NuHAG Numerical Harmonic Analysis Group

A Banach space version for the kernel theorem describing linear operators

Hans G. Feichtinger, Univ. Vienna & Charles Univ. Prague
hans.feichtinger@univie.ac.at
www.nuhag.eu

Bratislava, 13. Sept. 2018, CSASC-2018

Key aspects of my talk

1. Linear operators and matrices
2. $a_{j,k} = T(\mathbf{e}_k)(\mathbf{e}_j) = \langle \mathbf{e}_j, T(\mathbf{e}_k)\rangle_{\mathbb{C}^m}$;
3. continuous variables: integral operators:

$$Tf(x) = \int_{\mathbb{R}^d} K(x,y)f(y)dy.$$

4. Identity with $K(x,y) = \delta_x(y) = \delta(y-x)$ (!sifting);
5. Can we have $K(x,y) = \delta_y(T(\delta_x))$??
6. indirect description: $T(f)(g) = \int_{\mathbb{R}^{2d}} K(x,y)f(y)g(x)dydx$;
7. the Schwartz kernel theorem using $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ and $\boldsymbol{\mathcal{S}}'(\mathbb{R}^d)$;
8. The Banach Gelfand Triple $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$;
9. The Idea of **Conceptual Harmonic Analysis**

Function Spaces for Fourier Analysis

Most of the mathematical books on Fourier analyis first introduce various Lebesgue spaces, i.e. the space $\left(\boldsymbol{L}^p(\mathbb{R}^d), \|\cdot\|_p\right)$, with their natural norm, for $1 \leq p \leq \infty$.
Among them of course $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$ is the most important, because it is a *Hilbert space* with respect to the usual scalar product

$$\langle f, g \rangle_{\boldsymbol{L}^2} := \int_{\mathbb{R}^d} f(x)\overline{g(x)}dx.$$

For $1 \leq p < \infty$ the dual space can be determined in a similar way as $\boldsymbol{L}^{p'}(\mathbb{R}^d)$, with $1/p + 1/p' = 1$. In particular there is a natural identifcation $(\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right))^* = \left(\boldsymbol{L}^\infty(\mathbb{R}^d), \|\cdot\|_\infty\right)$.

Convolution and Fourier Transforms, Schwartz spaces

There are various natural reasons for using these function spaces. First of all Lebesgue integration rules imply that they are Banach spaces, so in fact one could define them (for $p < \infty$) as *completions* of the space $\boldsymbol{C}_c(\mathbb{R}^d)$ of compactly supported, continuous complex-valued functions functions with respect to the p-norm $(\int_{\mathbb{R}^d} |f(x)|^p dx)^{1/p}$.
For example one can define convolution within $(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1)$:

$$[f * g](x) = \int_{\mathbb{R}^d} g(x-y)f(y)dy, \quad x \in \mathbb{R}^d \ a.e.;$$

and satisfies

$$\|f * g\|_{\boldsymbol{L}^1} \leq \|f\|_{\boldsymbol{L}^1}\|g\|_{\boldsymbol{L}^1}, \quad f, g \in \boldsymbol{L}^1(\mathbb{R}^d).$$

Notations and Conventions

Let us collect here the normalizations of the Fourier transform and relevant transformations of function spaces.

$$\hat{f}(\omega) = \int_{\mathbb{R}^d} f(t) \cdot e^{-2\pi i \omega \cdot t} \, dt. \qquad (1)$$

The inverse Fourier transform (resp. Fourier *synthesis*) then has the form

$$f(t) = \int_{\mathbb{R}^d} \hat{f}(\omega) \cdot e^{2\pi i t \cdot \omega} \, d\omega, \qquad (2)$$

which is valid at least for those continuous, integrable functions which have a Fourier transform $\widehat{f} \in \boldsymbol{L}^1(\mathbb{R}^d)$.

