Concepts of Approximation Theory

July 21, 2007

almost Chebyshev: A subspace M of a normed linear space X is said to be almost Chebyshev if the set of points of X that do not possess a unique best approximant from M is of first category.

apolar polynomials: Two polynomials

$$P(x) = \sum_{k=0}^{n} a_k \binom{n}{k} x^k$$

and

$$Q(x) = \sum_{k=0}^{n} b_k \binom{n}{k} x^k,$$

both of exact degree n, are called apolar if

$$\sum_{k=0}^{n} (-1)^k a_k b_{n-k} \binom{n}{k} = 0.$$

absolutely monotone: A real-valued infinitely differentiable function f on $[a, b]$ is said to be absolutely monotone on $[a, b]$ if

$$f^{(k)}(x) \geq 0, \qquad x \in [a, b],\ k = 0, 1, 2, \ldots.$$

alternation points for a function f defined on an interval I is any sequence $x_1 < x_2 < \cdots < x_m$ of points in that interval such that $|f(x_i)| = \|f\|_\infty$ and $\operatorname{sign} f(x_{i+1}) = -\operatorname{sign} f(x_i)$.

approximand is the element to be approximated.

approximant is the element that is doing the approximating.

approximation spaces consist of functions with prescribed rate of approximation. E.g. if $E_n(f)$ is the error of best approximation of f by polynomials of degree at most n, $\alpha > 0$ and $0 < q \le \infty$, then the collection of functions f for which

$$\left(\sum_{k=1}^{\infty} [k^{\alpha} E_{k-1}(f)]^q \frac{1}{k} \right)^{1/q} < \infty$$

is an approximation space (called A_q^{α}).

approximation with constraint: An additional requirement (like monotonicity, convexity, interpolation) has to be satisfied by the approximation process.

backward difference: Let f be a function defined, say, on an interval (a, b), and let h be a real number. The backward differences $\nabla_h^r f$ of f with (positive) **step-size** h are defined recursively as $\nabla_h^0 f(x) := f(x)$,

$$\nabla_h^r f(x) := \nabla_h^{r-1} f(x) - \nabla_h^{r-1} f(x-h) = \sum_{k=0}^{r} (-1)^k \binom{r}{k} f(x - kh)$$

whenever this expression has meaning. Without the subscript h, the step-size $h = 1$ is implied.

basic polynomials of Hermite interpolation: These are defined analogously to the *basic polynomials of Lagrange interpolation* as the polynomials l of degree $m_0 + \cdots + m_k - 1$ that match the special Hermite interpolation conditions $l^{(r)}(x_i) = 0$ for all $r < m_i$, $0 \le i \le k$, except for one, and for it $l^{(r)}(x_i) = 1$. See *Hermite interpolation*.

basic polynomials of Lagrange interpolation: Let $x_0, x_1, \ldots, x_n$ be different points on the real line or the complex plane. The basic Lagrange interpolating polynomials with respect to these points (see *Lagrange interpolation*) are the polynomials l_j, $j = 0, 1, \ldots, n$, of degree n, for which $l_j(x_j) = 1$ and $l_j(x_i) = 0$ for all other i. They have the explicit form

$$l_j(x) := \prod_{i \ne j} \frac{x - x_i}{x_j - x_i}.$$

Baskakov operators associate with a function f defined on $[0, \infty)$ the functions

$$V_n f(x) := \sum_{k=0}^{\infty} f\left(\frac{k}{n}\right) \binom{n+k-1}{k} x^k (1+x)^{-n-k}$$

provided these exist.

Bernstein constant is the value of the limit

$$\lim_{n\to\infty} nE_n(|x|),$$

where $E_n(|x|)$ denotes the error of best approximation to $|x|$ on $[-1,1]$ by polynomials of degree at most n.

Bernstein-Bézier form of a univariate polynomial p:

$$p(x) =: \sum_{k=0}^{n} a_k \binom{n}{k} x^k (1-x)^{n-k}.$$

For a polynomial p of degree $\le n$ in $\mathbf{x} = (x_1, \ldots, x_d)$, the Bernstein-Bézier form with respect to the $d+1$-set V in general position is

$$p(\mathbf{x}) =: \sum_{|\alpha|=n} a_\alpha B_\alpha(\boldsymbol{\xi}(\mathbf{x})),$$

with

$$B_\alpha(\boldsymbol{\xi}) := \binom{|\alpha|}{\alpha} \boldsymbol{\xi}^\alpha$$

and

$$\mathbf{x} =: \sum_{\mathbf{v}\in V} \mathbf{v}\xi_{\mathbf{v}}(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d,$$

i.e., $\boldsymbol{\xi}(\mathbf{x}) := (\xi_{\mathbf{v}}(\mathbf{x}) : \mathbf{v} \in V)$ are the barycentric coordinates of $\mathbf{x}$ with respect to V.

Bernstein inequality: If t_n is a real *trigonometric polynomial* of *degree at most* n and $|t_n(x)| \le 1$ for all $x \in \mathbb{R}$, then

$$|t_n'(x)| \le n, \qquad x \in \mathbb{R}.$$

Bernstein polynomials: Let f be a function on $[0,1]$. Its Bernstein polynomial of degree n is defined as

$$B_n f(x) := \sum_{k=0}^{n} f\left(\frac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k}.$$

Bernoulli splines are defined recursively as

$$\mathcal{B}_1(x) := (\pi - x)/2, \quad \mathcal{B}_m(x) := \int_0^x \mathcal{B}_{m-1}(t)\,\mathrm{d}t - \int_0^{2\pi} \mathcal{B}_{m-1}(t)\,\mathrm{d}t, \quad 0 \le x < 2\pi,$$

and these are extended 2π-periodically.

Besov spaces: Let $\alpha > 0$, $r = \lfloor \alpha \rfloor + 1$, and $0 < p, q \le \infty$. The corresponding Besov space $B_q^\alpha(L_p)$ (say, on an interval) consists of all functions for which the Besov norm

$$\|f\|_{B_q^\alpha(L_p)} := \|f\|_p + \left(\int_0^\infty \left[t^{-\alpha} \omega_r(f,t)_p \right]^q \frac{\mathrm{d}t}{t} \right)^{1/q}$$

is finite. For $q = \infty$, the second term on the right is understood to be

$$\sup_{t>0} t^{-\alpha} \omega_r(f,t)_p.$$

The definition in higher dimension is similar, just use the appropriate moduli of smoothness.

best approximant to f from M in a metric space X (with metric ρ) containing M is any element $g \in M$ for which $\rho(f,g)$ is minimal, i.e., $\rho(f,g) = \inf_{h \in M} \rho(f,h) =: \operatorname{dist}(f,M)$. Also called **best approximation**.

Bleimann-Butzer-Hahn operators associate with a function f defined on $[0,\infty)$ the functions

$$L_n f(x) := \frac{1}{(1+x)^n} \sum_{k=0}^n f\left(\frac{k}{n-k+1} \right) \binom{n}{k} x^k.$$

Birkhoff interpolation: Similar to Hermite interpolation, but one wants to match not necessarily consecutive derivatives at the given points.

More precisely, let there be given an **interpolation matrix**, i.e., $E = (e_{ij} : i = 0, \ldots, k; j = 0, \ldots, \ell)$ with entries from $\{0,1\}$. If $x_0, \ldots, x_k$ are points and $y_{i,j}$ are numbers, then Birkhoff interpolation with these data consists of finding a polynomial P with $P^{(j)}(x_i) = y_{i,j}$ whenever $e_{ij} = 1$. The degree of the polynomial P should be less than the number of nonzero entries in E.

breakpoint or **break** of a spline or piecewise polynomial f is a place across which f or one of its derivatives has a jump. The **(order of) smoothness** of f at a breakpoint is the smallest m for which $D^m f$ has a jump across it.

B-spline associated with the $k+1$ points $t_0 \le t_1 \le \cdots \le t_k$ (its *knots*) is

$$N(x|t_0, \ldots, t_k) := (t_k - t_0)\,\boldsymbol{\Delta}(t_0, \ldots, t_k)(\cdot - x)_+^{k-1},$$

where $x_+^{k-1} := \max(0,x)^{k-1}$ is a *truncated power*, and $\boldsymbol{\Delta}(t_0, \ldots, t_k)g$ is the *divided difference* of g on $t_0, \ldots, t_k$.

