V. Baire category and consequences

Pointwise convergence

We explore necessary and sufficient conditions for pointwise convergence of linear maps, particularly in the presence of completeness, i.e., when the domain and/or the target of the maps is a complete nls, i.e., a Banach space (=: Bs).

H.P.(1) Prove: If \( A \in bL(X) \) and \( \dim \ker(A) < \dim X \), then \( \|A - 1\| \geq 1 \). (Hint:(III.7) Riesz’ Lemma)

Conclude that, on an infinite-dimensional nls, the identity cannot be the (norm) limit of finite-rank lm’s.

Definition. The sequence \( (A_n) \) in \( bL(X,Y) \) converges (in norm) to \( A \in bL(X,Y) \):

\[
\lim A_n = A, \quad \text{i.e., } \lim \|A - A_n\| = 0.
\]

This is also written

\[ A_n \rightarrow A \quad \text{(or } A_n \Rightarrow A \text{ or } A_n \rightharpoonup A). \]

Such convergence is also at times termed “uniform” since it allows a uniform error estimate on any bounded set: Since \( \|Ax - A_nx\| \leq \|A - A_n\|\|x\| \), we have

\[
\forall \{x \in B\} \, \|Ax - A_nx\| \leq \|A - A_n\| \rightarrow 0.
\]

But in fact norm convergence is not uniform convergence, it is only uniform convergence on bounded sets; in particular, \( A_n \rightharpoonup A \) on \( B \).

A standard example of a norm-convergent numerical process is fixed point iteration, cf. (III.15).

** norm convergence is rare **

Numerical approximation processes on infinite-dimensional spaces rarely converge in norm (or, ‘uniformly’). For example, let \( (P_n) \) be a sequence of linear projectors of finite rank. Each is the solution of some LIP, and we are looking for convergence of the interpolant \( P_nx \) to \( x \) as \( n \rightarrow \infty \). This means that we would like \( \lim P_n = 1 \).

Note that \( (1 - P_n)x = x \) for all \( x \in \ker(1 - P_n) = \ker P_n \), hence \( \|1 - P_n\| \geq 1 \) unless \( \ker(1 - P_n) = \{0\} \), i.e., unless \( 1 = P_n \). Thus \( \lim P_n = 1 \) implies that \( P_n = 1 \) from some \( n \) on. Since each \( P_n \) is of finite rank, by assumption, this is impossible in case \( \dim X \not< \infty \).

** the next best thing is pointwise convergence **

Definition. The sequence \( (A_n) \) in \( Y^X \) converges strongly, or pointwise, to \( A \in Y^X \):

\[
\forall \{x \in X\} \, \lim A_n x = Ax, \text{ i.e., } \lim \|Ax - A_nx\| = 0.
\]

This is also written

\[ A_n \rightharpoonup A. \]

H.P.(2) Prove that if \( (A_n) \) is in \( bL(X,Y) \), then a pointwise limit \( A \in Y^X \) of \( (A_n) \) is necessarily in \( L(X,Y) \), and \( \|A\| \leq \liminf \|A_n\| \). Hence, if \( (A_n) \) is bounded, then \( A \in bL(X,Y) \).

H.P.(3) Prove that pointwise convergence of a bounded sequence in \( bL(X,Y) \) is uniform on totally bounded sets.
** approximate identity **

A bounded sequence \((A_n)\) in \(bL(X)\) converging pointwise to the identity is called an approximate identity. In Numerical Analysis, approximate identities \((P_n)\), with each \(P_n\) a linear projector of finite rank, are often used to “discretize” an “operator equation”, i.e., an equation \(A\|x = y\) with \(A \in bL(X,Y)\) and \(y \in Y\) given, by “projecting” the equation into the finite-rank equation \(P_nA\|x_n = P_ny\) for which solutions are sought in some lss \(X_n\) of \(X\) with \(\dim X_n = \dim \text{ran} P_n\). Chapter VIII supplies many examples.

H.P.(4) Prove that broken-line interpolation provides an approximate identity for \(C[a..b]\). Specifically, take \(P_nf\) to be the broken line on \([a..b]\) that agrees with \(f\) at its breakpoints \(t_i := a + (b - a)/n,\quad i = 0, \ldots, n\). (Hint: H.P.(II.15).)

** w-convergence, w*-convergence **

Weak convergence and weak*-convergence are both special cases of pointwise convergence. We say that \((x_n)\) converges weakly to \(x\) (and write \(x_n \overset{w}{\to} x\) or \(x_n \to x\)) in case \(\forall \{\lambda \in X^*\} \lim \lambda x_n = \lambda x\). Thus weak convergence in the nls \(X\) is pointwise convergence when we consider \(X\) as a subset of \(X^*\). If \(X\) happens to be \(Y^*\) for some nls \(Y\), then we can also consider pointwise convergence (on \(Y\)). If \(Y\) is reflexive, the same is the weak convergence, but in general it is weaker. For this reason, and as a distinction, pointwise convergence in \(X = Y^*\), i.e., pointwise convergence on \(Y\), is called weak*-convergence, and is denoted by \(x_n \overset{w^*}{\to} x\).

One uses these weaker forms of convergence in order to achieve compactness, i.e., the existence of limit points, when trying to prove the existence of solutions to variational problems by showing that minimizing sequences have limit points. For example, the closed unit ball of any \(X^*\) is weak*-compact (by (IV.4)Alaoglu’s Theorem), but fails to be (norm)compact in case \(X\) is not finite-dimensional (by H.P.(III.16)).

The norm topology on a nls \(X\) is often called the strong topology in distinction to the weak topology on \(X\), i.e., the topology of pointwise convergence on \(X^*\).

H.P.(5) Let \(X = Y^*\) for some nls \(Y\). Prove: Weak convergence is stronger than weak*-convergence, i.e., \(x_n \overset{w}{\to} x \implies x_n \overset{w^*}{\to} x\).

H.P.(6) The weak topology on a nls \(X\) is, by definition, the topology of pointwise convergence ‘on’ \(X^*\). In particular, the nbhdsystem for \(x \in X\) in this topology consists of the balls \(x + B_{r,L}\), with \(r > 0\), \(L\) any finite subset of \(X^*\), and \(B_{r,L} := \{y \in X : \max_{\lambda \in L} |\lambda y| < r\}\). Prove: Any weakly closed subset of a nls space is (norm)-closed. Also, give an example to show that the converse does not hold. (This will require \(\dim X \not< \infty\); also, a convex weakly closed subset is also norm-closed (see H.P.(VI.10))).

H.P.(7) Prove: For \(X = L_p\) or \(\ell_p\) with \(1 \leq p < \infty\), the collection \(B(f) := \{B_{r,\lambda}(f) := \{g \in X : |\lambda (g - f)| < r\}, r > 0, \lambda \in X^*\}\) is equivalent to the nbhd system for the topology of weak convergence.

** bounded pointwise convergence **

(1) Bounded Pointwise Convergence Theorem. If \((A_n)\) is bounded in \(bL(X,Y)\), \(A \in bL(X,Y)\) and \(A_n \overset{s}{\to} A\) on a dense subset \(Z\) of \(X\), then \(A_n \overset{s}{\to} A\).

Proof: Let \(x \in X\).

\[
\forall \{z \in Z\} \|Ax - A_n x\| \leq \|(A - A_n)z\| + \|(A - A_n)(x - z)\| \\
\leq \|(A - A_n)z\| + \left(\|A\| + \sup_{n} \|A_n\|\right)\|x - z\|.
\]
Since $\sup \|A_n\| < \infty$ and $x \in X = Z^-$, one can make the second term small by proper choice of $z$; then, for that $z$, the first term is small for all large $n$ (since $z \in Z$ and $A_n \overset{s}{\rightarrow} A$ on $Z$).

