Banach Spaces

Definition Banach space is a complete normed vector space B over a field $F=\{\mathbb{R},\mathbb{C}\}$,

Vector space: $\exists 0\in B$, $\forall f,g\in B$, $\alpha,\beta\in F \Rightarrow \alpha f+\beta g\in B$.
Norm:

i. $\|f\|\geq 0$ and $\|f\|=0 \Leftrightarrow f=0$.
ii. $f\in B$ only if $\|f\|_B<\infty$.
iii. $\|\alpha f\|=|\alpha|\|f\|$, $\forall\alpha\in F,\forall f\in B$.
iv. Triangle inequality $\|f+g\|\leq\|f\|+\|g\|$.

Complete: Every Cauchy sequence in B converges to an element of B.

Measure

In this course we only use the standard Lebesgue measure $\leftrightarrow$ the volume of a (measurable) set.

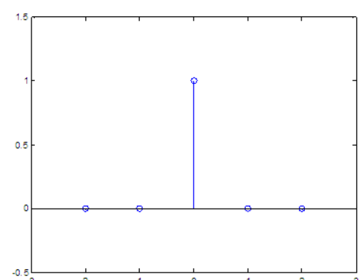
Example: $\Omega=[0,2]^n\subset\mathbb{R}^n$, $\mu(\Omega)=|\Omega|=2^n$.

We will need the notion of zero measure (volume). Example: a set of discrete points

Lp Spaces

$\Omega\subseteq\mathbb{R}^n$ domain. Examples: $\Omega=[a,b]\subset\mathbb{R}$, $\Omega=[0,1]^n\subset\mathbb{R}^n,\Omega=\mathbb{R}^n$.

$$\|f\|_{L_p(\Omega)}:=\begin{cases}\left(\int_\Omega|f(x)|^p\,dx\right)^{1/p}, & 0<p<\infty,\\ ess\sup\limits_{x\in\Omega}|f(x)|, & p=\infty.\end{cases}$$



$$ess\sup_x|f(x)|:=\sup_{A>0}\left\{A>0:\ \left|\{x:|f(x)|\geq A\}\right|>0\right\}.$$

For $1\leq p\leq\infty$, $L_p(\Omega)$ are Banach spaces.
For $0<p<1$, $L_p(\Omega)$ are Quasi-Banach spaces (quasi-triangle inequality holds)

$$\|f+g\|_p^p\leq\|f\|_p^p+\|g\|_p^p.$$

Theorem [Hölder] $1\leq p\leq\infty$, $f\in L_p, g\in L_{p'}$

$$\left|\int_\Omega fg\right|\leq\int_\Omega|fg|=\|fg\|_1\leq\|f\|_p\|g\|_{p'}\qquad \frac{1}{p}+\frac{1}{p'}=1.$$

Lemma Young's inequality for products,

$$ab \le \frac{a^p}{p} + \frac{b^{p'}}{p'}, \qquad \frac{1}{p} + \frac{1}{p'} = 1, \ \forall a,b \ge 0.$$

Proof of lemma The logarithmic function is concave. Therefore

$$\begin{aligned}\log\left(\frac{1}{p}a^p + \frac{1}{p'}b^{p'}\right) &= \log\left(\frac{1}{p}a^p + \left(1-\frac{1}{p}\right)b^{p'}\right) \\ &\ge \frac{1}{p}\log\left(a^p\right) + \frac{1}{p'}\log\left(b^{p'}\right) \\ &= \log(a) + \log(b) = \log(ab).\end{aligned}$$

Since the logarithmic function is increasing, we are done (or we take exp on both sides).

□

Proof of theorem If $p = \infty$

$$\int_\Omega |fg| \le \|f\|_\infty \int_\Omega |g| \le \|f\|_\infty \|g\|_1.$$

The proof is similar for $p = 1$. So, assume now $1 < p < \infty$ and $\|f\|_p = \|g\|_{p'} = 1$.

Integrating pointwise and applying Young's inequality almost everywhere, gives

$$\begin{aligned}\int_\Omega |f(x)g(x)|dx &\le \int_\Omega \left(\frac{|f(x)|^p}{p} + \frac{|g(x)|^{p'}}{p'}\right)dx \\ &= \frac{1}{p}\int_\Omega |f(x)|^p dx + \frac{1}{p'}\int_\Omega |g(x)|^{p'} dx \\ &= \frac{1}{p} + \frac{1}{p'} = 1\end{aligned}$$

Now assuming $f, g \ne 0$ (else, we are done)

$$\int_\Omega \frac{|f(x)|}{\|f\|_p}\frac{|g(x)|}{\|g\|_{p'}}dx \le 1 \Rightarrow \int_\Omega |fg| \le \|f\|_p \|g\|_{p'}$$

□

Schwartz inequality $p = 2$

$$\langle f, g\rangle_2 = \left|\int_\Omega f\bar{g}\right| \le \int_\Omega |fg| = \|fg\|_1 \le \|f\|_2 \|g\|_2.$$

The L_p spaces are not comparable on unbounded domains

Example We will use $\Omega = \mathbb{R}$. Assume $0 < q < p < \infty$

Choose

$$f(x) := \begin{cases} 0, & |x| \le 1, \\ \dfrac{1}{|x|^{1/q}}, & |x| > 1. \end{cases}$$

We have $f \in L_p(\mathbb{R})$, $f \notin L_q(\mathbb{R})$

Now choose

$$f(x) := \begin{cases} \dfrac{1}{|x|^{1/p}}, & |x| \le 1, \\ 0, & |x| > 1. \end{cases}$$

We have $f \in L_q(\mathbb{R})$, $f \notin L_p(\mathbb{R})$

Theorem If $|\Omega| < \infty$, $0 < q < p$, $f \in L_p(\Omega)$ then

$$\| f \|_{L_q(\Omega)} \le |\Omega|^{1/q - 1/p} \| f \|_{L_p(\Omega)} .$$

Proof For $p = \infty$

$$\| f \|_q = \left(\int_\Omega | f(x) |^q dx \right)^{1/q} \le \| f \|_\infty \left(\int_\Omega dx \right)^{1/q} = \| f \|_\infty |\Omega|^{1/q}$$

For $q < p < \infty$ define $r := p/q \ge 1$

$$\| f \|_q^q = \int_\Omega | f |^q = \int_\Omega | f |^q 1 \underset{Holder}{\le} \left(\int_\Omega \left(| f |^q \right)^r \right)^{1/r} \left(\int_\Omega 1^{r'} \right)^{1/r'}$$
$$= \left(\int_\Omega | f |^p \right)^{q/p} |\Omega|^{1-q/p}$$

□

Theorem [Minkowski] For $1 \le p \le \infty$, $\forall f, g \in L_p$,

$$\| f + g \|_p \le \| f \|_p + \| g \|_p \ .$$

Proof for $1 < p < \infty$ ($p = 1, \infty$ is easier). W.l.g $f, g \ge 0$. We apply Hölder twice,

$$\begin{aligned} \int (f+g)^p &= \int f (f+g)^{p-1} + \int g (f+g)^{p-1} \\ &\le \left(\| f \|_p + \| g \|_p \right) \left(\int (f+g)^{(p-1)p'} \right)^{1/p'} \\ &= \left(\| f \|_p + \| g \|_p \right) \left(\int (f+g)^p \right)^{1-1/p} \\ &= \left(\| f \|_p + \| g \|_p \right) \int (f+g)^p \left(\int (f+g)^p \right)^{-1/p} . \end{aligned}$$

□

Theorem For $0 < p < 1$, we have

(i) $\left\| \sum_k f_k \right\|_p^p \le \sum_k \| f_k \|_p^p$.

(ii) $\| f + g \|_p \le 2^{1/p-1} \left(\| f \|_p + \| g \|_p \right)$ or in general $\left\| \sum_{k=1}^N f_k \right\|_p \le N^{1/p-1} \sum_{j=1}^N \| f_k \|_p$.

Proof The quasi-triangle inequality (ii) is derived from (i). Observe first

$$\left.\begin{array}{c} 1 < r := \frac{1}{p} < \infty \\ \frac{1}{r} + \frac{1}{r'} = 1 \end{array}\right\} \Rightarrow r' = \frac{1}{1-p}$$

Then

$$\left\| \sum_{k=1}^{N} f_k \right\|_p \underset{(i)}{\leq} \left(\sum_{k=1}^{N} \| f_k \|_p^p \right)^{1/p} = \left(\sum_{k=1}^{N} 1 \cdot \| f_k \|_p^p \right)^{1/p} \underset{\substack{\text{Discrete} \\ \text{Holder}}}{\leq} \left(\sum_{k=1}^{N} 1^{\frac{1}{1-p}} \right)^{(1-p)1/p} \left(\sum_{k=1}^{N} \| f_k \|_p \right) = N^{1/p-1} \sum_{k=1}^{N} \| f_k \|_p$$

To prove (i), we need the following lemma

Lemma I For $0 < p \leq 1$ and any sequence of non-negative $a = \{a_k\}$,

$$\left(\sum_k a_k \right)^p \leq \sum_k a_k^p$$

Proof Observe that it is sufficient to prove $(a_1 + a_2)^p \leq a_1^p + a_2^p$ and then apply induction.

To prove the inequality use $h(t) := t^p + 1 - (t+1)^p$ for $t \geq 0$. $h(0) = 0$ and $h'(t) = pt^{p-1} - p(t+1)^{p-1} \geq 0$.

Therefore, $h(t) \geq 0$, for $t \geq 0$. This gives $t^p + 1 \geq (t+1)^p$. Setting $t = a_1 / a_2$ gives

$$\left(\frac{a_1}{a_2} \right)^p + 1 \geq \left(\frac{a_1}{a_2} + 1 \right)^p \Rightarrow a_1^p + a_2^p \geq (a_1 + a_2)^p .$$

□

Proof of Theorem (i) : Simply apply the lemma pointwise for $x \in \Omega$ and then Tonelli's theorem for the exchange of integration and sum

$$\left\| \sum_k f_k \right\|_p^p \leq \int_\Omega \left(\sum_k | f_k(x) | \right)^p dx \leq \int_\Omega \left(\sum_k | f_k(x) |^p \right) dx = \sum_k \int_\Omega | f_k(x) |^p \, dx = \sum_k \| f_k \|_p^p .$$

□

Definition The space $l_p(\mathbb{Z})$, $0 < p \leq \infty$, is the space of sequences $a = \{a_k\}_{k \in \mathbb{Z}}$, for which the following norm is finite

$$\|a\|_{l_p} := \begin{cases} \left(\sum_k |a_k|^p \right)^{1/p}, & 0 < p < \infty, \\ \sup_k |a_k|, & p = \infty. \end{cases}$$

Lemma II $l_p \subset l_q$ for $p \leq q$. That is, for any sequence $a = \{a_k\}$

$$\|a\|_{l_q} \leq \|a\|_{l_p} .$$

Proof Case of $q = \infty$, for any $j \in \mathbb{Z}$,

$$|a_j| = \left(|a_j|^p \right)^{1/p} \leq \left(\sum_k |a_k|^p \right)^{1/p} = \|a\|_{l_p} .$$

Therefore,

$$\|a\|_{l_\infty} = \sup_j |a_j| \leq \|a\|_{l_p} .$$

For $q<\infty$, we have

$$\left(\sum_k |a_k|^q\right)^{p/q} \le \sum_k \left(|a_k|^q\right)^{p/q} = \sum_k |a_k|^p \Rightarrow \left(\sum_k |a_k|^q\right)^{1/q} \le \left(\sum_k |a_k|^p\right)^{1/p} .$$

□

Hilbert spaces and $L_2(\Omega)$

Def Hilbert space H : Complete metric vector space induced by an inner product $\langle , \rangle : H \times H \to \mathbb{C}$.
Properties of the inner product:

i. symmetric $\langle f,g\rangle = \overline{\langle g,f\rangle}$,
ii. linear $\langle \alpha f_1 + \beta f_2, g\rangle = \alpha\langle f_1,g\rangle + \beta\langle f_2,g\rangle$,
iii. Positive definite $\langle f,f\rangle \ge 0$, with $\langle f,f\rangle = 0 \Leftrightarrow f=0$.

The natural norm $\|f\|_H := \langle f,f\rangle^{1/2}$ satisfies

(i) Cauchy-Schwartz

$$|\langle f,g\rangle| \le \|f\|_H \|g\|_H .$$

(ii) Triangle inequality (is based on Cauchy-Schwartz)

$$\|f+g\|^2 = \|f\|^2 + 2re\langle f,g\rangle + \|g\|^2 \le \|f\|^2 + 2\|f\|\|g\| + \|g\|^2 = \left(\|f\| + \|g\|\right)^2 .$$

So, a Hilbert space is a Banach space.

Examples

(i) $l_2(\mathbb{Z})$: $\langle \alpha,\beta\rangle_{l_2} := \sum_{i\in\mathbb{Z}} \alpha_i \bar{\beta}_i$, $\|\alpha\|_2 := \left(\sum_{i\in\mathbb{Z}} |\alpha_i|^2\right)^{1/2}$

(ii) $L_2(\Omega)$: f,g measurable, $\langle f,g\rangle := C_\Omega \int_\Omega f(x)\overline{g(x)}dx$,

$$\|f\|_{L_2(\Omega)} = \|f\|_2 = \langle f,f\rangle^{1/2} = \left(C_\Omega \int_\Omega |f(x)|dx\right)^{1/2} .$$

For $\Omega = \mathbb{R}^n, C_\Omega = 1$. For $\Omega = [-\pi,\pi]^n, C_\Omega = \frac{1}{(2\pi)^n}$.

Sobolev spaces

Multivariate derivatives: A partial derivative of order m

$$\alpha = (\alpha_1,\ldots,\alpha_n) \in \mathbb{Z}_+^n, \quad D^\alpha f = \frac{\partial^m f}{\partial x_1^{\alpha_1}\cdots\partial x_n^{\alpha_n}}, \quad |\alpha| := \sum_{i=1}^n \alpha_i = m .$$

Definition $C^m(\Omega)$: The space of all continuously differentiable functions of degree m in the classical sense.

$$\|f\|_{C^m(\Omega)} := \sum_{|\alpha|\le m} \|D^\alpha f\|_{L_\infty(\Omega)},$$

The ***semi-norm***

$$|f|_{C^m(\Omega)} := \sum_{|\alpha|=m} \|D^\alpha f\|_\infty$$

Examples $C^m([a,b])$ Then $\|f\|_{C^m([a,b])} = \sum_{k=0}^{m} \|f^{(k)}\|_\infty$ is a norm $|f|_{C^m(\mathbb{R})} = \|f^{(m)}\|_\infty$ is a semi-norm with the polynomials of degree $m-1$ as a null-space.

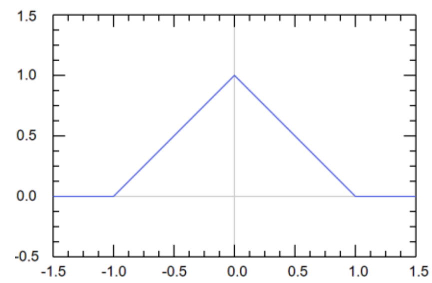
Definition Sobolev spaces $W_p^r(\Omega)$, $1 \le p \le \infty$

Def I For $1 \le p < \infty$, the closure of the compactly supported smooth functions $C_0^r(\Omega)$ with respect to the norm $\sum_{|\alpha|\le r} \|D^\alpha f\|_{L_p(\Omega)}$. For $p = \infty$, we take $W_\infty^r := C^r$.

Def II We define the space of ***test-functions*** $C_0^r(\Omega)$ - continuously differentiable with compact support in Ω. Let $f \in L_p(\Omega) \cap L_1(\Omega)$. Now for $\alpha \in \mathbb{Z}_+^d$, $|\alpha| = r$, $g := D^\alpha f$ is the ***distributional (generalized) derivative*** of f if for all $\phi \in C_0^r(\Omega)$

$$\int_\Omega g\phi = (-1)^{|\alpha|} \int_\Omega f D^\alpha \phi \ .$$

Example (assignment): For $H(x) := \begin{cases} x+1, & -1 \le x < 0, \\ 1-x, & 0 \le x \le 1, \\ 0, & else. \end{cases}$,



$$H'(x) = g(x) = \begin{cases} 1, & -1 \le x < 0, \\ -1, & 0 \le x \le 1, \\ 0, & else. \end{cases} .$$

So, in this sense $H \in W_p^1(\mathbb{R})$, $1 \le p < \infty$.

Now, we can define the Sobolev space by taking the norm below over the distributional derivatives.

The Sobolev norm and semi-norm. We require that the distributional derivatives exist as functions(!) in $L_p(\Omega)$ and

$$\|f\|_{W_p^r(\Omega)} := \sum_{|\alpha|\le r} \|D^\alpha f\|_{L_p(\Omega)} < \infty \qquad |f|_{W_p^r(\Omega)} := \sum_{|\alpha|=r} \|D^\alpha f\|_{L_p(\Omega)} .$$

Theorem W_p^r is a Banach space.

Theorem For a 'smooth' domain $\Omega \subseteq \mathbb{R}^n$, and any $0 < j < r$, there exist a constant(s) $C > 0$, such that for any $f \in W_p^r(\Omega)$,

$$|f|_{j,p} \le C\left(|f|_{r,p} + \|f\|_p\right),$$

$$\|f\|_{j,p} \le C\left(\|f\|_{r,p} + C\|f\|_p\right),$$

$$\|f\|_{j,p} \le C\|f\|_{r,p}^{j/r}\|f\|_p^{(r-j)/r}.$$

Remarks

(i) Sometimes one sees in textbooks a definition $\|f\|_{W_p^r(\Omega)} := \|f\|_{L_p(\Omega)} + |f|_{W_p^r(\Omega)}$, since by the theorem

$$\|f\|_{L_p(\Omega)} + |f|_{W_p^r(\Omega)} \le \sum_{|\alpha| \le r} \|D^\alpha f\|_{L_p(\Omega)} \le C\left(\|f\|_{L_p(\Omega)} + |f|_{W_p^r(\Omega)}\right).$$

(ii) The constants depend on the 'smoothness' of the boundary $\partial\Omega$.

Trigonometric polynomials and the Fourier Series

We now focus on the domain $\mathbb{T} = [-\pi, \pi]$ and 2π - periodic functions. They are extended $f(x + 2\pi k) = f(x)$, $k \in \mathbb{Z}$.

Periodic…what does it mean for us? Example, the function $f(x) = x$ is <u>not</u> continuous as a periodic function.

$L_2(\mathbb{T})$, is a Hilbert space equipped with the dot-product

$$\langle f, g \rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g(x)}dx.$$

The exponents are an **orthonormal basis**. For any $f \in L_2(\mathbb{T})$

$$f(x) \underset{L_2}{=} \sum_k \hat{f}(k)e^{ikx}, \quad \hat{f}(k) = \langle f, e^{ik\cdot} \rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-ikx}dx.$$

The partial Fourier sum

$$S_N(x) := \sum_{k=-N}^{N} \hat{f}(k)e^{ikx}.$$

Convergence in L_2 means that for any $f \in L_2(\mathbb{T})$,

$$\lim_{N\to\infty}\|f - S_N\|_{L_2(\mathbb{T})}^2 = \lim_{N\to\infty}\frac{1}{2\pi}\int_{-\pi}^{\pi}|f(x) - S_N(x)|^2 dx = 0.$$

Observe that we are not saying anything on the rate of convergence.

Observe **convergence is not pointwise**! There exists a continuous periodic function $f : \mathbb{T} \to \mathbb{R}$ such that

$$|S_N f(0)| = \left|\sum_{k=-N}^{N} \hat{f}(k)\right| \underset{N\to\infty}{\to} \infty.$$

There are even more exotic constructions! Conclusion: "Don't bring a knife to a gun fight" = Do not apply in L_∞ a Hilbert space/ L_2 tool.

Parseval identity

$$\|f\|_2^2 = \sum_k \left|\hat{f}(k)\right|^2 .$$

Also,

$$\sum_k \hat{f}(k)e^{ikx} = \hat{f}(0) + \sum_{k=1}^{\infty} \hat{f}(k)(\cos kx + i\sin kx) + \hat{f}(-k)(\cos kx - i\sin kx)$$
$$= \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos kx + b_k \sin kx$$

where

$$a_k = \hat{f}(k) + \hat{f}(-k) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(kx)\,dx$$
$$b_k = i\left(\hat{f}(k) - \hat{f}(-k)\right) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(kx)\,dx$$

Observe that if f is real, then the coefficients are also real.

BUT! In this course we are focused on approximation… not Fourier analysis. There is a big difference in the perspective!

From approximation theoretical perspective we are approximating a function from the space of trigonometric polynomials $\Pi_N(\mathbb{T}) := \left\{ \sum_{k=-N}^{N} a_k e^{ikx} \right\}$. The choice $a_k = \hat{f}(k)$ is the optimal(!) choice for $p = 2$. In some cases we need to choose different approximating trigonometric polynomials of degree N (e.g. approximation for $p \neq 2$).

Now let's try to say something about the rate of convergence. Here is a typical approximation theoretical result: the **Jackson-type estimate**.

Theorem There exists a constant $C(r) > 0$, such that for any $f \in W_p^r(\mathbb{T})$, $1 \le p \le \infty$,

$$E_N(f)_p := \min_{P \in \Pi_N} \|f - P\|_{L_p(\mathbb{T})} \le C(r) N^{-r} |f|_{r,p}, \qquad |f|_{r,p} = \left\| f^{(r)} \right\|_{L_p(\mathbb{T})}.$$

Right now, let us prove the Jackson-type estimate for the special case of $p = 2$

Theorem Let $f \in W_2^r(\mathbb{T})$ then

$$E_N(f)_2 \le N^{-r} |f|_{r,2}.$$

Proof

1. <u>Decay of the Fourier coefficients</u> - By Parseval, for any $g \in L_2(\mathbb{T})$

$$\|g\|_{L_2(\mathbb{T})}^2 = \sum_{k=-\infty}^{\infty} |\hat{g}(k)|^2 .$$

we have

$$E_N(f)_2 = \left\| f - S_N(f,x) \right\|_{L_2(\mathbb{T})} = \sqrt{\sum_{|k|\ge N+1} \left| \hat{f}(k) \right|^2}$$

Assume first that $f \in C^r(\mathbb{T})$. We will show $\left| \hat{f}(k) \right| = |k|^{-r} \left| \left(f^{(r)} \right)^{\wedge}(k) \right|$. Using the continuity of f as a periodic function, integration by parts yields,

$$\begin{aligned} \hat{f}(k) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-ikx} dx \\ &= \frac{1}{2\pi} \left(\underbrace{\left. \frac{f(x) e^{-ikx}}{-ik} \right|_{-\pi}^{\pi}}_{=0} + \frac{1}{ik} \int_{-\pi}^{\pi} f'(x) e^{-ikx} dx \right) \\ &= \frac{1}{ik} (f')^{\wedge}(k). \end{aligned}$$

By repeated application of the above

$$\left| \hat{f}(k) \right| = |k|^{-r} \left| \left(f^{(r)} \right)^{\wedge}(k) \right|.$$

2. <u>The estimate of the tail</u>

$$\begin{aligned} \left\| f(x) - S_N(f,x) \right\|_2^2 &= \sum_{|k|\ge N+1} \left| \hat{f}(k) \right|^2 \\ &\le N^{-2r} \sum_{|k|\ge N+1} |k|^{2r} \left| \hat{f}(k) \right|^2 \\ &= N^{-2r} \sum_{|k|\ge N+1} \left| \left(f^{(r)} \right)^{\wedge}(k) \right|^2 \\ &\le N^{-2r} \left\| f^{(r)} \right\|_2^2. \end{aligned}$$

$$\Rightarrow E_N(f)_2 = \left\| f(x) - S_N(f,x) \right\|_2 \le N^{-r} \left\| f^r \right\|_2.$$

For the general case $f \in W_2^r(\mathbb{T})$ we apply a **density** argument. Let $\{f_j\}_{j=1}^{\infty}$, $f_j \in C^r(\mathbb{T})$, such that

$$\left\| f - f_j \right\|_{W_2^r} \underset{j\to\infty}{\to} 0.$$

This implies

$$\left\| f - f_j \right\|_2 \underset{j\to\infty}{\to} 0, \quad \left\| f^{(r)} - f_j^{(r)} \right\|_2 \underset{j\to\infty}{\to} 0.$$

Therefore

$$\begin{aligned} \left\| f - S_N(f) \right\|_2 &\le \left\| f - f_j \right\|_2 + \left\| f_j - S_N(f_j) \right\|_2 + \left\| S_N(f_j) - S_N(f) \right\|_2 \\ &\le N^{-r} \left\| f_j^{(r)} \right\|_2 + \left\| f - f_j \right\|_2 + \left\| S_N(f - f_j) \right\|_2 \\ &\le N^{-r} \left\| f_j^{(r)} \right\|_2 + 2\left\| f - f_j \right\|_2 \underset{j\to\infty}{\to} N^{-r} \left\| f^{(r)} \right\|_2 \end{aligned}$$

□

Corollary [Approximation Spaces] Define $A_{\infty}^{\alpha}(L_2(\mathbb{T}))$ as the space of functions $f \in L_2(\mathbb{T})$ for which

$$|f|_{A_\infty^\alpha(L_2)} := \sup_{N\ge 1} N^\alpha E_N(f)_2 < \infty .$$

Then for any $r \ge \alpha$

$$W_2^r(\mathbb{T}) \subset A_\infty^\alpha(L_2(\mathbb{T})),$$

because for $f \in W_2^r(\mathbb{T})$ and $N \ge 1$

$$N^\alpha E_N(f)_2 \le N^r E_N(f)_2 \le |f|_{r,2} = \left\| f^{(r)} \right\|_2 .$$

Dirichlet kernel

$$D_N(x) := \sum_{k=-N}^{N} e^{ikx} = \frac{\sin((N+1/2)x)}{\sin(x/2)} .$$

$$\sum_{k=-N}^{N} e^{ikx} = e^{-iNx}\left(1+e^{ix}+\cdots+e^{i2Nx}\right) = e^{-iNx}\frac{e^{i(2N+1)x}-1}{e^{ix}-1} = \frac{e^{i(N+1)x}-e^{-iNx}}{e^{ix/2}\left(e^{ix/2}-e^{-ix/2}\right)}$$

$$= \frac{e^{i(N+1/2)x}-e^{-i(N+1/2)x}}{e^{ix/2}-e^{-ix/2}} = \frac{\sin((N+1/2)x)}{\sin(x/2)} .$$

Convolution over the torus

$$f*g(x) := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-y)g(y)dy .$$

Convolution with the Dirichlet kernel

$$\begin{aligned} D_N * f(x) &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\sum_{k=-N}^{N} e^{ik(x-y)}\right) f(y)dy \\ &= \sum_{k=-N}^{N}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-iky} f(y)dy\right)e^{ikx} \\ &= \sum_{k=-N}^{N} \hat{f}(k)e^{ikx} \\ &= S_N(f,x). \end{aligned}$$

This is a special case of

Assignment Let $f,g \in L_2(\mathbb{T})$. Prove that

(i) $f*g \in L_2(\mathbb{T})$.