Fourier Transform and Estimates

There are some simple estimates, the so-called *Riemann-Lebesgue Lemma*, with

$$\|\widehat{f}\|_\infty \leq \|f\|_1, \quad f \in \boldsymbol{L}^1(\mathbb{R}^d),$$

and *Plancherel's Theorem* describes the FT as a *unitary* operator on $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$:

$$\|f\|_{\boldsymbol{L}^2} = \|\widehat{f}\|_{\boldsymbol{L}^2}, \quad f \in \boldsymbol{L}^2(\mathbb{R}^d),$$

although in this context the inverse Fourier transform cannot be applied in a pointwise (a.e.) sense, using Lebesgue integrals.
By complex interpolation one obtains the Hausdorff-Young theorem, valid for any $p \in [1,2]$ (only):

$$\|\widehat{f}\|_{\boldsymbol{L}^{p'}} \leq \|f\|_p, \quad f \in \left(\boldsymbol{L}^p(\mathbb{R}^d), \|\cdot\|_p\right).$$

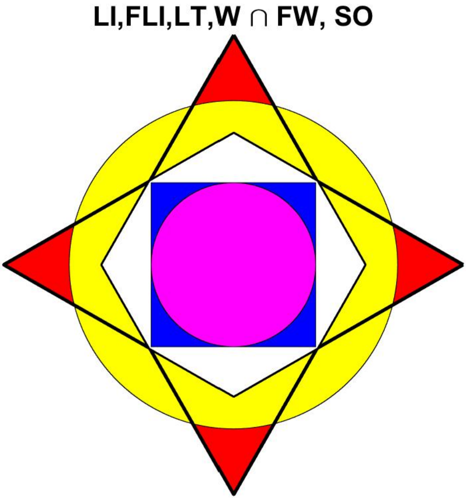
Disadvantages of Lebesgue spaces

One of the disadvantages of the scale of spaces $(\boldsymbol{L}^p(\mathbb{R}^d), \|\cdot\|_p)$, $1 \leq p \leq \infty$ is the fact that these space are not contained in each other, for any pair of different parameters!

As mentioned the inverse Fourier transform cannot be described by Lebesgue integration in $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ (the image of $\boldsymbol{L}^1(\mathbb{R}^d)$ within $\boldsymbol{C}_0(\mathbb{R}^d)$) or $\boldsymbol{L}^2(\mathbb{R}^d)$. In distribution theory the Schwartz space

$\mathcal{S}(\mathbb{R}^d)$ of *rapidly decreasing* functions is the starting points (any partial derivative of $f \in \mathcal{S}(\mathbb{R}^d)$ is decaying faster than any inverse of a polynomial). It is a *nuclear Frechet space* with the natural system of semi-norms.
The dual space (continuous linear functionals) $\mathcal{S}'(\mathbb{R}^d)$ is called the (Schwartz) space of *tempered distributions*.

Fourier transform for $W(L^1, \ell^2)(\mathbb{R}^d)$



BUPUs: Bounded Uniform Partitions of Unity

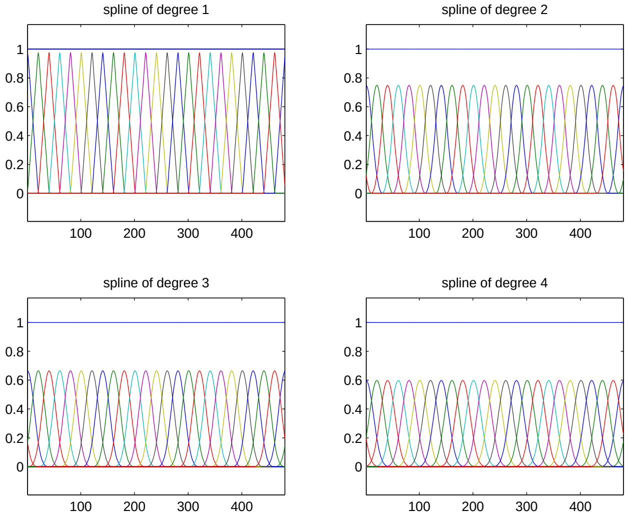


Figure: B-spline BUPUs of variable order 1, 2, 3, 4

The relevant space: Wiener's Algebra

It has turned out (there is meanwhile a long list of publications on the subject) that the most natural and simple condition on φ which allows to provide such estimates is in terms of Wiener's algebra $\left(\boldsymbol{W}(\boldsymbol{C}_0, \ell^1)(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{W}}\right)$.
This space (of bounded and continuous) functions on $\mathbb{R}^d$ can be described roughly as the linear space of all *absolutely Rieman integrable functions*, resp. the space of all continuous functions with finite upper Riemannian sum.

