These are classnotes for Math/CS 887, Spring '98 by Carl de Boor
corrections are welcome!

Approximation Theory overview

One considers special instances of the following

Problem. *Given some element f in a metric space X (with metric d), find a best approximation (=:**ba**) m^* to f from some given subset M of X, i.e., find*

$$m^* \in M \ \text{ s.t. } d(f, m^*) = \inf_{m \in M} d(f, m) =: \operatorname{dist}(f, M).$$

Abbreviation:

$$m^* \in \mathcal{P}_M(f).$$

Basic questions

- Existence: $\#\mathcal{P}_M(f) \geq 1$?
- Uniqueness: $\#\mathcal{P}_M(f) \leq 1$? More generally, $\#\mathcal{P}_M(f) =$?
- Characterization: how would one recognize a ba (other than by comparing it with all other candidates)? This is important for
- Construction:

The metric

The metric is almost always a **norm metric**, i.e., $d(x, y) := \|x - y\|$, and the set M is usually a finite-dimensional linear subspace. But, as the problem of approximating a curve shows, there are important practical instances in which linearity plays no role and there is no suitable norm in which to pose the problem.

curve approximation problem X is the set of 'smooth' closed curves, of finite length, say, in $\mathbb{R}^2$; it is a metric space with the **Hausdorff metric**

$$d(A, B) := \max\{\operatorname{dist}(A, B), \operatorname{dist}(B, A)\},$$

with

$$\operatorname{dist}(A, B) := \sup_{a \in A} \inf_{b \in B} \|a - b\|_2.$$

M is the set of ellipses in $\mathbb{R}^2$, say (or some other class of 'simple' curves).

Specific choices for normed X

(T is an interval or a suitable subset of $\mathbb{R}^d$)

- $\mathbf{L}_2(T)$, i.e., $\|x\| := \|x\|_2 := \left(\int_T |x(t)|^2\right)^{1/2} \mathrm{d}t$ **Least-squares**.
  More generally: inner product spaces.
- $C(T)$, i.e., $\|x\| := \|x\|_\infty := \sup_{t\in T} |x(t)|$ **uniform** (or, **Chebyshev**).
- $\mathbf{L}_1(T)$, i.e., $\|x\| := \|x\|_1 := \int_T |x(t)|\, \mathrm{d}t$ **Least-mean**.
- (weighted) $\mathbf{L}_p$ ...

Specific choices for M

Usually, M is a finite-dimensional linear subspace, i.e., of the form

$$M = \mathrm{ran}[f_1, \ldots, f_n] := \{\sum_{j=1}^n f_j\, a(j) : a(j) \in \mathbb{F}\}$$

with $\mathbb{F}$ either $\mathbb{R}$ or $\mathbb{C}$.

- $\Pi_k := \mathrm{ran}[()^j : j = 0, \ldots, k]$ (algebraic) polynomials of degree $\le k$. Here, $()^j : t \mapsto t^j$. More generally,
  $$\Pi_k = \Pi_k(\mathbb{R}^d) := \mathrm{ran}[()^\alpha : |\alpha| \le k]$$
  with $()^\alpha : \mathbb{R}^d \to \mathbb{F} : x \mapsto x^\alpha := x(1)^{\alpha(1)} \cdots x(d)^{\alpha(d)}$, $\alpha \in \mathbb{Z}_+^d$, and $|\alpha| := \|\alpha\|_1 = \sum_i \alpha(i)$. Also, for $\alpha \in \mathbb{Z}_+^d$, $\Pi_\alpha := \Pi_{\le\alpha} := \mathrm{ran}[()^\beta : \beta \le \alpha]$.
- $\mathcal{R}_{k,h} := \Pi_k/\Pi_h := \{p/q : p \in \Pi_k, q \in \Pi_h\}$ rational functions of degrees k, h.
- $\overset{\circ}{\Pi}_k := \mathrm{ran}[1, \sin(\nu\cdot), \cos(\nu\cdot) : \nu = 1, \ldots k]$ trigonometric polynomials of degree k. The natural domain for trigonometric polynomials is the circle, i.e., the interval $[0 \,..\, 2\pi]$ or $[-\pi \,..\, \pi]$ with the endpoints identified. Customary notation for this set:
  $$\mathbb{T} = \mathbb{T}^1.$$
  If also complex scalars are admitted, we get the simpler description
  $$\overset{\circ}{\Pi}_k = \mathrm{ran}[e_{\mathrm{i}j} : |j| \le k],$$
  with $\mathrm{i} := \sqrt{-1}$, and
  $$e_\theta : x \mapsto \exp(\theta\cdot x)$$
  the **exponential with frequency** θ, a definition which even makes sense for $x \in \mathbb{F}^d$ with θ also in $\mathbb{F}^d$ and
  $$\theta\cdot x := \sum_i \theta(i)x(i)$$
  the ordinary scalar product.
- $\mathrm{Exp}_\Theta := \mathrm{ran}[e_\theta : \theta \in \Theta]$ exponentials with frequencies $\theta \in \Theta$. We already noted that $\overset{\circ}{\Pi}_k(\mathbb{R}) \subseteq \mathrm{Exp}_{\mathrm{i}\{-k,\ldots,k\}}$. Also, for any $\Theta \subset \mathbb{R}$, $\Pi_k = \lim_{h\to 0} \mathrm{Exp}_{h\Theta}$ whenever $\#\Theta = k + 1$.
- $\Pi^\rho_{k,\boldsymbol{\xi}} :=$ piecewise polynomials (=: **pp**) in $C^{(\rho)}$, of degree $\le k$, with **break sequence** $\boldsymbol{\xi} = (\cdots < \xi_j < \xi_{j+1} < \cdots)$. (**splines**)

The first and third are linear. The second is nonlinear. The last two are linear or nonlinear depending on whether the frequencies Θ (resp. the break sequence $\boldsymbol{\xi}$) is fixed or variable.

Degree of approximation

considers the behavior of $h \mapsto \mathrm{dist}(f, M_h)$ as a function of the (discrete or continuous) parameter h. E.g., $k \mapsto \mathrm{dist}(f, \Pi_k)$ as $k \to \infty$, or $h \mapsto \mathrm{dist}(f, \Pi^\rho_{k,h\mathbf{t}})$ as $h \to 0$. Usually, one considers only

$$\mathrm{dist}(K, M_h) := \sup_{f\in K} \mathrm{dist}(f, M_h)$$

with K a class of functions sharing with the particular f of interest certain characteristics (e.g., all functions whose 14th derivatives are no bigger than 7 in absolute value). Only the behavior in the limit, as $h \to 0$ or $h \to \infty$ or whatever, is usually considered. If nothing is said, then $h \to 0$.

Jackson type theorems: $f \in K \quad\Longrightarrow\quad \mathrm{dist}(f, M_h) = O(h^\alpha)$

Bernstein type (or, inverse) theorems: $\mathrm{dist}(f, M_h) = O(h^\alpha) \quad\Longrightarrow\quad f \in K$.

Saturation theorems: $\mathrm{dist}(f, M_h) = o(h^\alpha) \quad\Longrightarrow\quad f \in K_0$ (for some appropriate α and with K_0 some very 'small' set).

Typically, $K := \{f \in X : \|f\|' \le 1\}$ for some stronger norm $\|\cdot\|'$. This leads to consideration of the **K-functional**

$$K_f : t \mapsto \inf_g (\|f - g\| + t\|g\|').$$

Related question: Is M a good choice for approximating f, given that we know that $f \in K$? Typical criterion involves the dimension of M, i.e., the degrees of freedom to be used. If $\dim M = n$, then one compares $\mathrm{dist}(K, M)$ with

$$d_n(K) := \inf_{\dim Y \le n} \mathrm{dist}(K, Y),$$

the **n-width** of K (in the sense of Kolmogorov). While it is not easy to find **optimal** subspaces, one can often find a 'ladder' (M_n) for which $\mathrm{dist}(K, M_n) \sim O(d_n(K))$ and $\dim M_n \sim n$.

Good approximation

In practice, best approximation is rarely used. Instead, one looks for (cheap but good) approximation schemes. E.g., if A is a linear map into M, then $\|(1 - A)_{|K}\| := \sup_{f\in K} \|f - Af\|$ may well be close to $\mathrm{dist}(K, M)$.

Special case: a **near-best** A is one for which, for some const and for all f,

$$\|f - Af\| \le \mathrm{const}\, \mathrm{dist}(f, M).$$

Any such A is necessarily a projector (onto M). Conversely, if A is a bounded linear projector onto M, then, for any f, and any $m \in M$, $\|f - Af\| = \|(1 - A)(f - m)\| \le \|1 - A\|\|f - m\|$, therefore

$$\|f - Af\| \le \|1 - A\|\, \mathrm{dist}(f, M). \tag{1}$$

course

The intent is to give a quick reading of these basics of AT, illustrated with the help of splines, thereby giving also a quick introduction to (univariate) splines.

Weierstraß, Korovkin, Lebesgue, Bernstein

Start off the course the way Lorentz starts off his book (the nicest book in classical AT) and the way Tikhomirov starts off his survey of AT, namely with

(2) Weierstraß (1885). *For any (finite) interval* $I = [a \mathinner{..} b]$, $\Pi|_I$ *is dense in* $C(I)$.

Since Π is invariant under **translation**

$$f \mapsto f(\cdot + t)$$

and **dilation**

$$f \mapsto f(\cdot/\delta),$$

it is sufficient to consider the special case $I = [0 \mathinner{..} 1]$.

We will give three proofs (at least). The first proof is Bernstein's, but done with Korovkin's theorem. The second is Lebesgue's, done with broken lines. The third is Stone's.

(3) Korovkin (1957). *T compact, (U_n) in $L(C(T))$, U_n positive, $U_n \xrightarrow{pw} 1$ on some finite set G. If there exists $(a_g : g \in G)$ in $C(T)$ so that*

$$p(t,s) := \sum_{g\in G} g(t)a_g(s) \ge 0 \quad \text{with equality iff } t = s,$$

then $U_n \xrightarrow{pw} 1$.

Explanations:

(i) For the U_n and U all maps on the same domain and into the same target, $U_n \xrightarrow{pw} U$ indicates that U_n converges **pointwise** to U, i.e., $\forall\{x \in X\} \quad \lim_{n\to\infty} U_n x = Ux$ (in the topology of the common target of the U_n and U). This is much weaker than the more elusive **uniform** convergence, denoted $U_n \xrightarrow{u} U$, which presupposes that the common target is normed and means that

$$\lim_{n\to\infty} \sup_{x\in X} \|U_n x - Ux\| = 0.$$

(ii) It is assumed here that $C(T)$ is the set of *real-valued* continuous maps on T. For those, there is a **natural (partial) order**, namely

$$f \le g :\Longleftrightarrow \forall t \in T \; f(t) \le g(t).$$

$U : C(T) \to C(T)$ is called **positive** (or, more precisely, **nonnegative**) if

$$0 \le f \quad \Longrightarrow \quad 0 \le Uf.$$

Observation: Assume that U is positive and linear. Then $f \le g \implies Uf \le Ug$. Further, with $|f| : T \to \mathbb{R} : t \mapsto |f(t)|$,

$$-|f| \le f \le |f| \quad \implies \quad -U(|f|) \le Uf \le U(|f|),$$

hence,

$$|U(f)| \le U(|f|).$$

Proof of Korovkin (3): By assumption, $U_n \xrightarrow{pw} 1$ on G, hence converges pointwise to 1 on

$$\mathrm{ran}[G] := \{\sum_{g \in G} gc(g) : c \in \mathbb{R}^G\} =: \mathrm{span}\, G.$$

The latter is finite-dimensional, therefore $U_n \to 1$ *uniformly on bounded subsets of* $\mathrm{ran}[G]$.

The rest is a very nice trick, in which the arbitrary $f \in C(T)$ to be approximated from Π is locally related to some element of $\mathrm{ran}[G]$, as follows. From the assumption, $\mathrm{ran}[G]$ contains (strictly) positive functions, e.g., the function $t \mapsto p(t,s) + p(t,s')$ for any $s \ne s'$. Let p^* be one such. For $s \in T$, set

$$f =: \frac{f(s)}{p^*(s)} p^* + h(\cdot, s).$$

Then

$$(U_n f)(s) = \frac{f(s)}{p^*(s)} (U_n p^*)(s) + (U_n h(\cdot, s))(s), \tag{4}$$

and

$$U_n p^* \to p^*.$$

Since T is compact, $\|1/p^*\|_\infty < \infty$, hence

$$\frac{f(s)}{p^*(s)} (U_n p^*)(s) \to \frac{f(s)}{p^*(s)} p^*(s) = f(s)$$

uniformly in s. Korovkin's result therefore follows from the following claim. □

Claim. $U_n(h(\cdot, s))(s) \to 0$ *uniformly in* s.

Proof: By the positivity of U_n, $|U_n h(\cdot, s)| \le U_n(|h(\cdot, s)|)$. Take any $\varepsilon > 0$. Then $|h| \le \varepsilon +$ a bound for $|h|$ on the set $\Delta_\varepsilon := \{(t,s) : |h(t,s)| \ge \varepsilon\}$. This set is closed since h is continuous, hence compact; it also does not contain the zero-set of p, i.e., the set $\{(t,t) : t \in T\}$, since h vanishes there. Therefore,

$$\delta := \inf p(\Delta_\varepsilon) \; > 0,$$

hence, $|h| \le (\|h\|_\infty / \delta) p$ on Δ_ε. So,

$$|h| \le \varepsilon + (\|h\|_\infty / \delta) p.$$

Consequently, for any s,

$$|U_n(h(\cdot,s))| \le U_n(|h(\cdot,s)|) \le \varepsilon M + (\|h\|_\infty/\delta) U_n(p(\cdot,s)),$$

with

$$M := \sup_n \|U_n()^0\|_\infty$$

finite since, by the strict positivity of p^*, there is some positive b so that $()^0 \le bp^*$, therefore

$$|U_n()^0| \le U_n()^0 \le bU_n(p^*) \to bp^*.$$

Now, $\{p(\cdot,s) : s \in T\}$ is a bounded subset of the finite-dimensional linear space $\mathrm{ran}[G]$, hence, on it, U_n converges to 1 *uniformly.* Since $p(s,s) = 0$, this means that, for large n, $U_n(p(\cdot,s))(s)$ is close to zero uniformly in s. Therefore, for all $n \ge n_\varepsilon$,

$$|U_n(h(\cdot,s))(s)| \le \varepsilon(M+1).$$

Since ε was arbitrary, this proves the Claim. □

From this, (2)Weierstraß for $I = [0\,..\,1]$ follows, with the choices $G = \{()^0, ()^1, ()^2\}$, $p : (t,s) \mapsto (t-s)^2$, and $U_n = B_n$, $n > 0$, the **Bernstein operator** (1912, introduced by Bernstein for a different proof of Weierstraß)

$$B_n : f \mapsto \sum_{j=0}^n \beta_{j,n-j} f(\frac{j}{n}),$$

with

$$\beta_{j,n-j} : t \mapsto \binom{n}{j} t^j (1-t)^{n-j}.$$

Indeed, B_n is linear, and is positive as a map on $C(I)$. Further,

$$DB_n f = \sum_{j=1}^n \binom{n}{j} j()^{j-1}(1-\cdot)^{n-j} f(\frac{j}{n}) - \sum_{j=0}^{n-1} \binom{n}{j} (n-j)()^j (1-\cdot)^{n-j-1} f(\frac{j}{n}),$$

therefore (and this is of independent interest)

$$DB_n f = n \sum_{j=0}^{n-1} \binom{n-1}{j} ()^j (1-\cdot)^{n-1-j} \Delta f(\frac{j}{n}), \tag{5}$$

with

$$\Delta f(\frac{j}{n}) := f(\frac{j+1}{n}) - f(\frac{j}{n}).$$

Since $\Delta^{k+1} f(\frac{j}{n}) := \Delta(\Delta^k f(\frac{j}{n}))$ vanishes identically for any $f \in \Pi_k$, it follows that

$$B_n(\Pi_k) \subseteq \Pi_k, \quad k \le n.$$

Since also

$$B_n f(t) = f(t), \quad t = 0, 1,$$

it follows that $B_n = 1$ on Π_1, therefore $B_n()^j \to ()^j$ for $j = 0, 1$ trivially. Also,

$$f : t \mapsto t(t-1)$$

vanishes at 0 and 1 and is quadratic, hence $B_n f$ must have the same properties, and therefore must equal $\alpha_n f$, with $\alpha_n := D^2 B_n f/2$. By (5), $D^2 B_n f/2 = n(n-1)\Delta^2 f/2$, while $\Delta^2 f(\frac{j}{n}) = \Delta^2()^2(\frac{j}{n}) = ((\frac{1}{n})^2 + (-\frac{1}{n})^2) = \frac{2}{n^2}$, therefore $\alpha_n = \frac{n-1}{n}$, i.e., $B_n f = f - f/n$, hence $B_n()^2 = ()^2 - f/n$. Therefore

$$\|B_n()^2 - ()^2\|_\infty \le \|f\|_\infty/n \longrightarrow 0.$$

□

(6) d-dim. Weierstraß. *The restriction $\Pi|_T$ of the polynomials in d arguments to any compact subset T of $\mathbb{R}^d$ is dense in $C(T)$.*

Proof: It is sufficient to prove the theorem for the special case $T = I := [0\,..\,1]^d$ since T is compact, hence contained in some axi-parallel box, which, after translation and scaling, we may assume to be I, and, by Tietze's extension theorem, $C(T)$ can be isometrically imbedded into $C(I)$.

Remember: The **Tietze** extension of $f \in C(T)$ to an $f_I \in C(I)$ is given by the rule

$$f_I : x \mapsto \begin{cases} f(x), & x \in T; \\ \alpha + \inf_{t\in T}(f(t) - \alpha)d(x,t)/d(x,T), & x \notin T, \end{cases}$$

with $\alpha := \inf f(T) - 1$, say, i.e., so that $f - \alpha()^0$ is strictly positive. A proof is usually given on the way to proving Urisohn's Lemma.

Now choose U_n as the **tensor product** $B_n \otimes \cdots \otimes B_n$ of d copies of the Bernstein operator, i.e.,

$$U_n f : x \mapsto \sum_{k\in\{0,\ldots,n\}^d} f(\frac{k}{n}) \prod_{i=1}^d \beta_{k(i),n-k(i)}(x(i)), \quad x \in [0\,..\,1]^d.$$

From the univariate argument, $U_n \xrightarrow{pw} 1$ on $\Pi_{2,\ldots,2}$ since $U_n()^\alpha = \otimes_{i=1}^d B_n()^{\alpha(i)} : x \mapsto \prod_i B_n()^{\alpha(i)}(x(i))$, and a suitable p is

$$p : (x,y) \mapsto \sum_i (x(i) - y(i))^2.$$

Korovkin also supplies a proof that the trigonometric polynomials are dense in $C(\mathbb{T})$, the space of continuous functions on the circle. In this case, the role of p is played by

$$p : (t,s) \mapsto 1 - \cos(t-s) = (e_0 - (e_{\mathrm{i}} + e_{-\mathrm{i}})/2)(t-s),$$

i.e., $\operatorname{ran}[G] = \overset{\circ}{\Pi}_1$, and the maps U_n are the **Fejér** operators, i.e.,

$$\sigma_n : f \mapsto \frac{1}{\pi} \int_{\mathbb{T}} F_n(\cdot - t) f(t)\, \mathrm{d}t$$

with

$$F_n(\theta) := \frac{1}{2(n+1)} \left(\frac{\sin((n+1)\theta/2)}{\sin(\theta/2)} \right)^2 .$$

The Fejér operator associates with f the average of its truncated Fourier series of orders $0, 1, \ldots, n$, i.e.,

$$\sigma_n f = \frac{1}{n+1} \sum_{j=0}^{n} s_j f$$

with

$$s_j f := \sum_{|\mu| \le j} e_{\mathrm{i}\mu} \int_{\mathbb{T}} \mathrm{e}^{-\mathrm{i}\mu t} f(t)\, \mathrm{d}t/(2\pi).$$

In particular, σ_n is a positive operator (which s_n is not), and

$$\sigma_n(e_{\mathrm{i}k}) = \frac{(n - |k| + 1)_+}{n+1} e_{\mathrm{i}k},$$

hence $\sigma_n(e_{\mathrm{i}k})$ converges to $e_{\mathrm{i}k}$ for any k, in particular for $|k| \le 1$. On the other hand, recall that $\|s_j\| \sim \ln j$ as $j \to \infty$, hence s_j fails to converge to 1 on $C(\mathbb{T})$.

Lebesgue's proof (1898) of (2)Weierstraß Let $I := [a \,..\, b]$. For any $f \in C(I)$ and any finite sequence $\boldsymbol{\xi} := (a = \xi_1 < \xi_2 < \cdots < \xi_{\ell+1} = b)$, the broken-line interpolant to f is, by definition, the unique element $P_{\boldsymbol{\xi}} f$ in $\Pi^0_{1,\boldsymbol{\xi}}$ which agrees with f at $\boldsymbol{\xi}$. For $\xi_j \le t \le \xi_{j+1}$,

$$f(t) - P_{\boldsymbol{\xi}} f(t) = (f(t) - f(\xi_j)) \frac{\xi_{j+1} - t}{\xi_{j+1} - \xi_j} + (f(t) - f(\xi_{j+1})) \frac{t - \xi_j}{\xi_{j+1} - \xi_j},$$

hence

$$\|f - P_{\boldsymbol{\xi}} f\|_\infty \le \omega_f(|\boldsymbol{\xi}|),$$

with ω_f the *(uniform) modulus of continuity of* f and

$$|\boldsymbol{\xi}| := \max_j |\xi_{j+1} - \xi_j|.$$

Conclusion: The collection

$$\Pi^0_1(I) := \bigcup_{\boldsymbol{\xi} \text{ in } I} \Pi^0_{1,\boldsymbol{\xi}}$$

of continuous broken lines on I is dense in $C(I)$.

Since each $\Pi^0_{1,\boldsymbol{\xi}}$ is contained in $\Pi_1 + \operatorname{ran}[|\cdot - \xi_j| : j = 2, \ldots, \ell]$, the following Claim therefore finishes the proof. □

(7) Claim. *On each bounded* $[a\,..\,b]$ *and for any* s, $|\cdot - s|$ *can be approximated uniformly by polynomials.*

Proof: Only the case $s \in (a..b)$ needs proof. For such s, we may choose $\sigma \in \Pi_1 \backslash \Pi_0$ that carries $[a\,..\,b]$ *into* $[-1\,..\,1]$ in such a way that $\sigma(s) = 0$. Further, if $p \in \Pi$ is close to $|\cdot|$ on $[-1\,..\,1]$, then $p \circ \sigma$ is a polynomial that, on $[a\,..\,b]$, is that close to $c|\cdot - s|$ for some nonzero c that depends only on σ. Hence it is sufficient to prove the Claim for $[a\,..\,b] = [-1\,..\,1]$ and $s = 0$. Since $|t| = (t^2)^{1/2}$, i.e.,

$$|\cdot| = \sqrt{\cdot} \circ ()^2,$$

and $()^2([-1\,..\,1]) = [0\,..\,1]$, and $\Pi \circ ()^2 \subset \Pi$, the following Claim finishes the proof. □

(8) Claim. *The square-root function,* $\sqrt{\cdot}$, *can be uniformly approximated on* $[0\,..\,1]$ *by polynomials.*

Proof: A standard proof uses the Taylor series expansion

$$\sqrt{1 - \cdot} = 1 - \sum_{n>0} ()^n \prod_{k=1}^{n} |3/2 - k|/k$$

wich converges uniformly on any compact subset of $(-1\,..\,1)$ (since the power coefficients are all bounded by 1). I prefer the following proof, from Dieudonné's Analysis.