(ii) For each $k \in \mathbb{Z}$, $(f*g)^\wedge(k) = \hat{f}(k)\hat{g}(k)$.

The problem with the Dirichlet kernel / Fourier series – optimal for $p=2$ …not adequate for $p=\infty$. Why? The short answer is $\{D_N\}$ are not the kernels of uniformly bounded operators on L_∞, $\|D_N\|_1 \ge C_1 + C_2 \log N$.

The origins of the Fourier Series (… which reveal how to generalize it)

The Heat equation over $\Omega \subset \mathbb{R}^n$, $t \ge 0$,

$$\begin{cases} \dfrac{\partial u}{\partial t} = \Delta u, \\ u(x,0) = f(x). \end{cases} \qquad \Delta u = \sum_{i=1}^{n} \frac{\partial^2 u}{\partial x_i^2}$$

For $\Omega = \mathbb{T}$, we have the **Laplace operator** $Lf := -\Delta f = -f''$.

$$\left\{e^{ikx}\right\} \Leftrightarrow k^2 \text{ , eigenvectors } \Leftrightarrow \text{ eigenvalues of } L.$$

$$Lf(x) = -\Delta f(x) = \sum_k k^2 \hat{f}(k) e^{ikx}, \qquad \forall f \in C^2(\mathbb{T}).$$

For $\varphi : \mathbb{R} \to \mathbb{R}$, even, define $\varphi(L) f(x) := \sum_k \varphi(k^2) \hat{f}(k) e^{ikx}$.

Spectral representation to solution of the heat equation with boundary condition f, is through **semi-group** $\varphi_t(u) := e^{-tu}$, $t > 0$. The solution is

$$u(x,t) = \varphi_t(L) f(x) = \sum_k e^{-tk^2} \hat{f}(k) e^{ikx}.$$

Example Let $\varphi(u) = 1_{[-1,1]}(u)$. Then,

$$\varphi\left(N^{-1}\sqrt{L}\right) f(x) = \sum_k \varphi\left(\frac{|k|}{N}\right) \hat{f}(k) e^{ikx} = S_N(f,x).$$

Fejér - The right convolution kernel for $p = \infty$

Def A ***summability kernel*** is a sequence $\{h_N\}$ satisfying:

(i) $\frac{1}{2\pi}\int_{-\pi}^{\pi} h_N(x)\,dx = 1$

(ii) $\frac{1}{2\pi}\int_{-\pi}^{\pi} |h_N(x)|\,dx \le C$.

(ii) For all $0 < \delta < \pi$, $\lim_{N\to\infty} \int_{|x|\ge\delta} |h_N(x)|\,dx = 0$

Remark For positive kernels we don't need (ii)

Theorem for a summability kernel $\{h_N\}$ and $f \in C(\mathbb{T})$,

$$\|f - h_N * f\|_{C(\mathbb{T})} = \max_{-\pi\le x\le\pi} |f(x) - h_N * f(x)| \underset{N\to\infty}{\to} 0.$$

Proof Assume $x = 0$. Let $\varepsilon > 0$. From the uniform continuity of f, there exists $0 < \delta < \pi$, such that $|x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$.

$$h_N * f(0) - f(0) = \frac{1}{2\pi}\int_{-\pi}^{\pi} h_N(t)\big(f(-t) - f(0)\big)dt$$

$$= \frac{1}{2\pi}\int_{-\delta}^{\delta} h_N(t)\big(f(-t) - f(0)\big)dt + \frac{1}{2\pi}\int_{|x|\ge\delta} h_N(t)\big(f(-t) - f(0)\big)dt$$

Now

$$\left|\frac{1}{2\pi}\int_{-\delta}^{\delta} h_N(t)\left(f(-t)-f(0)\right)dt\right| \le \max_{-\delta\le y\le\delta}\left|f(y)-f(0)\right|\frac{1}{2\pi}\int_{-\pi}^{\pi}\left|h_N(t)\right|dt$$

$$\le C\varepsilon.$$

Therefore

$$\left|h_N * f(0)-f(0)\right| \le C\varepsilon + 2\|f\|_\infty \frac{1}{2\pi}\int_{|x|\ge\delta}\left|h_N(t)\right|dt \underset{N\to\infty}{\to} C\varepsilon .$$

For $x\neq 0$, define $\tilde{f}(t)=f(t+x)$. Then

$$h_N * \tilde{f}(0)=\frac{1}{2\pi}\int \tilde{f}(0-y)h_N(y)dy=\frac{1}{2\pi}\int f(0-y+x)h_N(y)dy$$

$$=\frac{1}{2\pi}\int f(x-y)h_N(y)dy = h_N * f(x).$$

We now apply the first part of the proof for $\tilde{f}$ at 0, observing that $\|\tilde{f}\|_\infty = \|f\|_\infty$ and that for any $\varepsilon>0$, we can use the same $\delta>0$ we used for f . Hence, the approximation and convergence are in fact uniform for all $x\in\mathbb{T}$ ▫

Definition The Fejér kernel of degree $N-1$ is defined by averaging $D_0,...,D_{N-1}$

$$K_N(x) := \frac{1}{N}\sum_{n=0}^{N-1} D_n(x) = \frac{1}{N}\sum_{n=0}^{N-1}\sum_{k=-n}^{n} e^{ikx}$$

$$=\frac{1}{N}\left(N+(N-1)\left(e^{-ix}+e^{ix}\right)+\cdots\right)$$

$$=\sum_{k=-N}^{N}\left(1-\frac{|k|}{N}\right)e^{ikx}$$

$$K_N(x) := \sum_{k=-N}^{N}\left(1-\frac{|k|}{N}\right)e^{ikx} = \frac{1}{N}\left(\frac{\sin(Nx/2)}{\sin(x/2)}\right)^2 .$$

The partial Fejér series of f is

$$\sigma_N(f,x) := K_N * f(x) = \sum_{k=-N}^{N}\left(1-\frac{|k|}{N}\right)\hat{f}(k)e^{ikx} .$$

Let $\varphi(t)=1-|t|$, for $-1\le t\le 1$. Then,

$$\sigma_N(f,x)=\varphi\left(N^{-1}\sqrt{L}\right)f(x).$$

Theorem $\{K_N\}$ is a (positive) summability kernel

Proof

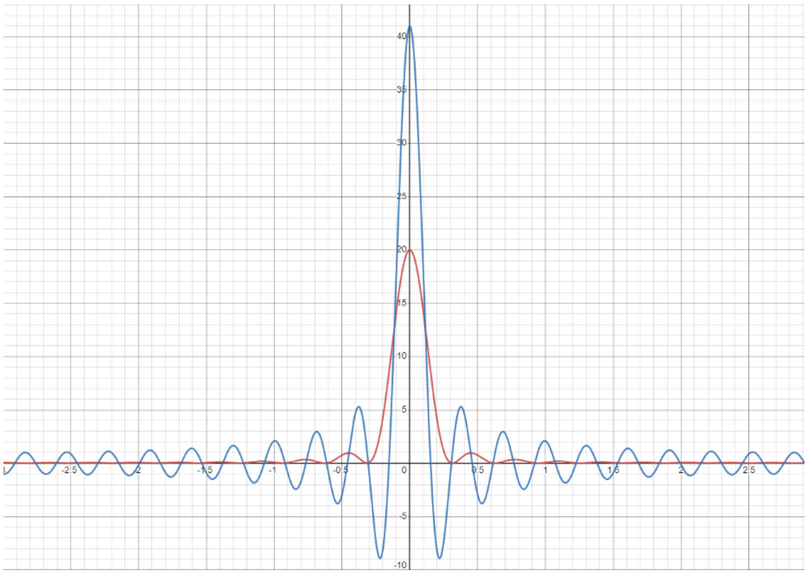
(i) and (ii) $\displaystyle \frac{1}{2\pi}\int_{-\pi}^{\pi} K_N(x)dx = \frac{1}{2\pi}\sum_{k=-N}^{N}\left(1-\frac{|k|}{N}\right)\int_{-\pi}^{\pi} e^{ikx}dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} dx = 1.$

(iii) Let $0<\delta<\pi$. Then $\displaystyle \int_{|x|\ge\delta}\left|K_N(x)\right|dx = \frac{1}{N}\int_{|x|\ge\delta}\left(\frac{\sin(Nx/2)}{\sin(x/2)}\right)^2 dx \le \frac{1}{N}\frac{2\pi}{\sin^2(\delta/2)} \underset{N\to\infty}{\to} 0$

Conclusions

(i) $\|f - K_N * f\|_{C(\mathbb{T})} \underset{N\to\infty}{\to} 0$.

(ii) The trigonometric polynomials are dense in $C(\mathbb{T})$.



Dirichlet (blue) & Fejér (red) kernels for $N = 20$

Fourier Integral

Def $f \in L_1(\mathbb{R}^n)$, then

$$\hat{f}(w) := \int_{\mathbb{R}^n} f(x) e^{-iw\cdot x} dx = \int_{\mathbb{R}^n} f(x) e^{-i\langle w,x\rangle} dx.$$

A rigorous method to define the Fourier integral is to first define it for Schwartz functions…
Definition $\phi \in \mathcal{S}$ is in the Schwartz class if it is in $C^\infty(\mathbb{R}^n)$ and for any $\alpha \in \mathbb{Z}_+^n$, $|\alpha| = m$ and $k \geq 0$, there exists $C_{\alpha,k}$ such that

$$\sup_{x\in\mathbb{R}^n} |\partial^\alpha \phi(x)|(1+|x|)^k \leq C_{\alpha,k}.$$

Example Gaussians are the pro-typical example for Schwartz functions, e.g. $\phi(x) = e^{-|x|^2}$.

We'll see examples where the "Schwartz is with us" soon…and then we will need the following
Assignment Show that for any $\phi \in \mathcal{S}$

(i) $\partial^\alpha \phi \in L_p(\mathbb{R}^n)$, for any $0 < p \leq \infty$, $\alpha \in \mathbb{Z}_+^n$.

(ii) $\hat{\phi} \in \mathcal{S}$ (Hint: you may prove the case $n = 1$. Use integration by parts of $\int_{-\infty}^{\infty} f(x) e^{-iwx} dx$).

Properties of the Fourier integral:

i. $\forall f \in L_1, \ \left\| \hat{f} \right\|_\infty \le \| f \|_1$.

ii. For $f \in L_1(\mathbb{R})$, $\hat{f}$ is uniformly continuous.

Proof For $\varepsilon > 0$, let $M > 0$ such that

$$\int_{\mathbb{R}^n \setminus [-M,M]^n} |f| < \frac{\varepsilon}{4}.$$

Let $\delta \in \mathbb{R}^n$, be sufficiently small, such that

$$\left| e^{-i\delta \cdot x} - 1 \right| < \frac{\varepsilon}{2\| f \|_1}, \quad \forall x \in [-M, M]^n$$

Then, for any $w \in \mathbb{R}^n$

$$\begin{aligned} \left| \hat{f}(w+\delta) - \hat{f}(w) \right| &= \left| \int_{\mathbb{R}^n} e^{-iwx} \left(e^{-i\delta x} - 1 \right) f(x) dx \right| \\ &\le \int_{\mathbb{R}^n} \left| e^{-i\delta x} - 1 \right| \left| f(x) \right| dx \\ &= \int_{[-M,M]^n} \left| e^{-i\delta x} - 1 \right| \left| f(x) \right| dx + \int_{\mathbb{R}^n \setminus [-M,M]^n} \left| e^{-i\delta x} - 1 \right| \left| f(x) \right| dx \\ &\le \max_{x \in [-M,M]^n} \left| e^{-i\delta x} - 1 \right| \| f \|_1 + 2\frac{\varepsilon}{4} < \varepsilon. \end{aligned}$$

□

iii. $f(\bullet - \alpha)^{\wedge}(w) = \int_{\mathbb{R}^n} f(x-\alpha) e^{-iwx} dx = \int_{\mathbb{R}^n} f(y) e^{-iw(y+\alpha)} dy = e^{-iw\alpha} \hat{f}(w)$

iv. For a Schwartz function the following calculations can be applied because $f, f' \in L_1$ and f decays to zero at $\pm\infty$.

$$\begin{aligned} (f')^{\wedge}(w) &= \int_{-\infty}^{\infty} f'(x) e^{-iwx} dx \\ &= -f(x) e^{-iwx} \Big|_{-\infty}^{\infty} + iw \int_{-\infty}^{\infty} f(x) e^{-iwx} dx \\ &= iw \int_{-\infty}^{\infty} f(x) e^{-iwx} dx = iw\hat{f}(w) \end{aligned}$$

Examples:

i. $f(x) = 1_{[0,1]}(x)$. Then, $\hat{f}(w) = \int_0^1 e^{-iwx} dx = \left. \frac{e^{-iwx}}{-iw} \right|_0^1 = \frac{e^{-iw} - 1}{-iw} = \frac{1 - e^{-iw}}{iw}$.

ii. $f(x) = \operatorname{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$. Notice that $f \notin L_1$, $f \in L_2$. We will see later $\hat{f}(w) = 1_{[-\pi,\pi]}(w)$.

iii. Gaussians $f(x) = e^{-\alpha x^2}$, $\hat{f}(w) = \sqrt{\frac{\pi}{\alpha}} e^{-\frac{w^2}{4\alpha}}$.

Proof Consider the function $g(y)=\int_{-\infty}^{\infty} e^{-\alpha x^2+xy}dx$, $y\in\mathbb{R}$. By completing the squares

$$g(y)=\int_{-\infty}^{\infty} e^{-\alpha\left(x-\frac{y}{2\alpha}\right)^2+\frac{y^2}{4\alpha}}dx=\frac{1}{\sqrt{\alpha}}e^{\frac{y^2}{4\alpha}}\int_{-\infty}^{\infty}e^{-z^2}dz$$
$$=\sqrt{\frac{\pi}{\alpha}}e^{\frac{y^2}{4\alpha}}$$

Can be extended to entire functions that agree on $\mathbb{R}$, we can set $y=-iw$ and then

$$\hat{f}(w)=g(-iw)=\int_{-\infty}^{\infty} e^{-\alpha x^2}e^{-iwx}dx=\sqrt{\frac{\pi}{\alpha}}e^{\frac{-w^2}{4\alpha}}.$$

□

Convolution on $\mathbb{R}^n$

$$f*g(x)=\int_{\mathbb{R}^n} f(x-y)g(y)dy.$$

Simple exercise $\|f*g\|_1\le\|f\|_1\|g\|_1$.

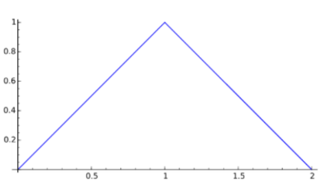
Theorem For $f,g\in L_1(\mathbb{R}^n)$, $(f*g)^{\wedge}(w)=\hat{f}(w)\hat{g}(w)$, $w\in\mathbb{R}^n$.

Proof

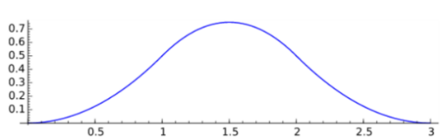
$$(f*g)^{\wedge}(w)=\int_{\mathbb{R}^n} f*g(x)e^{-iwx}dx=\int_{\mathbb{R}^n} e^{-iwx}dx\int_{\mathbb{R}^n} f(x-y)g(y)dy$$
$$=\int_{\mathbb{R}^n} g(y)dt\int_{\mathbb{R}^n} f(x-y)e^{-iwx}dx=\int_{\mathbb{R}^n} g(y)e^{-iwy}\hat{f}(w)dy$$
$$=\hat{f}(w)\hat{g}(w).$$

Examples: B-splines $N_1(x)=1_{[0,1]}(x)$

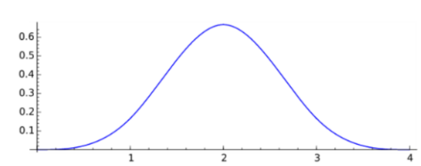
$$N_2(x)=N_1*N_1(x)=\int_0^1 1_{[0,1]}(x-t)dt=\int_{\max(x-1,0)}^{\min(x,1)}dt=\begin{cases}0 & x<0\\ x & 0\le x<1\\ 2-x & 1\le x\le 2\\ 0 & x>2\end{cases}$$



N_2



N_3



N_4

We define $N_r:=N_{r-1}*N_1$. Therefore, $(N_r)^{\wedge}(w)=\left(\frac{1-e^{-iw}}{iw}\right)^r$.

Inverse Fourier For $g \in L_1$ define

$$\mathcal{F}^{-1}g(x) := \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} g(w)e^{iwx}dw$$

Theorem For $\phi \in \mathcal{S}$

$$\phi(x) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\hat{\phi}(w)e^{iwx}dw, \qquad \forall x \in \mathbb{R}^n.$$

Theorem

A. For $f, h \in \mathcal{S}$ $\langle f, h\rangle = (2\pi)^{-n}\langle \hat{f}, \hat{h}\rangle$.

B. For $f \in \mathcal{S}$, we have $\|\hat{f}\|_2^2 = (2\pi)^n \|f\|_2^2$.

Proof (sketch for $n=1$) Let $g_\alpha(x) = \frac{1}{2\sqrt{\pi\alpha}}e^{-\frac{x^2}{4\alpha}}$. By the previous theorem $\hat{g}_\alpha(w) = e^{-\alpha w^2}$. We now compute

$$\begin{aligned}\int \hat{g}_\alpha(w)\,\hat{f}(w)\overline{\hat{h}(w)}dw &= \int \hat{g}_\alpha(w)\int f(x)e^{-iwx}dx\int \overline{h(y)}e^{iwy}dydw\\ &= \int f(x)\int \overline{h(y)}\left(\int \hat{g}_\alpha(w)e^{iw(y-x)}dw\right)dydx\\ &= 2\pi\int f(x)\left(\int \overline{h(y)}g_\alpha(y-x)\,dy\right)dx\\ &= 2\pi\int f(x)\left(\int \overline{h(y)}g_\alpha(x-y)\,dy\right)dx\end{aligned}$$

When we take limit $\alpha \to 0^+$

$$\begin{array}{ccc}\int \hat{g}_\alpha(w)\,\hat{f}(w)\overline{\hat{h}(w)}dw & = & 2\pi\int f(x)\left(\int \overline{h(y)}g_\alpha(x-y)\,dy\right)dx\\ \downarrow & & \downarrow\\ \int \hat{f}(w)\overline{\hat{h}(w)}dw & & = 2\pi\int f(x)\overline{h(x)}dx\end{array}$$

□

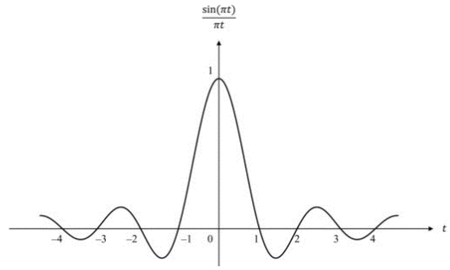
Definition Let $f \in L_2(\mathbb{R}^n)$. We define $\hat{f} := \lim_{k\to\infty}\hat{\varphi}_k$, where for all $1 \le k \le \infty$, $\varphi_k \in \mathcal{S}$ and $\varphi_k \underset{k\to\infty}{\to} f$ in L_2

Why is this well-defined? If $\{\varphi_k\}$ is a Cauchy sequence in L_2, then by the previous theorem, so is $\{\hat{\varphi}_k\}$. Since L_2 is complete there exists a limit which we define as $\hat{f}$.

Corollary We can extend the Fourier transform and its inverse to L_2 with $\mathcal{F}\mathcal{F}^{-1}f = f$.

Example The sinc function. $\hat{f}(w) = 1_{[-\pi,\pi]}(w)\ldots$ what is f?

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\hat{f}(w)e^{iwx}dw=\frac{1}{2\pi}\int_{-\pi}^{\pi}e^{iwx}dw=\frac{1}{2\pi}\frac{e^{iwx}}{ix}\bigg|_{-\pi}^{\pi}=\frac{1}{2\pi}\frac{e^{i\pi x}-e^{-i\pi x}}{ix}=\frac{\sin\pi x}{\pi x}.$$



The Laplace operator, the Heat equation and Fourier transform

$\Omega=\mathbb{R}^n$, Laplace operator

$$L=-\Delta:=-\sum_{k=1}^{n}\frac{\partial^2}{\partial x_k^2}.$$

On $\mathbb{R}$ we have that $L\left(e^{iwx}\right)=w^2e^{iwx}$, $\forall w\in\mathbb{R}$. The spectrum of the operator is the whole real line

$$Lf(x)=-\Delta f(x)=-\frac{1}{2\pi}\int_{-\infty}^{\infty}\left(f''\right)^{\wedge}e^{iwx}dw=\frac{1}{2\pi}\int_{-\infty}^{\infty}w^2\hat{f}(w)e^{iwx}dw, \qquad \forall f\in W_2^2(\mathbb{R}).$$

$$Lf(x)=-\Delta f(x)=\sum_k k^2\hat{f}(k)e^{ikx}, \qquad \forall f\in W_2^2(\mathbb{T}).$$

The Heat equation

$$\begin{cases}\dfrac{\partial u}{\partial t}=\Delta u,\\ u(x,0)=f(x).\end{cases}$$

The Gaussian (heat) Kernels satisfy the Heat equation

$$p_t(x):=\frac{1}{(4\pi t)^{n/2}}e^{-|x|^2/4t}, \qquad \int_{\mathbb{R}^n}p_t(x)dx=1, \qquad t>0.$$

Semi-group $p_t*p_s=p_{t+s}$, $t,s>0$.

Theorem If f is continuous and bounded then

$$u(x,t)=p_t*f(x),$$

solves the Heat equation with initial conditions f.

Sketch Easy to see

$$\left(\frac{\partial}{\partial t}-\Delta\right)u(x)=\int_{\mathbb{R}^n}\left(\frac{\partial}{\partial t}-\Delta\right)p_t(x-y)f(y)dy=0.$$

$$u(x,t)=p_t*f(x)\underset{t\to 0}{\longrightarrow}f(x).$$

Maximal function

$$M_\Phi f(x) := \sup_{t>0} |u(x,t)| = \sup_{t>0} |p_t * f(x)|.$$

Typical question If we know that $f \in L_p(\mathbb{R}^n)$, $0 < p \le \infty$, what can we say about the solution? In other words, can we bound $\|M_\Phi f\|_p$? This topic is discussed in the "Function Space Theory" course.

Spectral representation to solution of the Heat equation with boundary condition f

On $\mathbb{R}$ $$u(x,t) = e^{-tL} f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-tw^2} \hat{f}(w) e^{iwx} dw,$$

On $\mathbb{T}$ $$u(x,t) = e^{-tL} f(x) = \sum_k e^{-tk^2} \hat{f}(k) e^{ikx}.$$

We shall later in the course encounter approximation from shift-invariant spaces of the sinc function.

$$\operatorname{sinc}(x) = \frac{\sin(\pi x)}{\pi x}, \quad \operatorname{sinc}^\wedge(w) = 1_{[-\pi,\pi]}(w).$$

The approximation we shall use is equivalent to the following: Let $\varphi(u) = 1_{[-\pi,\pi]}(u)$. For any $h > 0$, we apply

$$\begin{aligned}\varphi\left(h\sqrt{L}\right) f(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi(h|w|) \hat{f}(w) e^{iwx} dw \\ &= \frac{1}{2\pi} \int_{-h^{-1}\pi}^{h^{-1}\pi} \hat{f}(w) e^{iwx} dw.\end{aligned}$$

Approximation using uniform piecewise constants (numerical integration)

The B-Spline of order one (degree zero, smoothness -1) $N_1(x) = \mathbf{1}_{[0,1]}(x)$.

Let $\Omega = \mathbb{R}$ or $\Omega = [a,b]$. We approximate from the space

$$S(N_1)^h := \left\{ \sum_{k\in\mathbb{Z}} c_k N_1\left(h^{-1}x - k\right) \right\} = \left\{ \sum_{k\in\mathbb{Z}} c_k \mathbf{1}_{[kh,(k+1)h]}(x) \right\}.$$

Theorem For $f \in W_p^1(\mathbb{R})$, $1 \le p \le \infty$,

$$E\left(f, S(N_1)^h\right)_{L_p(\mathbb{R})} := \inf_{g\in S(N_1)^h} \|f - g\|_{L_p(\mathbb{R})} \le h |f|_{W_p^1(\mathbb{R})}.$$

Proof First assume $f \in C^1(\mathbb{R}) \cap W_p^1(\mathbb{R})$. Let's take the interval $[kh, (k+1)h]$. Then, for $p = \infty$

$$|f(x) - f(kh)| = \left| \int_{kh}^{x} f'(u) du \right| \le h \sup_{kh \le u \le (k+1)h} |f'(u)|.$$

So select $c_k := f(kh)$ and you get the theorem for $p = \infty$ by using

$$g(x) = \sum_{k\in\mathbb{Z}} f(kh) N_1\left(h^{-1}x - k\right).$$

For $1 \le p < \infty$ we do something similar

$$\left| f(x) - f(kh) \right|^p \le \left(\int_{kh}^{(k+1)h} \left| f'(u) \right| du \right)^p, \qquad x \in \left[kh, (k+1)h \right].$$

Then

$$\begin{aligned} \int_{kh}^{(k+1)h} \left| f(x) - f(kh) \right|^p dx &\le h \left(\int_{kh}^{(k+1)h} \left| f'(u) \right| du \right)^p \\ &\le h \left(\| f' \|_{L_p([kh,(k+1)h])} \| 1 \|_{L_{p'}([kh,(k+1)h])} \right)^p \\ &= h h^{p/p'} \| f' \|^p_{L_p([kh,(k+1)h])} \\ &= h^p \| f' \|^p_{L_p([kh,(k+1)h])}. \end{aligned}$$

$$1 + \frac{p}{p'} = 1 + p \left(1 - \frac{1}{p} \right) = 1 + p - 1 = p$$

Therefore, with $g(x) := \sum_k f(kh) N_1 \left(h^{-1} x - k \right)$, we get

$$\| f - g \|_p^p = \int_{-\infty}^{\infty} \left| f(x) - g(x) \right|^p dx = \sum_k \int_{kh}^{(k+1)h} \left| f(x) - f(kh) \right|^p dx \le \sum_k h^p \| f' \|^p_{L_p([kh,(k+1)h])} = h^p \| f' \|_p^p.$$

Now assume $f \in W_p^1(\mathbb{R})$, $1 \le p < \infty$. There exist sequences $\{ f_k \}$, $f_k \in C^1(\mathbb{R}) \cap W_p^1(\mathbb{R})$, $\{ g_k \}$, $g_k \in S(N_1)^h$, such that $\| f - f_k \|_{W_p^r(\mathbb{R})} \underset{k \to \infty}{\to} 0$ and $\| f_k - g_k \|_{L_p(\mathbb{R})} \le h | f_k |_{W_p^1(\mathbb{R})}$. This gives

$$\begin{aligned} \| f - g_k \|_p &\le \| f - f_k \|_p + \| f_k - g_k \|_p \\ &\le \| f - f_k \|_p + h | f_k |_{1,p} \underset{k \to \infty}{\to} h | f |_{1,p} \end{aligned}$$

□

Modulus of smoothness

Def The ***difference operator*** Δ_h^r. For $h \in \mathbb{R}^n$ we define $\Delta_h(f,x) = f(x+h) - f(x)$. For general $r \ge 1$ we define

$$\Delta_h^r(f,x) = \underbrace{\Delta_h \circ \cdots \Delta_h}_{r}(f,x) = \sum_{k=0}^{r} \binom{r}{k} (-1)^{r-k} f(x+kh).$$

Remarks

1. For $\Omega \subset \mathbb{R}^n$, we in fact modify to $\Delta_h^r(f,x) := \Delta_h^r(f,x,\Omega)$, where $\Delta_h^r(f,x) = 0$, in the case $[x, x+rh] \not\subset \Omega$. So for $\Omega = [a,b]$, $\Delta_h^r(f,x) = 0$ on $[b - rh, b]$, for any function.
2. As an operator on $L_p(\Omega)$, $1 \le p \le \infty$, we have that $\left\| \Delta_h^r \right\|_{L_p \to L_p} \le 2^r$. Assume $\Omega = \mathbb{R}^n$, then

$$\left\| \Delta_h^r(f, \bullet) \right\|_p \le \sum_{k=0}^{r} \binom{r}{k} \| f(\bullet + kh) \|_p = \sum_{k=0}^{r} \binom{r}{k} \| f \|_p = 2^r \| f \|_p$$

Def The ***modulus of smoothness*** of order r of a function $f \in L_p(\Omega)$, $0 < p \le \infty$, at the parameter $t > 0$

$$\omega_r(f,t)_p := \sup_{|h|\le t} \left\| \Delta_h^r(f,x) \right\|_{L_p(\Omega)}.$$

For $r=1$ the modulus of smoothness is called the ***modulus of continuity***.