\section*{extension by continuity}

In the formulation of the BPC theorem, we assumed that $A_n \overset{s}{\rightarrow} A$ on a dense subset $Z$ of $X$ for some $A \in bL(X,Y)$. What if we only know that $A \in YZ$? Assuming $Z$ to be a lss of $X$, this implies that $A \in bL(Z,Y)$ (by H.P.(2), since we assumed that $(A_n)$ is bounded), and so the question is merely one of \textit{extending} $A \in bL(Z,Y)$ to some $C \in bL(X,Y)$. Such \textit{extension by continuity} is necessarily unique since, for $x \in X \setminus Z$, $\text{diam}(A(B_{1/n}(x) \cap Z)) \leq \|A\|(2/n) \to 0$ as $n \to \infty$, hence $\cap_n A(B_{1/n}(x) \cap Z)$ can have at most one element, and, by continuity, it would necessarily have to be $Cx$. But such an element need not exist, unless $Y$ is complete. So, assuming $Y$ is $B_s$, the extension of $A$ to all of $X$ is defined and uniquely so.

H.P.(8) Prove: If $Z$ is the completion of the nls $X$, then $Z^* \simeq X^*$, i.e., $Z^*$ and $X^*$ are linearly isometric via the map $\lambda \mapsto \lambda|_X$.

\section*{fundamental sets}

The BPC theorem requires pointwise convergence on some dense set $Z$. Actually, it is sufficient to have pointwise convergence on some \textbf{fundamental set} $F$, i.e., a set $F$ for which $Z := \text{ran}[F]$ is dense. For, if $\lim A_n f = Af$ for all $f \in F$, and $z \in Z$, then $z = \sum_{f \in G} a(f) f$ for some finite $G \subseteq F$ (and some $a \in \mathbb{R}^G$), hence

$$\lim A_n d = \sum a(f) \lim A_n f = \sum a(f) Af = A(\sum a(f)f) = Az.$$ 

To \textit{summarize}: Let $X$ nls, $Y$ $B_s$, $(A_n)$ bounded in $bL(X,Y)$, $A_n \overset{s}{\rightarrow} A$ on $F$, with $F$ fundamental in $X$. Then $A_n \overset{s}{\rightarrow} C$ for some $C \in bL(X,Y)$ (and necessarily $C|_F = A$).

The standard fundamental set for $C(T)$ or $L_p(T)$ with $T \subseteq \mathbb{R}^d$ is the set of monomials. We give a discussion of the basic quantitative aspects of fundamental sets under the heading of \textit{degree of approximation}, for which we need Baire category.

\section*{boundedness, though not necessary, is necessary}

Having $A_n \overset{s}{\rightarrow} A$ on some dense subset $Z$ of $X$ is obviously a necessary condition for strong convergence. But having a bound on $\|A_n\|$ uniformly in $n$ is, in general, too strong a requirement. After all, $\|A_n\|$ is dependent on the norm in $X$ while $A_n \overset{s}{\rightarrow} A$ depends only on the norm in $Y$. In particular, it may well be possible to change the norm in $X$ so that $A$ is still bounded while $A_n$ might not even be bounded, let alone have $(A_n)$ bounded. Here is an example.

Let $X = C^1[0..1]$, $Y = C[0..1]$, both with the sup-norm, and consider the problem of recovering $f \in X$ from its first derivative, $Df$, and its value at 0. We know that

$$f(t) = f(0) + \int_0^t (Df)(s) \, ds.$$ 

\textbf{Copyright} 2002 Carl de Boor
A simple approximation $A_nf$ to $f$ is provided by **Euler’s method**: Pick the uniform partition $(t_i := i/n : i = 0, \ldots, n)$ of $[0 \ldots 1]$ and set

$$(A_nf)(t) := \begin{cases} f(0), & t = 0, \\ (A_nf)(t_i) + Df(t_i)(t - t_i), & t_i < t \leq t_{i+1}. \end{cases}$$

Then $A_nf$ is a continuous broken line. Further,

$$D(f - A_nf)(t) = Df(t) - Df(t_i) \quad \text{for} \quad t_i < t < t_{i+1}, \quad \text{all} \quad i,$$

hence

$$kD(f - A_nf)k_\infty \leq \omega_Df(1/n).$$

Therefore

$$(f - A_nf)(t) = (f - A_nf)(0) + \int_0^t D(f - A_nf)(s) \, ds \leq t\|D(f - A_nf)\|_\infty \leq \omega_Df(1/n).$$

This shows that $\lim k f - A_nf k_\infty = \lim \omega_Df(1/n) = 0$ for all $f \in X = C^{(1)}[0 \ldots 1]$; i.e., $A_n \xrightarrow{s} 1$.

**H.P.** Prove that the linear maps $A_n$ in Euler’s method are bounded wrt to the “standard” norm $\|f\|^{(1)} := \max\{\|f\|_\infty, \|Df\|_\infty\}$ for $C^{(1)}$.

Nevertheless, none of the $A_n$ is even bounded since it is easy to make up a function $f \in X$ with $\|f\| \leq 1$ yet $Df(t_i) > c$, all $i$, for an arbitrary constant $c$, and, for such $f$, $(A_nf)(1) > c - 1$. (Having $\|f\|_\infty \leq 1$ doesn’t constrain $\|Df\|_\infty$ at all.)

Nevertheless, the boundedness of the sequence $(A_n)$ is necessary for pointwise convergence in two ways:

(i) Even if $A_n \xrightarrow{s} A$ can be established, this pointwise convergence cannot be realized numerically unless $\sup_n \|A_n\| < \infty$. For, in the presence of roundoff, we cannot hope to compute the element $A_nf$ exactly. Rather, we construct

$$(A_nf)_{\text{comp}} = A_n(f + r_n)$$

for some “small” error $r_n$. The best we can do is guarantee that $\sup_n \|r_n\| \leq \text{tol}$ for some positive tolerance $\text{tol}$. But then

$$\|Af - (A_nf)_{\text{comp}}\| \leq \|Af - A_nf\| + \|A_n\||\text{tol}$$

is the best estimate we can give for the error in $(A_nf)_{\text{comp}}$ as an approximation to $Af$. By assumption, $\|Af - A_nf\|$ can be made arbitrarily small by choosing $n$ large enough. But, if we do not have $\sup_n \|A_n\| < \infty$, then $(A_nf)_{\text{comp}}$ may have nothing to do at all with $Af$ no matter how large we take $n$ or how small we keep the computing tolerance $\text{tol}$.

(ii) When $X$ is a Bs, i.e., a complete nls, then boundedness of $(A_n)$ is necessary for its pointwise convergence, i.e., $A_n \xrightarrow{s} A \implies \sup \|A_n\| < \infty$. This is a consequence of the uniform boundedness principle, to be proved next. For this, we need a further piece of information about complete metric spaces, viz.
Baire category

This material could have been presented in the earlier discussion of metric spaces.

$X$ is nowhere dense $\iff (Y^-)^\circ = \emptyset$. $C \subseteq X$ is thin (or, meagre, or of first (Baire) category) $\iff C$ is the union of countably many nowhere dense sets. Otherwise, $C$ is called not thin (thick?) (or, not meagre, or of second category). Sometimes, a set is called everywhere dense if its complement is nowhere dense.

(2) Baire Category theorem. A complete metric space is not thin. Equivalently, in a complete ms, the countable intersection of everywhere dense sets is not empty.

