VIII. Compact perturbations of the identity

This is the first of several chapters in which the basic knowledge accumulated in the preceding chapters is used for the analysis of various numerical procedures.

Projection methods

The (Rayleigh-)Ritz method consists in minimizing the quadratic functional

$$\Phi_\lambda : x \mapsto \|x\|_A^2 - 2\,\mathrm{Re}\,\lambda x$$

over a sufficiently large lss F of the ips X. The minimizer f is the representer of the linear functional $\lambda|_F$ with respect to the inner product $\langle , \rangle_A := \langle \cdot, A\cdot \rangle$, hence an approximation to the solution u of the **Euler equation**

$$A? = y$$

associated with Φ_λ, in which $\langle \cdot, y\rangle = \lambda$. In fact, f is the best approximation to u from F wrto the A-norm, hence computable by interpolation, since it solves the LIP

$$? \in F, \quad u - ? \perp_A F.$$

This means that the interpolation functionals are of the special form

$$v^c A \;=\; \langle v, A\cdot \rangle, \quad v \in F.$$

Galerkin noticed that this procedure is capable of vast generalization. There is really no reason to insist that A be positive definite, i.e., that $\langle , \rangle_A$ be an inner product. We can use this LIP even when A is just any old lm. The price we pay for this is that we cannot be sure any more that the LIP is correct.

Further generalizations were offered by others to the point that **interpolation** is now a standard approach to solving functional equations of all kinds. The idea is this: Let $A \in L(X, Y)$. To approximate the solution u of the equation

$$A? = y,$$

pick some $V \in L(\mathbb{F}^n, X)$ and some $\Lambda \in L(\mathbb{F}^n, Y')$ and look for $f \in F := \mathrm{ran}\, V$ that solves

$$\Lambda^t A? = \Lambda^t y.$$

In other words, we approximate the solution u by an interpolant, i.e., by the solution to the $\mathrm{LIP}(\mathrm{ran}\, V, \mathrm{ran}\, A'\Lambda)$, making use of the fact that we are given y, i.e., Au, hence can compute $(A'\lambda)u = (\lambda A)u$ for any particular λ we care to (at least in principle), hence may try to match that information.

**** examples ****

(1) ODE-Example Consider the m-th order ODE

$$Au := D^m u - \sum_{j<m} a_j D^j u = y,$$

for which a solution u is sought in $X := C^{(m)}([a\,..\,b]) \cap \ker \mathrm{M}^t$, with $\mathrm{M}^t \in bL(C^{(m-1)}, \mathbb{R}^m)$ the data map that specifies the m (homogeneous) side conditions needed to select a unique solution from the m-dimensional solution set available for the ODE in $C^{(m)}([a\,..\,b])$. Then $Y = C([a\,..\,b])$ is an appropriate choice. Let F be some n-dimensional lss of X.

In **collocation**, one chooses $\Lambda^t : f \mapsto f|_U$ for some n-set U in $[a\,..\,b]$. In effect, the approximation f from F is chosen so as to satisfy the ODE exactly at the points of U.

In **Galerkin's method**, $\operatorname{ran}\Lambda$ consists of the lfl's

$$x \mapsto \int_a^b z(t)x(t)\,\mathrm{d}t, \qquad \text{all } z \in F.$$

This means that the **residual** $y - Af$ is made orthogonal to F.

In the **least-squares method**, one uses the lfl's

$$x \mapsto \int_a^b (Az)(t)x(t)\,\mathrm{d}t, \qquad \text{all } z \in F$$

instead. This means that the residual is minimized (over F) in the $\mathbf{L}_2$ sense.

In the **moment method**, one uses the linear functionals

$$x \mapsto \int_a^b t^j x(t)\,\mathrm{d}t, \qquad j = 0, \ldots, n-1,$$

thereby making the first n moments of the residual equal to zero. etc.

**** residual reduction ****

These methods are also called "**residual reduction methods**", since they can be viewed as an attempt to make the residual zero in a certain sense. Precisely, the residual is made to vanish on certain lfl's.

These methods are also called **projection methods**. This is *not* because they are based on interpolation, i.e., the approximation f to u is given as $f = Pu$, with P the lprojector given by F and $L := \operatorname{ran} A'\Lambda$. Rather, they got this name since f is sought as the solution in F of the **projected equation** $QAf = Qy$, with Q any l.projector on Y whose interpolation functionals are $\operatorname{ran}\Lambda$.

**** analysis by perturbation from a simple case ****

In each instance, one has to settle the correctness of the LIP, bound the resulting projector in suitable norms and then leave it to Lebesgue's Inequality and Approximation Theory to provide *a priori* error bounds. Correctness and norm bounds are usually obtained for a very simple instance A_0 of A. The general A is treated as a **perturbation** of the simple A_0, under the assumption that $A - A_0$ is "small" in some sense.

In the simplest case, one tries to use ${A_0}^{-1}$ as an *approximate inverse* for A. This requires, practically speaking, (cf. (III.15)Prop.) that

$$\|A - A_0\| < 1/\|{A_0}^{-1}\|,$$

hence has only limited applicability. As it turns out, it is often sufficient to assume that $A - A_0$ be *compact*, without any assumption on the norm of $A - A_0$.

Compact linear maps

(2) Definition. *$K \in L(X,Y)$ is* **compact** *(or,* **completely continuous***, or better but unconventional,* **totally bounded***) := K carries bounded sets to totally bounded sets, i.e., KB is totally bounded. I denote the collection of all compact linear maps from X to Y by*

$$cL(X,Y).$$

In particular, a compact lm is bounded. While this definition makes sense in more general topological spaces, we restrict attention to nls's X, Y.

**** examples ****

Any bounded finite-rank lm is compact (since any bounded set in a finite dimensional ls is totally bounded).