Example non continuous function. Let $\Omega = [-1,1]$. $f(x) = \begin{cases} 0 & x<0 \\ 1 & 0 \le x \end{cases}$.

Let's compute $\omega_r(f,t)_{L_p([-1,1])}$.

$$\Delta_h(f,x) = \begin{cases} 0 & -1 \le x \le -h \\ 1 & -h < x \le 0 \\ 0 & 0 < x \le 1 \end{cases}$$

For $p=\infty$ we get $\omega_1(f,t)_{L_\infty([-1,1])} = \sup_{|h|\le t} \left\| \Delta_h f \right\|_{L_\infty([-1,1])} = 1$.

For $p \ne \infty$ we get $\omega_1(f,t)_{L_p([-1,1])} = \sup_{|h|\le t} \left\| \Delta_h f \right\|_{L_p([-1,1])} = t^{1/p}$.

$$\Delta_h^2(f,x) = \Delta_h(\Delta_h f,x) = \begin{cases} 0 & -1 \le x \le -2h \\ 1 & -2h < x \le -h \\ -1 & -h < x \le 0 \\ 0 & 0 \le x \le 1 \end{cases}$$

We get $\omega_2(f,t)_{L_p([-1,1])} = (2t)^{1/p}$

In general, we'll get $\omega_r(f,t)_{L_p([-1,1])} \le C(r,p)t^{1/p}$

Quick jump into the "future" (Generalized Lipschitz / Besov smoothness)… for $\alpha < 1/\tau$, $r = \lfloor \alpha \rfloor + 1$,

$$|f|_{B^\alpha_{\tau,\infty}} := \sup_{t>0} t^{-\alpha}\omega_r(f,t)_\tau \le \sup_{0<t\le 2} t^{-\alpha}\omega_r(f,t)_\tau \le c \sup_{0<t\le 2} t^{1/\tau-\alpha} < \infty.$$

We then say that f has α (weak-type) smoothness. Observe that in this example α can be arbitrarily large as long as the integration takes place with τ sufficiently small.

<u>Properties</u>

1. For $1 \le p \le \infty$, $\omega_r(f,t)_p \le 2^r \|f\|_{L_p(\Omega)}$.
2. $\omega_r(f,t)_p$ is non-decreasing in t
3. For $1 \le p \le \infty$ the ***sub-linearity*** property

$$\left|\Delta_h^r(f+g,x)\right| \le \left|\Delta_h^r(f,x)\right| + \left|\Delta_h^r(g,x)\right|,$$

gives

$$\omega_r(f+g,t)_p \le \omega_r(f,t)_p + \omega_r(g,t)_p.$$

4. For $N \ge 1$, $\omega_r(f, Nt)_p \le N^r \omega_r(f,t)_p$, $1 \le p \le \infty$.

Proof For the proof, we need this property (**assignment**)

$$\Delta_{Nh}^r(f,x) = \sum_{k_1=0}^{N-1} \cdots \sum_{k_r=0}^{N-1} \Delta_h^r(f, x + k_1 h + \cdots + k_r h), \qquad \forall x \in \Omega, \text{ s.t. } [x, x + Nhr] \subset \Omega. \tag{*}$$

Let's see (*) for the case $r = 1$. We assume that $[x, x+Nh] \subset \Omega$, otherwise $\Delta_{Nh}(f,x) = 0$.

$$\begin{aligned}
\Delta_{Nh}(f,x) &= f(x+Nh) - f(x) \\
&= f(x+Nh) - f(x+(N-1)h) + f(x+(N-1)h) - \cdots + f(x+h) - f(x) \\
&= \sum_{k=0}^{N-1} \Delta_h(f, x+kh)
\end{aligned}$$

For $k := (k_1, \ldots, k_r)$, let

$$\Omega(k) := \{x \in \Omega : [x, x + k_1 h + \cdots + k_r h + rh] \subset \Omega\}.$$

Then,

$$\begin{aligned}
\left\|\Delta_h^r(f, \cdot + k_1 h + \cdots + k_r h)\right\|_{L_p(\Omega)}^p &= \int_{\Omega(k)} \left| \sum_{j=0}^r \binom{r}{j} (-1)^{r-j} f(x + k_1 h + \cdots + k_r h + jh) \right|^p dx \\
&= \int_{\Omega(k)} \left| \sum_{j=0}^r \binom{r}{j} (-1)^{r-j} f(y + jh) \right|^p dy \\
&\le \int_\Omega \left|\Delta_h^r(f,y)\right|^p dy = \left\|\Delta_h^r(f)\right\|_{L_p(\Omega)}^p
\end{aligned}$$

Then, assuming (*) for any $h \in \mathbb{R}^n$, $|h| \le t$

$$\begin{aligned}
\left\|\Delta_{Nh}^r(f,\cdot)\right\|_p &\le \sum_{k_1=0}^{N-1} \cdots \sum_{k_r=0}^{N-1} \left\|\Delta_h^r(f, \cdot + k_1 h + \cdots + k_r h)\right\|_p \\
&\le \sum_{k_1=0}^{N-1} \cdots \sum_{k_r=0}^{N-1} \left\|\Delta_h^r(f,\cdot)\right\|_p \le N^r \omega_r(f,t)_p.
\end{aligned}$$

Taking supremum over all $h \in \mathbb{R}^n$, $|h| \le t$, gives $\omega_r(f, Nt)_p \le N^r \omega_r(f,t)_p$.

□

It is easy to see that for $0 < p < 1$, the same proof yields $\omega_r(f, Nt)_p \le N^{r/p} \omega_r(f,t)_p$.

5. From (4) we get for $1 \le p \le \infty$,

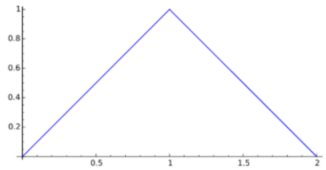
$$\omega_r(f, \lambda t)_p \le (\lambda + 1)^r \omega_r(f,t)_p, \qquad \lambda > 0$$

proof $\omega_r(f, \lambda t)_p \le \omega_r(f, \lfloor \lambda + 1 \rfloor t)_p \le (\lfloor \lambda + 1 \rfloor)^r \omega_r(f,t)_p \le (\lambda+1)^r \omega_r(f,t)_p$.

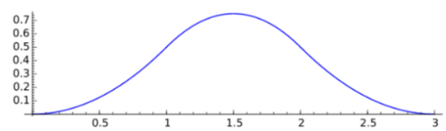
Theorem [connection between Sobolev and modulus] For $g \in W_p^r(\Omega)$, $1 \le p \le \infty$, we have that

$$\omega_r(g,t)_{L_p(\Omega)} \le C(r,n)t^r |g|_{W_p^r(\Omega)}, \qquad \forall t>0.$$

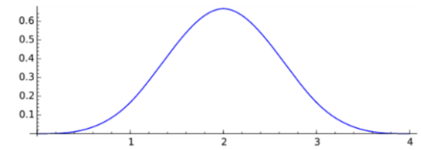
Proof for $\Omega=\mathbb{R}$. Recall the B-Splines, $N_1 = \mathbf{1}_{[0,1]^n}$. In general, $N_r := N_{r-1} * N_1 = \int_{\mathbb{R}^n} N_{r-1}(x-t)N_1(t)dt$.



N_2



N_3



N_4

- Properties:
  - Order r
  - Support $[0,r]^n$
  - Piecewise polynomial of degree $r-1$ with breakpoints (knots) at the integers
  - Smoothness $r-2$, thus in Sobolev W_p^{r-1}.
  - $\int_{\mathbb{R}^n} N_r(x)dx = 1$
  - Tensor-product in multivariate case

Let's see how we get the property of $\int_{\mathbb{R}^n} N_r(x)dx=1$. Let $f,g\in L_1(\mathbb{R}^n)$ then

$$\int_{\mathbb{R}^n} f*g(x)dx = \int_{\mathbb{R}^n} f \int_{\mathbb{R}^n} g$$

Here, we use the fact that for $h>0$, $\Delta_{-h}^r(f,x) = \Delta_h^r(f,x-rh)$. So W.L.G, for any $t>0$, we can work with $0<h\le t$. Define $N_r(x,h) := h^{-1}N_r(h^{-1}x)$, $h>0$. Let $g\in C^1(\mathbb{R})$. Then

$$\begin{aligned} h^{-1}\Delta_h(g,x) &= h^{-1}\left(g(x+h)-g(x)\right) \\ &= h^{-1}\int_x^{x+h} g'(u)du \\ &= \int_{\mathbb{R}} g'(x+u)N_1(u,h)du \end{aligned}$$

More generally, $g\in C^r(\mathbb{R})$

$$h^{-r}\Delta_h^r(g,x) = \int_{\mathbb{R}} g^{(r)}(x+u)N_r(u,h)du$$

To see that we apply induction

$$
\begin{aligned}
h^{-r}\Delta_h^r(g,x) &= h^{-1}h^{-(r-1)}\left(\Delta_h^{r-1}(g,x+h)-\Delta_h^{r-1}(g,x)\right)\\
&= h^{-1}\left(\int_{\mathbb{R}} g^{(r-1)}(x+h+u)N_{r-1}(u,h)\,du-\int_{\mathbb{R}} g^{(r-1)}(x+u)N_{r-1}(u,h)\,du\right)\\
&= h^{-1}\int_x^{x+h}\int_{-\infty}^{\infty} g^{(r)}(v+u)N_{r-1}(u,h)\,du\,dv\\
&= \int_{-\infty}^{\infty} N_{r-1}(u,h)\left(h^{-1}\int_x^{x+h} g^{(r)}(v+u)\,dv\right)du\\
&= \int_{-\infty}^{\infty} N_{r-1}(u,h)\left(\int_{-\infty}^{\infty} g^{(r)}(v+u)N_1(v-x,h)\,dv\right)du\\
&\overset{v+u=x+y}{=} \int_{-\infty}^{\infty} N_{r-1}(u,h)\left(\int_{-\infty}^{\infty} g^{(r)}(x+y)N_1(y-u,h)\,dy\right)du\\
&= \int_{-\infty}^{\infty} g^{(r)}(x+y)\left(\int_{-\infty}^{\infty} N_{r-1}(u,h)N_1(y-u,h)\,du\right)dy\\
&= \int_{-\infty}^{\infty} g^{(r)}(x+y)\mathrm{N}_r(y,h)\,\mathrm{dy}
\end{aligned}
$$

Now, let’s see the proof for $p=1$. Let $0<h\le t$

$$
\begin{aligned}
\int_{\mathbb{R}}\left|\Delta_h^r(g,x)\right|dx &\le h^r\int_{\mathbb{R}}\int_{\mathbb{R}}\left|g^{(r)}(x+u)\right|\left|N_r(u,h)\right|du\,dx\\
&\le t^r\underbrace{\int_{\mathbb{R}}\left|N_r(u,h)\right|du}_{=1}\int_{\mathbb{R}}\left|g^{(r)}(x)\right|dx\\
&\le t^r|g|_{W_1^r(\mathbb{R})}.
\end{aligned}
$$

For general $1\le p<\infty$ we need Minkowski’s inequality. It says that for measurable non-negative functions φ,ρ

$$
\left\{\int_A\left(\int_B\varphi(y)\rho(x,y)\,dy\right)^p dx\right\}^{1/p}\le\int_B\varphi(y)\left(\int_A\rho(x,y)^p\,dx\right)^{1/p}dy
$$

Or written differently

$$
\left\|\int_B\varphi(y)\rho(\cdot,y)\,dy\right\|_{L_p(A)}\le\int_B\varphi(y)\left\|\rho(\cdot,y)\right\|_{L_p(A)}dy
$$

Using it we have

$$\int_{\mathbb{R}} \left|\Delta_h^r(g,x)\right|^p dx \le h^{pr} \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \left|g^{(r)}(x+u)\right| \left|N_r(u,h)\right| du \right)^p dx$$
$$\le h^{pr} \left(\int_{\mathbb{R}} \left|N_r(u,h)\right| \left\|g^{(r)}(\cdot+u)\right\|_{L_p(\mathbb{R})} du \right)^p$$
$$\le h^{pr} \left(\int_{\mathbb{R}} \left|N_r(u,h)\right| \left\|g^{(r)}\right\|_{L_p(\mathbb{R})} du \right)^p$$
$$\le t^{pr} \left\|g^{(r)}\right\|_{L_p(\mathbb{R})}^p$$
$$= t^{pr} \left|g\right|_{W_p^r(\mathbb{R})}^p .$$

For a general function $g \in W_p^r(\mathbb{R})$ we use the density of $C^r(\mathbb{R}) \cap W_p^r(\mathbb{R})$ in $W_p^r(\mathbb{R})$.

□

Corollary For any $P \in \Pi_{r-1}$, $P(x) = \sum_{k=0}^{r-1} a_k x^k$,

$$h^{-r} \Delta_h^r(P,x) = \int_{\mathbb{R}} P^{(r)}(x+u) N_r(u,h) du = 0 \Rightarrow \Delta_h^r(P,x) = 0 \Rightarrow \omega_r(P,t)_p = 0$$

Marchaud inequalities

We know that for any $1 \le k < r$, $1 \le p \le \infty$,

$$\omega_r(f,t)_p = \sup_{|h| \le t} \left\|\Delta_h^r(f)\right\|_p = \sup_{|h| \le t} \left\|\Delta_h^{r-k} \Delta_h^k(f)\right\|_p \le 2^{r-k} \sup_{|h| \le t} \left\|\Delta_h^k(f)\right\|_p = 2^{r-k} \omega_k(f,t)_p .$$

The direct inverse cannot be true. If we take $\Omega = [a,b]$ and a polynomial $P \in \Pi_{r-1}$, then $\omega_r(P,t)_p = 0$, but we don't necessarily have $\omega_k(P,t)_p = 0$ for $0 \le k < r$.

Theorem. For any $1 \le k < r$, $1 \le p \le \infty$,

On $\Omega = \mathbb{R}^n$, $\qquad \omega_k(f,t)_p \le ct^k \int_t^{\infty} \frac{\omega_r(f,s)_p}{s^{k+1}} ds , \qquad \forall t > 0 .$

On $\Omega = [a,b]$, $\qquad \omega_k(f,t)_p \le ct^k \left(\int_t^{b-a} \frac{\omega_r(f,s)_p}{s^{k+1}} ds + \frac{\|f\|_p}{(b-a)^k} \right), \quad 0 < t \le \frac{b-a}{r} .$

Proof of the case $\Omega = \mathbb{R}^n$. We prove first for $r = k+1$ and then apply induction. Using induction on k , we get that

$$Q_k(x) := \frac{1-2^{-k}(x+1)^k}{x-1} \in \Pi_{k-1} .$$

This is by

$$Q_k(x) = \frac{1-2^{-k}(x+1)^k}{x-1} = \frac{1-\frac{x+1}{2}+\frac{x+1}{2}-2^{-k}(x+1)^k}{x-1} = Q_1(x)+\frac{x+1}{2}Q_{k-1}(x)$$

This gives

$$Q_k(x)(x-1) = 1-2^{-k}(x+1)^k \Rightarrow Q_k(x)(x-1)^{k+1} = (x-1)^k - 2^{-k}(x^2-1)^k$$

$$\Rightarrow (x-1)^k = 2^{-k}(x^2-1)^k + Q_k(x)(x-1)^{k+1}$$

With $T_h(f,x) := f(x+h)$ we have

$$(T_h - I)^k = 2^{-k}(T_{2h}-I)^k + Q_k(T_h)(T_h-I)^{k+1}.$$

It is evident that $\|Q_k(T_h)\|_{L_p \to L_p} \le M(k)$. Therefore, with $|h| \le t$

$$\begin{aligned}
\|\Delta_h^k f\|_p &\le 2^{-k}\|\Delta_{2h}^k f\|_p + M\|\Delta_h^{k+1} f\|_p \\
&\le 2^{-k}\left(2^{-k}\|\Delta_{4h}^k f\|_p + M\|\Delta_{2h}^{k+1} f\|_p\right) + M\|\Delta_h^{k+1} f\|_p \\
&\le \cdots \\
&\le M\sum_{j=0}^{m} 2^{-jk}\|\Delta_{2^j h}^{k+1} f\|_p + 2^{-km}\|\Delta_{2^m h}^k f\|_p \\
&\le M\sum_{j=0}^{m} 2^{-jk}\omega_{k+1}(f,2^j t)_p + 2^{-k(m-1)}\|f\|_p .
\end{aligned}$$

So if we let $m \to \infty$

$$\begin{aligned}
\omega_k(f,t)_p &\le M\sum_{j=0}^{\infty} 2^{-jk}\omega_{k+1}(f,2^j t)_p \\
&= Mt^k \sum_{j=0}^{\infty} (2^j t)^{-k}\omega_{k+1}(f,2^j t)_p \\
&\le c(k)t^k \sum_{j=0}^{\infty} \int_{2^j t}^{2^{j+1} t} \frac{\omega_{k+1}(f,s)_p}{s^{k+1}}\,ds \\
&= c(k)t^k \int_t^{\infty} \frac{\omega_{k+1}(f,s)_p}{s^{k+1}}\,ds
\end{aligned}$$

Using induction

$$
\begin{aligned}
\omega_k(f,t)_p &\le ct^k\int_t^\infty \frac{\omega_r(f,s)_p}{s^{k+1}}ds \\
&\le ct^k\int_t^\infty s^{r-k-1}ds\int_s^\infty \frac{\omega_{r+1}(f,u)_p}{u^{r+1}}du \\
&\le ct^k\int_t^\infty \frac{\omega_{r+1}(f,u)_p}{u^{r+1}}du\int_t^u s^{r-k-1}ds \\
&\le ct^k\int_t^\infty \frac{\omega_{r+1}(f,u)_p}{u^{r+1}}\left(u^{r-k}-t^{r-k}\right)du \\
&= ct^k\int_t^\infty \frac{\omega_{r+1}(f,u)_p}{u^{k+1}}du - ct^r\int_t^\infty \frac{\omega_{r+1}(f,u)_p}{u^{r+1}}du \\
&\le ct^k\int_t^\infty \frac{\omega_{r+1}(f,u)_p}{u^{k+1}}du.
\end{aligned}
$$

□

The K-functional

For two Banach spaces $X_1 \subset X_0$ the corresponding K-functional

$$K(f,t,X_0,X_1) := \inf_{f=f_0+f_1} \|f_0\|_{X_0} + t|f_1|_{X_1}$$

$$K\left(f,t,L_p(\Omega),W_p^r(\Omega)\right) := \inf_{g\in W_p^r(\Omega)} \|f-g\|_{L_p(\Omega)} + t|g|_{W_p^r(\Omega)}, \qquad 1\le p\le\infty .$$

Theorem [Equivalence of K-functional and modulus] For 'nice domains' $\Omega \subseteq \mathbb{R}^n$, $1\le p\le\infty$, $r\ge 1$, there exist $C_1, C_2 > 0$, such that for any $t>0$

$$C_1 K_r\left(f,t^r\right)_p \le \omega_r(f,t)_p \le C_2 K_r\left(f,t^r\right)_p .$$

It is easy to show that C_2 depends only on r, but the constant C_1 further depends on the geometry of Ω.

Proof of the easy direction Let $f\in L_p(\Omega)$ and let $g\in W_p^r(\Omega)$. Then

$$
\begin{aligned}
\omega_r(f,t)_p &\le \omega_r(f-g,t)_p + \omega_r(g,t)_p \\
&\le 2^r\|f-g\|_{L_p(\Omega)} + C(r)t^r|g|_{W_p^r(\Omega)} \\
&\le C(r)\left(\|f-g\|_{L_p(\Omega)} + t^r|g|_{W_p^r(\Omega)}\right).
\end{aligned}
$$

Taking infimum over all possible $g\in W_p^r(\Omega)$ we obtain the right-hand side.

□

Applications of K-functionals

The K-functional appears in many applications such as denoising. It provides a balance between approximation and smoothness.

1. Regularized Least Squares

$$\min_{g=\sum \alpha_k N_r(\cdot-k)} \|f-g\|_2^2 + t\left\|g^{(2)}\right\|_2^2 .$$

2. Denoising with Total Variation minimization over a bounded domain $\Omega \subset \mathbb{R}^n$

$$\min_{g \in W_2^1(\Omega)} \|f-g\|_2 + t|g|_{1,1}$$

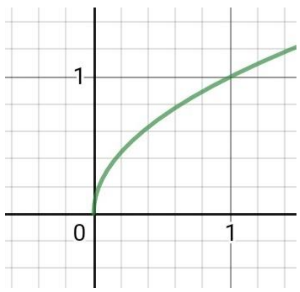
Lip spaces

Def For a domain $\Omega \subset \mathbb{R}^n$ and $0<\alpha \le 1$, we shall say that $f \in Lip(\alpha) = Lip(\alpha,\infty)$, if there exists $M>0$, such that $|f(x)-f(y)| \le M|x-y|^\alpha$, for all $x,y \in \Omega$. We shall denote $|f|_{Lip(\alpha)}$ by the infimum over all M satisfying the condition. Observe that we can replace the condition by

$$|\Delta_h(f,x)| \le M|h|^\alpha, \quad \forall h \in \mathbb{R}^n \Rightarrow \omega_1(f,t)_\infty \le Mt^\alpha \Rightarrow t^{-\alpha}\omega_1(f,t)_\infty \le M .$$

For $1 \le p \le \infty$, we can generalize by

$$|f|_{Lip(\alpha,p)} := \sup_{t>0} t^{-\alpha}\omega_1(f,t)_p .$$



Example For $f(x)=x^\alpha,\ 0<\alpha<1,\ f \in Lip(\alpha),\ f \notin Lip(\beta)\ ,\ \beta>\alpha$.

Proof

(i) Assume $f \in Lip(\beta),\ \beta>\alpha$. Then for $0<x\le 1$,

$$x^\alpha - 0^\alpha = x^\alpha \le M(x-0)^\beta = Mx^\beta \Rightarrow x^{\alpha-\beta} \le M \Rightarrow \text{contradiction}$$

(ii) We use the inequality $(a+b)^\alpha \le a^\alpha + b^\alpha$. Assume w.l.g $x \ge y$, we set $a=y, b=x-y$ and obtain $x^\alpha \le y^\alpha + (x-y)^\alpha \Rightarrow x^\alpha - y^\alpha \le (x-y)^\alpha$.

□

However, for any $0<\alpha\le 1,\ f(x)=x^\alpha \in Lip(1,1)$, because

$$\int_0^1 \left| f'(x) \right| dx = 1 \Rightarrow f' \in L_1$$

$$\Rightarrow \omega_1(f,t)_1 \le t\left|f\right|_1 = t\left\|f'\right\|_1 = t$$

$$\Rightarrow \left|f\right|_{Lip(1,1)} = \sup_{t>0} t^{-1}\omega_1(f,t)_1 \le 1.$$

Generalized Lip are a special case of Besov spaces. For any $\alpha > 0$, let $r := \lfloor \alpha \rfloor + 1$,

$$\left|f\right|_{B^{\alpha}_{p,\infty}} := \sup_{t>0} t^{-\alpha}\omega_r(f,t)_p .$$

Linear approximation of Lip functions

Theorem: Let $f \in Lip(\alpha)$. Approximation with piecewise constants over uniform knots gives

$$E_N(f)_{L_\infty([0,1])} := \inf_{\phi \in S(N_1)^{1/N}} \left\|f-\phi\right\|_\infty \le CN^{-\alpha}\left|f\right|_{Lip(\alpha)} .$$

Proof [Classic technique] Recall that for $g \in C^1[0,1]$, we constructed $\phi_g \in S(N_1)^{1/N}$, such that $E_N(g)_\infty \le \left\|g-\phi_g\right\|_\infty \le N^{-1}\left|g\right|_{1,\infty}$. Therefore, for any $g \in C^1[0,1]$

$$\begin{aligned}\left\|f-\phi_g\right\|_\infty &\le \left\|f-g\right\|_\infty + \left\|g-\phi_g\right\|_\infty \\ &\le \left\|f-g\right\|_\infty + N^{-1}\left|g\right|_{1,\infty} .\end{aligned}$$

For a sequence $\{g_k\}$, with $K_1(f,N^{-1})_\infty = \lim_{k\to\infty}\left\{\left\|f-g_k\right\|_\infty + N^{-1}\left|g_k\right|_{1,\infty}\right\}$, we get

$$\left\|f-\phi_{g_k}\right\|_\infty \le \left\|f-g_k\right\|_\infty + N^{-1}\left|g_k\right|_{1,\infty} \underset{k\to\infty}{\to} K_1(f,N^{-1})_\infty .$$

Using the equivalence of the modulus of smoothness and K-functional,

$$\begin{aligned}E_N(f)_\infty &\le K_1(f,N^{-1})_\infty \\ &\le C\omega_1(f,N^{-1})_\infty \\ &\le CN^{-\alpha}\left|f\right|_{Lip(\alpha)} .\end{aligned}$$

□

Inverse Theorem: Assume $E_N(f)_\infty \le MN^{-\alpha}$, $\forall N \ge 1$. Then, $f \in Lip(\alpha)$.