Define the sequence $(u_n : n = 0, 1, 2 \cdots)$ of polynomials by the iteration

$$u_{n+1}(t) := u_n(t) + (t - u_n(t)^2)/2, \quad n = 0, 1, 2, \ldots$$

with

$$u_0 := 0.$$

Claim: for all n,

$$u_n(t) \le \sqrt{t}, \quad t \in [0\,..\,1].$$

Indeed, assuming this already to be true for n (as it is for $n = 0$), we compute

$$\sqrt{t} - u_{n+1}(t) = (\sqrt{t} - u_n(t))(1 - (\sqrt{t} + u_n(t))/2) \ge (\sqrt{t} - u_n(t))(1 - \sqrt{t}) \ge 0$$

for $t \in [0\,..\,1]$.

It follows that $t - u_n(t)^2 \ge 0$ on $[0..1]$, hence (u_n) is monotone increasing, yet bounded on $[0\,..\,1]$, therefore pointwise convergent there, and its limit is necessarily a fixed point of the iteration used to define it, hence its limit is $\sqrt{\cdot}$. But since this limit function is continuous and $[0\,..\,1]$ is compact, $u_n \to \sqrt{\cdot}$ uniformly (by Dini's Theorem). □

Stone-Weierstraß

For an arbitrary set T, the collection $\mathbb{F}^T$ of all scalar-valued maps (whether real or complex) is not only closed under (pointwise) addition and multiplication by a scalar, but also closed under (pointwise) multiplication of two elements, i.e., for $f, g \in \mathbb{F}^T$, also

$$fg : T \to \mathbb{F} : t \mapsto f(t)g(t)$$

is in $\mathbb{F}^T$. This makes $\mathbb{F}^T$ a **ring with identity**. Since also

$$f(\alpha g) = (\alpha f)g = \alpha(fg), \quad \alpha \in \mathbb{F}, f, g \in \mathbb{F}^T,$$

$\mathbb{F}^T$ is an **algebra with identity**. Being an algebra, $\mathbb{F}^T$ is closed under formation of polynomials; i.e.,

$$\Pi \times \mathbb{F}^T \to \mathbb{F}^T : (p, f) \mapsto p \circ f$$

(with $p \circ f$ **map composition**, i.e., $p \circ f : T \to \mathbb{F} : t \mapsto p(f(t))$) is well-defined. Now assume that T is compact Hausdorff. Then $C(T)$ is a closed subalgebra of the space $b(T \to \mathbb{F})$ of bounded functions on T. Further, as the composition of continuous maps, $p \circ f \in C(T)$ if $f \in C(T)$.

The Stone(-Weierstraß) theorem employs the following notion: $A \subset \mathbb{F}^T$ is said to **separate points** if, for any two distinct points $s, t \in T$, there exists $a \in A$ with $a(s) \neq a(t)$. If A is a linear subspace and also contains the identity, then this is equivalent to the statement that the linear map

$$A \to \mathbb{F}^2 : a \mapsto (a(s), a(t))$$

is onto (since its range contains the vector $(1, 1)$ as well as some vector (α, β) with $\alpha \neq \beta$), hence the map has right inverses. In particular, for any $f \in \mathbb{F}^T$ and any $s, t \in T$, there exists $a_{f,s,t} \in A$ which agrees with f at s and t. It is this conclusion that is needed.

(9) Stone(-Weierstraß) (1937). *Let T be compact metric. The only closed subalgebra A of* real *$C(T)$ with identity which separates points is $C(T)$ itself.*

Proof: The range $a(T)$ of any $a \in A$ is a bounded subset of $\mathbb{R}$ (since T is compact), hence, by Claim 7, $|\cdot|$ can be approximated, uniformly on $a(T)$, by polynomials p, and A, being an algebra, contains $p \circ a$ for any polynomial. This implies that A, being closed, is closed under formation of absolute values; i.e.,

$$a \in A \quad \Longrightarrow \quad |a| \in A.$$

Since

$$\max\{f, g\} : T \to \mathbb{R} : t \mapsto \max\{f(t), g(t)\}$$

can also be written

$$\max\{f, g\} = ((f + g) + |f - g|)/2,$$

A is also closed under formation of the maximum of *finitely* many functions. Since

$$\min\{f,g\} = -\max\{-f,-g\},$$

A is also closed under formation of minimum.

Take $f \in C(T)$, $\varepsilon > 0$. Let $t \in T$. For each $s \in T$, there is $a_{s,t} \in A$ agreeing with f at s and t, hence, there is some neighborhood U_s of s on which $f - \varepsilon < a_{s,t}$. T being compact, there exists a *finite* set S for which $T = \cup_{s\in S} U_s$. Hence, the function

$$a_t := \max_{s\in S} a_{s,t}$$

is in A and satisfies

$$f - \varepsilon < a_t \quad \text{on } T, \qquad a_t(t) = f(t).$$

The latter implies that, on some neighborhood V_t of t, $f + \varepsilon > a_t$. Hence, with S a finite set for which $T = \cup_{t\in S} V_t$, the function

$$a := \min_{t\in S}\ a_t$$

is in A and satisfies

$$f + \varepsilon > a$$

as well as $f - \varepsilon < a$. Consequently, $\|f - a\|_\infty < \varepsilon$. As ε is arbitrary and A is closed, $f \in A$ follows. □

From this, even the multivariate Weierstraß theorem, (6), follows since Π is an algebra with identity which separates points.

Another consequence is the density of the *even* trigonometric polynomials, i.e., of $\operatorname{ran}[\cos(j\cdot) : j = 0, 1, 2, \ldots]$ in $C([0\,..\,\pi]$ (the function cos alone is enough to separate points). Note that this algebra fails to be dense in $C([0\,..\,\alpha])$ for any $\alpha > \pi$.

The restriction to *real*-valued functions is essential here. Not only does the proof of Stone make explicit use of it (it relies on the total ordering of the reals), but the polynomials fail to be dense in $C(\{z \in \mathbb{C} : |z| \le 1\})$ of all complex-valued continuous functions of one complex argument on the **unit disk** (even though they continue to form there an algebra with identity that separates points), since their closure consists of functions *analytic* in the interior of the disk.

On the other hand, Stone's theorem does have the following (weaker) immediate consequence for complex-valued functions which relies on nothing more than the fact that $C_{\mathbb{C}}(T) = C_{\mathbb{R}}(T) + \mathrm{i}C_{\mathbb{R}}(T)$.

(10) complex Stone. *Let T be compact metric. The only closed subalgebra A of* complex *$C(T)$ with identity which separates points and is closed under conjugation is $C(T)$ itself.*

Here, being closed under conjugation means that A contains, with f, also its conjugate, $\overline{f} : T \to \mathbb{C} : t \mapsto \overline{f(t)}$.

In particular, $\Pi|_D$ is dense in $C_{\mathbb{C}}(D)$ for $D = \{x \in \mathbb{R}^2 : \|x\|_2 \le 1\}$.

near-best approximation

While the questions concerning *best* approximations raised in the first lecture (such as existence, uniqueness and characterization) occupy a good part of a standard course in Approximation Theory, best approximations are hardly ever calculated because they can usually only be obtained as the limit of a sequence of approximations, and because of the easy availability of near-best approximations. To recall, the bounded linear map A on the nls X provides near-best approximations from the subset M if

$$\|f - Af\| \le \mathrm{const}\,\mathrm{dist}(f, M), \quad f \in X.$$

Such A is, necessarily, a linear projector onto M, i.e., $\mathrm{ran}\, A = M$ and $A^2 = A$, i.e., $A|_M = 1$. In particular, M is a linear subspace.

Conversely, if A is a bounded linear projector on X with range M, then $f - Af = (1 - A)f = (1 - A)(f - m)$ for all $m \in M$, hence

$$\mathrm{dist}(f, M) \le \|f - Af\| \le \|1 - A\|\,\mathrm{dist}(f, M), \quad f \in X. \tag{1}$$

This is Lebesgue's inequality.

This raises (at least) two questions:

(i) How small can one make $\|1 - A\|$ by proper choice of A?

(ii) What can be said about $\mathrm{dist}(f, M)$ in terms of some information we might have about f?

Example: broken line interpolation With $X = C[a\,.\,.\,b]$ for some finite interval $[a\,.\,.\,b]$, recall the broken line interpolant $A = P_{\boldsymbol{\xi}} f$ for given $\boldsymbol{\xi} = (a = \xi_1 < \cdots < \xi_{l+1} = b)$. Its range is the space

$$M = \Pi^0_{1,\boldsymbol{\xi}}$$

of broken lines with break sequence $\boldsymbol{\xi}$. Since $\|P_{\boldsymbol{\xi}}\| = 1$, and always $\|1 - A\| \le 1 + \|A\|$, we get in this case

$$\mathrm{dist}(f, \Pi^0_{1,\boldsymbol{\xi}}) \le \|f - P_{\boldsymbol{\xi}} f\|_\infty \le 2\,\mathrm{dist}(f, \Pi^0_{1,\boldsymbol{\xi}}).$$

In particular, the construction of a ba from $\Pi^0_{1,\boldsymbol{\xi}}$ will, at best, cut the maximum error in half.

Incidentally, the bound used here, $\|1 - P_{\boldsymbol{\xi}}\| \le 1 + \|P_{\boldsymbol{\xi}}\|$, is sharp as can be seen by looking at an f which is the broken line with breakpoints $\xi_1, \ldots, \xi_{l+1}$ at which it has the value 1 and one additional breakpoint at which it has the value -1.

In particular, we can get a good estimate for $\mathrm{dist}(f, \Pi^0_{1,\boldsymbol{\xi}})$ by looking at $\|f - P_{\boldsymbol{\xi}} f\|_\infty$. In Lebesgue's proof of Weierstraß, we already observed tnat

$$\|f - P_{\boldsymbol{\xi}} f\|_\infty \le \omega_f(|\boldsymbol{\xi}|).$$

In particular, if $f \in L^{(1)}_\infty[a\,.\,.\,b]$, i.e., f is absolutely continuous with essentially bounded first derivative, then $\omega_f(h) \le h\|Df\|_\infty$, hence

$$\|f - P_{\boldsymbol{\xi}} f\|_\infty \le |\boldsymbol{\xi}|\|Df\|_\infty, \quad f \in L^{(1)}_\infty[a\,.\,.\,b].$$

Further, since $P_{\boldsymbol{\xi}}$ reproduces all elements of $\Pi^0_{1,\boldsymbol{\xi}}$, we can replace here f by an arbitrary element of $\Pi^0_{1,\boldsymbol{\xi}}$ and so obtain, more precisely,

$$\|f - P_{\boldsymbol{\xi}} f\|_\infty \le |\boldsymbol{\xi}| \operatorname{dist}(Df, \Pi_{0,\boldsymbol{\xi}}), \quad f \in L^{(1)}_\infty[a\,..\,b].$$

Since $\operatorname{dist}(g, \Pi_{0,\boldsymbol{\xi}}) \le \omega_g(|\boldsymbol{\xi}|)$, we therefore obtain

$$\|f - P_{\boldsymbol{\xi}} f\|_\infty \le |\boldsymbol{\xi}| \omega_{Df}(|\boldsymbol{\xi}|), \quad f \in C^{(1)}[a\,..\,b],$$

and, so, finally,

$$\|f - P_{\boldsymbol{\xi}} f\|_\infty \le |\boldsymbol{\xi}|^2 \|D^2 f\|_\infty, \quad f \in L^{(2)}_\infty[a\,..\,b].$$

Actually, since, for $\xi_j \le t \le \xi_{j+1}$, $f(t) - P_{\boldsymbol{\xi}} f(t) = (t - \xi_j)(t - \xi_{j+1})[\xi_j, \xi_{j+1}, t]f$ (with $[\alpha, \ldots, \omega]f$ the divided difference of f at the point sequence $(\alpha, \ldots, \omega)$), we have the sharper estimate

$$\|f - P_{\boldsymbol{\xi}} f\|_\infty \le |\boldsymbol{\xi}|^2 \|D^2 f\|_\infty / 8, \quad f \in L^{(2)}_\infty[a\,..\,b].$$

You see how, as a function of the mesh size, $|\boldsymbol{\xi}|$, these estimates improve, i.e., go to zero faster as $|\boldsymbol{\xi}| \to 0$, when we restrict f to smoother and smoother function classes. However, further more restrictive smoothness assumptions will not lead to an increase in the rate at which the interpolation error goes to zero with $|\boldsymbol{\xi}|$.

This is a simple illustration of our next topic, degree of approximation.

If there is time, I may come back to item (i), i.e., the question of how small one can make $\|A\|$ or $\|1 - A\|$ for given M by appropriate choice of the linear projector A onto M. To whet your appetite, I mention that every linear projector on $C[a\,..\,b]$ onto Π_n has norm no smaller than $\sim \ln n$. This is related to the fact that the projector s_n providing the truncated Fourier series (mentioned earlier) is the unique projector of minimal norm on $C(\mathbb{T})$ onto $\mathring{\Pi}_n$, and $\|s_n\| \sim (2/\pi) \ln n + \text{const}$.

Degree of Approximation

Given a *sequence* $(M_n) = (M_n : n \in \mathbb{N})$ of approximating sets in the nls X and $f \in X$, one is interested in

$$n \mapsto E_n(f) := \operatorname{dist}(f, M_n)$$

as $n \to \infty$.

In this generality, nothing can be said. The following general model (from Chapter 7 of DeVore and Lorentz) covers most situations of practical, and even most situations of theoretical, interest.

(i) *$M_1 = \{0\}$, and (M_n) is increasing,* i.e., (M_n) is a **ladder**.

This guarantees that $E_n(f)$ is nonincreasing, with $E_1(f) = \|f\|$.

(ii) *$\cup_n M_n$ is dense in X.*

This is equivalent to having $\lim_n E_n(f) = 0$ for all $f \in X$.

In view of approximation by rationals, or exponentials with free frequencies, or splines with free knots, it would be too much to assume that the M_n are *linear* subspaces. But it is ok to assume

(iii) $\mathbb{F}M_n := \{\alpha m : \alpha \in \mathbb{F}, m \in M_n\} \subseteq M_n$.

(iv) $\exists\{c\}\forall\{n\}\ M_n + M_n \subset M_{cn}$.

Finally, although clearly not essential, the following assumption will avoid a certain amount of epsilontics.

(v) *Each M_n is an existence set.*

In particular, each M_n is closed.

To be sure, Assumption (iv) is somewhat restrictive and precludes some practically important ladders such as the following:

$$M_n := \Pi_{k,n}$$

consists of all piecewise polynomial functions, on $[-1\,..\,1]^2$ say, of some degree k and involving no more than n polynomial pieces. In other words, for each $f \in \Pi_{k,n}$ there is a partition of $[-1\,..\,1]^2$ into at most n subsets and, on each such subset, f agrees with some polynomial of degree $\le k$. For this example, we have only $M_n + M_n \subset M_{n^2}$. Finding the degree of approximation from this ladder $(\Pi_{k,n} : n \in \mathbb{N})$ is one of the outstanding problems in nonlinear approximation. (Strictly speaking, this particular example lacks that trivial initial space, $\{0\}$, which is really only used in the general theory to simplify notation. For that, we might have to set here $M_n = \Pi_{k,n-1}$. Such a switch, from n to $n \pm 1$, does not change the orders n^{-r} of interest here.)

Even with these assumptions in place, the best we can say about $E_n(f)$ for given f has already been said: $E_n(f)$ converges monotonely to 0. The question of interest is just how 'fast' or 'slow' this convergence is. As with all measuring, these terms are made precise by comparing $E_n(f)$ with certain *standard* sequences. The most popular of these are the sequences $(n^{-\alpha} : n \in \mathbb{N})$ for some real α. Thus we are looking for conditions on f which guarantee that

$$E_n(f) = O(n^{-r}),$$

i.e., $\limsup_n E_n(f)n^r < \infty$, or, perhaps, even

$$E_n(f) = o(n^{-r}),$$

i.e., $\lim_n E_n(f)n^r = 0$, or

$$\sum_n n^r E_n(f) < \infty.$$

Such conditions may single out a rather 'small' subset of X in case X is complete but not equal to any of the M_n. This is certainly so if the constant in (iv) is $c = 1$, i.e., in case the M_n are linear subspaces.

(11) Proposition. *If (M_n) is a sequence of proper closed linear subspaces (i.e., only (iii), (iv) with $c = 1$, (i) is weakened to 'proper', and (v) is weakened to 'closed'), then, for any real sequence α_n converging monotonely to 0, the set*

$$A_\alpha := \{f \in X : E_n(f) = O(\alpha_n)\}$$

is 'thin' in the sense that it is of first category, i.e., the countable union of nowhere dense sets. In particular, it cannot equal X in case X is complete.

Proof: $A_\alpha = \cup_{n_0,N\in\mathbb{N}} \cap_{n\ge n_0} B^-_{N\alpha_n}(M_n)$, with

$$B_{Nr}(M_n) := \{f \in X : \mathrm{dist}(f, M_n) \le Nr\} = NB^-_r(M_n)$$

since $\mathbb{F}M_n \subset M_n$. If now, for some $r > 0$ and some x, $B_r(x) \subset B_{\alpha_n}(M_n)$, then $B_r(0) \subseteq (B_r(x) + B_r(-x))/2 \subset (B_{\alpha_n}(M_n) + B(_{\alpha_n}(-M_n))/2$ and, even if we only knew (iv), we could now conclude that $B_r(0) \subset B_{\alpha_n}(M_{cn})$ and if, as we actually assume, the M_n are proper closed linear subspaces, this implies (by Riesz' Lemma) that $r \le \alpha_n$, hence $B_r(x) \subset \cap_{n\ge n_0} B^-_{\alpha_n}(M_n)$ implies $r \le \lim_{n\to\infty} \alpha_n = 0$, hence A_α is of first category. However, if the M_n are not linear spaces, then no such conclusion can be drawn. □

Here is a nice example, provided by Olga Holtz, to show that properties (i)-(v) by themselves are not strong enough to imply this proposition's conclusion. Let $X := \ell_\infty$. For any finite $V \subset \mathbb{F} := \mathbb{R}$, let

$$F_V := \{f \in X : \mathrm{ran}\, f \subset V\}.$$

Then, for any $f \in X$, $\mathrm{dist}(f, F_V) = \|f - f_V\|$, with $f_V : n \mapsto \mathrm{argmin}_{v\in V} |f(n) - v|$. Therefore,

$$\forall\{\varepsilon > 0\}\forall\{W \subset B_\varepsilon(V)\}\ \mathrm{dist}(f, F_V) \le \mathrm{dist}(f, F_W) + \varepsilon.$$

Let now

$$F_k := \cup_{\#V\le k} F_V,$$

and let (f_m) be a minimizing sequence in F_k for $\|f - \cdot\|$. Then $(\mathrm{ran}\, f_m)$ is a bounded sequence of subsets in $\mathbb{R}$ of cardinality $\le k$, hence, AGTASMAT there is some $V \subset \mathbb{R}$ with $\#V \le k$ for which $\varepsilon_m := \inf\{\varepsilon : \mathrm{ran}\, f_m \subset B_\varepsilon(V)\}$ goes to zero as $m \to \infty$. Therefore

$$\mathrm{dist}(f, F_k) \le \mathrm{dist}(f, F_V) \le \mathrm{dist}(f, F_{\mathrm{ran}\, f_m}) + \varepsilon_m \le \|f - f_m\| + \varepsilon_m \xrightarrow[m\to\infty]{} \mathrm{dist}(f, F_k),$$

showing f_V to be a ba to f from F_k, hence showing F_k to be an existence set. Also, with $V_k := \{1 - (2j-1)/k : j = 1, \ldots, k\}$, we have

$$\mathrm{dist}(f, F_k) \le \mathrm{dist}(f, F_{\|f\|V_k}) \le \|f\|/k, \qquad (*)$$

and, since also F_k is homogeneous, and is increasing with k, the sequence (F_k) satisfies conditions (i)-(v), except that, at best, $F_k + F_k \subset F_{k^2}$. But (iv) is satisfied (with $c = 2$) by its subsequence $M_n := F_{2^n}$, $n = 1, 2, \ldots$, (with $M_0 = \{0\}$, of course), along with the other conditions in (i)-(v), yet, by $(*)$,

$$\{f \in X : \mathrm{dist}(f, M_n) = O(2^{-n})\} = X,$$

and this is not of first category since X is a Bs.

As is customary in mathematics, we express suitable conditions on f as membership in some set Y. A typical choice for Y is a *semi*-normed ls whose semi-norm we denote by $\|\cdot\|_Y$ which is continuously imbedded in X. The standard example has $M_n = \mathring{\Pi}_n \subset X := X_p$, with

$$X_p := \begin{cases} \mathbf{L}_p(\mathbb{T}), & p < \infty; \\ C(\mathbb{T}), & p = \infty, \end{cases}$$

and

$$Y = X_p^r := \{f \in X_p : \|f\|_Y := \|D^r f\|_p < \infty\},$$

with r some positive integer.

One hopes to choose Y and r in such a way that, simultaneously,

$$\exists\{C_J\}\forall\{n \in \mathbb{N},\ f \in Y\}\ \ E_n(f) \le C_J\, n^{-r}\|f\|_Y \tag{12}$$

and

$$\exists\{C_B\}\forall\{n \in \mathbb{N},\ m \in M_n\}\ \|m\|_Y \le C_B\, n^r \|m\|. \tag{13}$$

For historical reasons, the former is called a **Jackson**, or **direct**, inequality or estimate, the latter a **Bernshtein**, or **inverse**, inequality or estimate. The Jackson inequality gives a *lower* bound on the speed with which $E_n(f)$ goes to zero, and, as we shall see, the Bernshtein inequality provides an upper bound, at least indirectly. The most natural bridge between Jackson and Bernshtein inequalities turns out to be **Peetre's K-functional**:

$$K(f,t) := K(f,t;X,Y) := \inf_{g\in Y}(\|f-g\| + t\|g\|_Y).$$

This functional measures how 'smooth' f is in the sense that it tells us how closely we can approximate f by 'smooth' elements. The larger t, the more stress we lay on the smoothness of g. As a function of t, $K(f,t)$ is the infimum of a collection of straight lines, all with nonnegative slope, hence $K(f,.)$ is also weakly increasing (i.e., nondecreasing), and, further, is *concave* (as would be the infimum of any collection of straight lines, increasing or not). This implies that $t \mapsto K(f,t)/t$ *is (weakly) decreasing* (i.e., nonincreasing).