Proof: Let $X$ be a complete ms. To show: $X \setminus \bigcup_{n=1}^{\infty} Y_n \neq \emptyset$ whenever all the $Y_n$ are nowhere dense. Might as well go over to the possibly larger sets $Y_n^-$, i.e., may assume that each $Y_n$ is closed, hence has no interior. This means that each $Y_n$ contains no open ball. This means that $\forall \{r > 0, x \in X\} B_r(x) \setminus Y_n$ is open and not empty, hence can find $s > 0$ and $y$ s.t. $B_s^-(y) \subseteq B_r(x) \setminus Y_n$.

So, starting with some $r_0 > 0$ and $x_0$, can choose inductively $r_n \in (0 \ldots 1/n)$ and $x_n$ s.t.

$$B_{r_n}^-(x_n) \subseteq B_{r_{n-1}}(x_{n-1}) \setminus Y_n, \quad n = 1, 2, \ldots$$

The resulting sequence $B_{r_n}^-(x_n)$ of closed sets is decreasing with diameters going to zero, hence, by completeness (see H.P.II.22), $\cap_n B_{r_n}^-(x_n)$ contains some point, necessarily in $X \setminus \bigcap_n Y_n$. \hfill $\square$

Before coming to the major application of Baire Category in this course, I show that all numerically relevant sets are thin, in a discussion of

Degree of approximation

** Weierstraß **

Dense subsets $D$ often come in the form $\cup_n Y_n$ of a union of finite dimensional lss’s $Y_n$ of $X$, and the corresponding fundamental set $F$ is made up of the (columns of) bases of the various $Y_n$. A favorite one for $X = C(T)$ with $T$ compact in $\mathbb{R}^m$ is the collection of monomials

$$\langle t^\alpha : t \mapsto t^\alpha := t(1)^{\alpha(1)} \cdots t(m)^{\alpha(m)},$$

with $\alpha$ a nonnegative integer $m$-vector. The fact that this collection is, indeed, fundamental for $C(T)$ is the content of

(3) Weierstrass’ Approximation Theorem. $\Pi_n = \text{ran}[\langle t^\alpha : |\alpha| \leq n]$, $T$ compact subset of $\mathbb{R}^m$. Then

$$\forall \{f \in C(T)\} \lim_{n \to \infty} d(f, \Pi_n) = 0.$$
**most of a Bs cannot be approximated well at all**

Knowledge of convenient fundamental sets is useful in reducing the labor involved in proving pointwise convergence. But one would have to know much more than that about a fundamental set before its use for the approximation of \( x \in X \) would be defensible. E.g., one would have to know, more precisely, just how to choose, for given \( \varepsilon > 0 \) and given \( x \in X \), a finite \( G \subseteq F \) so that \( d(x, \text{ran}[G]) < \varepsilon \), how big a \( G \) one actually has to choose, and the like.

The next proposition shows that all the elements having some positive degree of approximation from a sequence of proper closed subspaces form a thin set. Thus, by Baire category, most elements of a Bs cannot be approximated well at all.

**(4) Proposition (Harold S. Shapiro).** \((Y_n)\) a sequence of proper closed lss’s in nls \( X \), \((r_n)\) a real sequence converging to 0. Then

\[
A := \{ x \in X : d(x, Y_n) = O(r_n) \}
\]

is thin.

**Proof:**

\[
d(x, Y_n) = O(r_n) \iff \limsup_n d(x, Y_n)/r_n < \infty
\]

\[
\iff \exists\{m, M\} \forall\{n > m\} \ d(x, Y_n) \leq r_nM
\]

\[
\iff \exists\{m, M \in \mathbb{N}\} \ x \in M \bigcap_{n > m} B_{r_n}^{-}(Y_n) =: MA^{(m)}.
\]

Thus

\[
A = \bigcup_{m,M} MA^{(m)}
\]

and we are done once we show that \( A^{(m)} \) is nowhere dense. Since \( A^{(m)} \) is closed, need only show that it has no interior: Suppose that \( B_r(x) \subseteq A^{(m)} \). Then, \( \forall\{n > m\} \ B_r(x) \subseteq B_{r_n}^{-}(Y_n) \), hence \( r \leq r_n \), by (III.10)Corollary to Riesz’ Lemma, therefore \( r = 0 \).

**H.P.(10)** Deduce that totally bounded sets in an infinite-dimensional nls are thin. (Hint: Consider \( Y_n := \text{ran}[V_n] \), with \( V_n \) a finite \((1/n)\)-net for the totally bounded set in question.)

**(5) Corollary.** If \( X \) is complete, then, for any null sequence \((r_n)\), there exists \( x \in X \) so that \( d(x, Y_n) \) goes to zero even slower than does \( r_n \).

**degree of approximation**

These observations lead to a study of the classes

\[
\{ x \in X : d(x, Y_n) = O(r_n) \}
\]

for specific popular choices of \((Y_n)\) and standard sequences \((r_n)\) such as \((n^{-k})\) for some (positive) \( k \). A typical example is provided by the
**Hamel basis for inf.dim. Bs is uncountable**

By (III.10) Corollary to Riesz’ Lemma, a proper closed lss is nowhere dense, hence the countable union of proper closed lss’s is thin. It follows that an infinite-dimensional Bs $X$ cannot be of the form $\bigcup_n \mathrm{ran}\{x_1, \ldots, x_n\}$, and that says that any algebraic or Hamel basis for an infinite-dimensional Bs must be uncountable. This leads to consideration of a basis concept more suitable for infinite-dimensional lss’s in that it should permit one to work with infinite linear combinations.

**Schauder basis**

In a nls, the Schauder basis is the standard choice. Precisely, with $(v_n)$ an infinite sequence in the nls $X$ and $a$ a corresponding sequence of scalars, we define

$$\sum_n v_n a(n) := \lim_{j \to \infty} \sum_{n \leq j} v_n a(n)$$

if it exists. This sets up a lm

$$V : \text{dom} V \subset \mathbb{F}^\mathbb{N} \to X : a \mapsto \sum_n v_n a(n)$$

whose range we call the Schauder basis of $(v_n)$. We call the sequence $(v_n)$ Schauder independent in case $V$ is 1-1, and a Schauder basis (for $X$) in case $V$ is onto as well.

The domain of such a $V$ is, offhand, unclear; it consists of exactly those scalar sequences $a$ for which $\sum_n v_n a(n)$ exists in the above sense. The corresponding sequence $\lambda_i := \delta_i V^{-1}$ of lff’s on $X$ is dual to the sequence $(v_i)$ in the sense that $\lambda_i v_j = \delta_{ij}$.

All practically important Bs’s have Schauder bases, but not all Bs’s do.

**H.P.(11)** Prove:

(i) If $(v_n)$ is a Schauder basis for the nls $X$, then the sequence $(P_n)$ given by

$$P_n x := \sum_{j \leq n} v_j (V^{-1} x)(j), \quad \text{all } x \in X,$$

is an increasing (i.e., both sequences $(\text{ran } P_n : n \in \mathbb{N})$ and $(\text{ran } P'_n : n \in \mathbb{N})$ are increasing) sequence of projectors and converges pointwise to the identity.

(ii) If $(P_n)$ is bounded, hence an approximate identity, and $(v_n)$ is normalized, i.e., $\|v_n\| = 1$, all $n$, then $\sup_n \|\lambda_n\| < \infty$ for the corresponding dual sequence $(\lambda_n = \delta_n V^{-1})$, hence dom $V \subseteq \ell_\infty$.

