Three ages of function spaces: Generalized smoothness, Fourier characterization and Coorbit Spaces

Hans G. Feichtinger
hans.feichtinger@univie.ac.at
www.nuhag.eu

University of Marburg, Sept. 13th, 2010

Abstract

The purpose of this talk is to give a *historical perspective on some aspects of the theory of function spaces*, i.e. Banach spaces of functions (or distributions, when one looks at the dual spaces). The first approach to smoothness resulting in the definition of Sobolev spaces and Besov spaces (Besov, Taibleson, Stein) came from the idea of generalized smoothness, expressed by (higher order) difference expression and the corresponding *moduli of continuity*, e.g. describing smoothness by the decay of the modulus of continuity (via the membership in certain weighted Lq-spaces on (0,1]). Alternatively there is the line described in the book of S.Nikolksii characterizing smoothness (equivalently) by the degree of approximation using band-limited functions (S. M. Nikol'skij [7]). Fractional order Sobolev spaces can be expressed in terms of weighted Fourier transforms.

The second and third age

The second age is characterized by the Paley-Littlewood characterizations of Besov or Triebel-Lizorkin spaces using dyadic decompositions on the Fourier transform side, as used in the work of J.Peetre ([8]) and H.Triebel ([15, 16, 14, 10, 17]) , the masters of interpolation theory. Their contribution was to show that these families of function spaces are stable under (real and complex) interpolation methods.

The third age is - from our point of view - the characterization of function spaces in the context of coorbit spaces, using irreducible integrable group representations of locally compact groups.

Let us also remind that the concept of retracts plays an important role in the context of interpolation theory (see the book of Bergh-Loefstroem), and can be used to characterize Banach frames and Riesz projection bases.

Modulus of continuity

Definition

Assume that $(\mathbf{B} \| \cdot \|_{\mathbf{B}})$ is an isometrically translation invariant Banach space of locally integrable functions (i.e. $(\mathbf{B} \| \cdot \|_{\mathbf{B}}) \hookrightarrow \mathbf{L}^1_{loc}(\mathbb{R}^d)$) with

$$\| T_x f \|_{\mathbf{B}} = \| f \|_{\mathbf{B}} \quad \forall f \in \mathbf{B}.$$

In this situation we can define for every $f \in \mathbf{B}$ its *modulus of continuity with respect to* $\| \cdot \|_{\mathbf{B}}$ via

$$\omega_\delta(f) = sup_{|x| \leq \delta} \{ \| T_x f - f \|_{\mathbf{B}} \}.$$

In most cases ω is considered a function of δ for fixed $f \in \mathbf{B}$, but the notation is following the traditional one.

Modulus of continuity 2

For each such space it is easy to show that the elements with $\lim_{\delta \to 0} \omega_\delta(f) = 0$ are those for which $x \to T_x f$ is (uniformly) continuous from $\mathbb{R}^d$ into $\left(\mathbf{B} \| \cdot \|_{\mathbf{B}}\right)$. They form a *closed subspace* of $\left(\mathbf{B} \| \cdot \|_{\mathbf{B}}\right)$, which we denote by $\mathbf{B}_{cs}$ [1].
Within this class we can identify functions which have a higher degree of "smoothness", i.e. which are not just uniformly continuous, but behave better than the general function in $\mathbf{C}_{ub}(\mathbb{R}^d)$, because ω_δ tends to zero for $\delta \to 0$ at a given rate.
The so-called Lipschitz spaces $Lip(\alpha)$ are characterized by the property that there exists some constant $C > 0$ such that

$$\sup_{\delta > 0} \delta^{-\alpha} \omega_\delta(f) = C < \infty. \tag{1}$$

There are also so-called "small Lipschitz spaces" characterized by

$$\lim_{\delta \to 0} \delta^{-\alpha} \omega_\delta(f) = 0. \tag{2}$$

[1] "cs" standing for "continuous shift". for the case $(\mathbf{L}^\infty(\mathbb{R}^d), \| \cdot \|_\infty)$ we

Classical Lipschitz spaces

It is an easy exercise to show that $Lip\,(\alpha)$ is a Banach space with the norm

$$\|f\|_{Lip\,(\alpha)} := \|f\|_\infty + \sup_{\delta>0} \delta^{-\alpha}\omega_\delta(f), \tag{3}$$

and that $(lipa, \|\cdot\|_{Lip\,(\alpha)})$ is a *closed subspace* of $Lip\,(\alpha)$, in fact $Lip\,(\alpha)_{cs} = lipa$.

