2. METHODS OF SOLVING ILL-POSED PROBLEMS

2.1 VARIATIONAL REGULARIZATION

2.1.1 Pseudoinverse. Singular values decomposition

Consider linear Equation (1.3.1). Let $A : X \to Y$ be a linear closed operator, $D(A)$ and $R(A)$ be its domain and range, $X = H_1$ and $Y = H_2$ be Hilbert spaces, $N(A) := \{u : Au = 0\}$, A^* is the adjoint operator, $\overline{R(A)} \oplus N(A^*) = H_2$, the bar stands for the closure, and $\oplus$ is the orthogonal sum. If A is injective, i.e., $N(A) = \{0\}$, and surjective, i.e. $R(A) = H_2$, and $D(A) = H_1$, then the inverse operator A^{-1} is defined on H_2, $A^{-1}A = AA^{-1} = I$, I is the identity operator. A closed, linear, defined on all of H_1 operator, is bounded, so A is an isomorphism of H_1 onto H_2 if it is injective, surjective, and $D(A) = H_1$.

If $N(A) \neq \{0\}$, P is the orthoprojector onto $N(A)$ in H_1, and Q is the orthoprojector onto $\overline{R(A)}$, then one defines a *pseudoinverse* (generalized inverse) A^+: $D(A^+) := R(A) \oplus N(A^*)$, $A^+(N(A^*)) := \{0\}$, $A^+Au := u - Pu$. Thus, $AA^+A = A$, $A^+AA^+ = A^+$, and $AA^+u = Qu$ for $u \in D(A^+)$. The operator A^+ is bounded iff $R(A) = \overline{R(A)}$. If $f \in R(A)$, i.e., $f = Au_0$ for some u_0, then the problem $\|Au - f\| = \inf$ has a solution u_0, every element $u_0 + v$, $\forall v \in N(A)$, is also a solution, and there is a unique solution with minimal norm, namely the solution u_0 such that $u_0 \perp N(A)$, $u_0 = A^+f$. If $f \notin R(A)$, then the infimum of $\|Au - f\|$ is not attained. If A is bounded and $f \in R(A)$, then the element $u_0 = A^+f$ solves the equation $A^*Au = A^*f$ and is the minimal norm solution to this equation, i.e., $u_0 \perp N(A)$. Indeed, if $Au_0 = f$ and $u_0 \perp N(A)$, then $A^*Au = A^*f$. Conversely, if $A^*Au = A^*f$ and $f = Au_0$, with $u_0 \perp N(A)$, then $A^*A(u - u_0) = 0$, $(A^*A(u - u_0), u - u_0) = 0$, and $Au = Au_0$. Thus, $u = u_0$ if $u \perp N(A)$. One can prove the formula: $A^+f = \lim_{\alpha\to 0}(\alpha I + A^*A)^{-1}A^*f$, where $\alpha > 0$ is a regularization parameter (see Section 2.1.2) and $f \in R(A)$.

Let us define the singular value decomposition. Let $A : H_1 \to H_2$ be a linear compact operator, $B := A^*A : H_1 \to H_1$ is a compact selfadjoint operator, $B\varphi_j = s_j^2\varphi_j$, $\|\varphi_j\| = 1$,

$$(\varphi_j, \varphi_m) = \delta_{jm} := \begin{cases} 1 & j = m, \\ 0 & j \neq m, \end{cases} \qquad s_1 \geq s_2 \geq \cdots \geq 0,$$

s_j are called s-values of A. If $AA^* := T$, and $A\varphi_j/\|A\varphi_j\| := \psi_j$, then $T\psi_j = s_j^2\psi_j$, $(\psi_j, \psi_m) = (A\varphi_j, A\varphi_m)/(\|A\varphi_j\|\|A\varphi_m\|) = s_j^2\delta_{jm}/\|A\varphi_j\|^2 = \delta_{jm}$. Thus, $\|A\varphi_j\| = s_j$, $A\varphi_j = s_j\psi_j$, $A^*\psi_j = s_j\varphi_j$. If $u \in H_1$ is arbitrary, then $Au = \sum_{j=1}^\infty s_j(u, \varphi_j)\psi_j$ $(*)$, $\lim_{j\to\infty} s_j = 0$. Thus, an element $f \in R(A)$ iff $\sum_{j=1}^\infty |(f, \varphi_j|^2/s_j^2 < \infty$ (Picard's test). If $A = A^*$, then $s_j^2 = \lambda_j^2(A^2)$, where λ_j^2 are eigenvalues of A^2. Then $\psi_j = \varphi_j$, $A\varphi_j = \lambda_j\varphi_j$.

If $\dim H_1 = n < \infty$, $dim H_2 = m < \infty$, then A can be written as $A = VSU^*$, where A is an $m \times n$ matrix, U and V are unitary matrices ($n \times n$ and $m \times m$, respectively), whose columns are vectors φ_j and ψ_j respectively, and S is an $m \times n$ matrix with the diagonal elements s_j, $1 \leq j \leq r$, r is the rank of the matrix A, and other elements of S are zeros. The matrix A^+ can be calculated by the formula $A^+ = US^+V^*$, where S^+ is an $n \times m$ matrix with diagonal elements s_j^{-1}, $1 \leq j \leq r$, and other elements of S^+ are zeros.

2.1.2 Variational (Phillips-Tikhonov) regularization

Assume $A : H_1 \to H_2$, $\|A\| < \infty$ is linear, $f \in R(A)$, $\|f_\delta - f\| \leq \delta$, f_δ is not necessarily in $R(A)$. The problem $Au = f$ is assumed ill-posed (cf. Sec. 1.3). Consider the problem:

$$F(v) := \left\| Av - f_\delta \right\|^2 + \alpha\|v\|^2 = \inf, \tag{2.1.1}$$

where $\alpha > 0$ is a parameter.

Theorem 2.1.1. *Assume $Au = f$, and $u \perp N(A)$. Then:*

(i) The minimizer $u_{\alpha,\delta}$ of (2.1.1) *does exist and is unique*
(ii) If $\delta \to 0$ and $\alpha = \alpha(\delta)$ satisfies the condition $\delta^2/\alpha(\delta) \to 0$ as $\delta \to 0$, then $\lim_{\delta\to 0}\|u_\delta - u\| = 0$, where $u_\delta := u_{\alpha(\delta),\delta}$.

Proof. Functional (2.1.1) is quadratic. A necessary and sufficient condition for its minimizer is the Euler's equation:

$$Bv + \alpha v = A^* f_\delta, \quad B := A^* A \geq 0, \tag{2.1.2}$$

which has a unique solution $u_{\alpha,\delta} = (B + \alpha I)^{-1} A^* f_\delta$. Claim (i) is proved. One has $F(u_{\alpha\delta}) \leq F(u) = \delta^2 + \alpha \|u\|^2 = \alpha(\delta^2/\alpha + \|u\|^2) \leq c\alpha$, $c = \text{const} > 0$, if $\delta^2/\alpha \leq c_1$, $c_1 = \text{const}$. Thus $\|u_{\alpha,\delta}\| \leq c$. Below c stands for various positive constants. Choose $\alpha = \alpha(\delta)$ so that $\delta^2/\alpha(\delta) \to 0$ as $\delta \to 0$, and let $u_\delta := u_{\alpha(\delta),\delta}$. Then $\|u_\delta\| \leq c$, so $u_\delta \rightharpoonup u_0$ as $\delta \to 0$, and $Bu_\delta \to A^* f$. This implies $Bu_0 = A^* f$, and we claim that $\|u_0\| \leq \|w\|\ \forall w : Bw = A^* f$. This claim we prove later. Thus $u_0 = u$. Let us prove that $\lim_{\delta\to 0} \|u_\delta - u\| = 0$. One has $\|(B+\alpha)^{-1} A^* f_\delta - u\| \leq \|(B+\alpha)^{-1} A^* (f_\delta - f)\| + \|(B+\alpha)^{-1} A^* f - u\| \leq \|(B+\alpha)^{-1} A^*\|\delta + \|(B+\alpha)^{-1} Bu - u\| \leq \delta/(2\sqrt{\alpha}) + \eta(\alpha) \to 0$, $\alpha \to 0$. Here the estimate $\|(B+\alpha)^{-1} A^*\| \leq 1/(2\sqrt{\alpha})$ and the relation $\|[(B+\alpha)^{-1} B - I]u\| \to 0$ as $\alpha \to 0$, were used. To prove the *first estimate*, one uses the formula: $(B+\alpha)^{-1} A^* = A^*(T+\alpha)^{-1}$, $T := AA^*$, $B := A^* A$, and the polar representation of A^* yields, $A^* = VT^{1/2}$, where V is an isometry, $\|V\| \leq 1$. One has $\|T^{1/2}(T+\alpha)^{-1}\| = \max_{\lambda\geq 0} \lambda^{1/2}(\lambda+\alpha)^{-1} = 1/(2\sqrt{\alpha})$, where the spectral representation for T was used.

Let us prove the second relation: $\|[(B+\alpha)^{-1} B - I]u\|^2 = \int_0^{\|B\|} |\frac{\lambda}{\lambda+\alpha} - 1|^2 d(E_\lambda u, u) = \int_0^{\|B\|} \frac{\alpha^2}{(\lambda+\alpha)^2} d(E_\lambda u, u) \to \|Pu\|^2$ as $\alpha \to 0$, where P is the orthoprojector onto $N(A)$, and E_λ is the resolution of the identity of the self-adjoint operator B (see [KA]).

If $u \perp N(A)$, then $\lim_{\alpha\to 0} \|(B+\alpha)^{-1} Bu - u\| = 0$.

Finally, let us prove the claim used above.

Lemma 2.1.2. *If B is a monotone hemicontinuous operator in a Hilbert space H, $D(B) = H$, $Bv + \alpha(\delta)v = g_\delta$, $Bu_0 = g$, $v \rightharpoonup u_0$, $g_\delta \to g$, $\alpha(\delta) \to 0$ as $\delta \to 0$, then $Bv \to Bu_0 = g$.*

In our proof above, $B \geq 0$ is a linear operator, so B is monotone, that is, $(B(u) - B(v), u - v) \geq 0\ \forall u, v \in D(B)$. Recall that a nonlinear operator A is called hemicontinuous, if $(A(u + tv), w)$ is a continuous function of $t \in \mathbb{R}$ for any $u, v, w \in H$.

Proof of Lemma 2.1.2. Clearly, $Bv \to g$. If $Bv \to g$, $v \rightharpoonup u_0$, then $Bu_0 = g$, *that is, monotone operator is w-closed.* Indeed, if B is monotone, then $(Bv - B(u_0 - tw), v - u_0 + tw) \geq 0$ for any $w \in H$. Passing to the limit $v \rightharpoonup u_0$, $Bv \to g$, $t \to 0$, and using hemicontinuity of B, one gets: $(g - Bu_0, w) \geq 0\ \forall w \in H$. This implies $Bu_0 = g$, so Lemma 2.1.2 is proved. □

The claim $\|u_0\| \leq \|w\|\ \forall w : Bw = A^* f$ is proved for nonlinear monotone operators in Theorem 2.1.6 below. Theorem 2.1.1 is proved. □

2.1.3 Discrepancy principle

Theorem 2.1.1 gives an a priori choice of $\alpha = \alpha(\delta)$ which guarantees convergence $\lim_{\delta\to 0} u_\delta = u$. An a posteriori choice of α is given by Theorem 2.1.3 below.

Theorem 2.1.3. *(Discrepancy principle). Assume* $\|f_\delta\| > \delta$. *If* $\alpha = \alpha(\delta)$ *is the root of the equation*

$$\left\| A(B+\alpha)^{-1} A^* f_\delta - f_\delta \right\| = C\delta, \qquad C = \text{const} > 1, \tag{2.1.3}$$

then $\lim_{\delta\to 0} \|u_\alpha - u\| = 0$, $u_\delta := u_{\alpha(\delta),\delta}$.

Proof. First, let us prove that equation (2.1.3) has a unique solution. Write this equation as

$$C^2\delta^2 = \int_0^{\|B\|} \left| \frac{\lambda}{\lambda+\alpha} - 1 \right|^2 d\left(F_\lambda f_\delta, f_\delta\right) = \alpha^2 \int_0^{\|B\|} \frac{d(F_\lambda f_\delta, f_\delta)}{(\lambda+\alpha)^2} := I(\alpha, \delta),$$

where F_λ is the resolution of the identity of the selfadjoint operator $T := AA^*$, and the commutation formula $(B+\alpha)^{-1}A^* = A^*(T+\alpha)^{-1}$ was used. One checks this formula easily. If $\alpha \to +\infty$, then $I(\infty, \delta) = \|f_\delta\|^2 > \delta^2$. If $\alpha \to +0$, then $I(+0, \delta) = \|P_1 f_\delta\|^2 \le \delta^2$, where P_1 is the orthoprojector on $N(T) = N(A^*)$. Indeed, $\|P_1 f_\delta\| \le \|P_1(f_\delta - f)\| + \|P_1 f\| \le \delta$, because $\|P_1\| \le 1$, and $P_1 f = 0$ since $f \in R(A)$ and $R(A) \perp N(A^*)$. Thus, if $C^2 < \|f_\delta\|^2$ then equation (2.1.3) has a solution. This solution is unique because $I(\alpha, \delta)$ is a monotone increasing function of α for each fixed $\delta > 0$.

Now let us prove $\lim_{\delta\to 0} \|u_\delta - u\| = 0$. One has $\|Au_\delta - f_\delta\|^2 + \alpha(\delta)\|u_\delta\|^2 \le \delta^2 + \alpha(\delta)\|u\|^2$. Since $\|Au_\delta - f_\delta\|^2 > \delta^2$, it follows that $\|u_\delta\|^2 \le \|u\|^2$. Therefore (*) $\limsup_{\delta\to 0} \|u_\delta\| \le \|u\|$. If $\|u_\delta\| \le \|u\|$, then one can select a weakly convergent sequence $u_\delta \rightharpoonup u_0$ as $\delta \to 0$. In the proof of Theorem 2.1.1 it was proved that $u_0 = u$, where u is the unique minimal-norm solution of the equation $Bu = A^* f$. By the lower semicontinuity of the norm in H, one has $\|u\| \le \liminf_{\delta\to 0} \|u_\delta\|$. Together with (*), one gets $\lim_{\delta\to 0} \|u_\delta\| = \|u\|$. This and the weak convergence $u_\delta \rightharpoonup u$, imply $\lim_{\delta\to 0} \|u_\delta - u\| = 0$. □

Our proof is based on the following useful result.

Theorem 2.1.4. *If* $u_n \rightharpoonup y$ *and* $\|u_n\| \le \|y\|$, *then* $\lim_{n\to\infty} \|u_n - y\| = 0$.

Proof. If $u_n \rightharpoonup y$ then $\liminf_{n\to\infty} \|u_n\| \ge \|y\|$. Also one has $\limsup_{n\to\infty} \|u_n\| \le \|y\|$. Thus, $\lim_{n\to\infty} \|u_n\| = \|y\|$, and $\|u_n - y\|^2 = \|u_n\|^2 + \|y\|^2 - 2\Re(u_n, y) \to 0$ as $n \to \infty$. □

2.1.4 Nonlinear ill-posed problems

Lemma 2.1.5. *Assume that A in* (1.3.1) *is a closed, nonlinear, injective map. If K is a compactum, then the inverse operator A^{-1} is continuous on $A(K)$.*

Proof. Since A is injective, A^{-1} is well-defined on $A(K)$. Let $A(u_n) := f_n \to f$, $u_n \in K$. Then $u_n \to u \in K$, where u_n is a subsequence denoted again u_n. Since A is closed, $u_n \to u$ and $A(u_n) \to f$ imply $A(u) = f \in A(K)$, and, by the injectivity of A, one has $A^{-1}(f_n) \to A^{-1}(f)$. Lemma 2.1.5 is proved. □

Claim: Let us assume that $A: H \to H$ is monotone, continuous, $D(A) = H$, $A(u) = f$, and $A(u_\alpha) + \alpha\, u_\alpha = f$. Then $\|u_\alpha\| \leq \|u\|$.

Proof. Indeed, $A(u_\alpha) - A(u) + \alpha\, u_\alpha = 0$. Multiply this equation by $u_\alpha - u$ and use the monotonicity of A, to get $(u_\alpha, u_\alpha - u) \leq 0$. Thus, $\|u_\alpha\| \leq \|u\|$. Let $\alpha \downarrow 0$. Select a sequence, denoted again by u_α, such that $u_\alpha \rightharpoonup u_0$ as $\alpha \to 0$. Then $A(u_\alpha) \to f$. Since A is monotone, it is w-closed (see the proof of Theorem 2.1.1), so $A(u_0) = f$, and u_0 is the minimal norm solution to equation (1.3.1). □

Theorem 2.1.6. *If $A: H \to H$ is monotone and hemicontinuous, if $D(A) = H$, if $A(u_0) = f$, where u_0 is the minimal-norm solution to $A(u) = f$, and if $A(u_\alpha) + \alpha u_\alpha = f$, then the minimal-norm solution to* (1.3.1) *is unique and* $\lim_{\alpha\to 0} \|u_\alpha - u_0\| = 0$.

Proof. We have $\|u_\alpha\| \leq \|u\|$, $\forall u \in \{u : A(u) = f\} := N$. Thus, $\limsup_{\alpha\to 0} \|u_\alpha\| \leq \|u\|\ \forall u \in N$. Let $u_\alpha \rightharpoonup u_0$. Then $\|u_0\| \leq \liminf_{\alpha\to 0} \|u_\alpha\|$ and $\limsup \|u_n\| \leq \|u_0\|$. Thus, $\lim_{\alpha\to 0} \|u_\alpha\| = \|u_0\|$. This and the weak convergence $u_\alpha \rightharpoonup u$ imply strong convergence $u_\alpha \to u_0$ as in Theorem 2.1.3.

The minimal norm solution to (1.3.1) is unique if A is monotone and continuous, because in this case *the set of solutions N is convex and closed*. Its closedness is obvious, if A is continuous. Its convexity follows from the monotonicity of A and the following lemma:

Lemma 2.1.7 (Minty). *If A is monotone and continuous, then (a)* $(A(u) - f, v - u) \geq 0$ $\forall v$ *is equivalent to (b)* $(A(v) - f, v - u) \geq 0\ \forall v$.

Proof. If (a) holds, then $A(u) = f$, and (b) holds by the monotonicity of A. If (b) holds, then take $v = u + tw$, $t \geq 0$, where w is arbitrary, and get $(A(u + tw) - f, w) \geq 0\ \forall w$. Take $t \to 0$ and get $(Au - f, w) > 0\ \forall w$. Thus, $A(u) = f$, and (a) holds. Lemma 2.1.7 is proved. □

To prove that N is convex, one assumes that $u_1, u_2 \in N$ and derives that $tu_1 + (1 - t)u_2 \in N\ \forall t \in (0, 1)$. Indeed, if $u_j \in N$, then, by Lemma 2.1.7, $(Av - f, v - u_j) \geq 0\ \forall v$. Thus, $(Av - f, v - tu_1 - (1 - t)u_2) = t(Av - f, v - u_1) + (1 - t)(Av - f, v - u_2) \geq 0\ \forall v$. Thus, $tu_1 + (1 - t)u_2 \in N\ \forall t \in (0, 1)$.

To prove uniqueness of the minimal norm element of a convex and closed set N in a Hilbert space, one assumes that there are two such elements, u_1 and u_2. Then $\|u_1\| = \|u_2\| := m$, and $\|tu_1 + (1-t)u_2\| \le t\|u_1\| + (1-t)\|u_2\| = m$, so that any element of the segment joining u_1 and u_2 has minimal norm m. Since Hilbert space is strictly convex, this implies $u_1 = u_2$. Indeed, take $t = 1/2$. Then $\|(u_1+u_2)/2\|^2 = \|u_1\|^2 = \|u_2\|^2$. So $\Re(u_1, u_2) = \|u_1\|\|u_2\| = \|u_1\|^2 = \|u_2\|^2$. Thus, $u_1 = u_2$. Theorem 2.1.6 is proved. □

Consider the equation

$$A\big(u_\alpha\big) + \alpha\, u_\alpha = f_\delta. \tag{2.1.4}$$

Theorem 2.1.8. *Assume that A is monotone and continuous, and equation* (1.3.1) *has a solution. If $\alpha = \alpha(\delta) \to 0$ and $\delta/\alpha(\delta) \to 0$ as $\delta \to 0$, then the unique solution to* (2.1.4) *converges strongly to u, the unique solution to* (1.3.1) *of minimal norm.*

Proof. Because A is monotone and $\alpha > 0$, the equation $A(v_\alpha) + \alpha v_\alpha = f$ has a solution, and this solution is unique. Let $u_\delta := u_{\alpha(\delta)}$ solve (2.1.4), and $v_\delta := v_{\alpha(\delta)}$. One has $\|u - u_\delta\| \le \|u - v_\delta\| + \|v_\delta - u_\delta\|$. By Theorem 2.1.6, $\lim_{\delta\to 0}\|u - v_\delta\| = 0$. Let us prove $\lim_{\delta\to 0}\|v_\delta - u_\delta\| = 0$. We have $A(u_\delta) - A(v_\delta) + \alpha(u_\delta - v_\delta) = f_\delta - f$. Multiply this by $u_\delta - v_\delta$ and use the monotonicity of A to get $\alpha\|u_\delta - v_\delta\|^2 \le \delta$ $\|u_\delta - v_\delta\|$. This implies $\lim_{\delta\to 0}\|u_\delta - v_\delta\| = 0$, if $\lim_{\delta\to 0}\delta/\alpha(\delta) = 0$. Theorem 2.1.8 is proved. □

2.1.5 Regularization of nonlinear, possibly unbounded, operator

Assume that :

(1) $A : D(A) \to X$ is a closed, injective, possibly nonlinear, map in Banach space X.
(2) $\phi \ge 0$ is a functional such that the set $\{v : \phi(v) \le c\}$ is precompact in X for any constant $c > 0$,
(3) Equation (1.3.1) has a solution $y \in D(\phi)$, $A(y) = f$,
(4) $D(A) \subseteq D(\phi)$.
The last assumption can be replaced in some cases when A is an unbounded operator, by the assumption.
(4′) $D(\phi) \subseteq D(A)$.

Define the functional $F(u) = \|A(u) - f_\delta\| + \delta\phi(u)$, where $\delta > 0$ is a parameter, $\|f_\delta - f\| \le \delta$, $D(F) = D(A) \cap D(\phi)$. Consider the minimization problem:

$$F(u) = \inf, \qquad \inf_{u\in D(F)} F(u) := m = m(\delta). \tag{2.1.5}$$

Let $F(u_j) \le m + \frac{1}{j} \le m + \delta$, where $j = j(\delta)$ is the smallest integer satisfying the inequality $\frac{1}{j(\delta)} \le \delta$. Denote $u_\delta := u_{j(\delta)}$. One has $m \le F(u_\delta) \le m + \delta \le F(y) + \delta = \delta(2 + \phi(y)) := c\delta$, and $\phi(u_\delta) \le c$. By assumption (2), as $\delta \to 0$ one can select a convergent subsequence, denoted again $u_\delta, u_\delta \to u$, such that $\lim_{\delta\to 0} \|A(u_\delta) - f\| = 0$. Thus, $A(u) = f$ by the closedness of A, and $u = y$ by the injectivity of A. Since the limit of any subsequence u_δ is the same, namely y, it follows that $\lim_{\delta\to 0} \|u_\delta - y\| = 0$. We have proved:

Theorem 2.1.9. *If* (1.3.1) *has a solution, then, under the assumptions* (1)–(4) (*or* (4′)), *any sequence u_δ, such that $F(u_\delta) \le m(\delta) + \delta$, converges strongly to the solution y of* (1.3.1) *as $\delta \to 0$.*

Remark 2.1.10. *In the proof of Theorem 2.1.9 we do not need existence of the minimizer of the function* (2.1.5).

2.1.6 Regularization based on spectral theory

Assume that A in (1.3.1) is a linear bounded operator, $\|A\| < \infty$, $Ay = f$, and $y \perp N(A)$, i.e., y is the unique minimal-norm solution.

Lemma 2.1.11. *Solvable equation* (1.3.1) *with bounded linear operator A is equivalent to the equation*

$$Bu = f_1, \quad B := A^* A \ge 0, \quad f_1 := A^* f. \tag{2.1.6}$$

Proof. If u solves (1.3.1), apply A^* to (1.3.1) and get (2.1.6), so u solves (2.1.6). If u solves (2.1.6) and (1.3.1) is solvable, i.e., $f = Ay$, then $f_1 = By$, $B(u - y) = 0$, and $(B(u - y), u - y) = 0$. This implies $\|A(u - y)\| = 0$, so $Au = f$. Thus u solves (1.3.1). □

Equation (2.1.6) is a solvable equation with monotone, continuous operator, so Theorem 2.1.6 is applicable and yields the following theorem:

Theorem 2.1.12. *If $0 < \alpha(\delta) \to 0$, $\delta/\alpha(\delta) \to 0$ as $\delta \to 0$, $\|f_\delta - f\| \le \delta$, and y is the minimal-norm solution to* (1.3.1), *then $\lim_{\delta\to 0} \|u_\delta - y\| = 0$, where u_δ is the unique solution of the equation*

$$Bu_\delta + \alpha(\delta) u_\delta = A^* f_\delta.$$

Lemma 2.1.13. *Consider the elements $w_\delta := \int_0^{\|B\|} g(s, \alpha) d E_s A^* f_\delta$, where E_s is the resolution of the identity of $B = A^* A$, $|s g(s, \alpha)| \le c$, $c = \text{const} > 0$ does not depend on s and α, $\lim_{\alpha\downarrow 0} g(s, \alpha) = 1/s$ $\forall s > 0$, $\sup_s \sqrt{s} g(s, \alpha) := g(\alpha)$, and $g(s, \alpha)$ is a piecewise-continuous function. Let $\alpha = \alpha(\delta) \to 0$ so that $\delta g(\alpha(\delta)) \to 0$ as $\delta \to 0$. Then $\lim_{\delta\to 0} \|w_\delta - y\| = 0$.*

Proof. If $f = Ay$, then $A^* f = By$, and $y = \int_0^{\|B\|} s^{-1} dE_s\, A^* f$. Thus, $\|w_\delta - y\| \le \| \int_0^{\|B\|} g(s,\alpha) dE_s\, A^*(f_\delta - f)\| + \| \int_0^{\|B\|} s(g(s,\alpha) - s^{-1}) dE_s\, y\| := I + \eta(\alpha)$, where $\lim_{\alpha \downarrow 0} \eta(\alpha) = 0$, and $I \le \delta \| \int_0^{\|B\|} g(s,\alpha) dE_s\, A^* \|$. One has $\| \int_0^{\|B\|} g(s,\alpha) dE_s\, A^* \| = \|g(A^*A,\alpha)A^*\| = \|A^* g(AA^*,\alpha)\| \le \sup_s |\sqrt{s} g(s,\alpha)| = g(\alpha)$. Thus $\|w_\delta - y\| \le \delta g(\alpha) + \eta(\alpha) \to 0$ as $\delta \to 0$. If the rate of decay of $\eta(\alpha)$ and the rate of growth of $g(\alpha)$ can be estimated, then a quasioptimal choice of $\alpha = \alpha(\delta)$ can be made by minimizing $\delta g(\alpha) + \eta(\alpha)$ with respect to α for a fixed δ. □

Remark 2.1.14. *We have used the spectral theorem for a selfadjoint operator* B*, namely the formula* $g(B) = \int_{-\infty}^{\infty} g(s) dE_s$*, where* E_s *is the resolution of the identity of* B*,* $D(g(B)) = \{u : \int_{-\infty}^{\infty} |g(s)|^2 d(E_s u, u) < \infty\}$*,* $\|g(B)\| = \sup_{|s| \le \|B\|} |g(s)|$.