A sufficient condition for a continuous function f on $\mathbb{R}^d$ is:

$$|f(x)| \leq C(1+|x|)^{-(d+\varepsilon)}, \quad x \in \mathbb{R}^d.$$

The relevant space: Wiener's Algebra II

Among the main reasons, why Wiener's algebra is so important, we can identify these two most important ones:

1. The **atomic decomposition**: Every $f \in \boldsymbol{W}(\boldsymbol{C}_0, \ell^1)$ is the absolutely convergent sum of functions (in $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$) of functions with support in sets of the form of $x_n + Q$ (e.g. in the unit cube $Q = [0,1]^d$);
2. The **convolution relations** between the more general *Wiener amalgam spaces* and Wiener's algebra, e.g.

$$\boldsymbol{W}(\boldsymbol{M}, \ell^p) * \boldsymbol{W}(\boldsymbol{C}_0, \ell^1) \subset \boldsymbol{W}(\boldsymbol{C}_0, \ell^p).$$

Recalling the concept of Wiener Amalgam Spaces

Wiener amalgam spaces are a generally useful family of spaces with a wide range of applications in analysis. The main motivation for the introduction of these spaces came from the observations that the non-inclusion results between spaes $\left(\boldsymbol{L}^p(\mathbb{R}^d), \|\cdot\|_p\right)$ for different values of p are either of *local* or of *global* nature. Hence it makes sense to separate these to properties using BUPUs.

Definition

A bounded family $\Psi = (\psi_n)_{n\in\mathbb{Z}^d}$ in some Banach algebra $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ of continuous functions on $\mathbb{R}^d$ is called a regular **Uniform Partition of Unity** if $\psi_n = T_{\alpha n}\psi_0, n \in \mathbb{Z}^d$, $0 \leq \psi_0 \leq 1$, for some ψ_0 with compact support, and

$$\sum_{n\in\mathbb{Z}^d} \psi_n(x) = \sum_{n\in\mathbb{Z}^d} \psi(x - \alpha n) = 1 \quad \text{for all} \quad x \in \mathbb{R}^d.$$

Added in May 2018, hgfei

The definition of general **Wiener amalgam spaces** (originally called Wiener-type spaces, when introduced in 1980, see [3]) with global component ℓ^q is the following one. Assume that a $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ is a Banach space of (locally integrable) functions or distributions such that the action of the elements of the BUPU is uniformly bounded:

$$\|\psi_n \cdot f\|_{\boldsymbol{B}} \leq C_\Psi \|f\|_{\boldsymbol{B}}, \quad \forall f \in \boldsymbol{B}. \tag{3}$$

Definition

$$\boldsymbol{W}(\boldsymbol{B}, \ell^q) := \{f \in \boldsymbol{B}_{loc} \mid \|f\|_{\boldsymbol{W}(\boldsymbol{B},\ell^q)} := (\sum_{k \in \mathbb{Z}^d} \|\psi_n \cdot f\|_{\boldsymbol{B}}^q)^{1/q} < \infty\}$$

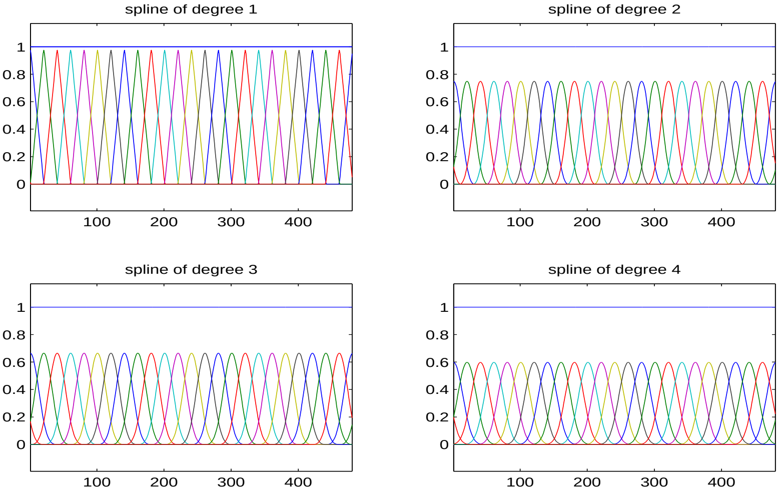
The usual boundedness

For the case of that $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}}) = \left(\boldsymbol{L}^p(\mathbb{R}^d), \|\cdot\|_p\right)$ it is sufficient to assume that the BUPU (ψ_n) is bounded in the pointwise sense, e.g. that