We also need the following somewhat complicated weighted sequence norm

$$\|a\|^{(r)} := \sum_{k=1}^{n} |k^r a(k)|/k, \quad a \in \mathbb{F}^n.$$

This norm is **monotone**, meaning that $|a| \le |b|$ implies $\|a\|^{(r)} \le \|b\|^{(r)}$. Further, for the following 'standard' null-sequence $a = (a(k) = k^{-(r-1+\alpha)} : k \le n)$, one computes

$$\|a\|^{(r)} = \sum_{k=1}^{n} k^{-\alpha} \le \int_0^n ()^{-\alpha} = n^{1-\alpha}/(1-\alpha), \tag{14}$$

as long as $\alpha \in [0\,..1)$. Here is a quick comment concerning the fact that, in the description of $\|a\|^{(r)}$, I did not combine the two powers of k appearing there. The reason is the following.

The argument to follow remains valid even when $\|\cdot\|$ and/or $\|\cdot\|_Y$ are only **quasi-norms**, meaning, in effect, that, instead of the triangle inequality, we only have

$$\|x+y\|^\mu \le \|x\|^\mu + \|y\|^\mu$$

for some μ and all x and y. In such a case, one considers, more generally, the weighted sequence norm

$$\|a\|_\mu^{(r)} := \Bigl(\sum_{k=1}^n |k^r a(k)|^\mu/k\Bigr)^{1/\mu}.$$

(15) Peetre's Theorem.
(i) (12) implies $\sup_n E_n(f)/K(f,n^{-r}) < \infty$.
(ii) (13) implies

$$K(f,n^{-r}) \le \mathrm{const}_r\, n^{-r}\|(E_k(f) : k \le n)\|^{(r)}. \tag{16}$$

In particular, we have $E_n(f) = O(n^{-(r-1+\alpha)})$ for some $\alpha \in [0\,..\,1)$ if and only if $K(f,n^{-r}) = O(n^{-(r-1+\alpha)})$. In the standard situation, the latter can be shown to be equivalent to having $f \in C^{(r-1)}$ with $D^{r-1}f$ in Lip_α.

Proof: (i): For any $g \in Y$, $E_n(f) \le \|f-g\| + E_n(g) \le C(\|f-g\| + n^{-r}\|g\|_Y)$, with $C := \max\{1, C_J\}$, and taking the inf over $g \in Y$ does it.

(ii): The trick here is the use of a telescoping series. (The proof in DeVore and Lorentz is a touch terse.) With $f_k \in \mathcal{P}_{M_k(f)}$, all k, we choose a sequence

$$1 = n_0 < n_1 < \cdots < n_\nu = n$$

and consider

$$h_k := f_{n_k} - f_{n_{k-1}} \in M_{cn_k}.$$

Then

$$\|h_k\| \le \|f - f_{n_k}\| + \|f - f_{n_{k-1}}\| \le 2E_{n_{k-1}}(f),$$

while $f_n = \sum_{k=1}^\nu h_k$. Therefore,

$$\begin{aligned}
K(f,n^{-r}) &\le \|f - f_n\| + n^{-r}\|f_n\|_Y \\
&\le E_n(f) + n^{-r}\sum_{k=1}^\nu \|h_k\|_Y \\
&\le n^{-r}\bigl(n^r E_n(f) + C_B\, c^r\, 2\sum_{k=1}^\nu n_k^r E_{n_{k-1}}(f)\bigr) \\
&\le n^{-r}\mathrm{const}_r \sum_{k=0}^{\nu-1} n_k^r E_{n_k}(f),
\end{aligned}$$

with $\mathrm{const}_r := 2C_B c^r \max_k (n_k/n_{k-1})^r$ a constant depending only on r (and the constants B_J and c) if we restrict attention to sequences (n_k) for which $\max_k n_k/n_{k-1} \le \mathrm{const}$. Since $k \mapsto E_k(f)$ is monotone decreasing, we have

$$\sum_{j<k\le J} k^{r-1} E_k(f) \ge E_J(f) \sum_{j<k\le J} k^{r-1},$$

while

$$\sum_{j<k\le J} k^{r-1} \ge \int_j^J ()^{r-1} = J^r(1/r)(1-(j/J)^r).$$

Hence, applying this with $j = n_k$, $J = n_{k+1}$, all k, we get

$$\sum_{k=0}^{\nu-1} n_k^r E_{n_k}(f) \le \mathrm{const}_r \sum_{k=1}^{n} k^{r-1} E_k(f)$$

provided we also insist that $\min_k (n_k/n_{k+1}) \le \mathrm{const} < 1$, i.e., the ratio n_k/n_{k-1} required earlier to be bounded must, on the other hand, be bounded away from 1. Choosing the sequence $(n_k : k = 1, \ldots, n)$ to be roughly dyadic, i.e., $n_k \sim 2^{k-1}$, all k, will do for both requirements. With this, we get (16), with const depending on C_B and r only. $\square$

The last argument proved that

$$\sum_{k=0}^{\nu-1} n_k^r E_{n_k}(f) \le \mathrm{const}_r \sum_{j=n_0}^{n_\nu} j^{r-1} E_j(f). \tag{17}$$

However, since also

$$\sum_{j\le k<J} k^{r-1} \le \int_j^J ()^{r-1} = (1/r)(J^r - j^r),$$

we also get

$$\sum_{j=n_0}^{n_\nu} j^{r-1} E_j(f) \le \mathrm{const}_r \sum_{k=0}^{\nu} n_k^r E_{n_k}(f)$$

with $\mathrm{const}_r = 2^r - 1$. Thus, up to r-dependent constants, we haven't lost anything when going from a sort-of dyadic sum to the full sum. We will make use of (17) in a moment, in Bernstein's proof that only smooth functions can be approximated well by trigonometric polynomials.

Remark The argument just given is very free and easy with constants. Its only purpose is to establish statements about the degree of convergence, such as the assertion that $o(n^{-r}) \ne E_n(f) = O(n^{-r})$. Rather different arguments are used to establish statements of the sort that

$$E_n(f) \le \mathrm{const}\, n^{-r} \|f\|_Y, \quad f \in Y,$$

with const the smallest possible constant independent of f and n.

The first such theorem was proved by Bernshtein, using

Bernshtein Inequality. *For any* $f \in \overset{\circ}{\Pi}_n \subset C(\mathbb{T})$,

$$\|Df\| \le n\|f\|.$$

Proof: Here is v.Golitschek's proof of Szegö's stronger inequality

$$(Df)^2 + (nf)^2 \le (n\|f\|)^2, \quad f \in \overset{\circ}{\Pi}_n. \tag{18}$$

As the inequality is homogeneous in f, we may assume WLOG that $\|f\| < 1$. (The reason for this assumption will be clear in a moment.) Further, since $\overset{\circ}{\Pi}_n$ is translation-invariant, it is sufficient to prove that

$$(Df(0))^2 + (nf(0))^2 \le (n\|f\|)^2, \quad f \in \overset{\circ}{\Pi}_n. \tag{19}$$

We may even assume WLOG that $Df(0) \ge 0$. Now let α be the unique point in $(-\pi \,..\, \pi)/(2n)$ at which $\sin n\alpha = f(0)$, and consider the trig.pol.

$$s := \sin n(\cdot + \alpha) - f.$$

Since $\|f\| < 1$, s alternates in sign at the extrema of $\sin n(\cdot + \alpha)$ and, as there are $2n$ such, s has exactly one zero between any two adjacent extrema of $\sin n(\cdot + \alpha)$. In particular, one of these zeros must be the point 0 since we chose α to make it so. If now $Df(0) > n\cos(n\alpha) = D\sin n(\cdot + \alpha)(0)$, then f would rise above $\sin n(\cdot + \alpha)$ as we move to the right from 0, yet is certain to be below it again by the time we reach the first extremum of $\sin n(\cdot + \alpha)$ to the right of 0 which is a maximum, and this would imply an impossible second zero of s between the two extrema which bracket 0. Consequently,

$$0 \le Df(0) \le n\cos n\alpha = n\sqrt{1 - (f(0))^2},$$

using the fact that, by our choice of α, $\sin n\alpha = f(0)$. Squaring and reordering terms gives

$$(Df(0))^2 + (nf(0))^2 \le n^2,$$

which isn't quite what we hoped for. However, for any $r > \|f\|$, we also have $\|f/r\| < 1$, and therefore $(Df(0))^2 + (nf(0))^2 \le (nr)^2$ for all such r, and taking now the infinum over all such r gives (19). □

We are ready to prove a sample inverse theorem.

(20) Theorem. *Let* $M_n = \overset{\circ}{\Pi}_n$*, as a subset of* $C(\mathbb{T})$*. Let* $f \in C(\mathbb{T})$ *and assume that, for some* r,

$$\|(E_n(f) : n \in \mathbb{N})\|^{(r)} < \infty \tag{21}$$

(as would be the case in case $E_n(f) = O(n^{-r-\varepsilon})$ *for some positive* ε*). Then* $f \in C^{(r)}(\mathbb{T})$.

Proof: Since differentiation is a closed operator, it is sufficient to show that f is the uniform limit of a sequence which is Cauchy in $C^{(r)}(\mathbb{T})$. With $p_n \in \mathcal{P}_{M_n}(f)$, all

n, we have $f = \lim_n p_n$; however, we have no way of knowing that the whole sequence $(p_n : n \in \mathbb{N})$ is Cauchy in $C^{(r)}$. The sequence $(p_{2^n} : n \in \mathbb{N})$, on the other hand, is seen to be Cauchy, as follows.

$$\begin{aligned}
\|D^r(p_{2^j} - p_{2^J})\| &\le \sum_{k=j+1}^{J} \|D^r(p_{2^k} - p_{2^{k-1}})\| \\
&\le \sum_{k=j+1}^{J} (2^k)^r \|p_{2^k} - p_{2^{k-1}}\| \\
&\le \sum_{k=j+1}^{J} (2^k)^r 2E_{2^{k-1}}(f) \\
&\le 2\mathrm{const}_r \sum_{k=2^j}^{2^J} k^{r-1} E_k(f),
\end{aligned}$$

the last inequality by (17). By assumption, this last sum goes to zero as $j \to \infty$. □

This theorem is quite remarkable since it says that even if f is very smooth *except* on a small subinterval of $\mathbb{T}$, it will be hard to approximate f well by trig.pols.

There are similar results for approximation by *algebraic* polynomials on $[a\,..\,b]$, except for the following hitch. In contrast to $\mathbb{T}$, the points in $[a\,..\,b]$ are *not* all equal. Taking for simplicity the max-norm on $[a\,..\,b]$, one has the Jackson inequality (see below)

$$\text{(22)} \qquad \mathrm{dist}_\infty(f, \Pi_n) =: E_n(f)_\infty \le \mathrm{const} n^{-1} \|Df\|_\infty.$$

However, the best inequality relating $\|f\|$ and $\|Df\|$ for $f \in \Pi_n \in C([a..b])$ is the **Markov Inequality**

$$\text{(23)} \qquad \|Df\|_\infty \le \frac{2}{|b-a|} n^2 \|f\|_\infty,$$

(note that, because of the possibility of dilating, the interval length must figure in exactly the position at which it appears, hence the inequality is sharp since it becomes equality for $[a\,..\,b] = [-1\,..\,1]$ and $f : t \mapsto \cos(n \cos^{-1} t)$ the Chebyshev polynomial of degree n) and these two inequalities do not at all match in the sense of Theorem 15. It is possible, though, to prove that $f \in C([a\,..\,b])$ must be in $C^{(r)}(I)$ for all closed subintervals I of $(a\,..\,b)$ in case $E_n(f)_\infty = O(n^{-r-\varepsilon})$ for some positive ε.

Because $\mathring{\Pi}_n$ is translation-invariant, it is easier to prove Jackson inequalities for $M_n = \mathring{\Pi}_n \subset C(\mathbb{T})$ than for $\Pi_n \subset C([a\,..\,b])$ (Π_n is translation-invariant only on $\mathbb{R}$ or $\mathbb{C}$). The standard argument uses integral operators of **convolution type**. With L_n functions still to be specified, one considers the approximation

$$L_n{*}f : t \mapsto \int_{\mathbb{T}} L_n(t-s) f(s)\,\mathrm{d}s = \int_{\mathbb{T}} f(t-s) L_n(s)\,\mathrm{d}s.$$

One makes the following assumptions:
(i) $L_n \in \mathring{\Pi}_n$.
Then, $\mathring{\Pi}_n$ being **translation-invariant**, i.e., $\forall\{p \in \mathring{\Pi}_n,\ s \in \mathbb{T}\}\ p(\cdot - s) \in \mathring{\Pi}_n$, we necessarily have

$$L_n(t-s) = \sum_j \varphi_j(t)\psi_j(s),$$

with $[\varphi_j : -n \le j \le n]$ any basis for $\mathring{\Pi}_n$. Therefore

$$L_n{*}f = \sum_j \varphi_j \int_{\mathbb{T}} \psi_j f \quad \in \mathring{\Pi}_n.$$

(ii) $\int_{\mathbb{T}} L_n = 1$, hence $L_n{*}()^0 = ()^0$.
Assuming that

$$\lambda_n := \int_{\mathbb{T}} L_n \ne 0,$$

this can always be achieved by then replacing L_n by L_n/λ_n. After that,

$$(f - L_n{*}f)(t) = \int_{\mathbb{T}} \big(f(t) - f(t-s)\big) L_n(s)\,\mathrm{d}s = \int_{\mathbb{T}} (\nabla_s f)(t)\, L_n(s)\,\mathrm{d}s,$$

with

$$\nabla_h f := f - f(\cdot - h) =: \Delta_h f(\cdot - h).$$

Note that

$$\|\nabla_s f\| \le \sup_{0\le k<|s|} \|\Delta_k f\| =: \omega(f, |s|),$$

with $\omega(f,\cdot)$ the **(uniform) modulus of continuity of** f. One could now try to get an error estimate involving $\omega(f, 1/n)$ by making use of the fact that

$$\omega(f,|s|) = \omega(f, n|s|{\cdot}1/n) \le (1 + n|s|)\omega(f, 1/n)$$

(using the obvious fact that $\omega(f, h+k) \le \omega(f,h) + \omega(f,k)$), hoping that

$$\sup_n \int_{\mathbb{T}} (1 + n|s|)|L_n(s)|\,\mathrm{d}s < \infty.$$

However, the L_n usually employed satisfy the following additional assumption:
(iii) L_n even.
and, with this, we even have

$$\begin{aligned} f(t) - L_n{*}f(t) &= \int_0^\pi \big(f(t) - f(t+s) \quad + \quad f(t) - f(t-s)\big) L_n(s)\,\mathrm{d}s \\ &= \int_0^\pi (-\Delta_s \nabla_s f)(t) L_n(s)\,\mathrm{d}s, \end{aligned}$$

with

$$\Delta_h \nabla_h f = \Delta^2 f(\cdot - h),$$

and

$$\|\Delta_s^2 f\| \le \sup_{h<|s|} \|\Delta_h^2 f\| =: \omega_2(f, |s|),$$

and $\omega_2(f,\cdot)$ called the **second (uniform) modulus of smoothness of** f. In this language, $\omega(f,\cdot)$ is called the **first (uniform) modulus of smoothness of** f. It is clear how one would define the rth modulus of smoothness, for any natural r, as is the bound

$$\omega_r(f,h) := \sup_{k<h} \|\Delta_k^r f\| \le 2^r \omega(f,h).$$

Since $\omega_2(f,h) = \omega_2(f,(nh)/n) \le (1+nh)^2 \omega_2(f,1/n)$, a natural assumption now is the following **moment condition**:
(iv) $\sup_n \int_0^\pi (ns)^k \, |L_n(s)| \, \mathrm{d}s < \infty$, i.e., $\int_0^\pi s^k |L_n(s)| \, \mathrm{d}s = O(n^{-k})$, for $k = 0, 1, 2$.
and with it, we get the typical **Jackson inequality**

$$\|f - L_n * f\| \le c\omega_2(f; 1/n).$$

It remains to find suitable L_n.

The simplest choice is $L_n = D_n / \int_{\mathbb{T}} D_n$ with

$$D_n(t) := \sum_{|j|\le n} e_{\mathrm{i}j}/2 = \frac{\sin((n+1/2)t)}{2\sin(t/2)}$$

Dirichlet's kernel. For it, $L_n \in \overset{\circ}{\Pi}_n$, $D_n * 1 = \pi$, and L_n even. However, L_n decays too slowly away from 0 to satisfy (iv); in fact, already $k=0$ gives trouble since $\int_{\mathbb{T}} |D_n| \sim \ln n$, hence, as already mentioned earlier, $L_n *$ cannot converge to 1 pointwise on $C(\mathbb{T})$, hence. a more sophisticated analysis wouldn't help here, either.

The next choice is $L_n = F_n / \int_{\mathbb{T}} F_n$, with

$$F_n(t) := \frac{1}{2(n+1)} \left(\frac{\sin(n+1)t/2}{\sin t/2} \right)^2$$

Fejér's kernel. One checks that

$$F_n = \sum_0^n D_k/(n+1),$$

hence $F_n \in \overset{\circ}{\Pi}_n$, and $F_n * 1 = \pi$. Also, $F_n \ge 0$, hence only need

$$\int_0^\pi (1+ns)^2 F_n(s) \, \mathrm{d}s = O(1).$$

However this, unfortunately, is not true. Thus, while, by Korovkin, $F_n{*}f$ converges to f uniformly for every $f \in C(\mathbb{T})$, we don't, offhand, get a Jackson estimate from it since (iv) does not hold for it.

Jackson's choice is the **Jackson kernel**

$$K_n(t) := \left(\frac{\sin mt/2}{\sin t/2}\right)^4 /\lambda_n, \quad m := \lfloor n/2 \rfloor + 1,$$

normalized to have $\int_{\mathbb{T}} K_n = 1$. It is a special case of the **generalized Jackson kernel**s:

$$K_{n,r}(t) := \left(\frac{\sin mt/2}{\sin t/2}\right)^{2r} /\lambda_{n,r}, \quad m := \lfloor n/r \rfloor + 1.$$

Since

$$\left(\frac{\sin mt/2}{\sin t/2}\right)^2 /(2m) = 1/2 + \sum_{k=1}^{m-1} (1 - k/m) \cos kt$$

(as one verifies!), it follows that $K_{n,r}$ is an even, nonnegative, trig.pol. of degree $\le n$. In particular $K_{n,r}{*}1 > 0$, hence the 0th moment condition is trivially satisfied. For the others, one may prove that $K_{n,r}$ satisfies the moment conditions

$$\sup_n \int_0^\pi (nt)^k K_{n,r}(t)\,\mathrm{d}t < \infty, \quad k = 0, 1, \ldots, 2r-2,$$

as follows. We'll use the convenient notation

$$a_n \sim b_n \quad :\Longleftrightarrow \quad 0 < \liminf_n a_n/b_n \le \limsup_n a_n/b_n < \infty.$$

Now, since $t/\pi \le \sin(t/2) \le t/2$ on $[0\,..\,\pi]$,

$$\begin{aligned}\int_0^\pi t^k \Big(\frac{\sin(mt/2)}{\sin t/2}\Big)^{2r}\,\mathrm{d}t \sim \int_0^\pi t^k \Big(\frac{\sin(mt/2)}{t}\Big)^{2r}\,\mathrm{d}t \\ &= \Big(\frac{2}{m}\Big)^{k-2r+1} \int_0^{m\pi/2} u^k \Big(\frac{\sin u}{u}\Big)^{2r}\,\mathrm{d}u \\ &\sim n^{2r-1-k} \int_0^\infty u^k \Big(\frac{\sin u}{u}\Big)^{2r}\,\mathrm{d}u \\ &\sim n^{2r-1-k}\end{aligned}$$

as long as $k - 2r < -1$, hence

$$\int_0^\pi (ns)^k K_{n,r}(s)\,\mathrm{d}s \sim n^k n^{2r-1-k}/n^{2r-1} \sim 1,$$

for $k = 0, \ldots, 2r-2$.

Thus, for $r \geq 2$, the sequence $L_n = K_{n,r}/\int_{\mathbb{T}} K_{n,r}$ satisfies all four conditions (i)–(iv) above.

For $r = 2$, Jackson got in this way the Jackson Inequality. For $r > 2$, the additional moment conditions provide, for $L_n := K_{n,r}/\int_{\mathbb{T}} K_{n,r}$, the inequalities (due to Stechkin):

$$\|f - L_n * f\|_p \leq C_r \omega_{2r-2}(f, 1/n)_p,$$

for any $f \in L_p(\mathbb{T})$ and any $1 \leq p \leq \infty$.

The resulting constants are far from sharp. Favard showed that, with S the unit ball in the semi-normed lss $Y := W_p^r(\mathbb{T})$ of $X_p(\mathbb{T})$, with semi-norm

$$\|f\|_Y := \|D^r f\|_p,$$

$$\sup_{f \in S} E_{n-1}(f) = E_{n-1}(B_p^r) \leq K_r/n^r,$$

with this inequality exact for $p = 1, \infty$, and the numbers K_r, the so-called **Favard constants**, given as the value of a fast converging series whose value, for large r, is indistinguishable from $\pi/4$.

The extremal function, B_p^r, is a **Bernoulli spline**, of which, perhaps, more anon.

It is standard to prove Jackson's theorem for *algebraic* polynomials from the one for trigonometric polynomials. For this, one first translates and dilates $[a \mathinner{..} b]$ into the interval $[-1 \mathinner{..} 1]$. Then one considers the map

$$F : C[-1 \mathinner{..} 1] \to C(\mathbb{T}) : f \mapsto Ff : \theta \mapsto f(\cos(\theta)).$$

(Draw the picture.) This is a linear map which carries $C[-1 \mathinner{..} 1]$ isometrically into

$$C(\mathbb{T})_e := \{g \in C(\mathbb{T}) : g(-\cdot) = g\},$$

the space of all **even** 2π-periodic functions. In particular, F maps Π_n onto

$$\mathring{\Pi}_{n,e} := \mathring{\Pi}_n \cap C(\mathbb{T})_e.$$

Also,

$$\omega(Ff, \cdot) \leq \omega(f, \cdot)$$

since the map $[-1 \mathinner{..} 1] \to [0 \mathinner{..} \pi] : t \mapsto \cos^{-1}(t)$ has slope ≥ 1 in absolute value everywhere. Hence

$$\operatorname{dist}(Ff, \mathring{\Pi}_n) \leq \text{const } \omega(Ff, 1/n) \leq \text{const } \omega(f, 1/n).$$

Therefore, with $m \in \mathcal{P}_{\mathring{\Pi}_n}(Ff)$,

$$\operatorname{dist}(f, \Pi_n) \leq \|f - F^{-1}m\| = \operatorname{dist}(Ff, \mathring{\Pi}_n) \leq \text{const } \omega(f, 1/n), \tag{24}$$

provided $F^{-1}m$ is defined and in Π_n. For this, observe that the map

$$C(\mathbb{T}) \to C(\mathbb{T})_e : g \mapsto g_e := (g + g(-\,\cdot))/2$$

is norm-reducing since $g \mapsto g(-\,\cdot)$ is an isometry. Hence, if $g \in C(\mathbb{T})$ (as is the case for $g = Ff$) and $m \in \mathcal{P}_{\mathring{\Pi}_n}(g)$, then m_e is an even trig.pol. in $\mathring{\Pi}_n$ and $\|Ff - m_e\| = \|g_e - m_e\| \le \|g - m\| = \operatorname{dist}(g, \mathring{\Pi}_n)$, hence also $m_e \in \mathcal{P}_{\mathring{\Pi}_n}(g)$, i.e., we may take $m \in \mathcal{P}_{\mathring{\Pi}_n}(Ff)$ to be even, hence in $F(\Pi_n)$. In fact, Π_n is Haar, hence we have just proved that the ba from $\mathring{\Pi}_n$ to an even function is even.