The (norm) limit K of any sequence (K_n) of compact lm's is compact. Indeed, $KB \subseteq (K - K_n)B + K_nB$, and, for $r > 0$, can choose n so that $(K - K_n)B \subseteq B_{r/2}$ and, for that n, can choose a finite $r/2$-net for K_nB.

In particular, all (norm) limits of bounded finite-rank lm's are compact, i.e., compact maps are the *only* maps uniformly (i.e., norm-) approximable by finite-rank maps, hence their importance in Numerical Analysis.

H.P.(1) Prove that, *if $k \in C(R \times T)$ for R, T compact in $\mathbb{R}^n$ and $1 \le p \le \infty$, then $K : \mathbf{L}_p(T) \to C(R) : f \mapsto \int_T k(\cdot,t) f(t)\,\mathrm{d}t$ is compact.*

Any finite linear combination of compact maps is compact, i.e., $cL(X,Y)$ is a lss of $L(X,Y)$. *If K is compact and A is bounded, then AK and KA are compact.* (In particular, $cL(X)$ is a (closed) two-sided ideal in $bL(X)$.)

**** compactness and convergence ****

The compactness of a map is used to force convergence or improve the mode of convergence: If (x_n) is a bounded sequence, then, by compactness of K, (Kx_n) is a totally bounded sequence, hence (cf. H.P.(II.33)) has a Cauchy subsequence. Consequently, if Y is complete, then some subsequence of (Kx_n) converges. This fact is usually used in the following form:

(3) Proposition (AGTASMAT). *If $K \in cL(X,Y)$ with Y a Bs, and (x_n) is bounded, then,* **A***fter* **G***oing* **T***o* **A** **S***ubsequence,* **M***ay* **A***ssume* **T***hat (Kx_n) converges.*

We also use the following consequence of the fact (cf. H.P.(V.3)) that bounded strong convergence is uniform on totally bounded sets:

(4) Proposition. *K compact, (A_n) bounded and $A_n \xrightarrow{s} A \quad \Longrightarrow \quad A_nK \to AK$.*

H.P.(2) Give an example of K compact, (A_n) bounded and $A_n \xrightarrow{s} A$ for which KA_n fails to converge (in norm) to KA. (Hint: Perhaps $A_n \in bL(X, \mathbb{R})$ and $K = 1$ will do it?)

(5) Corollary. *If $bL(X)$ contains an approximate identity, then $K \in bL(X)$ is compact iff it is the (norm) limit of finite-rank maps.*

An equivalent definition of compact lm involves weak convergence (cf. Chapter V).

(6) Proposition. *A compact map into a Bs carries weakly convergent sequences to norm convergent ones, i.e., $K \in cL(X, Y)$ with Y Bs, and $x_n \xrightarrow{w} x \quad \Longrightarrow \quad \lim Kx_n = Kx$.*

Proof: By (V.10)Corollary to Uniform Boundedness Principle, (x_n) is bounded, hence, AGTASMAT (i.e., by (3)), every subsequence of (Kx_n) has limit points. If y is such a limit point, i.e., $y = \lim Kx_{m(n)}$, then $\forall\{\mu \in Y^*\}$ $\mu y = \lim \mu K x_{m(n)} = \mu K x$ (since $\mu K \in X^*$), hence $y = Kx$ is the only limit point. Thus $Kx = \lim Kx_n$, by H.P.(II.27)(ii). □

** K compact $\Longrightarrow$ dual of K compact **

(7) Proposition. *$K \in bL(X, Y)$. K compact $\quad \Longrightarrow \quad K^*$ compact.*

Proof: We have to show that $K^* B_{Y^*} = B_{Y^*}K$ is totally bounded (with B_{Y^*} the unit ball for Y^*). This is equivalent to showing that $(B_{Y^*}K)|_B$ is totally bounded (in the Bs $b(B)$ of all bounded functions on $B :=$ unit ball in X). Since any ε-net $U|_{KB}$ for $(B_{Y^*})|_{KB}$ provides the ε-net $UK|_B$ for $(B_{Y^*}K)|_B$, it is sufficient to show that $(B_{Y^*})|_{KB}$ is totally bounded. Since K is compact, KB is totally bounded, hence, by (II.40)Lemma (essence of Arzela-Ascoli), it is sufficient to show that B_{Y^*} is bounded and equicontinuous on KB. But B_{Y^*} is bounded and equicontinuous on *any* bounded set. □

H.P.(3) Let $K \in cL(X, Y)$, 1-1, and set $E := K^{-1} \in L(\operatorname{ran} K, X)$. Prove:
(i) *if E is bounded, then* $\dim X < \infty$.
(ii) For all totally bounded T with $0 \notin T^-$ and T^- complete, $\|E|_{KT}\| := \sup_{y \in KT} \|Ey\|/\|y\| < \infty$.

** regularization **

The fact that a compact map turns bounded sets into totally bounded ones has its flip side, in that an equation $K? = y$ with K compact is troublesome, for the following reason. Even if K is invertible, it cannot be boundedly invertible unless its domain is finite-dimensional (cf. H.P.(3)). Hence, such an equation cannot be solved stably, i.e., it is **ill posed**.