This construction makes only sense for $\alpha \in (0,1]$, because the class becomes trivial for $\alpha > 1$[2].

There are two ways out, which turn out to be equivalent: Either one assumes that f is continuously differentiable and f' satisfies a Lipschitz condition, *or* one makes use of higher order difference operators e.g. for $k=2$ one expects decay of the sup-norm of the function $f(x-h) - 2f(x) + f(x+h)$ as $h \to 0$, with some order of h, up to order < 2 (also higher order differences).

[2]Only constant functions: because the assumption implies that the function *is differentiable everywhere* and that $f'(x) \equiv 0$.

Generalized Lipschitz spaces $\mathbf{Lip}(p, \alpha)$

Replacing in this traditional the sup-norm by an $\mathbf{L}^p$-norms and the corresponding modulus of continuity one arrives at the concept of the Lipschitz spaces $\mathbf{Lip}(p, \alpha)$ arise.
The next step towards a general theory of smoothness spaces was taken by **Besov**. Instead of considering just decay of a given order for the modulus of continuity (as a function on $(0, 1]$ or $\mathbb{R}^+$) he was making use of weighted $\mathbf{L}^q$-spaces with respect to the (natural $=$ Haar) measure dt/t on $\mathbb{R}^+$?
The corresponding norms (on $\mathbb{R}^+$ or $(0, 1]$) are of the form

$$\left[\int_0^1 (|H(t)|t^{-s})^q dt/t\right]^{1/q}. \tag{4}$$

Besov spaces

Note that the natural (say exponential function) isomorphism of $(\mathbb{R}, +)$ with $(\mathbb{R}^+, \cdot)$ via the exponential function transports functions H on $\mathbb{R}^+$ back into functions $h(t) := H(exp(t))$, so that the condition (4) is equivalent to the membership of h in the usual (polynomial) weighted $\mathbf{L}^q$-space,

$$\mathbf{L}^q_{w_s}(\mathbb{R}) := \{f \mid fw_s \in \mathbf{L}^q(\mathbb{R}^d)\}, \quad \text{with} \quad w_s(t) := (1 + |t|)^s. \quad (5)$$

which is a Banach space with its natural norm $\|f\|_{q,w_s} := \|fw_s\|_q$. The resulting family of spaces is then just the family of Besov spaces $\mathbf{B}^s_{p,q}(\mathbb{R}^d)$.
In the work of S. Nikolskij (still alive!? at age of 102?) the Besov spaces have been characterized by their approximation behaviour with respect to band-limited functions (in his work: entire functions of exponential type, [7]).

Sobolev spaces, fractional derivatives

On the other hand there was the idea of describing smoothness in the sense of differentiability in terms of the Fourier transform. The classical Sobolev spaces $\mathbf{W}^k(\mathbb{R}^d)$ or $(\mathcal{H}_s(\mathbb{R}^d), \|\cdot\|_{\mathcal{H}_s})$ or $\mathcal{L}^2_s$ are defined as the function having a derivative up to order k in $\mathbf{L}^2(\mathbb{R}^d)$. Of course it requires some care to explain in which sense this existence is to be interpreted. There are various natural options:

- assuming the existence of the classical (partial) derivatives a.e. and assuming that they define $\mathbf{L}^2$-functions;
- taking the derivative in the distributional sense and assume that those derivatives are *regular* distributions, i.e. can be represented by $\mathbf{L}^2$-functions;
- use Plancherel's theorem and make use of the fact that the differentiation corresponds to pointwise multiplication with polynomials on the Fourier transform side;

Fortunately these conditions are all *equivalent*!!

