5.4) — provided it is measurable so $x_* \in X$. We note the following.

LEMMA : *Let φ satisfy Carathéodory conditions and* (5.8).
Define $\psi(t, s)$ by (5.8), *setting $\psi(t, s) := a$ (arbitrary, say $a = 0$) when $\Psi(t, s) = \emptyset$.*
Then $\psi(\cdot, y(\cdot))$ is measurable for every measurable $y(\cdot)$.

Proof: For measurable u, v, let $u \bigcirc v := \chi_S u + (1 - \chi_S) v$ where $S := \{t : |v| \geqslant |u|\}$ so $u \bigcirc v$ is measurable. Now $\psi = \psi_+ \bigcirc \psi_-$ where $\psi_+(t, s) := \inf\{r \geqslant 0 : \varphi(t, r) = s\}$,

$\psi_-(t,s):=\sup\{r\leqslant 0,\ \varphi(t,r)=s\}$ (setting these 0 if the sets are empty). Suppose $y(\cdot)$ is measurable. For any (positive) α, using the Carathéodory conditions for φ,

$$\{t:\psi_+(t,y(t))<\alpha\}=\bigcap_n \bigcup_{\substack{r\in Q\\ 0\leqslant r<a}} \{t:|\varphi(t,r)-y(t)|<1/n\}.$$

Thus $\psi_+(\cdot,y)$ is measurable and, similarly, $\psi_-(\cdot,y)$ is measurable. ∎

To show that $\{x_k\}$, given by (5.7), etc., converges to x_* requires a modification of the proof of Theorem 2.3 to show that the particular nature of $\{x_k\}$ implies that if (a subsequence of) it converges weakly to some $\hat{x}$, then $f(\hat{x})=b$. We consider only the case in which φ, φ_k are defined on $\Omega\times R^+$ so f, f_k are defined on X^+ and all the functions x_k, x_*, etc., are nonnegative.

Assume we have $x_{k(i)}\to\hat{x}$. We modify the notation so as to continue to write simply $\{x_k\}$ as we select this and further subsequences. Let $y_k:=f_k(x_k)$ and $\hat{y}_k:=f(x_k)$. As with Theorem 5.2, $\hat{y}_k\to b$ and we may assume the convergence is also pointwise a.e. Define $\hat{x}_k$ by $\hat{x}_k(t):=\psi_+(t,\hat{y}_k(t))$ so also $f(\hat{x}_k)=y_k$. Since $\varphi(t,\cdot)$ is continuous (for a.e. $t\in\Omega$), we have $\psi_+(t,\cdot)$ lower semicontinuous (if $\hat{\varphi}:R^+\to R$ is continuous, $\hat{\psi}(s):=\min\{r:\hat{\varphi}(r)=s\}$, $s_k\to s$, $r_k:=\hat{\psi}(s_k)$, $\hat{r}:=\liminf r_k$, then, for a subsequence, $r_{k(i)}\to\hat{r}$, $s_{k(i)}=\hat{\varphi}(r_{k(i)})\to\hat{\varphi}(\hat{r})$ and $\hat{\varphi}(\hat{r})=s$ so $\hat{\psi}(s)\leqslant\hat{r}$). Thus, as $\hat{y}_k(t)\to b(t)$ one has

$$\hat{x}_*(t):=\liminf \hat{x}_k(t)\geqslant\psi_+(t,b(t))=x_*(t). \tag{5.9}$$

Note that (5.9) implies $\hat{x}(t)\geqslant\hat{x}_*(t)\geqslant x_*(t)$ a.e. However, as with Theorem 5.2, (5.7) implies that

$$\|\hat{x}\|\leqslant\liminf\|\hat{x}_n\|\leqslant\liminf\|x_k\|\leqslant x_*+\lim\delta_k=\|x_*\|$$

so $\hat{x}(t)=x_*(t)$ a.e. The remainder of the proof that $x_k\to x_*$ is as in the proof of Theorem 2.3. ∎

The case in which φ, φ_k are defined on all of $\Omega\times R$ can be treated similarly if one assumes that, for a.e. $t\in\Omega$,

$$r[\varphi(t,r)-\varphi(t,0] \text{ } \textit{has fixed sign for } r\in R \tag{5.10}$$

(as well as (5.8)). One can then proceed, analogously to the above, after dissecting Ω according to the sign of $x_*(t)$.

One can also treat multivariate Nemytsky operators ($\varphi:\Omega\times R^\mu\to R^\nu$) under appropriate hypotheses but the conditions and treatment become much more complicated unless, generalizing (5.3), one were to require, e.g., that for a.e. $t\in\Omega$, the map $\varphi(t,\cdot):R^\mu\to R^\mu$ ($\nu=\mu$) should be strictly monotone so $\varphi(t,\cdot)$ would have a continuous inverse.

6. Recovery of a Diffused Signal

Turning now to a more concrete application, consider (thermal) diffusion in a rod ($0\leqslant s\leqslant 1$) with diffusion coefficient D so

$$u=(Du_s)_s \quad t>0,\ 0<s<1. \tag{6.1}$$

For present purposes, we refer to the end $s=0$ as the *near* or *accessible* end and to $s=1$ as the *far* or *inaccessible* end. It is assumed that the value at the far end fluctuates with time:

$$u(t, 1)=x(t) \quad t\geqslant 0 \tag{6.2}$$

and we consider this unknown fluctuation $x(\cdot)$ as a *signal.* Assume that the near end is insulated.

$$u_x(t, 0)=0 \quad t\geqslant 0, \tag{6.3}$$

but that it is possible to observe the *transmitted signal*:

$$b(t):=u(t, 0) \quad t\geqslant 0 \tag{6.4}$$

which has been diffusively propagated along the rod. The problem is to recover the original signal x (for $0\leqslant t\leqslant T$) from the observed b (also for $0\leqslant t\leqslant T$). We treat this in somewhat less detail than in Section 5 as a more complete exposition has appeared in [65]. Compare with the treatments in [55], [7], obtaining u on $0\leqslant s\leqslant \leqslant s_0<1$ from b, given *a priori* bounds on x.

Before discussing any specific computational approaches to the problem, we first verify the condition (2.5). Actually we will show that exact knowledge of the observable $b(\cdot)$ uniquely determines the solution u of the Cauchy problem (6.1), (6.3), (6.4).

THEOREM 6.1: *Suppose the diffusion coefficient* $D=D(s, \omega)$ *satisfies*

$$\begin{aligned} &\text{(i)} \quad 0<d_1\leqslant D(s, \omega)\leqslant d_2 \text{ on } [0,1]\times \boldsymbol{R} \text{ and} \\ &\text{(ii)} \quad D, \text{ extended as an even function of } s, \text{ is analytic on } (-1,1)\times \boldsymbol{R}. \end{aligned} \tag{6.5}$$

Then the Cauchy problem (6.1), (6.3), (6.4) *has at most one solution* u *on* $[0, T]\times[0, 1]$.