Note, though, that we cannot bound $\omega(f, \cdot)$ in terms of $\omega(Ff, \cdot)$, and this is another indication that we cannot expect the same kind of paired direct/inverse theorems concerning the degree of approximation by algebraic polynomials.

Instead, one has theorems of the following kind.

(25) Theorem (Nikol'skii, Timan). *For some* const *and for all* $f \in C[-1\,..\,1]$ *and all* n, *there exists* $p_n \in \Pi_n$ *so that*

$$|f(t) - p_n(t)| \le \text{const } \omega(f, \Delta_n(t)),$$

with

$$\Delta_n : t \mapsto \max\{\frac{1}{n^2}, \frac{\sqrt{1-t^2}}{n}\}.$$

If one is not too worried about the constants involved, then Jackson's theorem provides the right order of the degree of approximation by polynomials to smooth functions, as follows: Starting with Jackson's theorem,

$$\operatorname{dist}(f, \Pi_n) \le \text{const } \omega(f, 1/n),$$

and adding to it the fact that, for $f \in C^{(1)}[a\,..\,b]$,

$$\omega(f, h) \le h\|Df\|,$$

we get, for any $p \in \Pi_n$,

$$\operatorname{dist}(f, \Pi_n) = E_n(f - p, \Pi_n) \le \text{const } (1/n)\,\|D(f - p)\|,$$

hence

$$\operatorname{dist}(f, \Pi_n) \le \text{const } (1/n) \operatorname{dist}(Df, \Pi_{n-1}).$$

Thus, by induction, for $n \gg r$,

$$\operatorname{dist}(f, \Pi_n) \le \text{const}_r\ (1/n)^r \omega(D^r f, 1/n).$$

Finally, here is a standard result relating a specific K-functional to the modulus of smoothness of a certain order.

(26) Theorem (Johnen). *If T is a closed interval (finite, infinite or all of $\mathbb{R}$) or $\mathbb{T}$, and $r \in \mathbb{N}$, then there exist positive constants c and C such that, for all $p \in [1\,..\,\infty]$ and all $f \in \mathbf{L}_p(T)$,*

$$c\omega_r(f,t)_p \le K(f,t^r;\mathbf{L}_p,W_p^{(r)}) \le C\omega_r(f,t)_p, \quad t>0. \tag{27}$$

Here, to recall,

$$\omega_r(f,t)_p := \sup_{h\le t} \|\Delta_h^r f\|_p,$$

with

$$\Delta_h^r f := \sum_{k=0}^{r} \binom{r}{k} (-1)^{r-k} f(\cdot + kh).$$

If T has an endpoint, then $\Delta_h^r f$ may, offhand, only be defined on some proper subset of T; it is taken to be zero wherever it is not defined.

An interlude on divided differences

While the *cardinal B-splines* already discussed in Amos Ron's lectures were, indeed, originally defined by Schoenberg in terms of their Fourier transform, B-splines for arbitrary knot sequences were introduced by Curry and Schoenberg as a divided difference of the truncated power function:

$$B_{j,k,\mathbf{t}} : x \mapsto (t_{j+k}-t_j)[t_j,\ldots,t_{j+k}](\cdot - x)^{k-1}.$$

This gives me an excuse for the following quick (and unusual) introduction to divided differences.

The sequence

$$(\cdot - c_1)(\cdot - c_2)\cdots(\cdot - c_{n-j}), \quad j=1,\ldots,n, \tag{28}$$

is linearly independent since no two terms have the same polynomial degree. Since the sequence contains n terms and each term is in $\Pi_{<n}$, a space of dimension n, it must necessarily be a basis for $\Pi_{<n}$, the **Newton basis with centers** $c_1,\ldots,c_{n-1}$. Correspondingly, there is, for every $p \in \Pi_{<n}$, exactly one sequence $a = a(p;c_1,\ldots,c_{n-1}) \in \mathbb{F}^n$ so that

$$p = \sum_{j=1}^{n} a_j \prod_{1\le i\le n-j} (\cdot - c_i). \tag{29}$$

This is the **Newton form centered at** $(c_1,\ldots,c_{n-1})$ for p. Its coefficients depend linearly on p.

Since the sequence (28) of **Newton polynomials** depends smoothly on the centers (and is a basis for $\Pi_{<n}$ for every choice of the centers), it follows that the $a_j = a_j(p; c_1, \ldots, c_{n-1})$ are also smooth functions of the centers. Indeed, with

$$V(c) : \mathbb{F}^n \to \Pi_n : a \mapsto \sum_{j=1}^{n} a_j \prod_{1 \le i \le n-j} (\cdot - c_i)$$

the invertible basis-map, we get, on differentiating the identity $V(c)V(c)^{-1} = 1$ as a function of c and rearranging, that

$$D(V(c)^{-1}) = V(c)^{-1} DV(c) V(c)^{-1}$$

is continuous and smooth. Actually, much more is true, as is already indicated by the following simple observations:

(i) If all the centers are 0, then the Newton form (29) reduces to the **power form**:

$$p = \sum_{j=1}^{n} a_j \, ()^{n-j}, \tag{30}$$

and, correspondingly,

$$a_j(p; 0, 0, \ldots) = D^{n-j} p(0)/(n-j)!, \quad j = 1, \ldots, n.$$

(ii) If, slightly more generally, all the centers are the same, for example $c_i = \zeta$, all i, then the Newton form (29) reduces to the **shifted power form**:

$$p = \sum_{j=1}^{n} a_j \, (\cdot - \zeta)^{n-j}, \tag{31}$$

and, correspondingly,

$$a_j(p; \zeta, \zeta, \ldots) = D^{n-j} p(\zeta)/(n-j)!, \quad j = 1, \ldots, n.$$

(iii) Since the first j terms in the Newton form all have the factor $(\cdot - c_1) \cdots (\cdot - c_{n-j})$, we can write

$$p = (\cdot - c_1) \cdots (\cdot - c_{n-j}) \, q \; + \; p_{n-j},$$

for some polynomial q and with

$$p_{n-j} := a_{j+1} (\cdot - c_1) \cdots (\cdot - c_{n-j-1}) + \textbf{l.o.t.}.$$

This identifies a_{j+1} as the **leading coefficient**, that is, the coefficient of $()^{n-j-1}$ in the power form, for the unique polynomial of degree $< n-j$ that **agrees with p at the sequence** $(c_1, \ldots, c_{n-j})$ in the sense that $p - p_{n-j}$ has the zeros $c_1, \ldots, c_{n-j}$, counting multiplicities. This condition does, indeed, determine p_{n-j} uniquely since, if also r is a polynomial of degree $< n-j$ for which $p-r$ has the zeros $c_1, \ldots, c_{n-j}$ (counting multiplicities), then the difference, $p_{n-j}-r$, must have these same $n-j$ zeros, hence $p_{n-j}=r$. It follows that a_{j+1} in (29) depends only on p and on $c_1, \ldots, c_{n-j}$. This is reflected in the standard notation for a_{j+1}:

$$a_{j+1} =: [c_1, \ldots, c_{n-j}]p.$$

For $j=0$, this reads $a_1 = [c_1, \ldots, c_n]p$, with the additional 'center' c_n arbitrary (since p is certainly the unique polynomial in $\Pi_{<n}$ that agrees with p at any n-sequence of points). Further, a_{j+1} does not depend on the *sequence* $(c_1, \ldots, c_{n-j})$, i.e., on the particular ordering of its entries, but only on the *multiset* formed by its entries. In other words,

$$(c_1, \ldots, c_{n-j}) \mapsto [c_1, \ldots, c_{n-j}]p$$

is a symmetric function of its arguments.

It remains to provide an algorithm for the construction of the numbers $[c_1, \ldots, c_j]p$, $j = 1, 2, \ldots$.

There's no problem when $c_1 = \cdots = c_j = z$, since, as we saw in (ii),

$$[\underbrace{z, z, \ldots, z}_{j \text{ terms}}]p = D^{j-1}p(z)/(j-1)!. \tag{32}$$

In the contrary case, we get the proper information from considering first the efficient evaluation of the Newton form (29), taking full advantage of the many common factors. Indeed,

$$p(x) = (x-c_1)\big((\cdots(x-c_{n-2})[(x-c_{n-1})a_1 + a_2] + \cdots) + a_{n-1}\big) + a_n,$$

leading to the following algorithm.

Nested Multiplication aka Horner's Method. *Input: The sequence* $a = (a_i : i = 1, \ldots, n)$ *of coefficients in the Newton form (29) for some* $p \in \Pi_{<n}$*, the relevant center sequence* $c = (c_i : i = 1, \ldots, n-1)$*, and some point* z*.*

$b_1 := a_1$
`for` $j = 1, 2, \ldots, n-1$`, do:`
  $b_{j+1} := (z - c_{n-j})b_j + a_{j+1}$
`enddo`

Output: The number $b_n = p(z)$*.*

The algorithm provides the value of p at z, at a cost of only $3n$ flops. More than that, the entire sequence $(b_1, \ldots, b_n)$ generated by this algorithm is valuable since it provides the coefficients in the Newton form

$$p = \sum_{j=1}^{n} b_j \prod_{j \le i < n} (\cdot - c_{i-1}) \tag{33}$$

centered at $(c_0 := z, c_1, \ldots, c_{n-2})$. Indeed, directly from the algorithm,

$$a_{j+1} = \begin{cases} b_1, & j = 0; \\ b_{j+1} + (c_{n-j} - z)b_j & j > 0. \end{cases} \tag{34}$$

On substituting these expressions for the a_j into (29), we find

$$\begin{aligned} p(x) &= b_1(x - c_1)\cdots(x - c_{n-1}) \\ &\quad + (b_2 + (c_{n-1} - z)b_1)\,(x - c_1)\cdots(x - c_{n-2}) \\ &\quad + (b_3 + (c_{n-2} - z)b_2)\,(x - c_1)\cdots(x - c_{n-3}) \\ &\quad + \cdots \\ &= b_1(x - c_1)\cdots(x - c_{n-2})(x - z) \\ &\quad b_2(x - c_1)\cdots(x - c_{n-3})(x - z) \\ &\quad + \cdots \end{aligned}$$

which proves (33). In particular, $b_{j+1} = [c_0, \ldots, c_{n-j-1}]p$, all j, hence (34) can be rewritten

$$[c_1, \ldots, c_{n-j}]p = [c_0, \ldots, c_{n-j-1}]p + (c_{n-j} - c_0)[c_0, \ldots, c_{n-j}]p.$$

For $c_{n-j} \neq c_0$, this gives the **recurrence relation**

$$\frac{[c_1, \ldots, c_{n-j}]p - [c_0, \ldots, c_{n-j-1}]p}{c_{n-j} - c_0} = [c_0, \ldots, c_{n-j}]p \tag{35}$$

which holds for arbitrary $c_0, \ldots, c_{n-j}$ (as long as $c_0 \neq c_{n-j}$). This recurrence relation led to the name **divided differences** for these numbers, and is the simple tool for the calculation, in a divided difference table, of the coefficients

$$a_{j+1} = [c_1, \ldots, c_{n-j}]p, \quad j = 0, \ldots, n-1, \tag{36}$$

needed for the Newton form (29) for $p \in \Pi_{<n}$, starting with the numbers

$$[\underbrace{\zeta, \ldots, \zeta}_{\rho \text{ terms}}]p \;=\; D^{\rho-1}p(\zeta)/(\rho-1)!, \quad 0 < \rho \leq \#\{i \in \{1, \ldots, n\} : \zeta = c_i\}. \tag{37}$$

The table is triangular and contains eventually all the numbers $[c_i, \ldots, c_j]p$, $1 \leq i \leq j \leq n$. Assuming $c_1 \leq \cdots \leq c_n$ for simplicity, each such entry $[c_i, \ldots, c_j]p$ either has $c_i = c_j$ in which case it is supplied by (37) (with $\zeta = c_i$ and $\rho = j + 1 - i$), or else $c_i < c_j$ in which case $[c_i, \ldots, c_j]$ can be computed from $[c_i, \ldots, c_{j-1}]p$ and $[c_{i+1}, \ldots, c_j]p$ by the recurrence relation.

When the table is filled in, the coefficients (36) for (29) can be found along one of its edges.

Having generated this table in the manner described (and assuming still for simplicity that $c_1 \le \cdots \le c_n$), we obtain from the table the coefficients of the Newton form (29) of the polynomial p for which the values

$$(38) \qquad y_j = D^\rho p(c_j)/\rho!, \quad \rho = \max\{j - i : c_i = c_j\}, \quad j = 1, \ldots, n,$$

are as entered into the table. But this means that, by entering arbitrary numbers y_j at these places in the table, and then completing the table via the recurrence, we obtain the unique polynomial p in $\Pi_{<n}$ that satisfies (38). In other words, we have completely solved the problem of polynomial interpolation, including osculatory (or Hermite) interpolation. If the y_i are computed from some (sufficiently smooth) function f (polynomial or not), then the table entries are denoted by

$$[c_i, \ldots, c_j]f$$

and called divided differences of the function f. Thus the

Definition. *For any sufficiently smooth function f, its* **divided difference at** $c_i, \ldots, c_j$, *denoted by*

$$[c_i, \ldots, c_j]f,$$

is the leading coefficient of the unique polynomial of degree $\le j - i$ that agrees with f at the sequence $(c_i, \ldots, c_j)$.

Theorem. *For any sufficiently smooth function f, and any sequence $(c_1, \ldots, c_n)$,*

$$p := \sum_{j=1}^{n} [c_1, \ldots, c_{n-j+1}]f \prod_{i=1}^{n-j} (\cdot - c_i)$$

is the unique polynomial of degree $< n$ that agrees with f at the sequence $(c_1, \ldots, c_n)$.

Univariate B-splines

Here is a very swift introduction to B-splines, and thereby to splines *aka* smooth piecewise polynomials. (A picturesque version of this material is available, by anonymous ftp, from the site `/ftp.cs.wisc.edu` in the subdirectory `Approx` as the `postscript` file `bsplbasic.ps`.)

Let $\mathbf{t} := \cdots \le t_{i-1} \le t_i \le t_{i+1} \le \cdots$ be a nondecreasing sequence. This sequence may be finite, infinite, or even bi-infinite. Two important (and extreme) special cases are

(i) $\mathbf{t} = \mathbb{Z}$, leading to **cardinal splines**;

(i) $\mathbf{t} = \mathbb{B} := (\cdots, 0, 0, 0, 1, 1, 1, \ldots)$, leading to the **Bernstein-Bézier** form (or, BB-form) for polynomials (restricted to $[0 \,..\, 1]$).

The associated (normalized) B-splines of **order** k for the **knot sequence t** are, by definition, the functions

$$B_j := B_{j,k} := B_{j,k,\mathbf{t}} : x \mapsto [t_j, \ldots, t_{j+k}](\cdot - x)_+^{k-1}(t_{j+k} - t_j), \tag{39}$$

with $[T]$ the divided difference functional introduced in the preceding section. The normalization ensures that the B_j form a (nonnegative) **partition of unity** (see (53) below). The B-spline B_j was originally denoted by $N_j = N_{j,k,\mathbf{t}}$ in order to distinguish it from the differently normalized B-spline

$$M_j = M_{j,k,\mathbf{t}} := \frac{k}{t_{j+k} - t_j} B_{j,k,\mathbf{t}},$$

for which

$$\int M_j = 1.$$

The latter normalization arises naturally when applying the divided difference to both sides of the Taylor identity:

$$f = \sum_{r<k} D^r f(a)(\cdot - a)^r/r! + \int_a^b k(\cdot - s)_+^{k-1} D^k f(s)\,\mathrm{d}s/k!$$

to obtain (under the assumption that $t_j, \ldots, t_{j+k} \in [a \,..\, b]$)

$$[t_j, \ldots, t_{j+k}]f = \int_{\mathbb{R}} M_{j,k,\mathbf{t}}\, D^k f/k!, \tag{40}$$

showing that $M_{j,k,\mathbf{t}}$ is the Peano kernel for the divided difference $[t_j, \ldots, t_{j+k}]$.

By using Leibniz' Rule

$$[\tau_0, \ldots, \tau_k](fg) = \sum_{i=0}^{k}([\tau_0, \ldots, \tau_i]f)([\tau_i, \ldots, \tau_k]g)$$

for the divided difference of a product, applied to the particular product $(\cdot - x)_+^{k-1} = (\cdot - x)(\cdot - x)_+^{k-2}$, one readily obtains the **recurrence relations**

$$B_{i,k} = \omega_{i,k} B_{i,k-1} \ + \ (1 - \omega_{i+1,k}) B_{i+1,k-1} \tag{41a}$$

with

$$\begin{aligned} B_{i,1} &:= \chi_{[t_i \,..\, t_{i+1})}; \\ \omega_{i,k}(x) &:= \begin{cases} \frac{x - t_i}{t_{i+k-1} - t_i}, & \text{if } t_i < t_{i+k-1} \\ 0, & \text{otherwise} \end{cases} \end{aligned} \tag{41b}$$

However, once one knows these recurrence relations, it is most efficient for the development of the B-spline theory to take (41) as the *starting point*, i.e., to *define* B-splines by (41). Equivalently, the development about to be given will make no use of divided differences, but will rely entirely on (41).

It follows at once that $B_{i,k}$ can be written in the form

$$B_{i,k} = \sum_{j=i}^{i+k-1} b_{j,k} \chi_{[t_j \,..\, t_{j+1})}, \tag{42}$$

with each $b_{j,k}$ a polynomial of degree $< k$ since it is the sum of products of $k-1$ linear polynomials. Therefore,

$$B_{i,k} \in \Pi_{<k,(t_i,\ldots,t_{i+k})} \qquad \text{on } [t_i \,..\, t_{i+k}],$$

while

$$B_{i,k} = 0 \qquad \text{off } [t_i \,..\, t_{i+k}]. \tag{43}$$

Further, $B_{i,k}$ depends only on the knots $t_i, \ldots, t_{i+k}$. For this reason, the alternative notation

$$B_{i,k} =: B(\cdot | t_i, \ldots, t_{i+k}) \tag{44}$$

is also customary. The actual smoothness of $B_{i,k}$ depends on the multiplicity with which each of the knots t_j, $i \le j \le i+k$, appears in its knot sequence $(t_i, \ldots, t_{i+k})$, as we will see in a moment.

Since both $\omega_{i,k}$ and $1 - \omega_{i+1,k}$ are positive on $(t_i \,..\, t_{i+k})$, it follows from (43) by induction on k that $B_{i,k}$ *is positive on* $(t_i \,..\, t_{i+k})$.

Since at most k of the $B_{i,k}$ are nonzero at any one point $x \in \mathbb{R}$, the definition

$$\sum a_i B_{i,k} : \mathbb{R} \to \mathbb{R} : x \mapsto \sum a_i B_{i,k}(x)$$

of $\sum_i a_i B_{i,k}$ as a pointwise sum makes sense for arbitrary a even when the sum has infinitely many terms. We call any such function a **spline of order k with knot sequence $\mathbf{t}$** and denote the collection of all such functions by

$$S_{k,\mathbf{t}}.$$

We deduce from the recurrence relation that

$$\sum a_i B_{i,k} = \sum \big(a_i \omega_{i,k} + a_{i-1}(1-\omega_{i,k})\big) B_{i,k-1}, \tag{45}$$

showing that the coefficients on the right are affine combinations of neighboring coefficients on the left.

Before proceeding with this, note the following technical difficulty. If $\mathbf{t}$ has a first knot, e.g., $\mathbf{t} = (t_1, \ldots)$, then the sum on the left in (45) starts with $i = 1$, while the sum on the right can only start at $i = 2$ since the coefficient a_{i-1} appears in it. To avoid such difficulty, we agree in this case (and the corresponding case when $\mathbf{t}$ has a last entry) to extend $\mathbf{t}$ in any way whatsoever to a bi-infinite knot sequence, denoting the extension again by $\mathbf{t}$. However, this increases the number of available B-splines, hence also increases the spline space. Since we are still interested only in our original spline space, we further agree to make use of results obtained from the extended situation only to the extent that they don't explicitly involve any of the additional knots. We can be sure none of the additional knots matters if we restrict attention to the largest interval not intersected by the interior of the support of any of the additional B-splines. We call this the **basic interval for $S_{k,\mathbf{t}}$** and denote it by

$$I_{k,\mathbf{t}} := (\mathbf{t}_- \,..\, \mathbf{t}_+)$$

with

$$\mathbf{t}_- := \begin{cases} t_k, & \text{if } \mathbf{t} = (t_1, \ldots); \\ \inf_i t_i, & \text{otherwise,} \end{cases} \qquad \mathbf{t}_+ := \begin{cases} t_{n+1}, & \text{if } \mathbf{t} = (\cdots, t_{n+k}); \\ \sup_i t_i, & \text{otherwise,} \end{cases}$$

In practice, the knot sequence is finite, having both a first and a last knot. In that case, one chooses $I_{k,\mathbf{t}}$ to be closed, and this is fine for the left endpoint, since the definition of B-splines makes them all continuous from the right. For this to work properly at the right endpoint, one modifies the above definition of B-splines to make them continuous from the left at the right endpoint of $I_{k,\mathbf{t}}$.

In summary, even if the given knot sequence is not biinfinite, we may always assume it to be biinfinite as long as we apply the results so obtained only to functions on the basic interval $I_{k,\mathbf{t}}$ determined by the given knot sequence.

With this, consider the special sequence

$$a_i := \psi_{i,k}(\tau) := (t_{i+1} - \tau) \cdots (t_{i+k-1} - \tau)$$

(with $\tau \in \mathbb{R}$). We find for $B_{i,k-1} \neq 0$, i.e., for $t_i < t_{i+k-1}$, that

$$\begin{aligned} a_i \omega_{i,k} + a_{i-1}(1-\omega_{i,k}) &= \psi_{i,k-1}(\tau)\Big((t_{i+k-1} - \tau)\omega_{i,k} + (t_i - \tau)(1-\omega_{i,k})\Big) \\ &= \psi_{i,k-1}(\tau)(\cdot - \tau) \end{aligned}$$

since $f(t_{i+k-1})\omega_{i.k} + f(t_i)(1-\omega_{i,k})$ is the straight line which agrees with f at t_{i+k-1} and t_i. This shows that

$$\sum \psi_{i,k}(\tau)B_{i,k} = (\cdot - \tau)\sum \psi_{i,k-1}(\tau)B_{i,k-1}, \tag{46}$$

hence, by induction, that

$$\sum \psi_{i,k}(\tau)B_{i,k} = (\cdot - \tau)^{k-1}\sum \psi_{i,1}(\tau)B_{i,1}. \tag{47}$$

This proves the following identity.