Remark 2.1.15. *Similarly, one can use the theory of spectral operators in place of the spectral theory of selfadjoint operators, in particular Riesz bases formed by the root vectors.*

2.1.7 On the notion of ill-posedness for nonlinear equations

If A is a linear operator, then problem (1.3.1) is ill-posed if either $N(A) \ne \{0\}$, or $f \notin R(A)$, or $R(A)$ is not closed, i.e. A^{-1} is unbounded. If A is nonlinear and Frechet differentiable, then there are several possibilities. If $A'(u)$ is boundedly invertible at some u, then $A(u)$ is a local homeomorphism at this point, but it may be not a global homeomorphism. If $A'(u)$ is not boundedly invertible, this does not imply, in general, that A is not a homeomorphism. For example, a homeomorphism $A(u)$ may have a compact derivative, so its linearization yields an ill-posed problem. On the other hand, $A(u)$ may be compact, so (1.3.1) is an ill-posed problem, but $A'(u)$ may be a finite-rank operator, so that the range of $A'(u)$ is closed. In spite of the above, we will often call a nonlinear equation problem (1.3.1) ill-posed if $A'(u)$ is not boundedly invertible, and well-posed if $A'(u)$ is boundedly invertible, deviating therefore from the usual terminology.

2.1.8 Discrepancy principle for nonlinear ill-posed problems with monotone operators

Assume that A in (1.3.1) is monotone, i.e., $(A(u) - A(v), u - v) \ge 0,\ \forall u, v \in D(A)$, $D(A) = H$, A is continuous, A^{-1} is unbounded or does not exist, so (1.3.1) is an ill-posed problem, $f \in R(A)$, $\|f_\delta - f\| \le \delta$. Consider the discrepancy principle for finding $\epsilon = \epsilon(\delta)$ assuming that A is nonlinear monotone:

$$\left\| A(u_{\delta,\epsilon}) - f_\delta \right\| = C\delta, \tag{2.1.7}$$

where $C = \text{const} > 1$, $u_{\delta,\epsilon}$ is any element such that $F(u_{\delta,\epsilon}) := \|A(u_{\delta,\epsilon}) - f_\delta\|^2 + \epsilon \|u_{\delta,\epsilon}\|^2 \le m(\delta,\epsilon) + \delta^2(C - 1 - b)$, where $m(\delta,\epsilon) := \inf_u F(u)$, and ϵ plays the role of the regularization parameter α. We need three lemmas.

Lemma 2.1.16. *If A is monotone and continuous, and the set $N_f := \{u : A(u) = f\}$ is nonempty, then it is convex and closed.*

Lemma 2.1.17. *If A is monotone and continuous, then it is w-closed, that is, $u_n \rightharpoonup u$ and $A(u_n) \to f$ imply $A(u) = f$, where $\rightharpoonup$ and $\to$ stand for the weak and strong convergence in H, respectively.*

Lemma 2.1.17 in a stronger form (hemicontinuity of A replaces continuity, and it is assumed in this case that the monotone operator A is defined on all of H) follows from the proof of Lemma 2.1.2.

Lemma 2.1.18. *If $u_n \rightharpoonup u$ and $\|u_n\| \le \|u\|$, then $u_n \to u$.*

Proof of Lemma 2.1.16. If $A(u_n) = f$ and $u_n \to u$, then $A(u) = f$, so N_f is closed. If A is monotone, $A(u) = f$ and $A(v) = f$, then $(A(z) - f, z - u) \ge 0$ and $(A(z) - f, z - v) \ge 0$ $\forall z$ and vice versa. Thus, for any $\lambda, \eta \ge 0$, $\lambda + \eta = 1$, the element $\lambda u + \eta v \in N_f$. □

Lemma 2.1.18 is Theorem 2.1.4.

Theorem 2.1.19. *Assume:*

(i) A is a monotone, continuous operator, defined on all of H,
(ii) equation $A(u) = f$ is solvable, y is its minimal-norm solution, and
(iii) $\|f_\delta - f\| \le \delta$, $\|A(0) - f_\delta\| > C\delta$, where $C > 1$ is a constant. Then:
(j) the equation

$$\left\| A(u_{\delta,\varepsilon}) - f_\delta \right\| = C\delta, \tag{2.1.8}$$

is solvable for ε for any fixed $\delta > 0$. Here $u_{\delta,\varepsilon}$ is any element satisfying inequality $F(u_{\delta,\varepsilon}) \le m + (C^2 - 1 - b)\delta^2$, where $F(u) := \|A(u) - f_\delta\|^2 + \varepsilon\|u\|^2$, $m = m(\delta, \varepsilon) := inf_u F(u)$, $b = const > 0$, and $C^2 > 1 + b$,
and
(jj) if $\varepsilon = \varepsilon(\delta)$ solves (2.1.8), and $u_\delta := u_{\delta,\varepsilon(\delta)}$, then $\lim_{\delta\to 0} \|u_\delta - y\| = 0$.

Remark 2.1.20. *The equation $A(v) + \varepsilon v = f_\delta$ is uniquely solvable for any $\varepsilon > 0$ and any $f_\delta \in H$. If $v := v_{\delta,\varepsilon}$ is its solution, and $\|A(0) - f_\delta\| > C\delta$, where $C = const > 1$, then equation* (2.1.8) *with $u_{\delta,\varepsilon}$ replaced by $v_{\delta,\varepsilon}$, is solvable for $\varepsilon > 0$. If $\varepsilon := \varepsilon(\delta)$ is its solution, then* $\lim_{\delta\to 0} \varepsilon(\delta) = 0$. *If A is injective, and if $v_\delta := v_{\delta,\varepsilon(\delta)}$, then* $\lim_{\delta\to 0} \|v_\delta - y\| = 0$, *where y solves the equation $A(y) = f$.*

If A is not injective, then it is not true, in general, that $\lim_{\delta\to 0} v_\delta = y$, *where y is the minimal-norm solution to the equation $A(u) = f$ even if one assumes that A is a linear operator.*

Proof of Theorem 2.1.19. If A is monotone, continuous and is defined on all of H, then the set $N_f := \{u : A(u) = f\}$ is convex and closed, so it has a unique minimal-norm element y. To prove the existence of a solution to (2.1.8), we prove that the function $h(\delta, \varepsilon) := \|A(u_{\delta,\varepsilon}) - f_\delta\|$ is greater than $C\delta$ for sufficiently large ε, and smaller than $C\delta$ for sufficiently small ε. If this is proved, then the continuity of $h(\delta, \varepsilon)$ with respect to ε on $(0, \infty)$ implies that the equation $h(\delta, \varepsilon) = C\delta$ has a solution.

Let us give the proof. As $\varepsilon \to \infty$, we use the inequality:

$$\varepsilon\left\|u_{\delta,\varepsilon}\right\|^2 \le F\left(u_{\delta,\varepsilon}\right) \le m + \left(C^2 - 1 - b\right)\delta^2 \le F(0) + \left(C^2 - 1 - b\right)\delta^2,$$

and, as $\varepsilon \to 0$, we use another inequality:

$$\begin{aligned}\left\|A\left(u_{\delta,\varepsilon}\right) - f_\delta\right\|^2 \le F\left(u_{\delta,\varepsilon}\right) &\le m + \left(C^2 - 1 - b\right)\delta^2 \le F(y) + \left(C^2 - 1 - b\right)\delta^2 \\ &= \varepsilon\|y\|^2 + \left(C^2 - b\right)\delta^2.\end{aligned}$$

As $\varepsilon \to \infty$, one gets $\|u_{\delta,\varepsilon}\| \le c/\sqrt{\varepsilon} \to 0$, where $c > 0$ is a constant depending on δ. Thus, by the continuity of A, one obtains $\lim_{\varepsilon\to\infty} h(\delta, \varepsilon) = \|A(0) - f_\delta\| > C\delta$.

As $\varepsilon \to 0$, one gets $h^2(\delta, \varepsilon) \le \varepsilon\|y\|^2 + (C^2 - b)\delta^2$. Thus, $\liminf_{\varepsilon\to 0} h(\delta, \varepsilon) < C\delta$. Therefore equation $h(\delta, \varepsilon) = C\delta$ has a solution $\varepsilon(\delta) > 0$.

Let us now prove that if $u_\delta := u_{\delta,\varepsilon(\delta)}$, then $\lim_{\delta\to 0} \|u_\delta - y\| = 0$. From the estimate

$$\left\|A\left(u_\delta\right) - f_\delta\right\|^2 + \varepsilon\left\|u_\delta\right\|^2 \le C^2\delta^2 + \varepsilon\|y\|^2,$$

and from the equation (2.1.8), it follows that $\|u_\delta\| \le \|y\|$. Thus, one may assume that $u_\delta \rightharpoonup U$, and from (2.1.8) it follows that $A(u_\delta) \to f$ as $\delta \to 0$. By w-closedness of monotone continuous operators (hemicontinuity in place of continuity would suffice), one gets $A(U) = f$, and from $\|u_\delta\| \le \|y\|$ it follows that $\|U\| \le \|y\|$. Because A is monotone, the minimal norm solution to the equation $A(u) = f$ in H is unique. Consequently, $U = y$. Thus, $u_\delta \rightharpoonup y$, and $\|u_\delta\| \le \|y\|$. By Theorem 2.1.4, it follows that $\lim_{\delta\to 0} \|u_\delta - y\| = 0$.

Note that $\|y\| > 0$, because $\|A(0) - f\| > 0$ due to the assumption $\|A(0) - f_\delta\| > C\delta$, where $C > 1$. Theorem 2.1.19 is proved. □

Proof of Remark 2.1.20. Let $v = v_\varepsilon$ solve the equation $A(v) + \varepsilon v = f$, let $w := v_{\delta,\varepsilon}$ solve the equation $A(w) + \varepsilon w = f_\delta$, $h := f_\delta - f$, $\|h\| \le \delta$, and $w - v := z$. Then $A(w) - A(v) + \varepsilon z = h$. Multiply this equation by z and use the monotonicity of A to get $\varepsilon\|z\| \le \delta$. The triangle inequality yields: $\varepsilon\|v\| - \delta \le \varepsilon\|w\| \le \varepsilon\|v\| + \delta$. Note that $\lim_{\varepsilon\to\infty} v = 0$, and $\delta/\varepsilon \to 0$ as $\varepsilon \to \infty$. Thus, $\lim_{\varepsilon\to\infty} \|w\| = 0$. Therefore $\lim_{\varepsilon\to\infty} \|A(w) - f_\delta\| = \|A(0) - f_\delta\| > C\delta$.

Fix $\delta > 0$ and let $\varepsilon \to 0$. Then $\lim_{\varepsilon\to 0} \|v - y\| = 0$ and $\|v\| \le \|y\|$, where y is the minimal-norm solution to the equation $A(u) = f$. If A is injective, then this equation has only one solution y. Since $\varepsilon\|w\| \le \varepsilon\|v\| + \delta$, one gets the inequality

$\limsup_{\varepsilon\to 0} \varepsilon\|w\| = \limsup_{\varepsilon\to 0} \|Aw - f_\delta\| \le \delta$. Consequently, equation (2.1.8), with w replacing $u_{\delta,\varepsilon}$, has a solution $\varepsilon(\delta) > 0$. We claim that $\lim_{\delta\to 0} \varepsilon(\delta) = 0$, in fact, $\varepsilon(\delta) = O(\delta)$ as $\delta \to 0$. Indeed, from (2.1.8), with w replacing $u_{\delta,\varepsilon}$, one gets $\varepsilon(\delta)\|w\| = C\delta$, and we prove below that $\liminf_{\delta\to 0} \|w_\delta\| > 0$, where $w_\delta := v_{\delta,\varepsilon(\delta)}$. This implies $\varepsilon(\delta) = O(\delta)$.

We now claim that the limit $\lim_{\delta\to 0} w_\delta := u$ does exist, that u solves the equation $A(u) = f$, and $\|u\| > 0$. It is sufficient to check that $\|w_\delta\| < c$ where $c = const$ does not depend on δ as $\delta \to 0$. Indeed, if $\|w_\delta\| < c$, then a subsequence, denoted w_δ again, converges weakly to an element y, $w_\delta \rightharpoonup y$, and (2.1.8) implies $\lim_{\delta\to 0} A(w_\delta) = f$. Since A is monotone, it is $w-$closed, so $A(y) = f$. By the injectivity of A, any subsequence w_δ converges weakly to the same element y, so $w_\delta \rightharpoonup y$. Consequently, $\liminf_{\delta\to 0} \|w_\delta\| \ge \|y\| > 0$ as claimed. The inequality $\|y\| > 0$ follows from the assumption $\|A(0) - f\| > 0$.

To prove the inequality $\|w_\delta\| < c$, note that $C\delta = \varepsilon\|w_\delta\| \le \varepsilon\|w_\delta - v_{\varepsilon(\delta)}\| + \varepsilon\|v_{\varepsilon(\delta)}\| \le \delta + \varepsilon\|y\|$, where $\varepsilon := \varepsilon(\delta)$. Since $C > 1$, this implies $\delta/\varepsilon(\delta) \le c_1$, where $c_1 := \|y\|/(C-1)$. Thus, $\|w_\delta\| \le c$, where $c := c_1 + \|y\|$.

The last statement of Remark 2.1.20 is illustrated by the following example:

Example 2.1.21. *Let $Aw = (w, p)p$, $\|p\| = 1$, $p \perp N(A)$, $f = p$, $f_\delta = p + q\delta$, where $(q, p) = 0$, $\|q\| = 1$, $Aq = 0$, $\|q\delta\| = \delta$. One has $Ay = p$, where $y = p$ is the minimal-norm solution to the equation $Au = p$. Equation $Aw + \varepsilon w = p + q\delta$, has the unique solution $w = q\delta/\varepsilon + p/(1+\varepsilon)$. Equation* (2.1.8) *is $C\delta = \|q\delta + (\varepsilon p)/(1+\varepsilon)\|$. This equation yields $\varepsilon = \varepsilon(\delta) = c\delta/(1 - c\delta)$, where $c := (C^2 - 1)^{1/2}$, and we assume $c\delta < 1$ (see the second inequality in the assumption (iii) of Theorem 2.1.19). Let $w_\delta = w(\delta, \varepsilon(\delta))$. Then, $\lim_{\delta\to 0} w_\delta = p + c^{-1}q := u$, and $Au = p$. Therefore $u = \lim_{\delta\to 0} w_\delta$ is not p, i.e., u is not the minimal-norm solution to the equation $Au = p$.*

Remark 2.1.20 is proved. □

Remark 2.1.22. *It is easy to prove that if conditions (i) and (ii) of Theorem 2.1.19 hold and $A(u_{\varepsilon,\delta}) + \varepsilon u_{\varepsilon,\delta} = f_\delta$, and if $\lim_{\delta\to 0} \delta/\varepsilon(\delta) = 0$, where $\lim_{\delta\to 0} \varepsilon(\delta) = 0$, $\varepsilon := \varepsilon(\delta)$, then $\lim_{\delta\to 0} \|u_\delta - y\| = 0$, where $u_\delta := u_{\varepsilon(\delta),\delta}$, and y is the minimal-norm solution to the equation $A(u) = f$. In particular, if $\varepsilon = \delta^a$, $0 < a < 1$, then $\lim_{\delta\to 0} \|u_\delta - y\| = 0$. Indeed, $\|u_\delta - u\| \le \|u_\delta - v_\delta\| + \|v_\delta - u\|$, where v_δ is the unique solution to the equation $A(v_\delta) + \varepsilon(\delta)v_\delta = f$. It is well known that $\lim_{\delta\to 0} \|v_\delta - y\| = 0$, provided that $\lim_{\delta\to 0} \varepsilon(\delta) = 0$, and, clearly, $\|u_\delta - v_\delta\| \le \delta/\varepsilon(\delta)$: one multiplies the identity $A(u_\delta) - A(v_\delta) + \varepsilon(\delta)(u_\delta - v_\delta) = f_\delta - f$ by $u_\delta - v_\delta$ and uses the monotonicity of A and the inequality $\|f_\delta - f\| \le \delta$.*

The result similar to the one in the above remark can be found in [ARy].

2.1.9 Regularizers for ill-posed problems must depend on the noise level

In this Section we prove the following simple claim:

Claim 2. *There is no regularizer independent of the noise level to a linear ill-posed problem. If such a regularizer exists, then the problem is well-posed.*

Let A be a linear operator in a Banach space X. Assume that A is injective and A^{-1} is unbounded, that equation

$$Au = f \tag{2.1.9}$$

is solvable, and g is such that

$$\|g - f\| < \delta, \tag{2.1.10}$$

where $\|g\|$ is the norm in X and $\delta > 0$ is the noise level. Nothing is assumed about the statistical nature of noise. In particular, we do not assume that the noise has zero mean value or finite variance.

Question: Can one find a linear operator R with the property:

$$\|Rg - u\| \longrightarrow 0 \text{ as } \delta \longrightarrow 0 \tag{2.1.11}$$

for any $f \in Ran(A)$, where $Ran(A)$ is the range of A, and any $g \in X$ satisfying (2.1.10)?

Answer: no.

Proof. If such an R is found, then, taking $g = f$ and using the fact that $f \in Ran(A)$ is arbitrary, one concludes that $R = A^{-1}$ on the range of A. Secondly, writing $g = f + w$, where

$$\|w\| < \delta, \tag{2.1.12}$$

and w is arbitrary otherwise, one concludes from (2.1.11) and from the fact that $Rf = A^{-1}f = u$, that

$$\|Rg - u\| = \|Rw\| \longrightarrow 0 \text{ as } \delta \longrightarrow 0, \tag{2.1.13}$$

for any w satisfying (2.1.12). Since R is linear, this implies that R is bounded, which contradicts the equation $R = A^{-1}$ on $Ran(A)$ and the unboundedness of A^{-1}, which is the necessary condition for the ill-posedness of (2.1.9). □

A similar result one can find in [LY].

2.2 QUASISOLUTIONS, QUASINVERSION, AND BACKUS-GILBERT METHOD

2.2.1 Quasisolutions for continuous operator

Assume that equation (1.3.1) is solvable, its solution $u \in K$, where K is a compactum in a Banach space X, and A is continuous. Consider the problem

$$\left\| A(u) - f_\delta \right\| = \inf := m(\delta), \quad u \in K, \tag{2.2.1}$$

where $m(\delta)$ is the infimum of the function of $\|A(u) - f_\delta\|$ and $\|f_\delta - f\| \le \delta$. A minimizer for (2.2.1) is called a quasisolution to (1.3.1) with $f = f_\delta$. Let u_j be a minimizing sequence for (2.2.1). Since K is a compactum, one may assume that $u_j \to u_\delta$ as $j \to \infty$, $\|A(u_j) - f_\delta\| \to m(\delta)$, $A(u_j) \to A(u_\delta)$.

Thus $\|A(u_\delta) - f_\delta\| = m(\delta)$, so u_δ is a minimizer for the problem (2.2.1). The above argument shows that if f replaces f_δ in (2.2.1), and if equation (1.3.1) is solvable and its solution belongs to K, then any minimizer for (2.2.1) with $f_\delta = f$, is a solution to (1.3.1). Let us prove that $\lim_{\delta\to 0} \|u_\delta - u\| = 0$, where u_δ is a minimizer for (2.2.1), and u is a solution to (1.3.1), where existence of a solution (1.3.1) is assumed.

Indeed, $u_\delta \in K$, so one may assume that $u_\delta \to u$ as $\delta \to 0$. By continuity of A, one has $\lim_{\delta\to 0} A(u_\delta) = A(u)$. Thus, $\|A(u) - f\| = \lim_{\delta\to 0} m(\delta) = 0$. The last conclusion follows from the solvability of (1.3.1), which yields $f = A(u)$ and from the inequality $m(\delta) \le \|A(u_\delta) - A(u)\| + \delta$. We have proved:

Theorem 2.2.1. *If equation* (1.3.1) *is solvable, K is a compactum containing all the solutions to* (1.3.1)*, and $\|f_\delta - f\| \le \delta$, then* (2.2.1) *has a minimizer u_δ, and* $\lim_{\delta\to 0} \|u_\delta - u\| = 0$ *for every minimizer u_δ and some solution u to* (1.3.1).

Remark. Suppose that X is strictly convex, i.e., if $\|u\| = \|v\| = \|(u + v)/2\|$, then $u = v$. For example Hilbert spaces H are strictly convex, the spaces $L^p(D)$, $p > 1$, are strictly convex, but $L^1(D)$ and $C(D)$ are not. Suppose that K is a convex compactum, i.e., convex closed compact set. The metric projection of an element $f \in X$ onto K is the element $P_K f \in K$ such that $\|P_K f - f\| = \inf_{u\in K} \|u - f\|$. If X is strictly convex, then $P_K f$ is unique, and if K is a convex compactum, then $P_K f$ depends continuously on f. If A is injective and closed, not necessarily linear, and K is a compactum, then A^{-1} is continuous on the set AK. Indeed, if $f_n = A(u_n)$, $u_n \in K$, and $f_n \to f$, then a subsequence, denoted again u_n, converges to u because K is compact, and if A is closed, then $A(u) = f$, which proves the claim. Therefore, if X is strictly convex and K is a convex compactum, and if A is an injective bounded linear operator, then the quasisolution $u(f) = A^{-1} P_{AK} f$ depends continuously on f in the norm of X.

2.2.2 Quasisolution for unbounded operators

Assume that A is closed, possibly nonlinear, injective, unbounded operator, A^{-1} is possibly, unbounded, equation (1.3.1) is solvable, assumptions (1)–(4) of Section 2.1.5 hold, and K is a compactum containing all the solutions to (1.3.1).

Theorem 2.2.2. *Under the above assumptions, if $\|A(w_\delta) - f_\delta\| \le m(\delta) + \delta$, where $w_\delta \in K$, then* $\lim_{\delta\to 0} \|w_\delta - u\| = 0$*, where u is a solution to* (1.3.1).

Proof. One can assume $w_\delta \to w$ as $\delta \to 0$, because K is a compactum. Note that $\lim_{\delta\to 0} m(\delta) = 0$, because $m(\delta) \le \|A(u) - f_\delta\| \le \delta$. Thus, $\lim_{\delta\to 0} \|A(w_\delta) - f\| = 0$.

By closedness of A, one gets $A(w) = f$, so w is a solution to (1.3.1), and one can denote w by u. $\square$

2.2.3 Quasiinversion

Let A be a linear bounded operator in (1.3.1), and (1.3.1) is solvable, consider the equation $(\varepsilon Q + B)u_\varepsilon = f_1$, $B := A^* A$, $f_1 := A^* f$, $\varepsilon > 0$ is a parameter, Q is an operator chosen so that $\|(\varepsilon Q + B)^{-1}\| \le c(\varepsilon)$, and $\lim_{\varepsilon \to 0} \|u_\varepsilon - u\| = 0$, where u is a solution to (1.3.1). If A is unbounded, a similar idea can be applied to equation (1.3.1): consider the equation $(\varepsilon Q + B)v_\varepsilon = f$, where Q is chosen so that $\|(\varepsilon Q + B)^{-1}\| \le c_1(\varepsilon)$, $\lim_{\varepsilon \to 0} \|v_\varepsilon - u\| = 0$. The problem is: how does one choose Q with these properties? If A is a linear bounded operator, then $Q = I$ can be used by Theorem 2.1.1. If A is unbounded, some assumptions on its spectrum are needed. See [LL] for details.

2.2.4 A Backus-Gilbert-type method: Recovery of signals from discrete and noisy data

We discuss in this section the following Problem 2. Let D and the bar denote variance and mean value respectively,

$$\int_a^b f \psi_j dy = f_j + \epsilon_j, \quad 1 \le j \le n, \; \overline{\epsilon_j^* \epsilon_p} = C_{jp}$$

Problem 1. *Given $f_j + \epsilon_j$, $1 \le j \le n$, estimate $f(x)$.*

The idea is as follows: the estimate is sought in the form

$$f_n(x) = \sum_{j=1}^n \varphi_j(x)(f_j + \epsilon_j).$$

The problem is to find $\varphi_j(x)$ such that:

(1) if $\epsilon_j = 0$ then $\|f_n - f\| \to 0$, $n \to \infty$;

(2) if $\overline{\epsilon_j^* \epsilon_p} = C_{jp}$ then $D[f_n - f] \le \sigma^2$; where $\sigma > 0$ is a certain number. We find $\{\varphi_j\}$, $1 \le j \le n$, optimal, in a certain sense. Namely, if $\epsilon_j = 0$, then φ_j are found from the requirements:

$$(A) \sum_{j=1}^n \varphi_j(x) \int_a^b \psi_j dy = 1, \quad \int_a^b \left(\sum_{j=1}^n \varphi_j(x)\psi_j(y) \right)^2 (x - y)^2 dy = \min.$$

Note that

$$(A) \Rightarrow \sum_{j=1}^n \varphi_j(x)\psi_j dy \equiv \delta_n(x, y) \longrightarrow \delta(x - y), \quad \text{in an optimal way, as } n \to \infty.$$

If $\epsilon_j \ne 0$ then φ_j are found from $\{(A)$ and $(C\varphi, \varphi) \le \sigma^2\}$ if this problem is solvable and, if not, one increases σ^2 so that this problem becomes solvable.

2.2.4.1. A typical problem we are concerned with is the problem of estimating the spectrum of a compactly supported function from the knowledge of the spectrum at a finite number of frequencies. More precisely, let

$$(2\pi)^{-1}\int_{-1}^{1} f(x)\exp(-i\omega x)dx = F(\omega). \tag{2.2.2}$$

Suppose that the numbers:

$$F_j = F(\omega_j), \quad 1 \le j \le n \tag{2.2.3}$$

are given. At the moment we assume that F_j are given exactly, i.e., there is no noise. The case when the data are noisy will be considered below.

Problem 2. *Given F_j, $1 \le j \le n$, find an estimate $\hat{f}_n(x)$ of $f(x)$ such that*

$$\hat{f}_n(x) \to f(x) \quad \text{as } n \to \infty \quad \text{(convergence)} \tag{2.2.4}$$

$$\text{The estimate } \hat{f}_n \quad \text{is to be optimal in the sense specified in (2.2.13).} \tag{2.2.5}$$

To be specific let us assume that $f(x) \in L^2[-1, 1]$ and that the estimate is of the form

$$\hat{f}_n(x) = \sum_{j=1}^{n} F_j h_j(x), \tag{2.2.6}$$

where the functions $h_j(x)$ will be chosen soon. From (2.2.5) and (2.2.2) it follows that

$$\hat{f}_n(x) = \int_{-1}^{1} A_n(x, y) f(y) dy, \tag{2.2.7}$$

and

$$A_n(x, y) = \sum_{j=1}^{n} h_j(x)\psi_j^*(y), \tag{2.2.8}$$

where $\psi_j(y) = (2\pi)^{-1}\exp(i\omega_j y)$, and the star denotes complex conjugate.

Property (2.2.4), convergence, holds if $A_n(x, y)$ is a delta-sequence, i.e.,

$$\left\| \int_{-1}^{1} A_n(x, y) f(y) dy - f(x) \right\| \to 0, \quad n \to \infty. \tag{2.2.9}$$

Let

$$Q(x) = Q(h(x)) = \int_{-1}^{1} |A_n(x, y)|^2 (x - y)^2 dy, \tag{2.2.10}$$

and let

$$h(x) = (h_1(x), \dots, h_n(x))$$

be a sequence of functions such that

$$\int_{-1}^{1} A_n(x, y) dy = \sum_{j=1}^{n} h_j(x) a_j = 1, \tag{2.2.11}$$

where

$$a_j = \int_{-1}^{1} \psi_j^*(y) dy. \tag{2.2.12}$$

One can interpret (2.2.11) as the requirement that the estimate $\hat{f}_n$ is exact for $f(x) =$ const. Given (2.2.11), the smaller $Q(x)$, the better is the quality of the delta-sequence $A_n(x, y)$. Thus we are led to the optimization problem: Find such a sequence $h_j(x)$, $1 \le j \le n$ that

$$Q(h(x)) = \min \quad \text{and (2.2.11) holds.} \tag{2.2.13}$$

Note that the general problem of the type

$$\int_{\mathscr{D}} f(x) \psi_j^*(x) dx = b_j, \quad 1 \le j \le n, \tag{2.2.14}$$

where $\{\psi_j\}$, $1 \le j \le n$, is a linearly independent set of functions, and $\mathscr{D}$ is a bounded domain in R^d, can be treated in exactly the same way as before.

If the problem (2.2.13) has the unique solution $h(x) = \{h_1(x), \dots, h_n(x)\}$ then (2.2.6) is the optimal estimate which, as we prove, has the convergence property (2.2.4).