$N(\cdot|t_0,\ldots,t_k)$ is zero outside the interval $[t_0,t_k]$ and is an element of any *Schoenberg space* $\mathcal{S}_k(b,m,\mathbb{R})$ whose *break* sequence b contains, for $0 \le i \le k$, some $b_j = t_i$ with the corresponding m_j at least as big as k minus the multiplicity with which b_j occurs in t.

For a general knot sequence t,

$$N_{ik} := N(\cdot|t_i,\ldots,t_{i+k}).$$

Sometimes, a different normalization is used. E.g.,

$$M(x|t_0,\ldots,t_k) := k\boldsymbol{\Delta}(t_0,\ldots,t_k)(\cdot - x)_+^{k-1}$$

are the B-splines normalized to have total integral 1 while $\sum_{i=r}^{s} N_{ik}(x) = 1$ on $[t_{r+k-1}, t_{s+1}]$.

capacity: Let K be a set in a metric space and, for $\varepsilon > 0$, let $M_\varepsilon(K)$ be the maximum number of points with mutual distances $> \varepsilon$. Then, $\log_2 M_\varepsilon(K)$ is called the **ε-capacity of** K.

cardinal spline interpolation is spline interpolation at all the integers, such that if the degree of the spline is odd then the knots of the spline are the integers, while if the degree is even, then the knots are the integers shifted by 1/2.

Cauchy index: Let r be a real rational function with a real pole at α. Its Cauchy index at α is 1 if $r(\alpha-) = -\infty$ and $r(\alpha+) = \infty$, it is -1 if $r(\alpha-) = \infty$ and $r(\alpha+) = -\infty$, and it is 0 otherwise.

The Cauchy index of r on an interval (a,b) is the sum of the Cauchy indices for the poles lying in (a,b).

Cauchy transform of a measure μ with (compact) support on the complex plane is the function

$$z \mapsto \int \frac{\mathrm{d}\mu(t)}{z-t}\ .$$

See also *Markov function*.

central difference: Same as *symmetric difference*.

Chebyshev alternation: See *alternation points*.

Chebyshev constant: See *Chebyshev polynomials on a compact set* $K \subset \mathbb{C}$.

Chebyshev numbers: See *Chebyshev polynomials on a compact set* $K \subset \mathbb{C}$.

Chebyshev polynomials: They are defined for $x \in [-1,1]$ as

$$T_n(x) := \cos(n \arccos x) = 2^{n-1}x^n + \cdots,$$

and for all complex z as

$$T_n(z) := \frac{1}{2}\left\{\left(z+\sqrt{z^2-1}\right)^n + \left(z-\sqrt{z^2-1}\right)^n\right\},$$

where the branch of the square root that is positive for positive z is selected. Sometimes, a different normalization is used (e.g., often $2^{-n+1}T_n$ are called Chebyshev polynomials).

Chebyshev polynomials of the second kind are defined as

$$U_n(x) := \frac{\sin((n+1)\arccos x)}{\sin(\arccos x)}, \qquad x \in [-1,1].$$

Chebyshev polynomials on a compact set $K \subset \mathbb{C}$ are polynomials $p_n(z) = z^n + \cdots$ with leading coefficient 1 that minimize the supremum norm on K. The nth root of the minimal norm is called the nth **Chebyshev number** of K, and the quantity

$$\lim_{n\to\infty} \|p_n\|^{1/n}$$

is called the **Chebyshev constant** of K.

Chebyshev space is the same as *unicity space*. Sometimes, the term "Chebyshev space" is used for the linear span of a *Chebyshev system*.

Chebyshev system is any sequence $(f_1, \ldots, f_n)$ of real-valued functions for which $\det(f_j(x_i) : i,j = 1,\ldots,n) > 0$ for all choices $x_1 < \cdots < x_n$.

Any Chebyshev system is a *Haar system*. For that reason, some people use the two terms interchangeably.

Christoffel functions of a measure μ on the complex plane are the functions

$$\lambda_n(z) := \inf_{p(z)=1} \int \|p\|^2 \, \mathrm{d}\mu,$$

where the infimum is taken over all polynomials p of degree at most n that take the value 1 at z.

Christoffel numbers are the same as *Cotes numbers*. See *quadrature formulæ*.

completely monotone: A real-valued infinitely differentiable function f on $[a,b]$ is said to be completely monotone on $[a,b]$ if

$$(-1)^k f^{(k)}(x) \geq 0, \qquad x \in [a,b],\ k = 0,1,2,\ldots.$$

collocation is another word for interpolation at given sites. Correspondingly, a **collocation matrix** is a matrix of the form $(f_j(\tau_i) : i, j = 1, \ldots, n)$, with $\tau_1, \ldots, \tau_n$ a sequence of points and $f_1, \ldots, f_n$ a sequence of functions.

complete Chebyshev system is any sequence $(f_1, \ldots, f_n)$ of functions for which $(f_1, \ldots, f_m)$ is a Chebyshev system for $m = 1, \ldots, n$. Sometimes called a **Markov system**.

complete orthogonal system: See *orthogonal system*.

condition number of an invertible square matrix A is $\|A\| \|A^{-1}\|$ (for some norm $\|\cdot\|$).

continuous selection is a continuous *metric selection*.

convex approximation: See *approximation with constraint*.

Cotes numbers: The coefficients in *quadrature formulæ*. Sometimes, these are called **Christoffel numbers**.

cubic splines are *splines* of *degree* 3, but it has become customary to refer to any spline of *order* 4 as a cubic spline.

de Boor-Fix functionals are the linear functionals

$$\lambda_{ik} : f \mapsto \sum_{\nu=1}^{k} (-D)^{k-\nu} \psi_{ik}(\tau_i) D^{\nu-1} f(\tau_i)/(k-1)!$$

with

$$\psi_{ik}(x) := (t_{i+1} - x) \cdots (t_{i+k-1} - x), \quad t_i+ \le \tau_i \le t_{i+k}-.$$

They are **dual** to the *B-spline*s in the sense that

$$\lambda_{ik} N_{jk} = \delta_{ij},$$

hence

$$Q : f \mapsto \sum_i (\lambda_{ik} f) N_{ik}$$

is a local linear *projector* onto the linear span of (N_{ik}); see *quasi-interpolant*.

decreasing rearrangement of a function f defined on a measurable set $A \subset \mathbb{R}$ is the function

$$f^*(t) := \inf\{y : \mu_f(y) \le t\},$$

where μ_f is the associated *distribution function*.

defect of a spline: See *splines*.

degree of a polynomial: See *polynomials*.

degree of a spline is the largest degree of any of its polynomial pieces.

density of a set X in a metric space Y means that every neighborhood of every point of Y contains an element of X.

Descartes system is any finite sequence of functions for which every subsequence is a *Chebyshev system*.

However, the literature ambiguous. For example, for some people, a Descartes system is defined as a finite sequence of functions for which every subsequence is a *Haar system*.

determinacy of the moment problem means that a measure μ is uniquely determined by its moments.

differences: See *forward*, *backward*, *symmetric*, *central*, or *divided difference*.

dilation-invariant space is a function space on $\mathbb{R}^d$ such that if f belongs to the space then so does the function $x \mapsto f(hx)$ for every positive h.

Dirac delta at x is the measure δ_x that puts mass 1 at x and 0 everywhere else.