(48) Marsden's Identity. *For any* $\tau \in \mathbb{R}$,

$$(\cdot - \tau)^{k-1} = \sum_i \psi_{i,k}(\tau)B_{i,k} \tag{49}$$

on $I_{k,\mathbf{t}}$, *with* $\psi_{i,k}(\tau) := (t_{i+1}-\tau)\cdots(t_{i+k-1}-\tau)$.

Since τ in (49) is arbitrary, it follows that $S_{k,\mathbf{t}}$ contains all polynomials of degree $< k$ (restricted to $I_{k,\mathbf{t}}$). More than that, we can even give an explicit expression for the required coefficients, as follows.

By differentiating (49) with respect to τ, we obtain the identities

$$\frac{(\cdot - \tau)^{k-\nu}}{(k-\nu)!} = \sum_i \frac{(-D)^{\nu-1}\psi_{i,k}(\tau)}{(k-1)!}\, B_{i,k}, \quad \nu > 0. \tag{50}$$

On using this identity in the Taylor formula

$$p = \sum_{\nu=1}^{k} \frac{(\cdot - \tau)^{k-\nu}}{(k-\nu)!} D^{k-\nu}p(\tau)$$

for a polynomial p of degree $< k$, we conclude that any such polynomial can be written in the form

$$p = \sum_i \lambda_{i,k}p\; B_{i,k}, \tag{51}$$

with $\lambda_{i,k}$ given by the rule

$$\lambda_{i,k}f := \lambda_{i,k,\mathbf{t}}f := \sum_{\nu=1}^{k} \frac{(-D)^{\nu-1}\psi_{i,k}(\tau)}{(k-1)!}\; D^{k-\nu}f(\tau). \tag{52}$$

Here are two special cases of particular interest. For $p = 1$, we get

$$1 = \sum_i B_{i,k} \tag{53}$$

since $D^{k-1}\psi_{i,k} = (-1)^{k-1}(k-1)!$, and this shows that the $B_{i,k}$ form a **partition of unity**. Further, since $D^{k-2}\psi_{i,k}$ is a linear polynomial which vanishes at

$$t_i^* := (t_{i+1} + \cdots + t_{i+k-1})/(k-1),$$

$$\ell = \sum_i \ell(t_i^*) B_{i,k}, \qquad \ell \in \Pi_1. \tag{54}$$

The identity (50) also gives us various **piecewise** polynomials contained in $S_{k,\mathbf{t}}$: Since $t_i < t_j < t_{i+k}$ implies that $D^{\nu-1}\psi_{i,k}(t_j) = 0$ in case $\nu \le \#t_j := \#\{s : t_s = t_j\}$, the choice $\tau = t_j$ in (50) leaves only terms with support either entirely to the left or else entirely to the right of t_j. This implies that

$$\frac{(\cdot - t_j)_+^{k-\nu}}{(k-\nu)!} = \sum_{i\ge j} \frac{(-D)^{\nu-1}\psi_{i,k}(t_j)}{(k-1)!}\ B_{i,k}, \quad 0 < \nu \le \#t_j. \tag{55}$$

Consequently,

$$(\cdot - t_j)_+^{k-\nu} \ \in\ S_{k,\mathbf{t}} \ \text{ for } \ 1 \le \nu \le \#t_j, \tag{56}$$

(on $I_{k,\mathbf{t}}$).

(57) Theorem. *If $t_i < t_{i+k}$, then the B-spline sequence $(B_{i,k} : i)$ is linearly independent and the space $S_{k,\mathbf{t}}$ coincides with the space $S := \Pi^{\rho}_{<k,\mathbf{t}}$ of all piecewise polynomials of degree $< k$ with breakpoints t_i that are $\rho(i) := k - 1 - \#t_i$ times continuously differentiable at t_i, all i. In particular, each $f \in S_{k,\mathbf{t}}$ is in $C^{(\rho(i))}$ near t_i, all i.*

Proof: It is sufficient to prove that, for any finite interval $I := [a\,..\,b]$, the restriction $S_{|I}$ of the space S to the interval I coincides with the restriction of $S_{k,\mathbf{t}}$ to that interval. The latter space is spanned by all the B-splines having some support in I, i.e., all $B_{i,k}$ with $(t_i\,..\,t_{i+k}) \cap I \ne \emptyset$. The space $S_{|I}$ has a basis consisting of the functions

$$(\cdot - a)^{k-\nu},\ \nu = 1,\cdots,k;\ (\cdot - t_i)_+^{k-\nu},\ \nu = 1,\cdots,\#t_i,\ \text{for } \ a < t_i < b, \tag{58}$$

This follows from the observations that (i) the sequence of functions in (58) is linearly independent; and (ii) a piecewise polynomial function f with a breakpoint at t_i which is $k - 1 - \#t_i$ times continuously differentiable there can be written uniquely as

$$f = p + \sum_{\nu=1}^{\#t_i} a_\nu (\cdot - t_i)_+^{k-\nu},$$

with p a suitable polynomial of degree $< k$ and suitable coefficients a_ν. Since each of the functions in (58) lies in $S_{k,\mathbf{t}}$, by (50) and (56), we conclude that

$$S_{|I} \subset (S_{k,\mathbf{t}})_{|I}. \tag{59}$$

On the other hand, the dimension of $S_{|I}$, i.e., the number of functions in (58), equals the number of B-splines with some support in I (since it equals $k + \sum_{a<t_i<b} \#t_i$), hence is an upper bound on the dimension of $(S_{k,\mathbf{t}})_{|I}$. This implies that equality must hold in (59), and that the set of B-splines having some support in I must be linearly independent over I. $\square$

(60) Corollary. *The sequence $(B_{j,k} : B_{j,k}|_I \neq 0)$ of B-splines having some support in a given (proper) interval I is linearly independent over that interval.*

(61) Corollary. *For all $p \in \Pi_{<k}$, $D_\tau \lambda_{j,k} p = 0$.*

Indeed, since the B-spline sequence is linearly independent, the coefficients $\lambda_{i,k}p$ appearing in (51) are uniquely determined, hence cannot change with τ. When convenient later on, we will choose τ in (52) in dependence on i, i.e., as τ_i.

(62) Corollary. *If $\hat{\mathbf{t}}$ is a refinement of the knot sequence $\mathbf{t}$, then $S_{k,\mathbf{t}} \subset S_{k,\hat{\mathbf{t}}}$.*

(63) Corollary. *The derivative of a spline in $S_{k,\mathbf{t}}$ is a spline of degree $< k-1$ with respect to the same knot sequence, i.e., $DS_{k,\mathbf{t}} \subseteq S_{k-1,\mathbf{t}}$.*

(64) Remark The word 'derivative' is used here in the pp sense: The 'derivative' of a pp f is, by definition, the pp with the same breakpoint sequence whose polynomial pieces are the first derivative of the corresponding polynomial pieces of f. This makes it possible in (63) to ignore the possibility that $t_i = t_{i+k-1}$ for some i, hence the elements of $S_{k,\mathbf{t}}$ will, in general, fail to be differentiable at such a t_i.

More generally, here and elsewhere, we do not exclude the possibility that some of the $B_{i,k}$ are trivial, i.e., that $t_i = t_{i+k}$ for some i. However, if we were to interpret $B_{i,k,\mathbf{t}}$ as *distributions*, we would have to proceed with more caution. For, by (40),

$$\lim_{t_i,\ldots,t_{i+k}\to\tau} M(\cdot|t_i,\ldots,t_{i+k}) \;=\; [\tau] \tag{65}$$

as a distribution, and that limit is quite different from 0. □

The identity (51) can be extended to all spline functions. For this, we agree, consistent with (41b), that all derivatives in (52) are to be taken as limits from the right in case τ coincides with a knot.

(66) Theorem. *If $\tau = \tau_i$ in definition (52) of $\lambda_{i,k}$ is chosen in the interval $[t_i \,..\, t_{i+k})$, then, with the* **understanding** *that $D^\nu f(\tau) := D^\nu f(\tau+)$ for any ν and any pp f,*

$$\lambda_{i,k}\Big(\sum_j a_j B_{j,k}\Big) = a_i. \tag{67}$$

It is remarkable that τ can be chosen arbitrarily in the interval $[t_i \,..\, t_{i+k})$. The reason behind this is Corollary (61).

Proof: Assume that $\tau \in [t_i \,..\, t_{i+k})$, hence $\tau \in [t_l \,..\, t_{l+1}) \subset [t_i \,..\, t_{i+k})$ for some l, and let p_j be the polynomial that agrees with $B_{j,k}$ on $(t_l \,..\, t_{l+1})$. Then

$$\lambda_{i,k}B_{j,k} = \lambda_{i,k}p_j.$$

On the other hand,

$$p_j = \sum_{m=l+1-k}^{l} \lambda_{m,k}p_j \; p_m,$$

since this holds by (51) on $[t_l \,..\, t_{l+1})$, while, by Corollary (60) or directly from (51), $(p_{l+1-k},\cdots,p_l)$ is linearly independent. Therefore necessarily $\lambda_{i,k}B_{j,k} = \lambda_{i,k}p_j$ equals 1 if $i = j$ and 0 otherwise. □

use of the dual functionals: differentiation; dependence on knots

Because of (67), the functionals $\lambda_{i,k}$ are called **the dual functionals for the B-splines**. Strictly speaking, there are many such functionals (of which more anon), but these particular ones have turned out to be quite useful in various contexts. Here are several examples.

Compare

$$\lambda_{i,k} f = \sum_{\nu=1}^{k} \frac{(-D)^{\nu-1}\psi_{i,k}(\tau)}{(k-1)!}\, D^{k-\nu} f(\tau) \tag{68}$$

with

$$\lambda_{i,k-1} Df = \sum_{\nu=1}^{k-1} \frac{(-D)^{\nu-1}\psi_{i,k-1}(\tau)}{(k-2)!}\, D^{k-1-\nu} Df(\tau).$$

Since $(t_i - \cdot)\psi_{i,k-1} = \psi_{i-1,k}$ and $(t_{i+k-1} - \cdot)\psi_{i,k-1} = \psi_{i,k}$, subtraction of the latter from the former gives

$$(t_{i+k-1} - t_i)\psi_{i,k-1} = \psi_{i,k} - \psi_{i-1,k}.$$

Hence

$$\lambda_{i,k-1} D = \frac{\lambda_{i,k} - \lambda_{i-1,k}}{(t_{i+k-1} - t_i)/(k-1)}. \tag{69}$$

Consequently, we get the **differentiation formula**

$$D \sum_i a_i B_{i,k} = \sum_i \frac{a_i - a_{i-1}}{(t_{i+k-1} - t_i)/(k-1)}\, B_{i,k-1}. \tag{70}$$

We now consider how $\lambda_{i,k}$ depends on the knot sequence $\mathbf{t}$. Perhaps surprisingly, although $B_{i,k}$ involves the knots $t_i, \cdots, t_{i+k}$, $\lambda_{i,k}$ only depends on the 'interior' knots, $t_{i+1}, \ldots, t_{i+k-1}$. Further, since $\lambda_{i,k}$ depends *linearly* on $\psi_{i,k}$, it depends *affinely* and *symmetrically* on the points $t_{i+1}, \cdots, t_{i+k-1}$. Indeed, for any α, x, y,

$$((\alpha x + (1-\alpha)y) - \cdot) = \alpha\,(x - \cdot) + (1-\alpha)(y - \cdot).$$

Hence, with

$$\lambda_k : \mathrm{I\!R}^{k-1} \to (C^{(k-1)})' : t \mapsto \lambda_{0,k},$$

we have

$$\lambda_k(\alpha x + (1-\alpha)y, s_2, \ldots, s_{k-1}) = \alpha\lambda_k(x, s_2, \ldots, s_{k-1}) + (1-\alpha)\lambda_k(y, s_2, \ldots, s_{k-1})$$

as well as

$$\lambda_k(s) = \lambda_k(s \circ \sigma)$$

for any permutation σ of order $k-1$.

Consider, in particular, $t_1 = \cdots = t_{k-1} = x$. Then $\psi_{0,k} = (x - \cdot)^{k-1}$, hence

$$(-D)^{\nu-1}\psi_{0,k}(\tau)/(k-1)! = (x-\tau)^{k-\nu}/(k-\nu)!$$

and therefore

$$\lambda_k(x,\ldots,x) = \sum_{\nu=1}^{k} \frac{(x-\tau)^{k-\nu}}{(k-\nu)!}[\tau]D^{k-\nu}.$$

In particular

$$\lambda_k(x,\ldots,x)p = p(x), \quad p \in \Pi_{<k}.$$

This meshes entirely with the fact that, if $x = t_{i+1} = \cdots = t_{i+k-1}$, then $B_{i,k}$ is the only $B_{j,k}$ which is nonzero at x, hence $B_{i,k}(x)$ must be 1 (since the $B_{j,k}$ form a partition of unity), and therefore

$$\sum_j a_j B_{j,k}(x) = a_i$$

in this case.

side issue: blossoms

It is worthwhile to point out that, associated with each $p \in \Pi_r$, there is a unique symmetric r-affine form called its **polar form** (in Algebra) or its **blossom** (in CAGD), denoted therefore here by

$$\overset{\omega}{p},$$

for which

$$\forall\{x \in \mathbb{R}\}\ p(x) = \overset{\omega}{p}(x,\ldots,x).$$

If $p = \sum_j ()^j c_j \in \Pi_r$, then

$$\overset{\omega}{p}(s_1,\ldots,s_r) := \sum_j \sum \{\prod_{i\in I} s_i : I \in \binom{\{1,\ldots,r\}}{j}\}/\binom{r}{j}\, c_j.$$

We deduce from the above that

$$\overset{\omega}{p}(t_1,\ldots,t_{k-1}) = \lambda_k(t_1,\ldots,t_{k-1})\,p, \quad p \in \Pi_{<k}.$$

In particular, the ith B-spline coefficient of a kth order spline with knot sequence $\mathbf{t}$ is the value at $(t_{i+1},\ldots,t_{i+k-1})$ of the blossom of *any* of the k polynomial pieces associated with the intervals $[t_j \,..\, t_{j+1})$, $j = i,\ldots,i+k-1$. This observation was made, in language incomprehensible to the uninitiated, by de Casteljau in the sixties. It was discovered independently and made plain (and given the nice name of 'blossom') by Lyle Ramshaw in the early eighties.

knot insertion

This has led to the *evaluation of a spline by knot sequence refinement*. Such a refinement can always be reached adding one knot at a time. So, suppose that the knot sequence $\hat{\mathbf{t}}$ has been obtained from the knot sequence $\mathbf{t}$ by the insertion of just one additional term, the point x say. Thus, for some j,

$$\hat{t}_i = \begin{cases} t_i, & \text{for } i \le j; \\ x, & \text{for } i = j+1; \\ t_{i-1}, & \text{for } i > j+1. \end{cases}$$

We saw already that, with

$$x = \alpha t_{i+k-1} + (1-\alpha) t_i,$$

i.e., with

$$\alpha = \omega(x) := \omega_{i,k}(x),$$

we have

$$\lambda_k(x, t_{i+1}, \ldots, t_{i+k-2}) = \alpha \lambda_k(t_{i+1}, \ldots, t_{i+k-1}) + (1-\alpha)\lambda_k(t_i, \ldots, t_{i+k-2}).$$

While this is true for arbitrary i, it matters here only when $i < j+1 < i+k$. Altogether, we find that

$$\hat{\lambda}_{i,k} = (1 - \widetilde{\omega}_{i,k}(x))\lambda_{i-1,k} + \widetilde{\omega}_{i,k}(x)\lambda_{i,k}, \quad \text{all } i,$$

with

(71) $$\widetilde{\omega}_{i,k}(x) := \max\{0, \min\{1, \omega_{i,k}(x)\}\}.$$

Correspondingly,

$$\sum_i a_i B_{i,k} = \sum_i \hat{a}_i \hat{B}_{i,k},$$

with

$$\hat{a}_i = (1 - \widetilde{\omega}_{i,k}(x))a_{i-1} + \widetilde{\omega}_{i,k}(x)a_i, \quad \text{all } i.$$

Note that this is exactly the way we computed coefficients by recurrence in (45). This intimate connection between the recurrence relation and knot insertion was first observed by Wolfgang Boehm. Note further that $k-1$-fold insertion of the knot x produces eventually a knot sequence $\tilde{\mathbf{t}}$, in which all the interior knots for $\tilde{B}_{j,k}$ are equal to x, hence, correspondingly,

$$\tilde{a}_j = (\sum_i a_i B_{i,k})(x).$$

Since also t_i^* is $k-1$-affine, it follows that also

$$\hat{t}_i^* = (1 - \widetilde{\omega}_{i,k}(x))t_{i-1}^* + \widetilde{\omega}_{i,k}(x)t_i^*.$$

This provides knot insertion with a very pretty geometric interpretation in terms of the **control polygon**

$$C_{k,\mathbf{t}} f$$

of the spline $f := \sum_j a_j B_{j,k}$. This is, by definition, the broken line whose vertices are the **control points** for $\sum_j a_j B_{j,k}$, i.e., the points

$$(P_i := (t_i^*, a_i) \in \mathbb{R}^2 : i).$$

We have just proved the following:

(72) Proposition. *If $\hat{\mathbf{t}}$ is obtained from $\mathbf{t}$ by the insertion of one additional knot, then, for any $f \in S_{k,\mathbf{t}}$, $C_{k,\hat{\mathbf{t}}}f$ interpolates, at its breakpoints, to $C_{k,\mathbf{t}}f$ (and is thereby uniquely determined).*

Below are two figures, to illustrate this geometric interpretation of knot insertion.

This implies that the map $a \mapsto f_a := \sum_i a_i B_{i,k}$ is **variation-diminishing**, meaning that f_a has no more variations in sign than does the sequence a, in a sense to be made precise next.

variation diminution

For a real sequence a without any zero entries,

$$S(a) := \#\{i : a(i)a(i+1) < 0\}$$

denotes the number of **sign changes** in it. It is less clear what this number should be in case a has some zero entries. The maximum number of sign changes obtainable in such a sequence by an appropriate choice in the sign of any zero entry is called the number of **weak sign change**s in it, denoted

$$S^+(a).$$

The minimum number so obtainable is denoted by

$$S^-(a)$$

and called the number of **strong sign change**s. If we think of choosing the sign of the zero entries as making very small perturbations of these entries, then we see that strong sign changes are impervious to small perturbations, while weak sign changes are not.

The following Lemma is immediate (from the geometric picture of knot insertion).

(76) Lemma (Lane, Riesenfeld). *If $\tilde{a}$ is the B-spline coefficient sequence obtained from the sequence a by the insertion of (zero or more) knots, then*

$$S^-(\tilde{a}) \le S^-(a).$$

If f is a function on an interval (including the interval $\mathbb{R}$), one defines

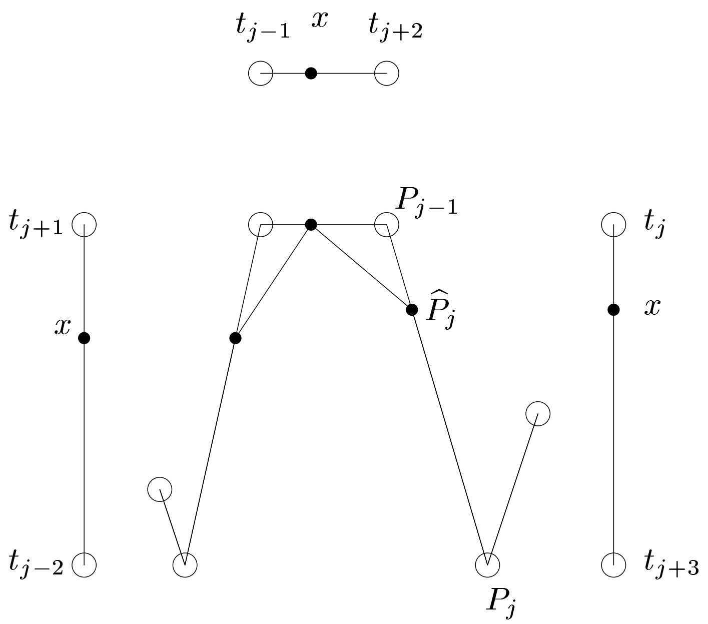
$$S^-(f) := \sup_{x_1 < \cdots < x_r} S^-(f(x_1), \ldots, f(x_r)).$$

(77) Proposition. $S^-(\sum_j a_j B_j) \le S^-(a)$.

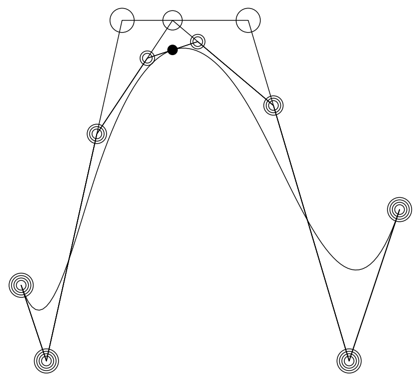
Proof: Insert into $\mathbf{t}$ each of the entries in a given increasing sequence (x_i) enough times to have them appear in the resulting knot sequence $\tilde{\mathbf{t}}$ at least $k-1$ times. Then $(f(x_1), \ldots, f(x_r))$ is a subsequence of the resulting B-spline coefficient sequence $\tilde{a}$, hence

$$S^-(f(x_1), \ldots, f(x_r)) \le S^-(\tilde{a}) \le S^-(a).$$

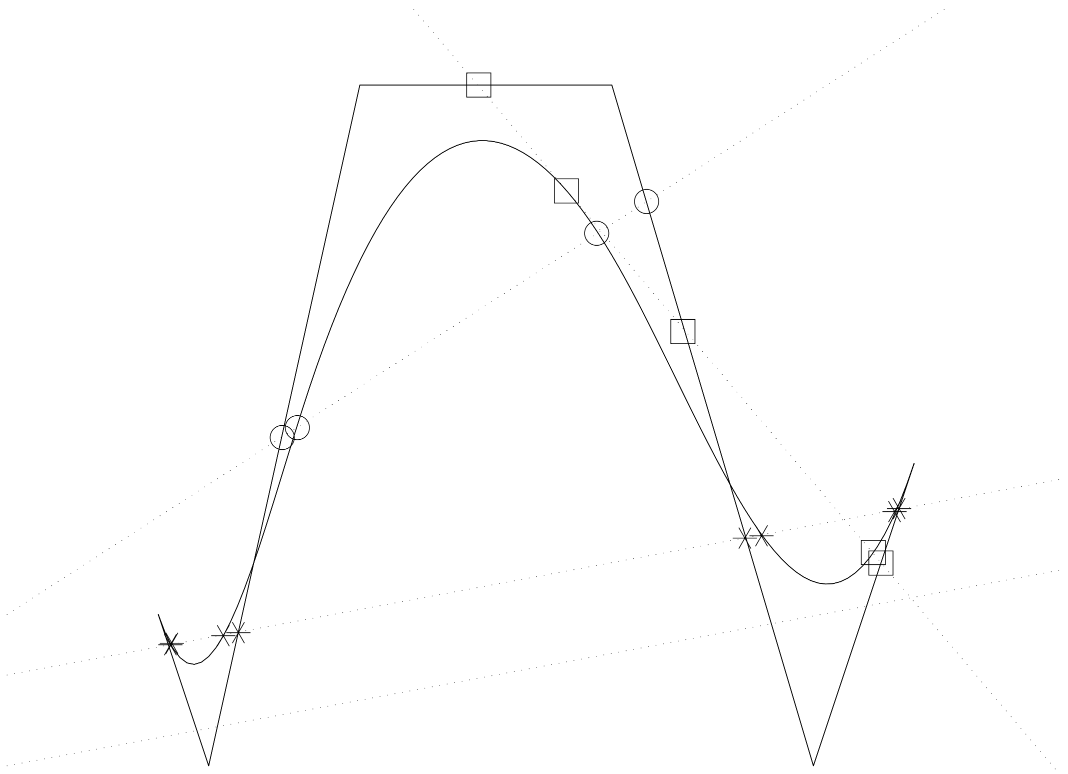
□



(73) Figure Insertion of $x = 2$ into the knot sequence $\mathbf{t} = (0,0,0,0,\mathbf{1},\mathbf{3},5,5,5,5)$, with $k = 4$.