$$0 \leq \psi_n(x) \leq 1, \quad \forall x \in \mathbb{R}^d, \forall n \geq .1$$

For spaces $\boldsymbol{B}$ describing some smoothness it is typically a good idea to assume that $\psi = \psi_0$ belongs to some $C^{(k)}$ space of k times continuously differentiable functions.
Finally the case $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}}) = \left(\mathcal{F}\boldsymbol{L}^p(\mathbb{R}^d), \|\cdot\|_p\right)$ is of great interest because it opens up the way to the definition of modulation spaces (spaces which are of the form $\boldsymbol{W}(\mathcal{F}\boldsymbol{L}^p, \ell^q)$ on the Fourier transform side). Since $\boldsymbol{L}^1 * \boldsymbol{L}^p \subset \boldsymbol{L}^p$ for $1 \leq p \leq \infty$ (together with the corresponding norm inequalities) it is enough assume that $\psi = \psi_0$ belongs to $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$, because translation is isometric in $\left(\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_{\mathcal{F}\boldsymbol{L}^1}\right)$!

Illustration of the B-splines providing BUPUs



Recalling the concept of Wiener Amalgam Spaces II

Note that one can **define the Wiener amalgam space** $\boldsymbol{W}(\boldsymbol{B}, \boldsymbol{\ell}^q)$ by the condition that the sequence $\|f\psi_n\|_{\boldsymbol{B}}$ belongs to $\boldsymbol{\ell}^q(\mathbb{Z}^d)$ and its norm is one of the (many equivalent) norms on this space.

Different BUPUs define the same space and equivalent norms. Moreover, for $1 \leq q \leq \infty$ one has Banach spaces, with natural inclusion, duality and interpolation properties.
Many known function spaces are also Wiener amalgam spaces:

- $\boldsymbol{L}^p(\mathbb{R}^d) = \boldsymbol{W}(\boldsymbol{L}^p, \boldsymbol{\ell}^p)$, same for weighted spaces;
- $\mathcal{H}_s(\mathbb{R}^d)$ (the Sobolev space) satisfies the so-called ℓ^2-puzzle condition (P. Tchamitchian): $\mathcal{H}_s(\mathbb{R}^d) = \boldsymbol{W}(\mathcal{H}_s, \boldsymbol{\ell}^2)$, and consequently for $s > d/2$ (Sobolev embedding) the pointwise multipliers (V. Mazya) equal $\boldsymbol{W}(\mathcal{H}_s, \ell^\infty)$.

Minimality of Wiener's algebra

The Wiener amalgam spaces are essentially a generalization of the original family $\boldsymbol{W}(\boldsymbol{L}^p, \ell^q)$, with local component $\boldsymbol{L}^p$ and global q-summability of the sequence of local $\boldsymbol{L}^p$ norms.
In contrast to the "scale" of spaces $\left(\boldsymbol{L}^p(\mathbb{R}^d), \|\cdot\|_p\right), 1 \leq p \leq \infty$ which do *not allow for any non-trivial inclusion relations* we have nice (and strict) inclusion relations for $p_1 \geq p_2$ and $q_1 \leq q_2$:

$$\boldsymbol{W}(\boldsymbol{L}^{p_1}, \ell^{q_1}) \subset \boldsymbol{W}(\boldsymbol{L}^{p_2}, \ell^{q_2}).$$

Hence $\boldsymbol{W}(\boldsymbol{L}^\infty, \ell^1)$ is the smallest among them, and $\boldsymbol{W}(\boldsymbol{L}^1, \ell^\infty)$ is the largest among them. The closure of the space of test functions, or also of $\boldsymbol{C}_c(\mathbb{R}^d)$ in $\boldsymbol{W}(\boldsymbol{L}^\infty, \ell^1)$ is just *Wiener's algebra* $\left(\boldsymbol{W}(\boldsymbol{C}_0, \ell^1)(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{W}}\right)$, which was one of Hans Reiter's list *Segal algebras*. It can also be characterized as the smallest of all *solid* Segal algebras.