2.2.4.2. If the data are noisy, that is $F_j + \epsilon_j$ are given in place of F_j, where $\epsilon = \{\epsilon_1, \dots, \epsilon_n\}$ random vector with the covariance matrix

$$C_{ij} = \overline{\epsilon_i^* \epsilon_j}, \quad \overline{\epsilon_j} = 0, \tag{2.2.15}$$

where the bar denotes the mean value, then the variance $D(f - \hat{f}_n)$ can be computed

$$D(f - \hat{f}_n) = D\left[f(x) - \sum_{j=1}^{n} F_j h_j(x) - \sum_{j=1}^{n} \epsilon_j h_j(x) \right]$$
$$= \sum_{i,j=1}^{n} C_{ij} h_j(x) h_i^*(x) = (Ch, h), \quad C = (c_{ij}), \tag{2.2.16}$$

where (,) is the inner product in $\mathbb{C}^n$. Let us fix $\sigma^2 > 0$ and require that

$$(Ch, h) \leq \sigma^2. \tag{2.2.17}$$

The optimization problem for finding the vector $h(x) = (h_1(x), \dots, h_n(x))$ can be formulated as follows:

$$\text{Minimize } Q(h(x)) \quad \text{under the restrictions } (h, a) = 1 \text{ and } (Ch, h) \leq \sigma^2. \tag{2.2.18}$$

Here $a = (a_1^*, a_2^*, \dots, a_n^*)$. Clearly, problem (2.2.18) is not solvable for all $\sigma^2 > 0$. We will discuss this important point below. If (2.2.18) is solvable, the solution is unique, and the optimal estimate is given by

$$\hat{f}_n = \sum_{j=1}^{n} (F_j + \epsilon_j) h_j(x), \tag{2.2.19}$$

This estimate has variance $\leq \sigma^2$.

Our arguments so far are close to the usual ones. The new point is our convergence requirement (2.2.4). We prove the convergence property of our estimate and give the rate of convergence. The case when the data are the finite number of moments is treated, and the optimization requirements are introduced.

The problem we discuss is of interest in geophysics and many other applications.

2.2.4.3. Here a solution of the estimation problems is given.

We start with problem (2.2.13). Let us write $Q(x)$ as a quadratic form

$$Q(x) = (Bh, h), \tag{2.2.20}$$

where

$$(Bh, h) = \sum_{i,j=1}^{n} b_{ij}(x) h_j(x) h_i^*(x) \tag{2.2.21}$$

and

$$b_{ij}(x) = b_{ij} = \int_{-1}^{1} \psi_i^*(y)\psi_j(y)(x-y)^2 dy \tag{2.2.22}$$

is a self-adjoint positive definite matrix.

Let us write (2.2.13) as

$$(Bh, h) = \min, \quad (h, a) = 1. \tag{2.2.23}$$

Using the Lagrange multiplier λ, one obtains the standard necessary and sufficient condition for the minimizer h_{opt}

$$Bh = \lambda a, \quad (h, a) = 1. \tag{2.2.24}$$

Therefore $h = \lambda B^{-1}a$, $\lambda = (B^{-1}a, a)^{-1}$,

$$h_{\text{opt}} = B^{-1}a/(B^{-1}a, a) \tag{2.2.25}$$

is uniquely defined by (2.2.25), since B^{-1} is positive definite, and the denominator in (2.2.25) does not vanish. The minimum of $Q(x)$ is

$$Q_{\min}(x) = (Bh_{\text{opt}}, h_{\text{opt}}) = (B^{-1}a, a)^{-1} := \alpha_n(x). \tag{2.2.26}$$

We assume that

$$\int_{-1}^{1} \alpha_n(x)dx \to 0 \quad \text{as } n \to \infty. \tag{2.2.27}$$

If (2.2.27) holds, then (2.2.4) holds. Indeed, (2.2.13) and (2.2.27) imply that

$$\begin{aligned} \|\hat{f}_n - f\|^2 &\le \left\| \int_{-1}^{1} \sum_{j=1}^{n} h_j(x)\psi_j(y)[f(y) - f(x)]dy \right\|^2 \\ &\le \int_{-1}^{1} dx \int_{-1}^{1} |A_n(x,y)|^2 (x-y)^2 dy \int_{-1}^{1} |f(y) - f(x)|^2 (x-y)^{-2} dy \\ &\le \int_{-1}^{1} \alpha_n(x)dx \sup_{-1\le x\le 1} \int_{-1}^{1} \left| \frac{f(y)-f(x)}{x-y} \right|^2 dy \to 0 \quad \text{as } n \to \infty. \end{aligned} \tag{2.2.28}$$

In this argument we assume that the function $f(x)$ satisfies the inequality

$$\sup_{-1\le x\le 1} \int_{-1}^{1} \left| \frac{f(x)-f(y)}{x-y} \right|^2 dy < \infty. \tag{2.2.29}$$

This inequality is satisfied if, for example, the derivative of $f(x)$ exists except at a finite number of points and is uniformly bounded.

Let us illustrate the assumption (2.2.27). Let $\psi_p = (2\pi)^{-1}\exp(ip\pi x)$, $p = 0, \pm 1, \dots, \pm n$. Then

$$a_p = \int_{-1}^{1} \psi_p^*(x)dx = \frac{\sin(p\pi)}{p\pi^2} = \begin{cases} 0 & p \neq 0, \\ \pi^{-1}, & p = 0 \end{cases}$$

$(B^{-1}a, a) = \pi^{-2}b_{00}^{(-1)}$, where $b_{pj} = \int_{-1}^{1}(x-y)^2\exp\{i(j-p)\pi y\}dy$, $B^{-1} = (b_{pj}^{(-1)}) = (B_{jp}(\det b_{pj})^{-1})$, where B_{jp} is the cofactor corresponding to the element b_{jp} of the matrix (b_{pj}).

One can show that (2.2.24) holds at any point at which $f(x)$ is differentiable and $\alpha_n(x) \to 0$. Indeed, if we do not take the h_{opt} but use $h = (\psi_1, \dots, \psi_n)$, then the error of the estimate will be not less that $\alpha_n(x)$. On the other hand, for this choice of h, the kernel (2.2.8) is the Dirichlet kernel. From the theory of the Fourier series one knows that (2.2.9) holds in L^2 and $f_n(x) \to f(x)$, $n \to \infty$, at any point at which $f(x)$ is differentiable.

In practice it is advisable to choose the system $\{\psi_j\}$ in such a way that $\alpha_n(x)$ tends rapidly to zero. Note that $\alpha_n(x)$ depends only on the system $\{\psi_j\}$, and therefore we can control this quantity to some extent by choosing the system $\{\psi_j\}$.

Let us note that one can estimate $f(x)$ at a given point x_0 optimally using the same procedure. In this case the convergence condition (2.2.4) will hold for x_0. If x_0 is fixed, we can choose the system ψ_j so that

$$b_{pj} = \delta_{pj} = \begin{cases} 0, & p \neq j, \\ 1, & p = j. \end{cases} \tag{2.2.30}$$

In this case

$$\alpha_n(x_0) := \alpha_n = \|a\|^{-2} = \left\{ \sum_{j=1}^{n} \left| \int_{-1}^{1} \psi_j^* dy \right|^2 \right\}^{-1}, \tag{2.2.31}$$

and we can choose ψ_j so that, in addition to (2.2.29), the condition

$$\|a\| \to \infty \quad \text{as } n \to \infty \tag{2.2.32}$$

is satisfied. For example, take $x_0 = 0$, then (2.2.29) reduces to

$$\int_{-1}^{1} y^2\psi_p(y)\psi_j^*(y)dy = \delta_{pj}, \tag{2.2.33}$$

and one can choose $\psi_p(y)$ which behave nearly like $|y|^{-1}$ in a small neighborhood of $y = 0$. Then $\int_{-1}^{1} \psi_j(y)dy$ can be made very large, and $\alpha_n \to 0$ as $n \to \infty$ in (2.2.31).

2.2.4.4. In this section we solve problem (2.2.18). As we have already mentioned, this problem may not be solvable for every $\sigma^2 > 0$, because there may be no h which satisfies both restrictions of (2.2.18). Since the set $\{h : (h, a) = 1, (Ch, h) \le \sigma^2\}$ is convex and $Q(h)$ is a strictly convex function of h, it is clear that the solution to (2.2.18) is unique when it exists. For the solution to exist it is necessary and sufficient that the set M of h, which satisfy the restrictions (2.2.18), be not empty.

Let us give an analytic solution to problem (2.2.18). If for the optimal h the inequality $(Ch, h) < \sigma^2$ holds, then the solution to (2.2.18) is the same as the solution to (2.2.13) and is given by formula (2.2.25). Therefore, first one checks if the function (2.2.25) satisfies the inequality

$$(CB^{-1}a, B^{-1}a) \cdot (B^{-1}a, a)^{-2} < \sigma^2. \tag{2.2.34}$$

If it does, then it is the solution to (2.2.18). If it does not, then the solution satisfies the equality

$$(Ch, h) = \sigma^2 \tag{2.2.35}$$

By the Lagrange method the necessary condition for the optimal h, for the solution to problem (2.2.18), is

$$Bh - \lambda Ch - \mu a = 0, \quad (h, a) = 1, \quad (Ch, h) = \sigma^2, \tag{2.2.36}$$

where λ and μ are the Lagrange multipliers. It follows from (2.2.36) that

$$(Bh, h) = \lambda^* \sigma^2 + \mu^*, \quad h = \mu(B - \lambda C)^{-1}a, \tag{2.2.37}$$

$$\mu = ((B - \lambda C)^{-1}a, a)^{-1}. \tag{2.2.38}$$

Taking the complex conjugate in the first equation (2.2.36) we see that

$$(Bh, h) = \lambda\sigma^2 + \mu \tag{2.2.39}$$

From (2.2.37) and (2.2.38) one gets

$$h = (B - \lambda C)^{-1}a/((B - \lambda C)^{-1}a, a). \tag{2.2.40}$$

Substituting (2.2.40) into

$$(Ch, h) = \sigma^2 \tag{2.2.41}$$

yields an equation for λ:

$$(C(B-\lambda C)^{-1}a,(B-\lambda C^{-1}a)=\sigma^2((B-\lambda C^{-1}a,a)^2. \tag{2.2.42}$$

The roots of Eq. (2.2.42), give μ by formula (2.2.38), and h by formula (2.2.37). Finally choose the h_{opt} for which $(Bh,h)=\min$. This h_{opt} solves problem (2.2.18).

2.2.4.5. One can simplify the solution to problem (2.2.18) in the following way. If $(Ch,h)=\sigma^2$ the problem (2.2.18) takes the form

$$(Bh,h)=\min,\quad (Ch,h)=\sigma^2,\quad (h,a)=1. \tag{2.2.43}$$

Let us choose the coordinate system so that

$$a_j=\delta_{jn}a_n,\quad \delta_{jn}=\begin{cases}0, & j\neq n,\\ 1, & j=n,\end{cases} \tag{2.2.44}$$

and normalize ψ so that

$$a_n=1. \tag{2.2.45}$$

In this case (2.2.43) can be written as

$$\begin{gathered}(\beta H,H)+2Re(H,\beta_n)+b_{nn}=\min,\\ (\gamma H,H)+2Re(H,\gamma_n)+C_{nn}=\sigma^2,\end{gathered} \tag{2.2.46}$$

where

$$\begin{aligned}&H=(h_1,\dots,h_{n-1}),\quad h_n=1;\quad \beta_{jp}=b_{jp},\quad 1\le j,p\le n-1,\\ &\beta_n=(b_{n1},\dots,b_{nn-1});\qquad \gamma_{pj}=C_{pj},\quad 1\le j,p\le n-1,\\ &\gamma_n=(C_{n1},\dots,C_{nn-1}).\end{aligned} \tag{2.2.47}$$

Thus, problem (2.2.43) in $\mathbb{C}^n$ with two constraints is reduced to problem (2.2.46) in $\mathbb{C}^{n-1}$ with one constraint. Problem (2.2.46) can be solved by the Lagrange multipliers method. One has

$$\beta H+\beta_n-\nu\gamma H-\nu\gamma_n=0, \tag{2.2.48}$$

where ν is the Lagrange multiplier. Thus,

$$H=-(\beta-\nu\gamma)^{-1}(\beta_n-\nu\gamma_n). \tag{2.2.49}$$

Substitute (2.2.49) into the constraint equation (2.2.46) to obtain an equation for ν. If this equation is solved then (2.2.49) gives the corresponding H. If there are several solutions then the H_{opt} is the one that minimizes the quadratic form (2.2.46).

2.3 ITERATIVE METHODS

There is a vast literature on iterative methods [VV], [BG], [R65]. First, we prove the following result.

Theorem 2.3.1. *Every solvable equation* (1.3.1) *with bounded linear operator A can be solved by a convergent iterative method.*

Proof. It was proved in Section 2.1.6 that if A is a bounded linear operator and equation (1.3.1) is solvable, then it is equivalent to the equation (2.1.6). By y we denote the minimal norm solution to (1.3.1), i.e., the solution orthogonal to $N(A)$, the null-space of A. Note that $N(B) = N(A)$. Without loss of generality assume $\|B\| \leq 1$ (if $\|B\| > 1$, then one can divide by $\|B\|$ equation (2.1.6)). Consider the iterations

$$u_{n+1} = u_n - (Bu_n - f_1), \quad u_0 = u_0, \quad f_1 := A^* f, \quad B := A^* A, \tag{2.3.1}$$

where u_0 is arbitrary. Denote $u_n - u := w_n$, and write (2.3.1) as $w_{n+1} = (I - B)w_n$, $w_0 := u_0 - u$.

Thus, $w_{n+1} = (I - B)^{n+1} w_0$. Since $0 \leq B \leq I$, one gets:

$$\|w_{n+1}\|^2 = \int_0^1 (1-\lambda)^{2n+2} d\left(E_\lambda w_0, w_0\right),$$

where E_λ is the resolution of the identity of B. Thus, $\lim_{n\to\infty} \|w_{n+1}\|^2 = \|Pw_o\|^2$, where P is the orthoprojector onto $N(B) = N(A)$. If one takes $u_0 \perp N(A)$, and $u = y$, $y \perp N(A)$, then $Pw_{n+1} = 0 \ \forall n$ by induction, and $\lim_{n\to\infty} \|w_n\| = 0$. In particular, if $u_0 = 0$, then $u_n \perp N(A) \ \forall n$ (by induction, since $u_1 = f_1 = A^* f \perp N(A)$), and $\lim_{n\to\infty} \|u_n - y\| = 0$, $y = A^+ f$, where A^+ is the pseudoinverse of A, defined in Section 2.1.1. □

Exercise 2.3.2. *Prove that if $f \notin D(A^+)$, then $\|u_n\| \to \infty$ in* (2.3.1).

Assume now that f_δ is given in place of f, $\|f_\delta - f\| \leq \delta$. Let us show that if one stops iterations (2.3.1), with f_δ in place of f and $u_0 = 0$, at $n = n(\delta)$, then $u_\delta := u_{n(\delta),\delta} \to y$ as $\delta \to 0$, if $n(\delta)$ is properly chosen, and $u_{n,\delta}$ is defined by (2.3.1). Let $u_n - u_{n,\delta} := w_n$. Then $w_{n+1} = w_n - (Bw_n - g)$, $g := f - f_\delta$, $\|g\| \leq \delta$, and $w_0 = 0$. Thus, $w_{n+1} = \sum_{j=0}^{n} (I - B)^j g$, $\|w_{n+1}\| \leq \delta n$, because $\|I - B\| \leq 1$ if $0 \leq B \leq I$, and we had assumed $\|B\| \leq 1$, which, together with $B \geq 0$, implies $0 \leq B \leq I$.

Therefore, if $n(\delta) \to \infty$ is chosen so that $\lim_{\delta\to 0} \delta n(\delta) = 0$, then $\lim_{\delta\to 0} \|u_\delta - y\| = 0$. Indeed, $\|u_\delta - y\| \leq \|u_\delta - u_{n(\delta)}\| + \|u_{n(\delta)} - y\|$. We have proved in Theorem 2.3.1

that $\lim_{\delta\to 0}\|u_{n(\delta)}-y\|=0$, and $\|u_\delta - u_{n(\delta)}\|\leq \delta n(\delta)\to 0$ as $\delta\to 0$. Let us summarize the result.

Theorem 2.3.3. *If* $\|f_\delta - f\|\leq\delta$, $\lim_{\delta\to 0} n(\delta)=\infty$, $\lim_{\delta\to 0}\delta n(\delta)=0$, *then* $\lim_{\delta\to 0}\|u_\delta - y\|=0$, *where* y *is the minimal norm solution to* (1.3.1), $u_\delta := u_{n(\delta),\delta}$, *and* $u_{n,\delta}$ *is obtained by iterations* (2.3.1), *with* f_δ *in place of* f *and* $u_0=0$.

A general approach to construction of convergent iterative methods for nonlinear problems is developed in Section 2.4.

2.4 DYNAMICAL SYSTEM METHOD (DSM)

2.4.1 The idea of the DSM

Consider the equation:

$$F(u)=0, \qquad F(u):=B(u)-f. \tag{2.4.1}$$

We assume in Section 2.4 that $F\in C^2_{\mathrm{loc}}$, that is,

$$\sup_{u\in B(u_0,R)}\left\|F^{(j)}(u)\right\|\leq M_j(R), \quad j=1,2, \tag{2.4.2}$$

where y is the minimal-norm solution to (2.4.1), $B(u_0,R):=\{u:\|u-u_0\|\leq R\}$, $F(y)=0$, $y\in B(u_0,R)$, $F:H\to H$, H is a real Hilbert space. Many of our results hold in reflexive Banach spaces X and $F:X\to X^*$, but we do not go into detail. The element u_0 in (2.4.2) will be specified later. In Section 2.4 we will call (2.4.1) a well-posed problem if

$$\sup_{u\in B(u_0,R)}\left\|\left[F'(u)\right]^{-1}\right\|\leq m(R), \tag{2.4.3}$$

and ill-posed if $F'(u)$ is not boundedly invertible. We assume existence of a solution to (2.4.1) unless otherwise stated, but uniqueness of the solution is not assumed.

If (2.4.3) holds, then one can construct Newton-type methods for solving (2.4.1). But if (2.4.3) fails, then it seems that there is no general approach to solving (2.4.1). One of our goals is to develop such an approach, which we call the Dynamical Systems Method (DSM). The DSM consists of finding a nonlinear locally Lipschitz operator $\Phi(t,u)$, such that the Cauchy problem:

$$\dot{u}=\Phi(t,u), \qquad u(0)=u_0, \tag{2.4.4}$$

has the following three properties:

$$\exists! u(t)\ \forall t\geq 0, \qquad \exists u(\infty), \quad F(u(\infty))=0, \tag{2.4.5}$$

that is, (2.4.4) is globally uniquely solvable, its unique solution has a limit at infinity $u(\infty)$, and this limit solves (2.4.1).

On first motivation is to develop a general approach to solving equation (2.4.1), especially nonlinear and ill-posed. Our second motivation is to develop a general approach to constructing convergent iterative methods for solving (2.4.1). We justify the DSM in the following cases

(1) For well-posed problems,
(2) For ill-posed linear problems with bounded linear operator A and $f \in R(A)$, and also for $f_\delta \notin R(A)$, $\|f_\delta - f\| \le \delta$.
(3) For ill-posed problem with monotone nonlinear A, for $f \in R(A)$, and also for $f_\delta \notin R(A)$, $\|f_\delta - f\| \le \delta$,
(4) For ill-posed problem with nonlinear A assuming some additional condition.

We give a general construction of convergent iterative schemes for well-posed nonlinear problems, and also for ill-posed nonlinear problems with monotone and non-monotone operators. These results are presented in subsections below.

2.4.2 DSM for well-posed problems

Consider (2.4.1), let (2.4.2) hold, and assume

$$\big(F'(u)\Phi(t,u), F(u)\big) \le -g_1(t)\big\|F(u)\big\|^a \quad \forall u \in B\big(u_0, R\big), \quad \int_0^\infty g_1 dt = \infty, \tag{2.4.6}$$

where $g_1 > 0$ is an integrable function, $a > 0$ is a constant. Assume

$$\big\|\Phi(t,u)\big\| \le g_2(t)\big\|F(u)\big\|, \quad \forall u \in B\big(u_0, R\big), \tag{2.4.7}$$

where $g_2 > 0$ is such that

$$G(t) := g_2(t)\exp\left(-\int_0^t g_1 ds\right) \in L^1\big(\mathbb{R}_+\big). \tag{2.4.8}$$

Remark 2.4.1. *Sometimes the assumption* (2.4.7) *can be used in the following modified form:*

$$\big\|\Phi(t,u)\big\| \le g_2(t)\big\|F(u)\big\|^b \quad \forall u \in B, \tag{2.4.7'}$$

where $b > 0$ is a constant. The statement and proof of Theorem 2.4.2 can be easily adjusted to this assumption.

Our first basic result is the folowing:

Theorem 2.4.2.

(i) If (2.4.6)–(2.4.8) *hold and*

$$\left\|F\left(u_0\right)\right\| \int_0^\infty G(t)dt \le R, \quad a = 2, \tag{2.4.9}$$

then (2.4.4) *has a global solution,* (2.4.5) *holds,* (2.4.1) *has a solution* $y = u(\infty) \in B(u_0, R)$*, and*

$$\|u(t) - y\| \le \|F(u_0)\| \int_t^\infty G(x)dx, \quad \|F(u(t))\| \le \|F(u_0)\| \exp\left(-\int_0^t g_1(x)dx\right). \tag{2.4.10}$$

(ii) If (2.4.6)–(2.4.8) *hold,* $0 < a < 2$*, and*

$$\left\|F\left(u_0\right)\right\| \int_0^T g_2 ds \le R, \tag{2.4.11}$$

where $T > 0$ *is defined by the equation*

$$\int_0^T g_1(s)ds = \left\|F\left(u_0\right)\right\|^{2-a} \Big/ (2-a), \tag{2.4.12}$$

then (2.4.4) *has a global solution,* (2.4.5) *holds,* (2.4.1) *has a solution* $y = u(\infty) \in B(u_0, R)$*, and* $u(t) = y$ *for* $t \ge T$*.*

(iii) If (2.4.6)–(2.4.8) *hold,* $a > 2$*, and*

$$\int_0^\infty g_2(s)h(s)ds \le R, \tag{2.4.13}$$

where

$$\left[\left\|F\left(u_0\right)\right\|^{2-a} + (a-2)\int_0^t g_1(s)ds\right]^{\frac{1}{2-a}} := h(t), \quad \lim_{t\to\infty} h(t) = 0, \tag{2.4.14}$$

then (2.4.4) *has a global solution,* (2.4.5) *holds,* (2.4.1) *has a solution* $y = u(\infty) \in B(u_0, R)$*, and*

$$\|u(t) - u(\infty)\| \le \int_t^\infty g_2(s)h(s)ds \longrightarrow 0$$

as $t \to \infty$*.*

Let us sketch the proof.

Proof of Theorem 2.4.2. The assumptions about Φ imply local existence and uniqueness of the solution $u(t)$ to (2.4.4). To prove global existence of u, it is sufficient to prove a uniform with respect to t bound on $\|u(t)\|$. Indeed, if the maximal interval of the existence of $u(t)$ is finite, say $[0, T)$, and $\Phi(t, u)$ is locally Lipschitz with respect to u, then $\|u(t)\| \to \infty$ as $t \to T$.

Assume $a = 2$. Let $g(t) := \|F(u(t))\|$. Since H is real, one uses (2.4.4) and (2.4.6) to get $g\dot{g} = (F'(u)\dot{u}, F) \leq -g_1(t)g^2$, so $\dot{g} \leq -g_1(t)g$, and integrating this inequality one gets the second inequality (2.4.10), because $g(0) = \|F(u_0)\|$. Using (2.4.7), (2.4.4) and the second inequality (2.4.10), one gets:

$$\left\|u(t) - u(s)\right\| \leq g(0) \int_s^t G(x)dx, \quad G(x) := g_2(x) \exp\left(-\int_0^x g_1(z)dz\right). \tag{2.4.10'}$$

Because $G \in L^1(R_+)$, it follows from (2.4.10') that the limit $y := \lim_{t\to\infty} u(t) = u(\infty)$ exists, and $y \in B$ by (2.4.9). From the second inequality (2.4.10) and the continuity of F one gets $F(y) = 0$, so y solves (2.4.1). Taking $t \to \infty$ and setting $s = t$ in (2.4.10') yields the first inequality (2.4.10). The inclusion $u(t) \in B$ for all $t \geq 0$ follows from (2.4.9) and (2.4.10'). The first part of Theorem 2.4.2 is proved. The proof of the other parts is similar. □

There are many applications of this theorem. We mention just a few, and assume that $g_1 = c_1 = const > 0$ and $g_2 = c_2 = const > 0$.

Example 2.4.3. *[Continuous Newton-type method]:* $\Phi = -[F'(u)]^{-1}F(u)$. *Assume that* (2.4.3) *holds, then* $c_1 = 1, c_2 = m_1$, (2.4.9) *takes the form (*)* $m_1(R)\|F(u_0)\| \leq R$, *and (*) implies that* (2.4.4) *has a global solution,* (2.4.5) *and* (2.4.10) *hold, and* (2.4.1) *has a solution in* $B(u_0, R)$. *This result belongs to Gavurin ([Gav]).*

Example 2.4.4. *[Continuous simple iterations method:] Let* $\Phi = -F$, *and assume* $F'(u) \geq c_1(R) > 0$ *for all* $u \in B(u_0, R)$. *Then* $c_2 = 1$, $c_1 = c_1(R)$, (2.4.9) *is:* $[c_1(R)]^{-1}$ $\|F(u_0)\| \leq R$, *and the conclusions of Example 1 hold.*

Example 2.4.5. *[Continuous gradient method:] Let* $\Phi = -[F']^*F$, (2.4.2) *and* (2.4.3) *hold,* $c_1 = m_1^{-2}$, $c_2 = M_1(R)$, (2.4.9) *is (**)* $M_1 m_1^2\|F(u_0)\| \leq R$, *and (**) implies the conclusions of Example 2.4.3.*

Example 2.4.6. *[Continuous Gauss-Newton method:] Let* $\Phi = -([F']^*F')^{-1}[F']^*F$, (2.4.2) *and* (2.4.3) *hold,* $c_1 = 1$, $c_2 = m_1^2 M_1$, (2.4.9) *is (***)* $M_1 m_1^2\|F(u_0)\| \leq R$, *and (***) implies the conclusions of Example 2.4.3.*

Example 2.4.7. *[Continuous modified Newton method:] Let* $\Phi = -[F'(u_0)]^{-1}F(u)$. *Assume* $\|[F'(u_0)]^{-1}\| \leq m_0$, *and let* (2.4.2) *hold. Then* $c_2 = m_0$. *Choose* $R = (2M_2 m_0)^{-1}$, *and* $c_1 = 0.5$. *Then* (2.4.9) *is* $2m_0\|F(u_0)\| \leq (2M_2m_0)^{-1}$, *that is,* $4m_0^2M_2\|F(u_0)\| \leq 1$. *Thus, if* $4m_0^2M_2\|F(u_0)\| \leq 1$, *then the conclusions of Example 2.4.3 hold.*

Example 2.4.8. *[Descent methods.] Let* $\Phi = -\frac{f}{(f',h)}h$, *where* $f = f(u(t))$ *is a differentiable functional* $f : H \to [0, \infty)$, *and* h *is an element of* H. *From* (2.4.4) *one gets* $\dot{f} = (f', \dot{u}) = -f$. *Thus,* $f = f_0e^{-t}$, *where* $f_0 := f(u_0)$. *Assume* $\|\Phi\| \leq c_2|f|^b$, $b > 0$. *Then* $\|\dot{u}\| \leq c_2|f_0|^b e^{-bt}$. *Therefore* $u(\infty)$ *does exist,* $f(u(\infty)) = 0$, *and* $\|u(\infty) - u(t)\| \leq ce^{-bt}$, $c = const > 0$.

If $h = f'$, and $f = \|F(u)\|^2$, then $f'(u) = 2[F']^*(u)F(u)$, $\Phi = -\frac{f}{\|f'\|^2}f'$, and (2.4.4) is a descent method. For this Φ one has $c_1 = \frac{1}{2}$, and $c_2 = \frac{m_1}{2}$, where m_1 is defined in (2.4.3). Condition (2.4.9) is: $m_1\|F(u_0)\| \leq R$. If this inequality holds, then the conclusions of Example 2.4.3 hold.

In Example 2.4.8 we have obtained some results from [Alb]. Our approach is more general than that in [Alb], since the choices of f and h do not allow one, for example, to obtain Φ used in Example 2.4.7.