(74) Figure The cubic spline and its control polygon from Figure (73) and the sequence of control polygons generated by three-fold insertion of the same knot. (The finest control polygon differs from its predecessor only by an additional vertex point.)



(75) Figure A cubic spline, its control polygon, and various straight lines intersecting them. The control polygon *exaggerates* the shape of the spline. The spline crossings are bracketed by the control polygon crossings.

(78) Definition. Schoenberg's (spline) operator *is, by definition, the linear map* $V = V_{k,\mathbf{t}}$ *given by the rule*

$$Vg := \sum_j g(t_j^*)B_j.$$

It is usually defined only when $\#t_i < k$, *all* i.

(79) Proposition. *Schoenberg's spline operator is* **variation-diminishing**. *Precisely, for any* g *and any* $\ell \in \Pi_1$,

(80) $$S^-(Vg - \ell) \le S^-(g - \ell).$$

Even more precisely,

(81) $$D^r g \ge 0 \quad \Longrightarrow \quad D^r Vg \ge 0, \quad r = 0, 1, 2,$$

and this holds even 'locally'.

Proof: By (54), $V\ell = \ell$ (on $I_{k,\mathbf{t}}$) for all $\ell \in \Pi_1$, hence $V(f-\ell) = Vf - \ell$, therefore, by Proposition (77), and by the strict increase in the sequence $t_i^* : i = 1, \ldots, n$ $S^-(Vf - \ell) \le S^-((f-\ell)(t_i^*) : i) \le S^-(f - \ell)$.

However, (80) by itself fails to imply the more precise statement (81), which follows from the nonnegativity of the B-splines along with the observation that

$$DVg = \sum_j [t_{j-1}^*, t_j^*]g\, B_{j,k-1},$$

hence

$$D^2Vg = \sum_j \frac{[t_{j-1}^*, t_j^*]g - [t_{j-2}^*, t_{j-1}^*]g}{t_{j+k-2} - t_j}\, B_{j,k-2}.$$

□

zeros of a spline, counting multiplicities

There is a complete theory that provides an upper bound on the number of zeros of spline, even counting multiplicity, in terms of the number of sign changes (strong and/or weak) in its B-spline coefficients, with **multiplicity** of a zero defined as the maximal number of distinct nearby zeros in a nearby spline (from the same spline space).

For the theory to give useful information, one has to assign a finite multiplicity to zero *intervals*, and this adds a further complication.

Multiplicity considerations are important when one wishes to consider osculatory spline interpolation, i.e., interpolation at possibly repeated points. Since I will not get to that topic in this course, I am content to state and prove only the following very useful proposition.

(82) Proposition. *If $f = \sum_j a_j B_{j,k,\mathbf{t}}$ vanishes at $x_1 < \cdots < x_r$, while $f^{\mathbf{t}} := \sum_j |a_j| B_{j,k,\mathbf{t}}$ does not, then $S^-(a) > r - 1$.*

Proof: Since $f^{\mathbf{t}}(x_i) > 0$, while $f(x_i) = 0$, the sequence $(a_j B_j(x_i) : B_j(x_i) \ne 0)$ must have at least one strong sign change, hence, so must the sequence $(a_j : B_j(x_i) \ne 0)$, by the nonnegativity of the B-splines. This gives altogether r strong sign changes in a, provided we can be sure that different x_i generate different sign changes. Off-hand, this may not be so, but can be guaranteed by inserting each of the points $(x_i + x_{i+1})/2$, $i = 1, \ldots, r-1$, into the knot sequence k times. If $\tilde{\mathbf{t}}$ and $\tilde{a}$ are the resulting knot and coefficient sequences, respectively, then still $f^{\tilde{\mathbf{t}}} > 0 = f$ on the x_i (since, from $f^{\mathbf{t}}(x_i) > 0$, we know that x_i is an isolated zero of f), while now $\{j : \tilde{B}_j(x_i) \ne 0\} \cap \{j : \tilde{B}_j(x_h) \ne 0\} = \emptyset$ for $i \ne h$, hence

$$S^-(a) \ge S^-(\tilde{a}) \ge r.$$

□

spline interpolation

We consider **spline collocation**, i.e., interpolation from $S_{k,\mathbf{t}}$ at the increasing sequence $\mathbf{x} = (x_i)$ of points. We consider this under the assumption that $\mathbf{t} = (t_i : i = 1, \ldots, n+k)$, with $\#t_i < k$, all i, hence $S_{k,\mathbf{t}} \subset C(\mathbb{R})$. This simplifying assumption avoids discussion otherwise needed in case some x_j agrees with a knot of multiplicity $\ge k$, in which case one would have to specify, in addition, whether it is x_j- or x_j+ one wants. With the assumption, none of the B-splines $B_{i,k}$ is trivial, hence $\dim S_{k,\mathbf{t}} = n$. For this reason, we assume, more precisely, that

$$\mathbf{x} = (x_1 < \cdots < x_n).$$

Hence, for given g, we are seeking $f \in S_{k,\mathbf{t}}$ with $f = g$ on x. Equivalently, we are seeking a solution to the linear system $A? = g|_{\mathbf{x}}$, with

$$A := (B_{j,k}(x_i) : i, j = 1, \ldots, n)$$

the so-called **collocation matrix**. There is exactly one interpolant to a given g iff A is invertible iff A is 1-1.

(83) Proposition. *If $A = (B_{j,k}(x_i))$ is invertible, then*

$$t_i < x_i < t_{i+k}, \quad \forall i.$$

Proof: If, for some i, $t_{i+k} \le x_i$, then the first i columns of A have nonzero entries only in the first $i-1$ rows, hence A cannot be invertible. Again, if $x_i \le t_i$, then columns $i, \ldots, n$ of A have nonzero entries only in rows $i+1, \ldots, n$, hence A cannot be invertible. $\square$

Note that the argument used nothing more than the fact that both the sequence $\mathbf{x}$ and the sequence of the supports of the B-splines are increasing. In particular, we have proved

(84) Corollary. *If $(B_{m_j,k}(s_i) : i, j = 1, \ldots, r)$ is invertible, with both (m_j) and (s_i) increasing, then $B_{m_i}(s_i) \ne 0$, all i.*

In other words, such a matrix is invertible only if its diagonal entries are nonzero. As it turns out, the converse also holds. The converse of the Proposition is

(85) Schoenberg-Whitney Theorem. *Let $\mathbf{t} = (t_1, \ldots, t_{n+k})$ with $B_{j,k} \ne 0$, all j, and let $x_1 < \cdots < x_n$. Then, $A = (B_{j,k}(x_i))$ is invertible iff $t_i < x_i < t_{i+k}$, all i.*

Proof: We only need to prove the 'if'. Since A is square, it is sufficient to prove that $Aa = 0$ implies $a = 0$. Consider $f = \sum_j a_j B_{j,k}$ with $Aa = 0$. If $a \ne 0$, then, by Corollary (60), $f \ne 0$. Let $I = (\alpha \,.\,.\, \beta)$ be a maximal open interval in

$$\operatorname{supp} \sum_j |a_j| B_{j,k} = \bigcup \{(t_j \,.\,.\, t_{j+k}) : a_j \ne 0\}.$$

It follows that $I = (t_\nu \,..\, t_{\mu+k})$ for some $1 \le \nu \le \mu \le n$, and that

$$f = f_I := \sum_{j=\nu}^{\mu} a_j B_{j,k} \qquad \text{on } \ I.$$

In particular, f_I has the distinct zeros $x_\nu, \ldots, x_\mu$, therefore, by Proposition (82),

$$S^-(a_\nu, \ldots, a_\mu) \ge \mu + 1 - \nu,$$

which is nonsense. □

The Schoenberg-Whitney Theorem has been generalized in at least two directions: (i) permission of coincidences in the sequence (x_i) correspondingly to osculatory interpolation; and (ii) consideration of a subsequence $(B_{m_j} : j = 1, \ldots, n)$ instead of the whole sequence (for a suitably longer knot sequence).

Any $\mathbf{x} = (x_1 < \cdots < x_n)$ satisfying the Schoenberg-Whitney conditions for $S_{k,\mathbf{t}}$ gives rise to the corresponding projector $P_\mathbf{x}$ that associates $g \in C$ with the unique $f = P_\mathbf{x} g \in S_{k,\mathbf{t}}$ that agrees with g at $\mathbf{x}$. Since $P_\mathbf{x}$ is a linear projector, we have

$$\operatorname{dist}(g, S_{k,\mathbf{t}}) \le \|g - P_\mathbf{x} g\| \le (1 + \|P_\mathbf{x}\|) \operatorname{dist}(g, S_{k,\mathbf{t}}),$$

hence $P_\mathbf{x}$ s a candidate for a “good” approximation scheme from $S_{k,\mathbf{t}}$ to the extent that $\|P_\mathbf{x}\|$ is “small”, i.e., not much larger than 1. We pursue this question in the context of the uniform norm, i.e., in the space

$$X = C[t_1 \,..\, t_{n+k}].$$

Since $P_\mathbf{x} g = \sum_j a_j B_j$ with $a = A^{-1}(g|_\mathbf{x})$, while (B_j) is a positive partition of unity,

(86) $$\|P_\mathbf{x} g\|_\infty \le \|A^{-1}(g|_\mathbf{x})\|_\infty \le \|A^{-1}\|_\infty \|(g|_\mathbf{x})\|_\infty \le \|A^{-1}\|_\infty \|g\|_\infty,$$

with $\|A^{-1}\|_\infty = \max_i \sum |A^{-1}(i,j)|$. Hence

$$\|P_\mathbf{x}\| \le \|A^{-1}\|_\infty.$$

While this bound is not sharp (in general), it is the only bound readily available. Hence, in search for a good approximation scheme from $S_{k,\mathbf{t}}$, we now look for $\mathbf{x}$ so that $\|A^{-1}\|_\infty$ is as small as possible. This search is aided by the knowledge that our collocation matrix A is totally positive, to be established next.

total positivity

We recall that an (m, n)-matrix C is **totally positive** if, for any strictly increasing index sets $I = (i_1 < \cdots < i_r)$ in $\{1, \ldots, m\}$ and and $J = (j_1 < \cdots < j_r)$ in $\{1, \ldots, n\}$, the determinant $\det C(I, J)$ of the (r, r)-submatrix

$$C(I, J) := \big(C(i_p, j_q) : p, q = 1, \ldots, r\big)$$

is nonnegative. The most immediately important fact concerning total positivity is the following.

(87) Fact. *If C is invertible and totally positive, then its inverse is* **checkerboard**, *meaning that $C^{-1}(i, j)(-1)^{i-j} \geq 0$, all i, j.*

Proof: By Cramer's rule,

$$C^{-1}(i, j) = (-1)^{i-j} \det C(\backslash j, \backslash i)/ \det C.$$

□

(88) Theorem. *For any $x_1 < \cdots < x_n$, any $k > 0$, and any knot sequence $\mathbf{t} = (t_i : i = 1, \ldots, n + k)$, the collocation matrix*

$$A := \big(B_j(x_i) : i, j = 1, \ldots, n\big)$$

is totally positive.

Proof: If $\hat{\mathbf{t}}$ is obtained from $\mathbf{t}$ by the insertion of just one knot, and $\hat{B}_j := B_{j,k,\hat{\mathbf{t}}}$, all j, then

$$B_j = (1 - \alpha_{j+1})\hat{B}_{j+1} + \alpha_j \hat{B}_j,$$

with all $\alpha_j \in [0\,.\,.\,1]$. Since the determinant of a matrix is a linear function of the columns of that matrix, we have, e.g.,

$$\det[\cdots, B_j(x), \ldots] = (1 - \alpha_{j+1}) \det[\cdots, \hat{B}_{j+1}(x), \ldots] + \alpha_j \det[\cdots, \hat{B}_j(x), \ldots],$$

with $\cdots$ unchanged in their respective places. It follows that, for any I, J,

$$\det A(I, J) = \sum_H \gamma_H \det \hat{A}(I, H),$$

with all the $\gamma_H \geq 0$, and the sum, offhand, over certain nondecreasing sequences, since only neighboring columns of $\hat{A}$ participate in a column of A. However, we may omit all H which are not strictly increasing, since the corresponding determinant is trivially zero. Therefore,

$$\det A(I, J) = \sum_H \gamma_H \det \hat{A}(I, H),$$

with the $\gamma_H \geq 0$ and all H strictly increasing.

Now insert each of the x_i enough times so that the resulting refined knot sequence $\tilde{\mathbf{t}}$ contains each x_i exactly $k-1$ times. By induction, we have

$$\det A(I,J) = \sum_H \gamma_H \det \tilde{A}(I,H),$$

with the $\gamma_H \ge 0$ and all H strictly increasing. However, in each row of $\tilde{A}$, there is exactly one nonzero entry, namely the entry belonging to $\tilde{B}_j$ with $\tilde{t}_j < x_i = \tilde{t}_{j+1}$, and that entry equals 1. Further, the sole nonzero entry in subsequent rows occurs to the right of this one. Thus $\tilde{A}(I,H)$ has a nonzero entry in every row iff $A(I,H)$ is the identity matrix. Consequently, $\det \tilde{A}(I,H) \ge 0$ for any H, and therefore $\det A(I,J) \ge 0$. □

spline interpolation (cont.)

We continue with the spline interpolation setup introduced earlier, considering interpolation at $\mathbf{x} = (x_1 < \cdots < x_n)$ from $S_{k,\mathbf{t}}$ with $\mathbf{t} = (t_1 \le \cdots \le t_{n+k})$ and $\#t_i < k$, all i.

The fact that the collocation matrix

$$A_{\mathbf{x}} := (B_j(x_i))$$

is totally positive implies that

$$\|A_{\mathbf{x}}^{-1}\|_\infty = \max_i \sum_j |A_{\mathbf{x}}^{-1}(i,j)| = \max_i \sum_j A_{\mathbf{x}}^{-1}(i,j)(-1)^{i-j},$$

hence

$$\|P_{\mathbf{x}}\|_\infty \le \|A_{\mathbf{x}}^{-1}((-1)^{n-i} : i=1,\ldots,n)\|_\infty = \|a^{\mathbf{x}}\|_\infty,$$

with

$$f_{\mathbf{x}} := \sum_j a_j^{\mathbf{x}} B_j$$

the unique element of $S_{k,\mathbf{t}}$ satisfying

$$f_{\mathbf{x}}(x_i) = (-1)^{n-i}, \quad i=1,\ldots,n.$$

Our search for argmin $\|A_{\mathbf{x}}^{-1}\|_\infty$ therefore is the search for the $\mathbf{x}$ that minimizes $\|a^{\mathbf{x}}\|_\infty$. For this, we use, in effect, Remes' (second) algorithm for the construction of a ba from Π_k.

First we note that

$$(-1)^{n-i} a_i^{\mathbf{x}} > 0, \text{all } i,$$

since $a_i^{\mathbf{x}} = \sum_j A_{\mathbf{x}}^{-1}(i,j)(-1)^{n-j} = (-1)^{n-i}\sum_j |A_{\mathbf{x}}^{-1}(i,j)|$ with the sum necessarily positive. Further, since $f_{\mathbf{x}}$ strictly alternates in sign at the n points $x_1,\ldots,x_n$ and vanishes outside $(t_1 \mathinner{\ldotp\ldotp} t_{n+k})$, it follows that $f_{\mathbf{x}}$ has n distinct local extrema $y_1 < \cdots < y_n$ with

$$(-1)^{n-i} f_{\mathbf{x}}(y_i) \ge 1, \quad \forall i.$$

This implies that $t_i < y_i < t_{i+k}$. all i. Indeed, if, e.g., $y_i \le t_i$ for some i, then $B_\mu(y_\nu) = 0$ for all $\nu \le i \le \mu$, showing that, on $y_1, \ldots, y_i$, $f_{\mathbf{x}}$ agrees with $\sum_{j<i} a_j^{\mathbf{x}} B_j$. In particular, $i-2 \ge S^-(a_1^{\mathbf{x}}, \ldots, a_{i-1}^{\mathbf{x}}) \ge S^-(\sum_{j<i} a_j^{\mathbf{x}} B_j) = i-1$, a contradiction.

Consequently, there is exactly one $f_{\mathbf{y}} =: \sum_j a_j^{\mathbf{y}} B_j$ with $f_{\mathbf{y}}(y_i) = (-1)^{n-i}$, all i. For any $\gamma < 1$, the difference, $f_{\mathbf{x}} - \gamma f_{\mathbf{y}}$, strictly alternates in sign on $y_1 < \cdots < y_n$, hence we must have $S^-(a^{\mathbf{x}} - \gamma a^{\mathbf{y}}) = n-1$, and therefore $(-1)^{n-i} a_i^{\mathbf{x}} \ge (-1)^{n-i} a_i^{\mathbf{y}} \ge 0$, all i.

In this way, we obtain a sequence $f_m := \sum_j a_j^m B_j$, $m = 1, 2, \ldots$, whose coefficients converge monotonely to some sequence a^*. Its corresponding sequences $y_1^m < \cdots < y_n^m$ of extrema of f_m therefore also converge, necessarily to a strictly increasing sequence $x_1^* < \cdots < x_n^*$ since f_m strictly alternates in sign on $y_1^m < \cdots < y_n^m$. Let $f_* := \sum_j a_j^* B_j$. Then,

(89) $$1 = (-1)^{n-i} f_*(x_i^*) = \|f_*\|_\infty, \quad i = 1, \ldots, n,$$

since $1 = (-1)^{n-i} f_m(y_i^{m-1})$, while $\|f_m\| = \max_i |f_m(y_i^m)|$ and f_* is the uniform limit of $(f_m : m)$.

(90) Lemma. *Let $f_* = \sum_j a_j^* B_{j,k,\mathbf{t}}$ satisfy (89) for some (strictly) increasing $\mathbf{x}^*$. Then*

$$(-1)^{n-i} a_i^* = 1/\operatorname{dist}(B_i, \operatorname{span}(B_j : j \ne i)), \quad i = 1, \ldots, n.$$

In particular, f^* is independent of the choice of the initial $\mathbf{x}$.

Proof: Let $A_* := (B_j(x_i^*))$ and set

$$\lambda_i : S_{k,\mathbf{t}} \to \mathrm{IR} : f \to \sum_j A_*^{-1}(i,j) f(x_j^*).$$

Then $\lambda_i B_j = \delta_{ij}$, hence

$$(-1)^{n-i} a^* i = (-1)^{n-i} \lambda_i f_* = \sum_j (-1)^{n-i} A_*^{-1}(i,j)(-1)^{n-j} = \sum_j |A_*^{-1}(i,j)| = \|\lambda_i\|,$$

while, for any linear functional λ on any nls X and for any $f \in X$,

$$|\lambda f| = \|\lambda\| \operatorname{dist}(f, \ker \lambda).$$

Hence, $1 = \operatorname{dist}(f_*, \ker \lambda_i)$, while $\ker \lambda_i = \operatorname{span}(B_j : j \ne i)$ and so $|a_i^*| \operatorname{dist}(B_i, \ker \lambda_i) = \operatorname{dist}(f_*, \ker \lambda_i) = 1$. □

The function f_* is, by definition, the **Chebyshov spline** for $S_{k,\mathbf{t}}$, i.e., the unique (up to scalar multiples) element that maximally equioscillates, i.e., satisfies (89). From Lemma (90), we can write it as

$$C_{k,\mathbf{t}} := \sum_j (-1)^{n-j} B_j / \operatorname{dist}(B_j, \operatorname{span}(B_i : i \neq j)).$$

To be sure, any $f \in S_{k,\mathbf{t}}$ can have at most $n-1$ sign changes, hence n is indeed the maximal number of equioscillations possible for $f \in S_{k,\mathbf{t}}$.

I claim that a^* is the sought-for minimizer of $\|a^{\mathbf{x}}\|_\infty$. Indeed, we just saw that, starting with any $\mathbf{x}$, we reach the Chebyshov spline in the limit in a process during which the B-spline coefficients decrease in absolute value. Hence $\|a^*\|_\infty \le \|a^{\mathbf{x}}\|_\infty$ for any $\mathbf{x}$.

The terminology *Chebyshov spline*, though apt in view of the fact that, for $t_1 = \cdots = t_k = -1, t_{k+1} = \cdots = t_{2k} = 1$ it is the Chebyshov polynomial of degree $k-1$, unfortunately clashes with the standard term 'Tschebyscheffian' spline, meaning a piecewise function whose pieces all come from the same Haar space. The notation $C_{k,\mathbf{t}}$ unfortunately clashes with the earlier notation concerning the control polygon of a spline. Although whole books have been written on the special case $k = n$, in which case $C_{k,\mathbf{t}}$ on $I_{k,\mathbf{t}}$ agrees with the (suitably scaled and translated) Chebyshov polynomial of degree $k-1$, the Chebyshov spline is largely unexplored territory (except for K. Mørken's Ph.D. Thesis).

(91) Corollary. *For* $\mathbf{x} = (x_1 < \cdots < x_n)$ *with* $t_i < x_i < t_{i+k}$*, all* i*, let* $f_{\mathbf{x}} = \sum_j a_j^{\mathbf{x}} B_j$ *be the unique element in* $S_{k,\mathbf{t}}$ *satisfying* $f(x_i) = (-1)^{n-i}$*, all* i*. Then* $\operatorname{argmin}_{\mathbf{x}} \|a^{\mathbf{x}}\|_\infty$ *equals the extreme-point sequence of* $C_{k,\mathbf{t}}$.

Offhand, some of the extreme points of $C_{k,\mathbf{t}}$ may lie outside the basic interval $I_{k,\mathbf{t}} = [t_k \,.\,.\, t_{n+1}]$. However, if we restrict attention to this interval, then we would choose $t_1 = t_k$ and $t_{n+1} = t_{n+k}$. This violates the assumption that $\#t_i < k$, all i. However, assume first that, e.g., $t_1 < t_2 = t_k$. Then f is strictly monotone on $[t_1 \,.\,.\, t_2]$, hence necessarily $t_2 \le y_1$. Thus we may choose $\alpha = t_2$ in this case. For the same reason, we may choose $\beta = t_{n+k-1}$ in case $t_{n+k-1} < t_{n+k}$. That choice made, the location of t_1 and t_{n+k} becomes irrelevant to $S_{k,\mathbf{t}}$ as restricted to $I_{k,\mathbf{t}}$ (though the B-spline coefficients will change as we vary t_1 and t_{n+k}). In particular, we may choose $t_1 = t_k$ and $t_{n+k} = t_{n+1}$.