Time and Frequency Shifts

$$[T_t f](x) = f(x-t), \quad x,t \in \mathbb{R}^d; \tag{4}$$

$$[M_\omega f](x) = e^{2\pi i \omega \cdot x} f(x) \quad x,\omega \in \mathbb{R}^d. \tag{5}$$

These operators show the following behavior under the FT

$$(T_x f)\widehat{\ } = M_{-x}\hat{f} \qquad (M_\omega f)\widehat{\ } = T_\omega \hat{f} \tag{6}$$

Combined, applying *first* the time-shift and *then* the frequency shift we get the TF-shifts for $\lambda = (t,\omega) \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d$:

$$[\pi(\lambda)f](x) = M_\omega T_t f(x) = e^{2\pi i \omega \cdot t} f(x-t). \tag{7}$$

Time and Frequency Shifts: on Time and Fourier Side

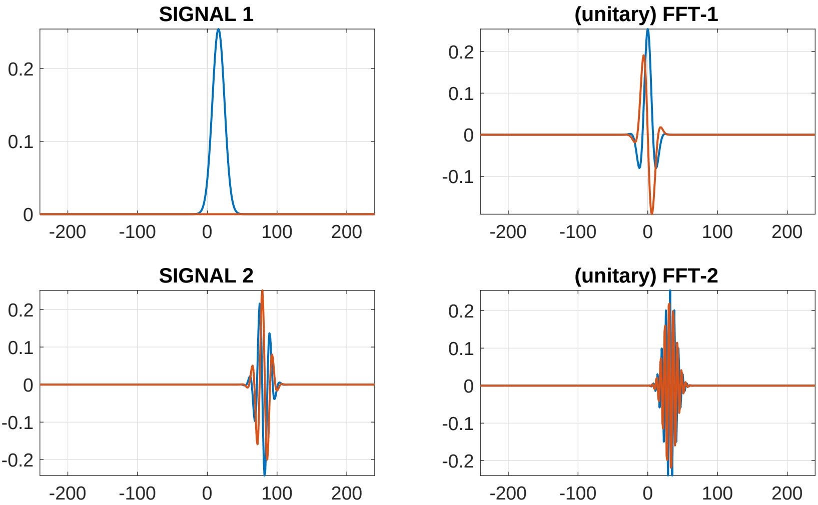


Figure: Demonstrating the effect of shifts on time or frequency side

A Summary of Operator Rules I

Although we will not use the theory based on Lebesgue integration it is still good to know what the standard rules are on the standard spaces, such as $\boldsymbol{L}^1(\mathbb{R}^d)$. We will come back to this space later on.

Operators		
T_z	$T_z f(x) = f(x-z)$	translation by z
M_s	$M_s f(x) = e^{2\pi i s \cdot x} f(x)$	modulation operator
St_ρ	$\mathrm{St}_\rho f(x) = \rho^{-d} f(x/\rho)$	stretching operator
D_ρ	$\mathrm{D}_\rho f(x) = f(\rho x)$	dilation operator
	$f^{\checkmark}(x) = f(-x)$	flip operator
	$f^*(x) = \overline{f(-x)}$	$\boldsymbol{L}^1$-involution
	$\overline{f}(x) = \overline{f(x)}$	conjugation operator

A Summary of Operator Rules II

Translation and modulation are isometric on *all the* $\boldsymbol{L}^p$-spaces, $1 \leq p \leq \infty$. The stretching operator is isometric an $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$, while D_ρ is isometric on $\left(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\right)$ hence on $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$ (or $\left(\boldsymbol{L}^\infty(\mathbb{R}^d), \|\cdot\|_\infty\right)$).