**H.P.(12)** Try to prove the following converse: If $(Q_n)$ is an increasing approximate identity for $X$ with $r := \sup_n (\text{rank } Q_{n+1} - \text{rank } Q_n) < \infty$, then there exists a Schauder basis $(v_j)$ for $X$ with ran $Q_n = \text{ran}\{v_{j \leq m(n)}\}$ for some strictly increasing sequence $m$. (Hint: Prove first that, with $R_n := Q_{n+1} - Q_n$, $Q_n = R_1 + \cdots + R_{n-1}$ with $R_i R_j = R_j R_i = \delta_{ij} R_i$, hence ran $Q_{n+1} = \text{ran } Q_n + \text{ran } R_n$ and, by (IV.13) Auerbach’s Theorem, $R_n = V_n M_n^t$ with $M_n^t := \Lambda_n^t R_n$ and $V_n$, $\Lambda_n$ normalized dual bases for ran $R_n$ and its dual.)
Uniform boundedness

Recall that a subset of a nls is **bounded** if it is contained in some $B_s$. We call a collection $F \subseteq Y^X$ of maps into the nls $Y$ **pointwise bounded** in case $\forall \{x \in X\} \ F x := \{f x : f \in F\}$ is bounded. Recall further that $f \in Y^X$ is **bounded** if it maps bounded sets to bounded sets, i.e., if $\forall \{r > 0\} \ \exists \{s > 0\} \ f B_r \subseteq B_s$. For this reason, one calls $F \subseteq Y^X$ **uniformly bounded** in case $\forall \{r > 0\} \ \exists \{s > 0\} \ F B_r \subseteq B_s$, i.e., the bound $s$ on $f B_r$ is **uniform** for all $f \in F$.

For $f \in L(X,Y)$, boundedness is equivalent to having $fB \subseteq B_s$ for some $s$, i.e., to having $\|f\| \leq s < \infty$. Therefore, a collection $F$ in $L(X,Y)$ is **uniformly bounded** if it is **bounded** as a subset of $bL(X,Y)$. Uniform boundedness of $F$ does not mean that $FX = \{f x : f \in F, x \in X\}$ is bounded.

(7) **Uniform Boundedness Principle.** For any $B_s X$, a pointwise bounded $A \subseteq bL(X,Y)$ is (uniformly, or norm) bounded. In symbols: ($\forall \{x \in X\} \ Ax \text{ bounded } \implies A$ is bounded).

**Proof:** For $n = 1, 2, \ldots$, $C_n := \{x \in X : Ax \subseteq B_n\} = \cap_{A \in A} A^{-1} B_n$ is closed (as the intersection of closed sets). Further, by the pointwise boundedness assumption, $X = \cup C_n$. By Baire category, not every $C_n$ can be nowhere dense, i.e., $\exists \{n, x, r > 0\} \ B_r(x) \subseteq C_n$, i.e., $A B_r \subseteq A B_r(x) - A x \in B_n + B_n \subseteq B_k$, hence $\|A\| \leq \frac{2n}{r}$.

H.P.(13) Prove the theorem in the following more general form. For this, recall that $p : X \rightarrow \mathbb{R}$ is **lower semicontinuous** $\iff \forall \{a \in \mathbb{R}\} \ p^{-1}(-\infty , a]$ is closed, and that a set $Z$ in a ls is **symmetric** if it contains $-Z$. (Hint: You’ll quickly find that $\exists \{r > 0\} \ \sup A B_r < \infty$. To conclude from this (uniform) boundedness of $A$, you may want to prove first that, for any subadditive $f \ p$, $\max \{p(-x), p(x)\} = \max \{|p(-x)|, |p(x)|\}$ and $\forall \{n \in \mathbb{N}\} \ p(n x) \leq np(x)$.)

(8) **Theorem.** A collection $A$ of lower semicontinuous subadditive functionals pointwise bounded on a symmetric non-thin subset of the nls $X$ is uniformly bounded.

**Proof:** Give an example to show that the symmetry assumption in (8)Theorem is, in general, necessary. (E.g., $X = \mathbb{R}, A = \{(-\alpha )_+ : \alpha > 0\}$.)

H.P.(15) Prove: $0$ is a limit point, in the weak topology, of the sequence $(v_n := \sqrt{n} e_n : n \in \mathbb{N})$ in $\ell_2$, yet no subsequence of this sequence converges weakly to $0$. (See H.P.(7) and prove first that, for $x \in \ell_2$, $\lim \inf_{n \rightarrow \infty} |v(n)|^2/n = 0$.)

** Banach-Steinhaus **

(9) **Corollary (Banach-Steinhaus).** $X B_s, (A_n) \in bL(X,Y)$, and $A_n \xrightarrow{s} A \in Y^X$. Then $\sup \|A_n\| < \infty$.

**Proof:** $\forall \{x \in X\} \ \{\lim A_n x \text{ exists } \implies \sup \|A_n x\| < \infty\}$. Hence $A := \{A_n : n \in \mathbb{N}\}$ is pointwise bounded.

H.P.(16) Prove

(10) **Corollary.** If $(x_n)$ in nls $X$ is w-convergent, or $w^*$-convergent with $X = Y^*$ and $Y B_s$, then $(x_n)$ is (norm)bounded.
Applications of uniform boundedness

** Schauder basis **

H.P.(17)  \( X \) Bs. Prove: \( X \) has Schauder basis \( \iff \) \( X \) has an increasing approximate identity \((Q_n)\) with \( \sup(\text{rank } Q_{n+1} - \text{rank } Q_n) < \infty. \)

** quadrature **

\( X = C(T), \ T \subseteq \mathbb{R}^d \) some compact domain, \( \lambda f := \int_T f(t) \, dt, \lambda_U := \sum_{u \in U} w_U(u) \delta_u \) a rule for \( \lambda, \) based on the nodes \( u \in U \) and the weights \( w_U \in \mathbb{R}^U. \) By H.P.(IV.6), \( \|\lambda_U\| = \|w_U\|_1, \) so

\[
\lambda_U \stackrel{s}{\to} \lambda \iff \lambda_U \stackrel{s}{\to} \lambda \ \text{on some fundamental set } F \text{ and } \sup \|w_U\|_1 < \infty.
\]

This is the downfall of the Newton-Cotes rules which, for the case \( T = [a..b], \) choose \( U = \{a + ih : i = 0, \ldots, n\} \) with \( h := (b - a)/n \) and then choose \( w_U \) by interpolation, i.e., such that \( \lambda_U = \lambda \) on \( \Pi_n. \) For such \( w_U, \|w_U\|_1 \sim 2^n. \)

On the opposite end of the scale are the quadrature rules with nonnegative weights. Now having \( \lambda_U \stackrel{s}{\to} \lambda \) just for the function \( 1 \) (i.e., for the function \( t \mapsto 1 \) ) is enough to get boundedness since it implies that

\[
\|w_U\| = \sum_U w_U(u) = \lambda U 1 \to \lambda 1.
\]

Thus, for such rules, only the convergence on some fundamental set (such as the polynomials) has to be checked. This is ensured by choosing \( \lambda_U \) to be exact on \( Y_U \) for some collection \((Y_U)\) of lss’s whose union is fundamental. Particular examples are the various Gauss rules. Examples of a different sort are provided by the observation (to be made in (VI.20)Proposition) that, for every \( n\)-dimensional lss \( Y \) of \( C(T) \) with \( T \) compact in \( \mathbb{R}^d \) and containing the constant function \( 1, \) there exists \( U \subseteq T \) with \( \#U \leq n \) and corresponding positive weight vector \( w \) so that \( \int_T \cdot = \sum_U w(u) \delta_u \) on \( Y. \)

** equivalence theorems **

Call the collection \((\lambda_U)\) of approximations to \( \lambda := \int \cdot \) consistent if \( \lambda_U \stackrel{s}{\to} \lambda \) (as \( \#U \to \infty) \) on some fundamental set, e.g., all “sufficiently smooth” functions. Call the collection convergent if \( \lambda_U \stackrel{c}{\to} \lambda \) on all of \( X. \) Call the collection stable if \( \sup \|\lambda_U\| < \infty. \) Then we have here a simple instance of an

(11) ** Equivalence Theorem.** consistency \( \implies \) (convergence \( \iff \) stability).

whose most famous instance is the Lax Equivalence Theorem for finite difference approximations to a parabolic PDE (or, more generally, any linear PDE that has associated with it a semigroup of solution operators). This is a prime topic in the course on the numerical solution of evolution equations to which I must refer you because of lack of time.