Proof: Let w on $[0,T]\times[-1,1]$ be u reflected as an even function of s. Then w is a weak solution of the equation

$$w_t=(D(\cdot,w)\, w_s)_s \tag{6.6}$$

So, by "interior regularity", w is analytic in $s\in(-1,1)$ for each fixed $t\in(0,T]$ and C^∞ on $(0,T]\times(-1,1)$ (in particular, this means that $b:=u(\cdot,0)=:w(\cdot,0)$ must be C^∞. Repeated differentiation of the differential equation $u_t=(Du_s)_s$ and use of the facts that $w_s(\cdot,0)=0$ (as w is even) and that $w(\cdot,0)=b$ now enable one to determine the coefficients of the power series expansion of $w(t,\cdot)$ around $s=0$, for each fixed $t\in(0,T]$, in terms of derivatives of b at t and the coefficients of the power series expansion of D around $s=0$, $\omega=b(t)$. Thus, w is uniquely determined by b, D on $\{t\}\times(-1,1)$ and so on $(0,T]\times(-1,1)$ as t varies. Taking traces, x and w_s are also uniquely determined and u is unique on $[0,T]\times[0,1]$. ■

REMARKS 6.2: The argument above could clearly be applied if (6.1) were replaced by the more general equation

$$u_t=(D(\cdot,u)\, u_s)_s-f(s, u, u_s) \tag{6.7}$$

under appropriate analyticity hypotheses. If one were to know *a priori* that u takes values in an interval $J \subset \boldsymbol{R}$ (e.g., if x, u_0 are known to be J-valued and the maximum principle can be applied), then (6.5) need hold only for $\omega \in J$. This is of particular relevance for nonlinear diffusion problems in which u represents a concentration whence, on physical grounds, it is known that $u \geqslant 0$ ($J := \boldsymbol{R}^+$). The argument above can also be used in connection with the two dimensional heat equation $u_t = \Delta u$ in an annulus with radial symmetry, in which case inversion ($r \leftrightarrow R^2/r$) would replace reflection ($s \leftrightarrow -s$) in defining "extension as an even function". In the *linear* case one might consider the Cauchy problem

$$u_t = (Du_s)_s - cu + f, \quad u(\cdot,0) = b, \quad u_s(\cdot,0) = b, \tag{6.8}$$

with D as in (6.5) and c also even and analytic on $(-1,1)$ since a reduction to the case $f=0$, $b_1=0$ is obtainable by subtracting the solution of

$$u_t = (Du_s)_s - cu + f, \quad u_s(\cdot,0) = b_1, \quad u(\cdot,1) = 0, \quad u(0,\cdot) = 0. \tag{6.9}$$

Such a reduction is not possible in the nonlinear case but the necessity that $b_1 = 0$ came from use of the "reflection trick" whose only purpose was to establish analyticity in s at $s=0$ for $u(t,\cdot)$. If, instead of a Cauchy problem one considers an *extrapolation problem* (cf. [29]): u satisfying (6.1) on $(0,T] \times (\alpha,\gamma)$ with observations available of $u(0,\cdot)$, $u(\cdot,\alpha)$, $u(\cdot,\beta)$ and with $u(\cdot,\gamma)$ desired ($\alpha < \beta < \gamma$), then u is analytic across $s=\beta$ (now in the interior) so uniqueness for the standard problem on $(0,T] \times (\alpha,\beta)$ extends to $(0,T] \times (\alpha,\gamma)$ whence, taking the trace, one has uniqueness of $x := u(\cdot,\gamma)$. ∎

For the linear problem (6.1)–(6.4) one may consider this as defining a linear map $(x, u_0) \mapsto b$. For simplicity we assume

$$u_0 := u(0,\cdot) = 0 \quad \textit{on} \quad [0,1] \tag{6.10}$$

and seek only to invert $f: x \mapsto b$. For "reasonable" spaces X, Y, this problem will be ill-posed since, as noted above, we have b always C^∞ for "arbitrary" x, making f compact.

Approach 6.3: Consider, first, $X := L^2(0,T)$ weighted by $D(1)$ and $Y := L^2(0,T)$ weighted by $D(0)$. Use the method of generalized interpolation. Given any total sequence $\{y_j\}$ of functions in $Y^* = Y$, solve the problems ($j = 1, 2, \ldots$):

$$-v_t = (Dv_s)_s, \quad v(T,\cdot) = 0, \quad v_s(\cdot,0) = y_k, \quad v(\cdot,1) = 0 \tag{6.11}$$

to obtain $\{v^j\}$ and set

$$u_j = v_s^j(\cdot,1), \tag{6.12}$$

$$a^{i,j} := \langle u_i, u_j \rangle := \int_0^T D(1)\, u_i\, u_j. \tag{6.13}$$

Given b, compute

$$\beta^j := \langle b, y_j \rangle := \int_0^T D(0)\, b y_j \quad j = 1, 2, \ldots \tag{6.14}$$

and for $k=1, 2, \ldots$, solve the system

$$\sum_{j=1}^{k} a^{i,j}\xi^j=\beta^i \quad i=1, \ldots, k \tag{6.15}$$

to obtain an approximant

$$x_k:=\sum_{j=1}^{k}\xi^j u_j \tag{6.16}$$

to x. It is easy to verify that $f: x\mapsto b$ is continuous from X to Y and that (6.11), (6.12) gives $y_j\mapsto u_j=f_* y_j$ so the construction (6.11)–(6.16) of x_k is just that of Theorem 4.2 which now proves convergence, $x_k\to x_*$. Suppose one would take $\{(u_k, \omega_k)\}$ to be the eigenvalues and (normalized) eigenvectors of $f_* f$ (so $y_j=fu_j/\omega_j$ for $j=1, 2, \ldots$). Then $((a^{ij}))$ is the identity matrix and (6.16) gives $x_k=\sum_{j=1}^{k}\beta^j u_j$ as the truncation of the orthonormal expansion of x. Further, (4.12) gives

$$\|x-\tilde{x}_k\|^2\leqslant\|x-sp\{u_1, \ldots, u_k\}\|^2+\|b-\tilde{b}_k\|^2/\omega_k \tag{6.17}$$

if $\tilde{x}_k$ is computed as above but using an approximant $\tilde{b}_k$ rather than b in computing (6.14).

Approach 6.4: An alternative, and frequently useful, viewpoint for a problem such as this is to consider (6.1), (6.2), (6.3), (6.10) as a *control problem*: determine a *control* $\hat{x}$ which is optimal in the sense of minimizing some suitable *cost criterion* J which forces the *output* $\hat{y}:=u(\cdot, 0)$ (corresponding to use of $\hat{x}$ in (6.2)) to match the observed b. As above, we consider this in the context of approximating b by $\tilde{b}=\tilde{b}_k$ ($\tilde{b}_k\to b$) and, now, also of approximating the map f by maps f_k defined by constructing a convergent sequence of discretizations for (6.1) using, e.g., the method of lines.

The construction of such discretizations for (6.1), whether in the linear or in the nonlinear case, is a standard problem in numerical analysis. For purposes of exposition we assume X can be taken to be a sufficiently high order Sobolev space $H^\nu[0,T]$ (actually, for suitable consistency with (6.10), take X to be the space of functions in $H^\nu(-\infty, T]$ vanishing on $\boldsymbol{R}^-$) to ensure that solutions are smooth enough for application of, e.g., Theorem 7.1 of [17] for data x in $\tilde{X}:=H^{\nu-\varepsilon}(0,T)$ and with continuous dependence of the solution u (with respect to suitable norms) on x in $\tilde{X}$. In the linear case the choice of ν poses no problems, at least if D is smooth up to the boundary $s=1$; in the nonlinear case additional care and restrictions on D are required to ensure the continuous dependence. For example, note that for the "standard" heat equation $u_t=u_{ss}$ and discretization of $\partial^2/\partial s^2$ using Galerkin's method with piecewise cubic splines it is adequate to take $\nu=2$ so $X=H^2[0,T]$.) In such a situation the maps f_k will converge to the continuous map $f: X_w\to Y:=L^2(0,T)$ uniformly on bounded sets so $\{f_k\}$ is graph subconvergent to f in $X_w\times Y$.