Intuition $0 \le y < x \le 1$**.** Let $x = y + h$, $(N+1)^{-1} \le h \le N^{-1}$. If $x, y \in \left[kN^{-1},(k+1)N^{-1}\right]$, then with the approximation constant approximation c_k in that interval,

$$\begin{aligned}\left|f(x)-f(y)\right| &\le \left|f(x)-c_k\right| + \left|f(y)-c_k\right| \\ &\le 2MN^{-\alpha} \\ &\le 2M\left|x-y\right|^\alpha\end{aligned}$$

However, since they might not fall in the same interval, there is a “mixing” argument.

So linear approximation is kind of limited when α is small. The problem is that we are not spending enough ‘budget’ in the vicinity of zero.

First glimpse to Adaptive / Nonlinear / Sparse approximation

Approximation using free-knot splines / non-uniform piecewise constants in $L_\infty([0,1])$

$$\Sigma_N := \left\{ \sum_{j=0}^{N-1} c_j \mathbf{1}_{[t_j,t_{j+1})} : T = \{t_j\},\ 0 = t_0 < t_1 < \cdots < t_N = 1 \right\}, \quad \sigma_N(f)_p := \inf_{g\in\Sigma_N} \|f-g\|_p .$$

$$Var(f) := \sup_T \left\{ \sum \left| f(t_{j+1}) - f(t_j) \right| \right\} .$$

If f' exists a.e., $Var(f) = \|f'\|_1$. Why?

$$\int_0^1 |f'(x)|\,dx = \lim_{h\to 0} \sum_k h \frac{|f((k+1)h) - f(kh)|}{h} .$$

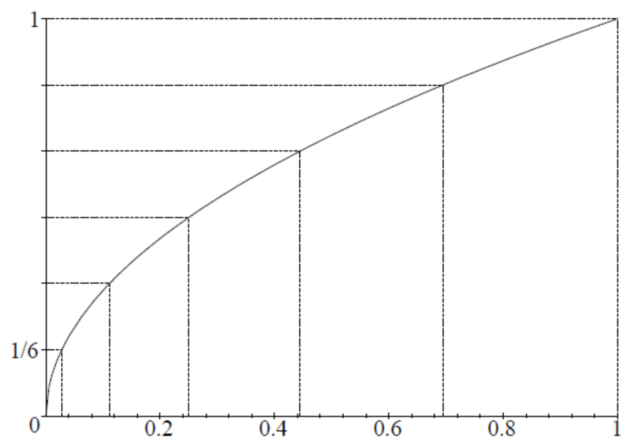
Now, create a partition where

$$Var_{[t_j,t_{j+1}]}(f) \le \frac{Var(f)}{N} .$$

For the example $f(x) = x^\alpha$, $0<\alpha<1$, this is equivalent to choosing

$$\int_0^1 f' = 1 \Rightarrow \int_{t_j}^{t_{j+1}} f'(u)\,du = \frac{1}{N}, \quad 0 \le j \le N-1 .$$

This equidistant partition of the range is achieved by choosing $t_j = \left(\frac{j}{N}\right)^{1/\alpha}$.



If a_j is the median value in $[t_j,t_{j+1}]$, then

$$|f(x) - a_j| \le \frac{Var_{[t_j,t_{j+1}]}(f)}{2} \le \frac{Var(f)}{2N}, \ \forall x \in [t_j, t_{j+1}] .$$

This gives a free knot piecewise constant $g \in \Sigma_N$ with

$$\|f - g\|_\infty \le \frac{Var(f)}{2N} .$$

Recall that earlier on, we promised that 'integration' of differences will be meaningful. Indeed, for the family $f(x) = x^\alpha$, $0 < \alpha < 1$, we see the smoothness

$$|f|_{Lip(1,1)} = \sup_{t>0} t^{-1}\omega_1(f,t)_1 = 1,$$

comes into play to show the advantage of nonlinear approximation over linear approximation

$$f \in Lip(\alpha,\infty), \qquad f \in Lip(1,1),$$

$$E_N(f)_\infty \sim N^{-\alpha} , \qquad \sigma_N(f)_\infty \sim N^{-1}.$$

Jackson Theorem for trigonometric polynomials

Denote $E_N(f)_p := \inf_{P\in\Pi_N} \|f - P\|_{L_p(\mathbb{T})}$. Here, we shall assume we are approximating real functions. This implies we can use real trigonometric polynomials. Our constructive approximation will in fact guarantee that.

Theorem For a periodic function $f \in L_p(\mathbb{T})$, $1 \le p \le \infty$ and any $r \ge 1$,

$$E_N(f)_p \le C(r)\omega_r(f, N^{-1})_p$$

Corollary for $f \in W_p^r(\mathbb{T})$ we obtain

$$E_N(f)_p \le C(r)N^{-r}|f|_{W_p^r(\mathbb{T})}.$$

Corollary Let $f:\mathbb{T}\to\mathbb{R}$ be piecewise algebraic polynomial of degree $r-1$ with the number of breakpoints $\#break$. Then, for large enough N

$$E_N(f)_p \le C\omega_r(f, N^{-1})_p \le C(p,r,\|f\|_\infty)\left(\frac{\#break}{N}\right)^{1/p}.$$

Proof of the Jackson theorem Recall the Fejér kernel of degree $m-1$

$$K_m(t) = \frac{1}{m}\left(\frac{\sin(mt/2)}{\sin(t/2)}\right)^2.$$

We construct the approximating trigonometric polynomial using the **Jackson kernel**

$$J_{N,r}(t) := \lambda_{N,r}\left(\frac{\sin(mt/2)}{\sin(t/2)}\right)^{2r}, \qquad \int_{\mathbb{T}} J_{N,r}(t)\,dt = 1, \qquad m := \left\lfloor \frac{N}{r} \right\rfloor + 1.$$

It is a positive, symmetric kernel, trigonometric polynomial of degree $\le N$ because

$$r(m-1) = r\left\lfloor \frac{N}{r} \right\rfloor \le r\frac{N}{r} = N .$$

Also, since it is an even trigonometric polynomial, we can write it as

$$J_{N,r}(x) = \sum_{k=0}^{N} a_k \cos(kx) .$$

Based on a theorem we proved (for domains of finite volume), if $f \in L_p(\mathbb{T})$, $1 \le p \le \infty$, then $f \in L_1(\mathbb{T})$. Observe that for any trigonometric polynomial $P \in \Pi_N$, we have $f * P \in \Pi_N$. Indeed, if $P(x) = \sum_{k=-N}^{N} a_k e^{ikx}$, then for $|k| > N$

$$(f * P)^{\wedge}(k) = \hat{f}(k)\underbrace{\hat{P}(k)}_{=0} = 0 .$$

The actual approximating polynomial is of a more sophisticated form of convolution

$$S_{N,r}(f,x) := \int_{\mathbb{T}} \left[(-1)^{r+1} \Delta_t^r(f,x) + f(x)\right] J_{N,r}(t)\,dt .$$

Notice that $\int_{\mathbb{T}} \big(-f(x) + f(x)\big) J_{N,r}(t)\,dt = 0$. This means that $S_{N,r}(f,x)$ is a combination of terms

$$\int_{\mathbb{T}} f(x+kt)\cos(lt)\,dt , \qquad k = 1,\ldots,r , \qquad l = 0,\ldots,N .$$

We want to show that $S_{N,r}(f,x) \in \Pi_N(\mathbb{T})$. Now $f(x+kt)$ as a function of t has period $2\pi/k$. This means that $\int_{\mathbb{T}} f(x+kt)\cos(lt)\,dt = 0$, unless k divides l. To see this let $g(t)$ have period $2\pi/k$. Then for any $l \in \mathbb{Z}$

$$\int_0^{2\pi} g(t)e^{ilt}dt = \int_{2\pi/k}^{2\pi+2\pi/k} g(t)e^{ilt}dt = \int_0^{2\pi} g(y+2\pi/k)e^{il(y+2\pi/k)}dy = \int_0^{2\pi} g(y)e^{il(y+2\pi/k)}dy = e^{i\frac{2\pi l}{k}}\int_0^{2\pi} g(y)e^{ily}dy$$

$$\Rightarrow e^{i\frac{2\pi l}{k}} = 1 \ \ or \ \ \int_0^{2\pi} g(y)e^{ily}dy = 0 .$$

Thus, we get for k that divides l

$$\begin{aligned}
\int_{\mathbb{T}} f(x+kt)\cos(lt)\,dt &= \frac{1}{k}\int_{x-\pi k}^{x+\pi k} f(y)\cos\left(\frac{l}{k}(y-x)\right)dy \\
&= \int_0^{2\pi} f(y)\cos\left(\frac{l}{k}(y-x)\right)dy \\
&= \left(\int_0^{2\pi} f(y)\cos\left(\frac{l}{k}y\right)dy\right)\cos\left(\frac{l}{k}x\right) + \left(\int_0^{2\pi} f(y)\sin\left(\frac{l}{k}y\right)dy\right)\sin\left(\frac{l}{k}x\right)
\end{aligned}$$

So $S_{N,r}(f,x)$ is composed of trigonometric polynomial terms of degree $\le N$. Now we use the following:

(i) $\omega_r(f,t)_p = \omega_r\left(f,\frac{Nt}{N}\right)_p \le (Nt+1)^r\,\omega_r\left(f,\frac{1}{N}\right)_p$

(ii) Lemma 7.2.1 in Constructive Approximation shows that $\int_0^\pi t^k J_{N,r}(t)\,dt \le C(r)N^{-k}$, $k=0,\ldots,2r-2$.

Therefore

$$\begin{aligned}
\left\|S_{N,r}(f,x)-f\right\|_p &= \left\|\int_{\mathbb{T}}\left((-1)^{r+1}\Delta_t^r(f,\cdot)+f(\cdot)-f(\cdot)\right)J_{N,r}(t)\,dt\right\|_p \\
&\le \left\|\int_{\mathbb{T}}\left|\Delta_t^r(f,\cdot)\right|J_{N,r}(t)\,dt\right\|_p \\
&\underset{\text{Minkowski}}{\le} \int_{\mathbb{T}}\omega_r\left(f,|t|\right)_p J_{N,r}(t)\,dt \\
&\le \omega_r\left(f,\frac{1}{N}\right)_p \int_{\mathbb{T}}\left(N|t|+1\right)^r J_{N,r}(t)\,dt \\
&= 2\omega_r\left(f,\frac{1}{N}\right)_p \sum_{k=0}^{r}\binom{r}{k}N^k\int_0^\pi t^k J_{N,r}(t)\,dt \\
&\le C(r)\,\omega_r\left(f,\frac{1}{N}\right)_p .
\end{aligned}$$

□

Besov Spaces

Continuous definition

Let $\alpha>0$, $0<q,p\le\infty$. Let $r\ge\lfloor\alpha\rfloor+1$. The Besov space $B_q^\alpha\left(L_p(\Omega)\right)$ is the collection of functions $f\in L_p(\Omega)$ for which

$$|f|_{B_q^\alpha(L_p(\Omega))} := \begin{cases}\left(\int_0^\infty\left[t^{-\alpha}\omega_r(f,t)_p\right]^q\frac{dt}{t}\right)^{1/q}, & 0<q<\infty,\\ \sup_{t>0} t^{-\alpha}\omega_r(f,t)_p, & q=\infty.\end{cases}$$

is finite. The norm is

$$\|f\|_{B_q^\alpha(L_p(\Omega))} := \|f\|_{L_p(\Omega)} + |f|_{B_q^\alpha(L_p(\Omega))}.$$

Theorem The space $B_q^\alpha\left(L_p(\Omega)\right)$ does not depend on the choice of $r\ge\lfloor\alpha\rfloor+1$.

Proof For $\Omega=\mathbb{R}^n$, $1\le q<\infty$. Let $r_2>r_1\ge\lfloor\alpha\rfloor+1$. We already know that for $1\le p\le\infty$, and any $t>0$, $\omega_{r_2}(f,t)_p\le 2^{r_2-r_1}\omega_{r_1}(f,t)_p$ (for $0<p\le 1$ with a different constant), so

$$\int_0^\infty\left[t^{-\alpha}\omega_{r_2}(f,t)_p\right]^q\frac{dt}{t}\le c\int_0^\infty\left[t^{-\alpha}\omega_{r_1}(f,t)_p\right]^q\frac{dt}{t} .$$

The other direction requires the Marchaud inequality

$$\int_0^\infty \left[t^{-\alpha}\omega_{r_1}(f,t)_p\right]^q \frac{dt}{t} \le c\int_0^\infty \left[t^{r_1-\alpha}\int_t^\infty \frac{\omega_{r_2}(f,s)_p}{s^{r_1+1}}ds\right]^q \frac{dt}{t}.$$

Denote $\theta := r_1-\alpha > 0$, and $\phi(s) := s^{-r_1}\omega_{r_2}(f,s)_p$. Then, we can apply the Hardy inequality [DL Theorem 2.3.1] for $1 \le q < \infty$, to the right-hand side

$$\begin{aligned}\int_0^\infty \left[t^{r_1-\alpha}\int_t^\infty \frac{\omega_{r_2}(f,s)_p}{s^{r_1+1}}ds\right]^q \frac{dt}{t} &= \int_0^\infty \left[t^{\theta}\int_t^\infty \frac{\phi(s)}{s}ds\right]^q \frac{dt}{t} \\ &\underset{\text{Hardy}}{\le} \frac{1}{\theta^q}\int_0^\infty \left[t^\theta \phi(t)\right]^q \frac{dt}{t} \\ &= \frac{1}{(r_1-\alpha)^q}\int_0^\infty \left[t^{r_1-\alpha}t^{-r_1}\omega_{r_2}(f,t)_p\right]^q \frac{dt}{t} \\ &= c\int_0^\infty \left[t^{-\alpha}\omega_{r_2}(f,t)_p\right]^q \frac{dt}{t}\end{aligned}$$

□

In certain cases, we will ask for the condition $r \ge \lfloor \alpha \rfloor + 1$. Otherwise, the space might be 'trivial'

Theorem (univariate case) For $r < \alpha$, $1 \le p \le \infty$, we get that $B_q^\alpha(L_p(\Omega)) = \Pi_{r-1}$ if $\Omega = [a,b]$ and $B_q^\alpha(L_p(\Omega)) = \{0\}$ if $\Omega = \mathbb{R}$.

Proof (sketch, see Proposition 2.7.1 in CA) If $f \in B_q^\alpha(L_p(\Omega))$, then $t^{-\alpha}\omega_r(f,t)_p \le C$, $0 < t \le 1$. This implies that $t^{-r}\omega_r(f,t)_p \le Ct^\varepsilon \underset{t\to 0}{\to} 0$, where $\alpha = r + \varepsilon$. The condition $t^{-r}\omega_r(f,t)_p \underset{t\to 0}{\to} 0$, in turn gives that $f^{(r)} = 0$ and so $f \in \Pi_{r-1}$.

□

Theorem For a bounded domain we can equivalently integrate the semi-norm on $[0,1]$. That is,

$$|f|_{B_q^\alpha(L_p(\Omega))} \sim \begin{cases} \left(\int_0^1 \left[t^{-\alpha}\omega_r(f,t)_p\right]^q \frac{dt}{t}\right)^{1/q}, & 0 < q < \infty, \\ \sup_{0<t\le 1} t^{-\alpha}\omega_r(f,t)_p, & q = \infty. \end{cases}$$

Proof If Ω is bounded, then we have $\omega_r(f,t)_p \equiv const$ for $t \ge diam(\Omega)$. Therefore for $1/2 \le t \le \infty$,

$$\omega_r(f,1/2)_p \le \omega_r(f,t)_p \le \omega_r(f,diam(\Omega))_p = \omega_r\left(f,\frac{2diam(\Omega)}{2}\right)_p \le (1+2diam(\Omega))^r \omega_r(f,1/2)_p.$$

This gives

$$\int_1^\infty \left[t^{-\alpha}\omega_r(f,t)_p\right]^q \frac{dt}{t} \le C\left(\omega_r(f,1/2)_p\right)^q \int_1^\infty t^{-q\alpha-1}dt$$
$$\le C\left(\omega_r(f,1/2)_p\right)^q$$
$$\le C(\alpha,q,\Omega)\int_{1/2}^1 \left[t^{-\alpha}\omega_r(f,t)_p\right]^q \frac{dt}{t}$$

□

Lemma For any domain taking the integral over $[0,1]$ gives a quasi-norm equivalent to $\|f\|_{B_q^\alpha(L_p(\Omega))}$

Proof We replace the integral over $[1,\infty]$ by

$$\int_1^\infty \left[t^{-\alpha}\omega_r(f,t)_p\right]^q \frac{dt}{t} \le C\|f\|_p^q \int_1^\infty t^{-q\alpha-1}dt$$
$$= C(\alpha,q)\|f\|_p^q .$$

Therefore

$$\|f\|_{B_q^\alpha(L_p(\Omega))} \sim \|f\|_p + \left(\int_0^1 \left[t^{-\alpha}\omega_r(f,t)_p\right]^q \frac{dt}{t}\right)^{1/q} .$$

□

Theorem $B_{q_1}^{\alpha_1}(L_p) \subseteq B_{q_2}^{\alpha_2}(L_p)$ if $\alpha_2 < \alpha_1$.

Proof $(q_1 = q_2)$ We may use $r_1 = \lfloor \alpha_1 \rfloor + 1 \ge \lfloor \alpha_2 \rfloor + 1$ to equivalently define $B_{q_2}^{\alpha_2}(L_p)$. This gives

$$\|f\|_{B_q^{\alpha_2}(L_p)} \le C\left(\|f\|_p + \left(\int_0^1 \left[t^{-\alpha_2}\omega_{r_1}(f,t)_p\right]^q \frac{dt}{t}\right)^{1/q}\right)$$
$$\le C\left(\|f\|_p + \left(\int_0^1 \left[t^{-\alpha_1}\omega_{r_1}(f,t)_p\right]^q \frac{dt}{t}\right)^{1/q}\right)$$
$$\le C\|f\|_{B_q^{\alpha_1}(L_p)}$$

□

Theorem $W_p^m \subseteq B_q^\alpha(L_p)$, $\forall \alpha < m$, $1 \le p \le \infty$, $0 < q \le \infty$.

Proof Let $g \in W_p^m(\Omega)$. This implies $g \in L_p(\Omega)$. We have that $r := \lfloor \alpha \rfloor + 1 \le m$. It is sufficient to take the integral over $[0,1]$.

$$\int_0^1 \left[t^{-\alpha}\omega_r(g,t)_p\right]^q \frac{dt}{t} \le C\int_0^1 \left[t^{-\alpha}t^r |g|_{r,p}\right]^q \frac{dt}{t}$$
$$\le C|g|_{r,p}^q \int_0^1 t^{(r-\alpha)q-1}dt$$
$$\le C|g|_{r,p}^q .$$

□

Discretization of the Besov semi-norm

Theorem One has the following equivalent form of the Besov semi-norm

$$|f|_{B_q^\alpha(L_p(\Omega))} \sim \begin{cases} \left(\sum_{k=-\infty}^{\infty} \left[2^{k\alpha} \omega_r\left(f, 2^{-k}\right)_p \right]^q \right)^{1/q}, & 0 < q < \infty. \\ \sup_{k \in \mathbb{Z}} 2^{k\alpha} \omega_r\left(f, 2^{-k}\right)_p, & q = \infty. \end{cases}$$

Proof Define $\varphi(t) := t^{-\alpha}\omega_r(f,t)_p$. Then we claim that for $t \in \left[2^{-k-1}, 2^{-k}\right]$, $k \in \mathbb{Z}$, we have

$$2^{-r}\varphi\left(2^{-k}\right) \le \varphi(t) \le 2^{\alpha}\varphi\left(2^{-k}\right).$$

To see that, we use the following properties:

(i) $\omega_r(f,t)_p$ is non-decreasing

(ii) $\omega_r(f,Nt)_p \le N^r \omega_r(f,t)_p$

The left-hand side

$$2^{-r}\varphi\left(2^{-k}\right) = 2^{k\alpha-r}\omega_r\left(f,2^{-k}\right)_p = 2^{k\alpha-r}\omega_r\left(f,22^{-k-1}\right)_p$$
$$\underset{\text{(ii)}}{\le} 2^{k\alpha-r}2^r\omega_r\left(f,2^{-k-1}\right)_p \underset{\text{(i)}}{\le} 2^{k\alpha}\omega_r(f,t)_p \le t^{-\alpha}\omega_r(f,t)_p$$

The right-hand side

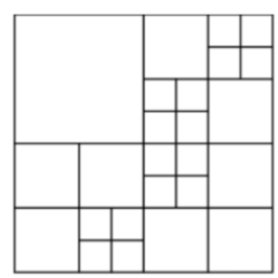
$$t^{-\alpha}\omega_r(f,t)_p \underset{\text{(i)}}{\le} t^{-\alpha}\omega_r\left(f,2^{-k}\right)_p \le 2^{(k+1)\alpha}\omega_r\left(f,2^{-k}\right)_p \le 2^{\alpha}\varphi\left(2^{-k}\right)$$

This gives us for $0 < q < \infty$, $k \in \mathbb{Z}$

$$\int_{2^{-k-1}}^{2^{-k}} \varphi(t)^q \frac{dt}{t} \sim \varphi\left(2^{-k}\right)^q \int_{2^{-k-1}}^{2^{-k}} \frac{dt}{t} \sim \varphi\left(2^{-k}\right)^q \Rightarrow \int_{2^{-k-1}}^{2^{-k}} \left(t^{-\alpha}\omega_r(f,t)_p\right)^q \frac{dt}{t} \sim \left[2^{k\alpha}\omega_r\left(f,2^{-k}\right)_p\right]^q.$$

□

Discretization over cubes



Definition [Dyadic cubes] Let $D := \left\{D_k : k \in \mathbb{Z}\right\}$

$$D_k := \left\{Q = 2^{-kn}\left[m_1, m_1+1\right] \times \cdots \times \left[m_n, m_n+1\right] : m \in \mathbb{Z}^n\right\}.$$

Observe that $Q \in D_k \Rightarrow |Q| = 2^{-kn}$.

For nonlinear/adaptive/sparse approximation in $L_p(\Omega)$, $\Omega \subseteq \mathbb{R}^n$, it is useful to use the special cases of Besov spaces

$$B_\tau^\alpha := B_\tau^\alpha\left(L_\tau(\Omega)\right), \qquad \frac{1}{\tau} = \frac{\alpha}{n} + \frac{1}{p}.$$

Theorem $\Omega = \mathbb{R}^n$. We have the equivalence

$$|f|_{B_\tau^\alpha} \sim \left(\sum_{k\in\mathbb{Z}} \left(2^{k\alpha} \omega_r \left(f, 2^{-k} \right)_\tau \right)^\tau \right)^{1/\tau} \sim \left(\sum_{Q\in D} \left(|Q|^{-\alpha/n} \omega_r \left(f, Q \right)_\tau \right)^\tau \right)^{1/\tau} ,$$

$$\omega_r \left(f, Q \right)_\tau := \sup_{h\in\mathbb{R}^n} \left\| \Delta_h^r \left(f, Q, \cdot \right) \right\|_{L_\tau(Q)} .$$

.

The following theorem generalizes what we showed for the univariate case

Theorem Let $f(x) = \mathbf{1}_{\tilde{\Omega}}(x)$, $\tilde{\Omega} \subset [0,1]^n$, a domain with smooth boundary. Then $f \in B_\tau^\alpha$, $\alpha < 1/\tau$.

Proof For $\Omega = [0,1]^n$, with $l(Q)$ denoting the level of the cube Q, we may take the sum over $k \ge 0$

$$|f|_{B_\tau^\alpha} \sim \left(\sum_{Q\in D, l(Q)\ge 0} \left(|Q|^{-\alpha/n} \omega_r \left(f, Q \right)_\tau \right)^\tau \right)^{1/\tau} .$$

For any Q, we have that $\omega_r \left(f, Q \right)_\tau = 0$, if $\partial\tilde{\Omega} \cap Q = \emptyset$. Otherwise, if $l(Q) = k$,

$$\omega_r \left(f, Q \right)_\tau \le C \| f \|_{L_\tau(Q)} \le C \left(\int_Q 1^\tau \right)^{1/\tau} = C |Q|^{1/\tau} = C 2^{-kn/\tau} .$$

Therefore,

$$\begin{aligned} |f|_{B_\tau^\alpha}^\tau &\le C \sum_{l(Q)\ge 0} \left(|Q|^{-\alpha/n} \omega_r \left(f, Q \right)_\tau \right)^\tau \\ &\le C \sum_{k=0}^{\infty} \left(2^{k\alpha} 2^{-kn/\tau} \right)^\tau \# \left\{ Q : l(Q) = k,\ Q \cap \partial\tilde{\Omega} \ne \emptyset \right\} \\ &= C \sum_{k=0}^{\infty} 2^{k(\alpha\tau - n)} \# \left\{ Q : l(Q) = k,\ Q \cap \partial\tilde{\Omega} \ne \emptyset \right\} \end{aligned}$$

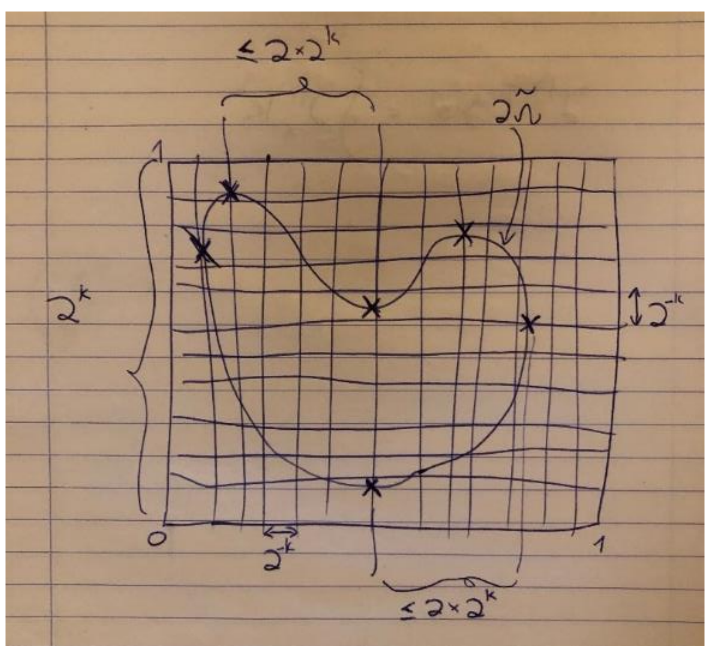
We argue that

$$\# \left\{ Q : l(Q) = k,\ Q \cap \partial\tilde{\Omega} \ne \emptyset \right\} \le c\left(\tilde{\Omega} \right) 2^{k(n-1)} . \quad (*)$$

This implies that if $\alpha < 1/\tau$

$$|f|_{B_\tau^\alpha}^\tau \le C \sum_{k=0}^{\infty} 2^{k(\alpha\tau - n)} 2^{k(n-1)} = C \sum_{k=0}^{\infty} 2^{k(\alpha\tau - 1)} < \infty .$$

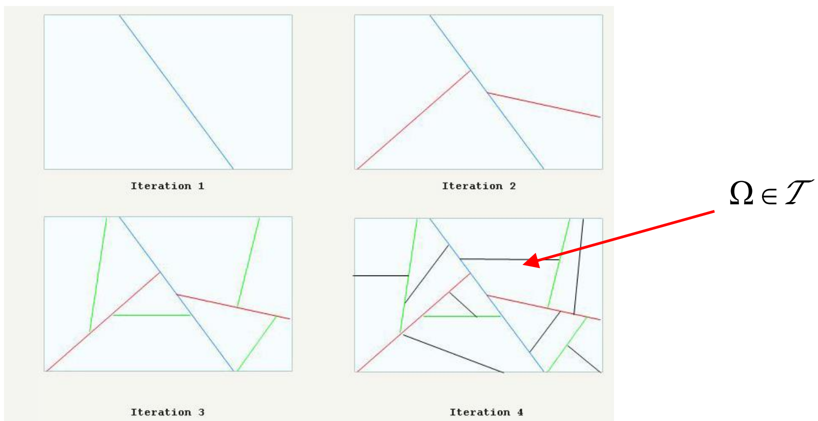
Let's get back to the estimate (*). Let use show a picture argument for $\tilde{\Omega} \subset [0,1]^2$. There is a finite number of points where the gradient of the boundary of the domain is aligned with one of the main axes. Between these points, the boundary segments are monotone in x_1 and x_2, and therefore can only intersect at most 2×2^k dyadic cubes.