** divergence **

The uniform boundedness principle is at its best when the sequence \((A_n)\) in \( bL(X, Y) \) fails to be bounded or stable, i.e., when \( \sup \|A_n\| = \infty. \) For it then asserts that the Bs \( X \)
must contain an \( x \) for which \( (A_n x) \) does not converge, more than that, for which \( (A_n x) \) is not even bounded.

A striking example is provided by the **Fourier series**: For this, we pick

\[
X = C := \{ f \in C[0 \ldots 2\pi] : f(0) = f(2\pi) \},
\]

the Bs of continuous \( 2\pi \)-periodic functions. It is convenient here to use the complex scalars \( \mathbb{C} \) rather than the reals; in particular, \( i := \sqrt{-1} \) in this example. The truncated Fourier series is provided by the map

\[
L_n := \sum_{|m| \leq n} [v_m]_m \lambda_m
\]

with

\[
v_m(t) := e^{imt}, \quad \lambda_m f := \int_0^{2\pi} f(t) e^{-imt} \, dt / 2\pi.
\]

This is a projector since

\[
\lambda_m v_k = \int_0^{2\pi} e^{ikt} e^{-imt} \, dt / 2\pi = \int_0^{2\pi} e^{it(k-m)} \, dt / 2\pi = \delta_{mk}.
\]

To analyze it, we observe that

\[
(L_n f)(t) = \sum_{|m| \leq n} e^{imt} \int_0^{2\pi} f(s) e^{-ims} \, ds / 2\pi = \int_0^{2\pi} D_n(t-s) f(s) \, ds / 2\pi
\]

with

\[
D_n(t) := \sum_{|m| \leq n} e^{itm}
\]

the **Dirichlet kernel**. Using the standard formula for summing a finite geometric series, this gives

\[
D_n(t) = \frac{e^{it(n+1)} - e^{-itn}}{e^{it} - 1} = \frac{e^{it(n+1/2)} - e^{-it(n+1/2)}}{e^{it/2} - e^{-it/2}} = \frac{\sin((n + 1/2)t)}{\sin t/2}.
\]

A careful estimate shows that \( \|D_n\|_1 = \int_0^{2\pi} |D_n(t)| \, dt \sim (4/\pi) \ln n \xrightarrow{n \to \infty} \infty. \) Since, for every \( t \), the linear functional

\[
\mu_n : f \mapsto (L_n f)(t)
\]

has norm \( \|\mu_n\| = \|D_n\|_1 / 2\pi, \) we conclude that, for every \( t \), there exists a continuous \( 2\pi \)-periodic function \( f \) so that \( (L_n f)(t) \) is unbounded and, in particular, fails to converge.

On the other hand, by Lebesgue’s inequality,

\[
(f - L_n f)(t) \leq \|1 - L_n\| d(f, \bar{\Pi}_n) \sim (\ln n) d(f, \bar{\Pi}_n)
\]

divergence
with
\[ \Pi_n := \text{ran}[e^{im} : |m| \leq n] =: \text{trigonometric polynomials of degree } \leq n. \]

By Jackson’s theorem,
\[ d(f, \Pi_n) \leq c\omega_f(1/n) \]
for some constant \( c \) independent of \( f \). Therefore \( \|f - L_n f\| \xrightarrow{n \to \infty} 0 \) for every \( f \) satisfying a Dini-Lipschitz condition, i.e., satisfying \( \omega_f(h) = o(\ln h) \).

**Remark.** Since \( D_n \) is real, we conclude that the situation is unchanged if we restrict ourselves to real functions and real scalars.

**H.P.**

Let \( I_n \) be the projector on \( X = C[-1, 1] \) given by \( I_n \) and \( \text{ran}[\delta_u : u \in U] \), with \( U \subseteq [-1, 1] \), \#\( U = n + 1 \). Then \( I_n = \sum_{u \in U}\left|\ell_u\right|\delta_u \) with \( \ell_u(t) := \prod_{w \neq u}(t - w)/(u - w) \).

(i) Prove that \( \|I_n\| = \|\ell\|_\infty \), with \( \ell := \sum_{u \in U}\left|\ell_u\right| \) the Lebesgue function of the process.

(ii) Now choose \( U \) equispaced, i.e., \( U = \{1 - ih : i = 0, \ldots, n\} \), with \( h := 2/n \). Prove that the value of \( \ell \) at \((1 - 1/n)\) is at least \( 2^n/(4n(n - 1/2)) \), showing that \( \|I_n\| \) grows ‘almost like’ \( 2^n \) as \( n \to \infty \). (Hint: \( 2^n = (1 + 1)^n = \sum_{j=0}^n(\binom{n}{j}) \).

(iii) For your information: In fact, for equispaced \( U \), \( \|I_n\| = 2^{n+1}/(2n\ln n) (1 + O(1)) \). On the other hand, if you use the expanded Chebyshev points
\[ U^c = \{\cos \frac{2j + 1}{2n + 2}\pi/\cos \frac{\pi}{2n + 2} : j = 0, \ldots, n\}, \]
you get \( \|I_n^c\| \leq (2/\pi)\ln n + 7 \), which grows so slowly with \( n \) that it doesn’t exceed 3 for \( n \leq 30 \).

**Remark.** The fact that \( \|L_n\|, \|I_n^c\| \xrightarrow{n \to \infty} \infty \) should not be taken too seriously since, in fact, they go to infinity so slowly that this hardly interferes with their effectiveness as good approximation schemes for the practical range of \( n \), say \( n < 1000 \). In addition, even as \( n \) goes to infinity, such a slow rate of growth is easily overcome by the decay to zero of \( d(f, \Pi_n) \) or \( d(f, \Pi_n^c) \) if \( f \) is suitably smooth (e.g., \( f \in C^{(1)} \)).

**Open mapping/closed graph**

There is one further basic f.a. result, also connected with Baire category, namely the Open Mapping theorem, and its corollary, the Closed Graph theorem. Its practical usefulness is less immediate, but it belongs into any basic course on f.a.

The open mapping theorem supplies a (necessary and) sufficient condition for \( A \in bL(X, Y) \) (with \( X, Y \) Bs’s) to be open, i.e., to map open sets to open sets. It is usually applied to \( A \) that is already known to be invertible as a linear map. In that case, having \( A \) open is precisely the same as saying that \( A^{-1} \) is continuous, hence bounded, since it says that the inverse image under \( A^{-1} \) of open sets is open.

**test for openness**

Whether or not \( A \) is invertible,

(12) Lemma. \( A \in L(X, Y) \) is open iff \( 0 \in (AB)^o \) iff \( B \subseteq AB_s \) for some \( s \).

**Proof:** If \( A \) is open, then \( AB \) is open, and, in particular, \( 0 \in (AB)^o \), hence, equivalently, \( B_s \subseteq AB \) for some \( r > 0 \) or, equivalently (with \( s = 1/r \)), \( B \subseteq AB_s \) for some \( s \). Conversely, having \( B \subseteq AB_s \) for some \( s \) implies that \( A \) is open, as follows: If \( O \) is an
open subset of \( X \) and \( y \in AO \), hence \( y = Ax \) for some \( x \in O \), hence \( B_r(x) \subseteq O \) for some \( r > 0 \), then \( y + B_{r/s} \subseteq Ax + AB_s = AB_r(x) \subseteq AO \), showing that \( y \in (AO)^o \). Since \( y \) is arbitrary, it follows that \( AO \) is open.