Dirichlet kernel of degree n is defined as

$$D_n(x) := \frac{1}{2}\sum_{k=-n}^{n} \mathrm{e}^{\mathrm{i}kx} = \frac{1}{2} + \cos x + \cos 2x + \cdots + \cos nx = \frac{\sin(n+\frac{1}{2})x}{2\sin(\frac{1}{2}x)}.$$

Sometimes, a different normalization is used.

direct theorems of approximation derive estimates on the rate of approximation from smoothness properties of the approximand.

distance of the *element* a, in the metric space X with metric ρ, from the set B in X is the number

$$\operatorname{dist}(a, B) := \inf_{b \in B} \rho(a, b).$$

The distance of the *set* A in the metric space X from the set B in X is the number

$$\operatorname{dist}(A, B) := \sup_{a \in A} \operatorname{dist}(a, B).$$

The corresponding **Hausdorff distance**, between A and B, is the number

$$\rho_H(A, B) := \max(\mathrm{dist}(A, B), \mathrm{dist}(B, A)).$$

It provides a metric for the set of compact subsets of X.

distribution function for a function f defined on a measurable set $A \subset \mathbb{R}$ is the function

$$\mu_f(y) := \mathrm{meas}(\{x \in A : |f(x)| > y\}).$$

divided difference: Let $t_0, t_1, \ldots, t_n$ be distinct points on the real line (complex plane), and let f be a function defined at these points. The divided difference $[t_0, \ldots, t_n]f$ of f on these points is defined recursively as $[t_0]f = f(t_0)$, and

$$[t_0, \ldots, t_n]f = \frac{[t_1, \ldots, t_n]f - [t_0, \ldots, t_{n-1}]f}{t_n - t_0}.$$

For repeated nodes, one has to take appropriate limits.

One can avoid having to figure out these limits by first defining divided differences of a *polynomial* p for any sequence $t_0, t_1, \ldots$, distinct or not, in one fell swoop as the unique coefficients in the expansion

$$p(x) =: \sum_{i \geq 0} (x - t_0) \cdots (x - t_{i-1}) \, [t_0, \ldots, t_i]p$$

of p into a **Newton series**, with the divided difference $[t_0, \ldots, t_n]f$ of any sufficiently smooth function f then being defined to be that of any polynomial that agrees with f at $t_0, \ldots, t_n$, counting multiplicities as in *Hermite interpolation*.

Here is an alternative, more literal, notation for the divided difference:

$$\Delta(t_0, \ldots, t_n) := [t_0, \ldots, t_n].$$

Durrmeyer operators associate with a function f defined on $[0, 1]$ the polynomials

$$\sum_{k=0}^{n} \binom{n}{k} x^k (1 - x)^{n-k} \left((n + 1) \int_0^1 f(t) t^k (1 - t)^{n-k} \, \mathrm{d}t \right).$$

dyadic splines are splines with dyadic knots $t_j = j/2^n$, $j = 1, \ldots, 2^n - 1$ (or, say, $j = 0, \pm 1, \ldots$).

elementary symmetric polynomials: The sth elementary symmetric polynomial in the variables $x_1, \ldots, x_d$ is

$$\sigma_s(x_1, \ldots, x_d) := \sum_{1 \leq i_1 < i_2 < \cdots < i_s \leq d} x_{i_1} x_{i_2} \cdots x_{i_s}.$$

entire function is an analytic function f on the complex plane. If for every $\varepsilon > 0$ there is a constant $C = C_\varepsilon$ such that

$$|f(z)| \le C\mathrm{e}^{C|z|^{\rho+\varepsilon}}, \qquad z \in \mathbb{C},$$

then f is called **of order** ρ. It is called **of exponential type** M if

$$|f(z)| \le C\mathrm{e}^{M|z|}, \qquad z \in \mathbb{C},$$

for some constant C.

entropy: Let K be a compact subset of a metric space X and $\varepsilon > 0$. Let $M_\varepsilon(K)$ be the mininum number of subsets of X in an ε**-covering for** K, i.e., in a collection of subsets, each of diameter $\le 2\varepsilon$, whose union covers K. Then, $\log_2 M_\varepsilon(K)$ is called the ε**-entropy of** K (also **metric entropy**, to distinguish it from the probabilistic entropy).

Let $N_\varepsilon(K)$ be the minimum number of points in an ε**-net for** K, i.e., in a set $C \subset X$ with $\mathrm{dist}(K, C) \le \varepsilon$. Then, $\log_2 N_\varepsilon(K)$ is called the ε**-entropy of** K **with respect to** X.

For each K, the sequence (2ε-*capacity*, ε-entropy, ε-entropy wrto X, ε-*capacity*) is nondecreasing.

ε**-capacity of a set:** See *capacity*.

ε**-covering:** See *entropy*.

ε**-entropy:** See *entropy*.

ε**-net:** See *entropy*.

equimeasurable functions have the same *distribution functions*.

Euler splines are defined recursively as

$$\mathcal{E}_0(x) := (-1)^\nu, \qquad \nu - \frac{1}{2} < x < \nu + \frac{1}{2},$$

$$\mathcal{E}_{m+1}(x) := \int_{-1/2}^{1/2} \mathcal{E}_m(x+t)\,\mathrm{d}t \Big/ \int_{-1/2}^{1/2} \mathcal{E}_m(t)\,\mathrm{d}t.$$

existence set is any subset M of a metric space X such that to each element of X there exists a *best approximant* from M.

exponential type: See *entire function*.

extended Chebyshev system is a *Chebyshev system* $(f_1, \ldots, f_n)$ for which $\det([t_1, \ldots, t_i]f_j : i, j = 1, \ldots, n) > 0$ for every $t_1 \le \cdots \le t_n$ in the common domain of the f_j. Equivalently, for every $t_1 \le \cdots \le t_n$,

$$\det(D^r f_j(t_i) : r := \max\{s : t_{i-s} = t_i\};\ i, j = 1, \ldots, n) > 0.$$

extended complete Chebyshev system is sequence $(f_1, \ldots, f_n)$ of functions for which $(f_1, \ldots, f_m)$ is an extended *Chebyshev system* for $m = 1, \ldots, n$. If the common domain of the f_j is an interval, then this is equivalent to having the *Wronskian* of each $(f_1, \ldots, f_m)$, $m = 1, \ldots, n$, be positive on that interval,

exponential sums are functions of the form

$$x \mapsto \sum_{j=1}^{n} a_j \mathrm{e}^{\lambda_j x}.$$

exponential type: See *entire function*.

Favard numbers are

$$K_m := \frac{4}{\pi} \sum_{j=0}^{\infty} \frac{(-1)^{(m+1)j}}{(2j+1)^{m+1}}.$$

Fejér kernel of degree n is the function

$$\frac{1}{n+1} \sum_{k=0}^{n} D_k(x) = \frac{1}{2(n+1)} \left(\frac{\sin \frac{n+1}{2} x}{\sin \frac{x}{2}} \right)^2,$$

where D_k is the kth Dirichlet kernel. Sometimes, a different normalization is used.

Fejér means or $(C,1)$-means of a Fourier series (say) are the arithmetic means of the first n partial sums $(n = 0, 1, \ldots)$. See *Fejér kernel*.

Fekete points: Let $K \subset \mathbb{C}$ be a compact set and $n > 1$ an integer. nth Fekete points are defined as the entries in an n-sequence $(z_1, \ldots, z_n)$ in K that maximizes the product

$$\prod_{0 \le i < j \le n} |z_j - z_i|$$

over all such sequences (they are not unique).

Fekete polynomials of a compact set $K \subset \mathbb{C}$ are the polynomials

$$F_n(z) = \prod_{0 \le j \le n} (z - z_j)$$

where $\{z_1, \ldots, z_n\}$ are Fekete points for K.

forward difference: Let f be a function defined, say, on an interval (a, b), and let h be a real number. The forward differences $\Delta_h^r f$ of f with (positive) step-size h are defined recursively as $\Delta_h^0 f(x) := f(x)$,

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

whenever this expression has meaning. Without the subscript h, the step-size $h = 1$ is implied.

Fourier coefficients: See *Fourier series*.

Fourier series: Let f be a 2π-periodic function. Its Fourier series in complex form is

$$f(x) \sim \sum_{k=-\infty}^{\infty} c_k(f) \mathrm{e}^{\mathrm{i}kx}$$

where

$$c_k(f) := \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \mathrm{e}^{-\mathrm{i}kt} \, \mathrm{d}t$$

are the **Fourier coefficients** of f.

The real form is

$$f(x) \sim \frac{a_0(f)}{2} + \sum_{k=1}^{\infty} (a_k(f) \cos kx + b_k(f) \sin kx)$$

where

$$a_k(f) := \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos kt \, \mathrm{d}t$$

are the **cosine Fourier coefficients** and

$$b_k(f) := \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin kt \, \mathrm{d}t$$

are the **sine Fourier coefficients** of f.