□

Note For the theory of geometric approximation in higher dimensions (Machine Learning!), we can generalize to anisotropic partitions of trees over $[0,1]^n$ (replacing dyadic cubes!)

$$|f|_{B_\tau^\alpha(\mathcal{T})} := \left(\sum_{\Omega \in \mathcal{T}} \left(|\Omega|^{-\alpha} \omega_r(f,\Omega)_\tau \right)^\tau \right)^{1/\tau}$$



Approximation from Shift-invariant spaces

Applied problem We want to re-sample an image: change its size or rotate it.

Input $\Omega = \mathbb{R}^n$, $f(k), k \in \mathbb{Z}^n$ (although the application is image processing and the boundaries need a special treatment)

Output Given a generating function ϕ, we look for coefficients $\{\alpha_k\}$ such that

$$f(x) \approx \sum_k \alpha_k \phi(x-k), \qquad x \in \mathbb{R}^n.$$

Application Example I – Image rotation Once we find $\{\alpha_k\}$, we can apply an affine transformation A by sampling

$$f_A(j) = f(A^{-1}j) \approx \sum_k \alpha_k \phi(A^{-1}j - k).$$

Fig. 4. Comparison between three interpolation methods of same order ($L = 4$), same support, and of decreasing asymptotic constant: 15 rotations of the top images by an angle of $2\pi/15$. The smaller the asymptotic constant, the better the quality. Remarks: The interpolation condition of the cubic I-MOMS is detrimental to accuracy; the cubic spline is the smoothest cubic MOMS, but does not provide the highest quality; the cubic O-MOMS minimizes the asymptotic constant among cubic MOMS and gives the best results, even though it is not smooth.

*T. Blu, P. Thévenaz and M. Unser, MOMS: maximal-order interpolation of minimal support, IEEE transactions on image processing 10 (2001), 1069-1080.

Application Example 2 - Image resizing. When we create a smaller version of an image, one can argue that we can simply sub-sample pixels at the correct rate. This leads to bad visual quality.

Simple example – Think of an image of a chess-board, with pattern of black and white at the pixel level. In zoom out it visually looks gray. Then, sub-sample every 4 pixels by the top left, to create a smaller image. You will arbitrarily get a black or white image instead of a gray image.

(a) Resizing by simple subsampling (b) Resizing via 'ideal' low-pass (followed by subsampling)

Definition For any $k \in \mathbb{Z}^n$ we denote the linear shift operator S_k by $S_k(f) := f(\cdot - k)$.

Definition Let V be a closed subspace of $L_p(\mathbb{R}^n)$, $1 \le p \le \infty$. We say that V is a ***shift invariant*** (SI) space if it is invariant under the operators $\{S_k \mid k \in \mathbb{Z}^n\}$. We say that a set Φ ***generates*** V if $V = S(\Phi) := \overline{span}\{\phi(\cdot - k) \mid \phi \in \Phi, \mathrm{k} \in \mathbb{Z}^n\}$. We say that V is a ***finite shift invariant*** (FSI) space, if there exists a finite generating set Φ, $|\Phi| = m$, such that $V = S(\Phi)$. In such a case we say that V is of ***length*** $\le m$. We denote $len(V) := \min\{|\Phi| \mid V = S(\Phi)\}$. An SI space V is called a ***principal shift invariant*** (PSI) space if $len(V) = 1$.

To approximate functions with arbitrary precision one uses dilates of shift invariant spaces. For a given subspace V and $h > 0$ we denote by V^h the dilated space

$$V^h := \{\phi(\cdot / h) \mid \phi \in V\}.$$

Assignment If $S(\phi)$ is a PSI space, then for $j \ge 0$, $S(\phi)^{2^{-j}}$ is a FSI space of length 2^{nj}. That is, as an integer shift invariant space, it is generated by at least 2^{nj} generators.

Properties of 'good' generators

1. Smoothness
2. Refinability (to be defined later. Only required for certain applications)
3. Localization properties - compact support or fast decay
4. **Approximation order**

Def We say that $V = S(\Phi)$ or Φ provide ***approximation order*** r if for any $g \in W_p^r(\mathbb{R}^n)$ and $h > 0$ we have

$$E\left(g, S(\Phi)^h\right)_p \le C h^r |g|_{r,p}.$$

Observe that this automatically gives for any $f \in L_p(\mathbb{R}^n)$

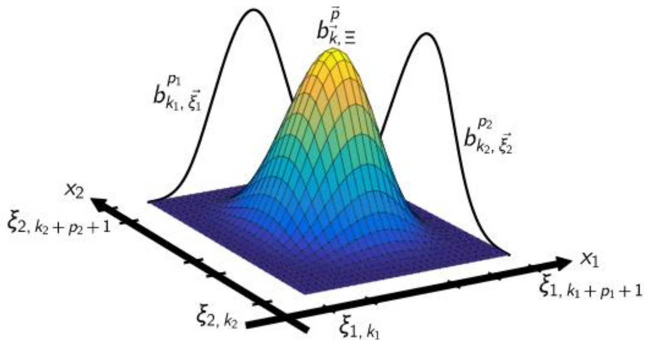
$$E\left(f, S(\Phi)^h\right)_p \le C\omega_r(f, h)_p.$$

We already proved the following Jackson-type estimate for piecewise constant approximation

Theorem $E\left(g, S(N_1)^h\right)_{L_p(\mathbb{R})} \le C h |g|_{W_p^1(\mathbb{R})}$.

In the multivariate case the B-spline N_r, $r \ge 1$, is defined as the 'tensor-product' of the one-dimensional B-spline $\tilde{N}_r$,

$$N_r(x) = N_r(x_1,\ldots,x_n) := \prod_{i=1}^{n} \tilde{N}_r(x_i).$$



Bivariate N_4

We want to show a generalized Jackson-type estimate for $r \geq 1$, $n \geq 1$,

$$E\left(g, S(N_r)^h\right)_{L_p(\mathbb{R}^n)} \leq Ch^r |g|_{W_p^r(\mathbb{R}^n)}.$$

To this end, we construct appropriate dual(s) to the multivariate B-spline that have sufficient decay and allow polynomial reproduction.

Polynomial reproduction

It is obvious that for $r = 1$

$$a = \sum_k aN_1(x-k), \qquad \forall x \in \mathbb{R}, a \in \mathbb{R} \text{ (or a.e).}$$

It is also easy to see that for $r = 2$ and any linear function $f(x) = ax + b$, we have

$$ax + b = \sum_k (ak+b) N_2(x+1-k).$$

The next theorem says that the shifts of the B-splines **reproduce polynomials**

Theorem [CA chapter 4] For any $r \geq 1$, there exist linear functionals $\{g_k\}$ on $\Pi_{r-1}(\mathbb{R})$, with support on $[k, k+r]$, such that for any univariate polynomial $P \in \Pi_{r-1}(\mathbb{R})$,

$$P(x) = \sum_{k\in\mathbb{Z}} g_k(P) N_r(x-k).$$

Theorem Suppose that for $r \geq 1$ and a bounded $\varphi : \mathbb{R} \to \mathbb{R}$, with sufficiently fast decay, there exist linear functionals $\{\tilde{g}_k\}$ on $\Pi_{r-1}(\mathbb{R})$, such that for any univariate polynomial $\tilde{P} \in \Pi_{r-1}(\mathbb{R})$,

$$\tilde{P}(x) = \sum_{k\in\mathbb{Z}} \tilde{g}_k(P) \varphi(x-k).$$

Then for $\phi(x) := \prod_{i=1}^{n} \varphi(x_i)$, there exist linear functionals $\{g_k\}$, such that for any $P \in \Pi_{r-1}(\mathbb{R}^n)$

$$P(x) = \sum_{k\in\mathbb{Z}^n} g_k(P)\phi(x-k).$$

Proof Let $x^\alpha := \prod_{i=1}^{n} x_i^{\alpha_i}$, with $\alpha_i \le r-1$. Since φ reproduces univariate polynomials, for $1 \le i \le n$,

$$x_i^{\alpha_i} = \sum_{k_i} \tilde{g}_{k_i}\left(x_i^{\alpha_i}\right)\varphi(x_i - k_i).$$

This gives

$$\begin{aligned} x^\alpha &= \prod_{i=1}^{n}\left(\sum_{k_i} \tilde{g}_{k_i}\left(x_i^{\alpha_i}\right)\varphi(x_i-k_i)\right) \\ &= \sum_{k\in\mathbb{Z}^n}\prod_{i=1}^{n} \tilde{g}_{k_i}\left(x_i^{\alpha_i}\right)\varphi(x_i-k_i) \\ &= \sum_{k\in\mathbb{Z}^n}\underbrace{\left(\prod_{i=1}^{n} \tilde{g}_{k_i}\left(x_i^{\alpha_i}\right)\right)}_{=:g_k(x^\alpha)}\phi(x-k) \\ &= \sum_{k\in\mathbb{Z}^n} g_k\left(x^\alpha\right)\phi(x-k), \end{aligned}$$

where we define

$$g_k\left(x^\alpha\right) := \left(\prod_{i=1}^{n} \tilde{g}_{k_i}\left(x_i^{\alpha_i}\right)\right).$$

Now, for any $P \in \Pi_{r-1}(\mathbb{R}^n)$, $P(x) = \sum_{|\alpha|<r} a_\alpha x^\alpha$, we define

$$g_k(P) := \sum_{|\alpha|<r} a_\alpha g_k\left(x^\alpha\right).$$

This gives

$$\begin{aligned} \sum_{k\in\mathbb{Z}^n} g_k(P)\phi(x-k) &= \sum_{k\in\mathbb{Z}^n}\left(\sum_{|\alpha|<r} a_\alpha g_k\left(x^\alpha\right)\right)\phi(x-k) \\ &= \sum_{|\alpha|<r} a_\alpha \sum_{k\in\mathbb{Z}^n} g_k\left(x^\alpha\right)\phi(x-k) \\ &= \sum_{|\alpha|<r} a_\alpha x^\alpha = P(x). \end{aligned}$$

□

Assume that for a generator ϕ, there exists a fast decaying **dual** $\tilde{\phi}$, reproducing polynomials. That is,

$$P(x) = \sum_{k\in\mathbb{Z}^n} g_k(P)\phi(x-k), \quad \text{with } g_k(P) = \left\langle P, \tilde{\phi}(\cdot - k)\right\rangle.$$

This leads to the construction of the **reproducing kernel:**

$$K(x,y) = \sum_{k\in\mathbb{Z}^n} \tilde{\phi}(y-k)\phi(x-k).$$

$$
\begin{aligned}
Tf(x) &:= \int_{\mathbb{R}^n} K(x,y) f(y)\,dy \\
&= \int_{\mathbb{R}^n} \left(\sum_{k\in\mathbb{Z}^n} \tilde{\phi}(y-k)\phi(x-k) \right) f(y)\,dy \\
&= \sum_{k\in\mathbb{Z}^n} \left(\int_{\mathbb{R}^n} \tilde{\phi}(y-k) f(y)\,dy \right) \phi(x-k) \\
&= \sum_{k\in\mathbb{Z}^n} \left\langle f, \tilde{\phi}(\cdot - k) \right\rangle \phi(x-k).
\end{aligned}
$$

Therefore,

$$
P(x) = \int_{\mathbb{R}^n} K(x,y) P(y)\,dy
$$

We then define

$$
K_h(x,y) := h^{-n} K\left(h^{-1}x, h^{-1}y\right), \; h>0,
$$

and

$$
T_h f(x) := \int_{\mathbb{R}^n} K_h(x,y) f(y)\,dy.
$$

Observe that $\{T_h\}_{h>0}$ are linear operators.

Assignment Prove $T_h f(x) \in S(\phi)^h$, $h>0$.

Theorem Assume a kernel operator $K(x,y)$ satisfies for $r \geq 1$

(i) $P(x) = \int_{\mathbb{R}^n} K(x,y) P(y)\,dy$, $\forall P \in \Pi_{r-1}\left(\mathbb{R}^n\right)$,

(ii) $\left|K(x,y)\right| \leq c \dfrac{1}{\left(1+|x-y|\right)^{n+r+\varepsilon}}$, for some $\varepsilon > 0$ and any $x, y \in \mathbb{R}^n$,

Then, for all $1 \leq p \leq \infty$ and $g \in W_p^r\left(\mathbb{R}^n\right)$

$$
\left\| g - T_h g \right\|_p \leq C h^r \left| g \right|_{r,p}, \qquad h > 0.
$$

Proof for $p=\infty$. Let $g \in C^r\left(\mathbb{R}^n\right) \cap W_p^r\left(\mathbb{R}^n\right)$.

Taylor polynomial $T_{r-1,x} g(y) := \sum_{|\alpha|<r} \dfrac{\partial^\alpha g(x)}{\alpha!}(y-x)^\alpha \in \Pi_{r-1}$,

The estimate of Taylor remainder $\left|R_{r,x} g(y)\right| \leq c\left|y-x\right|^r \max_{z\in B(x,|y-x|)} \max_{|\alpha|=r} \left|\partial^\alpha g(z)\right|$. For $p=\infty$

$$
\begin{aligned}
\left| g(x) - \int_{\mathbb{R}^n} K(x,y) g(y)\,dy \right| &= \left| g(x) - \int_{\mathbb{R}^n} K(x,y)\left(T_{r-1,x} g(y) + R_{r,x} g(y)\right) dy \right| \\
&= \left| g(x) - \int_{\mathbb{R}^n} K(x,y) T_{r-1,x} g(y)\,dy - \int_{\mathbb{R}^n} K(x,y) R_{r,x} g(y)\,dy \right| \\
&= \left| g(x) - T_{r-1,x} g(x) - \int_{\mathbb{R}^n} K(x,y) R_{r,x} g(y)\,dy \right| \\
&= \left| g(x) - g(x) - \int_{\mathbb{R}^n} K(x,y) R_{r,x} g(y)\,dy \right| \\
&= \left| \int_{\mathbb{R}^n} K(x,y) R_{r,x} g(y)\,dy \right|
\end{aligned}
$$

$$
\begin{aligned}
\left| g(x) - \int_{\mathbb{R}^n} K(x,y) g(y)\, dy \right| &\le \int_{\mathbb{R}^n} |K(x,y)| \left| R_{r,x} g(y) \right| dy \\
&\le C |g|_{r,\infty} \int_{\mathbb{R}^n} |K(x,y)| |y-x|^r \, dy \\
&\le C |g|_{r,\infty} \int_{\mathbb{R}^n} \frac{1}{\left(1+|y-x|\right)^{n+r+\varepsilon}} |y-x|^r \, dy \\
&\le C |g|_{r,\infty} \int_{\mathbb{R}^n} \frac{1}{\left(1+|y-x|\right)^{n+\varepsilon}} \, dy \\
&\le C |g|_{r,\infty} .
\end{aligned}
$$

Let $h > 0$. Then for $\tilde{g}(x) = g(hx)$

$$
\begin{aligned}
\left\| g(hx) - \int_{\mathbb{R}^n} K(x,y) g(hy)\, dy \right\|_\infty &= \left\| \tilde{g}(x) - \int_{\mathbb{R}^n} K(x,y) \tilde{g}(y)\, dy \right\|_\infty \\
&\le C |\tilde{g}|_{r,\infty} = C \left| g(h\cdot) \right|_{r,\infty} = C h^r |g|_{r,\infty}
\end{aligned}
$$

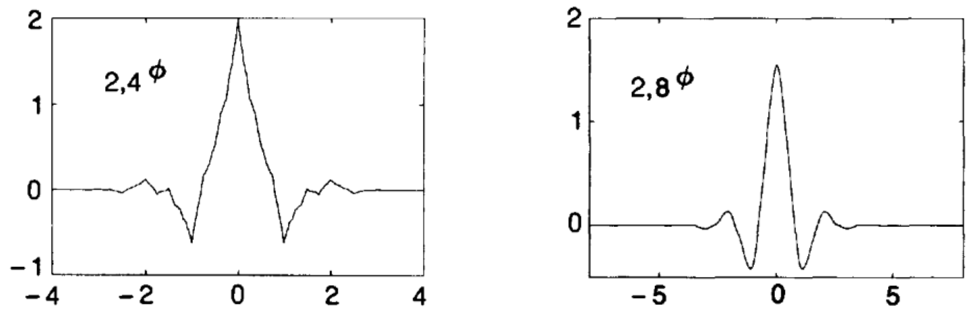
Therefore

$$
\begin{aligned}
\left\| g(x) - \int_{\mathbb{R}^n} K_h(x,y) g(y)\, dy \right\|_\infty &= \left\| g(x) - h^{-n} \int_{\mathbb{R}^n} K\left(h^{-1}x, h^{-1}y\right) g(y)\, dy \right\|_\infty \\
&= \left\| g(x) - \int_{\mathbb{R}^n} K\left(h^{-1}x, z\right) g(hz)\, dz \right\|_\infty \\
&= \left\| g(hx) - \int_{\mathbb{R}^n} K(x,z) g(hz)\, dz \right\|_\infty \\
&\le C h^r |g|_{r,\infty} .
\end{aligned}
$$

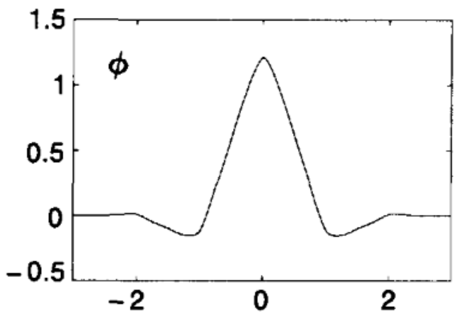
□

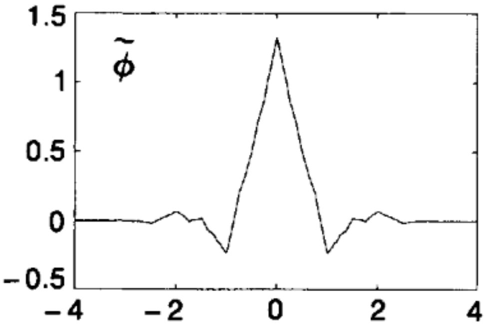
More examples:

- Duals to N_2



- Daubechies' famous [9,7] kernel for $r = 4$ (used in JPEG2000)





Multivariate Fourier series

$\mathbb{T}^n := [-\pi, \pi]^n$, $L_2(\mathbb{T}^n)$ is equipped with the dot-product

$$\langle f, g \rangle := \frac{1}{(2\pi)^n} \int_{\mathbb{T}^n} f(x)\bar{g}(x)\,dx .$$

$\{e^{ikx}\}_{k\in\mathbb{Z}^n}$ ortho-basis of $L_2(\mathbb{T}^n)$, $e^{ikx} := e^{i\sum_{j=1}^n k_j x_j}$, the orthonormality boils down to the univariate case,

$$\frac{1}{(2\pi)^n} \int_{\mathbb{T}^n} e^{ikx} e^{-ijx} dx = \prod_{m=1}^n \frac{1}{2\pi} \int_{\mathbb{T}} e^{ik_m x_m} e^{-ij_m x_m} dx_m = \prod_{m=1}^n \delta_{k_m, j_m} = \delta_{k,j} .$$

Fourier series

$$f(x) \underset{L_2}{=} \sum_{k\in\mathbb{Z}^n} \hat{f}(k) e^{ikx} , \qquad \hat{f}(k) := \frac{1}{(2\pi)^n} \int_{\mathbb{T}^n} f(x) e^{-ikx} dx .$$

The partial sum

$$S_N f(x) = \sum_{\|k\|_\infty \le N} \hat{f}(k) e^{ikx} , \qquad \|k\|_\infty = \max_{1\le m\le n} |k_m| .$$

Approximation from SI spaces in L_2

The sinc $\phi(x) = \phi(x_1, \ldots, x_n) = \prod_{i=1}^n \frac{\sin(\pi x_i)}{\pi x_i}$,

$$\hat{\phi}(w) = \hat{\phi}(w_1, \ldots, w_n) = 1_{[-\pi,\pi]^n}(w) = \prod_{i=1}^n 1_{[-\pi,\pi]}(w_i) .$$

Theorem The shifts $\{\phi(\cdot - k)\}_{k\in\mathbb{Z}^n}$ are an ortho-basis for $S(\phi)$.

Proof Observe that $\{\phi(\cdot - k)\}_{k=-\infty}^{\infty}$ is an orthonormal basis of its span $\Leftrightarrow \langle \phi(\cdot - k), \phi(\cdot - j) \rangle = \delta_{k,j}$, $\forall k, j \in \mathbb{Z}^n \Leftrightarrow$ $\langle \phi, \phi(\cdot + j) \rangle = \delta_{0,j}$, $\forall j \in \mathbb{Z}^n$. We now compute using Parseval

$$\begin{aligned}\langle \phi, \phi(\cdot + j)\rangle &= (2\pi)^{-n}\int_{\mathbb{R}^n}\hat{\phi}(w)\overline{(\phi(\cdot+j))^{\wedge}(w)}dw \\ &= (2\pi)^{-n}\int_{\mathbb{R}^n}\left|\hat{\phi}(w)\right|^2 e^{-ijw}dw \\ &= (2\pi)^{-n}\int_{[-\pi,\pi]^n} e^{-ijw}dw \\ &= \delta_{0,j}.\end{aligned}$$

□

Theorem Let $P_{S(\phi)^h}$ be the orthogonal projector onto $S(\phi)^h$, where ϕ is the sinc function. Then, for $f \in L_2(\mathbb{R}^n)$

$$\left(P_{S(\phi)^h} f\right)^{\wedge} = \hat{f}(w)1_{[-h^{-1}\pi,h^{-1}\pi]^n}(w), \quad h > 0.$$

Proof Since $\langle \phi(\cdot - k), \phi(\cdot - j)\rangle = \delta_{k,j}$, we have that $\phi_{h,k}(x) := h^{-n/2}\phi(h^{-1}x - k)$ satisfy $\langle \phi_{h,k}, \phi_{h,j}\rangle = \delta_{k,j}$. Thus, $\{\phi_{h,k}\}$ is an ortho-basis of $S(\phi)^h$.

$$\begin{aligned}\left(P_{S(\phi)^h} f\right)^{\wedge}(w) &= \left(\sum_k \langle f, \phi_{h,k}\rangle \phi_{h,k}\right)^{\wedge}(w) \\ &= h^{n/2}\hat{\phi}(hw)\sum_k \langle f, \phi_{h,k}\rangle e^{-ihkw} \\ &= \mathbf{1}_{[-h^{-1}\pi,h^{-1}\pi]^n}(w)\sum_k \frac{h^{n/2}}{(2\pi)^n}\left\langle \hat{f}, \hat{\phi}_{h,k}\right\rangle_{L^2(\mathbb{R}^n)} e^{-ihkw} \\ &= \mathbf{1}_{[-h^{-1}\pi,h^{-1}\pi]^n}(w)\underbrace{\sum_k \left(\frac{1}{(2h^{-1}\pi)^n}\int_{[-h^{-1}\pi,h^{-1}\pi]^n}\hat{f}(z)e^{ihkz}dz\right)e^{-ihkw}}_{=\hat{f}(w),\ \ w\in[-h^{-1}\pi,h^{-1}\pi]^n}\end{aligned}$$

Why? Observe that $\{e^{ihkw}\}_{k\in\mathbb{Z}^n}$ is an orthonormal basis of $L_2[-h^{-1}\pi, h^{-1}\pi]^n$ using the normalized dot-product

$$\langle h, g\rangle_h := \frac{1}{(2h^{-1}\pi)^n}\int_{[-h^{-1}\pi,h^{-1}\pi]^n} h(w)\overline{g(w)}dw.$$

Therefore, the restriction of $\hat{f}$ to $[-h^{-1}\pi, h^{-1}\pi]^n$ is represented by the Fourier series of $\{e^{ihkw}\}_{k\in\mathbb{Z}^n}$ and so, by the above computation

$$\left(P_{S(\phi)^h} f\right)^{\wedge} = \hat{f}(w)1_{[-h^{-1}\pi,h^{-1}\pi]^n}(w).$$

□

Therefore, the restriction of $\hat{f}$ to $[-h^{-1}\pi, h^{-1}\pi]^n$ is represented by the Fourier series of $\{e^{ihkw}\}_{k\in\mathbb{Z}^n}$ and so, by the above computation

$$\left(P_{S(\phi)^h} f\right)^{\wedge} = \hat{f}(w)1_{[-h^{-1}\pi,h^{-1}\pi]^n}(w).$$

□

Theorem The sinc has 'infinite' / spectral approximation order, i.e., $\forall r \ge 1$, $\forall f \in W_2^r(\mathbb{R}^n)$,

$$E\left(f, S(\phi)^h\right)_2 \le C(n,r) h^r |f|_{r,2}.$$

Remark Compare with the periodic case we proved for the Fourier series $\forall r \ge 1$, $\forall f \in W_2^r(\mathbb{T})$

$$\|f - S_N f\|_2 \le N^{-r} |f|_{r,2}.$$

Proof Let $f \in \mathcal{S}$. This means $f \in W_2^r(\mathbb{R}^n)$, $\forall r \ge 1$. We claim that

$$\frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \left|\hat{f}(w)\right|^2 |w|^{2r} dw \le C(n,r) |f|_{r,2}^2 = C(n,r) \left(\sum_{|\alpha|=r} \|D^\alpha f\|_2 \right)^2$$

Let's start with $n = 1$. In this case

$$\left(f^{(r)}\right)^\wedge (w) = (iw)^r \hat{f}(w) \Rightarrow \left|\left(f^{(r)}\right)^\wedge (w)\right| = |w|^r \left|\hat{f}(w)\right|.$$