The open mapping theorem states that \( A \in bL(X,Y) \), with \( X,Y \) Bs’s, is necessarily open if it is onto. Its standard proof contains the following lemma of practical interest.

** almost solvable stably \( \implies \) solvable stably **

(13) Lemma. Let \( A \in bL(X,Y) \), \( X \) Bs. If \( B \subseteq (AB_t)^- \) for some \( t \), then \( B^- \subseteq \cap_{s>t} AB_s \). In particular, \( A \) is onto.

Proof: Since \((AB_t)^-\) is closed, we conclude that \( B^- \subseteq (AB_t)^- \), hence, for any \( y \in Y \), \( y \in B^-_{\|y\|} \subseteq (AB_t\|y\|)^- \). In other words: For some \( t \) and any \( y \in Y \) and any \( \varepsilon > 0 \), we can find an \( x \in B_t\|y\| \) with \( \|y - Ax\| < \varepsilon \). From this, we wish to conclude that, for any \( s > t \) and any \( y \in Y \), we can find an \( x \in B_s\|y\| \) for which \( Ax = y \).

The proof uses the following variant of fixed point iteration. Let \( y \in Y \). With \( d_0 := y \), we pick, for \( n = 1, 2, \ldots \) and entitled by the fact that \( B^- \subseteq (AB_t)^- \), hence \( B_r^- \subseteq (AB_{rt})^- \), an \( x_n \) with \( \|x_n\| < t\|d_{n-1}\| \) so that

\[
d_{n-1} = Ax_n + d_n,
\]

with the norm of the residual \( d_n \) as small as we please. Thus, for any \( n \),

\[
y = Ax_1 + Ax_2 + \cdots + Ax_n + d_n = A(\sum_{j=1}^n x_j) + d_n,
\]

with

\[
\sum_{n} \|x_j\| \leq \sum_{n} \|x_j\| < t \sum_{n} \|d_{j-1}\|.
\]

Hence, choosing, as we may, the \( x_j \) so that \( \sum \|d_{j-1}\| \) is convergent, it follows that \( \lim_{n \to \infty} \|d_n\| = 0 \) and that \((\sum_{j<n} x_j)\) is a Cauchy sequence, hence converges to some \( x \in X \), therefore, \( A \) being bounded, we have \( y = Ax \). Further, \( \|x\| \leq \sum_{j} \|x_j\| < t \sum_{i} \|d_i\| = t(\|y\| + \varepsilon) \), with \( \varepsilon := \sum_{n=1}^\infty \|d_n\| \). Since \( y \) was arbitrary, this shows that \( B^- \subseteq A(B_{(1+\varepsilon)t}) \). However, also \( \varepsilon \) is arbitrary (positive) since we are free to make each \( \|d_n\| \) as small as we please. Therefore,

\[
B^- \subseteq \bigcap_{s>t} AB_s.
\]

(15) Corollary. \( X,Y \) Bs’s, \( A \in bL(X,Y) \) onto. Then \( B^- \subseteq AB_s \) for some \( s \).

Proof: Since \( A \) is onto, \( Y = \bigcup_{n=1}^\infty AB_n \), hence \( AB_n \) is somewhere dense for some \( n \), i.e., \( \exists \{z \in Y, r > 0\} B_r(z) \subseteq (AB_n)^- \). This implies \( B_r \subseteq (B_r(z) + B_r(-z))/2 \subseteq ((AB_n)^- + (AB_n)^-)/2 \subseteq (AB_n)^- \), or \( B \subseteq (AB_t)^- \) with \( t := n/r \), hence \( B^- \subseteq AB_s \) for some \( s \).
(16) Corollary. $X, Y$ Bs’s, $A \in bL(X,Y)$ onto. Then $A \mid : X/\ker A \to Y : \langle x \rangle \mapsto Ax$ is boundedly invertible.

Proof: By construction and assumption, $A \mid$ is invertible. By (III.11), $X/\ker A$ is a Bs with respect to the factor norm $\|\langle x \rangle\| := d(x, \ker A)$. With that, by (15)Corollary, there is an $s$ so that, for arbitrary $x \in X$, there is $x'$ with $Ax' = Ax$, i.e., $x' \in \langle x \rangle$, and $\|x'\| < s \|Ax\|$. But this says that $s\|A\langle x \rangle\| = s\|Ax\| > \|x'\| \geq d(x', \ker A) = \|\langle x \rangle\|$, showing $A \mid$ to be bounded below, hence $A^{-1} \mid$ is bounded.

(17) Corollary. $X, Y$ Bs’s, $A \in bL(X,Y)$, ran $A$ closed. Then $(\ker A)^\perp = \text{ran } A^*$. In particular, ran $A^*$ is closed.

Proof: By H.P.(IV.14), ran $A^*$ is contained in the closed lss $(\ker A)^\perp$. Hence it is sufficient to prove that $(\ker A)^\perp \subseteq \text{ran } A^*$. For this, observe that the im $A \mid : X/\ker A \to \text{ran } A : \langle x \rangle \mapsto Ax$ is boundedly invertible by (16)Corollary since ran $A$ is a closed lss of the Bs $Y$, hence a Bs. Now take $\lambda \perp \ker A$. Then (by (I.3)) $\lambda = \mu(\langle x \rangle)$ for $\mu : \langle x \rangle \mapsto \lambda x$, hence $\|\mu\| := \sup_x |\lambda x|/\|\langle x \rangle\| = \sup_x |\lambda x|/d(x, \ker A) \leq \sup_x |\lambda x|/d(x, \ker \lambda) = \|\lambda\|$, i.e., $\mu \in (X/\ker A)^*$, hence $\lambda = \mu(A \mid)^{-1}A \in \text{ran } A^*$. 

\[ \begin{array}{ccc} \lambda & A \\ \mathbb{F} & \leftarrow & X \\ \mu \downarrow & \downarrow & \rightarrow & A \mid \\ X/\ker A \end{array} \]

** example: an ordinary differential equation **

It follows that if the equation $Ax = y$ is uniquely solvable for every $y \in Y$, then the solution $x$ depends continuously on the datum $y$. A typical example involves the linear map

$$A : C^{(m)}[0 \ldots 1] \to C[0 \ldots 1] \times \mathbb{R}^m : f \mapsto (Lf, \Lambda^t f)$$

with

$$L := D^m + \sum_{j \leq m} a_j D^j, \quad a_j \in C[0 \ldots 1], \text{ all } j,$$

and $\Lambda^t = [\lambda_1, \ldots, \lambda_m]^t$ 1-1 on ker $L$ and made up of ifl’s continuous on $C^{(m)}[0 \ldots 1]$. Since dim ker $L = m$, this implies that $A$ is 1-1, and, assuming that the $\lambda_j$ are already continuous over $C^{(m-1)}$, it also implies that $A$ is onto. Since (see below) $C^{(m)}[0 \ldots 1]$ is complete, we conclude that the unique solution $f \in C^{(m)}[0 \ldots 1]$ of the linear BVP (:= boundary value problem)

$$\langle L, \Lambda^t \rangle? = (g, c)$$

depends continuously on the given $(g, c) \in C[0 \ldots 1] \times \mathbb{R}^m$, i.e.,

$$\exists \{M\} \forall \{(g, c) \in C[0 \ldots 1] \times \mathbb{R}^m\} \|f\|^{(m)} \leq M \max\{|g|, |c|\infty\}.$$

In other words, the BVP is stable.

This example is a bit of a sham since the proof that such $A$ is onto is usually given via Green’s function, in which case the continuous dependence of the solution on the data is explicit.
** open mapping theorem **

(18) Open Mapping Theorem. $X, Y$ Bs's, $A \in bL(X, Y)$ onto $\implies A$ is open.