Discretization of (6.1) by the method of lines produces a system of ordinary differential equations. Suppose we apply to this the penalty function approach of Section 3 (e.g., (3.6) with $p=p'=2$). Then Theorem 3.1 guarantees convergence to the true input x of the sequence $\{x_k\}$ of approximants. Here, each x_k is defined as the *optimal control* for the ODE system giving f_k with optimality defined as minimization of the *quadratic cost criterion*

$$J_k(x)=\|x\|_x^2+\lambda_k\int_0^T[\tilde{b}_k(t)-U_k(t,0)]^2\,dt \tag{6.18}$$

where U_k is the approximate solution of (6.1), (6.2), (6.3), (6.10) computed using the k-th discretization (i.e., $U_k(\cdot,0)=f_k(x)$). In the linear case the computation of x_k is then a standard problem in optimal control theory (cf., e.g., [4]). (The usual algorithm involves solution of a matrix Riccati equation and applying the result to $\tilde{b}_k$. There exists, however, a computationally more efficient "Chandrasekhar" algorithm (see, e.g., [10]) which may apply.)

APPROACH 6.5: For the nonlinear case it is somewhat more difficult to obtain continuity for the map:
$x\mapsto u(\cdot,0)\colon H^\nu[0,T]\to L^2(0,T)$ used in the approaches above. An interesting variant is to introduce a new map

$$f\colon u\mapsto(u(\cdot,0),\ [u_t-(D(u)u_s)_s]) \tag{6.19}$$

taking functions on $Q:=[0,T]\times[0,1]$ into pairs of functions (on $[0,T]$, on Q). In terms of this map we now seek a solution u of

$$f(u)=(b,0), \tag{6.20}$$

after which we can take the trace on $[0,T]\times\{1\}$ to recover $x:=u(\cdot,1)$. Note that while Theorem 6.1 does not give (2.5′) for this f, it does give uniqueness for right hand sides of the form $(b,0)$ — which is all that is required for the problem at hand.

To specify f as a map: $X\to Y$, one must adjoin to (6.19) a specification of the spaces X, Y. Take, for example,

$$X:=H^2(Q),\ Y:=Y_0\times Y_1:=L^2(0,T)\times H^{-1}(Q) \tag{6.21}$$

or, alternatively, with $2<p<\infty$ take

$$X:=W^{2,p}(Q),\ Y:=L^2(0,T)\times L^2(Q). \tag{6.21′}$$

Either of (6.21), (6.21′) gives continuity (indeed, compactness) of f from X_w to Y (*a fortiori*, to Y_w). Use of (6.21) gives $x\in H^{3/2}[0,T]$ on taking the trace and use of (6.21′) gives $x\in W^{2-1/p,p}[0,T]\subset H^{3/2}[0,T]$. Define approximants f_k to f by

$$f_k\colon u\mapsto(u(\cdot,0),\ [\tilde{u}_t-(D(\tilde{u})\tilde{u}_s)_s]) \tag{6.19′}$$

where, for each k and each $t\in[0,T]$, $\tilde{u}(t,\cdot)$ is the $H^2[0,1]-$orthogonal projection of $u(t,\cdot)$ on the space Sp^k of piecewise cubic splines with nodes $\{j/k\colon j=0,\ldots,k\}$. As with the previous approach, we have Y—convergence of $\{f_k\}$ to f uniformly on

bounded sets and so graph subconvergence of $\{f_k\}$ to f. Given a sequence $\{\tilde{b}_k\}$ of approximants to b one can then use estimates for $\{\|\tilde{b}_k - b\|\}$ and the convergence rate information for the discretization used in (6.19′) as applied to the solution u_* of (6.20) (i.e., one can make use of the known smoothness of solutions of (6.21)) to construct sets $B_k \downarrow \{(b,0)\}$ in Y so $f_k(u_*) \in B_k$ for each k. Then Theorem 2.3 can be used to give a convergent sequence of approximants $\{u_k\}$ to u_* and so to the desired input. Note that the minimization of the $H^2(Q)$—norm required by (2.8) for determining each u_k guarantees, by the form of (6.19′) that u_k will be in $H^2([0,T] \to Sp^k)$ so the computation can be done there. As in the previous approach, this can usefully be regarded as a control problem. In the nonlinear case, however, there is no such simple computational algorithm as noted there; the computational aspect will be discussed at greater length in [65]. If one were to use a suitable full space-time discretization of $u_t = (D(u)u_s)_s$, rather than the method of lines, then the entire minimization computation would take place in a finite-dimensional space as a nonlinear programming problem.

7. System Identification for Distributed Problems

Distributed parameter system theory is a particularly fertile source of ill-posed problems of practical significance. Problems of the determination of an inhomogeneity (= "forcing term", either in the equation or boundary conditions) may be referred to in system theoretic language as *signal recovery*, as in the preceding section (related ill-posed problems arise in integral geometry [42], as in X—ray tomography, or in the "sharpening" of fuzzy images [8], [28]). Problems of the determination of coefficients (again, either in the equation or boundary conditions; this may also include geometric considerations such as determination of the "affected" region Ω) are "classically" referred to as *inverse problems* but in system theoretic language are called *system identification* (see, e.g., [2], [23], [11], [24]).

A standard approach to system identification (cf., e.g., [20]) is least squares estimation via *history matching*: determine the coefficients so as to minimize a cost criterion for the deviation of the resulting predictions (assuming the initial state is known or estimable) from the observations. We diagram this, "quoting" Figure 1 of [12], as follows:

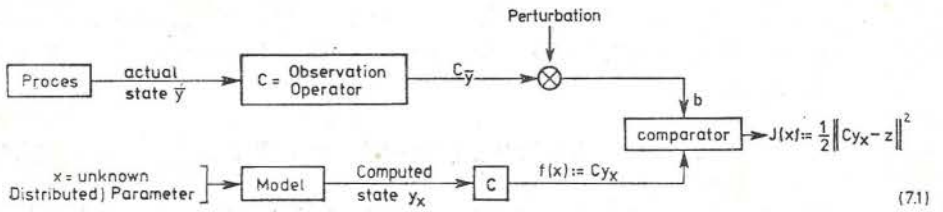

Fig. 1

For computational purposes in distributed problems this is used in a formulation involving a suitable (approximating) parametric representation for the unknown x (e.g., a discretization or spline approximation or a truncated series expansion). One then calculates, in a finite dimensional setting, the parameter values which minimize the resulting "cost". The term "*lumped parameter* system theory" may be taken to indicate the (implicit) use of such a formulation.

Thus the original operator equation

$$f(x)=b \tag{7.2}$$

is replaced by the variational problem (nonlinear least squares fit):

$$\textit{minimize } J(x):=\frac{1}{2}\|f(x)-b\|^2 \textit{ over } x\in\tilde{X}:=\{\tilde{x}(\lambda):\ \lambda\in\tilde{\Lambda}\} \tag{7.3}$$

which is treated computationally by minimizing $J(\lambda):=\tilde{J}(\tilde{x}(\lambda))$ over $\tilde{\Lambda}$. Here $\tilde{\Lambda}$ is a (finite-dimensional) parameter space associated with the parametric representation $\lambda\mapsto\tilde{x}(\lambda)$, b is the observational data and $x\mapsto f(x)$ is the mapping defined using the *direct* problem (i.e., the model) specified by x to predict the observations. (It is often fruitful to view the determination of x — or adaptation of the model — as an optimal control problem.) Much of the analytical effort is then devoted to showing that, for the resulting *approximate* problem, the map $\lambda\mapsto f(\tilde{x}(\lambda))$ is well-defined and suitably smooth on $\tilde{\Lambda}$ and that (7.3) has a unique solution which can be obtained by a feasible computational procedure.