Fourier series with respect to an orthogonal system: Let μ be a nonnegative measure and (φ_n) an orthogonal system in $L^2(\mu)$. The Fourier series of $f \in L^2(\mu)$ with respect to (φ_n) is

$$f(x) \sim \sum_n d_k(f)\varphi_n(x)$$

where

$$d_k(f) := \frac{1}{\gamma_n} \int f\overline{\varphi_n}\,\mathrm{d}\mu, \qquad \gamma_n := \int |\varphi_n|^2\,\mathrm{d}\mu$$

are the corresponding Fourier coefficients.

fundamental polynomials is a term used for the *basic polynomials of Lagrange interpolation*.

Gamma operators associate with a function f defined on $(0,\infty)$ the functions

$$G_n f(x) := \frac{x^{n+1}}{n!} \int_0^\infty \mathrm{e}^{-xu} u^n f\Big(\frac{n}{u}\Big)\,\mathrm{d}u$$

if these exist.

Gaussian quadrature formula: Let μ be a nonnegative measure on some interval $A \subset \mathbb{R}$, let $x_1, \ldots, x_n$ be the zeros of the nth *orthogonal polynomial* with respect to μ, and let l_j be the *basic polynomials of Lagrange interpolation* with respect to $\{x_i\}$. The quadrature formula

$$I(f) := w_1 f(x_1) + \cdots + w_n f(x_n),$$
$$w_j := \int l_j\,\mathrm{d}\mu = \int l_j^2\,\mathrm{d}\mu, \qquad j = 1, \ldots, n,$$

is called the corresponding Gaussian quadrature formula.

Alternatively, it can be defined as the unique quadrature formula on n points that is exact for polynomials of degree $< 2n$.

Gegenbauer polynomials are the same as *ultraspherical polynomials*, i.e., *Jacobi polynomials* with equal parameters. ($P_n^{(\alpha-1/2,\alpha-1/2)}$ is called the **Gegenbauer polynomial with parameter** α.)

generalized polynomials are functions of the form

$$f(z) = |a| \prod_{j=1}^n |z - z_j|^{m_j},$$

where $a, z_1, \ldots, z_n$ are complex numbers. $m_1 + \cdots + m_n$ is called the **degree** of f; since the m_j need not be integers, neither need the degree be.

Gram determinant: Let F be an inner product space. If $f_1, \ldots, f_n$ are in F, then their Gram determinant is $\det(\langle f_i, f_j \rangle : i, j = 1, \ldots, n)$.

Gram matrix is the matrix underlying the Gram determinant. More generally, any matrix of the form $(\lambda_i f_j)$, with $\lambda := (\lambda_1, \ldots, \lambda_m)$ a sequence of linear functionals on some vector space and $f := (f_1, \ldots, f_n)$ a sequence of elements of that vector space, is called the Gram matrix (for the sequences λ and f).

Haar functions are of the form $H(2^j x - k)/2^{j/2}$, with j and k integers and $H = \chi_{(0,1/2)} - \chi_{(1/2,1)}$.

Haar property: The defining property of a *Haar space*.

Haar space is any finite-dimensional space H of functions for which only the trivial element can vanish at $\dim H$ points. Equivalently, such a space provides a unique interpolant to arbitrary data at any $\dim H$ points in their common domain.

Haar system: any basis of a *Haar space*.

Hausdorff distance of two sets: see *distance*.

Hermite-Birkhoff interpolation: See *Birkhoff interpolation*.

Hermite interpolation: Let $x_0, x_1, \ldots, x_k$ be distinct points on the real line or the complex plane, and let

$$y_{i,r}, \quad 0 \le r < m_i, \;\; 0 \le i \le k,$$

be given sequences of values. The Hermite interpolating polynomial for these data is the unique polynomial p of degree $< m_0 + \cdots + m_k$ for which $p^{(r)}(x_i) = y_{i,r}$ for all $0 \le r < m_i$, $0 \le i \le k$.

Here is an alternative description of Hermite interpolation. Let $x_0, \ldots, x_n$ be points, distinct or not, on the real line or the complex plane, and let $y_0, \ldots, y_n$ be corresponding values. The Hermite interpolating polynomial for these data is the unique polynomial p of degree $\le n$ for which $p^{(r)}(x_i) = y_i$, with $r := r_i := \#\{j < i : x_j = x_i\}$, $i = 0, \ldots, n$.

Hermite polynomials are the orthogonal polynomials on $\mathbb{R}$ with respect to the weight function e^{-x^2}.

homogeneous polynomials in d variables are polynomials of the form

$$\sum_{|\mathbf{i}|=n} a_{\mathbf{i}} \mathbf{x}^{\mathbf{i}}$$

for some n, with $\mathbf{x} = (x_1, \ldots, x_d)$.

Hölder continuity/Hölder smoothness: See *Lipschitz continuity*.

Hölder classes: Classes of functions that satisfy a *Hölder smoothness* condition.

hypergeometric functions/series are of the form

$$ {}_2F_1(a, b; c; z) := \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{(c)_n} \frac{z^n}{n!}, $$

where $(a)_n$ is *Pochhammer's symbol*.

intermediary space: See *interpolation spaces*.

interpolation: See *Lagrange*, *Hermite*, or *Birkhoff interpolation*.

interpolation matrix: See *Birkhoff interpolation*.

interpolation of operators: See *interpolation spaces*.

interpolation spaces: Let X_0, X_1 be normed (or quasi-normed) spaces that are subsets of a larger linear space. X is called an interpolation space (or **intermediary space**) between them if any linear (or sublinear) operator that is bounded on both X_0 and X_1 is automatically bounded on X (example: if $1 \le p < r < q \le \infty$, then $L^r(\mu)$ is an interpolation space between $L^p(\mu)$ and $L^q(\mu)$).

inverse theorems of approximation derive smoothness properties from rates of approximation.

Jackson kernel: Let m be a positive integer. The corresponding Jackson kernel is defined as

$$ \frac{1}{d_m} \left(\frac{\sin \frac{n+1}{2} x}{\sin(x/2)} \right)^{2m}, $$

where

$$ d_m := \int_0^{2\pi} \left(\frac{\sin \frac{n+1}{2} t}{\sin(t/2)} \right)^{2m} \mathrm{d}t $$

is a normalizing constant.

Jacobi polynomials $P_n^{(\alpha,\beta)}$ with parameters $\alpha, \beta > -1$ are the *orthogonal polynomials* on $[-1, 1]$ with respect to the weight $(1-x)^{\alpha}(1+x)^{\beta}$.

K**-functional** between the spaces $X_1 \subset X_0$ equipped with the semi-norms $\|\cdot\|_{X_i}$, $i = 0, 1$, is defined as

$$K(f,t) := \inf_{g \in X_1} \left(\|f - g\|_{X_0} + t\|g\|_{X_1} \right).$$

Kantorovich polynomials: Let f be an integrable function on $[0,1]$. Its Kantorovich polynomial of degree n is

$$K_n f(x) := \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} \left((n+1) \int_{k/(n+1)}^{(k+1)/(n+1)} f \right).$$

knots of a spline f form the **knot sequence** $t = (t_i)$ in the representation $f =: \sum_i a_i N(\cdot|t_i, \dots, t_{i+k})$ of f as a weighted sum of *B-spline*s. Such a knot is **simple** if it occurs only once in t. Knots of a spline are also its *breakpoint*s or *break*s. But breaks have no multiplicity.

Lagrange interpolation: Let $x_0, x_1 \dots, x_n$ be different points on the real line or on the complex plane, and let $y_0, y_1, \dots, y_n$ be given values. The Lagrange interpolating polynomial for these data is the unique polynomial p of degree at most n such that $p(x_i) = y_i$ for $i = 0, \dots, n$.

Laguerre polynomials with parameter $\alpha > -1$ are the orthogonal polynomials on $[0, \infty)$ with respect to the weight $x^\alpha \mathrm{e}^{-x}$.

leading coefficient: See *polynomials*.

leading term: See *polynomials*, *polynomials in d variables*.