Nevertheless, people trying to solve real problems find themselves confronted with ill posed problems, and, in this unhappy situation, 'solve' such a problem numerically by converting it into a stable equation. This means that, instead of solving $K? = y$, they solve $(K - \alpha)? = y$ for some $\alpha > 0$ (such equations are well-posed as the discussion below will make clear). This approach (and similar ones) is called **regularization**, an idea first pushed by Tykhonov. The situation is further complicated by the fact that y may only be known approximately. Thus, if the computed solution, x_c, satisfies $(K - \alpha)? = y_c$, then the error, $x - x_c$, satisfies $(K - \alpha)(x - x_c) = y - y_c - \alpha x$, or, $\|x - x_c\| \le \|(K -$

$\alpha)^{-1}\|(\|y - y_c\| + \alpha\|x\|)$. As $\alpha \to 0$, $\|(K-\alpha)^{-1}\| \to \infty$, while, when $\alpha \to \infty$, then $\|x_c\| \to 0$, hence $\|x - x_c\| \to \|x\|$. There is, therefore, for given $\|y - y_c\|$, some *optimal* α, but its determination is not trivial; also, even with this optimal α in hand, the computed solution may have little in common with the (only theoretically defined) 'exact' solution.

A straightforward approach would be to recognize explicitly that the equation $K? = y$ fails to define its solution in any practical sense, hence to seek additional conditions (e.g., a certain smoothness, the smallness of certain seminorms, the vanishing of certain linear functionals, etc.) that, together with the equation $K? = y$, pin down a particular element *stably*. Regularization is certainly a particularly simple version of this approach, as long as it is understood in this way.

Compact perturbation of the identity

A **compact perturbation of the identity** is a map of the form $1 - K$ with K compact.

Such maps occur naturally in **integral equations of the second kind**:

$$f - \int_T k(\cdot, t) f(t)\,\mathrm{d}t = g,$$

which is to be solved for $f \in C(T)$, given $g \in C(T)$ and $k \in C(T^2)$. (Of the **second kind**, since it involves the unknown function f in *two* places, as opposed to the **first kind** integral equation

$$\int_T k(\cdot, t) f(t)\,\mathrm{d}t = g$$

in which the unknown function appears only once; a silly but standard nomenclature sanctified by long use.) As already noted, first-kind equations are hard to solve.

**** spectrum of a compact map ****

Compact perturbations of the identity also occur in the study of the **spectrum** of a compact map K, i.e., in studying $z \in \mathbb{C}$ for which $z - K$ is not boundedly invertible. For, if $z \neq 0$, then $z - K$ is invertible iff $1 - K/z$ is, while K/z is compact iff K is. In this connection:

(8) Proposition. *K compact* $\implies$ $\dim\ker(1-K) < \infty$.

Proof: $K = 1$ on $\ker(1-K)$, hence $B_{\ker(1-K)}$ is totally bounded, hence $\dim(\ker(1-K)) < \infty$, by H.P.(III.16). □

(9) Proposition. *X Bs and $K \in cL(X)$* $\implies$ $\operatorname{ran}(1-K)$ *closed.*

Proof: With $N := \ker(1-K)$, $\operatorname{ran}(1-K)$ is the range of the linear map $C := (1-K)_| : X/N \to X : \langle x\rangle \mapsto (1-K)x$, hence is closed in case C is bounded below. In the contrary case, $\inf_x \|Cx\|/d(x,N) = 0$, hence there is (x_n) in X with $\|\langle x_n\rangle\| = d(x_n, N) = 1$ and $\lim \|x_n - Kx_n\| = 0$. Without loss of generality, (x_n) is bounded, AGTASMAT $y := \lim_n Kx_n$ exists, hence $x_n \to y$ and therefore $d(y,N) = 1$, yet $(1-K)y = 0$, which is nonsense. □

(10) Fredholm Alternative. *X Bs, $K \in cL(X)$, $A := 1 - K$. Then*

$$A \text{ 1-1} \iff A \text{ onto}.$$

Proof: '$\Longleftarrow$', i.e. suppose A onto. Then, if $x \in \ker A^n \backslash \ker A^{n-1}$, there is y s.t. $Ay = x$, i.e., $y \in \ker A^{n+1} \backslash \ker A^n$. Hence, if A is *not* 1-1, then there is $x \in \ker A \backslash 0$, and then $(\ker A^n)$ is a *strictly* increasing sequence of closed lss's. By (III.7)Riesz' Lemma, $\forall\{r > 0, n\}\ \exists\{x_n \in B \cap \ker A^n\}\ d(x_n, \ker A^{n-1}) > 1 - r$. Hence, for $m < n$,

$$\|Kx_n - Kx_m\| = \|x_n - (Ax_n + x_m - Ax_m)\| > 1 - r$$

(since Ax_n, x_m, Ax_m are all in $\ker A^{n-1}$), showing that (Kx_n) is *not* totally bounded, even though K is compact and (x_n) is bounded.

'$\Longrightarrow$', i.e., suppose A 1-1. Since $\operatorname{ran} A$ is closed by (9)Prop., we can conclude that $\tilde{A} := A|^{\operatorname{ran} A}$ is boundedly invertible (by (V.16)Cor.2), hence $\forall\{\lambda \in X^*\}\ \lambda = \mu A$ with $\mu := \lambda(\tilde{A})^{-1} \in X^*$, therefore A^* is onto. Now, since K^* is compact (by (7)Prop.), the first part of the proof implies that A^* is 1-1, hence, since $\operatorname{ran} A$ is closed, we conclude (from H.P.(IV.14)(i)) that $\operatorname{ran} A = X$, i.e., A is onto. □

H.P.(4) Adjust the preceding argument to prove that $Ky_n = z_n y_n$, with $\|y_n\| = 1$, $n = 1, 2, \ldots$, and (z_n) a sequence of *distinct* points in $\mathbb{F}$, implies $\lim z_n = 0$. (Hint: Prove first that $Y_n := \operatorname{ran}[y_1, y_2, \ldots, y_n]$, $n = 1, 2, \ldots$ is a strictly increasing sequence and that $(z_n - K)Y_n \subset Y_{n-1}$.)