Often the results of simulation "experiments" or comparisons with real data are presented to suggest the merits of the procedure. In line with our present considerations we view this as embedded in a *computational scheme* in which the above is taken as one of a family of parametrizations of increasing dimensionality. If, as is typical, $\lambda\mapsto\tilde{x}(\lambda)$ is linear, then $\tilde{X}$ is a subspace and may be taken to parametrize itself so we consider $\tilde{X}=X_k$ where $\{X_k\}$ is an increasing sequence of subspaces becoming dense in X so that the desired solution x_* is potentially approximable with arbitrary accuracy. We then ask whether the sequence $\{x_k\}$ of approximations obtained by (7.3) using X_k for $\tilde{X}$ and b_k for b (with $b_k\to b$) converges to x_*. (For well-posed problems this approach (parametric reduction with least squares fit) converges and is an effective tool of numerical analysis; see, e.g., [3].) An analysis in the linear case [64] shows that in general *one cannot expect convergence from this procedure applied to ill-posed problems* (but note [56]). More precisely, one has the following from [64].

THEOREM 7.1: *Let X, Y be Hilbert spaces and $A: X\to Y$ linear, injective and compact. Let b be in the range of A so $b=Ax_*$. For b_k in Y and subspaces X_k of X, determine x_k in X_k by*

$$\textit{minimize } J_k(x):=\frac{1}{2}\|Ax-b_k\|^2 \textit{ over } X_k. \tag{7.4}$$

then, given any $\{X_k\}$ becoming dense in X there exist $\{b_k\}$ in Y with $b_k\to b$ but $\{x_k\}$ divergent (one can force $\|x_k\|\to\infty$ or $\{x_k\}$ bounded but $x_k\not\to x_*$). Even using exact

data ($b_k=b$ for each k), there is a sequence $\{X_k\}$ as above for which $\{x_k\}$ is unbounded (provided b is not a finite linear combination of eigenvectors of AA^*. ■

Note that the method of generalized interpolation is superficially similar but with quite different results (cf., Theorems 4.2, 4.5). In that case one assumes that each X_k is in the range of A^* (i.e., $X_k=A^* Y_k$) and, instead of (7.4), x_k satisfies

$$\text{minimize } J_k(x):=\frac{1}{2}\|\hat{P}_k(Ax-b_k)\|^2 \text{ over } X_k \tag{7.5}$$

where $\hat{P}_k$ is the orthoprojection onto Y_k. An alternate comparison is that (4.5) means

$$x_k \in X_k, \quad \hat{P}_k A x_k=\hat{P}_k b_k \tag{7.5'}$$

(this means $J_k(x_k)=0$ in (7.5), which is certainly the minimum) while (7.4) means

$$x_k \in X_k, \; P_k A^* A x_k=P_k A^* b_k \tag{7.4'}$$

where P_k is the orthoprojection onto X_k.

It should be emphasized that it is not the method of least squares which is causing the trouble asserted in Theorem 7.1 but its use in conjunction with parametric reduction. It is the use of approximating subspaces that are unrelated to the operator and the "cost criterion" which can introduce pathology: it can be shown that if each X_k is the span of eigenvectors of $A^* A$, then the sequence of approximants computed using the exact right hand side (each $b_k=b$) in (7.5) converges to x_*. Note that it is possible, in defining J_k for (7.5), to modify the cost criterion to be a new quadratic form such that the redefined adjoint A^* becomes such that X_k *is* (approximately) the span of eigenvectors of $A^* A$ (compare this with Galerkin's method in [58]). In that case we would again expect convergence.

Typically, the most difficult aspect of inverse problems (i.e., of coefficient determination) is demonstrating uniqueness. Theorem 2.3 only requires uniqueness of a *minimum norm* solution but this weaker requirement seems not to make things much easier for this class of problems. It does, however, seem likely, that in many such problems, it is *only* in such a weakened sense (and perhaps also with the advantage of an *a priori* restriction of the domain of f based on some knowledge of the desired solution) that one would actually have uniqueness.

EXAMPLES 7.2: Consider, first, the determination of the single parameter $x \in \boldsymbol{R}^+$ in

$$\begin{gathered} u_t=xu_{ss} \quad \text{for} \quad 0<t<T,\ 0<s<\pi, \\ u(\cdot,0)=u(\cdot,\pi)=0 \text{ on } [0,T],\ u_0:=u(0,\cdot)\in L^2(0,\pi) \end{gathered} \tag{7.6}$$

from observation of y given by

$$y(\cdot):=-xu_s(\cdot,0) \quad \text{on} \quad (0,T]. \tag{7.7}$$

Note that although u_0 is not given, if x were specified it would follow from [44] that $u(t,\cdot)$ is determined (indeed, well-posedly using L^2 topologies) by that and $y(\cdot)$ for every $t>0$ so that uniqueness for x is equivalent to uniqueness for the

augmented unknown $(x, u_0) \in \boldsymbol{R}^+ \times L^2(0,\pi)$. Suppose (x_*, u_{0*}) were a solution of (7.6), (7.7) (i.e., solving the direct problem (7.6) with x_* for x and u_{0*} for $u(0,\cdot)$, the y defined by (7.7) matches the observations). Then

$$y(t) = \sum_k (-x_* kc_k)\, e^{-x_* k^2 t} \quad \text{where } u_{0*} = \sum_k c_k \sin ks. \tag{7.8}$$

Now, set

$$\hat{x} := x_*/4, \qquad \hat{u}_0 := \sum_k 2c_k \sin 2ks \tag{7.9}$$

and observe that (7.6), (7.7) with $\hat{x}$ for x and $\hat{u}_0$ for $u(0,\cdot)$ gives

$$y(t) = \sum_k (-\hat{x}2k \cdot 2c_k)\, e^{-\hat{x}(2k)^2 t} \tag{7.8'}$$

which exactly matches (7.8) so that $(\hat{x}, \hat{u}_0)$ would equally well be a solution of (7.6), (7.7) and the solution could not be unique. Depending on the norm chosen for $X := \boldsymbol{R} \times L^2(0,\pi)$, there might or might not be a unique solution of minimum norm. Indeed, (relying on the equivalence of x and (x, u_0) as unknowns) one might plausibly topologize X by $x \in \boldsymbol{R}$ alone and then there would be a sequence of solutions attaining arbitrarily small positive norm. One would have to *assume*, for example, that it is known that the appropriate eigenfunction expansion of u_0 is not too lacunary (enough modes are excited) to permit reliable interpretation of the observed y as (7.8) rather than (7.8').

Next, consider the problem of determining the (distributed) unknown coefficient $x = x(\cdot)$ in

$$\begin{aligned} &u_t = u_{ss} - xu =: L_x u \quad \text{on} \quad \boldsymbol{R}^+ \times (0,\pi), \\ &u_s(\cdot,0) = u_s(\cdot,\pi) = 0 \ \text{on} \ \boldsymbol{R}^+, \quad u_0 := u(0,\cdot) \in L^2(0,\pi) \end{aligned} \tag{7.9}$$

from observation of y given by

$$y(t) := u(\cdot,0) \quad \text{on} \quad (0,T]. \tag{7.10}$$

As above (now using results of [57]) the determinations of x in $L^\infty(0,\pi)$ and of (x, u_0) in $L^\infty(0,\pi) \times L^2(0,\pi)$ are equivalent. Suppose (x_*, u_{0*}) is a solution of (7.9), (7.10) and that $\{(-\lambda_k^*, e_k^*)\}$ are the eigenpairs of L_{x_*}, i.e.,

$$e_k^{*\prime\prime} - x_* e_k^* = -\lambda_k^* e_k^* \quad \text{on} \quad (0,\pi), \quad e_k^{*\prime}(0) = e_k^{*\prime}(\pi) = 0 \tag{7.11}$$

for each k. Then

$$y(t) = \sum_k c_k e_k^*(0)\, e^{-\lambda_k^* t} \quad \text{where} \quad u_{0*} = \sum_k c_k e_k^*. \tag{7.12}$$

Now set

$$\begin{aligned} &\hat{x}(s) := x_*(\pi - s), \quad \hat{e}_k(s) := e_k^*(s), \quad \hat{\lambda}_k := \lambda_k^*, \\ &\hat{u}_0(s) := \sum_k \hat{c}_k \hat{e}_k \quad \text{with} \quad \hat{c}_k := [e_k^*(0)/e_k^*(\pi)]\, c_k. \end{aligned} \tag{7.13}$$