Remark 2.4.9. *A method for proving the existence of a solution to equation* (2.4.1) *can be stated as follows. Consider* (2.4.4) *with* $\Phi = -[F'(u)]^{-1}F(u)$, *and assume that* (2.4.4) *is locally solvable and* $\|[F'(u(t))]^{-1}\| \leq a(t)$, *where* $u(t)$ *solves* (2.4.4). *Let* $g(t) := \|F(u(t))\|$. *Then* $g\dot{g} = (F'(u(t))\dot{u}, F) = -g^2$, *so* $g(t) = g_0e^{-t}$, *and* $\|\dot{u}\| \leq g_0a(t)e^{-t}$. *Assume that* $a(t)e^{-t} \in L^1(0, \infty)$. *Then* $u(\infty)$ *does exist and* $\|u(t) - u(\infty)\| \leq g_0\int_t^\infty a(s)e^{-s} ds \to 0$ *as* $t \to \infty$. *Therefore* $F(u(\infty)) = \lim_{t\to\infty} F(u(t)) = 0$, *so* $u(\infty)$ *solves* (2.4.1).

The proof of Theorem 2.4.2 is given by this method and Theorem 2.4.20 below is an example of many applications of this method.

Conditions (2.4.7) *and* (2.4.9) *are essential: if* $F(u) = e^u$ *and* H *is the real line with the usual product of real numbers as the inner product and* $|u| := \|u\|$, *then condition* (2.4.7) *is not satisfied and equation* (2.4.1), *i.e.* $e^u = 0$, *does not have a solution in* H.

Exercise (cf. [R222]). Use the above Remark to prove the following result:

Assume $F : H \to H$, *(2.4.2)–(2.4.3) hold, and* $\limsup_{R\to\infty} \frac{R}{m(R)} = \infty$. *Then the map* F *is surjective.*

This result is related to Hadamard's theorem about homeomorphisms and its generalization by Meyer (see [OR], p. 139).

2.4.3 Linear ill-posed problems

We assume that (2.4.3) fails. Consider

$$Au = f. \tag{2.4.15}$$

Let us denote by (A) the folowing assumption:

(A) : *A is a linear, bounded operator in H, defined on all of H, the range R(A) is not closed, so* (2.4.15) *is an ill-posed problem, there is a y such that $Ay = f$, $y \perp N$, where N is the null-space of A.*

Let $B := A^*A$, $q := By = A^*f$, A^* is the adjoint of A. Every solution to (2.4.15) solves

$$Bu = q, \tag{2.4.16}$$

and, if $f = Ay$, then every solution to (2.4.16) solves (2.4.15). Choose a continuous function $\varepsilon(t) > 0$, monotonically decaying to zero on $\mathbb{R}_+$. Sometimes it is convenient to assume that

$$\lim_{t\to\infty} \left(\dot{\varepsilon}\varepsilon^{-2}\right) = 0. \tag{2.4.17}$$

For example, the functions $\varepsilon = c_1(c_0 + t)^{-b}$, $0 < b < 1$, where c_0 and c_1 are positive constants, satisfy (2.4.17). There are many such functions. One can prove the following:

Claim 3. *If $\varepsilon(t) > 0$ is a continuous monotonically decaying function on $\mathbb{R}_+$, $\lim_{t\to\infty}\varepsilon(t) = 0$, and* (2.4.17) *holds, then*

$$\int_0^\infty \varepsilon ds = \infty. \tag{2.4.17'}$$

In this Section we do not use assumption (2.4.17): in the proof of Theorem 2.4.9 one uses only the monotonicity of a continuous function $\varepsilon > 0$ and (2.4.17'). One can drop assumption (2.4.17'), but then convergence is proved in Theorem 2.4.9 to some element of N, not necessarily to the normal solution y, that is, to the solution orthogonal to N, or, which is the same, to the minimal-norm solution to (2.4.15). However, (2.4.17) is used (in a slightly weaker form) in the next section.

Consider problems (2.4.4) with

$$\Phi := -\left[Bu + \varepsilon(t)u - q\right], \quad \Phi_\delta = -\left[Bu_\delta + \varepsilon(t)u_\delta - q_\delta\right], \tag{2.4.18}$$

where $\|q - q_\delta\| \le \|A^*\|\delta := C\delta$. Without loss of generality one may assume that $C = \|A^*\| = 1$, which we do in what follows. Our main result is Theorem 2.4.9, stated below. It yields the following:

Conclusion: *Given noisy data f_δ, every linear ill-posed problem* (2.4.15) *under the assumptions (A) can be stably solved by the DSM.*

The result presented in Theorem 2.4.9 was essentially obtained in [R200], but the proof given here is different and much shorter.

Theorem 2.4.9. *Problem* (2.4.4) *with Φ from* (2.4.18) *has a unique global solution $u(t)$,* (2.4.5) *holds, and $u(\infty) = y$. Problem* (2.4.4) *with Φ_δ from* (2.4.18) *has a unique global* *solution $u_\delta(t)$. There exists t_δ, such that*

$$\lim_{\delta\to 0} \left\| u_\delta(t_\delta) - y \right\| = 0. \tag{2.4.19}$$

This t_δ can be chosen, for example, as a root of the equation

$$\varepsilon(t) = \delta^b, \quad b \in (0,1), \tag{2.4.20}$$

or of the equation (2.4.20'), *see below.*

Proof of Theorem 2.4.9. Linear equations (2.4.4) with bounded operators have unique global solutions. If $\Phi = -[Bu + \varepsilon(t)u - q]$, then the solution u to (2.4.4) is

$$u(t) = h^{-1}(t)U(t)u_0 + h^{-1}(t)\int_0^{\|B\|} \exp(-t\lambda)\int_0^t e^{s\lambda}h(s)ds\lambda dE_\lambda y, \tag{2.4.21}$$

where $h(t) := \exp(\int_0^t \varepsilon(s)ds) \to \infty$ as $t \to \infty$, E_λ is the resolution of the identity corresponding to the selfadjoint operator B, and $U(t) := e^{-tB}$ is a nonexpansive operator, because $B \geq 0$. Actually, (2.4.21) can be used also when B is unbounded, $\|B\| = \infty$.

Using L'Hospital's rule one checks that

$$\lim_{t\to\infty} \frac{\lambda\int_0^t e^{s\lambda}h(s)ds}{e^{t\lambda}h(t)} = \lim_{t\to\infty} \frac{\lambda e^{t\lambda}h(t)}{\lambda e^{t\lambda}h(t) + e^{t\lambda}h(t)\epsilon(t)} = 1 \quad \forall\lambda > 0, \tag{2.4.22}$$

provided only that $\varepsilon(t) > 0$ and $\lim_{t\to\infty}\varepsilon(t) = 0$. From (2.4.21), (2.4.22), and the Lebesgue dominated convergence theorem, one gets $u(\infty) = y - Py$, where P is the orthogonal projection operator onto the null-space of B. Under our assumptions (A), $Py = 0$, so $u(\infty) = y$. If $v(t) := \|u(t) - y\|$, then $\lim_{t\to\infty} v(t) = 0$. In general, the rate of convergence of v to zero can be arbitrarily slow for a suitably chosen f. Under an additional a priori assumption on f (for example, the source-type assumptions), this rate can be estimated.

Let us describe a method for deriving a stopping rule. One has:

$$\|u_\delta(t) - y\| \leq \|u_\delta(t) - u(t)\| + v(t). \tag{2.4.23}$$

Since $\lim_{t\to\infty} v(t) = 0$, any choice of t_δ such that

$$\lim_{t_\delta\to\infty} \|u_\delta(t_\delta) - u(t_\delta)\| = 0, \tag{2.4.24}$$

gives a stopping rule: for such t_δ one has $\lim_{\delta\to 0}\|u_\delta(t) - y\| = 0$.

To prove that (2.4.20) gives such a rule, it is sufficient to check that

$$\|u_\delta(t) - u(t)\| \leq \frac{\delta}{\varepsilon(t)}. \tag{2.4.25}$$

Let us prove (2.4.25). Denote $w := u_\delta - u$, $p := q_\delta - q$. Then

$$\dot{w} = -[Bw + \varepsilon w - p], \quad w(0) = 0, \quad \|p\| \leq \delta. \tag{2.4.26}$$

Integrating (2.4.26), and using the property $B \geq 0$, one gets (2.4.25).

Alternatively, multiply (2.4.26) by w, let $\|w\| := g$, use $B \geq 0$, and get $\dot{g} \leq -\varepsilon(t)g + \delta$, $g(0) = 0$. Thus, $g(t) \leq \delta \exp(-\int_0^t \varepsilon ds) \int_0^t \exp(\int_0^s \varepsilon d\tau) ds \leq \frac{\delta}{\varepsilon(t)}$. A more precise estimate, used at the end of the proof of Theorem 2.4.10 below, yields:

$$\|u_\delta(t) - u(t)\| \leq \frac{\delta}{2\sqrt{\varepsilon(t)}},$$

and the corresponding stopping time t_δ can be taken as the root of the equation:

$$2\sqrt{\varepsilon(t)} = \delta^b, \quad b \in (0, 1). \tag{2.4.20'}$$

Theorem 2.4.9 is proved. □

If the rate of decay of v is known, then a more efficient stopping rule can be derived: t_δ is the minimizer of the problem:

$$v(t) + \delta[\varepsilon(t)]^{-1} = \min. \tag{2.4.27}$$

For example, if $v(t) \leq c\varepsilon^a(t)$, then t_δ is the root of the equation $\varepsilon(t) = (\frac{\delta}{ca})^{\frac{1}{1+a}}$, that one gets from (2.4.27) with $v = c\varepsilon^a$.

One can also use a stopping rule based on an a posteriori choice of the stopping time, for example, the choice by a discrepancy principle.

A method, much more efficient numerically than Theorem 2.4.9, is given below in Theorem 2.4.12 and in Theorem 2.4.10 (see (2.4.29)).

For linear equation (2.4.16) with exact data this method uses (2.4.4) with

$$\Phi = -(B + \varepsilon(t))^{-1}[Bu + \varepsilon(t)u - q] = -u + (B + \varepsilon(t))^{-1}q, \tag{2.4.28}$$

and for noisy data it uses (2.4.4) with $\Phi_\delta = -u_\delta + (B + \varepsilon(t))^{-1}q_\delta$. The linear operator $B \geq 0$ is monotone, so Theorem 2.4.12 is applicable. For exact data, (2.4.4) with Φ, defined in (2.4.28), yields:

$$\dot{u} = -u + (B + \varepsilon(t))^{-1}q, \quad u(0) = u_0, \tag{2.4.29}$$

and (2.4.5) holds if $\varepsilon(t) > 0$ is monotone, continuous, decreasing to 0 as $t \to \infty$.

Let us formulate the result:

Theorem 2.4.10. *Assume (A), and let $B := A^*A$, $q := A^*f$. Assume $\varepsilon(t) > 0$ to be a continuous, monotonically decaying to zero function on* $[0, \infty)$. *Then, for any $u_0 \in H$, problem* (2.4.29) *has a unique global solution, $\exists u(\infty) = y$, $Ay = f$, and y is the minimal-norm solution to* (2.4.15). *If f_δ is given in place of f, $\|f - f_\delta\| \le \delta$, then* (2.4.19) *holds, with $u_\delta(t)$ solving* (2.4.29) *with q replaced by $q_\delta := A^* f_\delta$, and t_δ is chosen, for example, as the root of* (3.4.20') *(or by a discrepancy principle).*

Proof of Theorem 2.4.10. One has $q = Bz$, where $Az = f$, and the solution to (2.4.29) is

$$u(t) = u_0 e^{-t} + e^{-t} \int_0^t e^s (B + \varepsilon(s))^{-1} Bz ds := u_0 e^{-t} + \int_0^{\|B\|} j(\lambda, t) d E_\lambda z \tag{2.4.30}$$

where

$$j(\lambda, t) := \int_0^t \frac{\lambda e^s}{[\lambda + \varepsilon(s)] e^t} ds, \tag{2.4.31}$$

and E_λ is the resolution of the identity of the selfadjoint operator B. One has

$$0 \le j(\lambda, t) \le 1, \quad \lim_{t\to\infty} j(\lambda, t) = 1 \;\; \lambda > 0, \quad j(0, t) = 0. \tag{2.4.32}$$

From (2.4.30)–(2.4.32) it follows that $\exists u(\infty)$, $u(\infty) = z - P_N z = y$, where y is the minimal-norm solution to (2.4.15), $N := N(B) = N(A)$ is the null-space of B and of A, and P_N is the orthoprojector onto N in H. This proves the first part of Theorem 2.4.10.

To prove the second part, denote $w := u_\delta - u, g := f_\delta - f$, where we dropped the dependence on δ in w and g for brevity. Then $\dot{w} = -w + (B + \varepsilon(t))^{-1} A^* g$, $w(0) = 0$. Thus $w = e^{-t} \int_0^t e^s (B + \varepsilon(s))^{-1} A^* g ds$, so $\|w\| \le \delta e^{-t} \int_0^t \frac{e^s}{2\sqrt{\varepsilon(s)}} ds \le \frac{\delta}{2\sqrt{\varepsilon(t)}}$, where the known estimate was used: $\|(B + \varepsilon)^{-1} A^*\| \le \frac{1}{2\sqrt{\varepsilon}}$. Theorem 2.4.10 is proved. $\square$

2.4.4 Nonlinear ill-posed problems with monotone operators

There is a large body of literature on equations (2.4.1) and (2.4.4) with monotone operators. In the result we present, the problem is nonlinear and ill-posed, the new technical tool, Theorem 2.4.11, is used, and the stopping rules are discussed.

Consider (2.4.33) with monotone F under standard assumptions (2.4.2), and

$$\Phi = -A^{-1}_{\varepsilon(t)}(u)\big[F(u(t)) + \varepsilon(t)(u(t) - \tilde{u}_0)\big], \tag{2.4.33}$$

where $A = A(u) := F'(u)$, A^* is its adjoint, $\varepsilon(t)$ is the same as in Theorem 2.4.10, and in Theorem 2.4.12 $\varepsilon(t)$ is further specified, $\tilde{u}_0 \in B(u_0, R)$ is an element we can choose to improve the numerical performance of the method. If noisy data are given, then, as in Section 3.3, we take

$$F(u) := B(u) - f, \quad \Phi_\delta = -A^{-1}_{\varepsilon(t)}(u_\delta)\big[B(u_\delta(t)) - f_\delta + \varepsilon(t)(u_\delta(t) - \tilde{u}_0)\big],$$

where $\|f_\delta - f\| \le \delta$, B is a monotone nonlinear operator, $B(y) = f$, and u_δ solves (2.4.4) with Φ_δ in place of Φ.

To prove that (2.4.33) with the above Φ has a global solution and (2.4.5) holds, we use the following:

Theorem 2.4.11. *Let $\gamma(t), \sigma(t), \beta(t) \in C[t_0, \infty)$ for some real number $t_0 \ge 0$. If there exists a positive function $\mu(t) \in C^1[t_0, \infty)$ such that*

$$0 \le \sigma(t) \le \frac{\mu(t)}{2}\left[\gamma(t) - \frac{\dot{\mu}(t)}{\mu(t)}\right], \quad \beta(t) \le \frac{1}{2\mu(t)}\left[\gamma(t) - \frac{\dot{\mu}(t)}{\mu(t)}\right], \quad g_0\mu(t_0) < 1, \tag{2.4.34}$$

where g_0 is the initial condition in (2.4.35), then a nonnegative solution g to the following differential inequality:

$$\dot{g}(t) \le -\gamma(t)g(t) + \sigma(t)g^2(t) + \beta(t), \quad g(t_0) = g_0 \ge 0, \tag{2.4.35}$$

exists for all $t \ge t_0$ and satisfies the estimate:

$$0 \le g(t) \le \frac{1 - \nu(t)}{\mu(t)} < \frac{1}{\mu(t)}, \tag{2.4.36}$$

for all $t \in [t_0, \infty)$, where

$$0 < \nu(t) = \left(\frac{1}{1 - \mu(t_0)g(t_0)} + \frac{1}{2}\int_{t_0}^{t}\left(\gamma(s) - \frac{\dot{\mu}(s)}{\mu(s)}\right)ds\right)^{-1}. \tag{2.4.37}$$

There are several novel features in this result. First, differential equation, that one gets from (2.4.35) by replacing the inequality sign by the equality sign, is a Riccati equation, whose solution may blow up in a finite time, in general. Conditions (2.4.34) guarantee the global existence of the solution to this Riccati equation with the initial condition (2.4.35). Secondly, this Riccati differential equation cannot be integrated analytically by separation of variables. Thirdly, the coefficient $\sigma(t)$ may grow to infinity as $t \to \infty$, so that the quadratic term does not necessarily have a small coefficient, or the coefficient smaller than $\gamma(t)$. Without loss of generality one may assume $\beta(t) \ge 0$ in Theorem 2.4.11. This Theorem is proved in Section 2.4.9.

The main result in this Section is new. It claims a global convergence in the sense that no assumptions on the choice of the initial approximation u_0 are made. Usually one assumes that u_0 is sufficiently close to the solution of (2.4.1) in order to prove convergence. We take in Theorem 2.4.12 $\tilde{u}_0 = 0$ because in this theorem $\tilde{u}_0$ does not play any role. The proof is valid for any choice of $\tilde{u}_0$, but then the definition of r in Theorem 2.4.12 is changed.

Theorem 2.4.12. *If* (2.4.2) *holds,* $\tilde{u}_0 = 0$*,* $R = 3r$*, where* $r := \|y\| + \|u_0\|$*, and* $y \in N := \{z : F(z) = 0\}$ *is the (unique) minimal norm solution to* (2.4.1)*, then, for any choice of* u_0*, problem* (2.4.4) *with* Φ *defined in* (2.4.33)*,* $\tilde{u}_0 = 0$*, and* $\varepsilon(t) = c_1(c_0 + t)^{-b}$ *with some positive constants* c_1*,* c_0*, and* $b \in (0, 1)$*, specified in the proof of Theorem 2.4.12, has a global solution, this solution stays in the ball* $B(u_0, R)$ *and* (2.4.5) *holds. If* $u_\delta(t)$ *solves* (2.4.4) *with* Φ_δ *in place of* Φ*, then there is a* t_δ *such that* $\lim_{\delta\to 0} \|u_\delta(t_\delta) - y\| = 0$.

Proof of Theorem 2.4.12. Let us sketch the steps of the proof. Let V solve the equation

$$F(V) + \varepsilon(t)V = 0. \tag{2.4.38}$$

Under our assumptions on F, it is easy to prove that:

(i) (2.4.38) has a unique solution for every $t > 0$, and
(ii) $\sup_{t\geq 0} \|V\| \leq \|y\|$.
Indeed, $(F(V) - F(y), V - y) \geq 0$, $F(y) = 0, \varepsilon > 0$, so $(V, V - y) \leq 0$. This implies (ii).
If F is Fréchet differentiable, then V is differentiable, and $\|\dot{V}(t)\| \leq \|y\||\dot{\varepsilon}(t)|/\varepsilon(t)$.
It is also known (see Section 2.1.4) that if $F(y) = 0$, where y is the minimal-norm solution to (2.4.1), then $\lim_{t\to\infty} \|V(t) - y\| = 0$.

We will show that the global solution u to (2.4.4), with the Φ from (2.4.33), does exist, and $\lim_{t\to\infty} \|u(t) - V(t)\| = 0$. This is done by deriving a differential inequality for $w := u - V$, and by applying Theorem 2.4.11 to $g = \|w\|$. Since $\|u(t) - y\| \leq \|u(t) - V(t)\| + \|V(t) - y\|$, one obtains (2.4.5). We also check that $u(t) \in B(u_0, R)$, where $R := 3(\|y\| + \|u_0\|)$, for any choice of u_0 and a suitable choice of ε.

Let us derive the differential inequality for w. One has

$$\dot{w} = -\dot{V} - A^{-1}_{\varepsilon(t)}(u)\Big[F\big(u(t)\big) - F\big(V(t)\big) + \varepsilon(t)w\Big], \tag{2.4.39}$$

and $F(u) - F(V) = Aw + K$, where $\|K\| \leq M_2 g^2/2$, $g := \|w\|$ and M_2 is the constant from (2.4.2). Multiply (2.4.39) by w, use the monotonicity of F, that is, the property $A \geq 0$, and the estimate $\|\dot{V}\| \leq \|y\||\dot{\varepsilon}|/\varepsilon$, and get:

$$\dot{g} \leq -g + \frac{0.5Mg^2}{\varepsilon} + \|y\|\frac{|\dot{\varepsilon}|}{\varepsilon}, \tag{2.4.40}$$

where $M := M_2$. Inequality (2.4.40) is of the type (2.4.35): $\gamma = 1$, $\sigma = 0.5M/\varepsilon$, $\beta = \|y\|\frac{|\dot\varepsilon|}{\varepsilon}$. Choose

$$\mu(t) = \frac{2M}{\varepsilon(t)}. \tag{2.4.41}$$

Clearly $\mu \to \infty$ as $t \to \infty$. Let us check three conditions (2.4.34). One has $\frac{\dot\mu(t)}{\mu(t)} = \frac{|\dot\varepsilon|}{\varepsilon}$. Take $\varepsilon = c_1(c_0 + t)^{-b}$, where $c_j > 0$ are constants, $0 < b < 1$, and choose these constants so that $\frac{|\dot\varepsilon|}{\varepsilon} < \frac{1}{2}$, for example, $\frac{b}{c_0} = \frac{1}{4}$. Then the first condition (2.4.34) is satisfied. The second condition (2.4.34) holds if

$$8M\|y\|\,|\dot\varepsilon|\varepsilon^{-2} \le 1. \tag{2.4.42}$$

One has $\varepsilon(0) = c_1c_0^{-b}$. Choose

$$\varepsilon(0) = 4Mr. \tag{2.4.43}$$

Then

$$|\dot\varepsilon|\varepsilon^{-2} = bc_1^{-1}(c_0 + t)^{b-1} \le bc_0^{-1}c_1^{-1}c_0^b = \frac{1}{4\varepsilon(0)} = \frac{1}{16Mr}, \tag{2.4.44}$$

so (2.4.42) holds. Thus, the second condition (2.4.34) holds. The last condition (2.4.34) holds because

$$\frac{2M\|u_0 - V_0\|}{\varepsilon(0)} \le \frac{2Mr}{4Mr} = \frac{1}{2} < 1.$$

By Theorem 2.4.11 one concludes that $g = \|w(t)\| < \frac{\varepsilon(t)}{2M} \to 0$ when $t \to \infty$, and

$$\|u(t) - u_0\| \le g + \|V - u_0\| \le g(0) + r \le 3r. \tag{2.4.45}$$

This estimate implies the global existence of the solution to (2.4.4), because if $u(t)$ had a finite maximal interval of existence, $[0, T)$, then $u(t)$ could not stay bounded when $t \to T$, which contradicts the boundedness of $\|u(t)\|$, and from (2.4.45) it follows that $\|u(t)\| \le 4r$. We have proved the first part of Theorem 2.4.12, namely properties (2.4.5).

To derive a stopping rule we argue as in Section 2.4. One has:

$$\left\|u_\delta(t) - y\right\| \le \left\|u_\delta(t) - V(t)\right\| + \left\|V(t) - y\right\|. \tag{2.4.46}$$

We have already proved that $\lim_{t\to\infty} v(t) := \lim_{t\to\infty} \|V(t) - y\| = 0$. The rate of decay of v can be arbitrarily slow, in general. Additional assumptions, for example, the source-type ones, can be used to estimate the rate of decay of $v(t)$. One derives differential inequality (2.4.35) for $g_\delta := \|u_\delta(t) - V(t)\|$, and estimates g_δ using (2.4.36). The analog of (2.4.40) for g_δ contains additional term $\frac{\delta}{\varepsilon}$ on the right-hand side. If

$\frac{\delta}{\varepsilon^2} \leq \frac{1}{16M}$, then conditions (2.4.34) hold, and $g_\delta < \frac{\varepsilon(t)}{2M}$. Let t_δ be the root of the equation $\varepsilon^2(t) = 16M\delta$. Then $\lim_{\delta\to 0} t_\delta = \infty$, and $\lim_{\delta\to 0} \|u_\delta(t_\delta) - y\| = 0$ because $\|u_\delta(t_\delta) - y\| \leq v(t_\delta) + g_\delta$, $\lim_{t_\delta\to\infty} g_\delta(t_\delta) = 0$ and $\lim_{t_\delta\to\infty} v(t_\delta) = 0$, but the convergence can be slow. See [ARS3], [KNR] for the rate of convergence under source assumptions. If the rate of decay of $v(t)$ is known, then one chooses t_δ as the minimizer of the problem, similar to (2.4.27),

$$v(t) + g_\delta(t) = \min, \tag{2.4.47}$$

where the minimum is taken over $t > 0$ for a fixed small $\delta > 0$. This yields a quasioptimal stopping rule. Theorem 2.4.12 is proved. □

Let us give another result:

Theorem 2.4.13. *Assume that $\Phi = -F(u) - \varepsilon(t)u$, F is monotone, $\varepsilon(t)$ as in Theorem 2.4.10, and* (2.4.17)*, and* (2.4.2) *hold. Then* (2.4.5) *holds.*

Proof of Theorem 2.4.13. As in the proof of Theorem 2.4.12, it is sufficient to prove that $\lim_{t\to\infty} g(t) = 0$, where g, w, and V are the same as in Theorem 2.4.12, and u solves (2.4.4) with the Φ defined in Theorem 2.4.13. Similarly to the derivation of (2.4.39), one gets:

$$\dot{w} = -\dot{V} - \left[F(u) - F(V) + \varepsilon(t)w\right]. \tag{2.4.48}$$

Multiply (2.4.48) by w, use the monotonicity of F and the estimate $\|\dot{V}\| \leq \frac{|\dot{\varepsilon}(t)|}{\varepsilon(t)}\|y\|$, which was used also in the proof of Theorem 2.4.12, and get:

$$\dot{g} \leq -\varepsilon(t)g + \frac{|\dot{\varepsilon}(t)|}{\varepsilon(t)}\|y\|. \tag{2.4.49}$$

This implies

$$g(t) \leq e^{-\int_0^t \varepsilon(s)ds}\left[g(0) + \int_0^t e^{\int_0^s \varepsilon(x)dx}\frac{|\dot{\varepsilon}(s)|}{\varepsilon(s)}\|y\|ds\right]. \tag{2.4.50}$$

From our assumptions relation (2.4.17') follows, and (2.4.50) together with (2.4.17) imply $\lim_{t\to\infty} g(t) = 0$. Theorem 2.4.13 is proved. □

Remark 2.4.14. *One can drop the smoothness of F assumption* (2.4.2) *in Theorem 2.4.13 and assume only that F is a monotone hemicontinuous operator defined on all of H.*

Claim 4. *If $\varepsilon(t) = \varepsilon = const > 0$, then $\lim_{\varepsilon\to 0} \|u(t_\varepsilon) - y\| = 0$, where $u(t)$ solves* (2.4.4) *with $\Phi := -F(u) - \varepsilon u$, and t_ε is any number such that $\lim_{\varepsilon\to 0} \varepsilon t_\varepsilon = \infty$.*

Proof of the claim. One has $\|u(t)-y\| \leq \|u(t)-V_\varepsilon\| + \|V_\varepsilon - y\|$, where V_ε solves (2.4.38) with $\varepsilon(t) = \varepsilon = const > 0$. Under our assumptions on F, equation (2.4.38) has a unique solution, and $\lim_{\varepsilon\to 0}\|V_\varepsilon - y\| = 0$. So, to prove the claim, it is sufficient to prove that $\lim_{\varepsilon\to 0}\|u(t_\varepsilon) - V_\varepsilon\| = 0$, provided that $\lim_{\varepsilon\to 0}\varepsilon t_\varepsilon = \infty$. Let $g := \|u(t) - V_\varepsilon\|$, and $w := u(t) - V_\varepsilon$. Because $\dot{V}_\varepsilon = 0$, one has the equation: $\dot{w} = -[F(u) - F(V_\varepsilon) + \varepsilon w]$. Multiplying this equation by w, and using the monotonicity of F, one gets $\dot{g} \leq -\varepsilon g$, so $g(t) \leq g(0)e^{-\varepsilon t}$. Therefore $\lim_{\varepsilon\to 0} g(t_\varepsilon) = 0$, provided that $\lim_{\varepsilon\to 0}\varepsilon t_\varepsilon = \infty$. The claim is proved. □

Remark 2.4.15. *One can prove claims (i) and (ii), formulated below formula* (2.4.38)*, using DSM version presented in Theorem 2.4.20 below.*

Claim 5. *Assume that F is monotone,* (2.4.2) *holds, and $F(y) = 0$. Then claims (i) and (ii), formulated below formula* (2.4.38)*, hold.*

Proof. First, note that (ii) follows easily from (i), because the assumptions $F(y) = 0$, F is monotone, and $\varepsilon > 0$, imply, after multiplying $F(V) - F(y) + \varepsilon V = 0$ by $V - y$, the inequality $(V, V - y) \leq 0$, from which claim (ii) follows. Claim (i) follows from Theorem 2.4.20, proved below. □

Claim 6. *Assume that the operator F is monotone, hemicontinuous, defined on all of H, equation $F(u) = 0$ has a solution, possibly non-unique, y is the minimal-norm element of $N_F := \{z : F(z) = 0\}$, $\Phi = -F(u) - \varepsilon(t)u$, and $\varepsilon = c_1(c_0 + t)^{-b}$, $0 < b < 1$, where $c_0 > 0$, $c_1 > 0$ and b are constants. Then* (2.4.5) *holds for the solution to* (2.4.4).