Lebesgue constant is the supremum norm of the corresponding *Lebesgue function*, hence the norm of the corresponding linear map on the relevant space of continuous functions.

Lebesgue constants (of Fourier sums) are the norms of the nth partial sum operators for the Fourier series considered as operators from the space of continuous functions into itself, i.e., the numbers

$$\sup_{\|f\|_\infty \le 1} \|S_n f\|_\infty, \qquad n = 0, 1, \dots,$$

where $S_n f$ denotes the nth partial sum of the Fourier series associated with f.

The definition is similar if one uses Fourier expansion into an arbitrary *orthogonal system*.

Lebesgue function of a linear map L on $C(T)$ for some set T is the function on T whose value at $t \in T$ is the norm of the linear functional $f \mapsto (Lf)(t)$, i.e., the number

$$\sup_{\|f\|_\infty \le 1} |Lf(t)|.$$

The Lebesgue function is useful since its supremum norm, also called the *Lebesgue constant* for L, coincides with the norm of L.

Lebesgue function (of Lagrange interpolation) with respect to an $(n+1)$-set $\{x_0, \ldots, x_n\}$ equals the sum of the absolute values of the *basic Lagrange interpolation polynomials*, i.e.,

$$x \mapsto \sum_{i=0}^{n} |l_i(x)|.$$

Legendre polynomials are the *orthogonal polynomials* on $[-1, 1]$ with respect to linear (Lebesgue) measure.

lemniscate of a function f (usually of several variables or of a complex variable) are the level sets of the form

$$\{x : |f(x)| = c\}.$$

linear n-width of a set Y in a normed linear space X is given by

$$\delta_n(Y; X) = \inf_P \sup_{y \in Y} \|y - Py\|,$$

where the P vary over all continuous linear operators on X of rank at most n.

Lipschitz continuity: The literature is ambiguous. Sometimes, functions with the property

$$|f(x) - f(x')| \le C|x - x'|$$

(where C is a fixed constant) are called Lipschitz continuous, sometimes those that satisfy

$$|f(x) - f(x')| \le C|x - x'|^\alpha$$

with some $\alpha > 0$. In the latter case the expressions **Hölder-continuous with exponent α** or **Lipschitz α-continuous** are also often used.

Lipschitz spaces: Spaces of functions satisfying a *Lipschitz condition*.

logarithmic capacity of a compact set K is $\mathrm{e}^{-V(K)}$ where $V(K)$ denotes the *logarithmic energy* of K. If K is not compact, then its logarithmic capacity is the supremum of the logarithmic capacity of its compact subsets.

logarithmic energy of a measure μ is

$$V(\mu) := \int\int \log \frac{1}{|z-t|}\, \mathrm{d}\mu(t)\, \mathrm{d}\mu(z).$$

logarithmic energy $V(K)$ of a set K is the infimum of the logarithmic energies of all (Borel) measures of total mass 1 supported on K.

logarithmic potential of a measure μ is the function

$$U^\mu(z) := \int \log \frac{1}{|z-t|}\, \mathrm{d}\mu(t).$$

Lorentz degree of a polynomial p is the smallest k for which p admits a representation

$$p(x) = \pm \sum_{j=0}^{k} a_j (1-x)^j (1+x)^{k-j}, \qquad a_j \geq 0.$$

If there is no such k, then the Lorentz degree is ∞.

For a trigonometric polynomial T, the Lorentz degree with respect to a point $\omega \in (0, \pi]$ is the smallest k such that T admits a representation

$$T(t) = \pm \sum_{j=0}^{2k} a_j \sin^j \frac{\omega - t}{2} \sin^{2k-j} \frac{\omega + t}{2}, \quad a_j \geq 0,$$

and again it is ∞ if there is no such representation.

Lorentz spaces: Let φ be a locally integrable function on $[0, \infty)$ and let $1 \leq q < \infty$. The associated Lorentz spaces consist of those functions f for which the norms

$$\|f\|_{\Lambda(\varphi,q)} := \left(\int_0^\infty \varphi(t) f^*(t)^q \, \mathrm{d}t \right)^{1/q}$$

resp.

$$\|f\|_{M(\varphi,q)} := \sup_c \left(\frac{1}{\Phi(c)} \int_0^c f^*(t)^q \, \mathrm{d}t \right)^{1/q}$$

are finite, where

$$\Phi(c) := \int_0^c \varphi(t)\, \mathrm{d}t,$$

and where f^* denotes the *decreasing rearrangement* of f.

main part modulus of smoothness is formed as a *φ-modulus of smoothness*, but the norm of the corresponding difference is taken only on part of the interval so that the arguments do not get too close to the endpoints. E.g. if $I = [0, 1]$, $\varphi(x) = \sqrt{x(1-x)}$, then the rth φ-modulus is

$$\omega_{r,\varphi}(f,t)_p = \sup_{h \le t} \left(\int |\delta^r_{h\varphi(x)} f(x)|^p \, \mathrm{d}x \right)^{1/p},$$

where the integration is taken for all values for which the integrand is defined, i.e., for all values of x such that $x \pm rh\varphi(x)/2 \in [0,1]$, while the main part modulus is defined as

$$\Omega_{r,\varphi}(f,t)_p := \sup_{h \le t} \left(\int_{2h^2}^{1-2h^2} |\delta^r_{h\varphi(x)} f(x)|^p \, \mathrm{d}x \right)^{1/p},$$

i.e., a small interval around the endpoints is omitted.

Markov function is of the form

$$z \mapsto \int \frac{\mathrm{d}\mu(t)}{z-t},$$

where μ is a compactly supported measure on the real line. See also *Cauchy transform*.

Markov inequality: If p is a real algebraic polynomial of degree at most n and $|p(x)| \le 1$ for all $x \in [-1,1]$, then

$$|p'(x)| \le n^2, \qquad x \in [-1,1].$$

This classic inequality was proven by A. A. Markov. A similar inequality for higher order derivatives was proven by his younger brother, V. A. Markov.

Markov system: See *complete Chebyshev system*.

mean periodic function is a function in $C(\mathbb{R}^d)$ the span of whose translates are not dense in $C(\mathbb{R}^d)$ in the topology of uniform convergence on compacta.

metric selection is a map P_M on a metric space X that associates with each element $x \in X$ a *best approximant* to x from the given subset M of X.

Meyer-König and Zeller operators associate with a function f defined on $[0,1)$ the functions

$$M_n f(x) := \sum_{k=0}^{\infty} f\left(\frac{k}{n+k}\right) \binom{n+k}{k} x^k (1-x)^{n+1}$$

if these exist.

minimax approximation is best approximation with respect to the uniform norm.

modulus of continuity is a nonnegative nondecreasing function ω defined in a right neighborhood of the origin with the properties that $\omega(0+) = 0$ and $\omega(t_1 + t_2) \le \omega(t_1) + \omega(t_2)$.

modulus of continuity of a function f defined, say, on an interval I is

$$\omega(f, t) := \sup_{x, x' \in I,\ |x - x'| \le t} |f(x) - f(x')|.$$

It can be written as

$$\omega(f, t) = \sup_{0 \le h \le t} \|\Delta_h^1 f\|, \tag{1}$$

where $\Delta_h^1 f$ is the first difference of f and $\|\cdot\|$ denotes the supremum norm. Note that the modulus of continuity of a function is a *modulus of continuity* iff the function is uniformly continuous.

The definition in other norms (like in L^p) is similar, just use that particular norm in (1).

modulus of smoothness: Let f be a function given, say, on an interval I. Its modulus of smoothness is defined as

$$\omega_2(f, t) := \sup_{x-h, x+h \in I,\ 0 \le h \le t} |f(x - h) - 2f(x) + f(x + h)|.$$

It can be written as

$$\omega_2(f, t) = \sup_{0 \le h \le t} \|\delta_h^2 f\|, \tag{2}$$

where $\delta_h^2 f$ is the second symmetric difference of f and $\|\cdot\|$ denotes the supremum norm. The definition in other norms (like in L^p) is similar, just use that particular norm in (2).