Clearly $\{(-\hat{\lambda}_k, \hat{e}_k)\}$ are the eigenpairs of $L_{\hat{x}}$. It follows from Lemma 1 of [57] that $\{e_k^*(\pi)\}$ is bounded away from 0, so, if x_* is such that $\{e_k^*(0)\}$ is bounded, as is entirely likely, then u_{0*} in $L^2(0, \pi)$ will automatically imply that $\hat{u}_0$ is also in $L^2(0, \pi)$. In any case, given x_* there will always be nontrivial u_{0*} in $L^2(0, \pi)$ (define by (7.9), (7.10) the observable y to be considered) for which $\hat{u}_0$, defined by (7.13), also happens to be in $L^2(0, \pi)$. Using $(\hat{x}, \hat{u}_0)$ in (7.9), (7.10) gives

$$y(t) = \sum_k \hat{c}_k \hat{e}_k(0) e^{-\hat{\lambda}_k t} \tag{7.12'}$$

which, from (7.13), just matches (7.12), precluding uniqueness. This example is closely related to the need, in [23], for presentation, in general, of *two* spectra (associated with distinct boundary conditions) to determine the unknown coefficient $x(\cdot)$ uniquely. ■

EXAMPLE 7.3: Consider a distributed system whose evolution is governed by a partial differential equation of the form:

$$\begin{aligned} &u_t = \nabla \cdot x_0 \nabla u - x_1 u \quad \text{on} \quad (0,T)\times\Omega, \\ &u = 0 \quad \text{on} \quad (0,T)\times\partial\Omega, \quad u(0,\cdot) = u_0 \end{aligned} \tag{7.14}$$

with observations of the form:

$$y := x_0 \frac{\partial u}{\partial \nu} \quad \text{on} \quad (0,T]\times\Gamma. \tag{7.15}$$

The bounded domain Ω in $\boldsymbol{R}^m$ is given, as also is the observable portion of the boundary, $\Gamma \subset \partial\Omega$. It is desired to determine the pair of coefficients $x := (x_0, x_1)$ from the observation $\tilde{y}$. The initial state u_0 is not given but estimation of u_0 or of u is of no *direct* interest here.

Our principal assumption is that the coefficient pair x_* can be restricted, on the basis of *a priori* knowledge, to a subset $D \subset X$ for which one has the following:

(7.16)

(a) $X := X_0 \times X_1$ *is topologized at least as strongly as in* $C^{\nu+1}(\Omega)\times C^{\nu}(\Omega)$ *with* $\nu \geqslant 1$,

(b) D *is compact in* X,

(c) *there is a unique* x_* *in* D *consistent with* (7.14), (7.15) *and the observed* $\tilde{y}$,

(d) Ω, Γ *and* D *are such that the hypotheses of Theorem* 5.2 *of* [63] *are satisfied for each* x *in* D.

The condition (7.16b) could, of course, be verified by assuming that the uncertainty in x requires only specification of finitely many parameters in some suitable representation (note that here we are assuming that the desired solution x_* is such that this is *exact*). On the other hand, one continues to consider a true distributed

parameter x if the compactness of D results from *a priori* knowledge of extra *smoothness* of x_* with a bound on, say, some higher-order Sobolev norm. The "difficult" assumption, as noted earlier, is (7.16c). From the Examples 7.2 it is clear that this means that specification of D must involve considerable *a priori* knowledge about the unknown coefficients.

Introduce a *state space* U of suitably smooth functions on Ω vanishing on $\partial\Omega$ for which the operators $L_x\,(x \in D)$ determined by

$$L_x v := \nabla \cdot x_0 \nabla v - x_1 v \quad on \quad \Omega \tag{7.17}$$

are infinitesimal generators of C_0 semigroups $S(\cdot;X)$ on U. Then (7.14), (7.15) become

$$y(t) := B(x)\, S(t; x)\, u_0 \tag{7.18}$$

where $B(x)$ denotes the observation map (from U to functions on Γ) given by (7.15).

Since u_0 is not given we note that (7.16d) implies existence of an "estimator" $E(\tau: x)$ for which u, y satisfying (7.14), (7.15) give

$$E(\tau; x)\, y(\cdot) = u(\tau, \cdot) \quad \tau > 0$$

(note that E requires no knowledge of u_0) with $\{E(\tau; x): x \in D\}$ bounded uniformly in operator norm: $L^2([0,T]\times\Gamma) \to H^K(\Omega)$ for arbitrarily large K restricted only by the selection of ν in (7.16a)); see [63]. Taking $\omega := E(\tau; x)\,\tilde{y}$ and sufficiently large K that the embedding of $H^k(\Omega) \hookrightarrow U$ is compact, we have *a priori* that $\omega = u(\tau, \cdot)$ is in a fixed compact set $U_* \subset U$ determined by $\tilde{y}$ and D. We now replace (7.18) by

$$y(t) := B(x)\, S(t-\tau; x)\, \omega \quad \tau \leqslant t \leqslant T \tag{7.18'}$$

which again is equivalent to (7.14), (7.15).

In practice one must work with computational approximations to the governing partial differential equations (7.14). Introduce a sequence of (increasingly accurate) discretizations and let $\{S_k(\cdot; x)\}$ be the semigroups associated with these discretized versions of (7.14). Each family of semigroups $\{S_k(\cdot; x): x \in D\}$ "lives" on a finite-dimensional space U_k which may be taken to be embedded in the original state space U with corresponding projections $P_k: U \to U_k$, uniformly bounded and converging strongly to the identity. Consistency of the discretizations means that

$$S_k(t; x)\, P_k\, \hat{u} \to S(t; x)\, \hat{u} \quad \hat{u} \in U,\ x \in D. \tag{7.19}$$

(If one discretizes in "time" as well, then additional minor modifications are required but this causes no extra difficulty). Note that U_{**}, the closure of $\{P_k\,\omega: \omega \in U_*,\ k = 1, 2, \ldots\}$, will also be compact since $\{P_k\}$ is equicontinuous and convergent and U_* is compact. Discretization will similarly involve approximating $B(x)$ by $B_k(x)$ and again consistency means equicontinuity and pointwise convergence as $k \to \infty$.

The original problem of seeking x_* consistent with the observation $\bar{y}$ and (7.14), (7.15) or, equivalently, (7.18) or (7.18′) is now replaced by the sequence of approximating problems:

$$\begin{gathered} \textit{minimize } J_k(x,\omega) := \frac{1}{2}\int_\tau^T |\bar{y}(t) - B_k(s)\, S_k(t-\tau;x)\,\omega|^2\, dt \\ \textit{over } D\times P_k U_* \end{gathered} \tag{7.20}$$

The consistency of the discretization means that $J_k(x_*, P_k\omega_*)\to 0$ ($\omega_* := u(\tau,\cdot)$ for the "true" state). Let (x_k,ω_k) be approximate solutions of (7.20) as $k\to\infty$ so $J_k(x_k,\omega_k)\leqslant\varepsilon_k\to 0$, i.e.,

$$y_k := B_k(x_k)\, S_k(\cdot-\tau;x_k)\,\omega_k \to \bar{y}|_{[\tau,T]} \textit{ in } L^2([\tau,T]\times\Gamma). \tag{7.21}$$

Since $\{(x_k,\omega_k)\}$ is in the compact set $D\times U_{**}$, any subsequence contains a subsubsequence converging to, say, $(\hat{x},\hat{\omega})$. By equicontinuity and pointwise convergence, the corresponding subsubsequence of $\{y_k\}$ converges to $\hat{y} := B(\hat{x})\, S(\cdot-\tau;\hat{x})\hat{\omega}$ but (7.21) then gives $\hat{y}=\bar{y}|_{[\tau,T]}$ so the assumption (7.16c) implies $\hat{x}=x_*$, $\hat{\omega}=\omega_*$. (Note: specification of $\bar{y}|_{[\tau,T]}$ uniquely specifies $\bar{y}$ on $(0,T]$ since Theorem 5.3 of [63] guarantees the analyticity in t of $\bar{y}$ for $Re\, t>0$.) This shows, in particular, that $x_k\to x_*$ in X so the approximation scheme (7.20) is convergent.