Proof: From (15)Corollary, we know that $B \subset AB_s$ for some $s$, and, by (12), that is equivalent to $A$ being open. \hfill \Box

H.P.(19) Where in the proof of the OMT is the continuity of $A$ used?

** illustrations **

(19) Corollary. If $X$ is Bs and $A \in bL(X, Y)$ has finite rank, then $A$ maps open sets to sets relatively open in ran $A$.

Proof: Since $\dim \text{ran } A < \infty$, ran $A$ is Bs, hence $A| : X \to \text{ran } A : x \mapsto Ax$ is open. \hfill \Box

(20) Corollary. Every nontrivial bounded lfl on a Bs is an open map.

(21) Corollary. If $Y$ is a closed lss of Bs $X$, then $\langle \rangle : X \to X/Y$ is open.

(22) Corollary. $X, Y$ Bs's, $A \in bL(X, Y)$, 1-1, onto $\implies A^{-1} \in bL(Y, X)$.

H.P.(20) Suppose $A \in bL(X, Y)$, with $X, Y$ Bs's. Prove: ran $A$ is closed and ker $A = \{0\} \iff A$ is bounded below.

** equivalence of norms **

For example, suppose that $X$ is Bs with respect to two norms, $\| \cdot \|$ and $\| \cdot \|'$.

Suppose further that $1 : X \to X : x \mapsto x$ is bounded as a map from $(X, \| \cdot \|)$ to $(X, \| \cdot \|')$. This says that $\sup \|x\|'/\|x\| < \infty$, i.e.,

$$\exists \{M\} \forall \{x \in X\} \|x\|' \leq M\|x\|.$$

Then (22)Corollary implies that $1$ is also a bounded map from $(X, \| \cdot \|')$ to $(X, \| \cdot \|)$, i.e.,

$$\exists \{m\} \forall \{x \in X\} \|x\| \leq m\|x\|'.$$

Thus, the equivalence of two complete norm topologies can be checked by merely checking whether one is stronger than the other.

** completeness is necessary here **

The fact that $X$ must be complete wrt both norms is crucial here. For example, $X := C^{(1)}[0,1]$ is nls wrt the max norm $\| \cdot \|\infty$, but also wrt the norm $\| \cdot \|'$ given by

$$\|f\|' := \max\{|f(0)|, \|Df\|\infty\}.$$ 

Further,

$$\|f\|\infty \leq 2\|f\|'.$$

But a bound the other way is impossible since $\sup \|Df\|\infty/\|f\|\infty = \infty$. Since $(C^{(1)}, \| \cdot \|')$ is a Banach space (cf. below), this also shows that $(C^{(1)}, \| \cdot \|\infty)$ cannot be complete.

closed graphs \hfill ©2002 Carl de Boor
** closed graphs **

Other more or less useful consequences of the open mapping theorem, hence ultimately of the Baire category theorem, are connected with the cartesian product of nls’s. You should verify that the cartesian product \( X \times Y \) of two nls’s \( X \) and \( Y \) is indeed a linear space with addition and scalar multiplication taken pointwise, i.e.,

\[
(x_1, y_1) + (x_2, y_2) := (x_1 + x_2, y_1 + y_2) \\
\alpha(x, y) := (\alpha x, \alpha y).
\]

(In fact, the cartesian product \( \times_{t \in T} X_t \) of an ‘assignment’ \((X_t : t \in T)\) of nls’s \( X_t \) with an arbitrary index set \( T \) is naturally a ls wrto pointwise vector operations.) Further, \( X \times Y \) is nls with respect to the norm

\[
\|(x, y)\| := \|(x\|, \|y\|)\|
\]

with \(| \cdot |\) any norm on \( \mathbb{R}^2 \), and is complete if both \( X \) and \( Y \) are. Usually, \( \|(x, y)\| := \max(\|x\|, \|y\|) \). (In the case of an arbitrary Cartesian product, its subspace \( \{ x \in \times_{t \in T} X_t : \|t \to \|x(t)\|_{X_t}\| < \infty \} \), with \( \| \cdot \| \) any (extended) norm on \( \mathbb{R}^T \), is a nls, and is complete in case each \( X_t \) is complete and, e.g., we use the max-norm on \( \mathbb{R}^T \).

H.P.(21) Prove the following Corollary to (15)Corollary: X Bs, \( X = Y + Z \), with \( Y, Z \) closed lss’s. Then \( \exists \{\text{const} \} \ \forall \{x\} \ \exists \{(y, z) \in Y \times Z\} \ x = y + z \) and \( \text{const}\|x\| \geq \|y\| + \|z\| \).

H.P.(22) Use H.P.(21) to prove: X Bs, \( P \in L(X) \), \( P^2 = P \) (i.e., \( P \) is l.projector). Then \( P \in blL(X) \iff 1 P, \text{ran} \ P \) are closed.

(23) Closed Graph Theorem. X, Y Bs’s, \( A \in L(X, Y) \). A closed \( \implies A \) bounded.

Here, \( A \in L(X, Y) \) is called closed if it is closed as a subset of the nls \( X \times Y \). (This may be confusing since, a little while ago, we called a lm open for quite a different reason.)

Now a curiosity: The graph of \( A \in Y^X \) is, by definition, the subset \( \Gamma(A) \) of \( X \times Y \) defined by

\[
\Gamma(A) := \{(x, y) \in X \times Y : y = Ax\}.
\]

This notation is customary but, of course, superfluous since, what is a map if it isn’t what is called here its graph? So, forget about the whole thing, but get used to the idea that \( A \in Y^X \) is a subset of \( X \times Y \).

As a training in this way of thinking, you should verify that, for \( A \in L(X, Y) \), \( A \) is a lss of \( X \times Y \).

What does it mean for \( A \) to be closed? It means that if \( (x, y) = \lim(x_n, y_n) \) (i.e., \( x = \lim x_n, \ y = \lim y_n \)) for some sequence \( ((x_n, y_n)) \) in \( A \) (i.e., \( x_n \in X, \ y_n = Ax_n, \) all \( n \)), then \( (x, y) \in A \) (i.e., \( y = Ax \)).

For example, if \( A \in blL(X, Y) \), then \( A \) is closed since then \( Ax = A(\lim x_n) = \lim Ax_n = \lim y_n = y \). The closed graph theorem is a kind of converse.

**Proof of the closed graph theorem.** Since \( A \) is closed and linear, it is a closed linear subspace of the Bs \( X \times Y \), hence itself a Bs. The map \( A \to X : (x, Ax) \mapsto x \) is linear, onto, and 1-1 since \( A \) is a map, and is bounded since \( \|x\| \leq \max\{\|x\|, \|Ax\|\} = \|(x, Ax)\| \).
By the open mapping theorem, its inverse, i.e., the map $x \mapsto (x, Ax)$, is therefore also bounded, i.e.,

$$\infty > \sup_x \| (x, Ax) \|/\|x\| = \sup_x \max \{ 1, \| Ax \|/\|x\| \} = \max \{ 1, \| A \| \}.$$

\[\square\]

**nls’s of smooth functions**

A clean way to think of functions with a certain number of derivatives is as just that, i.e., as a collection or vector of functions related by differential operators. This point of view is evident when one looks at just how such a space of differentiable functions is normed and how its dual is constructed (cf. H.P.(24) below). As an illustration, here is the simplest possible example, the ls $X = C^{(1)}[0, 1]$. I’ll drop the reference to the interval $[0, 1]$; in effect any closed interval would do.