The Fourier and Littlewood-Paley age

To my knowledge it have been mostly the two pioneers in interpolation theory, namely Jaak Peetre and Hans Triebel. The most important alternative description of Besov (and also Bessel potential spaces $(\mathcal{H}_s(\mathbb{R}^d), \|\cdot\|_{\mathcal{H}_s})$, which are special cases of the more general Triebel-Lizorkin spaces) is through *dyadic partitions of unity*, typically in the form of dilation of a fixed function ψ which is assumed to be such that one can control all of its derivatives.
The classical description of Besov spaces in the books of Triebel makes use of terms such as

$$\| \mathcal{F}^{-1}[\widehat{f} \cdot \psi(2^k \cdot)]\|_p \qquad (6)$$

Since we are working with Banach spaces (such as $\mathbf{L}^p(\mathbb{R}^d)$ etc.) within the tempered distributions $\mathcal{S}'(\mathbb{R}^d)$ anyway, I prefer to rather *take the* $\mathbf{L}^p$*-norm over to the Fourier transform side*, rather than jumping between time- and frequency side all the time.

The Fourier age

This means, that I prefer to use dilation operators

$$[D_\rho h](z) = h(\rho z), \quad \rho > 0 \tag{7}$$

and define for $h = \widehat{f}$, with $f \in \mathbf{L}^p(\mathbb{R}^d)$:

$$\|h\|_{\mathcal{F}\mathbf{L}^p} := \|f\|_p. \tag{8}$$

Dilation on the Fourier transform side using D_ρ corresponds to $\mathbf{L}^1$-norm preserving dilation on the time side using:

$$St_\rho f(z) = \rho^{-d} f(z/\rho), \quad \text{for} \rho \neq 0, \tag{9}$$

we find that $\|D_\rho f\|_{\mathcal{F}\mathbf{L}^1} = \|f\|_{\mathcal{F}\mathbf{L}^1} \quad \text{for} \quad \rho \neq 0.$
Consequently (6) is equivalent to

$$\|\widehat{f} \cdot D_{2^k}\psi\|_{\mathcal{F}\mathbf{L}^p} \tag{10}$$

with the side condition that $\sum_{k\in\mathbb{Z}} D_{2^k}\psi(x) \equiv 1$ on $\mathbb{R}^d \setminus \{0\}$. This is what we call a dyadic decomposition of unity.

NEXT

In fact, the smoothness assumptions on ψ can easily be translated into an uniform boundedness condition of the family $(\psi_k) := (D_{2^k}\psi)_{k\in\mathbb{Z}}$ in $\left(\mathcal{F}\mathbf{L}^1(\mathbb{R}^d), \|\cdot\|_{\mathcal{F}\mathbf{L}^1}\right)$.
There is a deep result from analysis which helps to also characterize the Triebel-Lizorkin spaces (recall, you have among them the $\mathbf{L}^p$-potential spaces, obtained by applying the *Fourier multiplier* $w_{-s}(\xi) = (1+|\xi|^2)^{-s/2}$ to $\mathcal{F}\mathbf{L}^p$, are among them) using also the sequence of functions $(\mathcal{F}^{-1}\widehat{f}\cdot D_{2^k}\psi)_{k\in\mathbb{Z}}$, using the so-called Paley-Littlewood decomposition. It allows to express the $\mathbf{L}^p$-norm equivalently by the $\mathbf{L}^p$-norm of the function

$$h(t) = \left(\sum_{k\in\mathbb{Z}} |\,\mathcal{F}^{-1}(\widehat{f}\cdot\psi_k)(t)|^2\right)^{-1/2} \qquad (11)$$

continued ..