Proof. The steps of the proof are:

1) we prove that $\sup_{t\geq 0}\|u(t)\| < \infty$ for the solution to (2.4.4) with $\Phi(t,u) := -F(u) - \epsilon(t)u$, where $0 < \epsilon(t) \searrow 0$, $\int_0^\infty \epsilon(s)ds = \infty$. Inequality $\sup_{t\geq 0}\|u(t)\| < \infty$ implies that from any sequence $t_n \to \infty$ one can select a subsequence, denoted again t_n, such that $u(t_n) := u_n \rightharpoonup v$, where $v \in H$ is some element. We prove that $F(v) = 0$.
2) we prove that the solution $u(t)$ to the equation

$$\dot{u} = -F(u) - \epsilon(t)u, \quad u(0) = u_0, \tag{*}$$

satisfies the relation: $g(t) := \|u(t+h) - u(t)\| \to 0$ as $t \to \infty$, where $h > 0$ is an arbitrary fixed number.
3) we prove that $\|\dot{u}\| = O(\frac{1}{t})$ as $t \to \infty$.
4) passing to the limit $t_n \to \infty$ in equation (*) yields $F(v) = 0$. We prove that $u(t) \to v$ as $t \to \infty$, and $v = y$.

Let us give the details of the proof.

1) Let $F(y) = 0$, $u - y := w$, where u solves (*).Then $\dot{w} = -[F(u) - F(w)] - \epsilon(t)w - \epsilon(t)y$. Multiply this equation by w, use the monotonicity of F, denote $\|w\| := z(t)$, and get $z\dot{z} \leq -\epsilon(t)z^2 + \epsilon(t)\|y\|z$. Because $z \geq 0$, this implies

$\dot{z} \leq -\epsilon(t)z + \epsilon(t)\|y\|$, so $z(t) \leq e^{-\int_0^t \epsilon(s)ds}[z(0) + \int_0^t e^{\int_0^s \epsilon(s')ds'}\epsilon(s)ds\,\|y\|]$. This inequality implies $\sup_{t\geq 0}\|u(t)\| < \|z(0)\| + \|y\| < \infty$. Thus $u(t_n) \rightharpoonup v$ as $t_n \to \infty$.

2) Denote $\eta := u(t+h) - u(t)$, where u solves (*), and $\|\eta\| := g(t)$. Then $\dot{\eta} = -[F(u(t+h)) - F(u(t))] - [\epsilon(t+h)u(t+h) - \epsilon(t)u(t)]$. Multiply this equation by η and use the monotonicity of F to get: $g\dot{g} \leq -\epsilon(t)g^2 + |\dot{\epsilon}(t)|h\|u(t+h)\|g$. Because $g \geq 0$ and $\|u(t+h)\| < c$, one gets

$$\dot{g} \leq -\epsilon(t)g + |\dot{\epsilon}(t)|hc. \tag{2.4.51}$$

This implies

$$g(t) \leq e^{-\int_0^t \epsilon(s)ds}\left[g(0) + hc\int_0^t e^{\int_0^s \epsilon(s')ds'}|\dot{\epsilon}(s)|ds\right]. \tag{2.4.52}$$

Under our assumptions about $\epsilon(t)$, one can check that $\int_0^t \epsilon(s)ds = O(t^a)$, as $t \to \infty$, where $0 < a := 1 - b < 1$. Also $e^{-\int_0^t \epsilon(s)ds}\int_0^t e^{\int_0^s \epsilon(s')ds'}|\dot{\epsilon}(s)|ds = O(\frac{1}{t})$. Thus $g(t) = O(\frac{1}{t})$ as $t \to \infty$.

3) Denote $G(t) := \|\dot{u}\|$. Divide (2.4.52) by h and let $h \to 0$. Then one gets $G(t) = O(\frac{1}{t})$, so one has $\|\dot{u}(t)\| = O(\frac{1}{t})$ as $t \to \infty$.

4) Passing to the limit $t = t_n \to \infty$ in equation (*), yields $F(v) = 0$. The limit $\lim_{t_n\to\infty} F(u(t_n)) = F(v)$ exists because $\epsilon(t_n) \to 0$, $\sup_n \|u(t_n)\| < \infty$, $\lim_{t_n\to\infty}\|\dot{u}(t_n)\| = 0$, and $u(t_n) \rightharpoonup v$ as $t_n \to \infty$, so that Lemma 2.1.2 implies $F(v) = 0$. Let us prove that $u(t_n) \to v$. Since $u(t_n) \rightharpoonup v$, one gets $\|v\| \leq \liminf_{n\to\infty}\|u(t_n)\|$. If $\limsup_{n\to\infty}\|u(t_n)\| \leq \|v\|$, then $\lim_{n\to\infty}\|u(t_n)\| = \|v\|$, and together with the weak convergence $u(t_n) \rightharpoonup v$ this implies strong convergence $u(t_n) \to v$.

Let us prove that $\limsup_{n\to\infty}\|u(t_n)\| \leq \|v\|$. One has $(F(u(t_n)) - F(v), u(t_n) - v) + \epsilon(t_n)(u(t_n), u(t_n) - v) = -(\dot{u}(t_n), u(t_n) - v)$. Since F is monotone, $\|\dot{u}(t)\| = O(\frac{1}{t})$ and $\|u(t_n) - v\| \leq c$, it follows that $(u(t_n), u(t_n) - v) \leq \frac{c}{t_n\epsilon(t_n)}$. Thus, $\limsup_{n\to\infty}\|u(t_n)\| \leq \|v\|$, because $\lim_{n\to\infty} t_n\epsilon(t_n) = \infty$.

Let us prove that $v = y$. Replacing v by y in the above argument yields $(u(t_n), u(t_n) - y) \leq \frac{c}{t_n\epsilon(t_n)}$, so $\|v\| = \lim_{n\to\infty}\|u(t_n)\| \leq \|y\|$. Since y is the unique minimal-norm solution to (2.4.1) and v solves (2.4.1), it follows that $v = y$.

Since the limit $\lim_{n\to\infty} u(t_n) = v = y$ is the same for every subsequence $t_n \to \infty$, for which the weak limit of $u(t_n)$ exists, one concludes that the strong limit $\lim_{t\to\infty} u(t) = y$. Indeed, assuming that for some sequence $t_n \to \infty$ the limit of $u(t_n)$ does not exist, one selects a subsequence, denoted again t_n, for which the weak limit of $u(t_n)$ does exist, and proves as before that this limit is y, thus getting a contradiction. Claim 6 is proved. □

For convenience of the reader let us prove the global existence and uniqueness of the solution to (2.4.4) with $\Phi = -F$, where F is a monotone, hemicontinuous operator in H (cf [Dei]). Uniqueness of the solution is trivial: if there are two solutions, u and v, then their difference $w := u - v$ solves the problem $\dot{w} = -[F(u) - F(v)]$, $w(0) = 0$.

Multiply this by w and use the monotonicity of F to get $\dot{g} \le 0, g(0) = 0$, where $g := \|w(t)\|$. Thus, $g = 0$, so $w = 0$, and uniqueness is proved.

Let us prove the global existence of the solution to (2.4.4) with $\Phi = -F$. Consider the equation:

$$w_n(t) = w_0 - \int_0^t F\left(w_n\left(s - \frac{1}{n}\right)\right) ds, \quad t > 0, \qquad w_n(t) = w_0, \quad t \le 0. \tag{**}$$

We wish to prove that

$$\lim_{n\to\infty} w_n(t) = w(t), \qquad \forall t > 0,$$

where $\dot{w}(t) = -F(w)$. Our assumptions (monotonicity and hemicontinuity of F) imply demicontinuity of F. Fix an arbitrary $T > 0$, and let $B(w_0, r)$ be the ball centered at w_0 with radius $r > 0$. Let $\sup_{u\in B(w_0,r)} \|F(u)\| := c$. Then $\|w_n(t) - w_0\| \le ct$. If $t \le r/c$, then $w_n(t) \in B(w_0, r)$, and $\|\dot{w}_n(t)\| \le c$. Define

$$z_{nm}(t) := w_n(t) - w_m(t), \qquad \|z_{nm}(t)\| := g_{nm}(t).$$

From (**) one gets:

$$g_{nm}\dot{g}_{nm} = -\left(F\left(w_n\left(t - \frac{1}{n}\right)\right) - F\left(w_m\left(t - \frac{1}{m}\right)\right), w_n(t) - w_m(t)\right) := I.$$

One has:

$$\begin{aligned} I = &-\left(F\left(w_n\left(t - \frac{1}{n}\right)\right) - F\left(w_m\left(t - \frac{1}{m}\right)\right), w_n\left(t - \frac{1}{n}\right) - w_m\left(t - \frac{1}{m}\right)\right) \\ &-\left(F\left(w_n\left(t - \frac{1}{n}\right)\right) - F\left(w_m\left(t - \frac{1}{m}\right)\right), w_n(t) - w_n\left(t - \frac{1}{n}\right)\right. \\ &\left.-\left(w_m(t) - w_m\left(t - \frac{1}{m}\right)\right)\right). \end{aligned}$$

Using the monotonicity of F, the estimate $\sup_{w\in B(w_0,r)} \|F(w)\| \le c$, and the estimate $\|\dot{w}_n(t)\| \le c$, one gets:

$$I \le 2c^2\left(\frac{1}{n} + \frac{1}{m}\right).$$

Therefore

$$g_{nm}\dot{g}_{nm} \le 2c^2\left(\frac{1}{n} + \frac{1}{m}\right) \to 0 \quad as \quad n, m \to \infty.$$

This implies

$$\lim_{n,m\to\infty} g_{nm}(t) = 0, \quad 0 \le t \le \frac{r}{c}.$$

Therefore there exists the strong limit $w(t)$:

$$\lim_{n\to\infty} w_n(t) = w(t), \quad 0 \le t \le \frac{r}{c}.$$

This function w satisfies the integral equation:

$$w(t) = w_0 - \int_0^t F(w(s))ds,$$

and solves the Cauchy problem

$$\dot{w} = -F(w), \quad w(0) = w_0. \tag{$***$}$$

If F is continuous, then this Cauchy problem and the preceding integral equation are equivalent. If F is demicontinuous, then they are also equivalent, but the derivative in the Cauchy problem should be understood in the weak sense. We have proved the existence of the unique local solution to(***). To prove that the solution to (***) exists for any $t \in [0, \infty)$, let us assume that the solution exists on $[0, T)$, but not on a larger interval $[0, T + d)$, $d > 0$, and show that this leads to a contradiction. It is sufficient to prove that the finite limit: $\lim_{t\to T} w(t) := W$ does exist, because then one can solve locally, on the interval $[T, T + d)$, equation (***) with the initial data $w(T) = \lim_{t\to T} w(t)$, and construct the solution to (***) on the interval $[0, T + d)$, thus getting a contradiction.

To prove that W exists, consider

$$w(t + h) - w(t) := z(t), \quad \|z\| := g.$$

One has $\dot{z} = -[F(w(t + h)) - F(w(t))]$. Using the monotonicity of F, one gets $(z, \dot{z}) \le 0$. Thus,

$$\|w(t + h) - w(t)\| \le \|w(h) - w(0)\|.$$

The right-hand side of the above inequality tends to zero as $h \to 0$. This, and the Cauchy test imply the existence of W. The proof is complete. □

2.4.5 Nonlinear ill-posed problems with non-monotone operators

Assume that $F(u) := B(u) - f$, B is a non-monotone operator, $A := F'(u)$, $\tilde{A} := F'(y)$, $T := A^* A$, $\tilde{T} := \tilde{A}^* \tilde{A}$, $T_\varepsilon := T + \varepsilon I$, where I is the identity operator, ε is as in Theorem 2.4.10 and $\frac{|\dot{\varepsilon}(t)|}{\varepsilon(t)} < 1$,

$$\Phi := -T_\varepsilon^{-1}(u)[A^*(B(u) - f) + \varepsilon(u - \tilde{u}_0)], \quad \varepsilon = \varepsilon(t) > 0, \tag{2.4.53}$$

and Φ_δ is defined similarly, with f_δ replacing f and u_δ replacing u.

The main result of this Section is:

Theorem 2.4.16. *If* (2.4.2) *holds,* $u, u_0 \in B(y, R)$, $y - \tilde{u}_0 = \tilde{T}z$, $\|z\| \ll 1$, *(that is,* $\|z\|$ *is sufficiently small), and* $R \ll 1$, *then problem* (2.4.4) *has a unique global solution and* (2.4.5) *holds. If* $u_\delta(t)$ *solves* (2.4.4) *(with* Φ_δ *in place of* Φ*), then there exists a* t_δ *such that* $\lim_{\delta\to 0} \|u_\delta(t_\delta) - y\| = 0$.

The derivation of the stopping rule, that is, the choice of t_δ, is based on the ideas presented in Section 2.4.4.

Sketch of proof of Theorem 2.4.16. Proof of Theorem 2.4.16 consists of the following steps.

First, we prove that $g := \|w\| := \|u(t) - y\|$ satisfies a differential inequality (2.4.35), and, applying (4.4), conclude that $g(t) < \mu^{-1}(t) \to 0$ as $t \to \infty$. A new point in this derivation (compared with that for monotone operators) is the usage of the source assumption $y - u_0 = \tilde{T}z$.

Secondly, we derive the stopping rule. The source assumption allows one to get a rate of convergence [KNR]. Details of the proof are technical and are not included. One can see [KNR] for some proofs.

Let us sketch the derivation of the differential inequality for g. Write $B(u) - f = B(u) - B(y) = Aw + K$, where $\|K\| \le \frac{M_2}{2} g^2$, and $\varepsilon(u - \tilde{u}_0) = \varepsilon w + \varepsilon(y - \tilde{u}_0) = \varepsilon w + \varepsilon\tilde{T}z$. Then (2.4.53) can be written as

$$\Phi = -w - T_\varepsilon^{-1}A^*K - \varepsilon T_\varepsilon^{-1}\tilde{T}z, \quad \varepsilon := \varepsilon(t). \tag{2.4.54}$$

Multiplying (2.4.4), with Φ defined in (2.4.54), by w, one gets:

$$g\dot{g} \le -g^2 + \frac{M_2}{2}\left\| T_{\varepsilon(t)}^{-1}A^*\right\| g^3 + \varepsilon(t)\left\| T_{\varepsilon(t)}^{-1}\tilde{T}\right\| \|z\| g. \tag{2.4.55}$$

Since $g \ge 0$, one obtains:

$$\dot{g} \le -g + \frac{M_2}{4\sqrt{\varepsilon(t)}} g^2 + \varepsilon(t)\|T_\varepsilon^{-1}\tilde{T}\|\|z\|, \tag{2.4.56}$$

where the estimate $\|T_\varepsilon^{-1}A^*\| \le \frac{1}{2\sqrt{\varepsilon}}$ was used. Clearly,

$$\|T_\varepsilon^{-1}\tilde{T}\| \le \|(T_\varepsilon^{-1} - \tilde{T}_\varepsilon^{-1})\tilde{T}\| + \|\tilde{T}_\varepsilon^{-1}\tilde{T}\|, \quad \|\tilde{T}_\varepsilon^{-1}\tilde{T}\| \le 1, \quad \varepsilon\|T_\varepsilon^{-1}\| \le 1,$$

and

$$T_\varepsilon^{-1} - \tilde{T}_\varepsilon^{-1} = T_\varepsilon^{-1}(A^*A - \tilde{A}^*\tilde{A})\tilde{T}_\varepsilon^{-1}.$$

One has:

$$\|A^*A - \tilde{A}^*\tilde{A}\| \le 2M_2M_1g, \quad \|z\| \ll 1.$$

Let $2M_1M_2\|z\| \le \frac{1}{2}$. This is possible since $\|z\| \ll 1$. Using the above estimates, one transforms (2.4.56) into the following inequality:

$$\dot{g} \le -\frac{1}{2}g + \frac{M_2}{4\sqrt{\varepsilon(t)}}g^2 + \|z\|\varepsilon. \tag{2.4.57}$$

Now, apply Theorem 2.4.11 to (2.4.57), choosing

$$\mu = \frac{2M_2}{\sqrt{\varepsilon}}, \quad \frac{|\dot{\varepsilon}|}{\varepsilon} < \frac{1}{2}, \quad 16M_2\|z\|\sqrt{\varepsilon(0)} < 1, \quad and \quad \frac{2M_2\|u_0 - y\|}{\sqrt{\varepsilon(0)}} < 1.$$

Then conditions (2.4.34) are satisfied, and Theorem 2.4.11 yields the estimate:

$$g(t) < \frac{\sqrt{\varepsilon(t)}}{2M_2}.$$

This is the main part of the proof of Theorem 2.4.16. □

Remark. The assumption $y - \tilde{u}_0 = \tilde{T}z$, $\|z\| \ll 1$, is satisfied under very weak assumption on $\tilde{T}$, namely, $\tilde{T}B_r \cap (B_a \setminus \{0\}) \ne \emptyset$, i.e. the image of a ball $B_r := \{u : \|u\| \le r\}$ $\forall r \in (0, r_0)$, $r_0 = \text{const} > 0$, intersects a punctured ball $B_a \setminus \{0\}$, where $a = \text{const} > 0$. This implies the existence of an element $\tilde{u}_0$ such that $y - \tilde{u}_0 = \tilde{T}z$, where $z \in B_r$ and $y - \tilde{u}_0 \in B_a \setminus \{0\}$. The condition $\|z\| \ll 1$ is satisfied because $r \in (0, r_0)$ can be chosen arbitrarily small.

2.4.6 Nonlinear ill-posed problems: avoiding inverting of operators in the Newton-type continuous schemes

In the Newton-type methods for solving well-posed nonlinear problems, for example, in the continuous Newton method with $\Phi = -[F'(u)]^{-1}F(u)$, the difficult and expensive part of the solution is inverting the operator $F'(u)$. In this section we give a method to avoid inverting of this operator. This is especially important in the ill-posed problems, where one has to invert some regularized versions of F', and to face more difficulties than in the well-posed problems.

Consider problem (2.4.1) and assume (2.4.2), and (2.4.3). Thus, we discuss our method in the simplest well-posed case.

Replace (2.4.4) by the following Cauchy problem (dynamical system):

$$\dot{u} = -QF, \quad u(0) = u_0, \tag{2.4.58}$$

$$\dot{Q} = -TQ + A^*, \quad Q(0) = Q_0, \tag{2.4.59}$$

where $A := F'(u)$, $T := A^*A$, and $Q = Q(t)$ is a bounded operator in H.

First, let us state our new technical tool: an operator version of the Gronwall inequality.

Theorem 2.4.17. *Let*

$$\dot{Q} = -T(t)Q(t) + G(t), \quad Q(0) = Q_0, \tag{2.4.60}$$

where $T(t)$, $G(t)$, and $Q(t)$ are linear bounded operators on a real Hilbert space H. If there exists $\varepsilon(t) > 0$ such that

$$\big(T(t)h, h\big) \geq \varepsilon(t)\|h\|^2 \quad \forall h \in H, \tag{2.4.61}$$

then

$$\|Q(t)\| \leq e^{-\int_0^t \varepsilon(s)ds}\left[\|Q(0)\| + \int_0^t \|G(s)\| e^{\int_0^s \varepsilon(x)dx}\, ds\right]. \tag{2.4.62}$$

A simple proof of Theorem 2.4.17 is left to the reader. It can be found in [RSm4].

Let us turn now to the proof of Theorem 2.4.18, formulated at the end of this Section. This theorem is the main result of Section 2.4.6.

Applying (2.4.62) to (2.4.59), and using (2.4.2) and (2.4.3), which implies

$$\big(T(t)h, h\big) \geq c\|h\|^2 \quad \forall h \in H, \quad c = \text{const} > 0, \tag{2.4.63}$$

one gets:

$$\|Q(t)\| \leq e^{-ct}\left[\|Q(0)\| + \int_0^t M_1 e^{cs}\, ds\right] \leq \left[\|Q_0\| + M_1 c^{-1}\right] := c_1, \tag{2.4.64}$$

as long as $u(t) \in B(u_0, R)$.

Let $u(t) - y := w$, $\|w\| := g$, $\tilde{A} := F'(y)$. Since $F(y) = 0$, one has $F(u) = \tilde{A}w + K$, where $\|K\| \leq 0.5M_2 g^2 := c_0 g^2$, and M_2 is the constant from (2.4.2). Rewrite (2.4.58) as

$$\dot{w} = -Q[\tilde{A}w + K]. \tag{2.4.65}$$

Let $\Lambda := I - Q\tilde{A}$. Multiply (2.4.65) by w and get

$$g\dot{g} \leq -g^2 + (\Lambda w, w) + c_0 g^3, \quad c_0 = \text{const} > 0. \tag{2.4.66}$$

We prove below that

$$\sup_{t\geq 0} \|\Lambda\| \leq \lambda < 1. \tag{2.4.67}$$

From (2.4.66) and (2.4.67) one gets the following differential inequality:

$$\dot{g} \leq -\gamma g + c_0 g^2, \quad 0 < \gamma < 1, \ \gamma := 1 - \lambda, \tag{2.4.68}$$

which implies:

$$g(t) \leq r e^{-\gamma t}, \quad r := g(0)[1 - g(0)c_0]^{-1}, \tag{2.4.69}$$

provided that

$$g(0)c_0 < 1. \tag{2.4.70}$$

Inequality (2.4.70) holds if u_0 is sufficiently close to y.

From (2.4.69) it follows that $u(\infty) = y$. Thus, (2.1.6) holds.

The trajectory $u(t) \in B(u_0, R)$, $\forall t > 0$, provided that

$$\int_0^\infty \|\dot{u}\| dt = \int_0^\infty \|\dot{w}\| dt \leq r + \frac{c_0 r^2}{2\gamma} \leq R. \tag{2.4.71}$$

This inequality holds if u_0 is sufficiently close to y, that is, r is sufficiently small.

To complete the argument, let us prove (2.4.67). One has:

$$\dot{\Lambda} = -\dot{Q}\tilde{A} = -T\Lambda + A^*(A - \tilde{A}), \tag{2.4.72}$$

and one has $\|A - \tilde{A}\| \leq M_2 g$. Using (2.4.69) and Theorem 2.4.17, one gets

$$\|\Lambda\| \leq e^{-ct}\left[\|\Lambda_0\| + r M_1 M_2 \int_0^t e^{(c-\gamma)s} ds\right]. \tag{2.4.73}$$

Thus,

$$\|\Lambda\| \leq \|\Lambda_0\| + Cr := \lambda, \quad C := M_1 M_2 \sup_{t>0} \frac{e^{-\gamma t} - e^{-ct}}{c - \gamma}. \tag{2.4.74}$$

If u_0 is sufficiently close to y and Q_0 is sufficiently close to $\tilde{A}^{-1}$, then $\lambda > 0$ can be made arbitrarily small. We have proved:

Theorem 2.4.18. *If* (2.4.2), *and* (2.4.3) *hold,* Q_0 *and* u_0 *are sufficiently close to* $\tilde{A}^{-1}$ *and* y, *respectively, then problem* (2.4.58)–(2.4.59) *has a unique global solution,* (2.4.5) *holds, and* $u(t)$ *converges to* y, *which solves* (2.4.1), *exponentially fast.*

In [RSm4] a generalization of Theorem 2.4.18 is given for ill-posed problems.

Exercise. 1. Let $\epsilon(t) > 0$, $\dot{\epsilon} < 0$, $|\dot{\epsilon}|\epsilon^{-1} < \gamma$, where $\gamma > 0$ is a constant, (2.4.2) holds in $B(y, R\epsilon(0))$, where y is a solution to (2.4.1) and $R > 0$ is a constant. Assume that there exists a $z \in B(y, R\epsilon(0))$, such that $y - z = \tilde{T}w$, where w is some element, and $\tilde{T} := (A^*A)|_{u=y}$. Assume that:

$$\dot{u} = -Q(t)[A^*F(u(t)) + \epsilon(t)(u(t) - u_0)], \quad A := F'(u(t)),$$
$$\dot{Q}(t) = -T_{\epsilon(t)}Q(t) + I, \quad u(0) = u_0, \ \ Q(0) = Q_0, \quad T_\epsilon := T + \epsilon I.$$

Then, if R, y and w are sufficiently small, then the above Cauchy problem has a unique global solution $\begin{pmatrix} u(t) \\ Q(t) \end{pmatrix}$, and $\|u(t) - y\| \leq R\epsilon(t)$. Precise meaning of the above smallness condition is explained in [RSm4].

2. If F is monotone, then the assumption $y - z = \tilde{T}w$ can be dropped, and, under the assumption $|\dot{\epsilon}|\epsilon^{-3}$ is sufficiently small, one derives the existence and uniqueness of the global solution to the Cauchy problem of n.1 of this Exercise. The method of the proof is the same as in [RSm4].

2.4.7 Iterative schemes

In this section we present a method for constructing convergent iterative schemes for a wide class of well-posed equations (2.4.1). Some methods for constructing convergent iterative schemes for a wide class of Ill-posed problems are given in [AR1]. There is an enormous literature on iterative methods ([BG], [VV]).

Consider a discretization scheme for solving (2.4.4) with $\Phi = \Phi(u)$, so that we assume no explicit time dependence in Φ:

$$u_{n+1} = u_n + h\Phi(u_n), \quad u_0 = u_0, \quad h = \text{const} > 0. \tag{2.4.75}$$

One of the results from [AR1], concerning the well-posed equations (2.4.1) is Theorem 2.4.19, formulated below. Its proof is shorter than in [AR1].

Theorem 2.4.19. *Assume* (2.4.2), (2.4.3), (2.4.6)–(2.4.9) *with* $a = 2$, $g_1 = c_1 = \text{const} > 0$, $g_2 = c_2 = \text{const} > 0$, $\|\Phi'(u)\| \leq L_1$, *for* $u \in B(y, R)$. *Then, if* $h > 0$ *is sufficiently small, and* u_0 *is sufficiently close to* y, *then* (2.4.75) *produces a sequence* u_n *for which*

$$\|u_n - y\| \leq Re^{-chn}, \quad \|F(u_n)\| \leq \|F_0\|e^{-chn}, \tag{2.4.76}$$

where $R := \frac{c_2\|F_0\|}{c_1}$, $F_0 = F(u_0)$, $c = \mathit{const} > 0$, $c < c_1$.