Higher order moduli of smoothness are analogously defined as

$$\omega_r(f, t) := \sup_{0 \le h \le t} \|\delta_h^r f\|,$$

with possible variations in what difference is used.

moments of a measure μ are the numbers

$$\int x^n \, \mathrm{d}\mu(x), \qquad n = 0, 1, \ldots.$$

Sometimes, n runs through (positive) real numbers.

moments of a functional λ are the numbers $\lambda(g_n)$ where $g_n(x) = x^n$.

monic polynomials are polynomials with *leading coefficient* 1.

monospline is a function of the form $a_m x^m + S(x)$, with $a_m \neq 0$ and S a *spline* of degree $< m$. Monosplines arise as Peano kernels for the error in quadrature formulæ.

monotone approximation: See *approximation with constraint*.

multi-index is a vector with nonnegative integer entries. The sum of the entries of the multi-index $\mathbf{i}$ is called its **degree** and is denoted $|\mathbf{i}|$.

multiple orthogonality: Let $\mu_1, \dots, \mu_k$ be measures on the complex plane. Multiple orthogonality with respect to these measures means that a function S is orthogonal to some given set of functions $\varphi_{0,j}, \dots, \varphi_{\nu_j - 1, j}$ with respect to μ_j for all $j = 1, \dots, k$:

$$\int \varphi_{i,j}(t) S(t) \, \mathrm{d}\mu_j(t) = 0, \qquad 0 \le i < \nu_j, \quad j = 1, \dots, k.$$

E.g., if $n = \nu_1 + \cdots + \nu_k$, then the polynomial p of degree at most n and multiply orthogonal with respect to the measures $\mu_1, \dots, \mu_k$ satisfies

$$\int t^i p(t) \, \mathrm{d}\mu_j(t) = 0, \qquad 0 \le i < \nu_j, \quad j = 1, \dots, k.$$

multiplicity of a zero t of a univariate function f is the smallest k for which $f^{(k)}(t) \neq 0$.

multipoint Padé approximation: Let $a_1, \dots, a_k$ be k complex numbers, $\nu_1, \dots, \nu_k$ nonnegative integers, and let f be a function analytic in a neighborhood of each a_j. Its (m, n) multipoint *Padé approximant* $[m/n]f$ at the points $a_1, \dots, a_k$ and of order $\nu_1, \dots, \nu_k$, respectively, is a rational function p_m/q_n of numerator degree at most m and denominator degree at most n such that $m + n + 1 = \nu_1 + \cdots + \nu_k$, and

$$q_n(z) f(z) - p_m(z)$$

has a zero at a_j of order $\ge \nu_j$ for each $j = 1, \dots, k$.

Müntz approximation is approximation by linear combinations of some system $\{x^{\lambda_k}\}$ of (possibly fractional) powers of x.

Müntz polynomial is of the form

$$\sum_{k=0}^{m} a_k x^{\lambda_k}.$$

Müntz rationals are ratios of *Müntz polynomials*.

Müntz space associated with a sequence $(\lambda_k : k = 0, \ldots, m)$ is the set of all *Müntz polynomials*

$$\sum_{k=0}^{m} a_k x^{\lambda_k}.$$

n-width in the sense of Kolmogorov of a set Y in a normed linear space X is given by

$$d_n(Y;X) := \inf_Z \sup_{y \in Y} \inf_{z \in Z} \|y - z\|,$$

the left-most infimum being taken over all n-dimensional subspaces Z of X.

n-width in the sense of Gel'fand of a set Y in a normed linear space X is given by

$$d^n(Y;X) := \inf_Z \sup_{y \in Y \cap Z} \|y\|,$$

the infimum being taken over all subspaces Z of X of codimension n.

natural spline is a *spline* of *order* $2m$ whose derivatives of order $m, m+1, \ldots, 2m-2$ vanish at the endpoints of the relevant interval.

Newton interpolation adds interpolation points one at a time, resulting in a *Newton series* for the interpolant.

Newton series: See *divided difference*.

numerical quadrature: See *quadrature formulæ or Gaussian quadrature*.

one-sided approximation occurs when the *approximant* is required to be everywhere no bigger (or no smaller) than the *approximand*.

operator semigroup is a one-parameter family of operators T_t, $t > 0$, with the property $T_t T_s = T_{t+s}$.

order of an entire function: See *entire function*.

order of a polynomial: A polynomial of order k is any polynomial of degree less than k. The collection of all (univariate) polynomials of order k is a vector space of dimension k. (The collection of all polynomials of degree k is not even a vector space.)

order of a spline: See *splines*.

orthogonal polynomials: Let μ be a (nonnegative) measure on the complex plane for which

$$\int |z|^m \,\mathrm{d}\mu(z), \quad m = 0, 1, 2, \ldots,$$

are finite. If (φ_n) is an *orthogonal system* with respect to μ and φ_n is a polynomial of *degree* n, then (φ_n) is called an **orthogonal polynomial system**. If $\varphi_n(x) = x^n + \cdots$, then it is called the **nth monic orthogonal polynomial**, while if (φ_n) is an *orthonormal system*, then the φ_n's are called **orthonormal polynomials**.

Orthogonal polynomials with respect to a *weight* function w are the same as orthogonal polynomials with respect to the measure $w(x)\,\mathrm{d}x$.

orthogonal system: Let μ be a (nonnegative) measure. A sequence (φ_n) of functions from $L^2(\mu)$ is called an orthogonal system if

$$\int \varphi_n \overline{\varphi_m} \,\mathrm{d}\mu = 0 \quad \Leftrightarrow \quad n \neq m.$$

It is called an **orthonormal system** if, in addition,

$$\int |\varphi_n|^2 \,\mathrm{d}\mu = 1$$

for all n.

(φ_n) is called **complete in some space** X if there is no $g \in X$ other than the zero element that is orthogonal to every φ_n.

orthonormal system: See *orthogonal system*.

Padé approximation: Let f be a function analytic in a neighborhood of the origin. Its (m, n) Padé approximant $[m/n]f$ at the origin is a rational function p_m/q_n of numerator degree at most m and denominator degree at most n such that

$$q_n(z)f(z) - p_m(z)$$

has a zero at the origin of order $\geq m+n+1$. Padé approximants at other points (including infinity) are defined analogously.

Padé table for f is the infinite matrix $([m/n]f : m, n \geq 0)$ consisting of the $[m/n]f$ Padé approximants.

Peano kernel: Let λ be a continuous linear functional on $C[a, b]$ that vanishes on all polynomials of *order* n. Then, for n times differentiable functions f,

$$\lambda(f) = \int_a^b K(t) f^{(n)}(t) \, \mathrm{d}t$$

for some function K, which is called the Peano kernel associated with λ.

Peetre's functional: See *K-functional*.

perfect spline is a spline with simple knots whose highest nontrivial derivative is constant in absolute value.

φ-moduli of smoothness: Let φ and f be functions on an interval I. The rth order φ-modulus of smoothness of f in L^p is defined as

$$\omega_{r,\varphi}(f, t)_p := \sup_{0 \leq h \leq t} \|\delta^r_{h\varphi(\cdot)} f(\cdot)\|_p,$$

with $\delta^r_{h\varphi(s)}$ the rth *symmetric difference* of f with step $h\varphi(s)$; other differences may also be used.

Pochhammer's symbol $(a)_n$ stands for

$$\frac{a(a+1)\cdots(a+n-1)}{n!}.$$

They are also called **rising factorials**.

Poisson kernels are the members of the family

$$p_r(t) := \frac{1}{2\pi} \frac{1 - r^2}{1 - 2r\cos t + r^2}, \quad 0 \leq r < 1.$$

polar set is a set of *logarithmic capacity* 0.