Putting weights into the sum, i.e. using the functions

$$h_s(t) = \left(\sum_{k \in \mathbb{Z}} |w_s(2^{-k}) \cdot \mathcal{F}^{-1}(f \cdot \psi_k)(t)|^2 \right)^{-1/2} \qquad (12)$$

we find (cf. work of E.Stein, Triebel etc.) that the p-Bessel potential norm or order s of f is equivalent to $\|h_s\|_p$.
At first sight it looks that the difference between the two types (Sobolev or Besov spaces) consists in the order in which the continuous $\mathbf{L}^p$-norm resp. the discrete ℓ^q-norm are applied.
However, there are also other mixtures, e.g. a *completely continuous* characterization, where finally only the order in which the summation is realized is relevant. For $p = 2 = q$ we just have the classical $\mathbf{L}^2$ Sobolev spaces.

The method of Frazier-Jawerth: atomic decompositions

The approach taken by Frazier and Jawerth (certainly heavily influenced by the work of Jaak Peetre) established a connection between the characterization of the different *function spaces* (to use Triebel's terminology) with dyadic decompositions in order to arrive at atomic decomposition of these spaces resp. characterizations of function spaces by the coefficients.
In a nutshell the dyadic decompositions allow to decompose a function (or tempered distribution) into contributions sitting in dyadic *frequency bands* which in turn can expanded into series of shifted atoms (suitably chosen) making use of (dilated versions) of Shannon's sampling theorem (for each of the blocks).

Comments on those early atomic decompositions

The atomic decompositions proposed in the work of Frazier-Jawerth claim that there are function spaces (in fact pairs of functions, matching well to each other, but different from each other) such that one could be used for analysis, i.e. in order to generate a set of coefficients, while the other is used for synthesis. An important point is the fact that these atoms (used for analysis and synthesis) are transformed jointly (using dyadic dilations and essentially integer translations), and make sure that for each of the classical function spaces there is an appropriate (solid) Banach space of sequences, allowing to characterize the distributions by the coefficients arising in the decomposition.

Connection to Wavelet Theory

With the advent of wavelet theory it was found, that all those *function spaces* (Besov-Triebel-Lizorkin spaces) have a characterization in terms of the CWT (continuous wavelet transform), which is defined over the upper half-plan (parameterized by the parameters $a > 0, b \in \mathbb{R}$), better viewed as the "ax+b"-group G, which is a locally compact group with left (and different from it) right Haar measure.
The correct characterization of function spaces is in terms of mixed norm spaces (mixed $\mathbf{L}^p - \mathbf{L}^q$-norms over G), with a weight depending only on the scale variable $a > 0$ in a natural way.
Anisotropic and weighted spaces can be characterized by alternative weight functions depending on *a and b* as well.

Calderon's reproducing formula

Reinterpretation of older results in the light of wavelet theory shows that the characterization of function spaces by *higher order differences* is more or less using a wavelet transform with respect to some very "rough" wavelet, namely a weighted sum of Dirac-measures (e.g. $\delta_{-1} - 2\delta_0 + \delta_1$ or its convolution powers), which are however satisfying the admissibility by having the *correct behaviour of their Fourier transform* near the origin.
The role of the partition of unity property (only valid for specific Schwartz functions) for dyadic partitions on the FT-side is taken by the more flexible continuous analogue, the so-called Calderon Reproducing formula, which can be seen as a direct consequence of the fact that the CWT is isometric from $\mathcal{H} = \mathbf{L}^2(\mathbb{R}^d)$ into $\mathbf{L}^2(G)$. Hence the inverse operator on the range of the CWT is just its adjoint. This allows to characterize all those function spaces using arbitrary *admissible* wavelets in $\mathcal{S}(\mathbb{R}^d)$.

Orthonormal Wavelet Bases

One of the important developments in wavelet theory has been the construction of orthonormal wavelet basis due to Yves Meyer, Lemarie, and above all Ingrid Daubechies, who was the first to construct orthonormal wavelet bases with compact support and a given degree of smoothness. They cannot be used to characterize all the function spaces, but e.g. Besov spaces $\mathbf{B}^s_{p,q}(\mathbb{R}^d)$ up to some order $|s| \leq s_0$.