Proof of Theorem 2.4.19. The proof is by induction. For $n = 0$ estimates (2.4.76) are clear. Assuming these estimates for $j \leq n$, let us prove them for $j = n + 1$. Let $F_n := F(u_n)$, and let $w_{n+1}(t)$ solve problem (2.4.4) on the interval (t_n, t_{n+1}), $t_n := nh$, with $w_{n+1}(t_n) = u_n$. By (2.4.10) (with $G = c_2e^{-c_1t}$) and (2.4.76) one gets:

$$\|w_{n+1}(t) - y\| \leq \frac{c_2}{c_1}\|F_n\|e^{-c_1t} \leq Re^{-cnh-c_1t}, \quad t_n \leq t \leq t_{n+1}. \tag{2.4.77}$$

One has:

$$\|u_{n+1} - y\| \le \|u_{n+1} - w_{n+1}(t_{n+1})\| + \|w_{n+1}(t_{n+1}) - y\|, \tag{2.4.78}$$

and

$$\begin{aligned}\|u_{n+1} - w_{n+1}(t_{n+1})\| &\le \int_{t_n}^{t_{n+1}} \|\Phi(u_n) - \Phi(w_{n+1}(s))\| ds \\ &\le L_1 c_2 h \int_{t_n}^{t_{n+1}} \|F(w_{n+1}(t))\| dt \le L_1 c_1 h^2 Re^{-cnh},\end{aligned} \tag{2.4.79}$$

where we have used the formula $R := \frac{c_2\|F_0\|}{c_1}$, and the estimate:

$$\|F(w_{n+1}(t))\| \le \|F_n\| e^{-c_1(t-t_n)} \le \|F_0\| e^{-cnh-c_1(t-t_n)}. \tag{2.4.80}$$

From (2.4.77)–(2.4.80) it follows that:

$$\|u_{n+1} - y\| \le Re^{-cnh}(e^{-c_1 h} + c_1 L_1 h^2) \le Re^{-c(n+1)h}, \tag{2.4.81}$$

provided that

$$e^{-c_1 h} + c_1 L_1 h^2 \le e^{-ch}. \tag{2.4.82}$$

Inequality (2.4.82) holds if h is sufficiently small and $c < c_1$. So, the first inequality (2.4.76), with $n+1$ in place of n, is proved if h is sufficiently small and $c < c_1$.

Now

$$\|F(u_{n+1})\| \le \|F(u_{n+1}) - F(w_{n+1}(t))\| + \|F(w_{n+1}(t))\|, \quad t_n \le t \le t_{n+1}. \tag{2.4.83}$$

Using (2.4.2) and (2.4.79), one gets:

$$\|F(u_{n+1}) - F(w_{n+1}(t_{n+1}))\| \le M_1 \|u_{n+1} - w_{n+1}(t_{n+1})\| \le M_1 c_2 L_1 h^2 \|F_0\| e^{-cnh}. \tag{2.4.84}$$

From (2.4.83) and (2.4.84) it follows that:

$$\|F(u_{n+1})\| \le \|F_0\| e^{-cnh}(e^{-c_1 h} + M_1 c_2 L_1 h^2) \le \|F_0\| e^{-c(n+1)h}, \tag{2.4.85}$$

provided that

$$e^{-c_1 h} + M_1 c_2 L_1 h^2 \le e^{-ch}. \tag{2.4.86}$$

Inequality (2.4.86) holds if h is sufficiently small and $c < c_1$. So, the second inequality (2.4.76) with $n+1$ in place of n is proved if h is sufficiently small and $c < c_1$. Theorem 2.4.19 is proved. □

In the well-posed case, if $F(y) = 0$, the discrete Newton's method

$$u_{n+1} = u_n - [F'(u_n)]^{-1} F(u_n), \; u_0 = u(0),$$

converges superexponentially if u_0 is sufficiently close to y. Indeed, if $v_n := u_n - y$, then $v_{n+1} = v_n - [F'(u_n)]^{-1}[F'(u_n)v_n + K]$ where $\|K\| \leq \frac{M_2}{2}\|v_n\|^2$. Thus, $g_n := \|v_n\|$ satisfies the inequality: $g_{n+1} \leq q g_n^2$, where $q := \frac{m_1 M_2}{2}$. Therefore $g_n \leq q^{2^n - 1} g_0^{2^n}$, and if $0 < q g_0 < 1$, then the method converges superexponentially.

If one uses the iterative method $u_{n+1} = u_n - h[F'(u_n)]^{-1} F(u_n)$, with $h \neq 1$, then, in the well-posed case, assuming that this method converges, it converges exponentially, that is, slower than in the case $h = 1$.

The continuous analog of the above method

$$\dot{u} = -a[F'(u)]^{-1} F(u), \quad u(0) = u_0,$$

where $a = const > 0$, converges at the rate $O(e^{-at})$. Indeed, if $g(t) := \|F(u(t))\|$, then $g\dot{g} = -ag^2$, so $g(t) = g_0 e^{-at}$, $\|\dot{u}\| \leq a m_1 g_0 e^{-at}$. Thus

$$\|u(t) - u(\infty)\| \leq m_1 g_0 e^{-at}, \quad \textit{and} \quad F(u(\infty)) = 0.$$

In the continuous case one does not have superexponential convergence no matter what $a > 0$ is (see [R210]).

2.4.8 A spectral assumption

In this section we introduce the spectral assumption which allows one to treat some nonlinear non-monotone operators.

Assumption S: *The set* $\{r, \varphi : \pi - \varphi_0 < \varphi < \pi + \varphi_0, \; \varphi_0 > 0, 0 < r < r_0\}$, *where* φ_0 *and* r_0 *are arbitrarily small, fixed numbers, consists of the regular points of the operator* $A := F'(u)$ *for all* $u \in B(u_0, R)$.

Assumption S implies the estimate:

$$\|(F'(u) + \varepsilon)^{-1}\| \leq \frac{1}{\varepsilon \sin \varphi_0}, \quad 0 < \varepsilon < r_0(1 - \sin \varphi_0), \tag{2.4.87}$$

because $\|(A - z)^{-1}\| \leq \frac{1}{dist(z, s(A))}$, where $s(A)$ is the spectrum of a linear operator A, and $dist(z, s(A))$ is the distance from a point z of a complex plane to the spectrum. In our case, $z = -\varepsilon$, and $dist(z, s(A)) = \varepsilon \sin \varphi_0$, if $\varepsilon < r_0(1 - \sin \varphi_0)$.

Theorem 2.4.20. *If* (2.4.2) *and* (2.4.87) *hold, and* $0 < \varepsilon < r_0(1 - \sin \varphi_0)$, *then problem* (2.4.38), *with* $\varepsilon(t) = \varepsilon = const > 0$, *is solvable, problem* (2.4.4), *with* Φ *defined in* (2.4.33) *and* $\tilde{u}_0 = 0$, *has a unique global solution,* $\exists u(\infty)$, *and* $F(u(\infty)) + \varepsilon u(\infty) = 0$. *Every solution to the equation* $F(V) + \varepsilon V = 0$ *is isolated.*

Proof of Theorem 2.4.20. Let $g = g(t) := \|F(u(t)) + \varepsilon u(t)\|$, where $u = u(t)$ solves locally (2.4.4), where Φ is defined in (2.4.33) and $\tilde{u}_0 = 0$. Then:

$$g\dot{g} = -((F'(u) + \varepsilon)(F'(u) + \varepsilon)^{-1}(F(u) + \varepsilon u), F(u) + \varepsilon u) = -g^2, \tag{2.4.88}$$

so

$$g = g_0 e^{-t}, \ g_0 := g(0); \quad \|\dot{u}\| \le \frac{g_0}{\varepsilon \sin\varphi_0} e^{-t}. \tag{2.4.89}$$

Thus,

$$\|u(t) - u(\infty)\| \le \frac{g_0}{\varepsilon \sin\varphi_0} e^{-t}, \quad \|u(t) - u_0\| \le \frac{g_0}{\varepsilon \sin\varphi_0}, \quad F(u(\infty)) + \varepsilon u(\infty) = 0. \tag{2.4.90}$$

Therefore equation

$$F(V) + \varepsilon V = 0, \quad \varepsilon = \text{const} > 0, \tag{2.4.91}$$

has a solution in $B(u_0, R)$, where $R = \frac{g_0}{\varepsilon \sin\varphi_0}$.

Every solution to equation (2.4.91) is isolated. Indeed, if $F(W) + \varepsilon W = 0$, and $\psi := V - W$, then $F(V) - F(W) + \varepsilon\psi = 0$, so $[F'(V) + \varepsilon]\psi + K = 0$, where $\|K\| \le \frac{M_2}{2}\|\psi\|^2$. Thus, using (2.4.87), one gets $\|\psi\| \ge \frac{2\varepsilon \sin\varphi_0}{M_2}$ unless $\|\psi\| = 0$. Consequently, if $\|\psi\|$ is sufficiently small, then $\psi = 0$. Theorem 2.4.20 is proved. $\square$

2.4.9 Nonlinear integral inequality

The main result of this section is Theorem 2.4.22, which is used extensively in Sections 2.4.4–2.4.8.

The following lemma is a version of some known results concerning integral inequalities.

Lemma 2.4.21. *Let $f(t, w)$, $g(t, u)$ be continuous on region $[0, T) \times D$ $(D \subset R, T \le \infty)$ and $f(t, w) \le g(t, u)$ if $w \le u$, $t \in (0, T)$, $w, u \in D$. Assume that $g(t, u)$ is such that the Cauchy problem*

$$\dot{u} = g(t, u), \quad u(0) = u_0, \quad u_0 \in D \tag{2.4.92}$$

has a unique solution. If

$$\dot{w} \le f(t, w), \quad w(0) = w_0 \le u_0, \quad w_0 \in D, \tag{2.4.93}$$

then $u(t) \ge w(t)$ for all t for which $u(t)$ and $w(t)$ are defined.

Proof of Lemma Lemma 2.4.21.

Step 1. Suppose first $f(t, w) < g(t, u)$, if $w \le u$. Since $w_0 \le u_0$ and $\dot{w}(0) \le f(t, w_0) < g(t, u_0) = \dot{u}(0)$, there exists $\delta > 0$ such that $u(t) > w(t)$ on $(0, \delta]$. Assume that for some $t_1 > \delta$ one has $u(t_1) < w(t_1)$. Then for some $t_2 < t_1$ one has

$$u(t_2) = w(t_2) \quad \text{and} \quad u(t) < w(t) \quad \text{for} \quad t \in (t_2, t_1].$$

One gets

$$\dot{w}(t_2) \geq \dot{u}(t_2) = g(t, u(t_2)) > f(t, w(t_2)) \geq \dot{w}(t_2).$$

This contradiction proves that there is no point t_2 such that $u(t_2) = w(t_2)$.

Step 2. Now consider the case $f(t, w) \leq g(t, u)$, if $w \leq u$. Define

$$\dot{u}_n = g(t, u_n) + \varepsilon_n, \quad u_n(0) = u_0, \quad \varepsilon_n > 0, \quad n = 0, 1, \dots,$$

where ε_n tends monotonically to zero. Then

$$\dot{w} \leq f(t, w) \leq g(t, u) < g(t, u) + \varepsilon_n, \quad w \leq u.$$

By Step 1 $u_n(t) \geq w(t)$, $n = 0, 1, \dots$. Fix an arbitrary compact set $[0, T_1]$, $0 < T_1 < T$.

$$u_n(t) = u_0 + \int_0^t g(\tau, u_n(\tau))d\tau + \varepsilon_n t. \tag{2.4.94}$$

Since $g(t, u)$ is continuous, the sequence $\{u_n\}$ is uniformly bounded and equicontinuous on $[0, T_1]$. Therefore there exists a subsequence $\{u_{n_k}\}$ which converges uniformly to a continuous function $u(t)$. By continuity of $g(t, u)$ we can pass to the limit in (2.4.94) and get

$$u(t) = u_0 + \int_0^t g(\tau, u(\tau))d\tau, \quad t \in [0, T_1]. \tag{2.4.95}$$

Since T_1 is arbitrary (2.4.95) is equivalent to the initial Cauchy problem that has a unique solution. The inequality $u_{n_k}(t) \geq w(t)$, $k = 0, 1, \dots$ implies $u(t) \geq w(t)$. If the solution to the Cauchy problem (2.4.92) is not unique, the inequality $w(t) \leq u(t)$ holds for the maximal solution to (2.4.92). □

The following theorem was stated earlier as Theorem 2.4.11. For convenience of the reader, we repeat its formulation.

Theorem 2.4.22. *Let $\gamma(t), \sigma(t), \beta(t) \in C[t_0, \infty)$ for some real number t_0. If there exists a positive function $\mu(t) \in C^1[t_0, \infty)$ such that*

$$0 \leq \sigma(t) \leq \frac{\mu(t)}{2}\left(\gamma(t) - \frac{\dot{\mu}(t)}{\mu(t)}\right), \quad \beta(t) \leq \frac{1}{2\mu(t)}\left(\gamma(t) - \frac{\dot{\mu}(t)}{\mu(t)}\right), \tag{2.4.96}$$

then a non-negative solution to the following inequalities:

$$\dot{v}(t) \leq -\gamma(t)v(t) + \sigma(t)v^2(t) + \beta(t), \quad v(t_0) < \frac{1}{\mu(t_0)}, \tag{2.4.97}$$

satisfies the estimate:

$$v(t) \leq \frac{1 - \nu(t)}{\mu(t)} < \frac{1}{\mu(t)}, \tag{2.4.98}$$

for all $t \in [t_0, \infty)$*, where*

$$\nu(t) = \left(\frac{1}{1 - \mu(t_0)v(t_0)} + \frac{1}{2}\int_{t_0}^{t} \left(\gamma(s) - \frac{\dot{\mu}(s)}{\mu(s)} \right) ds \right)^{-1}. \tag{2.4.99}$$

Remark 2.4.23. *Without loss of generality one can assume* $\beta(t) \geq 0$.

In [Alb1] a differential inequality $\dot{v} \leq -A(t)\psi(v(t)) + \beta(t)$ was studied under some assumptions which include, among others, the positivity of $\psi(v)$ for $v > 0$. In Theorem 2.4.22 the term $-\gamma(t)v(t) + \sigma(t)v^2(t)$ (which is analogous to some extent to the term $-A(t)\psi(v(t)))$ can change sign. Our Theorem 2.4.22 is not covered by the result in [Alb1]. In particular, in Theorem 2.4.22 an analog of $\psi(v)$, for the case $\gamma(t) = \sigma(t) = A(t)$, is the function $\psi(v) := v - v^2$. This function goes to $-\infty$ as v goes to $+\infty$, so it does not satisfy the positivity condition imposed in [Alb1]. Lemma 1 in [Alb1] is wrong. Its corrected version is given in [ARS3], where a counterexample to Lemma 1 from [Alb1] is constructed. In [ARS3] the following result is proved:

Lemma. *Let* $u \in C^1[0, \infty), u \geq 0$*, and*
$\dot{u} \leq -a(t)f(u(t)) + b(t)$*,* $u(0) = u_0$.
Assume:
1) $a(t), b(t) \in C[0, +\infty)$*,* $a(t) > 0$*,* $b(t) \geq 0$ *for* $t > 0$*,*
2) $\int^{+\infty} a(t)dt = +\infty$*,* $\frac{b(t)}{a(t)} \to 0$ *as* $t \to +\infty$*,*
3) $f \in C[0, +\infty)$*,* $f(0) = 0$*,* $f(u) > 0$ *for* $u > 0$*,*
4) there exists $c > 0$ *such that* $f(u) \geq c$ *for* $u \geq 1$.
Then $u(t) \to 0 \quad as \quad t \to +\infty$.

Unlike in the case of Bihari integral inequality ([BB]) one cannot separate variables in the right hand side of the first inequality (2.4.97) and estimate $v(t)$ by a solution of the Cauchy problem for a differential equation with separating variables. The proof below is based on a special choice of the solution to the Riccati equation majorizing a solution of inequality (2.4.97).

Proof of Theorem 2.4.22. Denote:

$$w(t) := v(t)e^{\int_{t_0}^{t} \gamma(s)ds}, \tag{2.4.100}$$

then (2.4.97) implies:

$$\dot{w}(t) \leq a(t)w^2(t) + b(t), \quad w(t_0) = v(t_0), \tag{2.4.101}$$

where

$$a(t) = \sigma(t)e^{-\int_{t_0}^t \gamma(s)ds}, \quad b(t) = \beta(t)e^{\int_{t_0}^t \gamma(s)ds}.$$

Consider Riccati's equation:

$$\dot{u}(t) = \frac{\dot{f}(t)}{g(t)}u^2(t) - \frac{\dot{g}(t)}{f(t)}. \tag{2.4.102}$$

One can check by a direct calculation that the the solution to problem (2.4.102) is given by the following formula [Kam, eq. 1.33]:

$$u(t) = -\frac{g(t)}{f(t)} + \left[f^2(t)\left(C - \int_{t_0}^t \frac{\dot{f}(s)}{g(s)f^2(s)}ds\right)\right]^{-1}. \tag{2.4.103}$$

Define f and g as follows:

$$f(t) := \mu^{\frac{1}{2}}(t)e^{-\frac{1}{2}\int_{t_0}^t \gamma(s)ds}, \quad g(t) := -\mu^{-\frac{1}{2}}(t)e^{\frac{1}{2}\int_{t_0}^t \gamma(s)ds}, \tag{2.4.104}$$

and consider the Cauchy problem for equation (2.4.102) with the initial condition $u(t_0) = v(t_0)$. Then C in (2.4.103) takes the form:

$$C = \frac{1}{\mu(t_0)v(t_0) - 1}.$$

From (2.4.96)) one gets

$$a(t) \leq \frac{\dot{f}(t)}{g(t)}, \quad b(t) \leq -\frac{\dot{g}(t)}{f(t)}.$$

Since $fg = -1$ one has:

$$\int_{t_0}^t \frac{\dot{f}(s)}{g(s)f^2(s)}ds = -\int_t^t \frac{\dot{f}(s)}{f(s)}ds = \frac{1}{2}\int_{t_0}^t \left(\gamma(s) - \frac{\dot{\mu}(s)}{\mu(s)}\right)ds.$$

Thus

$$u(t) = \frac{e^{\int_{t_0}^t \gamma(s)ds}}{\mu(t)}\left[1 - \left(\frac{1}{1-\mu(t_0)v(t_0)} + \frac{1}{2}\int_{t_0}^t \left(\gamma(s) - \frac{\dot{\mu}(s)}{\mu(s)}\right)ds\right)^{-1}\right]. \tag{2.4.105}$$

It follows from conditions (2.4.96) and from the second inequality in (2.4.97) that the solution to problem (2.4.102) exists for all $t \in [0, \infty)$ and the following inequality holds with $\nu(t)$ defined by (2.4.99):

$$1 > 1 - \nu(t) \geq \mu(t_0)\nu(t_0). \tag{2.4.106}$$

From Lemma 2.4.21 and from formula (2.4.105) one gets:

$$\nu(t)e^{\int_{t_0}^t \gamma(s)ds} := w(t) \leq u(t) = \frac{1-\nu(t)}{\mu(t)} e^{\int_{t_0}^t \gamma(s)ds} < \frac{1}{\mu(t)} e^{\int_{t_0}^t \gamma(s)ds}, \tag{2.4.107}$$

and thus estimate (2.4.98) is proved. □

To illustrate conditions of Theorem 2.4.22 consider the following examples of functions γ, σ, and β, satisfying (2.4.96) for $t_0 = 0$.

Example 2.4.24. *Let*

$$\gamma(t) = c_1(1+t)^{\nu_1}, \quad \sigma(t) = c_2(1+t)^{\nu_2}, \quad \beta(t) = c_3(1+t)^{\nu_3}, \tag{2.4.108}$$

where $c_2 > 0$, $c_3 > 0$. *Choose* $\mu(t) := c(1+t)^{\nu}$, $c > 0$. *From* (2.4.96), (2.4.97) *one gets the following conditions*

$$\begin{aligned} c_2 &\leq \frac{cc_1}{2}(1+t)^{\nu+\nu_1-\nu_2} - \frac{c\nu}{2}(1+t)^{\nu-1-\nu_2}, \\ c_3 &\leq \frac{c_1}{2c}(1+t)^{\nu_1-\nu-\nu_3} - \frac{\nu}{2c}(1+t)^{-\nu-1-\nu_3}, \quad c\nu(0) < 1. \end{aligned} \tag{2.4.109}$$

Thus, one obtains the following conditions:

$$\nu_1 \geq -1, \quad \nu_2 - \nu_1 \leq \nu \leq \nu_1 - \nu_3, \tag{2.4.110}$$

and

$$c_1 > \nu, \quad \frac{2c_2}{c_1 - \nu} \leq c \leq \frac{c_1 - \nu}{2c_3}, \quad c\nu(0) < 1. \tag{2.4.111}$$

Therefore for such γ, σ *and* β *a function* μ *with the desired properties exists if*

$$\nu_1 \geq -1, \quad \nu_2 + \nu_3 \leq 2\nu_1, \tag{2.4.112}$$

and

$$c_1 > \nu_2 - \nu_1, \quad 2\sqrt{c_2 c_3} \leq c_1 + \nu_1 - \nu_2, \quad 2c_2\nu(0) < c_1 + \nu_1 - \nu_2. \tag{2.4.113}$$

In this case one can choose $\nu = \nu_2 - \nu_1$, $c = \frac{2c_2}{c_1+\nu_1-\nu_2}$. *However in order to have* $v(t) \to 0$ *as* $t \to +\infty$ *one needs the following conditions:*

$$\nu_1 \geq -1, \quad \nu_2 + \nu_3 \leq 2\nu_1, \quad \nu_1 > \nu_3, \tag{2.4.114}$$

and

$$c_1 > \nu_2 - \nu_1, \quad 2\sqrt{c_2 c_3} \leq c_1, \quad 2c_2\, v(0) < c_1. \tag{2.4.115}$$

Example 2.4.25. *If*

$$\gamma(t) = \gamma_0, \quad \sigma(t) = \sigma_0\, e^{\nu t}, \quad \beta(t) = \beta_0 e^{-\nu t}, \quad \mu(t) = \mu_0\, e^{\nu t},$$

then conditions (2.4.96), (2.4.97) *are satisfied if*

$$0 \leq \sigma_0 \leq \frac{\mu_0}{2}\big(\gamma_0 - \nu\big), \quad \beta_0 \leq \frac{1}{2\mu_0}\big(\gamma_0 - \nu\big), \quad \mu_0\, v(0) < 1.$$

Example 2.4.26. *Here* log *stands for the natural logarithm. For some* $t_1 > 0$

$$\gamma(t) = \frac{1}{\sqrt{\log(t + t_1)}}, \quad \mu(t) = c \log\big(t + t_1\big),$$

conditions (2.4.96), (2.4.97) *are satisfied if*

$$0 \leq \sigma(t) \leq \frac{c}{2}\left(\sqrt{\log(t+t_1)} - \frac{1}{t+t_1}\right),$$

$$\beta(t) \leq \frac{1}{2c\log^2(t+t_1)}\left(\sqrt{\log(t+t_1)} - \frac{1}{t+t_1}\right), \quad v(0)c\log t_1 < 1.$$

In all considered examples $\mu(t)$ tends to infinity as $t \to +\infty$ and provide a decay of a nonnegative solution to integral inequality (2.4.97) even if $\sigma(t)$ tends to infinity. Moreover, in the first and the third examples $v(t)$ tends to zero as $t \to +\infty$ when $\gamma(t) \to 0$ and $\sigma(t) \to +\infty$.

2.4.10 Riccati equation

An alternative approach to a study of Riccati equation (2.4.101) with non-negative coefficients $a(t)$ and $b(t)$ is based on the iterative method for solving integral equation $w(t) = w_0 + \int_0^t b(s)ds + \int_0^t a(s)w^2(s)ds$, $w_0 = \text{const} \geq 0$. Let $B(t) := w_0 + \int_0^t b\,ds$, $A(t) := \int_0^t a(s)ds$. Then $w(t) = B(t) + \int_0^t w^2(s)\,dA(s)$. Assume $a(t) \geq 0$, $B \geq 0$, and consider the process $w_{n+1}(t) = B(t) + \int_0^t w_n^2\,dA$, $w_0(t) = B(t)$, $w_1(t) \geq w_0(t)$. By induction, $w_{n+1}(t) \geq w_n(t)\ \forall n$. If $w_n(t) \leq c(T)\ \forall t \in [0, T]$, where $T > 0$ is an arbitrary number, then $\lim_{n\to\infty} w_n(t) := w(t)$ exists and solves (2.4.101).

Assume

$$\int_0^t dA\left(\int_0^t B^2(s)\,dA\right)^2 \le c_0\int_0^t B^2(s)\,dA. \tag{2.4.116}$$

Then

$$w_n(t) \le B(t) + c_1\int_0^t B^2(s)\,dA, \tag{2.4.117}$$

where c_1 is a constant estimated below. Inequality (2.4.117) holds for $n = 0$, and, by induction,

$$\begin{aligned} w_{n+1}(t) \le B(t) + \int_0^t dA\left(B + c_1\int_0^t B^2\,dA\right)^2 &\le B(t) \\ &+ 2\left[\int_0^t B^2\,dA + \int_0^t c_1^2\left(\int_0^t B^2\,dA\right)^2 dA\right]. \end{aligned}$$

Using (2.4.116) one gets:

$$w_{n+1}(t) \le B + 2\int_0^t B^2\,dA + 2c_1^2c_0\int_0^t B^2\,dA = B + 2\left(1 + c_1^2c_0\right)\int_0^t B^2\,dA.$$

If

$$2\left(1 + c_1^2c_0\right) \le c_1, \tag{2.4.118}$$

then (2.4.117) has been proved by induction. Inequality (2.4.118) holds if $\frac{1}{4c_0} - \sqrt{\frac{1}{b\,c_0^2} - \frac{1}{c_0}} < c_1 < \frac{1}{4c_0} + \sqrt{\frac{1}{16c_0^2} - \frac{1}{c_0}}$, and $0 < c_0 < \frac{1}{16}$.

2.5 EXAMPLES OF SOLUTIONS OF ILL-POSED PROBLEMS

2.5.1 Stable numerical differentiation: when is it possible?

In many applications one has to estimate a derivative f', given the noisy values of the function f to be differentiated. As an example we refer to the analysis of photo-electric response data. The goal of that experiment is to determine the relationship between the intensity of light falling on certain plant cells and their rate of uptake of various substances. Rather than measuring the uptake rate directly, the experimentalists measure the amount of each substance not absorbed as a function of time, the uptake rate being defined as minus the derivative of this function. As for the other example, one can mention the problem of finding the heat capacity c_p of a gas as a function of temperature T. Experimentally one measures the heat content $q(T) = \int_{T_0}^T c_p(\tau)d\tau$, and the heat capacity is determined by numerical differentiation.

One can give many other examples of practical problems in which one has to differentiate noisy data. In navigation problems one selects the direction of the motion of a ship by the maximum of a certain univalent curve, called the navigation characteristic. This direction can be obtained by differentiation of this curve. Since the navigation characteristic is communicated with some errors, one has to differentiate it numerically in order to find its maximum. In [R83, p. 94], the shape of a convex obstacle is found by differentiation of a support function of this obstacle. The support function is found from the experimentally measured scattering data, and by this reason the support function is noisy. In [R121, pp. 81–92], optimal estimates for the derivatives of random functions are obtained. In [RKa, p. 438], numerical differentiation of functions, contaminated by random noise is discussed. The noise has zero mean value and finite variance, and is identically distributed independently of the point x. It is proved that in this case the error of an optimal formula of numerical differentiation can be made $O(p^{-0.25}\epsilon)$, where p is the number of observation points and ϵ is the error for a noise which is non-random (see the precise formulation of the result in [RKa]). Section 2.5.1 is essentially paper [RSm5].

The differentiation of noisy data is an ill-posed problem: small (in some norm) perturbations of a function may lead to large errors in its derivative. Indeed, if one takes $f_\delta = f + \delta \sin(\frac{t}{\delta^2})$, $f' \in L^\infty(0,1)$, then $\|f_\delta - f\|_\infty = \delta$ and $\|f'_\delta - f'\|_\infty = \frac{1}{\delta}$, so that small in $L^\infty(0,1)$-norm perturbations of f result in large perturbations of f' in $L^\infty(0,1)$-norm.