EXAMPLE 7.4: A very similar approach can be used for a comparable identification problem associated with a delay differential equation. Consider an autonomous linear hereditary system of the form:

$$\begin{gathered} \dot{u}(t) = x_0\, u(t) + x_1\, u(t-\xi) + \int_0^\infty x(s)\, u(t-s)\, ds + \varphi(t), \\ u(0)=u_0, \quad u|_{(-\infty,0)}=\omega \end{gathered} \tag{7.22}$$

with the observations of the form:

$$y(t) := Cu(t) \quad \textit{on} \quad (0,T]. \tag{7.23}$$

Here $u(\cdot)$ takes values in $\boldsymbol{R}^m$ so x_0, x_1 and the values of $x(\cdot)$ are in the space M_m of $m\times m$ matrices while ξ is in $R^+ := (0,\infty)$. The unknown is then $x := (\xi, x_0, x_1, x(\cdot))$ in some subset D of $X = R\times M_m\times M_m\times X_0$ where X_0 is a suitable space of M_m—valued functions on R^+. It is the presence of $x(\cdot)$ as an unknown function in X_0 which makes this a distributed problem and ill-posed in the sense under discussion. Note that, even with a reduction to an assumed parametric representation, the *dynamics* of (7.22) involve an infinite-dimensional state space and so would require analysis of computational approximation; see, e.g., [14].

It is assumed that C in (7.23), the input φ and the initial data (u_0,ω) are known. (Analogously with the discussion in Example 7.3, above, it would also be plausible

to consider ω as unknown or only partly known but with (7.22), (7.23) observable; i.e., such that knowledge of x would permit determination of the state, including ω, from φ, $\tilde{y}$). If one were to have $\omega=0$ (and one also can take $u_0=0$ with no loss of generality by including a δ—function in φ) as well as suitable growth conditions on $x(\cdot)$, then taking Laplace transforms in (7.22) gives

$$[\hat{x}(\sigma)+x_0+e^{-\sigma\xi}x_1-\sigma]\,\hat{u}(\sigma)=\hat{\varphi}(\sigma). \tag{7.24}$$

If φ, u were known on all of R^+ ($T=\infty$ in (7.23) and, say, $C=I$), one would have $\hat{\varphi}$, $\hat{u}$ and, under reasonable conditions, this would determine $[\hat{x}(\sigma)+x_0+e^{-\sigma\xi}x_1]$ (e.g., in the scalar case one just divides by $\hat{u}$). Since $[\hat{x}(\sigma)+e^{-\sigma\xi}x_1]\to 0$ as $\sigma\to\infty$, x_0 is then determined. Note that $[\hat{x}(\sigma)+e^{-\sigma\xi}x_1]$ is the Laplace-Stieltjes transform of $[x(s)+\delta(s-\xi)x_1]$ so, taking the inverse transform (assuming suitable decay of $x(s)$ as $s\to\infty$), $x(\cdot)$, ξ and x_1 are determined. This analysis is, of course, totally unreasonable in that one would never have $T=\infty$ in practice or, presumably, $C=I$. On the other hand, as with Examples 7.2, this gives some indication of the difficulties involved in establishing uniqueness and the nature of the ill-posedness involved in the determination of x.

As with (7.16), we now assume that the unknown x can *a priori* be restricted to a subset $D\subset X$ and that we have the following

(7.25)
- (a) $X:=R\times M_m\times M_m\times X_0$ *with* X_0 *embeddable in* $C([0,\infty)\to M_m)$, *i.e.,* X_0*—convergence implies uniform convergence on compact intervals,*
- (b) D *is compact in* X,
- (c) *there is a unique* x_* *in* D *consistent with* (7.22), (7.23) *and the observed* $\tilde{y}$,
- (d) D *and* ω *are such that* $\int_0^\infty x(t+s)\,\omega(-s)\,ds$ *is absolutely and uniformly convergent for* $t\geqslant 0$ *and all* x *in* D.

It follows from (7.25) that (7.22) defines a solution $u(\cdot\,;x)$ for x in D and that $y=$ $=y(\cdot\,;x):=Cu(\cdot\,;x)$ will depend continuously (in sup norm) on x in D for fixed C, φ, u_0, ω.

A variety of approximation methods are available for such equations as (7.22); see, e.g., [68]. If D imposes suitable smoothness and growth conditions on x, then one will have equicontinuity and convergence for the sequence of discretizations. Slightly more generally, let $\{D_k\}$ be a sequence of subsets of X (e.g., associated with approximating parametric representations) and assume one has approximate solutions

$$\begin{aligned} &u_k(t;\hat{x}_k) \quad \textit{defined for} \quad t\in T_k,\ \hat{x}_k\in D_k,\\ &u_k(t_k;\hat{x}_k)\to u(t;\hat{x}) \quad \textit{as} \quad t_k\to t,\ \hat{x}_k\to\hat{x}\in D. \end{aligned} \tag{7.26}$$

This assumes that T_k is discrete but "becomes dense" in $[0,T]$ and that each $\hat{x}$ in D is a limit of some associated sequence $\{\hat{x}_k\in D_k\}$; we also require that $\bigcup_k D_k$ is precompact in X and $\{D_k\}$ subconvergent to D.

For each k we define the quadratic cost criterion

$$J_k(x)=\frac{1}{2}I_k(|y-Cu_k(\cdot;x)|^2) \quad for \quad x\in D_k \tag{7.27}$$

where I_k is an operation of numerical integration over $[0,T]$ defined in terms of values at $t\in T_k$ (i.e., $I_k(\hat{f})\to\int_0^T \hat{f}\,dt$ for "smooth" $\hat{f}$). We then can consider x_k defined by the least squares problem:

$$minimize\ J_k(x)\ over\ D_k. \tag{7.28}$$

The precompactness of $\bigcup_k D_k$ and subconvergence to D imply existence of a convergent subsequence: $x_{k(j)}\to\hat{x}\in D$. Clearly (for $k=k(j)$)

$$J_k(x_k)\to\frac{1}{2}\int_0^T|\bar{y}(t)-Cu(t;\hat{x})|^2dt=:J_*(\hat{x}). \tag{7.29}$$

By assumption one can find $\{\tilde{x}_k\in D_k\}$ such that $\tilde{x}_k\to x_*$ so $J_k(\tilde{x}_k)\to J_*(x_*)=0$; by minimality $0\leqslant J_k(x_k)\leqslant J_k(\tilde{x}_k)\to 0$ so $J_*(\hat{x})=0$ and $\hat{x}$ is a solution. By (7.25c), this means $\hat{x}=x_*$ and $x_{k(j)}\to x_*$. As earlier, we conclude that $x_k\to x_*$.