Pólya system: Let $w_0, w_1, \ldots, w_n$ be strictly positive functions on an interval I and $c \in I$. The following sequence of functions is called the corresponding Pólya system:

$$\begin{aligned}
u_0(x) &:= w_0(x) \\
u_1(x) &:= w_0(x) \int_c^x w_1(t_1) \, \mathrm{d}t_1 \\
&\vdots \\
u_n(x) &:= w_0(x) \int_c^x w_1(t_1) \int_c^{t_1} w_2(t_2) \cdots \int_c^{t_{n-1}} w_n(t_n) \, \mathrm{d}t_n \cdots \mathrm{d}t_1.
\end{aligned}$$

polynomial form: One of several ways of writing a polynomial. Most are in terms of some basis, like the Newton form, the (local) power form, the Lagrange form, the Bernstein-Bézier form. In the univariate case, there is at least one other useful form, the root form.

polynomials are functions that can be written in **power form**

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0$$

where the a_i's are real or complex numbers and x is a real or complex variable. If $a_n \neq 0$, then n is called the **degree** of the polynomial, a_n its **leading coefficient**, and $a_n x^n$ its **leading term**.

polynomials in d variables are functions that can be written in **power form**

$$p(\mathbf{x}) = \sum_{|\mathbf{i}| \le n} a_{\mathbf{i}} \mathbf{x}^{\mathbf{i}}$$

where $\mathbf{i} = (i_1, \ldots, i_d)$ is a *multi-index*, the **coefficients** $a_{\mathbf{i}}$ are real or complex numbers, and

$$\mathbf{x}^{\mathbf{i}} := x_1^{i_1} \cdots x_d^{i_d},$$

with $\mathbf{x} = (x_1, \ldots, x_d)$ the d-vector whose entries are the d (real or complex) variables for the polynomial. If $a_{\mathbf{i}} \neq 0$ for at least one $\mathbf{i}$ with $|\mathbf{i}| = n$ then n is called the **degree** of the polynomial, and the *homogeneous polynomial*

$$p_{\uparrow}(x) := \sum_{|\mathbf{i}| = n} a_{\mathbf{i}} \mathbf{x}^{\mathbf{i}}$$

is called its **leading term**.

positive operators map nonnegative functions to nonnegative functions.

Post-Widder operators associate with a function f defined on $(0, \infty)$ the functions

$$P_\lambda f(x) := \frac{(\lambda/x)^\lambda}{\Gamma(\lambda)} \int_0^\infty \mathrm{e}^{-\lambda u/x} u^{\lambda-1} f(u) \, \mathrm{d}u$$

provided these exist.

projection constants: The **(relative) projection constant** $\lambda(V, X)$ of a vector space V relative to a normed space X containing it is the infimum of the norms of linear *projector*s from X onto V. The **absolute projection constant** $\lambda(V)$ of such a vector space V is the infimum over all projection constants relative to any space X into which V can be isometrically embedded.

projector is an operator T that is **idempotent**, i.e., it coincides with its square: $T(T(f)) = T(f)$ for all f.

proximinal set: See *existence set*.

quadratic splines are *splines* of *degree* 2, but it has become customary to refer to any spline of *order* 3 as a quadratic spline.

quadrature formulæ: Let μ be a nonnegative measure on some set A, $x_0, \ldots, x_n$ points in A, and $w_0, \ldots, w_n$ real or complex numbers. These define a quadrature formula

$$I(f) := w_0 f(x_0) + \cdots + w_n f(x_n)$$

to approximate the integral

$$\int f \, \mathrm{d}\mu.$$

If $A \subset \mathbb{R}$ (or $A \subset \mathbb{C}$), then the quadrature is called **exact of degree** m if

$$w_0 f(x_0) + \cdots + w_n f(x_n) = \int f \, \mathrm{d}\mu$$

for each f that is a *polynomial* of *degree* at most m.

For given $x_0, \ldots, x_n$, there is only one such quadrature formula that is exact of degree n, and it necessarily has the weights

$$w_j = \int l_j \, \mathrm{d}\mu, \qquad j = 0, 1, \ldots, n,$$

where l_j are the *basic polynomials of Lagrange interpolation*. Such a quadrature formula is called **interpolatory**, and the w_j's are called the associated **Cotes numbers**.

quasi-analytic classes: Let (a, b) be some interval and (n_k) a subsequence of the natural numbers. The corresponding quasi-analytic class consists of all infinitely differentiable functions f such that

$$|f^{(n_k)}(x)|^{1/n_k} \le M n_k, \qquad x \in (a, b), \ k = 1, 2, \ldots,$$

where M is a fixed constant (which depends on f).

quasi-interpolants are bounded linear operators Q on some normed linear space of functions (on some domain Ω in $\mathbb{R}$ or $\mathbb{R}^d$) that **reproduce** all polynomials p (i.e., $Qp = p$) of a certain *order* and are **local** in the sense that, for some $r > 0$ and every f and every set $C \subset \Omega$, Qf vanishes on C if f vanishes on $C + rB$ (with B the unit ball).

radial basis function is a function in $\mathbb{R}^d$ of the form

$$\mathbf{x} \mapsto g(\|\mathbf{x} - \mathbf{a}\|)$$

where g is a univariate function, $\mathbf{a}$ is a point in $\mathbb{R}^d$, and $\|\cdot\|$ denotes the Euclidean norm in $\mathbb{R}^d$.

rational functions are ratios of polynomials.

rearrangement-invariant space is a function space X with the property that if $f \in X$ and $|g|$ is *equimeasurable* with $|f|$, then $g \in X$.

regularity of Birkhoff interpolation: Let $X = (x_1, \ldots, x_m)$ be a sequence of distinct points and E a corresponding *interpolation matrix*. The pair E, X is called **regular** if the corresponding *Birkhoff interpolation* is always solvable regardless of what values are prescribed. E is called **regular** if E, X is regular for each sequence X of m distinct nodes.

reproducing kernels for an orthonormal system (p_n) are the functions

$$K_m(x, y) := \sum_{k=0}^{m} p_k(x) p_k(y), \quad m = 0, 1, \ldots .$$

ridge function is a function in $\mathbb{R}^d$ of the form $f(\mathbf{x}) = g(\mathbf{x} \cdot \mathbf{a})$ where g is a univariate function, $\mathbf{a}$ is a nonzero vector in $\mathbb{R}^d$, and

$$\mathbf{x} \cdot \mathbf{a} = x_1 a_1 + \cdots + x_d a_d$$

is the inner product.

rising factorials are the same as the *Pochhammer symbol*s.

saturation: A sequence (L_n) of approximation processes/operators is said to be **saturated of order** (ρ_n) if

$$\|L_n f - f\| = o(\rho_n)$$

holds only for functions in a small special class (called **trivial class**), while

$$\|L_n f - f\| = O(\rho_n)$$

holds for functions in a much larger class (called **saturation class**).

Schauder basis for a normed linear space X is a sequence $(f_1, f_2, ...)$ such that every element $f \in X$ has a unique representation of the form $f = \sum_n a_n f_n$ (with the sum converging in norm).

Schoenberg space $\mathcal{S}_k(b, m, I)$, with the *breakpoint* sequence b strictly increasing (finite or not), consists of all functions f on the interval I (finite or not) that, on each interval (b_i, b_{i+1}), agree with some polynomial of *order* k, and have *smoothness* $\geq m_i$ *at* b_i in the sense that the jump of $D^j f$ across b_i is zero for $0 \leq j < m_i$, all i. It has become customary to define such f (and its derivatives) at a breakpoint by right continuity, except when that breakpoint is the right-most one (if any), in which case f and its derivatives are defined there by left continuity.

The elements of the Schoenberg space are called *splines* or **piecewise polynomials**. By a basic result (of Curry and Schoenberg), if I and the part of b in I are finite, then the elements of $\mathcal{S}_k(b, m, I)$ are splines of order k with knot sequence t, with t the smallest nondecreasing sequence that contains $b_i \in I$ exactly $k - m_i$ times, all i, and contains the end points of I k times.

self-reciprocal polynomials are of the form

$$\sum_{k=0}^{n} a_k x^k$$

with $a_k = a_{n-k}$.

shape preservation means that certain properties (like monotonicity or convexity) are preserved by the approximating process.

Shepard operators associate with a function f defined on $[0, 1]$ the rational functions

$$\frac{\sum_{k=0}^{n} f(k/n)(x - k/n)^{-2}}{\sum_{k=0}^{n} (x - k/n)^{-2}}.$$

Sometimes, nodes other than k/n or other powers $|x - k/n|^{-p}$, $p > 0$, are used.

shift-invariant space is a function space on $\mathbb{R}$ (or $\mathbb{R}^d$) such that if f belongs to the space then so does $f(\cdot - n)$ for every integer (vector) n.

simultaneous approximation means approximating a function and some of its derivatives by another function and its corresponding derivatives.

smoothness of spline: See *breakpoint*.