H.P.(5) Adjust the preceding argument to prove the following more complete version of the **Fredholm Alternative**: *For $K \in cL(X)$ and $z \in \mathbb{F}\backslash 0$, there exists $n \in \mathbb{N}$ so that $\ker(z - K)^n = \ker(z - K)^{n+r}$ for all $r \in \mathbb{N}$. Hence, if X is a Bs, then $X = \ker(z - K)^n \dot{+} \operatorname{ran}(z - K)^n$.* (Hint: H.P.(I.38))

For *example*, the second kind integral equation

$$f - \int_T k(\cdot, t) f(t)\, \mathrm{d}t = g$$

with $k \in C(T^2)$ has a (unique) solution $f \in C(T)$ for every $g \in C(T)$ iff it has only the trivial solution $f = 0$ when $g = 0$, and, in that case, f depends continuously on g, by (V.18)OMT.

H.P.(6) Prove the **full Fredholm Alternative**: *If X is Bs, and $K \in cL(X)$, then (*$\operatorname{ran}(1 - K)$ *is closed and) both* $\ker(1 - K)$ *and* $\ker(1 - K^*)$ *have the same finite dimension.* (Hint: H.P.(I.38).)

We are now prepared for

The standard compact perturbation argument

Assume that X is a Bs, $K \in cL(X)$, and $A := 1 - K$ 1-1 or onto, hence boundedly invertible. To approximate the solution of the equation

$$(1 - K)? = g$$

(for given $g \in X$), we pick a finite-dimensional lss F and a corresponding finite-dimensional lss L of X^* s.t. $\operatorname{LIP}(F, L)$ is correct, hence gives rise to a bounded lprojector P. Consider the **projected equation**:

$$\text{find } f \in F \text{ s.t. } L \perp (1 - K)f - g.$$

Since $L \perp h$ iff $Ph = 0$, this is equivalent to

$$\text{(11)} \qquad \text{find } f \in F \text{ s.t. } P(1-K)f = Pg.$$

Now an adjustment: If $f \in F$, then $Pf = f$, hence we infer from (11) that

$$\text{(12)} \qquad (1 - PK)f = Pg.$$

Conversely, an f satisfying (12) can be written $f = Pg + PKf \in \operatorname{ran} P = F$, hence $Pf = f$, and therefore also $P(1-K)f = Pg$ for such an f. This shows that (11) and (12) are equivalent. It is more convenient to consider (12) since it gives the projected equation in a form rather more close to the original problem.

Now assume that $P = P_n$ with $P_n \xrightarrow{s} 1$. Since X is Bs, this implies that (P_n) is bounded (by (V.7)UBP). Therefore, since K is compact, we conclude from (4)Prop. that $P_n K \to K$, hence $1 - P_n K \to 1 - K$. Since $1 - K$ is boundedly invertible, this makes $(1-K)^{-1}$ an approximate inverse for $1 - P_n K$ for all sufficiently large n, hence implies that, for all n sufficiently large, $1 - P_n K$ is invertible, and $\|(1 - P_n K)^{-1}\| \to \|(1-K)^{-1}\|$.

In other words, if $P = P_n$ with $P_n \xrightarrow{s} 1$, and n sufficiently large, then the projected equation (12) has a unique solution f_P which, in terms of the solution f of the original problem, can be written as

$$f_P = (1 - PK)^{-1} Pg = (1 - PK)^{-1} P(1-K)f.$$

Since $P(1-K) = P - 1 \ + \ 1 - PK$, this can also be written

$$f_P = \Big((1 - PK)^{-1}(P - 1) + 1\Big) f$$

or

$$\text{(13)} \qquad f - f_P = (1 - PK)^{-1}(1 - P)f,$$

and this goes to 0 with $\|f - Pf\|$. Note the resulting error estimate

$$\|f - f_P\| \le \|(1 - PK)^{-1}\| \|f - Pf\| \sim \|(1-K)^{-1}\| \|f - Pf\|$$

which bounds the error $f - f_P$ in the approximate solution f_P in terms of the error in the interpolant Pf.

**** example: second kind Fredholm integral equation ****

A typical example might be the second kind Fredholm integral equation

$$(1-K)f := f - \int_a^b k(\cdot, t) f(t)\,\mathrm{d}t = g,$$

with k a (piecewise) continuous kernel. A typical choice for F might be the space $\Pi^0_{1,\Delta}$ of continuous piecewise linear functions with vertices at the points of the partition $\Delta : a = t_0 < \cdots < t_n = b$ (and nowhere else). Choosing $\Lambda^t : f \mapsto (f(t_i) : i = 0, \ldots, n)$ results in the correct $\mathrm{LIP}(F, \operatorname{ran} \Lambda)$ whose corresponding lprojector P is broken line interpolation at the t_i, hence $\|P\| = 1$ and $P \xrightarrow{s} 1$ on $X := C([a \,..\, b])$ as $|\Delta| := \max \Delta t_i \to 0$. Since $d(f, F) = O(|\Delta|^2)$ for all smooth f (see (IV.11)Example), this provides a second order approximation to the solution of the integral equation, in case Δ is uniform.

Quadrature methods and uniform compactness

**** Nystrom's method ****

Compactness is also of great help in analyzing quadrature (and other) methods for second kind integral equations which are not projection methods.