Since, for all i, $\lambda_{i,k} = \sum_j A_*^{-1}(i,j)[x_j^*]$ on $S_{k,\mathbf{t}}$, and the latter linear functional takes on its norm on $C_{k,\mathbf{t}}$, it follows that *the map*

$$S_{k,\mathbf{t}} \to \mathrm{I\!R} : \sum_j a_j B_j \mapsto \|a\|_\infty / \|\sum_j a_j B_j\|$$

takes on its maximum at $C_{k,\mathbf{t}}$, and that maximum is $\|a^*\|_\infty$. This maximum determines the condition of the B-spline basis, to be discussed next.

The condition of the B-spline basis

The **condition**

$$\kappa(V) := \|V\|\|V^{-1}\|$$

of a basis V (i.e., an invertible linear map from some $\mathbb{F}^n$ to the normed linear space $\operatorname{ran} V$) measures the extent to which the relative changes in the coordinates a of an element Va may be close to the resulting relative change in Va itself. The closer $\kappa(V)$ to 1, the more closely do these two relative sizes correspond.

For the B-spline basis, using the max-norm both in $\mathbb{R}^n$ and in $S_{k,\mathbf{t}}$, we have

$$\kappa_\infty([B_j : j]) = \sup_a \frac{\|a\|_\infty}{\|\sum_j a_j B_j\|_\infty}$$

since the B_j form a partition of unity, hence trivially $\sup_a \frac{\|\sum_j a_j \varphi_j\|_\infty}{\|a\|_\infty} = 1$.

Further,

$$\kappa_\infty([B_j : j]) = \max_i \|\lambda_{i,k}\|.$$

Hence the following lemma is of interest.

(92) Lemma. *The number*

$$d_k := \sup_{\mathbf{t}} \sup_i \sup_{f \in S_{k,\mathbf{t}}} |\lambda_{i,k,\mathbf{t}} f| / \|f\|_\infty$$

is finite.

Proof: Let $f \in S_{k,\mathbf{t}}$, hence

$$\lambda_{i,k} f = \sum_{\nu=1}^{k} (-D)^{\nu-1} \psi_i(\tau_i)/(k-1)!\, D^{k-\nu} f(\tau_i),$$

and let $I := [t_\ell \,..\, t_{\ell+1}]$ be any knot interval in $[t_i \,..\, t_{i+k}]$, and choose, as we may, $\tau_i \in I$. Then, as both ψ_i and f are polynomials of degree $< k$ on I, Markov's inequality (23) implies that

$$|(-D)^{\nu-1} \psi_i(\tau_i)||D^{k-\nu} f(\tau_i)| \le \text{const}_{k,\nu} \|\psi\|_\infty(I)/|I|^{\nu-1}\ \|f\|_\infty(I)/|I|^{k-\nu}.$$

On the other hand, by choosing, as we may, I to be a largest such knot interval, we can ensure that

$$|t_j - \tau_i| \le k|I|, \quad j = i, \ldots, i+k,$$

therefore $\|\psi_i\|_\infty(I) \le \text{const}_k |I|^{k-1}$. Therefore, altogether, $|\lambda_{i,k} f| \le \text{const}'_k \|f\|_\infty(I)$, which is even stronger than the claim to be proved. □

A more careful quantitative analysis shows that $d_k = O(9^k)$. Better results can be obtained with the aid of the following

(93) Claim. $d_k = \sup_{\mathbf{s}} \|\lambda_{k-1,k}|_{S_{k,\mathbf{s}}}\|([-1\,..\,1])$, *with* $\|\lambda\|(I) := \sup_{f\in \mathrm{dom}\,\lambda} |\lambda f|/\|f\|_\infty(I)$, *and* $\mathbf{s}$ *any knot sequence of the type*

$$s_1 = \cdots = s_k = -1 \le s_{k+1} \le \cdots \le s_{2k-3} \le 1 = s_{2k-2} = \cdots = s_{3k-3}.$$

Proof: Consider any particular $\lambda_i := \lambda_{i,k}|_{S_{k,\mathbf{s}}}$. If $t_{i+1} = t_{i+k-1}$, then (assuming WLOG that $t_i < t_{i+1}$), $\lambda_i = [t_{i+1}-]$, therefore $\|\lambda_i\| = 1 \le d_k$. In the contrary case, we may assume, after a suitable linear change of the independent variable, that $-1 = t_{i+1}$, $t_{i+k-1} = 1$. Let $\hat{\mathbf{t}}$ be the knot sequence obtained from $\mathbf{t}$ by inserting both -1 and 1 enough times to increase their multiplicity to $k-1$, and let $\hat{\imath}$ be such that $\hat{t}_{\hat{\imath}+j} = t_{i+j}$ for $j = 1,\ldots,k-1$. The corresponding spline space $S_{k,\hat{\mathbf{t}}}$ may well be larger than the space $S_{k,\mathbf{t}}$ we started with, but $\lambda_i = \hat{\lambda}_{\hat{\imath}}$ since, by (52), λ_i only depends on the knots $t_{i+1},\ldots,t_{i+k-1}$. This shows that

$$\|\lambda_i\| := \sup_{f\in S_{k,\mathbf{t}}} |\lambda_i f|/\|f\|_\infty \le \sup_{f\in S_{k,\hat{\mathbf{t}}}} |\hat{\lambda}_{\hat{\imath}} f|/\|f\|_\infty = \|\hat{\lambda}_{\hat{\imath}}\| = \|\hat{\lambda}_{\hat{\imath}}\|([-1\,..\,1]),$$

the last equality since $\hat{\lambda}_{\hat{\imath}} f$ only depends on $f|_{[-1..1]}$.

Consequently, $d_k \le \sup_{\mathbf{s}} \|\lambda_{k-1,k}\|([-1\,..\,1])$. The opposite equality is trivial. □

It is believed that $d_k \sim 2^k$. However, earlier hopes that the extremal knot configuration in the Claim would have no interior knots (i.e., would have an $\mathbf{s}$ with all its entries from $\{-1,1\}$) were dashed by the following simple counter-example: The cubic Chebyshov polynomial can be shown to have maximum B-coefficient 5 when written as an element of $S_{k,\mathbf{t}}$ with $\mathbf{t} = (-1,-1,-1,-1,1,1,1,1)$; however, the cubic Chebyshov spline for the knot sequence $(-1,-1,-1,-1,0,1,1,1,1)$ has B-coefficients $(1,-7/2,11/2,-7/2,1)$. Nevertheless, by considering the related extremum problem

$$d_{k,1} := \sup_{\mathbf{t}} \sup_i \sup_{f\in S_{k,\mathbf{t}}} |\lambda_{i,k,\mathbf{t}} f|\,|t_i - t_{i+k}|/\|f\|_1([t_i\,..\,t_{i+k}]), \tag{94}$$

Karl Scherer and Aleksei Shadrin were recently able to show that $d_k = O(k2^k)$.

It follows that

$$\|a\|_\infty/d_k \le \|\sum_i B_{i,k} a_i\|_\infty \le \|a\|_\infty$$

for any knot sequence $\mathbf{t}$ and any coefficient sequence a.

For such an estimate in $\mathbf{L}_p$, observe that, for any p and with $1/p + 1/p^* = 1$ and by Hölder's Inequality,

$$|\sum_j B_j(x) a_j| \le (\sum_j B_j(x)|a_j|^p)^{1/p} (\sum_j B_j(x))^{1/p^*},$$

hence

$$\|\sum_j B_j a_j\|_p^p \le \sum_j |a_j|^p (t_{j+k} - t_j)/k$$

(using the fact that $\int B_j = (t_{j+k} - t_j)/k$). On the other hand, from (94) and Hahn-Banach, we deduce the following

(95) Proposition. *There exists h_j with $|h_j| \le d_{k,1}/(t_{j+k}-t_j)$ and support in $[t_j \,..\, t_{j+k}]$ for which*
$$\int_{\mathbb{R}} h_j f = \int_{t_j}^{t_{j+k}} h_j f = \lambda_{j,k} f, \qquad f \in S_{k,\mathbf{t}}.$$
In particular,
$$d_k \le d_{k,1}.$$

Consequently, if $f = \sum_j B_j a_j$ and with $\int_j := \int_{t_j}^{t_{j+k}}$, then
$$a_j = \int_j h_j f \le (\int_j |h_j|^{p^*})^{1/p^*} (\int_j |f|^p)^{1/p},$$
while
$$(\int_j |h_j|^{p^*})^{1/p^*} \le \frac{d_{k,1}}{(t_{j+k}-t_j)} (t_{j+k}-t_j)^{1/p^*} = d_{k,1}/(t_{j+k}-t_j)^{1/p}.$$
Therefore,
$$\sum_j |a_j|^p (t_{j+k}-t_j)/k \le (d_{k,1})^p \sum_j \int_j |f|^p/k \ \le (d_{k,1})^p \|f\|_p^p,$$
the last inequality since at most k of the intervals $[t_j \,..\, t_{j+k}]$ have some given knot interval $[t_\ell \,..\, t_{\ell+1}]$ in common. It follows that, for any $1 \le p \le \infty$,

$$\|c\|_p / d_{k,1} \le \|\sum_i B_{i,k,p} c_i\|_p \le \|c\|_p, \tag{96}$$

for any knot sequence $\mathbf{t}$ and any coefficient sequence c, with

$$B_{i,k,p} := (k/(t_{i+k}-t_i))^{1/p} B_{i,k}. \tag{97}$$

Degree of approximation by splines

quasiinterpolants

It follows from Theorem (66) that the linear map
$$P : g \mapsto \sum_j \lambda_{j,k} g \, B_{j,k}$$
is a linear projector, with range $S = S_{k,\mathbf{t}} = \Pi^\rho_{k,\mathbf{t}}$, provided (as we have already assumed) that, for each i, we choose τ in (52) equal to some $\tau_i \in [t_i \,..\, t_{i+k})$. It is also local, since $\lambda_{j,k} g$ depends only on the behavior of g near τ_j, hence, by our choice of τ_j, only on the behavior of g on $\operatorname{supp} B_{j,k}$.

However, P is defined only for sufficiently smooth functions. In order to get such a projector on all of $\mathbf{L}_1(I_{k,\mathbf{t}})$, we make use of (95) in order to obtain the linear functional

$$\mu_j g := \int h_j g$$

with $\operatorname{supp} h_j \subset [t_j \mathinner{..} t_{j+k}]$ and $\|h_j\|_\infty \le d_{k,1}/(t_{j+k} - t_j)$ which, on $S_{k,\mathbf{t}}$, agrees with $\lambda_{j,k,\mathbf{t}}$. This implies that, for any $g \in \mathbf{L}_1$ and for $1 \le p \le \infty$,

$$|\mu_j g| \le \frac{d_{k,1}}{(t_{j+k} - t_j)^{1/p}} \|g\|_p([t_j \mathinner{..} t_{j+k}]).$$

The corresponding linear map

$$Q : g \mapsto \sum_j \mu_j g \, B_{j,k}$$

is defined for any $g \in \mathbf{L}_1(I_{k,\mathbf{t}})$, hence for any $g \in \mathbf{L}_p(I_{k,\mathbf{t}})$. Further, it is the identity on its range, hence a linear projector, its norm is bounded by $d_{k,1}$, and it is **local**, in the sense that, for $x \in [t_l \mathinner{..} t_{l+1})$,

(98)

$$\begin{aligned} Qg(x) = \sum_{j=l+1-k}^{l} \mu_j g \; B_j(x) &\le \sum_{j=l+1-k}^{l} |\mu_j g| \; B_j(x)/(t_{j+k} - t_j)^{1/p} \\ &\le d_{k,1} \|g\|_p([t_{l+1-k} \mathinner{..} t_{l+k}]) \sum_{j=l+1-k}^{l} B_j(x)/(t_{j+k} - t_j)^{1/p} \\ &= d_{k,1} \|g\|_p([t_{l+1-k} \mathinner{..} t_{l+k}]) \max_{j \in \{l+1-k,\ldots,l\}} (t_{j+k} - t_j)^{-1/p}. \end{aligned}$$

Since also $g - Qg = (1 - Q)(g - q)$ for any $q \in \Pi_{<k}$, we conclude that, for any g,

$$\|g - Qg\|_p([t_l \mathinner{..} t_{l+1}]) \le (1 + d_{k,1}) \operatorname{dist}(g, \Pi_{<k})_p([t_{l+1-k} \mathinner{..} t_{l+k}]) \tag{99}$$

(since $(t_{l+1} - t_l)/(t_{j+k} - t_j) \le 1$ for $j = l+1-k, \ldots, l$). In fact, this conclusion can already be reached if we only take care that μ_j agree with λ_j on $\Pi_{<k}$, all j.

The resulting map Q is called a **good quasiinterpolant of order** k, with 'good' referring to the uniform localness expressed by (98), and 'order k' referring to the fact that Q reproduces $\Pi_{<k}$. The term 'quasiinterpolant' was chosen by finite-element people once they realized that approximation order could be ascertained with the aid of maps Q that did not actually interpolate at the 'nodal points' of their elements, but merely matched enough information to give reproduction of certain polynomial spaces.

The error bound (99) is local; it is in terms of how well g can be approximated locally from polynomials of order k. That local distance is best estimated with the aid of

(100) Whitney's Theorem. *For any finite interval I,*

$$\operatorname{dist}(g, \Pi_{<k})_p(I) \sim \omega_k(g, |I|)_p.$$

Proof: For all $f \in \Pi_{<k}$, $\omega_k(g,|I|)_p = \omega_k(g-f,|I|)_p \le 2^k \|g-f\|_p(I)$, hence $\omega_k(g,|I|)_p \le \mathrm{const}_k \, \mathrm{dist}(g,\Pi_{<k})_p(I)$.

For the converse inequality, let $I =: [a \,..\, b]$ and start with an arbitrary $f \in W_p^{(k)}(I)$. With

$$T_k f := \sum_{j<k} D^j f(a)(\cdot - a)^j / j!$$

its truncated Taylor series, we find

$$\begin{aligned}
|f(x) - T_k f(x)| &= |\int_I (x-\cdot)_+^{k-1}/(k-1)! \, D^k f| \\
&\le \|D^k f|_I\|_p/(k-1)! \, \big(\int_I |(x-\cdot)_+^{k-1}|^{p^*}\big)^{1/p^*} \\
&\le \|D^k f|_I\|_p/(k-1)! \, |I|^{k-1/p},
\end{aligned}$$

hence

$$\|f - T_k f\|_q(I) \le \|D^k f|_I\|_p/(k-1)! \, |I|^{k-1/p+1/q}.$$

Consequently,

$$\mathrm{dist}(g,\Pi_{<k})_p(I) \le \|g - T_k f\|_p(I) \le \|g-f\|_p + |I|^k \|D^k f|_I\|_p/(k-1)!,$$

and, as $f \in W_p^{(k)}(I)$ is arbitrary here, we get

$$\mathrm{dist}(g,\Pi_{<k})_p(I) \le \mathrm{const}_k K(g,|I|^k; \mathbf{L}_p, W_p^{(k)}(I)),$$

while, from Theorem 26, we know that $\omega_k(g,t)_p \sim K(f,t^k;\mathbf{L}_p,W_p^{(k)})$. □

We conclude from Whitney's theorem and from (99) that

$$\mathrm{dist}(g,S_{k,\mathbf{t}})_p \le \mathrm{const}_k \, \omega_k(g,|\mathbf{t}|)_p, \quad 1 \le p \le \infty. \tag{101}$$

More than that, since

$$\omega_k(g,h)_p(I) = O(h^k) \quad \Longleftrightarrow \quad g \in W_p^{(k)}(I),$$

with

$$\sup_h \frac{\omega_k(g,h)_p}{h^k} = \begin{cases} \|D^k g\|_p(I), & p>1; \\ \mathrm{Var}(D^{k-1}g), & p=1, \end{cases}$$

we have the local bound

$$\|g - Qg\|_p([t_l \,..\, t_{l+1}]) \le \mathrm{const}_k |t_{l+k} - t_{l+1-k}|^k \|D^k g\|_p([t_{l+1-k} \,..\, t_{l+k}]).$$

For $p=\infty$, this suggests that, in approximating some g that is smooth except for some isolated singularities, the knot sequence $\mathbf{t}$ be chosen so as to make

$$l \mapsto |t_l - t_{l+1}|^k \|D^k g\|_\infty([t_l \,..\, t_{l+1}])$$

approximately constant. This is equivalent to making the map

$$l \mapsto |t_l - t_{l+1}|\bigl(\|D^k g\|_\infty([t_l \,..\, t_{l+1}])\bigr)^{1/k}$$

approximately constant, or, at least for a large knot sequence, making

$$l \mapsto |t_l - t_{l+1}||D^k g([t_l \,..\, t_{l+1}])|^{1/k}$$

approximately constant. If there are to be n knot intervals, then we can achieve this (approximately), by choosing $t_0 = 0 < t_1 < \cdots < t_n = 1$ so that

$$l \mapsto \int_{t_l}^{t_{l+1}} |D^k g|^{1/k}$$

is constant. With that choice,

$$\frac{1}{n^k}\|D^k g\|_{1/k} = \Bigl(\int_{t_l}^{t_{l+1}} |D^k g|^{1/k}\Bigr)^k \sim |t_{l+1} - t_l|^k \|D^k g\|_\infty([t_l \,..\, t_{l+1}]),$$

hence

$$\|g - Qg\|_\infty \le \mathrm{const}_k n^{-k} \|D^k g\|_{1/k}^k.$$

While this argument lacks some details, it makes the following essential point: *In order to achieve approximation order n^{-k} from the set of splines of order k with n interior knots, it is sufficient to have*

$$\|D^k g\|_{1/k}$$

finite. For example, this norm is finite for functions such as $\sqrt{\cdot}$ on $[0\,..\,1]$, ensuring therefore approximations by splines of order k whose error behaves like $O(n^{-k})$, with n the degrees of freedom used. In contrast, the error in best polynomial approximation from Π_n to $\sqrt{\cdot}$ on $[0\,..\,1]$ cannot be better than $O(n^{-1/2})$, by Bernshtein's Inverse Theorem, hence approximation by splines with n equally spaced knots cannot be better, either, as we show in the next section.

Jackson and Bernshtein for splines with uniform knot sequence

We know from (101) that, for every $g \in W_p^{(k)}$,

$$\mathrm{dist}(g, S_{k,\mathbf{t}})_p \le \mathrm{const}_k |\mathbf{t}|^k \|D^k g\|_p. \tag{102}$$

However, having $\mathrm{dist}(g, S_{k,\mathbf{t}})_p = O(|\mathbf{t}|^k)$ is, in general, no guarantee that $g \in W_p^{(k)}$ unless the knot sequences $\mathbf{t}$ involved are sufficiently 'generic'. Indeed, if every knot sequence considered contains the point $1/2$, then the function $g := (\cdot - 1/2)_+^{k-1}$ can be approximated without error from $S_{k,\mathbf{t}}$ even though g fails to have a kth derivative. But it is true that (102) cannot hold for every knot sequence $\mathbf{t}$ unless $g \in W_p^{(k)}$. In fact, this conclusion can already be reached if (102) holds for every *uniform* knot sequence. The essential point is that, for every point in the interval of approximation, there must be many knot sequences among those considered for which that point falls somewhere in the middle of a knot interval. Here are the details.

(103) Theorem (DeVore, Richards). *For $m \in \mathbb{N}$, let*

$$\mathbf{t}^{(m)} := (\ldots, 0, 0, 1/m, 2/m, \ldots, 1, 1, \ldots),$$

and set

$$E_m(g)_p := \operatorname{dist}(g, S_{k,\mathbf{t}^{(m)}})_p(I)$$

with $I := [0\,..\,1]$. Then,

$$\omega_k(g,\delta)_p \;\le\; \mathrm{const}_k \begin{cases} \left(\frac{1}{n}\sum_{m=n}^{2n} E_m(g)_p^p\right)^{1/p}, & p<\infty; \\ \max_{n\le m\le 2n} E_m(g)_\infty, & p=\infty, \end{cases}$$

with

$$n := \lfloor 1/\delta \rfloor.$$

Proof: Since

$$|\Delta_h^k g(x)| = |\Delta_h^k (g-f)(x)| \le 2^k \|g-f\|_\infty(x\,..\,x+kh)$$

for every $f \in \Pi_{<k}$, the heart of the proof is in the (nontrivial) observation (known as a **mixing lemma**) that, for all n, there exists $m \in \{n,\ldots,2n\}$ so that $\operatorname{dist}(x,\mathbf{t}^{(m)}) \ge 1/(16n)$, hence, for our n and for all $0 < h \le \delta_1 := 1/(16kn)$, since $kh \le k\frac{1}{16kn} = 1/(16n)$, there is $m \in \{n,\ldots,2n\}$ so that $[x\,..\,x+kh] \cap \mathbf{t}^{(m)} = \emptyset$, implying that any $f \in S_{k,\mathbf{t}^{(m)}}$ is a polynomial on $[x\,..\,x+kh]$, hence, with f a ba to g from $S_{k,\mathbf{t}^{(m)}}$, we get

$$|\Delta_h^k g(x)| \le 2^k E_m(g)_\infty.$$

It follows that

$$\omega_k(g,\delta_1) \le 2^k \max_{n\le m\le 2n} E_m(g)_\infty.$$

Since $\delta \le 1/n = 16k\delta_1$, this implies the claimed result for $p=\infty$.

As to $p<\infty$, the above-mentioned mixing lemma implies that

$$\frac{1}{n}\sum_{m=n}^{2n} \chi_{I_m}(x) \ge \frac{1}{64},$$

with

$$I_m := \{x \in [0\,..\,1-kh] : \operatorname{dist}(x,\mathbf{t}^{(m)}) \ge \frac{1}{16n}\}.$$

Therefore,

$$\frac{1}{64}\int_0^{1-kh} |\Delta_h^k g(x)|^p \,\mathrm{d}x \;\le\; \frac{1}{n}\sum_{m=n}^{2n}\int_{I_m} |\Delta_h^k g(x)|^p \,\mathrm{d}x \le \frac{2^{kp}}{n}\sum_{m=n}^{2n} E_m(g)_p^p.$$

This proves the result for $p<\infty$. □

We conclude that

$$E_n(g)_p \sim \omega_k(g, 1/n)_p.$$

In particular,

$$E_n(g)_p = O(n^{-k}) \iff g \in \begin{cases} W_p^{(k)} & 1 < p \le \infty; \\ W^{(k-1)}(BV) & p = 1. \end{cases}$$

Further, we get the *saturation result*:

$$E_n(g)_p = o(n^{-k}) \iff g \in \Pi_{<k}.$$

For the characterization of other rates of convergence, we make use of the following standard way to measure convergence behavior.

Approximation Spaces

It has become standard to measure the decay of $E_n(g) := \mathrm{dist}(g, M_n)$ of the distance of $g \in X$ from a given sequence $(M_n : n \in \mathbb{N})$ of subsets by comparing it to the sequence $(n^{-\alpha} : n \in \mathbb{N})$ for various values of α. Most simply, one might ask to have

$$E_n(g) = O(n^{-\alpha}),$$

with the supremum over all such α then being the **approximation order** for g provided by $(M_n : n \in \mathbb{N})$. However, in order to be able to *characterize* the class of functions g with a given approximation order, a somewhat more subtle way of measuring approximation order turns out to be often needed.