So, by Parseval

$$|f|_{r,2}^2 = \left\|f^{(r)}\right\|_2^2 = \frac{1}{2\pi} \int_{\mathbb{R}} \left|\left(f^{(r)}\right)^\wedge (w)\right|^2 dw = \frac{1}{2\pi} \int_{\mathbb{R}} \left|\hat{f}(w)\right|^2 |w|^{2r} dw.$$

For $n \ge 2$, repeated application, coordinate by coordinate, each step similar to the univariate case, gives

$$\left(D^\alpha f\right)^\wedge (w) = (iw)^\alpha \hat{f}(w) \Rightarrow \left|\left(D^\alpha f\right)^\wedge (w)\right| = \left|w^\alpha\right| \left|\hat{f}(w)\right|.$$

Example for $n = 2$, $\alpha = (1,1)$

$$\begin{aligned}
\left(\frac{\partial}{\partial x_1}\frac{\partial}{\partial x_2} f\right)^\wedge (w) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\partial}{\partial x_1}\frac{\partial}{\partial x_2} f(x) e^{-iw_1x_1} e^{-iw_2x_2} dx_1 dx_2 \\
&= \int_{-\infty}^{\infty} e^{-iw_2x_2} \left(\int_{-\infty}^{\infty} \frac{\partial}{\partial x_1} \left(\frac{\partial}{\partial x_2} f(x) \right) e^{-iw_1x_1} dx_1 \right) dx_2 \\
&= iw_1 \int_{-\infty}^{\infty} e^{-iw_2x_2} \left(\int_{-\infty}^{\infty} \frac{\partial}{\partial x_2} f(x) e^{-iw_1x_1} dx_1 \right) dx_2 \\
&= iw_1 \int_{-\infty}^{\infty} e^{-iw_1x_1} \left(\int_{-\infty}^{\infty} \frac{\partial}{\partial x_2} f(x) e^{-iw_2x_2} dx_2 \right) dx_1 \\
&= -w_1 w_2 \int_{-\infty}^{\infty} e^{-iw_1x_1} \left(\int_{-\infty}^{\infty} f(x) e^{-iw_2x_2} dx_2 \right) dx_1 \\
&= -w_1 w_2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x) e^{-iw_1x_1} e^{-iw_2x_2} dx_1 dx_2 \\
&= -w_1 w_2 \hat{f}(w).
\end{aligned}$$

This gives

$$\|D^\alpha f\|_2^2 = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \left|\left(D^\alpha f\right)^\wedge (w)\right|^2 dw = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \left|\hat{f}(w)\right|^2 \left|w^\alpha\right|^2 dw.$$

Now, for $n \ge 2$

$$|w|^{2r} = \left(\sum_{m=1}^{n} w_m^2\right)^r = \left(w_1^r\right)^2 + n\left(w_1^{r-1}w_2\right)^2 + \cdots + \left(w_n^r\right)^2 = \sum_{|\alpha|=r} a_\alpha \left(w^\alpha\right)^2 .$$

Thus, we obtain

$$\begin{aligned}
\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\left|\hat{f}(w)\right|^2 |w|^{2r}\,dw &= \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\left|\hat{f}(w)\right|^2 \sum_{|\alpha|=r} a_\alpha \left|w^\alpha\right|^2 dw \\
&\le C(n,r)\sum_{|\alpha|=r}\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\left|\hat{f}(w)\right|^2 \left|w^\alpha\right|^2 dw \\
&= C(n,r)\sum_{|\alpha|=r}\left\|D^\alpha f\right\|_2^2 \\
&\underset{\|\cdot\|_{l_2}\le\|\cdot\|_{l_1}}{\le} C(n,r)\left(\sum_{|\alpha|=r}\left\|D^\alpha f\right\|_2\right)^2 \\
&= C(n,r)\left|f\right|_{r,2}^2 .
\end{aligned}$$

Now, for $h>0$, $w\notin[-\pi/h,\pi/h]^n \Rightarrow h|w|\ge\pi$. Therefore

$$\begin{aligned}
E\left(f,S(\phi)^h\right)_2^2 = \left\|f - P_{S(\phi)^h} f\right\|_2^2 &= \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n\setminus[-\pi/h,\pi/h]^n}\left|\hat{f}(w)\right|^2 dw \\
&\le \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n\setminus[-\pi/h,\pi/h]^n}|hw|^{2r}\left|\hat{f}(w)\right|^2 dw \\
&\le h^{2r}\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}|w|^{2r}\left|\hat{f}(w)\right|^2 dw \\
&\le C(n,r)h^{2r}\left|f\right|_{r,2}^2 .
\end{aligned}$$

The general case of $f\in W_2^r(\mathbb{R}^n)$ is obtained by a density argument $\{f_k\}_{k=1}^\infty$, $f_k\in\mathcal{S}$, $\|f-f_k\|_{W_2^r}\underset{k\to\infty}{\to}0$.

□

Refinable Shift Invariant Spaces

Def A refinable SI space satisfies $S(\Phi)\subset S(\Phi)^{1/2}$. In the PSI case, $S(\phi)\subset S(\phi)^{1/2}$ is equivalent to the existence of a ***two-scale relation***, i.e., the existence of a set of coefficients $\{p_k\}$ such that

$$\phi(x) = \sum_k p_k \phi(2x-k).$$

A refinable SI space provides a ***multiresolution analysis***:

$$\cdots\subset S(\Phi)^2\subset S(\Phi)\subset S(\Phi)^{1/2}\subset S(\Phi)^{1/4}\subset\cdots$$

Under certain conditions (such as compact support and reproduction of constants) we have for $1\le p\le\infty$,

$$\overline{\bigcup_j S(\Phi)^{2^{-j}}} = L_p(\mathbb{R}^n).$$

In some cases, we also require

$$\bigcap_j S(\Phi)^{2^{-j}} = 0 .$$

Theorem For $r \geq 1$, the univariate B-spline satisfies the two-scale relation

$$N_r(x) = \sum_{k=0}^{r} 2^{1-r} \binom{r}{k} N_r(2x-k) .$$

Proof Recall that $N_r = \underbrace{N_1 * \cdots * N_1}_{r \text{ times}}$ and therefore

$$\hat{N}_r(w) = \left(\frac{1-e^{-iw}}{iw} \right)^r .$$

Assume that there exist $\{p_{r,k}\}_{k=0}^{r}$, such that

$$N_r(x) = \sum_{k=0}^{r} p_{r,k} N_r(2x-k) .$$

This implies that

$$\begin{aligned}
&\hat{N}_r(w) = \sum_{k=0}^{r} p_{r,k} \left(N_r(2\cdot - k) \right)^{\wedge}(w) \\
&\Leftrightarrow \left(\frac{1-e^{-iw}}{iw} \right)^r = \frac{1}{2} \left(\frac{1-e^{-i(w/2)}}{i(w/2)} \right)^r \sum_{k=0}^{r} p_{r,k} e^{-ikw/2} \\
&\Leftrightarrow \left(\left(1+e^{-i(w/2)}\right)\left(1-e^{-i(w/2)}\right) \right)^r = 2^{r-1} \left(1-e^{-i(w/2)}\right)^r \sum_{k=0}^{r} p_{r,k} e^{-ikw/2} \\
&\Leftrightarrow 2^{1-r} \left(1+e^{-i(w/2)}\right)^r = \sum_{k=0}^{r} p_{r,k} e^{-ikw/2} \\
&\Leftrightarrow p_{r,k} = 2^{1-r} \binom{r}{k}, \quad 0 \leq k \leq r.
\end{aligned}$$

□

Application: High quality image resize (→ smaller)

Goal: Create a ‘faithful’, high quality low resolution digital copy of $\{f(k)\}$. The low resolution is determined by a scale $h > 1$.

We use the sinc ϕ as a platform for an ‘ideal low pass’.

Let $\{f(k)\}$ be input samples. We assume they are samples of a band-limited function $f \in L_2(\mathbb{R}^n)$, $\text{supp}(\hat{f}) \subseteq \mathbb{T}^n$.
Therefore

$$f(x) = \sum_{k \in \mathbb{Z}^n} f(k) \phi(x-k) .$$

This is because

$$\left\langle f, \phi(\cdot - k) \right\rangle = (2\pi)^{-n} \left\langle \hat{f}, \phi(\cdot - k)^{\wedge} \right\rangle$$
$$= \frac{1}{(2\pi)^n} \int_{[-\pi,\pi]^n} \hat{f}(w) e^{ikw} dw \overset{\text{inverse Fourier}}{=} f(k).$$

Algorithm: Project f onto $S(\phi)^h$ and then subsample the projection at the rate h. On the Fourier side, the projection is simply $\hat{f}(w) 1_{[-\pi/h,\pi/h]}(w)$.

Digital implementation I ('low' dimensional signals):

1. Apply a discrete Fourier transform on the signal f of dimension M^n.
2. Leave the $(h^{-1}M)^n$ lowest frequency Fourier coefficients and set the rest to zero to obtain $\hat{f}_h$.
3. Apply an inverse Fourier transform $\hat{f}_h \underset{\mathcal{F}^{-1}}{\Rightarrow} f_h$.
4. Subsample f_h in each coordinate direction at the rate h to obtain a lower resolution version of f.

Note: In step 3, one can apply the appropriate inverse Fourier transform only on the lower resolution coefficients and after scaling obtain directly the lower resolution.

Resizing large images

For large images, applying the Fourier transform on the whole image is time consuming. We would like to apply 'local' filtering prior to the subsampling and obtain results which are a good approximation to the ideal low-pass with equivalent visual quality. To this end, we go back to the 'functional model'. Recall that $\hat{f}(w) = \sum_k f(k) e^{-ikw}$, $w \in [-\pi, \pi]^n$. We now compute the Fourier series of the 'filter':

$$1_{[-\pi/h,\pi/h]}(w) = \sum_k \alpha_h(k) e^{-ikw}, \qquad h > 1.$$

$$\alpha_h(k) = \frac{1}{(2\pi)^n} \int_{[-\pi/h,\pi/h]} e^{ikx} dx = \prod_{j=1}^n \frac{1}{2\pi} \int_{-\pi/h}^{\pi/h} e^{ik_j x_j} dx_j = \prod_{j=1}^n \frac{\sin(\pi k_j / h)}{\pi k_j}.$$

Therefore, using also the notation for the discrete sequences $f^d := \{ f(k) \}_{k \in \mathbb{Z}^n}$, $\alpha_h^d := \{ \alpha_h(k) \}_{k \in \mathbb{Z}^n}$,

$$\begin{aligned}
\hat{f}(w) 1_{[-\pi/h,\pi/h]}(w) &= \sum_k f(k) e^{-ikw} \sum_j \alpha_h(j) e^{-ijw} \\
&= \sum_{k,j} f(k) \alpha_h(j) e^{-i(k+j)w} \\
&= \sum_l \left(\sum_k f(k) \alpha_h(l-k) \right) e^{-ilw} \\
&= \sum_l \left(f^d * \alpha_h^d \right)(l) e^{-ilw}.
\end{aligned}$$

When we translate this back to the ‘time’ domain, we obtain that the ideal low-pass can be achieved through the convolution of the filter α_h^d with the samples f^d. In practice (MatLab, PhotoShop, etc.) we apply a finite filter that approximates the action of α_h^d. After the convolution we may safely subsample the values.

Wavelets: Efficient adaptive/sparse/nonlinear approximation

Motivation

Recall approximation using non-uniform piecewise constants

$$\Sigma_N := \left\{ \sum_{j=0}^{N-1} c_j \mathbf{1}_{[t_j, t_{j+1})} : T = \{t_j\},\ 0 = t_0 < t_1 < \cdots < t_N = 1 \right\}, \quad \sigma_N(f)_p := \inf_{g \in \Sigma_N} \|f - g\|_p .$$

$$Var(f) := \sup_T \left\{ \sum \left| f(t_{j+1}) - f(t_j) \right| \right\} .$$

To approximate a function of bounded variation in $[0,1]$, in the maximum norm from Σ_N, we created a partition where $Var_{[t_j, t_{j+1}]}(f) \le \frac{Var(f)}{N}$. We then selected a_j as the median value in $[t_j, t_{j+1}]$, to obtain

$$\left| f(x) - a_j \right| \le \frac{Var_{[t_j, t_{j+1}]}(f)}{2} \le \frac{Var(f)}{2N}, \ \forall x \in [t_j, t_{j+1}] .$$

Questions:

1. Can one implement this in practice? In higher dimensions? Not really. One then resorts to the greedy algorithms of adaptive binary partition (aka decision trees, Random Forest in machine learning)
2. Higher degree polynomial pieces?
3. What about approximation in L_2?

Wavelet construction starts with a multiresolution

Let $\phi \in L_2(\mathbb{R}^n)$, $\|\phi\|_2 = 1$, with

$$V_0 := S(\phi) \subset S(\phi)^{1/2} =: V_1$$

$$V_j := S(\phi)^{2^{-j}} = \overline{span}\{\phi_{j,k}\},\ \phi_{j,k}(x) := 2^{jn/2}\phi(2^j x - k),\ j \in \mathbb{Z},\ k \in \mathbb{Z}^n .$$

$$\cdots \subset V_{-(j+1)} \subset V_{-j} \subset \cdots \subset V_0 \subset V_1 \subset \cdots \subset V_j \subset V_{j+1} \subset \cdots$$

Remarks

1. In signal processing books, the notation is ‘reversed’ $\phi_{j,k}(x) := 2^{-jn/2}\phi(2^{-j}x - k)$ and $V_j \subset V_{j-1}$ because the discrete wavelet transform takes signals in V_0 and decomposes into lower resolutions.
2. $\|\phi_{j,k}\|_2 = 1,\ j \in \mathbb{Z},\ k \in \mathbb{Z}^n$.

Haar wavelets – Introduction using the discrete wavelet transform

$$\phi = N_1\,,\ \phi_{j,k} := 2^{jn/2}\phi\left(2^j \cdot - k\right)\,,\ \left\|\phi_{j,k}\right\|_2 = 1\,,\ \forall j,k \in \mathbb{Z}\,.$$

Suppose we are given a discrete univariate signal $f^0 = \left\{ f(k) \right\}_{k\in\mathbb{Z}}$ which we want to compress. In practice one typically gets a finite number of samples, e.g. $f^0 = \left\{ f(k) \right\}_{k=0}^{1023}$.

We assume the 'approximation'

$$f(k) \approx \left\langle f, N_1(\cdot - k) \right\rangle = \left\langle f, \phi(\cdot - k) \right\rangle =: f_k^0\,.$$

Compression by 'averaging'? $f_k^{-1} = \frac{1}{\sqrt{2}}\left(f_{2k}^0 + f_{2k+1}^0 \right)$. But what did we really do? We projected onto $S(\phi)^{-1}$

$$f_k^{-1} = \frac{1}{\sqrt{2}}\left(f_{2k}^0 + f_{2k+1}^0 \right) = \left\langle f, \frac{1}{\sqrt{2}} N_1\left(2^{-1} \cdot - k\right) \right\rangle = \left\langle f, \phi_{-1,k} \right\rangle.$$

But we 'lost' data. This can be recovered by using the **Haar Wavelet**

$$\psi(x) := \begin{cases} 1, & 0 \le x < 1/2, \\ -1, & 1/2 \le x \le 1, \\ 0, & else. \end{cases}$$

and computing

$$\alpha_k^{-1} = \frac{1}{\sqrt{2}}\left(f_{2k}^0 - f_{2k+1}^0 \right) = \left\langle f, \frac{1}{\sqrt{2}} \psi\left(2^{-1} \cdot - k\right) \right\rangle = \left\langle f, \psi_{-1,k} \right\rangle.$$

So, with $W_j := S(\psi)^{2^{-j}} = \overline{\text{span}}\left\{\psi_{j,k}\right\}, \psi_{j,k}(x) := 2^{j/2}\psi\left(2^j x - k\right)$, we shall see that we have

$$V_0 = V_{-1} \oplus W_{-1}\,.$$

The Crux - Sparsity The wavelet coefficients $\left\{\alpha_k^{-1}\right\}$ serve as the 'difference' between the two scales. In many applications a large number of the coefficients $\left\{\alpha_k^{-1}\right\}$ are 'insignificant', e.g with absolute value below some threshold.

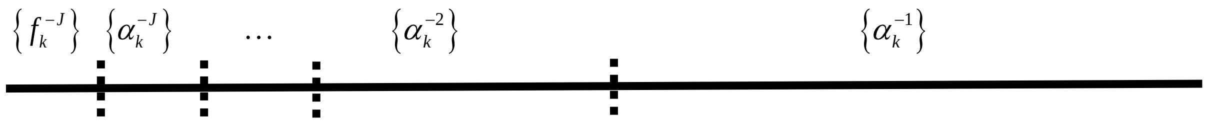
We continue recursively with the decomposition algorithm and now compute

$$f_k^{-2} = \frac{1}{\sqrt{2}}\left(f_{2k}^{-1} + f_{2k+1}^{-1} \right) = \left\langle f, \frac{1}{2} N_1\left(2^{-2} \cdot - k\right) \right\rangle = \left\langle f, \phi_{-2,k} \right\rangle.$$

So, we computed a projection onto $V_{-2} = S(\phi)^{-2}$. We also compute wavelet coefficients

$$\alpha_k^{-2} = \frac{1}{\sqrt{2}}\left(f_{2k}^{-1} - f_{2k+1}^{-1} \right) = \left\langle f, \frac{1}{2}\psi\left(2^{-2}\cdot - k\right) \right\rangle = \left\langle f, \psi_{-2,k} \right\rangle.$$

After J steps, we get the discrete wavelet decomposition



Lemma

$$W_j \oplus V_j = V_{j+1},\ \forall j \in \mathbb{Z}.$$

Proof

Observe that $W_0 \subset V_1$. Indeed $\psi = \frac{1}{\sqrt{2}}\left(\phi_{1,0} - \phi_{1,1}\right)$ and similarly to the assignment $W_0 = S(\psi) \subset S(\phi)^{1/2} = V_1$. In this particular case it is easy to see

$$\psi_{0,k} = \frac{1}{\sqrt{2}}\left(\phi_{1,2k} - \phi_{1,2k+1}\right).$$

Next we show that $V_0 \perp W_0$. It is easy to see that $\left\langle \phi_{0,k}, \psi_{0,j} \right\rangle = 0$ for $k \neq j$, simply because their support does not overlap. For $k = j$, we have $\left\langle \phi_{0,k}, \psi_{0,k} \right\rangle = \left\langle \phi_{0,0}, \psi_{0,0} \right\rangle = 1/2 - 1/2 = 0$. This gives that $V_0 \perp W_0$.

To show that the sum of the spaces gives V_1, it is sufficient to see that

$$\phi_{1,0} = \frac{1}{\sqrt{2}}\left(\phi_{0,0} + \psi_{0,0}\right),\ \phi_{1,1} = \frac{1}{\sqrt{2}}\left(\phi_{0,0} - \psi_{0,0}\right).$$

This gives

$$\phi_{1,2k} = \phi_{1,0}(x-k) = \frac{1}{\sqrt{2}}\left(\phi_{0,k} + \psi_{0,k}\right).$$

and

$$\phi_{1,2k+1} = \phi_{1,1}(x-k) = \frac{1}{\sqrt{2}}\left(\phi_{0,k} - \psi_{0,k}\right).$$

Therefore, we may conclude that $W_0 \oplus V_0 = V_1$. By scaling, this gives

$$W_j \oplus V_j = V_{j+1},\ \forall j \in \mathbb{Z}.$$

□

Theorem The Haar wavelets are an orthonormal basis of $L_2(\mathbb{R})$.

Proof

(i) $\overline{\bigcup_j V_j} = L_p(\mathbb{R})$. How do we see that? Several ways. If $f \in L_p(\mathbb{R})$, $1 \le p \le \infty$, then $\omega_r(f,t)_p \underset{t\to 0}{\to} 0$. To prove that observe that for any $\varepsilon > 0$, one can construct $g \in W_p^1(\mathbb{R})$ through a convolution of a Gaussian with f, such that $\|f-g\|_p < \varepsilon$. Then,

$$\begin{aligned}\omega_r(f,t)_p &\le C\omega_1(f,t)_p \\ &\le CK_1(f,t)_p \\ &\le C\left(\|f-g\|_p + t|g|_{1,p}\right) \\ &\le C\left(\varepsilon + t|g|_{1,p}\right) \underset{t\to 0}{\to} C\varepsilon.\end{aligned}$$

This implies that

$$E(f,V_j)_p := E\left(f, S(\phi)^{2^{-j}}\right)_p \le C\omega_1\left(f,2^{-j}\right)_p \underset{j\to\infty}{\to} 0.$$

Now, we use the fact that for any $J \in \mathbb{Z}$,

$$V_J = V_{J-1} \oplus W_{J-1} = V_{J-2} \oplus W_{J-2} \oplus W_{J-1} = \cdots = \bigoplus_{j=-\infty}^{J-1} W_j.$$

And so,

$$\bigoplus_{j=-\infty}^{\infty} W_j = \overline{\bigcup_j V_j} = L_2.$$

(ii) Orthonormality of $\{\psi_{j,k}\}$, $\psi_{j,k}(x) := 2^{j/2}\psi\left(2^j x - k\right)$: Let $\psi_{j_1,k_1}, \psi_{j_2,k_2}$ be two Haar wavelets. If $j_1 = j_2$, then simply by consideration the intersections of supports, $\operatorname{supp}\left(\psi_{j,k}\right) = \left[2^{-j}k, 2^{-j}(k+1)\right]$, we have that $\left\langle \psi_{j_1,k_1}, \psi_{j_1,k_2}\right\rangle = \delta_{k_1,k_2}$. If $j_1 \ne j_2$, then w.l.g. $j_1 < j_2$. Then, we claim that $\psi_{j_1,k_1} \equiv \text{const}$ on the support of ψ_{j_2,k_2}. Indeed, if the supports do not intersect, then $\psi_{j_1,k_1} \equiv 0$, on the support of ψ_{j_2,k_2}. Else, there are two possibilities: $\psi_{j_1,k_1} \equiv 2^{j_1/2}$ or $\psi_{j_1,k_1} \equiv 2^{-j_1/2}$, on the support of ψ_{j_2,k_2}. In any case,

$$\left\langle \psi_{j_1,k_1}, \psi_{j_2,k_2}\right\rangle = c\int_{\mathbb{R}} \psi_{j_2,k_2} = 0$$

□

There are many constructions of univariate orthonormal multiresolution and wavelet systems. Unlike the Haar wavelet, they do not have an analytic definition.

The cascade algorithm

Suppose ϕ is given by the two-scale equation $\phi = \sum_m p_m \phi_{1,m}$. Then,

$$\begin{gathered}\phi(x) = \sqrt{2}\sum_m p_m \phi(2x-m) \Rightarrow \\ \sqrt{2}\phi(2x-k) = 2\sum_m p_m\phi\left(2(2x-k)-m\right) = 2\sum_m p_m \phi\left(4x-(2k+m)\right) \underset{l=2k+m}{\Rightarrow} \\ \phi_{1,k} = \sum_l p_{l-2k}\phi_{2,l}\end{gathered}$$

And in general

$$\phi_{j,k} = \sum_l p_{l-2k}\phi_{j+1,l}$$

Therefore

$$\phi = \sum_k p_k \sum_l p_{l-2k}\phi_{2,l} = \sum_l \left(\sum_k p_k p_{l-2k} \right)\phi_{2,l} \ .$$

Recursively, this gives explicit coefficients for any $J \geq 1$,

$$\phi = \sum_l p_{J,l}\phi_{J,l} \ , \qquad \phi_{J,l}(x) = 2^{J/2}\phi(2^J x - l).$$

This gives a **subdivision** algorithm, that starts with an initial sequence $\{\delta_{0,k}\}_{k\in\mathbb{Z}}$ at the level 0, and at the level J has values $\{p_{J,l}\}$. If we plot the values $2^{J/2}p_{J,l}$ at the dyadic points $2^{-J}l$, we will see an 'approximation' to ϕ.

Vanishing Moments

The main idea is that we would like $S(\phi)$ to have higher approximation order r, by reproducing polynomials of order r (degree $r-1$). If ϕ has compact support and $\{\phi(\cdot - k)\}$ are an ortho-basis of $V_0 = S(\phi)$, then for any polynomial $P \in \Pi_{r-1}(\mathbb{R})$

$$P(x) = \sum_k \langle P, \phi(\cdot - k)\rangle \phi(x-k).$$

If ψ has compact support and $S(\psi) \perp S(\phi)$, this gives

$$\int_{\mathbb{R}} P(x)\psi(x-j)\,dx = \sum_k \langle P, \phi(\cdot - k)\rangle \underbrace{\langle \phi(\cdot - k), \psi(\cdot - j)\rangle}_{=0} = 0 \,, \qquad \forall j \in \mathbb{Z}.$$

So the wavelets will have r **vanishing moments**

$$\int_{\mathbb{R}} x^m \psi(x)\,dx = 0 \,, \qquad m = 0,\ldots,r-1,$$

and

$$\int_{\mathbb{R}} P(x)\psi_{j,k}(x)\,dx = 0 \,, \qquad \forall P \in \Pi_{r-1}(\mathbb{R}),\ \forall j,k \in \mathbb{Z}.$$

This provides in many cases better sparsity! We shall see that this is the efficient alternative to adaptive piecewise polynomial approximation of order r (degree $r-1$).

Discrete (orthonormal) wavelet transforms

Suppose $S(\phi) \subset S(\phi)^{1/2}$, where $\{\phi(\cdot - k)\}$ are an orthonormal basis of $S(\phi)$ and that $S(\phi) \oplus S(\psi) = S(\phi)^{1/2}$, with $\{\psi_{j,k} := 2^{j/2}\psi(2^j \cdot - k)\}$ an orthonormal basis for $L_2(\mathbb{R})$.