A standard numerical method for the solution of the second kind equation

$$(14) \qquad (1-K)f := f - \int_a^b k(\cdot,t)f(t)\,\mathrm{d}t = g$$

makes use of quadrature rules

$$\lambda_U := \sum_{u\in U} w(u)\delta_u,$$

with U some finite subset of $[a\,..\,b]$ and $w \in \mathbb{R}^U$ an appropriate weight vector, to approximate the integral. This leads to the (simpler) equation

$$(14_U) \qquad (1-K_U)f_U := f_U - \sum_{u\in U} k(\cdot,u)w(u)f_U(u) = g.$$

On evaluating both sides on U, we obtain the finite square linear system

$$(14_{UU}) \qquad f_U(v) - \sum_{u\in U} k(v,u)w(u)f_U(u) = g(v), \qquad \text{all } v\in U,$$

for the vector $f_{U|U}$, and this system is uniquely solvable if and only if $(1-K_U)$ is invertible. For, once $f_{U|U}$ satisfies (14_{UU}), then f_U, given by

$$f_U := g + \sum_{u\in U} k(\cdot,u)w(u)f_U(u)$$

satisfies (14_U). This is **Nystrom's method**.

**** use better approximate inverses ****

This leaves the question of the invertibility of $1-K_U$. In the case of projection methods, we were able to settle this question of invertibility of $1-PK =: 1-\widetilde{K}$ by showing that $(1-K)^{-1}$ can serve as an approximate inverse for it. Precisely, since $P := P_n \xrightarrow{s} 1$ and K is compact, $\widetilde{K} = PK$ converges uniformly to K, therefore

$$(15) \qquad E = 1-(1-K)^{-1}(1-\widetilde{K}) = (1-K)^{-1}(1-K-1+\widetilde{K}) = (1-K)^{-1}(\widetilde{K}-K)$$

eventually becomes < 1 in norm. By contrast, for quadrature methods, we cannot get uniform convergence to K, even if, as one of course assumes, $\lambda_U \xrightarrow{s} \int\cdot$. The best one can conclude is that, therefore (see below), $K_U \xrightarrow{s} K$. In fact, it can be shown that $\liminf \|K-K_U\| \ge 2\|K\|$, so that uniform convergence is out of the question. There is an ingenious, involved way in the Russian literature (e.g., Krasnoselski et al.) to interpret the quadrature method after all as a projection method and so make use of the earlier analysis. But, there is a very simple direct approach (now associated with the name **Anselone** and the *terminus technicus* **collectively compact**), which started with **Brakhage**'s observation that there might be more suitable approximate inverses available. Specifically, since

$$(1-K)^{-1} = (1-K)^{-1}(1-K\ +\ K) = 1 + (1-K)^{-1}K,$$

he proposed using $1+(1-K)^{-1}\widetilde{K}$ as an approximate inverse for $(1-\widetilde{K})$. One computes

$$\begin{aligned} E = 1-(1+(1-K)^{-1}\widetilde{K})(1-\widetilde{K}) &= (1-K)^{-1}\big(1-K\ \underbrace{-(1-K+\widetilde{K})(1-\widetilde{K})}_{-(1-K)+(1-K)\widetilde{K}-\widetilde{K}(1-\widetilde{K})}\ \big) \\ &= (1-K)^{-1}(\widetilde{K}-K)\widetilde{K}. \end{aligned} \tag{16}$$

This looks just like (15), but with the crucial difference that now $(\widetilde{K}-K)\widetilde{K}$ needs to be small in norm, and this can be deduced from the pointwise convergence $\widetilde{K} \xrightarrow{s} K$ if $\widetilde{K}$ is compact, as is the case. Well, actually, we are in a slightly tricky situation in that the totally bounded set $\widetilde{K}B$ on which we do get uniform convergence of $\widetilde{K}$ to K *changes* with $\widetilde{K}$. This can be handled, though, for our case $\widetilde{K} = K_U$ because it can be shown (see below) that $\mathbf{K} := \{K_U\}$ is **uniformly compact**, i.e., $\mathbf{K}B := \cup_{A\in\mathbf{K}} AB$ is totally bounded. Anselone has coined the term **collectively compact** for this useful property. With this, $\widetilde{K} = K_U \xrightarrow{u} K$ even on $\mathbf{K}B$, hence $\|(K_U-K)K_U\| \ \to\ 0$.

**** uniformly compact strongly convergent perturbations of 1 ****

Once this is recognized, it is possible to prove results concerning uniformly compact approximations K_n to K without explicit reference to an approximate inverse, with the concomitant loss of the *a posteriori* error bound provided by such approximate inverse. Here is a sample:

(17) (Anselone's Theorem). *X Bs, (K_n) in $bL(X)$, uniformly compact, $K_n \xrightarrow{s} K$ compact. Then,*

$$1-K \text{ is boundedly invertible} \quad\iff\quad \limsup{}_n \|(1-K_n)^{-1}\| < \infty.$$

Further, when one or the other of these conditions holds, then $(1-K_n)^{-1} \xrightarrow{s} (1-K)^{-1}$.

Here, having $\limsup \|(1-K_n)^{-1}\|$ finite means that $1-K_n$ is invertible for all sufficiently large n and these inverses are bounded uniformly in n.

I give the *proof* to show that it is really a *standard* argument.

For any compact K on a Bs, having $1-K$ boundedly invertible is equivalent, by (10)FA and (V.18)OMT, to having $1-K$ 1-1. Hence the conclusion of the theorem reads

$$1-K \text{ is 1-1} \quad\iff\quad \liminf{}_{n\to\infty} \inf \|(1-K_n)S\| > 0$$

(with $S := \partial B$ the unit sphere and $\inf \|Z\| := \inf\{\|z\| : z\in Z\}$).

'$\Longrightarrow$' If not $\liminf_n\ \inf\|(1-K_n)S\| > 0$, then $\exists\{(x_n) \text{ in } S\}\ \liminf \|(1-K_n)x_n\| = 0$. AGTASMAT $\lim(1-K_n)x_n = 0$. Since, by the uniform compactness of (K_n), (K_nx_n) lies in a totally bounded set, AGTASMAT $\lim K_n x_n = x_\infty$ for some $x_\infty \in X$, therefore also $\lim x_n = x_\infty$, hence $\|x_\infty\| = 1$. By H.P.(V.3), bounded pointwise convergence is uniform on totally bounded sets while $\{x_n : n \in \mathbb{N}^+\}$ is totally bounded; hence $Kx_\infty = \lim Kx_n = \lim K_n x_n$. This shows that $(1-K)x_\infty = \lim(x_n - K_n x_n) = 0$, i.e., $1-K$ fails to be 1-1.