Sobolev spaces: Let I be an interval on $\mathbb{R}$, $r \geq 1$ an integer and $1 \leq p \leq \infty$. The Sobolev space $W^r(L^p(I))$ is the family of all functions f on I that have an absolutely continuous $(r-1)$st derivative, and for which the Sobolev norm $\|f\|_p + \|f^{(r)}\|_p$ is finite.

The definition is similar in several variables, just there one has to add the norms of all mixed derivatives up to order r.

splines: A (univariate) spline of **order** k with **knot sequence** $t = (t_i)$ is any weighted sum of the corresponding *B-splines* $N(\cdot|t_i, \ldots, t_{i+k})$. Each knot t_i may be a breakpoint for such a spline f, with the *smoothness* of f at t_i no smaller than the order k minus the multiplicity with which the number t_i occurs in t. This multiplicity is sometimes called the **defect** of the spline at t_i.

Steklov means of a function f defined on an interval are the averages

$$\frac{1}{h}\int_x^{x+h} f(u)\,\mathrm{d}u$$

or their variants when the integration is going from $x - h$ to x, etc.

strong unicity of best approximation means that if Pf is the best approximation to f from a given subspace X, then for every f there is a constant $\gamma > 0$ such that

$$\|f - g\| \geq \|f - Pf\| + \gamma\|g - Pf\|$$

for every $g \in X$.

supremum norm of a function f on a set A is

$$\|f\|_A := \sup_{x\in A} |f(x)|.$$

If A is understood from the context, then

$$\|f\|_\infty$$

is used.

symmetric difference: Let f be a function defined, say, on an interval (a, b), and let h be a real number. The symmetric differences $\delta_h^r f(x)$ of f with **stepsize** h are defined recursively as $\delta_h^0 f(x) := f(x)$,

$$\begin{aligned}\delta_h^r f(x) &:= \delta_h^{r-1} f(x + h/2) - \delta_h^{r-1} f(x - h/2)\\ &= \sum_{k=0}^{r} (-1)^k \binom{r}{k} f(x + (r/2 - k)h)\end{aligned}$$

whenever this expression has meaning. Sometimes, the symmetric difference is called **central difference**.

Szász-Mirakjan operators associate with a function f defined on $[0, \infty)$ the functions

$$S_n f(x) := \mathrm{e}^{-nx} \sum_{k=0}^{\infty} f\left(\frac{k}{n}\right) \frac{(nx)^k}{k!}$$

provided they exist.

three-term recursion: Orthonormal polynomials $\{p_n\}$ on the real line with respect to any measure satisfy a three-term recursion formula

$$xp_n(x) = \frac{\gamma_n}{\gamma_{n+1}} p_{n+1}(x) + b_n p_n(x) + \frac{\gamma_{n-1}}{\gamma_n} p_{n-1}(x),$$

where γ_n is the *leading coefficient* of p_n and b_n are some real numbers.

total positivity of a function K on $[a,b] \times [c,d]$ means that all the determinants $\det(K(s_i,t_j) : 0 \le i,j \le n)$, $n = 0,1,\ldots$, are nonnegative, where $a \le s_0 < s_1 < \cdots < s_n \le b$ and $c \le t_0 < t_1 < \cdots < t_n \le d$.

total positivity of a matrix means that all its minors (i.e., determinants of its square submatrices) are nonnegative.

transfinite diameter of a set K in a normed space is defined as the limit as $n \to \infty$ of

$$\sup_{x_1,\ldots,x_n \in K} \left(\prod_{i<j} \|x_i - x_j\| \right)^{2/n(n-1)}.$$

translation-invariant space is a function space on $\mathbb{R}^d$ such that if f belongs to the space then so does $f(\cdot - a)$ for every $a \in \mathbb{R}^d$.

trigonometric polynomials: Functions of the form

$$\sum_{k=-n}^{n} c_k \mathrm{e}^{\mathrm{i}kx}$$

are called trigonometric polynomials of **degree** $\le n$. Their real form is

$$\sum_{k=0}^{n} (a_k \cos kx + b_k \sin kx).$$

The **degree** is n if $c_n \ne 0$ or $c_{-n} \ne 0$ ($a_n \ne 0$ or $b_n \ne 0$).

trigonometric polynomials in d **variables:** Functions of the form

$$\mathbf{x} \mapsto \sum_{|\mathbf{j}| \le n} c_{\mathbf{j}} \mathrm{e}^{\mathrm{i}\,\mathbf{j}\cdot\mathbf{x}}$$

with $\mathbf{j}$ a multi-index and $\mathbf{x} = (x_1,\ldots,x_d)$, are called trigonometric polynomials of d variables of **degree** $\le n$. Their real form is a linear combination of terms

$$\sigma(j_1x_1)\sigma(j_2x_2)\cdots\sigma(j_dx_d)$$

with $|\mathbf{j}| \le n$, where σ stands for sin or cos independently at each occurrence.

trivial class: See *saturation*.

truncated power function is defined as $x \mapsto (\max(0,x))^k =: x_+^k$.

ultraspherical polynomials (often called *Gegenbauer polynomials*) are *Jacobi polynomials* with equal parameters $\alpha = \beta$.

unicity space: Y is a unicity space in a normed linear space X if every $f \in X$ has a unique best approximant from Y. Sometimes, a unicity space is referred to as a *Chebyshev space*.

Voronovskaya: Let (L_n) be an operator sequence that is *saturated* of order (ρ_n). If for an f the limit

$$\lim_{n\to\infty} \rho_n^{-1}(L_n(f,x) - f(x)) =: \varphi(x)$$

exists, then it is called the Voronovskaya of f with respect to (L_n) and (ρ_n).

variation diminishing property of a process taking f to g is the property that g has no more sign changes than f.

wavelet is a function ψ on $(-\infty,\infty)$ such that the system

$$\psi(2^k x - j), \qquad k, j = 0, \pm 1, \pm 2, \dots \tag{3}$$

forms a basis in $L^2(-\infty,\infty)$. It is called an **orthogonal wavelet** if (3) constitutes an orthogonal basis in $L^2(-\infty,\infty)$.

weak Chebyshev system is a sequence $(\varphi_1,\dots,\varphi_n)$ of functions on an interval I such that for no nontrivial linear combination σ of these functions are there $n+1$ points $x_0 < x_1 < \dots < x_n$ in I with $\sigma(x_i)\sigma(x_{i+1}) < 0$, $0 \le i < n$.

Weierstrass operators (or integrals) associate with a function f defined on $(-\infty,\infty)$ the function

$$(W_\sigma f)(x) := \frac{1}{\sigma\sqrt{\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{-(x-t)^2/\sigma^2} f(t)\,\mathrm{d}t$$

(provided it exists), where σ is a positive parameter.

weight usually means a nonnegative function.

weighted approximation means that a weight is used in the norm.

weighted moduli of smoothness are *moduli of smoothness* formed with *weights*. E.g., if w is a *weight* then

$$\omega_r(f,t)_{p,w} := \sup_{0\le h\le t} \|w\delta_h^r f\|_p$$

is the corresponding rth weighted modulus of smoothness in L_w^p with possible variations in what difference is used.

In a similar fashion, weighted φ-moduli are defined as

$$\omega_{r,\varphi}(f,t)_{p,w} = \sup_{0\le h\le t} \|w\delta_{h\varphi(\cdot)}^r f\|_p,$$

with possible variations in what difference is used.

Wronskian of a sequence $(\varphi_0,\ldots,\varphi_m)$ of m times differentiable functions on an interval I is the determinant $\det(\varphi_j^{(i)}(t) : 0\le i,j\le m)$ (as a function of t).

zero counting measure is the measure that puts mass k at every zero of *multiplicity* k.

Zygmund class: The class of functions with $\omega_2(f,t) = O(t)$, where $\omega_2(f,t)$ is the (second order) *modulus of smoothness* of f.