Therefore, there exist coefficients $\{p_k\},\{q_k\}$, such that

$$\phi = \sum_k p_k \phi_{1,k} \ , \ \psi = \sum_k q_k \phi_{1,k} \ .$$

We are given again, discrete samples $\left\{ f\left(k\right) \right\}_{k\in\mathbb{Z}}$ which we want to compress. We assume again that

$$f\left(k\right) \approx \left\langle f, \phi\left(\cdot - k\right) \right\rangle =: f_k^0 .$$

We want to compress by projecting onto V_{-1}. We use

$$\begin{aligned}
\phi\left(x\right) &= \sqrt{2}\sum_m p_m \phi\left(2x-m\right) \Rightarrow \\
\frac{1}{\sqrt{2}}\phi\left(2^{-1}x-k\right) &= \sum_m p_m \phi\left(2\left(2^{-1}x-k\right)-m\right) = \sum_m p_m \phi\left(x-\left(2k+m\right)\right) \underset{l=2k+m}{\Rightarrow} \\
\phi_{-1,k} &= \sum_l p_{l-2k}\phi_{0,l} \\
f_k^{-1} &= \left\langle f, \phi_{-1,k} \right\rangle = \left\langle f, \sum_l p_{l-2k}\phi_{0,l} \right\rangle = \sum_l p_{l-2k} \left\langle f, \phi_{0,l} \right\rangle = \sum_l p_{l-2k} f_l^0 .
\end{aligned}$$

In the same manner we compute the wavelet/difference coefficients

$$\alpha_k^{-1} = \left\langle f, \psi_{-1,k} \right\rangle = \left\langle f, \sum_l q_{l-2k}\phi_{0,l} \right\rangle = \sum_l q_{l-2k} \left\langle f, \phi_{0,l} \right\rangle = \sum_l q_{l-2k} f_l^0 .$$

We then continue to process $\left\{ f_k^{-1} \right\}$ to compute $\left\{ f_k^{-2} \right\}$ and $\left\{ \alpha_k^{-2} \right\}$… etc…

Multivariate wavelet bases via tensor-products

Assignment Assume ϕ^*, ψ^* generate a univariate orthonormal multiresolution and a wavelet system, respectively. Then, for $n=2$ the following is an orthonormal basis of $L_2\left(\mathbb{R}^2\right)$:

$$\left\{ \psi_{j,k}^e \right\}, \qquad \psi_{j,k}^e\left(x\right) := 2^j \psi^e\left(2^j x - k\right), \qquad e=1,2,3,\ \ j\in\mathbb{Z}, \quad k\in\mathbb{Z}^2,$$

where

$$\psi^1\left(x_1,x_2\right) := \phi^*\left(x_1\right)\psi^*\left(x_2\right), \quad \psi^2\left(x_1,x_2\right) := \psi^*\left(x_1\right)\phi^*\left(x_2\right), \quad \psi^3\left(x_1,x_2\right) := \psi^*\left(x_1\right)\psi^*\left(x_2\right).$$

We can abbreviate notation using the notation $I := \left(j,k,e\right)$

$$f \underset{L_2}{=} \sum_I \left\langle f, \psi_I \right\rangle \psi_I .$$

Biorthogonal wavelets

Orthonormality of wavelet bases is a limitation. For example, one cannot construct symmetric wavelets/filters.

Def Riesz Basis $\left\{ \psi_I \right\}$ is a Riesz basis for $L_2\left(\mathbb{R}^n\right)$, if there exist $0 < A < B < \infty$, such that for any $\left\{ c_I \right\} \in l_2$

$$A\left\|\{c_I\}\right\|_{l_2}^2 \le \left\|\sum_I c_I \psi_I\right\|_{L_2}^2 \le B\left\|\{c_I\}\right\|_{l_2}^2$$

Notice this is a Parseval-type equivalence. One can construct a dual basis $\{\tilde{\psi}_I\}$ (not unique), such that:

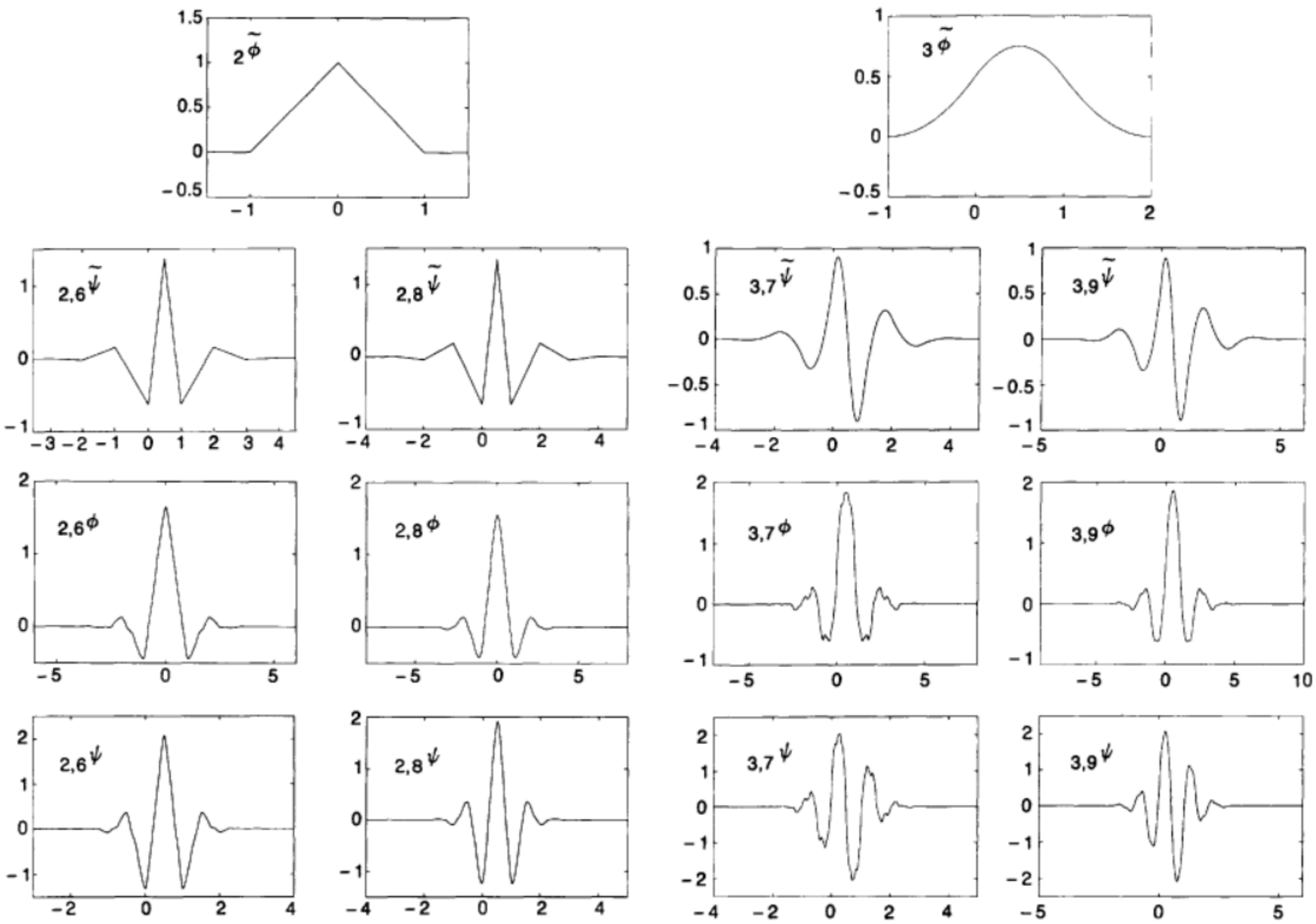
i. $\langle \tilde{\psi}_{\tilde{I}}, \psi_I \rangle = \delta_{\tilde{I},I}$, (biorthogonality),

ii. for $f \in L_2(\mathbb{R}^n)$,

$$f(x) \underset{L_2}{=} \sum_I \langle f, \tilde{\psi}_I \rangle \psi_I(x),$$

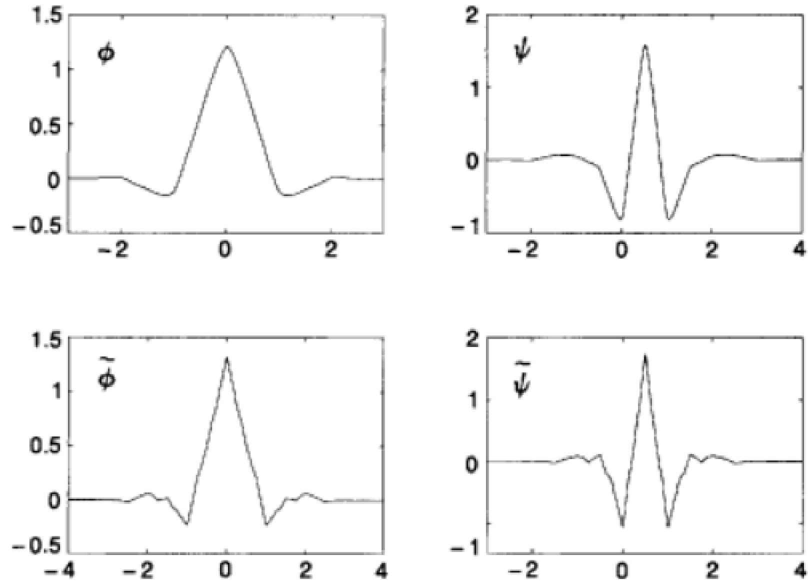
iii. $\|f\|_{L_2} \sim \left\|\{\langle f, \tilde{\psi}_I \rangle\}\right\|_{l_2}$.

This means 'stability' of the representation. The application is mostly in nonlinear approximation.



Example of biorthogonal duals of B-splines and corresponding dual wavelets

$N, \tilde{N}$	n	coefficient of $e^{-in\xi}$ in m_0	coefficient of $e^{-in\xi}$ in $\tilde{m}_0$
$N = 4$	0	.557543526229	.602949018236
$\tilde{N} = 4$	1, −1	.295635881557	.266864118443
	2, −2	−.028771763114	−.078223266529
	3, −3	−.045635881557	−.016864118443
	4, −4	0	.026748757411



The 'famous' 9/7 scaling function and wavelet duals

Nonlinear approximation with wavelets

We define the nonlinear 'manifold' of all N-term wavelets

$$\Sigma_N := \left\{ \sum_{\# I \le N} c_I \psi_I \right\}.$$

The degree of nonlinear/adaptive approximation with wavelets

$$\sigma_N(f)_p := \inf_{\varphi \in \Sigma_N} \| f - \varphi \|_p .$$

How to choose a good approximant from Σ_N ? We use the **'greedy' algorithm** and reorder the wavelet coefficients by

$$\left| \left\langle f, \tilde{\psi}_{I_1} \right\rangle \right| \ge \left| \left\langle f, \tilde{\psi}_{I_2} \right\rangle \right| \ge \left| \left\langle f, \tilde{\psi}_{I_3} \right\rangle \right| \ge \cdots,$$

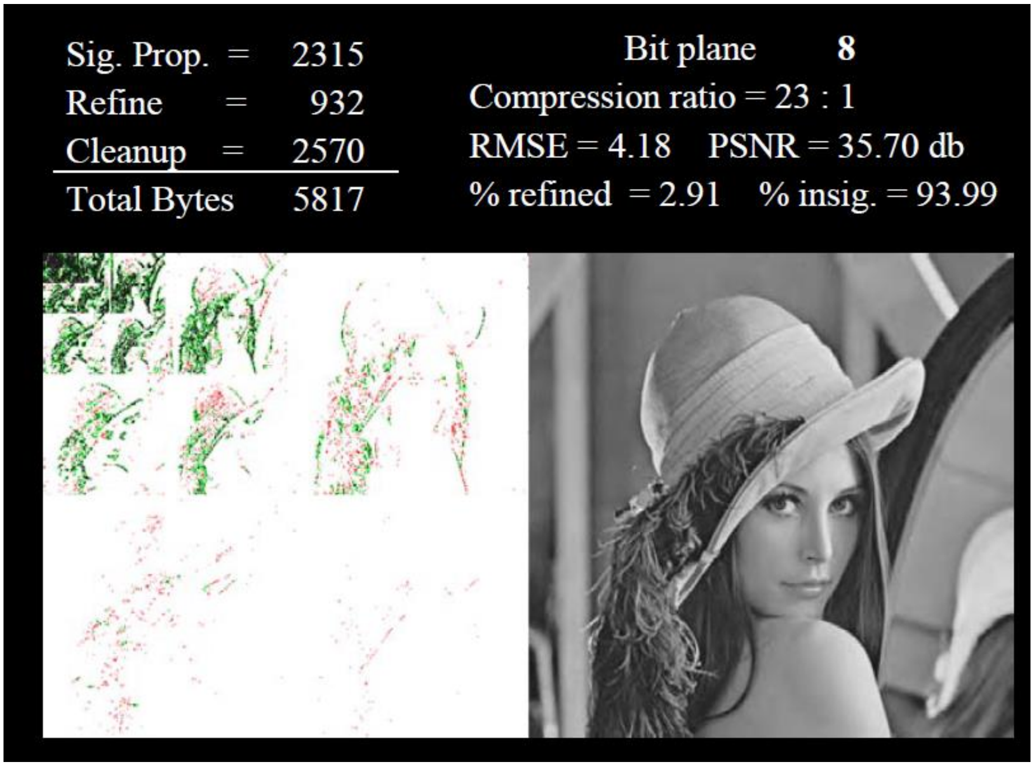
and create the N-term sum

$$f_N(x) := \sum_{j=1}^{N} \left\langle f, \tilde{\psi}_{I_j} \right\rangle \psi_{I_j} \in \Sigma_N \ .$$

For biorthogonal bases, the error can be estimated by

$$\| f - f_N \|_2^2 = \left\| \sum_{j=N+1}^{\infty} \left\langle f, \tilde{\psi}_{I_j} \right\rangle \psi_{I_j} \right\|_2^2 \sim \sum_{j=N+1}^{\infty} \left| \left\langle f, \tilde{\psi}_{I_j} \right\rangle \right|^2 .$$

This is our efficient, multivariate version of non-uniform piecewise polynomial approximation!



Jackson theorem for wavelets

We know that if $f \in B_\tau^\alpha(\mathbb{R}^n) = B_{\tau,\tau}^\alpha(\mathbb{R}^n)$, then $f \in L_\tau(\mathbb{R}^n)$. The next theorem says that the 'knowledge' of the additional smoothness of f gives more

Theorem For the parameters $1 \le p < \infty$, $\alpha > 0$ and

$$\frac{1}{\tau} = \frac{\alpha}{n} + \frac{1}{p},$$

we have the embedding

$$B_\tau^\alpha(\mathbb{R}^n) = B_{\tau,\tau}^\alpha(\mathbb{R}^n) \subset L_p(\mathbb{R}^n), \qquad f \in B_\tau^\alpha(\mathbb{R}^n) \Rightarrow f \in L_p(\mathbb{R}^n).$$

The following Jackson theorem is a special case that is sufficient for our purpose.

Theorem [Jackson theorem] Let $\{\psi_I\}, \{\tilde{\psi}_I\}$ be dual Riesz wavelet bases for $L_2(\mathbb{R}^n)$ where:

(i) $\operatorname{supp}(\psi^e), \operatorname{supp}(\tilde{\psi}^e) \subseteq [-M, M]^n$, $e \in E$.

(ii) $\psi^e, \tilde{\psi}^e \in W_\infty^r$, $e \in E$, $r > \alpha$,

(iii) $\psi^e, \tilde{\psi}^e$, $e \in E$, have $r > \alpha$ vanishing moments.

Let $p = 2$ and

$$\frac{1}{\tau} = \frac{\alpha}{n} + \frac{1}{2}.$$

Then, for $f \in B_\tau^\alpha\left(\mathbb{R}^n\right)$

$$\sigma_N\left(f\right)_2 \le \left\| f - f_N \right\|_2 \le CN^{-\alpha/n} \left| f \right|_{B_\tau^\alpha} .$$

We will prove most parts of the theory that leads to the Jackson theorem for wavelets.

Lemma 1 For the case of L_2-normalized wavelet $\psi_I\left(x\right) := 2^{jn/2}\psi^e\left(2^j x - k\right)$, $I = \left(e, j, k\right)$, we have that $\left|\operatorname{supp}\left(\psi_I\right)\right| = \left(2^{-j} 2M\right)^n \sim 2^{-jn}$ and

$$\left\|\psi_I\right\|_\infty = 2^{jn/2}\left\|\psi^e\right\|_\infty \le C\left|\operatorname{supp}\left(\psi_I\right)\right|^{-1/2} .$$

Lemma 2 Let $F\left(x\right) = \sum_{j=1}^J c_{I_j}\psi_{I_j}$, where $\left|c_{I_j}\right| \le L$. Then,

$$\left\|F\right\|_2 \le CLJ^{1/2}$$

Remark Notice the triangle inequality gives a weaker estimate

$$\left\|F\right\|_2 \le \sum_{j=1}^J \left\|c_{I_j}\psi_{I_j}\right\|_2 = \sum_{j=1}^J \left|c_{I_j}\right| \le LJ .$$

Proof We apply Lemma 1

$$\begin{aligned} \left\|F\right\|_2 &= \left\|\sum_{j=1}^J c_{I_j}\psi_{I_j}\right\|_2 \\ &\le \left\|\sum_{j=1}^J \left|c_{I_j}\psi_{I_j}\right|\right\|_2 \\ &\le \left\|\sum_{j=1}^J \left|c_{I_j}\right| \left\|\psi_{I_j}\right\|_\infty \mathbf{1}_{\operatorname{supp}\left(\psi_{I_j}\right)}\left(\cdot\right)\right\|_2 \\ &\le CL\left\|\sum_{j=1}^J \left|\operatorname{supp}\left(\psi_{I_j}\right)\right|^{-1/2} \mathbf{1}_{\operatorname{supp}\left(\psi_{I_j}\right)}\left(\cdot\right)\right\|_2 . \end{aligned}$$

Define

$$\Gamma\left(x\right) := \begin{cases} \min\limits_{1 \le j \le J}\left\{\left|\operatorname{supp}\left(\psi_{I_j}\right)\right| \,:\, x \in \operatorname{supp}\left(\psi_{I_j}\right)\right\}, & x \in \bigcup_{j=1}^J \operatorname{supp}\left(\psi_{I_j}\right), \\ 0, & \text{else.} \end{cases}$$

Any point $x \in \mathbb{R}^n$ can only be contained only in a finite number of supports of wavelets at the same level and as we go to lower levels, the size of the support grows by a factor of 2^n. This means that for any $x \in \mathbb{R}^n$,

$$\begin{aligned} \sum_{j=1}^J \left|\operatorname{supp}\left(\psi_{I_j}\right)\right|^{-1/2} \mathbf{1}_{\operatorname{supp}\left(\psi_{I_j}\right)}\left(x\right) &\le C\left(\Gamma\left(x\right)^{-1/2} + \left(2^n\Gamma\left(x\right)\right)^{-1/2} + \cdots\right) \\ &\le C\Gamma\left(x\right)^{-1/2} . \end{aligned}$$

So,

$$
\begin{aligned}
\|F\|_2 &\le CL\left\|\Gamma(\cdot)^{-1/2}\right\|_2 \\
&= CL\left(\int_{\bigcup_{j=1}^{J}\operatorname{supp}(\psi_{I_j})}\Gamma(x)^{-1}\,dx\right)^{1/2} \\
&\le CL\left(\sum_{j=1}^{J}\int_{\operatorname{supp}(\psi_{I_j})}\left|\operatorname{supp}(\psi_{I_j})\right|^{-1}dx\right)^{1/2} \\
&= CLJ^{1/2}.
\end{aligned}
$$

□

Theorem [Wavelet characterization of Besov spaces] Let $\{\psi_I\},\{\tilde{\psi}_I\}$ be dual Riesz wavelet bases with properties as in the Jackson theorem. Let $p=2$ and

$$\frac{1}{\tau}=\frac{\alpha}{n}+\frac{1}{2}.$$

For $f\in B_\tau^\alpha(\mathbb{R}^n)$

$$|f|_{B_\tau^\alpha}\sim\mathcal{N}_\tau(f):=\left(\sum_I\left|\langle f,\tilde{\psi}_I\rangle\right|^\tau\right)^{1/\tau}.$$

Remark Here is where the proof uses all of the properties of the wavelets – compact support, sufficient smoothness, vanishing moments

Weak l_τ space

Recall the 'strong' norm $\|\beta\|_{l_\tau}=\left(\sum_k|\beta_k|^\tau\right)^{1/\tau}$.

Def For a sequence $\beta=\{\beta_k\}$,

$$\|\beta\|_{wl_\tau}:=\sup_{\varepsilon>0}\#\{\beta_k:\ |\beta_k|>\varepsilon\}^{1/\tau}\varepsilon.$$

For any $\varepsilon>0$, $\#\{\beta_k:\ |\beta_k|>\varepsilon\}\varepsilon^\tau\le\sum_{k,|\beta_k|>\varepsilon}|\beta_k|^\tau\le\|\beta\|_{l_\tau}^\tau$. This implies that $\|\beta\|_{wl_\tau}\le\|\beta\|_{l_\tau}$ and $l_\tau\subset wl_\tau$.

Example Typical example for $\beta\in wl_1$, $\beta\notin l_1$ is: $\beta_k=1/k$, since

$$\#\left\{\beta_k:\ |\beta_k|>\frac{1}{N}\right\}N^{-1}=\frac{N-1}{N}\le 1\Rightarrow\|\beta\|_{wl_1}=1$$

Proof of Jackson theorem Let $f\in B_\tau^\alpha(\mathbb{R}^n)$. Then, by embedding $f\in L_2(\mathbb{R}^n)$. By the wavelet characterization $|f|_{B_\tau^\alpha}\sim\mathcal{N}_\tau(f):=\left(\sum_I\left|\langle f,\tilde{\psi}_I\rangle\right|^\tau\right)^{1/\tau}$. For $\nu=1,2,\ldots$, denote

$$\Lambda_\nu:=\left\{I:\ 2^{-\nu}\mathcal{N}_\tau(f)<\left|\langle f,\tilde{\psi}_I\rangle\right|\le 2^{-\nu+1}\mathcal{N}_\tau(f)\right\}.$$

We have, using the weak l_τ space

$$
\begin{aligned}
\#\Lambda_m &\le \sum_{\nu\le m}\#\Lambda_\nu \\
&= \#\bigcup_{\nu\le m}\Lambda_\nu \\
&= \#\left\{I:\ \ 2^{-m}\mathcal{N}_\tau(f) < \left|\left\langle f,\tilde{\psi}_I\right\rangle\right|\right\} \\
&= \left(\#\left\{I:\ \ \underbrace{2^{-m}\mathcal{N}_\tau(f)}_{\varepsilon} < \left|\left\langle f,\tilde{\psi}_I\right\rangle\right|\right\}\underbrace{2^{-m\tau}\mathcal{N}_\tau(f)^\tau}_{\varepsilon^\tau}\right)2^{m\tau}\mathcal{N}_\tau(f)^{-\tau} \\
&\le \left\|\left\{\left|\left\langle f,\tilde{\psi}_I\right\rangle\right|\right\}\right\|_{wl_\tau}^\tau 2^{m\tau}\mathcal{N}_\tau(f)^{-\tau} \\
&\le \left\|\left\{\left|\left\langle f,\tilde{\psi}_I\right\rangle\right|\right\}\right\|_{l_\tau}^\tau 2^{m\tau}\mathcal{N}_\tau(f)^{-\tau} \\
&= \mathcal{N}_\tau(f)^\tau 2^{m\tau}\mathcal{N}_\tau(f)^{-\tau} = 2^{m\tau}
\end{aligned}
$$

Let $N := \sum_{\nu\le m}\#\Lambda_\nu$. $f_N := \sum_{I\in\Lambda_\nu,\nu\le m}\left\langle f,\tilde{\psi}_I\right\rangle\psi_I$. Denote $F_\nu := \sum_{I\in\Lambda_\nu}\left\langle f,\tilde{\psi}_I\right\rangle\psi_i$. We apply Lemma 2 to obtain

$$
\begin{aligned}
\|f-f_N\|_2 &\le \left\|\sum_{\nu=m+1}^{\infty}F_\nu\right\|_2 \\
&\le \sum_{\nu=m+1}^{\infty}\|F_\nu\|_2 \\
&\le C\sum_{\nu=m+1}^{\infty}2^{-\nu+1}\mathcal{N}_\tau(f)\left(\#\Lambda_\nu\right)^{1/2} \\
&\le C\mathcal{N}_\tau(f)\sum_{\nu=m+1}^{\infty}2^{-\nu+1}2^{\nu\tau/2} \\
&\le C\mathcal{N}_\tau(f)\sum_{\nu=m+1}^{\infty}2^{-\nu(1-\tau/2)} \\
&\underset{\tau<2}{\le} C\mathcal{N}_\tau(f)2^{-m(1-\tau/2)} \\
&= C\mathcal{N}_\tau(f)2^{-m\tau(1/\tau-1/2)} \\
&\underset{N\le 2^{m\tau}}{\le} C\mathcal{N}_\tau(f)N^{-\alpha/n} \\
&\le CN^{-\alpha/n}|f|_{B_\tau^\alpha}.
\end{aligned}
$$

<u>Assignment:</u> Prove the case $\sum_{\nu\le m-1}\#\Lambda_\nu < N < \sum_{\nu\le m}\#\Lambda_\nu$.

□

Approximation Spaces

Let $\Phi := \{\Phi_N\}_{N\ge 0}$, each Φ_N is a set of functions in a (quasi) Banach space X , satisfying:

(i) $0\in\Phi_N$, $\Phi_0 := \{0\}$,

(ii) $\Phi_N \subset \Phi_{N+1}$,

(iii) $a\Phi_N = \Phi_N$, $\forall a \neq 0$,

(iv) $\Phi_N + \Phi_N \subset \Phi_{cN}$, for some constant $c(\Phi)$,

(v) $\overline{\bigcup_N \Phi_N} = X$,

(vi) Each $f \in X$ has a near best approximation from Φ_N. That is, there exists a constant $C(\Phi)$, such that for any N, one has $\varphi_N \in \Phi_N$,

$$\|f - \varphi_N\|_X \le C E_N(f)_X, \qquad E_N(f)_X := \inf_{\varphi \in \Phi_N} \|f - \varphi\|_X .$$

Examples for Φ_N

<u>Linear</u>

- Trigonometric polynomials of degree $\le N$, $X = L_p\left([-\pi,\pi]^n\right)$.
- Algebraic polynomials of degree $\le N$, $X = L_p[-1,1]$.
- Uniform dyadic knot piecewise polynomials over pieces of length 2^{-N}, of fixed order r, $X = L_p[0,1]$.
- Shift invariant refinable spaces $\Phi_N := S(\phi)^{2^{-N}}$, $S(\phi) \subset S(\phi)^{1/2}$, $X = L_p\left(\mathbb{R}^n\right)$.

<u>Nonlinear/Adaptive</u>

- Rational functions of degree $\le N$, $X = L_p[-1,1]$,
- Free knot piecewise polynomials of fixed order r over N non-uniform intervals, $X = L_p[0,1]$.
- N-term wavelets $\Phi_N = \Sigma_N := \left\{ \sum_{\#I \le N} c_I \psi_I \right\}$, $X = L_2\left(\mathbb{R}^n\right)$.