'$\Longleftarrow$' For every $x \in S$, $\|(1-K)x\| = \lim_n \|(1-K_n)x\|$ while $\|(1-K_n)x\| \ge \inf \|(1-K_n)S\|$, hence $\inf \|(1-K)S\| \ge \liminf \|(1-K_n)S\|$. (This argument only uses the strong convergence.) □

**** back to Nystrom ****

(18) Corollary. *Assume that (14) has only the trivial solution when $g = 0$, and that the "kernel" k is continuous. If $\lambda_U \xrightarrow{s} \int \cdot$, then $\{K_U\}$ is compact uniformly in U, and $K_U \xrightarrow{s} K$, hence (14_U) or (14_{UU}) is uniquely solvable for all sufficiently fine U, and the corresponding solution f_U converges to the unique solution of (14).*

Proof: Let $I := [a \,.\,.\, b]$, and consider, for given $\lambda \in (C([a\,.\,.\,b]))^*$, the map K_λ given on $C([a\,.\,.\,b])$ by the rule

$$K_\lambda f : s \mapsto \lambda(k(s,\cdot)f).$$

For example, $K = K_{\int_a^b \cdot}$, while $K_U = K_{\lambda_U}$.

Since k is continuous, the family $k(I,\cdot) := \{k(s,\cdot) : s \in I\}$ is bounded and equicontinuous, hence totally bounded in $C(I)$, by (II.38)Arzela-Ascoli, hence so is the family $k(I,\cdot)f$ for any fixed $f \in C(I)$. Consequently, $\lambda \xrightarrow{s} \mu$ implies that $\lambda \xrightarrow{u} \mu$ on $k(I,\cdot)f$ for any particular $f \in C(I)$, therefore that $K_\lambda \xrightarrow{s} K_\mu$. In particular, $K_U \xrightarrow{s} K$.

Further, $\|K_\lambda f\| \le \|\lambda\|\|k\|\|f\|$, and

$$|(K_\lambda f)(s) - (K_\lambda f)(t)| \le \|\lambda\|\|k(s,\cdot)f - k(t,\cdot)f\| \le \|\lambda\|\omega_k(|s-t|)\|f\|,$$

showing that $\omega_{K_\lambda f} \le \|\lambda\|\|f\|\omega_k$. By (II.38)Arzela-Ascoli, this implies that $K_\lambda B$ is totally bounded, hence K_λ is compact, uniformly in $\|\lambda\|$. In particular, $K = K_{\int \cdot}$ is compact. Since $\lambda_U \xrightarrow{s} \int \cdot$, hence $\sup_U \|\lambda_U\| < \infty$, this also shows that $\{K_U\}$ is uniformly compact.

Thus, by (17)Anselone's theorem, $(1-K_U)^{-1}$ exists for all sufficiently fine U, and $(1-K_U)^{-1} \xrightarrow{s} (1-K)^{-1}$. □

The hypotheses of this corollary can be relaxed to deal with less smooth kernels, making sure only that the required equicontinuity is preserved.

(19) Example The typical weakening uses the fact that $|Kf(s) - Kf(s')| \le \int |k(s,t) - k(s',t)|\,\mathrm{d}t\, \|f\|_\infty$, hence

$$\omega_{KB} \le \omega_{k,\mathbf{L}_1} : h \mapsto \sup_{|s-s'|<h} \|k(s,\cdot) - k(s',\cdot)\|_1.$$

In particular, $\omega_{KB}(0+) = 0$ (i.e., equicontinuity of KB) is easily proved if k is *Green's function*, hence (see (22) below)

$$k(s,t) = \begin{cases} k_l(s,t), & s < t; \\ k_r(s,t), & s > t, \end{cases}$$

with both k_l and k_r in $C(I \times I)$. In this case,

$$|k(s,t) - k(s',t)| \le \begin{cases} \omega_{k_l}(|s-s'|), & s < s' < t; \\ \|k_l - k_r\|_\infty, & s < t < s'; \\ \omega_{k_r}(|s-s'|), & t < s < s'. \end{cases}$$

Therefore,

$$\|k(s,\cdot) - k(s',\cdot)\|_1 \le |b-a| \max\{\omega_{k_l}(|s-s'|), \omega_{k_r}(|s-s'|)\} + \|k_l - k_r\|_\infty \, |s-s'|,$$

hence

$$\omega_{KB} \le \omega := |b-a| \max\{\omega_{k_l}, \omega_{k_r}\} + \|k_l - k_r\|_\infty ()^1,$$

and evidently $\omega(0+) = 0$.

Further weakening can be achieved because of the fact that the typical quadrature rule is of the form $\lambda_U = \int P_U \cdot$, and P_U need not just be some standard interpolation scheme, nor need it even converge strongly to 1; after all, we only need that $\int P_U \cdot \xrightarrow{s} \int \cdot$.