It was certainly an important property of wavelets (aside from the fact that they came early on together with efficient algorithms) that they could be used to *characterize* most of the important function spaces known at that time, using the wavelet coefficients. Again, the quality of the atoms g (typically a combination of decay and smoothness conditions) are relevant for the range of parameters they could handle.

Coorbit Theory: the third age

Coorbit theory gives a group theoretical framework to all those statements, using a group theoretical point of view.
It started out as an attempt to understand the similarities between known results in the theory of function spaces, wavelet transforms, including orthonormal expansions.
The analogy between Besov spaces and modulation spaces (introduced in the early 80s, imitating the definition of $\mathbf{B}^s_{p,q}$ by replacing the dyadic BAPUs by uniform partitions of unity (BUPUs) in order to get to the $\mathbf{M}^s_{p,q}$-family) was quite obvious and motivated the search for their common properties and analogies.

Coorbit Theory, group representation theory

The insight was, that one only needs an integrable group representation of some locally compact group (such as the "ax+b" or the reduced Heisenberg group), say $\pi(x)$ on some Hilbert space $\mathcal{H}$, in order to come up with a continuous *voice transform*

$$V_g f(x) = \langle f, \pi(x) g \rangle_{\mathcal{H}}, \quad x \in G. \tag{13}$$

Then one can use Moyal's formula (a kind of Plancherel theorem for non-commutative groups) in order to come up with (the weak form) of a reproducing formula, allowing to write any element $f \in \mathcal{H}$ as a *"continuous" superposition* of elements of the form $\pi(x)g$, for suitable (admissible) atoms $g \in \mathcal{H}$. There is an abundance of such situations, shearlet theory being the most recent one. Margit Pap is studying the Moebius group.

Function spaces from the Coorbit point of view

Already a first step towards a continuous characterization is the reinterpretation of the Calderon reproducing formula which - in a modern interpretation - shows that the family $\pi(x)g, x \in G$ defines a continuous frame (at least for *admissible atoms* $g \in \mathcal{H}$).
Coorbit theory unifies various aspects and exhibits analogies between different families of spaces, such as modulation spaces (linked to the *Schrödinger representation of the (reduced) Heisenberg group*) or Besov-Triebel-Lizorkin spaces, linked to the affine group ("ax+b"-group).
While it is possible to have wavelet orthonormal bases (i.e. orthonormal bases of the form $(\pi(\lambda_i)g)_{i\in I}$, where (λ_i) is a discrete set in $''ax + b''$ nothing similar is possible in the case of modulation spaces (despite D. Gabor's original hope and suggestion).

Literature and relevant citations

Taibleson: Fourier analysis on local fields ([13])
Stein-Weiss [12] Introduction to Fourier Analysis on Euclidean Spaces.
E. Stein:[11] Singular Integrals and Differentiability Properties of Functions.
J. Peetre [8] New Thoughts on Besov Spaces.
H. Triebel (lots of books), such as: [15, 16, 14, 10, 17]
S. M. Nikol'skij [7] Approximation of Functions of Several Variables and Imbedding Theorems. Translated from the Russian by J. M. Danskin.
H.G. Feichtinger and K. Gröchenig: coorbit papers [2, 3]
I. Daubechies, A. Grossmann and Y. Meyer [1] Painless nonorthogonal expansions
M. W. Frazier, B. D. Jawerth and G. Weiss [6] Littlewood-Paley Theory and the Study of Function Spaces.
M. Frazier and B. Jawerth [5] A discrete transform and decompositions of distribution spaces.
M. Frazier and B. Jawerth [4] Decomposition of Besov spaces.