Def Approximation spaces For $\alpha > 0$, $0 < q \le \infty$, $f \in X$,

$$|f|_{A_q^\alpha} := \begin{cases} \left(\sum_{N=1}^{\infty} \left[N^\alpha E_N(f) \right]^q \frac{1}{N} \right)^{1/q}, & 0 < q < \infty, \\ \sup_{N \ge 1} N^\alpha E_N(f), & q = \infty. \end{cases}$$

$$\|f\|_{A_q^\alpha} := \|f\|_X + |f|_{A_q^\alpha} .$$

One can show

$$|f|_{A_q^\alpha} \sim \begin{cases} \left(\sum_{m=0}^{\infty} \left[2^{m\alpha} E_{2^m}(f) \right]^q \right)^{1/q}, & 0 < q < \infty, \\ \sup_{m \ge 0} 2^{m\alpha} E_{2^m}(f), & q = \infty. \end{cases}$$

Goal: Fully characterize approximation spaces by smoothness spaces (iff)
How? It's pretty heavy lifting:

1. Jackson & Bernstein machinery,
2. Interpolation spaces,

3. Equivalence of interpolation spaces & Smoothness (Besov) spaces.

Jackson & Bernstein machinery

We assume there is a linear quasi-Banach space $Y := Y(r)$, or $Y := Y(\alpha)$, $Y \subset X$, so that the following two inequalities are valid for $\Phi := \{\Phi_N\}_{N\geq 0}$, for $N \geq N_0$, for some fixed N_0,

[Jackson] $\quad E_N(g)_X \leq CN^{-r}|g|_Y$, $\quad g \in Y$.
[Bernstein] $\quad |\varphi|_Y \leq CN^r \|\varphi\|_X$, $\quad \varphi \in \Phi_N$.

We have seen a few examples for Jackson- type estimates

(i) $\quad E_N(g)_p \leq CN^{-r}|g|_{r,p}$, $\quad \Phi_N$ trigonometric polynomials, $g \in Y(r) = W_p^r[-\pi,\pi] \subset L_p[-\pi,\pi] = X$,

(ii) $\quad E\left(g, S(\phi)^{2^{-N}}\right)_p \leq C2^{-Nr}|g|_{r,p}$, $\Phi_N = S(\phi)^{2^{-N}}$ shift invariant spaces,
$g \in Y(r) = W_p^r(\mathbb{R}^n) \subset L_p(\mathbb{R}^n) = X$,

(iii) $\quad E_N(g)_\infty \leq CN^{-\alpha}|g|_{Lip(\alpha)}$, Φ_N piecewise constants over N uniform intervals, $0 < \alpha \leq 1$,
$Y = Lip(\alpha) \subset L_\infty[0,1] = X$,

(iv) $\quad \sigma_N(g)_\infty \leq N^{-1}|g|_{1,1}$, $\quad \Phi_N$ adaptive non-uniform piecewise constants, $Y = W_1^1[0,1] \subset L_\infty[0,1]$,

(v) $\quad \sigma_N(f)_p \leq CN^{-\alpha}|f|_{B_\tau^\alpha}$, $\quad \Phi_N = \Sigma_N$ N-term wavelets $Y = Y(\alpha) = B_\tau^\alpha \subset L_p = X$.

Let's see a few Bernstein estimates. We begin with trigonometric polynomials

Theorem For $r \geq 1$ and $1 \leq p \leq \infty$, for any real trigonometric polynomial $T_N \in \Pi_N(\mathbb{T})$,

$$|T_N|_{r,p} = \left\|T_N^{(r)}\right\|_p \leq N^r \|T_N\|_p .$$

We shall prove the case $p = \infty$ and show something stronger.

Theorem For any real $T_N \in \Pi_N(\mathbb{T})$,

$$T_N'(x)^2 + N^2 T_N(x)^2 \leq N^2 \|T_N\|_\infty^2, \qquad x \in \mathbb{T}.$$

Corollary $\|T_N'\|_\infty \leq N\|T_N\|_\infty$, and by repeated applications $\left\|T_N^{(r)}\right\|_\infty \leq N^r \|T_N\|_\infty$.

Proof of theorem First assume $\|T_N\| < 1$. W.l.g we can assume $x = 0$, and that $T_N'(0) \geq 0$. Let α, $|\alpha| < \pi/(2N)$ such that $\sin N\alpha = T_N(0)$ and define,

$$S_N(y) := \sin N(y+\alpha) - T_N(y) \in \Pi_N .$$

At the points

$$y_k := -\alpha + \frac{(2k-1)\pi}{2N}, \qquad k = 0, \pm 1, \pm 2, \ldots, \pm N ,$$

$$\operatorname{sign}\left(S_N\left(y_k\right)\right)=\operatorname{sign}\left(\underbrace{\sin\frac{(2k-1)\pi}{2}}_{=(-1)^{k+1}}-\underbrace{T_N\left(y_k\right)}_{\left|T_N\left(y_k\right)\right|<1}\right)=(-1)^{k+1}.$$

This means S_N has $2N$ zeros, with a unique zero at each interval $\left(y_k, y_{k+1}\right)$. Next,

$$\left.\begin{array}{c}\left(y_0, y_1\right)=\left(-\alpha-\dfrac{\pi}{2N},-\alpha+\dfrac{\pi}{2N}\right)\\ |\alpha|<\dfrac{\pi}{2N}\end{array}\right\}\Rightarrow 0\in\left(y_0, y_1\right)$$

Also $S_N(0):=\sin N\alpha-T_N(0)=0$, $S_N\left(y_1\right)>0$. If $S_N'(0)<0$, then there must be another zero in $\left(0, y_1\right)$, which is a contradiction. Hence $S_N'(0)\geq 0$ and

$$\begin{aligned}0\leq T_N'(0)&=N\cos N\alpha-S_N'(0)\\ &\leq N\cos N\alpha=N\sqrt{1-\sin^2 N\alpha}=N\sqrt{1-T_N(0)^2}\end{aligned}.$$

This gives

$$T_N'(0)^2\leq N^2\left(1-T_N(0)^2\right)\Rightarrow T_N'(0)^2+N^2T_N(0)^2\leq N^2.$$

Now take arbitrary $T_N\neq 0$ and $\lambda>\left\|T_N\right\|_\infty$, and apply this relation to T_N/λ. Then

$$\begin{aligned}\frac{T_N'(0)^2}{\lambda^2}+N^2\frac{T_N(0)^2}{\lambda^2}\leq N^2&\Rightarrow T_N'(0)^2+N^2T_N(0)^2\leq N^2\lambda^2\\ &\underset{\lambda\to\left\|T_N\right\|}{\Rightarrow} T_N'(0)^2+N^2T_N(0)^2\leq N^2\left\|T_N\right\|^2.\end{aligned}$$

□

Bernstein for non-uniform piecewise polynomials. Let

$$\Sigma_{N,r}:=\left\{\sum_{j=0}^{N-1}P_j\mathbf{1}_{\left[t_j, t_{j+1}\right)}: T=\left\{t_j\right\},\ 0=t_0<t_1<\cdots<t_N=1,\quad P_j\in\Pi_{r-1}\right\}.$$

Lemma For any algebraic polynomial $P\in\Pi_{r-1}\left(\mathbb{R}^n\right)$, bounded convex $\Omega\subset\mathbb{R}^n$, and $0<p,q\leq\infty$,

$$\|P\|_{L_q(\Omega)}\sim|\Omega|^{1/q-1/p}\|P\|_{L_p(\Omega)},$$

with constants of equivalency that depend on p,q,n,r but not on the polynomial or domain.

Proof By John’s Lemma, there exists an affine transformation, $Ax=Mx+b$, such that

$$B(0,1)\subseteq A^{-1}(\Omega)\subseteq B(0,n).$$

Observe that

$$A\left(B(0,1)\right)\subseteq\Omega\subseteq A\left(B(0,n)\right)\Rightarrow|\Omega|\sim|\det M|.$$

By the equivalency of same finite dimensional (quasi) Banach spaces, there exist constants depending only on p,q,n,r, such that for any $\tilde{P}\in\Pi_{r-1}\left(\mathbb{R}^n\right)$, $\left\|\tilde{P}\right\|_{L_p(B(0,1))}\sim\left\|\tilde{P}\right\|_{L_q(B(0,n))}$. Now, for $P\in\Pi_{r-1}$, denote $\tilde{P}:=P(A\cdot)\in\Pi_{r-1}$. Then,

$$\begin{aligned}\|P\|_{L_q(\Omega)} &= |\det M|^{1/q}\left\|\tilde{P}\right\|_{L_q\left(A^{-1}(\Omega)\right)}\\ &\le |\det M|^{1/q}\left\|\tilde{P}\right\|_{L_q(B(0,n))}\\ &\le C|\det M|^{1/q}\left\|\tilde{P}\right\|_{L_p(B(0,1))}\\ &\le C|\det M|^{1/q}\left\|\tilde{P}\right\|_{L_p\left(A^{-1}(\Omega)\right)}\\ &\le C|\Omega|^{1/q-1/p}\|P\|_{L_p(\Omega)}.\end{aligned}$$

□

Theorem For $\varphi\in\Sigma_{N,r}$, $\frac{1}{\tau}=\alpha+\frac{1}{p}$, $0<\alpha<r$,

$$|\varphi|_{B_\tau^\alpha}\le CN^\alpha\|\varphi\|_p.$$

Proof Let $P_j\in\Pi_{r-1}$ and $t>0$. If $t<\left(t_{j+1}-t_j\right)/r$, we have seen we can estimate

$$\omega_r\left(P_j\mathbf{1}_{[t_j,t_{j+1})},t\right)_\tau\le C\left\|P_j\mathbf{1}_{[t_j,t_{j+1})}\right\|_\infty t^{1/\tau}.$$

For $t\ge\left(t_{j+1}-t_j\right)/r$, we can bound by

$$\begin{aligned}\omega_r\left(P_j\mathbf{1}_{[t_j,t_{j+1})},t\right)_\tau &\le C\left\|P_j\mathbf{1}_{[t_j,t_{j+1})}\right\|_\tau\\ &= C\left\|P_j\right\|_{L_\tau[t_j,t_{j+1}]}\\ &\le C\left\|P_j\right\|_{L_\infty[t_j,t_{j+1}]}\left(t_{j+1}-t_j\right)^{1/\tau}.\end{aligned}$$

Therefore, for $0<\tau\le 1$ (the case $1<\tau<\infty$ is similar)

$$\begin{aligned}\omega_r(\varphi,t)_\tau^\tau &\le \sum_{j=0}^{N-1}\omega_r\left(P_j\mathbf{1}_{[t_j,t_{j+1})},t\right)_\tau^\tau\\ &\le C\sum_{j=0}^{N-1}\left\|P_j\mathbf{1}_{[t_j,t_{j+1})}\right\|_\infty^\tau \min\left(t,\left(t_{j+1}-t_j\right)/r\right).\end{aligned}$$

We apply the lemma for $q=\infty$,

$$
\begin{aligned}
|\varphi|_{B_\tau^\alpha}^\tau &= \int_0^\infty \left(t^{-\alpha}\omega_r(\varphi,t)_\tau\right)^\tau \frac{dt}{t} \\
&\le C\sum_{j=0}^{N-1}\left\|P_j \mathbf{1}_{[t_j,t_{j+1})}\right\|_\infty^\tau \int_0^\infty t^{-\alpha\tau}\min\left(t,\left(t_{j+1}-t_j\right)/r\right)\frac{dt}{t} \\
&\le C\sum_{j=0}^{N-1}\left\|P_j \mathbf{1}_{[t_j,t_{j+1})}\right\|_p^\tau \left(t_{j+1}-t_j\right)^{-\tau/p}\left(\int_0^{(t_{j+1}-t_j)/r} t^{-\alpha\tau}dt+\left(t_{j+1}-t_j\right)\int_{(t_{j+1}-t_j)/r}^\infty t^{-\alpha\tau-1}dt\right) \\
&\le C\sum_{j=0}^{N-1}\left\|P_j \mathbf{1}_{[t_j,t_{j+1})}\right\|_p^\tau \left(t_{j+1}-t_j\right)^{-\tau/p}\left(t_{j+1}-t_j\right)^{1-\alpha\tau} \\
&\underset{1=\alpha\tau+\tau/p}{=} C\sum_{j=0}^{N-1}\left\|P_j \mathbf{1}_{[t_j,t_{j+1})}\right\|_p^\tau \\
&\underset{Holder\ p>\tau}{\le} C\left(\sum_{j=0}^{N-1}\left(\left\|P_j \mathbf{1}_{[t_j,t_{j+1})}\right\|_p^\tau\right)^{p/\tau}\right)^{\tau/p} N^{1-\tau/p} \\
&= C\left(\sum_{j=0}^{N-1}\left\|P_j \mathbf{1}_{[t_j,t_{j+1})}\right\|_p^p\right)^{\tau/p} N^{1-\tau/p} \\
&= C\|\varphi\|_p^\tau N^{1-\tau/p}.
\end{aligned}
$$

Therefore

$$|\varphi|_{B_\tau^\alpha} \le C\|\varphi\|_p N^{1/\tau-1/p} = CN^\alpha\|\varphi\|_p .$$

□

Theorem [Bernstein for wavelets] For $\varphi(x)=\sum_{j=1}^N c_{I_j}\psi_{I_j}$, $1<p<\infty$, $\tau^{-1}=\alpha/n+1/p$, we have

$$|\varphi|_{B_\tau^\alpha} \le CN^{\alpha/n}\|\varphi\|_p .$$

Theorem Let Y,X , $r>0$, and $\Phi:=\{\Phi_N\}$ as above. For the K-functional

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

(i) If the Jackson inequality is satisfied then

$$E_N(f)_X \le CK\left(f,N^{-r}\right),\quad f\in X\ ,\ N=1,2,\ldots .$$

(ii) If the Bernstein inequality is satisfied then (for $|\cdot|_Y$ semi-norm)

$$K\left(f,2^{-mr}\right)\le C2^{-mr}\sum_{k=0}^m 2^{kr}E_{2^k}(f).$$

Proof

(i) Let $f\in X$. By the Jackson theorem, for any $g\in Y$

$$E_N(f)_X \le \|f-g\|_X+E_N(g)_X \le C\left(\|f-g\|_X+N^{-r}|g|_Y\right).$$

We then take infimum over $g\in Y$.

(ii) Let $\varphi_k \in \Phi_{2^k}$, such that $\|f-\varphi_k\| \le CE_{2^k}(f)$, $k=0,1,2\ldots$. Denote $\psi_0 = \varphi_0 = 0$, $\psi_k := \varphi_k - \varphi_{k-1}$, $k \ge 1$. Observe that by properties (iii), (iv), $\psi_k \in \Phi_{c2^k}$. Using the fact that $\{\varphi_k\}$ are near-best approximants

$$\|\psi_k\| \le \|f-\varphi_k\| + \|f-\varphi_{k-1}\| \le 2CE_{2^{k-1}}(f),\ k \ge 1.$$

Since $\varphi_m = \sum_{k=0}^{m} \psi_k$, $|\psi_0|_Y = 0$, it follows that

$$\begin{aligned}
K\left(f,2^{-mr}\right) &\le \|f-\varphi_m\|_X + 2^{-mr}|\varphi_m|_Y \\
&\le C\left(E_{2^m}(f) + 2^{-mr}\sum_{k=1}^{m}|\psi_k|_Y\right) \\
&\le C\left(E_{2^m}(f) + 2^{-mr}\sum_{k=1}^{m}2^{kr}\|\psi_k\|_X\right) \\
&\le C\left(E_{2^m}(f) + 2^{-mr}\sum_{k=1}^{m}2^{kr}E_{2^{k-1}}(f)\right) \\
&\le C2^{-mr}\sum_{k=0}^{m}2^{kr}E_{2^k}(f).
\end{aligned}$$

□

Interpolation spaces

For $0<\theta<1$, $0<q\le\infty$, $X,Y(r)$,

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

$$|f|_{(X,Y)_{\theta,q}} = |f|_{\theta,q} := \begin{cases} \left(\int_0^1 \left[t^{-\theta}K(f,t)\right]^q \frac{dt}{t}\right)^{1/q}, & 0<q<\infty, \\ \sup_{0<t\le 1} t^{-\theta}K(f,t), & q=\infty. \end{cases}$$

$$\|f\|_{\theta,q} := \|f\|_X + |f|_{\theta,q}.$$

It is convenient to discretize at $t_m = \left(2^r\right)^{-m} = 2^{-mr}$, $m \ge 0$,

$$|f|_{\theta,q} \sim \begin{cases} \left(\sum_{m=0}^{\infty}\left[2^{mr\theta}K\left(f,2^{-mr}\right)\right]^q\right)^{1/q}, & 0<q<\infty, \\ \sup_{m\ge 0} 2^{mr\theta}K\left(f,2^{-mr}\right), & q=\infty. \end{cases}$$

Def We call the space of functions for which $\|f\|_{\theta,q}$ is finite the **interpolation space** $(X,Y)_{\theta,q}$.

Observe that by definition $(X,Y)_{\theta,q} \subset X$, while for $0<\theta<1$, we have that $Y \subset (X,Y)_{\theta,q}$. To see this, let $g \in Y$. Then,

$$\sum_{m=0}^{\infty}\left[2^{mr\theta}K\left(g,2^{-mr}\right)\right]^q \le \sum_{m=0}^{\infty}\left[2^{mr\theta}2^{-mr}|g|_Y\right]^q$$
$$= |g|_Y^q \sum_{m=0}^{\infty} 2^{mrq(\theta-1)}$$
$$\le C(r,q,\theta)|g|_Y^q .$$

Also, for $\theta_2 \le \theta_1$, $(X,Y)_{\theta_1,q} \subseteq (X,Y)_{\theta_2,q}$. So, for $0<\theta<1$, we have a 'scale' of spaces between Y and X.

Example $\left(L_p, W_p^r\right)_{\theta,q} = B_q^{\alpha}\left(L_p\right)$, $\alpha=\theta r$, $0<\theta<1$, $1\le p\le\infty$.

Proof

$$\int_0^1\left[t^{-\theta}K(f,t)\right]^q\frac{dt}{t} = \int_0^1\left[t^{-\theta}K_r(f,t)_p\right]^q\frac{dt}{t}$$
$$\sim \int_0^1\left[t^{-\alpha/r}\omega_r\left(f,t^{1/r}\right)_p\right]^q\frac{dt}{t} \qquad s=t^{1/r} \Rightarrow ds=\frac{1}{r}t^{1/r-1}dt \Rightarrow s^{-1}ds=\frac{1}{r}t^{-1}dt$$
$$\underset{s=t^{1/r}}{\sim} \int_0^1\left[s^{-\alpha}\omega_r(f,s)_p\right]^q\frac{ds}{s}$$
$$= |f|_{B_q^{\alpha}(L_p)}^q .$$

□

Lemma [Discrete Hardy inequality] For a sequence $a=\{a_m\}_{m\in\mathbb{Z}}$, we define

$$\|a\|_{\alpha,q} := \begin{cases}\left(\sum_{m\in\mathbb{Z}}\left(2^{m\alpha}|a_m|\right)^q\right)^{1/q}, & 1\le q<\infty,\\ \sup_{m\in\mathbb{Z}} 2^{m\alpha}|a_m|, & q=\infty.\end{cases}$$

If $\{a_m\},\{b_m\}$ are sequences and we know that for some $\lambda>\alpha$

$$|b_m| \le C2^{-m\lambda}\sum_{k=-\infty}^{m}2^{k\lambda}|a_k|,\ \forall m,$$

then, $\|b\|_{\alpha,q} \le C\|a\|_{\alpha,q}$.

Proof W.l.g assume the sequences are non-negative. We prove for $1\le q<\infty$. Pick $\alpha<\beta<\lambda$. With $\frac{1}{q}+\frac{1}{q'}=1$,

$$\sum_{k=-\infty}^{m}2^{k\lambda}a_k = \sum_{k=-\infty}^{m}2^{k(\lambda-\beta)}a_k 2^{k\beta}$$
$$\le \left(\sum_{k=-\infty}^{m}2^{k(\lambda-\beta)q'}\right)^{1/q'}\left(\sum_{k=-\infty}^{m}a_k^q 2^{k\beta q}\right)^{1/q}$$
$$\le C2^{m(\lambda-\beta)}\left(\sum_{k=-\infty}^{m}a_k^q 2^{k\beta q}\right)^{1/q}$$

Therefore,

$$
\begin{aligned}
\sum_{m=-\infty}^{\infty} 2^{m\alpha q} b_m^q &\le C \sum_{m=-\infty}^{\infty} 2^{m\alpha q} 2^{-m\lambda q} 2^{m(\lambda-\beta)q} \sum_{k=-\infty}^{m} a_k^q 2^{k\beta q} \\
&= C \sum_{m=-\infty}^{\infty} 2^{m(\alpha-\beta)q} \sum_{k=-\infty}^{m} 2^{k\alpha q} a_k^q 2^{k(\beta-\alpha)q} \\
&= C \sum_{m=-\infty}^{\infty} \sum_{k=-\infty}^{m} 2^{k\alpha q} a_k^q \left(2^{m(\alpha-\beta)q} 2^{k(\beta-\alpha)q}\right) \\
&= C \sum_{k=-\infty}^{\infty} 2^{k\alpha q} a_k^q \sum_{m=k}^{\infty} 2^{(m-k)(\alpha-\beta)q} \\
&\le C \sum_{k=-\infty}^{\infty} 2^{k\alpha q} a_k^q
\end{aligned}
$$

□

Theorem If the Jackson and Bernstein inequalities are valid for $X, Y(r), \Phi$, then, for $0 < \theta < 1$, $\alpha := \theta r$ $0 < q \le \infty$, then

$$
A_q^\alpha(X) = A_q^{\theta r}(X) \sim (X, Y)_{\theta,q}.
$$

Proof Assume $f \in (X, Y)_{\theta,q}$. By a previous theorem, the Jackson inequality gives $E_{2^m}(f)_X \le CK(f, 2^{-mr})$. Therefore, with $\alpha := \theta r$,

$$
\begin{aligned}
|f|_{A_q^\alpha}^q &\le C \sum_{m=0}^{\infty} \left[2^{m\alpha} E_{2^m}(f)\right]^q \\
&\le C \sum_{m=0}^{\infty} \left[2^{mr\theta} K(f, 2^{-mr})\right]^q \\
&\le C |f|_{\theta,q}^q .
\end{aligned}
$$

Now assume $f \in A_q^\alpha(X)$. We shall use the discrete Hardy inequality. In our case, we have $a_m = E_{2^m}(f)$, $b_m = K(f, 2^{-mr})$, for $m \ge 0$, $a_m, b_m = 0$, for $m < 0$. Using the Bernstein inequality, we proved that with $\lambda := r > \alpha = \theta r$

$$
K(f, 2^{-mr}) \le C 2^{-mr} \sum_{k=0}^{m} 2^{kr} E_{2^k}(f).
$$

Therefore,

$$
\begin{aligned}
|f|_{\theta,q} &\le C \left(\sum_{m=0}^{\infty} \left[2^{mr\theta} K(f, 2^{-mr})\right]^q \right)^{1/q} \\
&\le C \left(\sum_{m=0}^{\infty} \left[2^{m\alpha} E_{2^m}(f)\right]^q \right)^{1/q} \\
&\le C |f|_{A_q^\alpha} .
\end{aligned}
$$

□

Characterization of approximation spaces

1. Trigonometric polynomials

- $X = L_p[-\pi,\pi]$, $Y = W_p^r[-\pi,\pi]$, $r \ge 1$, $1 \le p \le \infty$.
- We proved the Jackson theorem
$$E_N(f)_p \le \omega_r\left(f,\frac{1}{N}\right)_p, \ \forall f \in L_p \Rightarrow E_N(g)_p \le CN^{-r}|g|_{r,p}, \ \forall g \in W_p^r.$$
- We proved the Bernstein inequality,
$$|T_N|_{r,p} = \left\|T_N^{(r)}\right\|_p \le N^r \|T_N\|_p, \ \forall T_N \in \Pi_N(\mathbb{T}).$$

- So we can apply the Jackson-Bernstein machinery to obtain for $0<\theta<1$, $\alpha := \theta r$, $0 < q \le \infty$,

$$A_q^\alpha(L_p) = A_q^{\theta r}(L_p) \sim (L_p, W_p^r)_{\theta,q} \sim B_q^\alpha(L_p).$$

2. Dyadic piecewise polynomials approximation

- $X = L_p[0,1]$, $Y = W_p^r[0,1]$, $r \ge 1$, $1 \le p \le \infty$.

- We proved the Jackson theorem for refinable shift invariant spaces $S(\phi) \subset S(\phi)^{1/2}$, which include B-splines that are piecewise polynomials

$$E\left(g, S(\phi)^{2^{-j}}\right)_p \le C2^{-jr}|g|_{r,p}, \ \forall g \in W_p^r.$$
- For $0<\theta<1$, $\alpha < r-1+1/p$, $0<q\le\infty$,
$$A_q^\alpha(L_p) \sim B_q^\alpha(L_p).$$

3. Wavelets

- $X = L_p(\mathbb{R}^n)$, $1<p<\infty$, $Y = B_\tau^\alpha$, $\frac{1}{\tau} = \frac{\alpha}{n} + \frac{1}{p}$.
- Jackson
$$\sigma_N(f)_p \le CN^{-\alpha/n}|f|_{B_\tau^\alpha}.$$
- Bernstein $\varphi(x) = \sum_{j=1}^N c_{I_j}\psi_{I_j}$
$$|\varphi|_{B_\tau^\alpha} \le CN^{\alpha/n}\|\varphi\|_p.$$

- Approximation space with parameter $r = \alpha/n$, $0<\theta<1$,

$$A_q^{\theta\alpha/n}(L_p) \sim (L_p, B_\tau^\alpha)_{\theta,q}.$$

- Reiteration theorem (from interpolation theory) If

$$\frac{1}{q} = \frac{\theta\alpha}{n} + \frac{1}{p}, \quad 0 < \theta < 1,$$

then

$$\left(L_p, B_\tau^\alpha\right)_{\theta,q} \sim B_q^{\alpha\theta}$$

- So, we can circle back to the original notation, define $\tilde{\alpha} = \theta\alpha$, $\tilde{\tau}$ by

$$\frac{1}{\tilde{\tau}} = \frac{\tilde{\alpha}}{n} + \frac{1}{p},$$

then

$$A_{\tilde{\tau}}^{\tilde{\alpha}/n}\left(L_p\right) \sim B_{\tilde{\tau}}^{\tilde{\alpha}}.$$

4. Adaptive non-uniform univariate piecewise polynomial approximation

- Same approximation space as wavelets. So, wavelets are a better way to implement the adaptive approach, certainly in higher dimensions.

Important observation So, we get characterizations with Besov spaces of different indices…but they all look the same. Not so! Let's go back to the example of $\Omega = [-1,1]$,

$$f(x) = \begin{cases} 0 & x < 0 \\ 1 & 0 \le x \end{cases}$$

We showed that $\omega_r(f,t)_\tau \sim t^{1/\tau}$. So, if we want to compute the Besov semi-norm in $B_q^\alpha(L_\tau)$, it is sufficient to compute

$$\int_0^1 \left(t^{-\alpha}\omega_r(f,t)_\tau\right)^q \frac{dt}{t} \sim \int_0^1 t^{(1/\tau-\alpha)q-1} dt.$$

We see that only if $1/\tau > \alpha$, this is finite, e.g., we need to integrate the modulus at 'smaller' index τ for higher smoothness. This means nonlinear approximation, can provide a good degree of approximation (fast decay rate), for functions that only exist in Besov spaces with 'smaller' index τ.