The iterated projection method

Uniform compactness also helps to explain the success of the iterated projection method for solving the second kind integral equation

$$(1-K)f := f - \int_a^b k(\cdot,t) f(t) \, \mathrm{d}t = g.$$

In this method, the solution f_P of the projected equation $(1-PK)f_P = Pg$ is improved (so one hopes and often finds to be the case) by the standard iteration for second kind equations, i.e., by constructing

$$h := g + Kf_P.$$

Why should h be better than f_P? Well, $h - Kf_P = g$, while $Ph = Pg + PKf_P = f_P$. Therefore, h satisfies the equation

$$h - KPh = g.$$

This looks just like the equation for f_P except that K is approximated by KP rather than by PK (and Pg is replaced by g). At first sight, this doesn't look like an improvement at all, since we already know from H.P.(2) that KP need not converge uniformly to K even if $P := P_n \xrightarrow{s} 1$. But, if $P_n \xrightarrow{s} 1$ and $r := \sup \|P_n\| < \infty$, then $KP_n \xrightarrow{s} K$ and (KP_n) is uniformly compact (since $KP_nB \subseteq KB_r = rKB$), hence (17)Anselone's Theorem gives

$$\limsup \|(1-KP_n)^{-1}\| < \infty \quad \text{and} \quad h \xrightarrow[n\to\infty]{} f.$$

Actually, more is true. Apply the identity (cf. H.P.(III.19))

$$A^{-1} - C^{-1} = A^{-1}(C-A)C^{-1} = C^{-1}(C-A)A^{-1}$$

to the error $f - h$. This gives

$$\begin{aligned} f - h =&(1-K)^{-1}g - (1-KP)^{-1}g \\ =&(1-KP)^{-1}\big(1-KP-(1-K)\big)(1-K)^{-1}g \\ =&(1-KP)^{-1}K(1-P)f. \end{aligned}$$

This looks just like the error formula

(13) $$f - f_P = (1-PK)^{-1}(1-P)f$$

with one important difference: The interpolation error $(1-P)f$ now appears with the map K applied to it. This allows at times the conclusion that the error $f-h$ is of **higher order** than the error $f - f_P$, since $1 - P = (1-P)^2$, hence

$$\|K(1-P)f\| \le \|K(1-P)\|\|(1-P)f\|,$$

and it may happen that $\|K(1-P)\| \ \to\ 0$.

To see why this might be so, assume, specifically, that the underlying nls is $X = C([a\,..\,b])$. Then, any $\lambda : f \mapsto \int_a^b \varphi f$ with $\varphi \in \mathbf{L}_1$ is in X^*, and $\|\lambda\| = \|w\|_1$. Therefore, in this case,

$$\|K(1-P)\| = \sup_t \|\delta_t K(1-P)\| \le \sup_t d(\delta_t K, L)\|1-P\|,$$

with $d(\delta_t K, L) = d_1(\delta_t K, L)$ the distance in the $\mathbf{L}_1$-norm, of the function $k(t,\cdot)$ (which represents $\delta_t K$) from the space of representers of the elements of $L := \operatorname{ran} P'$, and this may be small uniformly in t because the functions $k(t,\cdot)$ may be smooth uniformly in t, hence, if also the elements of L are smooth, in the sense that they are given (or representable) as integration against nice functions, we could expect that $d(\delta_t K, L)$ is small uniformly in t.

**** example: Galerkin's method... ****

Take P to be $\mathbf{L}_2$-approximation from F (hence the projection method in question would be Galerkin's method). Then

$$Pf = \int_a^b p(\cdot, s) f(s)\,\mathrm{d}s,$$

with $p(t,s) := \sum_i u_i(t)\lambda_i(s)$, and (u_i) some basis for F and (λ_i) the corresponding dual basis for $L = F$. Then, taking the underlying nls to be $C([a..b])$, hence using the max-norm to measure function size,

$$d(\delta_t K, L) = \inf_c \|k(t,\cdot) - \sum_i c(i)\lambda_i\|_1,$$

and we would expect this to be small uniformly in t provided k is uniformly smooth and the function family $L = F$ has some approximation power.

**** ...using broken lines ****

Specifically, consider Galerkin's method using $F := \Pi^0_{1,\Delta}$, i.e., continuous broken lines on some partition $\Delta := (t_i)_0^n$ of $[a\,..\,b]$. Then L is (integration against elements of) F, hence, for $\varphi \in C^{(2)}$,

$$d(\int_a^b \varphi\cdot, L) \le |a-b| d_\infty(\varphi, F) \le \text{const}|\Delta|^2 \|D^2\varphi\|_\infty,$$

with $|\Delta| := \max \Delta t_i$. Therefore, if $k \in C^{(2)}([a\,..\,b]^2)$, we get that

$$\|K(1-P)\| = \sup_t d(\delta_t K, L) \ge |a-b| \sup_t d_\infty(k(t,\cdot), \Pi^0_{1,\Delta}) \le \text{const}|\Delta|^2.$$

This implies that $\|f-h\|_\infty = O(|\Delta|^4)$ (in case $f \in C^{(2)}([a\,..\,b])$) while the best we can say for f_P is that $\|f - f_P\|_\infty = O(|\Delta|^2)$.

**** splitting ****

Consider now the more general case of finding $u \in X$ for which

$$Au = g$$

with $A \in bL(X,Y)$, X, Y Bs's, and $g \in Y$ given. Suppose that we seek to solve this equation by projection, seeking

$$u_n \in U_n \text{ s.t. } L_n \perp Au_n - g,$$

with $U_n \subseteq X$, $L_n \subseteq Y^*$. This means that we are considering the $\text{LIP}(AU_n, L_n)$. The standard way to carry out the analysis is to split A into N, with M boundedly invertible and so that the $\text{LIP}(MU_n, L_n)$ is "easily" analyzable. If such an analysis succeeds in showing that, for all sufficiently large n, $\text{LIP}(MU_n, L_n)$ is correct and that the resulting projectors P_n converge boundedly and strongly to 1, then we can conclude the same for the original $\text{LIP}(AU_n, L_n)$ *provided* NM^{-1} is compact. For, we then are, in effect, solving $(1-K)? = g$, with $K := NM^{-1} \in bL(Y)$ compact, by projecting it, i.e., by considering $P_n(1-K)? = P_n g$.