If (7.25b) were omitted and (7.25c) were weakened to require existence of a unique *minimum norm* solution x_*, then approaches along the lines of Section 3 and 4 would still provide convergent approximation methods. One would then have to assume that X_0 is, e.g., uniformly convex and gives such smoothness conditions that the sequence of maps: $x\mapsto Cu_k(\cdot;x): D_k\to L^2([0,T]\to\boldsymbol{R}^{m'})=:Y$ defined by the discretizations employed should be graph subconvergent to the map: $x\mapsto Cu(\cdot;x): D\to Y$ (for the given C, φ, u_0, ω). For example, one might replace (7.27) by

$$J_k(x):=\frac{1}{2}\|x\|_X^2+\frac{\lambda_k}{2}I_k(|\bar{y}_k-Cu_k(\cdot;x)|^2) \quad for \quad x\in D_k \tag{7.27'}$$

where $\lambda_k\to\infty$ at a suitable rate (depending on the rates of convergence of $\bar{y}_k$ to $\bar{y}$, of u_k to u, uniformly over bounded subsets of D, and I_k to the integral). Alternatively, one might replace (7.28) by the generalized interpolation procedure:

$$minimize\ \|x\|_X \quad over \quad \{x\in D_k: |\bar{y}_k(t)-Cu_k(t;x)|\leqslant\varepsilon_k \quad for \quad t\in T_k\} \tag{7.28'}$$

with $\varepsilon_k\to 0$ at a suitable rate. ■

References

[1] Aubin J.-P. Approximation of Elliptic Boundary-Value Problems, Wiley-Interscience, N.Y. 1972.

[2] Backus G. Inference from inadequate and inaccurate data, in: Mathematical Problems in the Geophysical Sciences, Lect. in Appl. Math. **14**, Amer. Math. Soc., Providence, 1971, 1–106.

[3] BRAMBLE J. and SCHATZ A. H. On the numerical solution of elliptic boundary value problems by least squares approximation of the data, in: Numerical Solution of Partial Differential Equations, II, (B. E. Hubbard, editor), Acad. Press, N. Y., 1971, 107–131.
[4] BROCKETT R. W. Finite Dimensional Linear Systems, Wiley, N. Y., 1970.
[5] BRYNIELSON L. On Fredholm integral equations of the first kind with convex constraints, *SIAM J. Math. Anal.* **5** (1974), 955–962.
[6] CANNON J. R. and DOGRU A. H. Estimation of permeability and porosity from well test data, paper SPE 5345, AIME-Soc. Petrol. Eng. meeting, Ventura (Calif.), April, 1975.
[7] CANNON J. R. and EWING E. E. A direct numerical procedure for the Cauchy problem for the heat equation, *J. Math. Anal. Appl.* **56** (1976), 7–17.
[8] CARASSO A. S., SANDERSON J. G. and HYMAN J. M. Digital removal of random media image degradations by solving the diffusion equation backwards in time, *SIAM J. Numer. Anal.* **15** (1978), 344–367.
[9] CARASSO A. and STONE A. P. (editors), Improperly Posed Boundary Value Problems, Pitman, London, 1975.
[10] CASTI J. Matrix Riccati equations, dimensionality reduction and generalized X—Y functions, *Utilitas Math.* **6** (1974), 95–110.
[11] CHAVENT G. Analyse Fonctionelle et Identification de Coefficients Répartis dans les Equations aux Dérivées Partielles, doctoral thesis, Paris, 1971.
[12] — , Identification of functional parameters in partial differential equations, in [24], 31–48.
[13] CHEWNING W. C. and SEIDMAN T. I. A convergent scheme for boundary control of the heat equation, *SIAM J. Control & Opt.* **15** (1977), 64–72.
[14] CLIFF E. M. and BURNS J. A. Parameter identification for linear hereditary systems via an approximation technique, (preprint) VPI & SU, Blacksburg (Va.), March, 1978.
[15] CRAVEN P. and WAHBA G. Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of cross-validation, (preprint) *Tech. Report* 445, Univ. Wisc. (Dept. Stat.), Madison, 1977.
[16] DAY M. M. Normed Linear Spaces, Springer-Verlag, Berlin, 1966.
[17] DOUGLAS J. jr., and DUPONT T. Galerkin methods for parabolic equations, *SIAM J. Numer. Anal.* **7** (1970), 575–626.
[18] DUNFORD N. and SCHWARTZ J. Linear Operators, vol. II, Interscience, N. Y., 1963.
[19] EFIMOV N. V. and STEČKIN S. B. Approximative compactness and Čebyšev sets, *Dokl. Akad. Nauk SSSR* **140** (1961), 522–524 (Transl., Sov. Math. Dokl. 2 (1961)).
[20] EYKHOFF P. System Identification, Wiley-Interscience, London, 1974.
[21] FRANKLIN J. N. On Tikhonov's method for ill-posed problems, *Math. of Comp.* **28** (1974), 889–907.
[22] GANDER W. How to apply the dual problem to solve $\|x\|_2 = \min$ subject to $\|Ax - b\|_2 \leqslant a$, (preprint) SIAM spring meeting, Madison, 1978.
[23] GEL'FAND I. M. and LEVITAN B. M. On the determination of a differential equation from its spectral function, *Izv. Akad. Nauk SSSR Ser. Mat.* **15** (1951), 309–360.
[24] GOODSON R. E. and POLIS M. (editors), Identification of Parameters in Distributed Systems, (1974 JACC) Amer. Soc. Mech. Eng., N. Y., 1975.
[25] GROETSCH C. W. Sequential regularization of ill-posed problems involving unbounded operators, *Comm. Math. U. Carol.* **18** (1977), 489–498.
[26] GROETCH C. W. and KING J. T. Extrapolation and the method of regularization for generalized inverses, (preprint) U. Cincinnatti.
[27] HADAMARD J. Lectures on the Cauchy Problem in Linear Partial Differential Equations, Yale Univ. Press, New Haven, 1923.
[28] HERMAN G. T. and NAPARSTEK A. Fast image reconstruction based on a Radon inversion formula appropriate for rapidly collected data, *SIAM J. Appl. Math.* **33** (1977), 511–533.
[29] IMBER M. and KHAN J. Prediction of transient temperature distributions with embedded thermocouples, *AIAA J.* **10** (1972), 784–789.

[30] IVANOV V. V. Algorithms of optimal accuracy for the approximate solution of operator equations of the first kind, *Ž. Vyčisl. Mat. i Mat. Fiz.* **15** (1975), 3–11 (Transl., *USSR Comp. Math. & Math. Phys.* **15** (1975), 1–9).

[31] JEROME J. W. On the L_2 n-width of certain classes of functions of several variables, *J. Math. Anal. Appl.* **20** (1967), 110–123.

[32] — , Asymptotic estimates of the L_2 n-width, *J. Math. Anal. Appl.* **22** (1968), 449–464.

[33] JOHN F. Numerical solution of heat conduction for preceding times, *Ann. Mat. Pura Appl.* **40** (1955), 129–142.

[34] — , Plane Waves and Spherical Means, (Chapter IV), Interscience, N. Y., 1955.

[35] KAC M. Can one hear the shape of a drum? *Amer. Math. Monthly* **73** (1966), 1–23.

[36] KATO T. Perturbation Theory for Linear Operators, Springer-Verlag, Berlin, 1966.

[37] KHUDAK Yu. N. On regularizing algorithms, *Ž. Vyčisl. Mat. i Mat. Fiz.* **15** (1975), 12–18 (Transl., *USSR Comp. Math. & Math. Phys.* **15** (1975), 9–16).

[38] KNOPS R. J. (editor), Symposium on Non-Well-Posed Problems and Logarithmic Convexity, Lect. Notes in Math. **316**, Springer-Verlag, Berlin, 1973.

[39] KRASNOSELSKII M. A. Topological Methods in the Theory of Non-Linear Integral Equations, (Transl.) Macmillan, N. Y., 1964.

[40] KRYANEV A. V. An iterative method for solving incorrectly posed problems, *Ž. Vyčisl. Mat. i Mat. Fiz.* **14** (1974) (Transl., *USSR Comp. Math. & Math. Phys.* **14** (1974), 24–33).

[41] LATTES R. and LIONS J.-L. The Method of Quasi-Reversibility, Amer. Elsevier, N. Y., 1969.