Various methods have been developed for stable numerical differentiation of f given f_δ, $\|f_\delta - f\| \le \delta$. We mention three groups of methods:

(1) regularized difference methods with a step size $h = h(\delta)$ being a regularization parameter, see [R18], where this idea was proposed for the first time, and [R156], [R58], [R168]. As an example of such a method one may consider:

$$R_{h(\delta)} f_\delta(x) := \begin{cases} \dfrac{1}{h}(f_\delta(x+h) - f_\delta(x)), & 0 < x < h, \\ \dfrac{1}{2h}(f_\delta(x+h) - f_\delta(x-h)), & h \le x \le 1-h, \\ \dfrac{1}{h}(f_\delta(x) - f_\delta(x-h)), & 1-h < x < 1,\ h > 0. \end{cases} \tag{2.5.1}$$

If $f_\delta \in L^\infty(0,1)$, and $f \in W^{2,p}(0,1)$, where $W^{n,p}(0,1)$ is the Sobolev space of functions whose n-th derivative belongs to $L^p(0,1)$, $\|f_\delta - f\|_p \le \delta$, then

$$\begin{aligned} \left\| R_{h(\delta)} f_\delta - f' \right\|_p &\le \left\| R_{h(\delta)}\big(f_\delta - f\big) \right\|_p + \left\| R_{h(\delta)} f - f' \right\|_p \\ &\le \frac{2\delta}{h} + \frac{N_{2,p} h}{2}, \end{aligned} \tag{2.5.2}$$

where $N_{2,p}$ is an estimation constant: $\|f''\|_p \le N_{2,p}$. The error in the interval $h \le x \le 1-h$ can be estimated slightly better (by a quantity $\frac{\delta}{h} + \frac{N_{2,p}h}{2}$). In this paper by $\|\cdot\|_p$ we denote $\|\cdot\|_{L^p(0,1)}$. The right-hand side of (2.5.2) attains the absolute minimum $2\sqrt{\delta N_{2,p}}$ at $h = h_{2,p}(\delta) := 2(\frac{\delta}{N_{2,p}})^{\frac{1}{2}}$, while if one uses the error estimate for the interval $h \le x \le 1-h$, then one gets the absolute minimum $\sqrt{2\delta N_{2,p}}$ at $h = (\frac{2\delta}{N_{2,p}})^{\frac{1}{2}}$. When the function $f \in W^{3,p}(0,1)$, one can modify (2.5.1) near the ends so that it has the order $O(h^2)$ of the error of approximation as $h \to 0$, and results in an algorithm of order $\delta^{2/3}$. For example one can take

$$R_{h(\delta)} f_\delta(x) := \begin{cases} \frac{1}{2h}(4f_\delta(x+h) - f_\delta(x+2h) - 3f_\delta(x)), & 0 < x < 2h, \\ \frac{1}{2h}(f_\delta(x+h) - f_\delta(x-h)), & 2h < x < 1-2h, \\ \frac{1}{2h}(3f_\delta(x) + f_\delta(x-2h) - 4f_\delta(x-h)), & 1-2h < x < 1. \end{cases} \tag{2.5.3}$$

The difference methods use only local values of the function f_δ, which is natural when one estimates a derivative, and these methods are very simple, which is an advantage.

(2) An alternative approach is first to smooth f_δ by a mollification with a certain kernel, for example with a Gaussian kernel, or to use a mollification by splines, and then to differentiate the resulting smooth approximation, see e.g. [VA]. If one applies mollification with the Gaussian kernel $w_h(x) := \frac{1}{h\sqrt{\pi}} \exp(-\frac{x^2}{h^2})$, $x \in \mathbb{R}$, $h > 0$, then $(M_{h(\delta)})' : L^2(0,1) \to L^2(0,1)$,

$$\big(M_{h(\delta)}\big)' f_\delta(x) := \big(w_h' \star f_\delta\big)(x) := \int_0^1 w_h'(x-s) f_\delta(s)\,ds, \tag{2.5.4}$$

where $\star$ stands for the convolution, $f_\delta \in L^2(0,1)$, and $\|f_\delta - f\|_2 \le \delta$. Assume that $f \in H^1(0,1)$ with $f(0) = f(1) = 0$ and $\|f''\|_2 \le N_{2,2}$. One has

$$\big\| \big(M_{h(\delta)}\big)' f_\delta - f' \big\|_2 \le \big\| \big(M_{h(\delta)}\big)' \big(f_\delta - f\big) \big\|_2 + \big\| \big(M_{h(\delta)}\big)' f - f' \big\|_2 \tag{2.5.5}$$

From the Cauchy inequality the first term in the right-hand side of (2.5.5) can be estimated as follows:

$$\begin{aligned} \Big\| \big(M_{h(\delta)}\big)' \big(f_\delta - f\big) \Big\|_{L^2(0,1)} &= \Big\| w_h' \star \big(f_\delta - f\big) \Big\|_{L^2(0,1)} \le \Big\| w_h' \star \big(f_\delta - f\big) \Big\|_{L^2(\mathbb{R})} \\ &\le \big\| w_h' \big\|_{L^1(\mathbb{R})} \big\| f_\delta - f \big\|_{L^2(0,1)} \le \frac{2\delta}{h\sqrt{\pi}}, \end{aligned} \tag{2.5.6}$$

because $\|w_h'\|_{L^1(\mathbb{R})} = -2\int_0^\infty w_h'(s)\,ds = 2w_h(0) = \frac{2}{h\sqrt{\pi}}$. By a partial integration one gets:

$$\big(w_h' \star f\big)(x) = \int_0^1 w_h'(x-s) f(s)\,ds = \int_0^1 w_h(x-s) f'(s)\,ds = \big(w_h \star f'\big)(x). \tag{2.5.7}$$

To complete the argument one uses the inequality

$$\left\| w_h \star z - z \right\|_{L^2(\mathbb{R})} \leq h \|z'\|_{L^2(0,1)} \tag{2.5.8}$$

for every $z \in H^1(0,1)$ with $z(0) = z(1) = 0$. Here the above functions z are extended from $[0,1]$ to $\mathbb{R}$ by zero.

To verify (2.5.8) define the Fourier transform by

$$(\mathscr{F}z)(t) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z(s) e^{ist} ds, \quad t \in \mathbb{R}.$$

Using Parseval's equation, one gets:

$$\begin{aligned} \left\| w_h \star z - z \right\|_{L^2(\mathbb{R})} &= \left\| \mathscr{F}\big(w_h \star z - z\big) \right\|_{L^2(\mathbb{R})} = \left\| \big(\sqrt{2\pi}\,\mathscr{F}w_h - 1\big)\mathscr{F}z \right\|_{L^2(\mathbb{R})} \\ &= \left\| \frac{1}{-it}\Big(\sqrt{2\pi}\,\mathscr{F}w_h - 1\Big)\mathscr{F}z' \right\|_{L^2(\mathbb{R})} \leq \left\| \frac{1}{-it}\Big(\sqrt{2\pi}\,\mathscr{F}w_h - 1\Big) \right\|_{L^\infty(\mathbb{R})} \|z'\|_{L^2(0,1)}. \end{aligned} \tag{2.5.9}$$

Since

$$\varphi_h(t) := \frac{1}{it}\Big(1 - \sqrt{2\pi}\,\mathscr{F}w_h\Big) = \frac{1}{it}\big[1 - e^{-h^2t^2/4}\big], \quad t \in \mathbb{R},$$

and $\frac{1-e^{-\tau^2}}{\tau} \leq 2$, for all $\tau > 0$, estimate (2.5.9) yields inequality (2.5.8). Thus, from (2.5.7) and (2.5.9) one obtains

$$\left\| (M_{h(\delta)})' f - f' \right\|_2 \leq \left\| w_h \star f' - f' \right\|_2 \leq h N_{2,2}. \tag{2.5.10}$$

Finally, combining (2.5.5), (2.5.6) and (2.5.10) one gets

$$\left\| (M_{h(\delta)})' f_\delta - f' \right\|_2 \leq \frac{2\delta}{h\sqrt{\pi}} + h N_{2,2} := \varepsilon_2.$$

The choice $h = h_{2,2}(\delta) = \sqrt{\frac{2\delta}{N_{2,2}\sqrt{\pi}}}$ leads to the estimate $\varepsilon_2 \leq 2\sqrt{2/\sqrt{\pi}}\sqrt{\delta N_{2,2}}$.

(3) The third group of methods uses variational regularization for solving ill-posed problems ([Phi], [IVT]). One applies variational regularization to a Volterra integral equation

$$Au(x) := \int_0^x u(s) ds = f(x). \tag{2.5.11}$$

For example, if the noisy data f_δ are given, $\|f_\delta - f\|_2 \leq \delta$, then one minimizes the functional

$$F_0(u) := \left\| Au - f_\delta \right\|_2^2 + \alpha \|u\|_2^2$$

or

$$F_m(u) := \|Au - f_\delta\|_2^2 + \alpha \|u^{(m)}\|_2^2, \quad m > 0,$$

where $\alpha >$ is a regularization parameter. One proves that for a suitable choice of α, $\alpha = \alpha(\delta)$, the above functionals have a unique minimizer u_δ and $\|u_\delta - f'\|_2 \to 0$ as $\delta \to 0$. An optimal choice of the regularization parameter α in this approach is a nontrivial problem. Some methods for choosing $\alpha = \alpha(\delta)$ are presented in Section 2.1.

The above methods have been discussed in the literature, and their analysis is not our goal. Our goal is to study two principally different statements of the problem of stable numerical differentiation, and to understand when it is possible in principle to get a stable approximation to f' given noisy data f_δ. In Problem I a new notion of regularizer is introduced. Our treatment of the stable differentiation is an example of application of this new notion. In Section 2.1 a regularization method for unbounded linear and nonlinear operators is discussed.

Statements of the problem of stable numerical differentiation

First, we recall some standard definitions (cf. Section 1.3). The problem of finding a solution u to the equation

$$A(u) = f, \quad A : X \longrightarrow Y, \tag{2.5.12}$$

where X and Y are Banach spaces, A is an operator, possibly nonlinear, is *well-posed* (in the sense of J. Hadamard) if the following conditions hold:

(a) for every element $f \in Y$ there exists a solution $u \in X$;
(b) this solution is unique;
(c) the problem is stable under small perturbations of the initial data in the sense:

$$\|u_\delta - u\|_X \longrightarrow 0 \quad \text{if} \quad \|f_\delta - f\|_Y \longrightarrow 0, \quad \text{where} \quad A(u_\delta) := f_\delta. \tag{2.5.13}$$

If at least one of the conditions (a), (b) or (c) is violated, then the problem is called *ill-posed*. The problem of numerical differentiation can be written as

$$A(u) := \int_0^x u(s)ds = f, \quad A : X = L^p(0,1) \longrightarrow L^p(0,1), \quad f(0) = 0. \tag{2.5.14}$$

We study the cases $p = 2$ and $p = \infty$ in detail. Problem (2.5.14) is solvable only if $f' \in X$. So, condition (a) is not satisfied, condition (c) is not satisfied either, and condition (b) is satisfied. Therefore, problem (2.5.14) is ill-posed.

Practically, one does not know f and the only information available for computational processing is f_δ together with an *a priori* information about f, for example, one may know that $f \in K(p, \delta, a)$, where

$$K(p,\delta,a) := K := \{f : f \in W^{a,p}(0,1),\ \|f^{(a)}\|_p \le N_{a,p} < \infty,\ \|f_\delta - f\|_p \le \delta\}, \quad (2.5.15)$$

$a = 0$, $a = 1$, or $1 < a \le 2$. For $1 < a < 2$

$$\left\|f^{(a)}\right\|_p := \|f'\|_p + \sup_{x,y\in(0,1),\, x\neq y} \frac{\|f'(x) - f'(y)\|_p}{|x-y|^{a_0}}, \quad a = 1 + a_0, \quad 0 < a_0 < 1. \quad (2.5.16)$$

Therefore given $\delta > 0$ and f_δ one has to estimate f' for *any* $f \in K(p, \delta, a)$ and the problem of stable numerical differentiation has to be understood in the following sense:

Problem I:

Find a linear or nonlinear operator $R_{h(\delta)}$ such that

$$\sup_{f\in K(p,\delta,a)} \|R_{h(\delta)} f_\delta - f'\|_p \le \eta(\delta) \longrightarrow 0 \quad \textit{as} \quad \delta \longrightarrow 0, \quad (2.5.17)$$

where $\eta(\delta)$ is some positive continuous function of $\delta \in (0, \delta_0)$, and $\delta_0 > 0$ is some number.

The traditional formulation of the problem of stable numerical differentiation is different from the above:

Problem II:

Find a linear or nonlinear operator $R_{h(\delta)}$ such that

$$\sup_{f_\delta\in\mathscr{B}^p_{\delta,f}} \|R_{h(\delta)} f_\delta - f'\|_p \le \eta(\delta, f) \longrightarrow 0 \quad \textit{as} \quad \delta \longrightarrow 0, \quad (2.5.18)$$

where $\mathscr{B}^p_{\delta,f} := \{f_\delta : \|f_\delta - f\|_p \le \delta\}$ and $f \in K(p, \delta, a)$ is **fixed***,*

or even in a weaker form:

Find $R_{h(\delta)}$ such that

$$\left\|R_{h(\delta)} f_\delta - f'\right\|_p \le \eta(\delta, f) \longrightarrow 0 \quad \textit{as} \quad \delta \longrightarrow 0 \quad (2.5.20')$$

for a **fixed** *$f \in K(p, \delta, a)$ and* **fixed** *family $f_\delta \in \mathscr{B}^p_{\delta,f}$.*

Note the **principal difference** in the statements of Problems I and II of stable numerical differentiation: in Problem I the data are $\{f_\delta, N_{a,p}\}$, f is arbitrary in the set $K(p, \delta, a)$, and we wish to find a stable approximation of f', which is valid uniformly with respect to $f \in K(p, \delta, a)$. On the other hand, in Problem II it is assumed that $f \in K(p, \delta, a)$ is fixed and the approximation of f' is either uniform with respect to $f_\delta \in \mathscr{B}^p_{\delta,f}$ or holds for a particular family $f_\delta \in \mathscr{B}^p_{\delta,f}$. Since in practice we do not know f and we do know only one family f_δ, Problem I is much more important practically than Problem II. In this Section we show when one can obtain, in principle, a stable approximation of f' in the sense formulated in Problems I and II, and when it is not possible, in principle, to obtain such an approximation of f' from noisy data.

The main result on the stable numerical differentiation problem in the first formulation is stated in Theorem 2.5.1:

Theorem 2.5.1. *There does not exist an operator* $R_{h(\delta)} : L^p(0,1) \to L^p(0,1)$*, linear or nonlinear, for* $p = 2$ *and* $p = \infty$*, such that inequality* (2.5.17) *holds for* $a \le 1$*. If* $a \le 1$*, then*

$$\gamma^{\infty}_{\delta,a} := \inf_{R_{h(\delta)}:L^p(0,1)\longrightarrow L^p(0,1)} \sup_{f\in K(p,\delta,a)} \| R_{h(\delta)} f_\delta - f' \|_p \ge c > 0, \quad p = 2, \infty. \tag{2.5.19}$$

If $a > 1$ *and* $p \ge 1$*, then there does exist an operator* $R_{h(\delta)}$ *such that 2.5.17 holds. For example, one can use* $R_{h(\delta)}$ *defined in 2.4.92 with*

$$h = h_a(\delta) := \begin{cases} \left(\frac{2\delta}{a_0 N_{a,p}}\right)^{\frac{1}{a}}, & a = 1 + a_0, \quad 0 < a_0 < 1, \\ 2\left(\frac{\delta}{N_{2,p}}\right)^{\frac{1}{2}}, & a = 2. \end{cases} \tag{2.5.20}$$

The error of the corresponding differentiation formula is

$$\eta(\delta) := \begin{cases} a\,(N_{a,p})^{\frac{1}{a}} \left(\frac{2\delta}{a_0}\right)^{\frac{a_0}{a}}, & a = 1 + a_0, \quad 0 < a_0 < 1, \\ 2(\delta N_{2,p})^{\frac{1}{2}}, & a = 2. \end{cases} \tag{2.5.21}$$

The main result on the stable numerical differentiation problem in the second formulation is stated in Theorem 2.5.2:

Theorem 2.5.2. *If* $a = 1$*, then there exists an operator* $R_{h(\delta)} : L^2(0,1) \to L^2(0,1)$*, such that inequality* (2.5.18) *holds.*

The principal difference is: for $a = 1$ one cannot differentiate stably in the sense formulated in Problem I. In the sense of Problem II Stable differentiation is possible in principle. However the approximation error, $\| R_{h(\delta)} f - f' \|_2$, cannot be estimated, and this error $\eta(\delta, f)$ may go to zero arbitrarily slowly as $\delta \to 0$. This is in sharp contrast with the practically computable error estimate given in (2.5.21). Moreover, no matter how small the error bound $\delta > 0$ is, there exists a function $f \in \mathscr{K}^2_{\delta,1}$, such that not only $R_{h(\delta)}$ (with any fixed function $h(\delta)$), defined in (2.5.35), but *any other operator* $R_{h(\delta)}$, linear or nonlinear, will satisfy the inequality $\| R_{h(\delta)} f - f' \|_2 \ge c > 0$, where $c > 0$ does not depend on δ. This follows from (2.5.19).

Proofs

Proof of Theorem 2.5.1. First, consider the case $p = \infty$. Take

$$f_1(x) := -\frac{M}{2} x(x - 2h), \quad 0 \le x \le 2h. \tag{2.5.22}$$

Extend $f_1(x)$ from $(0, 2h)$ to $(2h, 1)$ by zero and denote the extended function by $f_1(x)$ again. Then $f_1(x) \in W^{1,\infty}(0, 1)$ and the norms $\|f^{(a)}\|_\infty$, $a = 0$ and $a = 1$ are preserved. Define $f_2(x) = -f_1(x)$, $x \in (0, 1)$. Note that

$$\sup_{x \in (0,1)} |f_k(x)| = \frac{Mh^2}{2}, \quad k = 1, 2. \tag{2.5.23}$$

Choose $h = h_\infty := h_\infty(\delta) := \sqrt{\frac{2\delta}{M}}$, so that

$$\frac{Mh_\infty^2}{2} = \delta, \tag{2.5.24}$$

Then for $f_\delta(x) \equiv 0$ one has: $\|f_k - f_\delta\|_\infty = \delta$, $k = 1, 2$. Let $(R_{h(\delta)} f_\delta)(0) = (R_{h(\delta)} 0)(0) := b$. One gets

$$\begin{aligned} \gamma_{\delta,a}^\infty &:= \inf_{R_{h(\delta)}} \sup_{f \in \mathscr{K}_{\delta,a}^\infty} \|R_{h(\delta)} f_\delta - f'\|_\infty \geq \inf_{R_{h(\delta)}} \max_{k=1,2} \|R_{h(\delta)} f_\delta - f_k'\|_\infty \\ &\geq \inf_{R_{h(\delta)}} \max_{k=1,2} \|R_{h(\delta)} f_\delta(0) - f_k'(0)\|_\infty \\ &= \inf_{b \in \mathbb{R}} \max \{|b - Mh_\infty|, |b + Mh_\infty|\} = Mh_\infty. \end{aligned} \tag{2.5.25}$$

If $h_\infty = \sqrt{\frac{2\delta}{M}}$, then $Mh_\infty = \sqrt{2\delta M}$. If $a = 0$, then (2.5.23) implies that $f_k \in \mathscr{K}_{\delta,0}^\infty$, $k = 1, 2$, with $N_{0,\infty} := \frac{Mh_\infty^2}{2} = \delta$. For any fixed $\delta > 0$ and $N_{0,\infty} = \delta$ the constant M in (2.5.22) can be chosen arbitrary. Therefore inequality (2.5.25) proves that (2.4.10) is false in the class $\mathscr{K}_{\delta,0}^\infty$ and, in fact, $\gamma_{\delta,0}^\infty \to \infty$ as $M \to \infty$.

Suppose now that $f \in \mathscr{K}_{\delta,1}^\infty$. One has

$$\|f_1'\|_\infty = \|f_2'\|_\infty = \sup_{0 \le x \le 2h_\infty} |M(x - h_\infty)| = Mh_\infty. \tag{2.5.26}$$

Thus, for given δ and $N_{1,\infty}$ one can take $h = h_\infty := \sqrt{\frac{2\delta}{M}}$, so that $\|f_k - f_\delta\|_\infty = \delta$, $k = 1, 2$, holds, and then take M so that $N_{1,\infty} = \sqrt{2\delta M}$. For these h_∞ and M the functions $f_k \in \mathscr{K}_{\delta,1}^\infty$, $k = 1, 2$. One obtains from (2.5.25) the following inequality

$$\gamma_{\delta,1}^\infty \geq N_{1,\infty} > 0 \quad \text{as} \quad \delta \longrightarrow 0, \tag{2.5.27}$$

which implies that estimate (2.4.10) is false in the class $\mathscr{K}_{\delta,1}^\infty$.

Now let $p = 2$. For f_1 defined in (2.5.22) one has

$$\|f_1\|_{L^2(0,2h)} = \frac{2}{\sqrt{15}} Mh^{\frac{5}{2}}, \quad \|f_1'\|_{L^2(0,2h)} = \sqrt{\frac{2}{3}} Mh^{\frac{3}{2}}. \tag{2.5.28}$$

Extend $f_1(x)$ from $(0, 2h)$ to $(2h, 1)$ by zero and denote the extended function $f_1(x)$ again. Then $f_1 \in W^{1,2}(0,1)$, $\|f_1\|_{L^2(0,1)} = c_1 M h^{\frac{5}{2}}$, and $\|f_1'\|_{L^2(0,1)} = c_2 M h^{\frac{3}{2}}$. Define $f_2(x) = -f_1(x)$, $f_\delta(x) \equiv 0$, $x \in (0,1)$.

Choose $h = h_2 := (\frac{\delta}{c_1 M})^{\frac{2}{5}} = \delta$ to satisfy the identity

$$c_1 M h_2^{\frac{5}{2}} = \delta, \tag{2.5.29}$$

then $\|f_k - f_\delta\|_{L^2(0,1)} = \delta$, $k = 1, 2$. Thus,

$$\begin{aligned} \gamma_{\delta,a}^2 &:= \inf_{R_{h(\delta)}} \sup_{f \in \mathscr{K}_{\delta,a}^2} \|R_{h(\delta)} f_\delta - f'\|_2 \geq \inf_{R_{h(\delta)}} \max_{k=1,2} \|R_{h(\delta)} f_\delta - f_k'\|_2 \\ &= \inf_{\varphi \in \mathscr{L}} \max\left\{\|\varphi - f_1'\|_2,\ \|\varphi + f_1'\|_2\right\}, \end{aligned}$$

where $\mathscr{L} := \{\varphi :\ \varphi = c f_1' + \psi,\ \ \psi \perp f'\}$. Therefore

$$\begin{aligned} \gamma_{\delta,a}^2 &\geq \inf_{c \in \mathbb{R},\, \psi \perp f'} \max\left\{\sqrt{(1-c)^2\|f_1'\|_2^2 + \|\psi\|_2^2},\ \sqrt{(1+c)^2\|f_1'\|_2^2 + \|\psi\|_2^2}\right\} \\ &= \inf_{c \in \mathbb{R}} \max\left\{|1-c|\|f_1'\|_2,\ |1+c|\|f_1'\|_2\right\} = \|f_1'\|_2 \\ &= c_2 M h_2^{\frac{3}{2}} = c_2 M^{\frac{2}{5}} \left(\frac{\delta}{c_1}\right)^{\frac{3}{5}}. \end{aligned} \tag{2.5.30}$$

If $a = 0$, then (2.5.29) yields $f_k \in \mathscr{K}_{\delta,0}^2$, $k = 1, 2$, with $N_{0,2} := c_1 M h^{\frac{5}{2}} = \delta$, and one gets $\gamma_{\delta,0}^2 \to \infty$ as $M \to \infty$.

Given constants δ and $N_{1,2}$ (in the case $a = 1$), one takes $h = h_2 := \left(\frac{\delta}{c_1 M}\right)^{\frac{2}{5}}$ so that $\|f_k - f_\delta\|_2 = \delta$, and then takes M so that $C_2 M^{\frac{2}{5}} \left(\frac{\delta}{c_1}\right)^{\frac{3}{5}} = N_{1,2}$. With this choice of h_2 and M the functions $f_k \in \mathscr{K}_{\delta,1}^2$, $k = 1, 2$, and one obtains from (2.5.30)

$$\gamma_{\delta,1}^2 \geq N_{1,2} > 0 \quad \text{as} \quad \delta \longrightarrow 0. \tag{2.5.31}$$

Finally, consider $a \in (1,2)$, $p \geq 1$. For the operator $R_{h(\delta)}$ defined by (2.4.92) one gets using the Lagrange formula:

$$\begin{aligned} \|R_{h(\delta)} f_\delta - f'\|_p &\leq \|R_{h(\delta)}(f_\delta - f)\|_p + \|R_{h(\delta)} f - f'\|_p \\ &\leq \frac{2\delta}{h} + N_{a,p} h^{a_0}. \end{aligned} \tag{2.5.32}$$

Minimizing the right-hand side of (2.5.32) with respect to $h \in (0, \infty)$ yields

$$h_a(\delta) = \left(\frac{2\delta}{a_0 N_{a,p}}\right)^{\frac{1}{a}}, \quad \eta(\delta) = a (N_{a,p})^{\frac{1}{a}} \left(\frac{2\delta}{a_0}\right)^{\frac{a_0}{a}}, \quad a = 1 + a_0, \quad 0 < a_0 < 1.$$

The case $a = 2$ is treated in the introduction (see estimate 2.4.93). So one arrives at 2.4.13–2.4.14. This completes the proof. □

Proof of Theorem 2.5.2. We give two proofs based on quite different methods.

The first proof uses the construction of the regularizing operator $R_{h(\delta)}$ defined in (2.4.92). The right-hand side of the error estimate of the type (2.4.93) is now replaced by $E(h) := \frac{2\delta}{h} + w(h)$, where $w \to 0$ as $h \to 0$, provided that $a = 1$. Minimizing E with respect to h for a fixed δ, one obtains a minimizer $h(\delta) \to 0$ as $\delta \to 0$ and the error estimate $E(h(\delta)) \to 0$ as $\delta \to 0$. Therefore one gets (2.6). Alternatively, if one chooses $h_1(\delta) \to 0$ as $\delta \to 0$, such that $\frac{2\delta}{h_1} = w(h_1)$, then $E(h_1(\delta)) \le 2w(h_1(\delta)) \to 0$. The first proof is completed. □

Remark 2.5.3. *The statement of Theorem 2.5.2 with* (2.4.11) *replaced by* (2.4.12) *is obvious: since f is fixed, one may take $R_{h(\delta)} f_\delta = f'$. This is, of course, of no practical use because f' is unknown.*

The second proof is based on the DSM. The ideas of this proof have an advantage of being applicable to a wide variety of ill-posed problems, and not only to stable numerical differentiation. By this reason we give this proof in detail. In order to show that for $a = 1$ there exists an operator $R_{h(\delta)} : L^2(0, 1) \to L^2(0, 1)$, ($L^2(0, 1)$ is a *real* Hilbert space) such that (2.5.18) holds we will use the DSM (dynamical systems method) (see Section 2.4). This approach consists of the following steps:
Step 1. Solve the Cauchy problem:

$$\dot{v} = -[Av + h(t)v - f_\delta], \quad v(0) = v_0 \in L^2(0, 1), \tag{2.5.33}$$

where A is defined in (2.4.7), $p = 2$, $\dot{v} := \frac{dv}{dt}$, $\|f_\delta - f\| \le \delta$ and

$$h(t) \in C^1[0, +\infty), \quad h(t) > 0, \quad h(t) \searrow 0, \quad \frac{\dot{h}(t)}{h^2(t)} \longrightarrow 0 \text{ as } t \longrightarrow +\infty. \tag{2.5.34}$$

Step 2. Calculate $v(t_\delta)$, where $t_\delta > 0$ is a number such that $t_\delta \to +\infty$ and $\frac{\delta}{h(t_\delta)} \to 0$ as $\delta \to 0$. Then

$$R_{h(\delta)} f_\delta := v(t_\delta) \tag{2.5.35}$$

and

$$\sup_{f_\delta \in \mathscr{B}^2_{\delta, f}} \|R_{h(\delta)} f_\delta - f'\|_2 \le \eta(\delta, f) \longrightarrow 0 \text{ as } \delta \longrightarrow 0 \tag{2.5.36}$$

with $\eta(\delta, f)$ given by (2.5.49) below and $f \in K(2, \delta, 1)$.