**** example: m-th order linear ODE ****

The m-th order ODE

$$Au := (D^m - \sum_{j<m} a_j D^j)u = g \quad \in C([a\,..\,b]) =: Y \tag{20}$$

is to be solved for $u \in X := C^{(m)}([a\,..\,b]) \cap \ker \mathrm{M}^t$, with $\mathrm{M}^t \in bL(C^{(m-1)}([a\,..\,b]), \mathbb{R}^m)$ 1-1 on $\ker A$ and supported on $\{a, b\}$ (for simplicity). Here, D^m is the proper candidate for the "easy" map M if we assume, for simplicity, that M^t is 1-1 on $\Pi_{<m} = \ker D^m$ (as a map on all of $C^{(m)}([a\,..\,b])$). Then $T := (\mathrm{M}^t|_{\Pi_{<m}})^{-1}\mathrm{M}^t$ is a lprojector (the linear projector given by $\Pi_{<m}$ and $\operatorname{ran} \mathrm{M}$), and

$$u = Tu + \int_a^b G(\cdot,t) D^m u(t)\,\mathrm{d}t, \quad \text{all } u \in C^{(m)}([a\,..\,b]). \tag{21}$$

Here, **Green's function** G is obtained in the form

$$G(\cdot,t) := (1-T)(\cdot - t)_+^{m-1}/(m-1)!, \tag{22}$$

as can be seen by applying $1-T$ to both sides of the Taylor identity

$$u = \sum_{j<m} (D^j u)(a)(\cdot - a)^j/j! + \int_a^b (\cdot - t)_+^{m-1}(D^m u)(t)\,\mathrm{d}t/(m-1)!. \tag{23}$$

H.P.(7) Prove (21), (22).

This allows us to rewrite (20) in terms of $f := D^m u$ as

$$f - \int_a^b k(\cdot,t)f(t)\,\mathrm{d}t = g \qquad \text{with} \ \ k(\cdot,t) := \sum_{j<m} a_j D^j G(\cdot,t).$$

This kernel is only piecewise continuous (if $a_{m-1} \neq 0$), but the continuity is uniform since the pieces all come from the same *finite-dimensional* lss of $C([a\,..\,b])$. This is enough to conclude (cf. (19)Example) that the corresponding map

$$K : C([a\,..\,b]) \to C([a\,..\,b]) : f \mapsto \int_a^b k(\cdot,t)f(t)\,\mathrm{d}t$$

is compact, and the earlier analysis is applicable.

If the side conditions are not homogeneous, it is always possible to modify the problem (by solving for $u-v$ instead of for u, for an appropriate v) so as to make the side conditions homogeneous. If M^t fails to be 1-1 on $\Pi_{<m}$, a slight variant has to be played. In this variant, $X := C^{(m)}([a\,..\,b])$, and one considers

$$C : X \to Y := C([a\,..\,b]) \times \mathbb{R}^m : u \mapsto (Au, \mathrm{M}^t u),$$

and the problem to be solved is to find $u \in X$ so that $Cu = y$ for given $y \in Y$. Now a suitable M is

$$M : X \to Y : u \mapsto (D^m u, u(a), \ldots, (D^{m-1}u)(a))$$

and M^{-1} is given by (23), i.e., by the Taylor identity. The compact map K is again NM^{-1}, with $N := M - C$, of course. Its compactness is seen as before; after all, we have only added something finite-dimensional to the earlier construction.

H.P.(8) Prove the compactness of the map $K = NM^{-1}$ defined in the preceding paragraph.

** superconvergence **

The idea of iterated projection methods does not work so well when solving a differential equation since, in that case, the "kernel" k fails to be smooth. In fact, k (cf. (22)) can be expected to be only piecewise smooth. Yet the same idea does work there (in fact came from there), if looked at properly.

Recall that the approximation u_Q from U to the solution u of the mth order ODE (20) is constructed in the form

$$u_Q = \int_a^b G(\cdot, s) f_Q(s)\,\mathrm{d}s$$

with $f_Q = Qf$ the interpolant to $f := D^m u$ from $F := D^m U$ using the interpolation functionals $L(1-K)$. This means that

$$u - u_Q = \int_a^b G(\cdot, s)(f - f_Q)(s)\,\mathrm{d}s,$$

with $L(1-K) \perp f - f_Q$. Therefore

$$u(t) - u_Q(t) = \int_a^b \big(G(t,s) - p(t,s)\big)(f - f_Q)(s)\,\mathrm{d}s$$

with $p(t,\cdot)$ any representer of an element of $L(1-K)$. Now $G(t,\cdot)$ is only piecewise smooth, precisely, in $C^{(m-2)}$, with a jump discontinuity in its $(m-1)$st derivative across t. This means that it is often possible to choose $p(t,\cdot)$ so that $\|G(t,\cdot) - p(t,\cdot)\| = O(|\Delta|^m)$ when using for U piecewise polynomial functions on some partition Δ. In fact, if t is one of the partition points, then even

$$\|G(t,\cdot) - p(t,\cdot)\|_1 = O(|\Delta|^{m+r})$$

is possible for some $r > 0$. This is called **superconvergence**.

** PDEs are much tougher **

For PDEs, the same arguments are applied sometimes. But it becomes less interesting there since the role of D^m, i.e., of the essential and simple part, is now played by many different operators, and it isn't so simple any more to establish correctness of even the simple LIP or its stability and convergence. Also, superconvergence is harder to achieve since the singularity of Green's function is usually much more subtle than just a jump in some derivative, hence the approximating space is much less likely to model certain sections of Green's function accurately.