The standard definition is

$$C^{(1)} := \{ f \in C : Df \in C \}.$$

But already the standard norm

$$\| f \|^{(1)}_{\infty} := \max \{ \| f \|_{\infty}, \| Df \|_{\infty} \}$$

isn’t just the norm of one function. Rather, it is the norm of the pair $(f, Df)$ as an element of $C \times C$. In effect, $C^{(1)} = D$, as we now make clear.

We note in passing that, earlier, we used the norm

$$\| f \|' := \max \{ \| f(0) \|, \| Df \|_{\infty} \}$$

but this is equivalent to $\| \cdot \|^{(1)}_{\infty}$ since $f = f(0) + \int_0^1 (s - u)_{+} Df(u) \, du$ for any $f \in C^{(1)}[0, 1]$, hence, for any such $f$, $\| f \|' \leq \| f \|^{(1)}_{\infty} \leq 2\| f \|'$.

We now show that $C^{(1)}$ is a Bs (hence the equivalence of these two norms already follows from one of the inequalities just mentioned). We do this by thinking of $C^{(1)}$ as $D$, i.e., as a linear subspace of the Bs $C \times C$, hence require nothing more than that $D$ be closed. This latter fact is most easily proved by considering the Volterra operator or map

$$V : C \to C : f \mapsto \int_0^t f(s) \, ds = \int_0^1 (s - u)_{+} f(u) \, du.$$

Since $Vf(t) - Vf(u) = \int_u^t f(s) \, ds \leq |t - u|\|f\|_{\infty}$, it follows that ran $V \subset C$ and that $V$ is bounded; in fact, $\| V \| = 1$. (More than that, $V(B)$ is totally bounded since $\omega_{Vf}(h) \leq h\|f\|_{\infty}$.) Further, $DV = 1$, hence $VD$ is a projector; in particular, $f = f(0) + VDf$ for all $f \in C^{(1)}$.

This says that $C^{(1)} = \{ (f, g) \in C \times C : g = Df \} = (\Pi_0, 0)_{+} V^{-1}$, with

$$V^{-1} = \{ (Vf, f) : f \in C \}$$

closed since $V$ is bounded, hence closed. But this implies that $C^{(1)} = D$ is itself closed, as the sum of a closed and a finite-dimensional lss (cf. H.P.(III.13)).

nls’s of smooth functions

©2002 Carl de Boor
H.P.(23) Show that $1 - V$ does not take on its norm (as a map on $C$).

Here is the same argument in more conventional terms.

To show that $C^{(1)}$ is closed as a subspace of $C \times C$, consider $(f_n, Df_n) \sim N_{n-\infty} (f, g)$. Then $f_n = f_n(0) + V Df_n$ and $f - f(0) = \lim (f_n - f_n(0)) = \lim V Df_n = \lim V g$ (using the continuity of $V$). But this says that $f = f(0) + V g$, hence $g = Df$, i.e., $(f, g) \in C^{(1)}$.

In just the same way, the space $C^{(m)}$ is identified with

$$\{(f_r)_0^m \in C^{m+1} : f_r = Df_{r-1}, \ r = 1, \ldots, m\}$$

and is shown to be closed (as the intersection of closed sets), hence complete. In particular, the norms

$$\|f\|^{(m)} := \max \{\|D^r f\|_\infty : r = 0, \ldots, m\}$$

and

$$\|f\| := \max \{|f(0)|, \ldots, |D^{m-1} f(0)|, \|D^m f\|_\infty\}$$

are equivalent.

In the same way, the Sobolev space $W^{(m)}_p(G)$ of all functions on some domain $G \subseteq \mathbb{R}^d$ with all partial derivatives of order $\leq m$ in $\mathbf{L}_p(G)$ is identified with a closed linear subspace of $(\mathbf{L}_p(G)^N, \|\cdot\|_p^N)$. In this generality, the definition of $D$ is “weak”, i.e., $(f, g) \in \mathbf{L}_p^2$ in $D_y$ iff $\int_G \varphi g = -f \int_D \varphi f$ for all “test functions” $\varphi$. This raises some questions concerning $f, g$ “on” $\partial G$ in case $G$ is not all of $\mathbb{R}^d$.

Also, it is slightly more work to prove that the standard norm

$$\|f\| := \|(\|D^j f\|_p(G) : j = 0, \ldots, m)\|_p$$

on $W^{(m)}_p(G)$, with

$$\|D^j f\|_p(G) := \|(\|D^\alpha f\|_p(G) : |\alpha| = j\|_p,$$

is equivalent to the norm

$$f \mapsto \|P f\| + \|D^m f\|_p(G),$$

with $P$ any linear projector onto the (finite-dimensional) kernel of $f \mapsto \|D^m f\|_p(G)$, and an arbitrary (fixed) norm taken on that kernel.

However, if $G$ is a bounded, open, connected domain with Lipschitz boundary, then that kernel can be shown to be $\Pi_{<m}(G)$, i.e., the space of polynomials in $d$ variables of degree $< m$, restricted to $G$. Consequently, with $P$ any of the many available bounded linear projectors on $W^{(m)}_p(G)$ (e.g., $\text{ran} P' = \Pi_{<m}(G)$ would do), one obtains the

**Bramble-Hilbert Lemma.** If $L$ is any bounded linear map, from $X := W^{(m)}_p(G)$ to some $\text{nls}$ $Y$ (with $G$ bounded, open, connected, with Lipschitz boundary), and $\Pi_{<m} \subseteq \ker L$, then,

$$\|Lf\| \leq \|L\| \text{const}_G \|D^m f\|_p(G), \quad f \in X.$$  

Indeed, $\|Lf\| \leq \|L\|d(f, \ker L) \leq \|L\|d(f, \Pi_{<m})$, while $\text{dist}(f, \Pi_{<m}) \leq \|f - Pf\| \leq \text{const}_G (\|P(f - Pf)\| + \|D^m f - Pf\|_p(G)) = \text{const}_G \|D^m f\|_p(G)$, with the second inequality by the norm-equivalence (and const$_G$ the corresponding constant).
H.P.(24) Prove that every \( \lambda \in (C(m)[0..1])^* \) has a representation of the form

\[
\lambda : f \mapsto \sum_{j=0}^{m} \int_{0}^{1} f_{j}(s) \, dx_j(s),
\]

with \( x := (x_j : j = 0, \ldots, m) \in (NBV[0..1])^{m+1} \).

** example: an ordinary differential equation initial value problem **

As a further illustration, consider the \( m \)-th order ODE Initial Value Problem in which we seek \( f \in C^{(m)}[0..1] \) that satisfies

\[
(D^m f)(t) = F(t, f(t), Df(t), \ldots, (D^{m-1} f)(t)) \quad \text{for } 0 \leq t \leq 1
\]

together with the initial conditions

\[
(D^j f)(0) = c(j), \quad j = 0, \ldots, m - 1.
\]

In terms of the vector-valued function \( f := (D^j f)^{m-1}_{j=0} =: (f_j)^{m-1}_{j=0} \), the problem is to find \( f \in C^{(1)}([0..1] \to \mathbb{R}^m) \) that satisfies

\[
Df(t) = \tilde{F}(t, f(t)) \quad \text{for } 0 \leq t \leq 1,
\]

\[
f(0) = c,
\]

with \( \tilde{F} : [0..1] \times \mathbb{R}^m \) defined by

\[
\tilde{F}(t, x)(j) := \begin{cases} 
  x(j+1), & j < m; \\
  F(t, x(1), \ldots, x(m)), & j = m.
\end{cases}
\]

If \( \mathbb{R}^m \to \mathbb{R}^m : x \mapsto \tilde{F}(t, x) \) is Lipschitz continuous uniformly for \( t \in [0..1] \), then Picard iteration (see Chapter II) provides a proof for the existence of a solution for this first-order system, hence for the equivalent original \( m \)th order equation.