To verify (2.5.36) consider the problem

$$Aw + h(t)w - f = 0. \tag{2.5.37}$$

Since A is monotone in $L^2(0,1)$:

$$(A\phi,\phi) = \int_0^1 \left(\int_0^x \phi(\tau)d\tau\right)\phi(x)dx = \frac{1}{2}\int_0^1 \left[\left(\int_0^x \phi(\tau)d\tau\right)^2\right]' dx \geq 0, \tag{2.5.38}$$

for any $\phi \in L^2(0,1)$ and $h(t) > 0$, the solution $w(t)$ to (2.5.37) exists, is unique, and admits the estimate

$$(A(w - f'), w - f') + h(t)\|w\|_2^2 = h(t)(w, f'), \quad \|w\|_2 \leq \mathcal{N}_{1,2}. \tag{2.5.39}$$

Differentiate (2.5.37) with respect to t (this is possible by the implicit function theorem) and get

$$[A + h(t)I]\dot{w} = -\dot{h}(t)w, \quad \|\dot{w}\|_2 \leq \frac{|\dot{h}(t)|}{h(t)}\|w\|_2 \leq \frac{|\dot{h}(t)|}{h(t)}\mathcal{N}_{1,2}, \tag{2.5.40}$$

where (2.5.39) was used. Denote

$$z(t) := v(t) - w(t). \tag{2.5.41}$$

From (2.5.37) and (2.5.33) one obtains

$$\dot{z}(t) = -\dot{w} - [A + h(t)I]z + f_\delta - f, \quad z(0) = v_0 - w(0). \tag{2.5.42}$$

Multiply (2.5.42) by $z(t)$ and get

$$(\dot{z}, z) = -(\dot{w}, z) - (Az, z) - h(t)(z, z) + (f_\delta - f, z). \tag{2.5.43}$$

Let $\|z(t)\|_2 := g(t)$, then (2.5.39) and (2.5.43) imply

$$g\dot{g} \leq (\|\dot{w}\|_2 + \delta)g - h(t)g^2. \tag{2.5.44}$$

Since $g \geq 0$, from (2.5.44) and (2.5.40) it follows that

$$\dot{g} \leq \mathcal{N}_{1,2}\frac{|\dot{h}(t)|}{h(t)} + \delta - h(t)g(t), \quad g(0) = \|v_0 - w(0)\|. \tag{2.5.45}$$

So,

$$g(t) \leq e^{-\int_0^t h(s)ds}\left[g(0) + \int_0^t e^{\int_0^\tau h(s)ds}\left(\mathcal{N}_{1,2}\frac{|\dot{h}(\tau)|}{h(\tau)} + \delta\right)d\tau\right]. \tag{2.5.47'}$$

Under assumption (2.5.34), one has

$$\int_0^\infty h(t)dt = \infty. \tag{2.5.46}$$

Indeed, (2.5.34) implies $\frac{|\dot h|}{h^2} \le c$, $c = \text{const} > 0$, so $\frac{d}{dt}\frac{1}{h} \le c$, $\frac{1}{h(t)} - \frac{1}{h(0)} \le ct$, $\frac{1}{h(t)} \le c_0 + ct$, $c_0 := [h(0)]^{-1} > 0$, and $h(t) \ge \frac{1}{c_0+ct}$. Conclusion (2.5.46) follows.

If one chooses $t = t_\delta$ so that $t_\delta \to +\infty$ and $\frac{\delta}{h(t_\delta)} \to 0$ as $\delta \to 0$, then by (2.5.47') and (2.5.46), using L'Hospital's rule one obtains

$$\|v(t_\delta) - w(t_\delta)\|_2 := g(t_\delta) \longrightarrow 0 \quad \text{as} \quad \delta \longrightarrow 0. \tag{2.5.47}$$

The existence of the solution to (2.5.33) on $[0, +\infty)$ is obvious, since equation (2.5.33) is linear with a bounded operator.

We claim that

$$\|w(t_\delta) - f'\|_2 \longrightarrow 0 \quad \text{as} \quad \delta \longrightarrow 0 \tag{2.5.48}$$

For convenience of the reader this claim is established below. Equations (2.5.34), (2.5.45), (2.5.47), and (2.5.48) imply:

$$\sup_{f_\delta \in \mathscr{B}^2_{\delta,f}} \|v(t_\delta) - f'\|_2 \le \|w(t_\delta) - f'\|_2 + e^{-\int_0^{t_\delta} h(s)ds}\left[g(0) + \int_0^{t_\delta} e^{\int_0^\tau h(s)ds}\left(\mathscr{N}_{1,2}\frac{|\dot h(\tau)|}{h(\tau)} + \delta\right)d\tau\right]$$

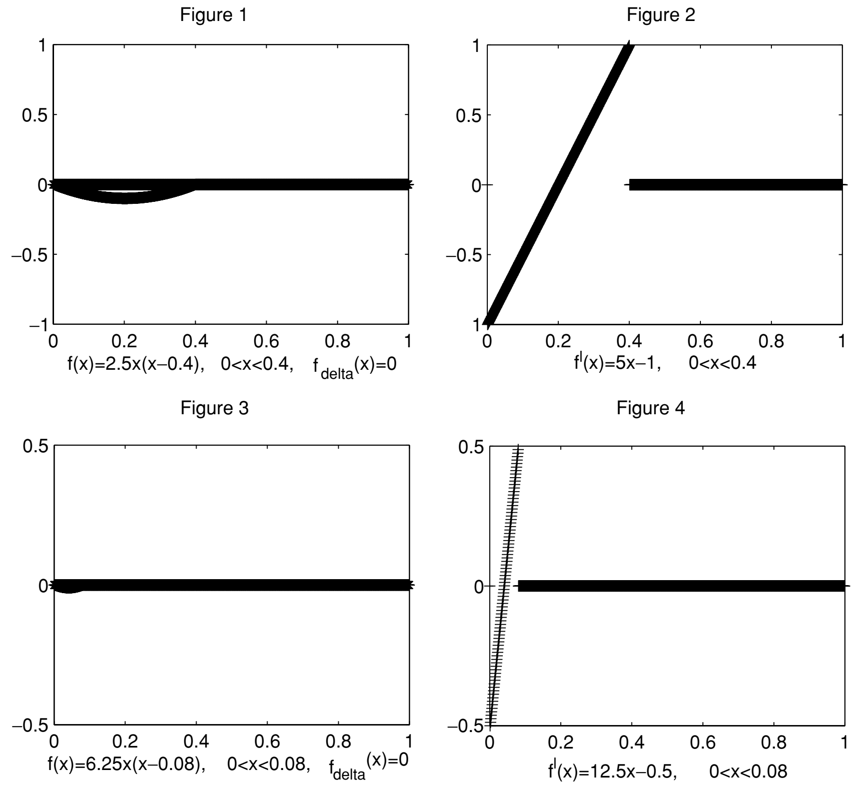
$$:= \eta(\delta, f) \longrightarrow 0 \quad \text{as} \quad \delta \longrightarrow 0. \tag{2.5.49}$$

Let us now prove (2.5.48). Because $f' \in L^2(0,1)$ and $f(0) = 0$, one can rewrite (2.5.37) as $A(w - f') + h(t)w = 0$. This and the monotonicity of A imply $h(t)(w, w - f') \le 0$, so, since $h(t) > 0$, one gets $(w, w - f') \le 0$, and $\|w\|_2 \le \|f'\|_2$. Thus, w converges weakly in $L^2(0,1)$ to some element ψ, $w \rightharpoonup \psi$ as $t \to \infty$. Because A is monotone, it is w-closed, that is $w \rightharpoonup \psi$ and $A(w - f') \to 0$ imply $A(\psi - f') = 0$, so $\psi = f'$ and $w \rightharpoonup f'$. The inequality $(w, w - f') \le 0$ can be written as $\|w - f'\|_2^2 \le (f', f' - w)$, and $(f', w - f') \to 0$ because $w - f' \rightharpoonup 0$. Therefore the claim (2.5.48) is proved and the second proof is completed.

Numerical aspects

Figures 1–4 illustrate the impossibility to differentiate stably a function, which does not have a bound on $f^{(a)}$, $a > 1$. If one takes the function

$$f(x) := \begin{cases} \frac{\mathscr{N}_{1,\infty}^2}{4\delta} x\left(x - \frac{4\delta}{\mathscr{N}_{1,\infty}}\right), & 0 \le x \le 4\delta \\ 0, & 4\delta < x \le 1, \end{cases} \tag{2.5.50}$$



and $f_\delta \equiv 0$, then

$$f(x) \in \left\{ f : \ f \in W^{1,\infty}(0,1), \ \|f'\|_\infty < \mathcal{N}_{1,\infty}, \ \|f - f_\delta\|_\infty \leq \delta \right\},$$

and any formula of numerical differentiation will give error not going to zero as $\delta \to 0$, because, by (2.5.27), one has:

$$\inf_{R_h(\delta)} \|R_h(\delta) f_\delta - f'\|_\infty \geq \mathcal{N}_{1,\infty}.$$

In Figure 1 one can see $f(x)$ given by (2.5.50) with $\delta = 0.1$ and $\mathcal{N}_{1,\infty} = 1$:

$$f(x) := \begin{cases} 2.5x(x-0.4), & 0 \leq x \leq 0.4 \\ 0, & 0.4 < x \leq 1. \end{cases} \tag{2.5.51}$$

Figure 2 presents

$$f'(x) := \begin{cases} 5x - 1, & 0 \le x \le 0.4 \\ 0, & 0.4 < x \le 1. \end{cases} \tag{2.5.52}$$

Figure 3 shows the case $\delta = 0.01$ and $\mathcal{N}_{1,\infty} = 0.5$:

$$f(x) := \begin{cases} 6.25x(x - 0.08), & 0 \le x \le 0.08 \\ 0, & 0.08 < x \le 1. \end{cases} \tag{2.5.53}$$

The derivatives are given in Figure 4:

$$f'(x) := \begin{cases} 12.5x - 0.5, & 0 \le x \le 0.4 \\ 0, & 0.4 < x \le 1. \end{cases} \tag{2.5.54}$$

Even if the bound on $f^{(a)}$ in some norm is given, one can experience difficulties with stable differentiation. Namely, if δ is fixed and $\mathcal{N}_{a,p}$ is very large, then h_{opt} in finite difference scheme (2.4.92) is very small, and practitioners might not have sufficiently many observation points. Another difficulty is: the estimated error $\varepsilon_{a,p}$ in such a case is very big and does not give any information regarding the accuracy of computations. This is illustrated in Table 2.1 below for the function $f(x) = \sin((\pi x)^n)$ and $\delta = 0.1$.

Figures 5–8 show the exact and computed derivatives of $f(x) = \sin((\pi x)^5)$. The derivatives of this function were computed in the presence of the noise function

$$e(x) = \delta\Big(\cos(2x) + \cos(3x^2)\Big)\Big/2, \tag{2.5.55}$$

and with different step sizes. One can see in Figure 5 that for h_{opt} the computed derivative is very accurate. However as h grows, the accuracy decreases.

Table 2.1

n	$\mathcal{N}_{2,p}$	h_{opt}	$\varepsilon_{a,p}$
5	$1.24 \cdot 10^6$	$5.67 \cdot 10^{-4}$	$7.04 \cdot 10^2$
10	$3.27 \cdot 10^{11}$	$1.11 \cdot 10^{-6}$	$3.62 \cdot 10^5$
15	$4.41 \cdot 10^{16}$	$3.01 \cdot 10^{-9}$	$1.33 \cdot 10^8$
20	$7.94 \cdot 10^{21}$	$7.10 \cdot 10^{-12}$	$5.63 \cdot 10^{10}$
25	$2.62 \cdot 10^{27}$	$1.24 \cdot 10^{-14}$	$3.24 \cdot 10^{13}$

Figure 5

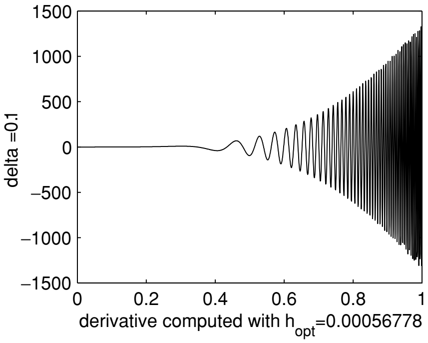


Figure 6

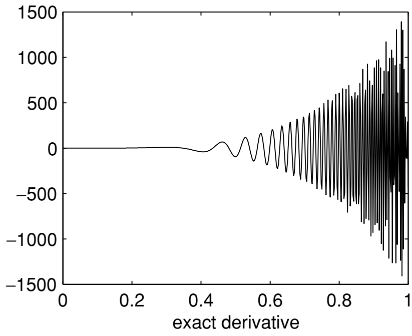


Figure 7

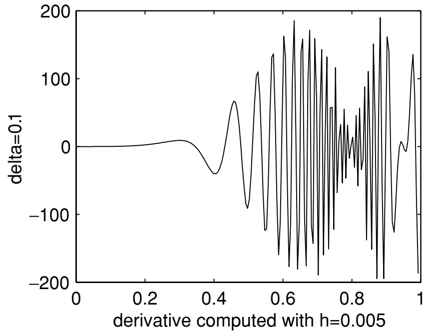
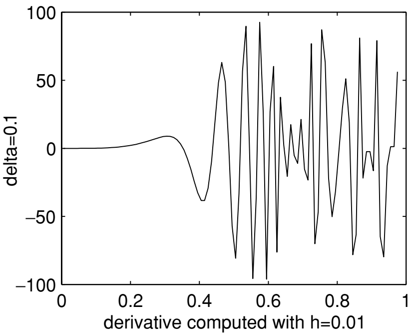


Figure 8



2.5.2 Stable summation of the Fourier series and integrals with perturbed coefficients

Assume that $f(x)$ is a smooth 2π-periodic function

$$f(x) = (2\pi)^{-\frac{1}{2}} \sum_{l=-\infty}^{\infty} f_l \exp(ilx), \tag{2.5.56}$$

where

$$f_l = (2\pi)^{-\frac{1}{2}} \int_{-\pi}^{\pi} f(x) \exp(-ilx) dx, \tag{2.5.57}$$

and

$$(1 + l^2)^{\frac{s}{2}} |f_l| \leq M_s, \quad s > \frac{3}{2}. \tag{2.5.58}$$

Assume that $f_{\delta l}$ are given such that

$$|f_{\delta l} - f_l| < \delta, \quad l = 0, \pm 1, \pm 2, \ldots, \tag{2.5.59}$$

and f_l are not known.

The problem is: given $f_{\delta l}$, calculate stably $f'(x)$. In other words, calculate F_δ such that

$$\| F_\delta - f' \| < \eta(\delta) \longrightarrow 0 \quad \text{as} \quad \delta \longrightarrow 0. \tag{2.5.60}$$

The norm here can depend on the particular problem. Let us assume that this is the $L^2[(-\pi, \pi)]$ norm. Let us look for F_δ of the form:

$$F_\delta(x) = (2\pi)^{-1/2} \sum_{l=-N}^{N} i l f_{\delta l} \exp(i l x). \tag{2.5.61}$$

One has, using Parseval's equality,

$$\begin{aligned} \| F_\delta - f' \|^2 &= \sum_{l=-N}^{N} l^2 |f_{\delta l} - f|^2 + \sum_{|l|>N} l^2 |f_l|^2 \\ &\leq \delta^2 \frac{N(N+1)(2N+1)}{3} + 2M_s^2 \sum_{l=N+1}^{\infty} \frac{l^2}{(1+l^2)^s} \\ &\leq \frac{2}{3}(N+1)^3 \delta^2 + \frac{2M_s^2}{2s-3}(N+1)^{-2s+3} \\ &:= c_0 \nu^3 \delta^2 + c_1 \nu^{-2s+3}, \qquad \nu := N+1, \end{aligned} \tag{2.5.62}$$

where the constants c_0 and c_1 are defined by the last equation. Minimizing the right-hand side of (2.5.62) with respect to $\nu > 1$, with $\delta > 0$ being fixed, we find the optimal $\nu := \nu(\delta)$:

$$\nu(\delta) = \left\{ \frac{(2s-3)c_1}{3c_0} \right\}^{\frac{1}{2s}} \delta^{-\frac{1}{s}} := c_2 \delta^{-\frac{1}{s}}, \tag{2.5.63}$$

and the error estimate:

$$\| F_\delta - f' \| \leq c_3 \delta^{\frac{2s-3}{s}} := \eta(\delta). \tag{2.5.64}$$

If $s = 2$, then $\eta(\delta) = c_3 \delta^{\frac{1}{2}}$. Let us summarize the result.

Theorem 2.5.4. *Assume that $f_{\delta l}$ are given such that* (2.5.58) *and* (2.5.59) *hold. Define F_δ by formula* (2.5.61) *with $N = N(\delta) = \nu(\delta) - 1$, where $\nu(\delta)$ is defined in* (2.5.63). *Then* (2.5.60) *holds with $\eta(\delta)$ defined by* (2.5.64).

The above arguments are applicable also to Fourier integrals with perturbed Fourier transforms, which play the role of the perturbed coefficients.

2.5.3 Stable solution of some Volterra equations of the first kind

Consider equation (1.3.1) with $Au := Vu = \int_0^x V(x, y)u(y)dy$. If $V(x, y) \geq c > 0$ is a bounded function, and $V_x(x, y)$ is a kernel for which $\max_{0\leq x\leq B} \int_0^x |V_x(x, y)|^2 dy \leq c_1$, then (1.3.1), after a differentiation, yields

$$u(x) + \int_0^x \frac{V_x(x, y)u\, dy}{V(x, x)} = \frac{f'(x)}{V(x, x)}, \quad 0 \leq x \leq l, \tag{2.5.65}$$

provided that $f' \in L^2(0, b)$. This is a Fredholm second-kind integral equation, for which the Fredholm alternative is valid in $L^2(0, b)$.

If $V(x, y)$ is not differentiable with respect to x, or $V(x, x)$ may vanish, then equation (1.3.1) with $A = V$ can be solved by the methods discussed in Section 2.1–2.4.

A different general approach to stable solution of the equation $Vu = f$ consists of the factorization $Vu = Q(I + S)u = f$, where S is a compact operator, I is the identity, the null-space $N(I + S) = \{0\}$ is trivial, so that by Fredholm's alternative the equation $(I + S)u = w$ can be stably solved by a projection method, and the operator Q is such that w can be stably found from the equation $Qw = f$. If the noisy data f_δ are given in place of f, then a stable solution to the equation $Qw = f_\delta$ is given by a formula $w_\delta = R_\delta f_\delta$, and a stable solution of the equation $Vu = f$ with noisy data f_δ is given by the formula $u_\delta = (I + S)^{-1} R_\delta f_\delta$. A numerical implementation of this scheme is given in [RSm6]. If $V(x, y) = V(x - y)$ and $V(0) \neq 0$, then the above scheme leads to the operator Q whose stable inversion is equivalent to stable differentiation, $Qw = \int_0^x w(y)dy$.

A discretization method for stable solution of Volterra integral equations of the first kind is proposed and justified in [RG], where the error estimates for the proposed method are also obtained.

2.5.4 Deconvolution problems

Equation (1.3.1) with $Au = \int_D A(x, y)u(y)dy$ can be solved by the methods of Sections 2.1–2.4, provided that $u \in H$. In the special case, when the kernel

$$A(x, y) = R(x, y), \quad \text{and} \quad QR(x, y) = P\delta(x - y) \quad \text{in} \quad \mathbb{R}^r, \quad r \geq 1, \tag{2.5.66}$$

where Q and P are elliptic operators, $\delta(x - y)$ is the delta-function, $\Re Q \geq c > 0$, one can use the theory from [R121], [R189], and prove, under suitable assumptions, that the operator R, $Rh := \int_D R(x, y)h(y)dy$, is an isomorphism of $H^{-a}(D)$ onto $H^a(D)$, where $a = \frac{n-m}{2}$, $n = \text{ord}\, Q$, $m = ord P$, $n \geq m$, $H^a(D)$ is the usual Sobolev space, $H^{-a}(D)$ is its dual with respect to $L^2(D)$ inner product, $H^{-a}(D)$ is, in other words, the closure of $C_0^\infty(D)$ in the norm of $H^{-a}(\mathbb{R}^r)$, that is, the subset of the elements of $H^{-a}(R^r)$ with support $\overline{D}$, $\overline{D} := D \cup S$, $S := \partial D$.

If L is a selfadjoint elliptic operator in $L^2(R^2)$, ord $L := s$ and $Q(\lambda)$ and $P(\lambda)$ are positive polynomials, deg $Q = n$, deg $P = m \le n$, then $Q := Q(L)$ and $P := P(L)$ are elliptic operators, and if $R(x, y)$ solves (2.5.1) then $a = \frac{s(n-m)}{2}$. Equation $Rh = f$ has a unique solution in $H^{-a}(D)$ for any $f \in H^a(D)$, and $R: H^{-a}(D) \to H^a(D)$ is an isomorphism. In [R121], [R189] one finds analytical formulas for the solution h. Under the above assumptions, the equation (*) $Rh = f$ in $\bar{D}$ does not have integrable solutions, in general. It has only distributional solutions of finite order of singularity, in general. Finding a solution of minimal order of singularity (mos solution) is a well posed problem ([R121]). The minimal order of singularity is equal to a. Since R is an isomorphism of $H^{-a}(D)$ onto $H^a(D)$ the problem of solving equation (*) is well-posed. One does not need regularization methods for finding the solution to (*).

Example 2.5.1.

$$Rh = \int_{-1}^{1} e^{-|x-y|} h(y)dy = f(x), \quad -1 \le x \le 1.$$

In this example $L = \frac{1}{i}\frac{d}{dx}$, $P(\lambda) = 1$, $Q(\lambda) = \frac{\lambda^2+1}{2}$, $a = 1$, $e^{-|x|} = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{e^{i\lambda x} d\lambda}{\lambda^2+1}$, $h(x) = \frac{-f''(x)+f(x)}{2} + \delta(x-1)\frac{f'(1)+f(1)}{2} + \frac{-f'(-1)+f(-1)}{2}\delta(x+1)$.

One can see that generically $h(x)$ is not an integrable function, it is a distribution, $h \in H^{-1}(-1, 1)$. If and only if $f'(1) + f(1) = 0$ and $f'(-1) = f'(1)$ and $f'' \in L^1(-1, 1)$, is $h \in L^1(-1, 1)$ Thus, equation

$$Rh_\delta = f_\delta, \quad \| f_\delta - f \|_{H^1} \le \delta, \quad f \in H^1,$$

has a solution h_δ which depends on f_δ continuously in the norm $H^{-1}(-1, 1)$: $\|h_\delta - h\|_{H^{-1}(-1,1)} \le c\delta$ where $c = \| R^{-1} \|_{H^1 \to H^{-1}} < \infty$. Details one can see in [R121].

2.5.5 Ill-conditioned linear algebraic systems

Let $A: \mathbb{R}^n \to \mathbb{R}^n$ in (1.3.1). Then A can be represented by a matrix (a_{ij}). If A is non-singular, define its condition number

$$\gamma(A) := \|A\| \left\| A^{-1} \right\|.$$

One has (see (1.4.3)) $\frac{\|\Delta u\|}{\|u\|} \frac{\|\Delta f\|}{\|f\|} \le \gamma(A) \le \frac{\|A^{-1}\|\|\Delta f\|}{\|u\|} \Big/ \frac{\|\Delta f\|}{\|f\|} = \|A^{-1}\|\|A\| = \gamma(A)$. Thus, if $\gamma(A)$ is large then small relative errors $\frac{\|\Delta f\|}{\|f\|}$ in the data may lead to large relative errors $\frac{\|\Delta u\|}{\|u\|}$ in the solution. Solving linear algebraic ill-conditioned is an ill-posed problem practically. Note that $\gamma(A) = sup_{u\neq 0}(\frac{\|Au\|}{\|u\|}) / \inf_{v\neq 0}(\frac{\|Av\|}{\|v\|})$, because $\inf_{v\neq 0} \frac{\|Av\|}{\|v\|} = \|A^{-1}\|^{-1}$, if A is not singular (not degenerate: det $A \neq 0$) Also, $\gamma(A) = \frac{s_{\max}}{s_{\min}}$, where s_j^2 are eigenvalues of A^*A.

Examples below are taken from [VA].

Example 2.5.2. $A = \begin{pmatrix} 4.1 & 2.8 \\ 9.7 & 6.6 \end{pmatrix}$, $f = (4.1, 9.7)$, $u = (1, 0)$, $Au = f$. *Let* $g = (4.11, 9.70)$. *Then* $v = (0.34, 0.97)$. *Here* $\mu(A) = 2, 249.5$.

Example 2.5.3. *Hilbert matrix* $a_{ij} = (i + j - 1), 1 \le i, j \le n$, $\mu(A) = 1.510^7$ *if* $n = 6$, $\mu(A) = 1, 6.10^{13}$ *if* $n = 10$.

2.6 PROJECTION METHODS FOR ILL-POSED PROBLEMS

Let $A : H \to H$ be a linear, injective, bounded operator in a Hilbert space H, $A^{-1} : R(A) \to H$ is unbounded, so that the problem of solving the equation $Au = f$ is ill-posed.

Let $H_n \subset H_{n+1} \subset \dots$ be a sequence of finite-dimensional subspaces of H which is limit-dense in H, that is, for any $f \in H$ one has $\lim_{n\to\infty} \rho(f, H_n) = 0$, where $\rho(f, H_n) = \inf_{v \in H_n} \|f - v\|$ is the distance from f to H_n. Let P_n be the orthoprojection operators on H_n. Assume that f_δ is given such that $\|f_\delta - f\| \le \delta$. The problem consists of finding $u_\delta := u_{m(\delta)}$, which solves the equation (*) $P_m Au_{m(\delta)} = P_m f_\delta, u_{m(\delta)} \in H_m$, and such that $\lim_{\delta\to 0} \|u_{m(\delta)} - u\| = 0$. Equation (*) is an equation of a projection method. A general approach to finding a stable approximation $u_{m(\delta)}$ is the following one: let R_h be a regularizer in the sense $\lim_{h\to 0} R_h Au = u$ $\forall u \in H$, and define $u_\delta = u_\delta(h, m) := R_h P_m AP_m f_\delta$. Then $\|u_\delta - u\| \le \|R_h P_m f_\delta - R_h f_\delta\| + \|R_h f_\delta - R_h f\| + \|R_h f - u\| \le \|R_h\|(\|(P_m - I)f_\delta\| + \|f_\delta - f\|) + \|R_h Au - u\| := a(h)l(m, \delta) + p(h)$, where $p(h) := \|R_h Au - u\| \to 0$ as $h \to 0$, $a(h) := \|R_h\| \to \infty$ as $h \to 0$, $l(m, \delta) := \|f_\delta - f\| + \|P_m f_\delta - f_\delta\| \to 0$ as $\delta \to 0$ and $m = m(\delta)$ is chosen suitably. More precisely, for any fixed f_δ one has $\lim_{m\to\infty} \|P_m f_\delta - f_\delta\| = 0$, because the sequence H_n is limit-dense. Thus, one can choose $m = m(\delta)$ so that $\|P_m f_\delta - f_\delta\| \le \delta$, then $l(m(\delta), \delta) \le 2\delta$, and for a fixed $\delta > 0$, one can choose $h = h(\delta)$ so that $2a(h)\delta + p(h) = \min := m(\delta) \to 0$ as $\delta \to 0$. Therefore $\|R_{h(\delta)} P_{m(\delta)} f_\delta - u\| \le m(\delta) \to 0$ as $\delta \to 0$, so that u is stably approximated.

There are many ways to choose the regularizer R_h: one can use a Variational Regularization, Quasisolutions, Iterative Regularization, and the DSM. One can also use the method developed in Section 2.5.2 for constructing a convergent projection method for solving (1.3.1). Assume that $X = Y = H$ in (1.3.1), A is a linear compact operator, $\|A\| \le 1$, and (1.3.1) is solvable, i.e, $f = Ay$ for some y. We assume that $y \perp N(A)$, that is, y is the minimal-norm solution. By Lemma 2.1.11, a solvable equation (1.3.1) is equivalent to equation (2.1.6). Denote $g := f_1 := A^* f$, assume that g_δ is given in place of g, and $\|g_\delta - g\| \le \delta$. Let $B\varphi_j = s_j^2 \varphi_j$, where φ_j, $j = 1, 2, \dots\dots\dots$, is an orthonormal system of eigenvectors of the selfadjoint compact operator B, and $s_j^2 > 0$ are the corresponding eigenvalues. Let $c_j := \frac{g_{\delta,j}}{s_j^2}$, and $u_\delta := \sum_{j=1}^N c_j \varphi_j$. The element u_δ is a regularizer if $N = N(\delta)$ is chosen properly, as in Section 2.5.2. In this case one has $\lim_{\delta\to 0} \|u_\delta - y\| = 0$.

See more on this topic in [IVT], [VA].

http://www.springer.com/978-0-387-23195-2

Inverse Problems
Mathematical and Analytical Techniques with
Applications to Engineering
Ramm, A.G.
2005, XX, 442 p., Hardcover
ISBN: 978-0-387-23195-2