[42] LAVRENTIEV M. M., ROMANOV V. G. and VASILIEV V. G. Multidimensional Inverse Problems for Differential Equations, Lect. Notes in Math. **167**, Springer-Verlag, Berlin, 1970.

[43] LAX P. D. and RICHTMYER R. D. Survey of the stability of linear finite difference equations, *Comm. Pure Appl. Math.* **9** (1956), 267–293.

[44] MIZEL V. J. and SEIDMAN T. I. Observation and prediction for the heat equation, *J. Math. Anal. Appl.* **28** (1969), 303–312.

[45] MOROZOV V. A. Choice of parameter for the solution of functional equations by the regularization method, (Transl.) *Sov. Math. Dokl.* **8** (1967), 1000–1003.

[46] — , Theory of splines and the problem of stable computation of the values of an unbounded operator, *Ž. Vyčisl. Mat. i Mat. Fiz.* **11** (1971), 545–558 (Transl., *USSR Comp. Math. & Math. Phys.* **11** (1971), 1–19).

[47] — , Regular Methods for the Solution of Incorrectly Posed Problems, Izdat. Moskov. Univ., Moscow, 1974.

[48] MOROZOV V. A. and GREBENNIKOV A. I. Optimal methods for the solution of ill-posed problems, *Gamoqeneb. Math. Inst. Sem. Mohsen. Anotacie.* Vyp. **11** (1976), 51–55 (MR55 # 4712).

[49] NASHED M. Z. Generalized inverses, normal solvability and iteration for singular operator equations, in: Nonlinear Functional Analysis and Applications (L. B. Rall, editor), Acad. Press, N. Y., 1971, 311–359.

[50] OPIAL Z. Lecture Notes on Nonexpansive and Monotone Mappings in Banach Spaces, Lect. Notes 67–1, Brown Univ., Providence, 1967.

[51] PAYNE L. E. Improperly Posed Problems in Partial Differential Equations, CBMS-NSF Regional Conf. Series, 1975; note also „Some general remarks on ..." in [36].

[52] PETRYSHYN W. V. On the approximation-solvability of nonlinear equations, *Math. Ann.* **177** (1968), 156–164.

[53] POGOŽAEV S. I. Normal solvability of nonlinear mappings in uniformly convex Banach spaces, *Funk. Anal. Prilozen.* **3** (1969), 80–84 (Transl., *Fnl. Anal. Appl.* **3** (1969), 147–151).

[54] POLYAK B. T. The convergence rate of the penalty function method, *Ž. Vyčisl. Mat. i Mat. Fiz.* **11** (1971), 3–11 (Transl., *USSR Comp. Math. & Math. Phys.* **11** (1971), 1–12).

[55] PUCCI C. Alcune limitazioni per la soluzioni di equazioni paraboliche, *Ann. Mat. Pura Appl.* (Ser. IV) **58** (1959), 161–172.

[56] RICHTER G. R. Numerical solution of integral equations of the first kind with nonsmooth kernels, *SIAM J. Numer. Anal.* **15** (1978), 511–522.

[57] SEIDMAN T. I. Observation and prediction for one-dimensional diffusion equations, *J. Math. Anal. Appl.* **51** (1975), 165–175.

[58] — , Solution of singular problems, I: linear problems in Hilbert space, *Pac. J Math.* **61** (1975), 513–520.

[59] — , The solution of singular problems, II: nonlinear problems, *Math. Res. Report* 75–10, Univ. of MD, Balto. Co., Baltimore, 1975.

[60] — , Computational approaches to ill-posed problems: general considerations, Proc. 1976 Conf. Inf. Sci. and Systems, Johns Hopkins Univ., Baltimore, 1976, 258–262.

[61] — , Observation and prediction for the heat equation, IV: patch observability and controllability, *SIAM J. Control and Opt.* **15** (1977), 412–427.

[62] — , Parameter and state estimation for a diffusion process, Proc. IFAC 2-nd Symp. on Control of Distrib. Param. Systems (U. Warwick, 1977), Pergamon Press, 1978.

[63] — , Exact boundary control for some evolution equations, *SIAM J. Control and Opt.* **16** (1978), 979–999.

[64] — , Nonconvergence results for the application of least squares estimation to ill-posed problems, *J. Opt. Th. Appl.* **30** (1980), 535–547.

[65] — , Recovery of a diffused signal, Séminaires IRIA-Analyse et contrôle de systèmes, Rocquencourt, 1979, 71–82.

[66] STUMMEL F. Discrete convergence of mappings, in: Proc. Conf. on Numer. Anal. (Dublin, 1972), Acad. Press, N. Y., 1973.

[67] TANANA V. P. Optimal algorithms for solving nonlinear unstable problems, *Sibirskii Mat. Ž.* **17** (1976), 1116–1128 (Transl., *Sib. Math. J.* **5** (1976), 825–834).

[68] THOMPSON R. J. Difference approximations for some functional differential equations, Lect. Notes in Math. **333**, Springer-Verlag, Berlin, 1973, 511–533.

[69] TIKHONOV A. N. Solution of incorrectly formulated problems and the regularization method, *Dokl. Akad. Nauk SSSR* **151** (1963), 501–504 (Transl., *Sov. Math. Dokl.* **4** (1963), 1035–1038).

[70] VASIN V. V. and TANANA V. P. On the stability of projection methods for solving incorrectly posed problems, *Ž. Vyčisl. Mat. i Mat. Fiz.* **15** (1975), 19–29 (Transl., *USSR Comp. Math. & Math. Phys.* **15** (1975), 16–26).

[71] WAHBA G. On the optimal choice of nodes in the collocation projection method for solving linear operator equations, *J. Approx. Theory* **16** (1976), 175–186.

[72] — , Practical approximate solutions to linear operator equations when the data are noisy, *SIAM J. Numer. Anal.* **14** (1977), 651–667.

Received, January 1981

Zbieżne metody aproksymacji zagadnień niepoprawnie postawionych

Praca stanowi obszerne studium aproksymacyjnych metod rozwiązywania rozmaitych zagadnień niepoprawnie postawionych. Wyniki dotyczące zbieżności proponowanych metod są przedstawione łącznie z dowodami.

Część I jest poświęcona zagadnieniom abstrakcyjnym. Zawiera ona zasadnicze rezultaty dotyczące zbieżności proponowanych algorytmów aproksymacji, a także dyskusję przydatności rozmaitych technik (funkcje kary i minimalizacja, regularyzacja). W szczególności omawiane są algorytmy rozwiązywania problemów niepoprawnie postawionych w sytuacji jedynie przybliżonej znajomości danych będących wynikami pomiarów, wykorzystujące tzw. uogólnioną interpolację.

Część II jest poświęcona rozmaitym zastosowaniom. Omawia się algorytmy aproksymacji równań zawierających operatory Niemyckiego. Przedstawione są metody rozwiązywania zagadnień odwrotnych dla równań różniczkowych cząstkowych typu parabolicznego (odtwarzanie sygnału wejściowego) oraz ogólniejszych zagadnień identyfikacji układów o parametrach rozłożonych.

Сходящиеся методы аппроксимации некорректных задач

В работе предлагаются аппроксимационные алгоритмы решения различных некорректных задач. Изучается сходимость представленных алгоритмов.

В части I исследуется абстрактная формулировка задач. Вводятся основные алгоритмы, доказывается их сходимость и рассматривается эффективность различных подходов (функция штрафа и минимизация, регуляризация). Анализируются алгоритимы решения некорректных задач, со знанием данных (результатов измерений) лишь в дискретных точках, использующие обобщенную интерполяцию.

В части II обсуждаются некоторые прикладные вопросы. Рассматривается аппроксимация уравнений, содержащих операторы Немыцкого. Представлены методы решения обратных задач для параболических уравнений и более общих проблем идентификации систем с распределенными параметрами.