Define

$$\|a\|_q^{(\alpha)} := \begin{cases} \left(\sum_n (n^\alpha |a(n)|)^q/n\right)^{1/q} & 0 < q < \infty; \\ \sup_n n^\alpha |a(n)| & q = \infty. \end{cases}$$

With this, we define

$$A_q^{(\alpha)} := \{a \in \mathbb{R}^{\mathbb{N}} : \Delta|a| \le 0, \|a\|_q^{(\alpha)} < \infty\},$$

and make the following observations. First,

$$\|a\|_\infty^{(\alpha)} < \infty \iff a(n) = O(n^{-\alpha}),$$

i.e., $A_\infty^{(\alpha)}$ consists exactly of all the antitone sequences a that go to zero (at least) to order $-\alpha$. Further, $\|a\|_q^{(\alpha)} < \infty$ for some finite q implies that $a(n) = o(n^{-\alpha})$. (Indeed, if $a(n) \ne o(n^{ga})$, then $m(n)^\alpha |a(m(n))| \ge M > 0$ for some strictly increasing $m : \mathbb{N} \to \mathbb{N}$, and AGTASMAT $m(n-1) < m(n)/2$; thus, $(\|a\|_q^{(\alpha)})^q \ge \sum_n \sum_{m(n)/2 \le j \le m(n)} (j^\alpha |a(j)|)^q/n \ge \sum_n (2^{-\alpha} M)^q \sum_{m(n)/2 \le j \le m(n)} 1/n \to \infty$ (since $j^\alpha |a(j)| \ge (j/m(n))^\alpha m(n)^\alpha |a(m(n))|$ for $j < m(n)$, and $\sum_{m(n)/2 \le j \le m(n)} \sim \ln 2$), a contradiction.) In particular

$$N := \|()^\alpha a\|_\infty < \infty,$$

hence

$$\infty > (\|a\|_q^{(\alpha)})^q =: N^q \sum_n |b(n)|^q/n \ge N^q \sum_n |b(n)|^{q+r}/n$$

for any $r > 0$ since $|b(n)| \le 1$, all n. In particular,

$$\|a\|_{q+r}^{(\alpha)} = N\big(\sum_n |b(n)|^{q+r}/n\big)^{1/(q+r)}$$

is finite for any $r > 0$. In other words,

$$q < r \quad \Longrightarrow \quad A_q^{(\alpha)} \subset A_r^{(\alpha)}.$$

Finally, under the same assumption, and for any positive ε and any q,

$$\sum_n |n^{\alpha-\varepsilon} a(n)|^r/n = \sum_n |n^\alpha a(n)|^r/n^{1+r\varepsilon} \le N^r \sum_n n^{-1-r\varepsilon} < \infty.$$

Hence,

$$\beta < \alpha \quad \Longrightarrow \quad A_q^{(\alpha)} \subset A_\infty^{(\alpha)} \subset A_r^{(\beta)}, \quad \text{all } q, r.$$

For a given sequence $(M_n : n \in \mathbb{N})$ of subsets of the nls X, one defines, correspondingly, the **approximation classes**

$$A_q^{(\alpha)}(X, (M_n)) := \{g \in X : \|(E_n(g) : n \in \mathbb{N})\|_q^{(\alpha)} < \infty\}$$

that single out all the elements g of X for which $E_n(g)$ goes to zero in a certain way. By the earlier discussion of $(\alpha, q) \mapsto \|a\|_q^{(\alpha)}$, we conclude that $(\alpha, q) \mapsto A_q^{(\alpha)}$ is antitone in α for arbitrary q and isotone in q for fixed α, i.e.,

$$\alpha < \beta \Longrightarrow A_q^{(\alpha)} \supset A_r^{(\beta)},$$

$$q < r \Longrightarrow A_q^{(\alpha)} \subset A_r^{(\alpha)}.$$

There is a corresponding way to quantify and compare the speed at which, for a function f, $f(t)$ approaches 0 as $t \to 0$. Precisely, for such a function, one defines

$$\|f\|_q^{(\alpha)} := \begin{cases} \big(\int_I |t^{-\alpha} f(t)|^q \, \mathrm{d}t/t\big)^{1/q} & q < \infty; \\ \sup_I |t^{-\alpha} f(t)| & q = \infty, \end{cases}$$

with $I = [0 \,..\, a]$ some suitable interval, e.g., $I = [0 \,..\, 1]$. Usually, f is isotone. In that case, $\|f\|_q^{(\alpha)}$ is equivalent to (i.e., bounded above and below by certain f-independent positive multiples of) the following discrete versions, in which I is replaced by the sequence (a) $I = (1/n : n \in \mathbb{N})$ or (b) $I = (2^{-n} : n = 0, 1, \ldots)$, and, correspondingly, on the interval $[t_{i+1} \,..\, t_i]$, $\mathrm{d}t/t$ is replaced by $\sim (t_i - t_{i+1})/t_i$, which, for (a), is $1/(i+1)$, and, for (b), is a constant. For $q < \infty$, this gives the equivalent discrete versions

$$\|f\|_q^{(\alpha)}(1/\mathbb{N}) := \|(f(1/n) : n \in \mathbb{N})\|_q^{(\alpha)}$$

and

$$\|f\|_q^{(\alpha)}(2^{-\mathbb{N}}) := \big(\sum_n (2^{n\alpha}f(2^{-n}))^q\big)^{1/q},$$

respectively.

With these definitions in place, consider again the earlier result that

$$E_n(g)_p := \operatorname{dist}(g, S_{k,\mathbf{t}^{(n)}})_p \sim \omega_k(g, 1/n)_p,$$

hence

$$g \in A_q^{(\alpha)}(\mathbf{L}_p, (S_{k,\mathbf{t}^{(n)}}, n \in \mathbb{N})) \quad\iff\quad \|\omega_k(g,\cdot)_p\|_q^{(\alpha)} < \infty.$$

The latter condition appeared some time ago in the study of approximation order from trigonometric polynomials, leading to what we now call **Besov spaces**. Precisely,

$$B_q^{(\alpha)}(\mathbf{L}_p) := \{g \in \mathbf{L}_p : |g|_{B_q^{(\alpha)}(\mathbf{L}_p)} := \|\omega_{\lfloor\alpha\rfloor+1}(g,\cdot)_p\|_q^{(\alpha)} < \infty\}, \quad 0 < \alpha, 0 < q \le \infty.$$

These are complete metric spaces with the metric given by the (quasi-)norm

$$\|g\|_{B_q^{(\alpha)}(\mathbf{L}_p)} := \|g\|_p + |g|_{B_q^{(\alpha)}(\mathbf{L}_p)}.$$

(This is only a **quasi-**norm when $p < 1$ since then one only has $\|a+b\| \le \operatorname{const}(\|a\| + \|b\|)$ for some const instead of the triangle inequality.)

It is convenient to know that, for any $r > \alpha$,

$$g \mapsto \|\omega_r(g,\cdot)_p\|_q^{(\alpha)}$$

provides a (quasi-)seminorm on $B_q^{(\alpha)}(\mathbf{L}_p)$ equivalent to $|\cdot|_{B_q^{(\alpha)}(\mathbf{L}_p)}$. One direction of this claim is obvious since

$$\omega_{r+m}(g,t)_p \le 2^m \omega_r(g,t)_p.$$

For the other direction, one needs a result like

(104) Marchaud's Theorem. *For $g \in \mathbf{L}_p(I)$,*

$$\omega_r(g,t)_p \le \operatorname{const}_r t^r \big(\int_t^{|I|} \frac{\omega_{r+m}(g,s)_p}{s^{r+1}}\,\mathrm{d}s + \frac{\|g\|_p}{|I|^r}\big).$$

along with

(105) Hardy's Inequality. *For $\alpha > 0$, and $1 \le q \le \infty$, and f any positive measurable function,*

$$\int_0^\infty \big[t^\alpha \int_t^\infty f(s)\frac{\mathrm{d}s}{s}\big]^q \frac{\mathrm{d}t}{t} \le \frac{1}{\alpha^q}\int_0^\infty \big[t^\alpha f(t)\big]^q \frac{\mathrm{d}t}{t}.$$

With this, we reach the conclusion that

$$\text{(106)}\quad A_q^{(\alpha)}(\mathbf{L}_p(0\,.\,.\,1), (S_{k,\mathbf{t}^{(n)}} : n \in \mathbb{N})) \;=\; B_q^{(\alpha)}(\mathbf{L}_p(0\,.\,.\,1)), \qquad 0 < q \le \infty,\ 0 < \alpha < k.$$

But we have to question just what we have gained by this somewhat formal exercise. The gain is substantial to the extent that we understand

Besov spaces

Such understanding comes from knowing more about 'typical' elements and/or from knowing alternative characterizations of such spaces. Among these will be those derived from the fact that Besov spaces turn out to be approximation spaces for various other sequences of approximating sets, among these $S_{k,n}$:= splines with n free knots and Π_n/Π_n := rational functions with numerator and denominator degree $\le n$, both classical examples of *nonlinear* approximating sets, to be discussed next.

As for specific examples, consider the Heaviside function

$$g := ()_+^0,$$

as an element of $\mathbf{L}_p[-1\,..\,1]$, say. For $0 < h < 1$, $\|\Delta_h g\|_p^p = h$, hence

$$\omega_1(g,h)_p = h^{1/p}.$$

Therefore, for $0 < \alpha < 1$ and $0 < q < \infty$,

$$(\|\omega(g,\cdot)_p\|_q^{(\alpha)})^q \sim \int_0^1 (t^{-\alpha} t^{1/p})^q \,\mathrm{d}t/t = \int_0^1 ()^{(-\alpha+1/p)q-1},$$

and this is finite iff $\alpha < 1/p$. For $q = \infty$, we look instead at

$$\sup_t t^{-\alpha} t^{1/p},$$

and this is finite iff $\alpha \le 1/p$. Consequently,

$$()_+^0 \in B_q^{(\alpha)}(\mathbf{L}_p) \quad \Longleftrightarrow \quad \begin{cases} \alpha < 1/p, & \text{if } 0 < q < \infty; \\ \alpha \le 1/p, & \text{if } q = \infty. \end{cases}$$

Note the minor, yet decisive, role played by the parameter q here. Note also that α can be large provided we are willing to consider $p < 1$. Finally, note the implication that any function with *finitely many* jump discontinuities but which is otherwise smooth lies in $B_q^{(\alpha)}(\mathbf{L}_p)$ for exactly the same triples (α, q, p).

More generally,

$$(\Delta_h^k ()_+^{k-1})(x) = k! h^k [0, h, \ldots, kh](\cdot - x)_+^{k-1} = (k-1)! h^{k-1} B(x|0, h, \ldots, kh),$$

hence, for $0 < q < \infty$,

$$(\|\omega_k(()_+^{k-1}, \cdot)_p\|_q^{(\alpha)})^q \sim \int_0^1 (t^{-\alpha} t^{k-1+1/p})^q \,\mathrm{d}t/t = \int_0^1 ()^{(-\alpha+k-1+1/p)q-1}$$

while

$$\|\omega_k(()_+^{k-1}, \cdot)_\infty\|_\infty^{(\alpha)} \sim \sup_t t^{-\alpha+k-1+1/p}.$$

Therefore,

$$()_+^{k-1} \in B_q^{(\alpha)}(\mathbf{L}_p) \quad\iff\quad \begin{cases} \alpha < k-1+1/p, & \text{if } 0<q<\infty; \\ \alpha \le k-1+1/p, & \text{if } q=\infty. \end{cases}$$

Other examples include:

$$B_\infty^{(r)}(\mathbf{L}_p(I)) \supset \begin{cases} W_p^{(r)}(I), & \text{if } 1<p<\infty; \\ W^{(r-1)}BV(I), & \text{if } p=1, \end{cases}, \quad r \in \mathbb{N},$$

i.e., for $1 \le p < \infty$ and integer r, $B_\infty^{(r)}(\mathbf{L}_p(I))$ contains the **Sobolev space** of all functions on I with absolutely continuous $(r-1)$st derivative and (a) rth derivative in $\mathbf{L}_p$ if $p>1$; (b) $(r-1)$st derivative of bounded variation, if $p=1$. There is no equality here, the only related equality being

$$B_2^{(r)}(\mathbf{L}_2) = W_2^{(r)}.$$

For $p=\infty$, one has

$$B_\infty^{(r)}(C(I)) = \mathrm{Lip}(r+1, C(I)),$$

with the special case $B_\infty^{(1)}(C(I))$ equal the **Zygmund space**, i.e., slightly larger than $\mathrm{Lip}_1(I)$. Note the somewhat more subtle description in the extreme cases $p=1,\infty$.

More generally, for nonintegral α, and with $\alpha =: r-1+\beta$ for some $r \in \mathbb{N}$ and $\beta \in (0\,..\,1)$,

$$B_\infty^{(\alpha)}(\mathbf{L}_p(I)) = \mathrm{Lip}(\alpha, X_p(I)),$$

the space of all functions on I with absolutely continuous $(r-1)$st derivative and rth derivative in $\mathrm{Lip}_\beta(I)_p$. In view of the fact that

$$\mathrm{Lip}_\beta(I)_p := \{g \in \mathbf{L}_p(I) : \sup_h \|\Delta_h g\|_p / h^\beta < \infty\},$$

this is certainly just a tautology.

Besov spaces are helpful also because they appear as the 'right' spaces in interpolation between standard spaces and in the Sobolev embedding theorem and its generalization. For the discussion of these matters, it is very helpful to follow Ron DeVore's advice and view the whole situation by representing *all* the spaces $(B_q^{(\alpha)}(\mathbf{L}_p) : 0 < q \le \infty)$ by the point

$$(1/p, \alpha)$$

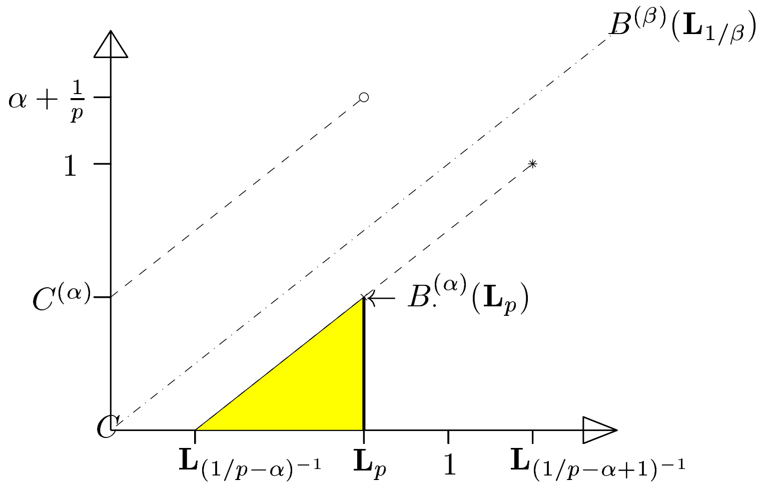
in $\mathbb{R}_+^2$, as in Figure (107).

Besov spaces arise naturally in interpolation between the spaces $\mathbf{L}_p$ and $W_p^{(r)}$. Briefly, for any pair (X_0, X_1) of (quasi-)normed spaces with

$$X_1 \hookrightarrow X_0,$$

i.e., X_1 continuously imbedded in X_0, one obtains a two-parameter continuum of spaces

$$X_1 \subset (X_0, X_1)_{\theta,q} \subset X_0$$



(107) Figure DeVore's diagram associates with the point $(\frac{1}{p}, \alpha)$ the whole family $B_{\cdot}^{(\alpha)}(\mathbf{L}_p) := (B_q^{(\alpha)}(\mathbf{L}_p[0\,..\,1]) : 0 < q \le \infty)$ of Besov spaces, to facilitate discussion (and retention) of basic facts about embeddings, interpolation in smoothness spaces, as well as the essential difference between linear *vs* nonlinear approximation.

The shaded triangle comprises all the Besov spaces into which the marked space $B_q^{(\alpha)}(\mathbf{L}_p)$ is continuously embedded, with the precise choice of the secondary parameter, q, of import only along the slanted edge.

as

$$X_{\theta,q} := (X_0, X_1)_{\theta,q} := \{g \in X_0 : |g|_{\theta,q} := \|K(g,\cdot)\|_q^{(\theta)} < \infty\},$$

with

$$K(g,t) := K(g,t; X_0, X_1) := \inf_{f \in X_1} (\|g - f\|_{X_0} + t\|f\|_{X_1})$$

the K-functional for the pair (X_0, X_1).

The main result concerning this interpolation of spaces is the following

(108) Theorem. *Let $X_1 \hookrightarrow X_0$ and $Y_1 \hookrightarrow Y_0$ be pairs of complete (quasi-)normed spaces and assume that the linear map U maps X_i boundedly into Y_i, $i = 0, 1$. Then, for $0 < q \le \infty$ and $0 < \theta < 1$, U also maps each $X_{\theta,q}$ boundedly into $Y_{\theta,q}$ and, with*

$$M_i := \|U : X_i \to Y_i\|, \quad i = 0, 1,$$

one has

$$\|U : X_{\theta,q} \to Y_{\theta,q}\| \le M_0^{1-\theta} M_1^{\theta}.$$

(See Theorem 7.1 in Chapter 6 of DeVore-Lorentz.)

As an application, take $X_0 = L_p(I)$ and $X_1 = W_p^{(k)}(I)$. Then, for any $0 < \alpha < k$,

$$(X_0, X_1)_{\alpha/k,q} = B_q^{(\alpha)}(\mathbf{L}_p),$$

(using Theorem (26)). Further, with $Y_0 = \mathbf{L}_p = Y_1$, we take for U the error $1 - Q$ in the quasi-interpolant Q introduced earlier. From (99) and Whitney's Theorem (100), we know that then

$$M_1 := \|(1 - Q) : X_1 \to X_0\| \le \mathrm{const}_k |\mathbf{t}|^k,$$

while

$$M_0 := \|(1 - Q) : X_0 \to X_0\| \le 1 + d_{k,1} < \infty.$$

Therefore, for any $0 < \alpha < k$,

$$\|(1 - Q) : B_q^{(\alpha)}(\mathbf{L}_p) \to \mathbf{L}_p\| \le \mathrm{const}_k \mathbf{t}^\alpha,$$

thus providing (a lower bound on) the approximation order from $S_{k,\mathbf{t}}$ to elements of $B_q^{(\alpha)}(\mathbf{L}_p)$.

Nonlinear approximation

We know from Theorem (103) that we cannot have $\mathrm{dist}(g, S_{k,\mathbf{t}})_p = O(\mathbf{t}^k)$ for all $\mathbf{t}$ unless $g \in W_p^{(k)}$. Yet, we already observed that some functions without a kth derivative, like $g := \surd$ on $I := [0\,.\,.\,1]$, can nevertheless be approximated to $O(n^{-k})$ by a spline of order k with n suitably chosen interior knots.

The basic results here are the following. Let

$$M_n := S_{k,n}$$

be the space of all splines of order k on I with $< n$ interior knots, hence with at most n polynomial pieces. Then, M_n is scale-invariant but fails to be closed under addition. However,

$$M_n + M_n \subset M_{2n},$$

hence Theorem (15) is applicable here provided we can produce compatible Jackson and Bernshtein inequalities. These were obtained not all that long ago by Petrushev.

(109) Theorem (Petrushev). *Let $0 < p < \infty$ and, correspondingly,*

$$B^{(\alpha)} := B_\gamma^{(\alpha)}(\mathbf{L}_\gamma(0\,.\,.\,1)), \qquad \gamma := 1/(\alpha + 1/p).$$

then, for $0 < \alpha < k$,

(i) $\forall g \in \mathbf{L}_p, \quad \mathrm{dist}(g, S_{k,n}) \le \mathrm{const}_k n^{-\alpha} |g|_{B^{(\alpha)}}$.

(ii) $\forall m \in S_{k,n}, \quad \|m\|_{B^{(\alpha)}} \le \mathrm{const}_k n^\alpha \|m\|_p$.

(See Theorem 8.2 in Chapter 12 of DeVore-Lorentz.)

As a quick example, consider the simplest possible case, that of approximation from $M_n := S_{1,n}$, the (nonlinear) space of all step functions on $[0\,.\,.\,1]$ with at most n different values. Already in 1961, Kahane proved the following neat result:

(110) Proposition (Kahane). *For* $g \in C([0\,..\,1])$, $\operatorname{dist}(g, S_{1,n}) \le M/(2n)$ *for all* $n \in \mathbb{N}$ *if and only if* $\operatorname{Var}(g) \le M$.

Proof: There is nothing to prove if M or $\operatorname{Var}(g)$ are infinite, hence assume that both are finite.

Choose $0 = t_0 < \cdots < t_n = 1$ so that $\operatorname{Var}(g)(t_i\,..\,t_{i+1}) \le \operatorname{Var}(g)/n$, all i. Let $f \in S_{1,n}$ be such that, on $(t_i\,..\,t_{i+1})$, f equals the midpoint of the interval $g([t_i\,..\,t_{i+1}])$, all i. Then, $\|g - f\|_\infty \le \frac{1}{2}\operatorname{Var}(g)/n$.

Conversely, for any $f \in S_{1,n}$, and any $0 = x_0 < \cdots < x_m = 1$,

$$|g(x_i) - g(x_{i-1})| \le 2\|g - f\|_\infty \# f([t_{i-1}\,..\,t_i]),$$

therefore

$$\sum_i |g(x_{i-1}) - g(x_i)| \le 2\|g - f\|_\infty \sum_{i=1}^m \# f([x_{i-1}\,..\,x_i]) \le 2(n+m)\|g - f\|_\infty.$$

Hence, if $\operatorname{dist}(g, S_{1,n}) \le M/(2n)$, then, for any $\varepsilon > 0$,

$$\sum_i |g(x_{i-1}) - g(x_i)| \le (M + \varepsilon)(1 + m/n),$$

therefore, by letting $n \to \infty$, $\operatorname{Var}(g) \le M + \varepsilon$. $\square$

Thus, while $\operatorname{dist}(g, S_{1,\mathbf{t}^{(n)}})_\infty = O(1/n)$ requires g to lie in $W_\infty^{(1)}$, getting $\operatorname{dist}(g, S_{1,n}) = O(1/n)$ only requires $g \in \mathrm{BV}$, i.e., in $B_\infty^{(1)}(\mathbf{L}_1)$. In the DeVore diagram (107), $W_\infty^{(1)}$ lies vertically above C, while the latter space lies on the 45-degree line emanating from C. Petrushev's result shows this to hold for splines of general order k. Namely, the approximation order from splines of order k cannot exceed k, but that order is achieved by nonlinear approximation (i.e., approximation from $(S_{k,n} : n \in \mathbb{N})$) to much rougher functions than is possible by linear approximation (i.e., approximation from $(S_{k,\mathbf{t}^{(n)}} : n \in \mathbb{N})$).

See DeVore's survey article on Nonlinear Approximation, in *Acta Numerica, 1998.*