END OF THE PRESENTATION in Marburg

Question (A) from the audience: Where did the name modulation space come from:
Answer: While Besov spaces and other function spaces can be characterized by the rate of convergence by which the solution of the heat equation approaches the initial value f, i.e. by

$$\|(St_\rho h) * f - f\|_{\mathbf{L}^p} = \|[St_\rho(h - \delta_0)] * f\|_{\mathbf{L}^p}$$

(where h is the Gauss function, with $\int_{\mathbb{R}^d} h(x)dx = 1$), we can reformulate the growth conditions of $V_g(f)$ over $\mathbb{R}^d \times \widehat{\mathbb{R}}^d$ equivalently by looking at the decay of $\|M_t g * f\|_{\mathbf{L}^p}$ for $t \to \infty$ (which can be seen as a kind of quantitative variant of the *Riemann-Lebesgue Lemma*, according to which $\widehat{f} \in \mathbf{C}_0(\mathbb{R}^d)$ for $f \in \mathbf{L}^1(\mathbb{R}^d)$). The name is based on the fact that $M_t g$ is a modulated version of g.

END OF THE PRESENTATION in Marburg

Question B from the audience: Where does the name coorbit space come from.
Answer: This is related to terminology already used in a more general setting by Jaak Peetre in his paper [9]:
Jaak Peetre [pe85] Paracommutators and minimal spaces. In "Operators and Function Theory", Proc. NATO Adv Study Inst, Lancaster/Engl 1984, NATO ASI Ser, Ser C 153, 163-224, (1985)
There are certainly motivations coming from the two equivalent descriptions of the real interpolation method, namely the K- and the J-method, which are also kind of dual to each other.

THANK you very much for your attention!

WARNING: USING ENDBIBL!!!

I. Daubechies, A. Grossmann, and Y. Meyer.
Painless nonorthogonal expansions.
J. Math. Phys., 27(5):1271–1283, May 1986.

H. G. Feichtinger and K. Gröchenig.
Banach spaces related to integrable group representations and their atomic decompositions, I.
J. Funct. Anal., 86(2):307–340, 1989.

H. G. Feichtinger and K. Gröchenig.
Banach spaces related to integrable group representations and their atomic decompositions, II.
Monatsh. Math., 108(2-3):129–148, 1989.

M. Frazier and B. Jawerth.
Decomposition of Besov spaces.
Indiana Univ. Math. J., 34:777–799, 1985.

M. Frazier and B. Jawerth.
A discrete transform and decompositions of distribution spaces.
J. Funct. Anal., 93(1):34–170, 1990.

M. W. Frazier, B. D. Jawerth, and G. Weiss.
Littlewood-Paley Theory and the Study of Function Spaces.
American Mathematical Society, Providence, RI, 1991.

S. M. Nikol'skij.
Approximation of Functions of Several Variables and Imbedding Theorems. Translated from the Russian by J. M. Danskin.
Die Grundlehren der mathematischen Wissenschaften. Band 205. Berlin-Heidelberg-New York: Springer-Verlag. VIII, 420 p. DM 108.00, 1975.

J. Peetre.
New thoughts on Besov spaces.

Duke University Mathematics Series, No. 1. Mathematics Department, Duke University, 1976.

J. Peetre.
Paracommutators and minimal spaces.
In *Operators and Function Theory, Proc NATO Adv Study Inst, Lancaster/Engl 1984, NATO ASI Ser, Ser C 153, 163-224*. 1985.

H.-J. Schmeisser and H. Triebel.
Topics in Fourier Analysis and Function Spaces.
Akademische Verlagsgesellschaft, Leipzig, 1987.

E. M. Stein.
Singular Integrals and Differentiability Properties of Functions.
Princeton University Press, Princeton, N.J., 1970.

E. M. Stein and G. Weiss.
Introduction to Fourier Analysis on Euclidean Spaces.
Princeton University Press, Princeton, N.J., 1971.

M. H. Taibleson.
Fourier Analysis on Local Fields.
Princeton University Press, Princeton, NJ, 1975.

H. Triebel.
Fourier Analysis and Function Spaces. Selected topics.
B. G. Teubner, Leipzig, 1977.

H. Triebel.
Theory of Function Spaces., volume 78 of *Monographs in Mathematics*.
Birkhäuser, Basel, 1983.

H. Triebel.
The Structure of Functions.
Birkhäuser, Basel, 2001.

H. Triebel.

Theory of Function Spaces III, volume 100 of *Monographs in Mathematics*.
Birkhäuser, 2006.