Compatibility of Operators	
$T_z \circ M_s = e^{-2\pi i s \cdot z} M_s \circ T_z$	translation with modulation
$M_s(g * f) = M_s f * M_s g$	modulation and convolution
$T_x(h \cdot f) = T_x h \cdot T_x f$	translation and multiplication
$\mathrm{D}_\rho(h \cdot f) = \mathrm{D}_\rho h \cdot \mathrm{D}_\rho f$	dilation and multiplication
$\mathrm{St}_\rho(g * f) = \mathrm{St}_\rho f * \mathrm{St}_\rho g$	stretching and convolution
$(f * g)^* = g^* * f^*$	convolution and involution
$T_z(f * g) = [T_z f] * g$	convolution and translation
$\overline{h \cdot f} = \overline{h} \cdot \overline{f}$	multiplication and conjugation

A Summary of Operator Rules III

We have for $1 \leq p \leq \infty$

Operators	
$\|T_z f\|_p = \|f\|_p$	translation by z
$\|M_s f\|_p = \|f\|_p$	modulation operator
$\|\mathrm{St}_\rho f\|_1 = \|f\|_1$	stretching operator
$\|\mathrm{D}_\rho f\|_\infty = \|f\|_\infty$	dilation operator

Compatibility with Fourier Transform	
$\mathcal{F} \circ M_s = T_s \circ \mathcal{F}$	translation and Fourier
$\mathcal{F} \circ T_{-x} = M_s \circ \mathcal{F}$	modulation and Fourier
$\mathcal{F} \circ \mathrm{St}_\rho = \mathrm{D}_\rho \circ \mathcal{F}$	stretching and Fourier
$\mathcal{F} \circ \mathrm{D}_\rho = \mathrm{St}_\rho \circ \mathcal{F}$	stretching and Fourier
$\mathcal{F}(f^*) = \overline{\hat{f}}.$	involution and Fourier

A Summary of Operator Rules IV

We also have a couple of adjointness relationship (adjoint operators in the sense of the Hilbert space $(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2)$ an its standard scalar product, given by $\langle f, g\rangle = \int_{\mathbb{R}^d} f(x)\overline{g(x)}dx$. For example $T_x' = T_{-x}, M_s' = M_{-s}$, and $\mathrm{D}_\rho{}' = \mathrm{St}_\rho$ resp. (equivalently) $\mathrm{St}_\rho{}' = \mathrm{D}_\rho$.
Sometimes also the $\boldsymbol{L}^2$-isometric dilated version is used (e.g. in *wavelet theory*, which suggest this form of the *scaling opeator*:

$$S_\rho f(z) = \rho^{-d/2} f(z/\rho), \rho \neq 0.$$

Then one has $S_\rho' = S_{1/\rho}$ (adjoint operator), and

$$\|S_\rho f\|_2 = \|f\|_2, \quad \text{and} \quad \mathrm{supp}(S_\rho f) = \rho \cdot \mathrm{supp}(f). \tag{8}$$

A Summary of Operator Rules V

Definition (Banach spaces of continuous functions on $\mathbb{R}^d$)

$$\boldsymbol{C}_b(\mathbb{R}^d) := \{f : \mathbb{R}^d \mapsto \mathbb{C}, \text{ continuous and bounded, } \ldots$$

$$\text{with norm } \|f\|_\infty = sup_{x\in\mathbb{R}^d}|f(x)| \}$$

The spaces $\boldsymbol{C}_{ub}(\mathbb{R}^d)$ and $\boldsymbol{C}_0(\mathbb{R}^d)$ are defined as the subspaces of $\boldsymbol{C}_b(\mathbb{R}^d)$ consisting of functions which are *uniformly continuous* (and bounded) resp. *decaying at infinity*, i.e.

$$f \in \boldsymbol{C}_0(\mathbb{R}^d) \quad \text{if and only if} \quad \lim_{|x|\to\infty} |f(x)| = 0.$$

The Short-Time Fourier Transform

The Short-Time Fourier Transform

$$V_g f(\lambda) = \langle f, M_\omega T_t g \rangle = \langle f, \pi(\lambda) g \rangle = \langle f, g_\lambda \rangle, \;\; \lambda = (t, \omega);$$

We also need dilation operators:

$$[\mathrm{St}_\rho g](x) = \rho^{-d} g(x/\rho), \quad \rho \neq 0, \tag{9}$$

and the value preserving *dilation operator*

$$[\mathrm{D}_\rho h](x) = h(\rho x), \quad \rho \neq 0. \tag{10}$$

Creating Dirac sequences

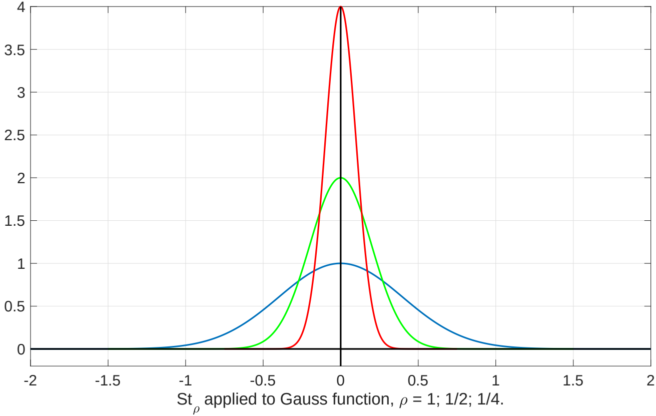


Figure: The stretching operator applied to a standard Gauss function, with "compression" factors of 1 (blue),1/2 (green),1/4 (red).

Summability kernels by dilation

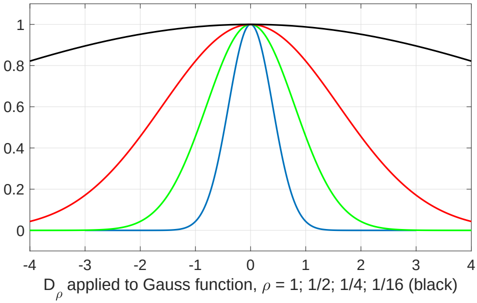


Figure: Dilation corresponding to this on the Fourier transform side, for $\rho \to 0$, exactly: $\rho = 1$ (blue), $1/2$ (green), $1/4$ (red), $1/16$ (black).

The Main Subject of the Course

The main subject of this course will be a triple of Banach spaces, namely $(\mathbf{S}_0, \mathbf{L}^2, \mathbf{S}_0')(\mathbb{R}^d)$, or a so-called **Banach Gelfand Triple** or *rigged Hilbert space*, because it is the (usual) Hilbert space $(\mathbf{L}^2(\mathbb{R}^d), \|\cdot\|_2)$, "surrounded" by a pair of spaces, namely the Banach space of (continuous and Riemann integrable) *test functions* $(\mathbf{S}_0(\mathbb{R}^d), \|\cdot\|_{\mathbf{S}_0})$ and it dual $(\mathbf{S}_0'(\mathbb{R}^d), \|\cdot\|_{\mathbf{S}_0'})$. Thus

$$\mathbf{S}_0(\mathbb{R}^d) \hookrightarrow \mathbf{L}^2(\mathbb{R}^d) \hookrightarrow \mathbf{S}_0'(\mathbb{R}^d) \qquad (11)$$

with two continuous embeddings, and density of $\mathbf{S}_0(\mathbb{R}^d)$ in $(\mathbf{L}^2(\mathbb{R}^d), \|\cdot\|_2)$ and w^*-density of $\mathbf{S}_0(\mathbb{R}^d)$ or $\mathbf{L}^2(\mathbb{R}^d)$ in $\mathbf{S}_0'(\mathbb{R}^d)$, i.e. for any $\sigma \in \mathbf{S}_0'(\mathbb{R}^d)$ there exists a sequency of test functions (h_n) in $\mathbf{S}_0(\mathbb{R}^d)$, such that for any given $g \in \mathbf{S}_0(\mathbb{R}^d)$ on has

$$\int_{\mathbb{R}^d} g(x) h_n(x) dx \to \sigma(g), \quad \text{for } n \to \infty.$$

The Overall Perspective

We could give longer courses on the following goals:

- Motivate the necessity (originally coming from applications) of allowing objects which are not “proper functions”, like the *so-called* **Dirac function** $\delta(t)$ or δ_0.
- Go through the technicalities of topological vector spaces and explain the concept of $\mathcal{S}'(\mathbb{R}^d)$, the *tempered distributions* and then work within that larger reservoir;
- Doing things from scratch and provide all the *functional analytic* details we would have a solid basis but would not get far enough to present interesting applications;
- **INSTEAD I plan to provide BACKGROUND information, BASIC FACTS and describe TYPICAL APPLICATION SITUATIONS.**

The Course Structure

I will try to follow roughly the following plan:

1. Provide a list of motivating properties, why do we need Banach algebras of *test functions*;
2. Define then Banach space $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ (and similar spaces) and show its basic properties;
3. Derive the basic properties of the *dual space* $\boldsymbol{S}_0'(\mathbb{R}^d)$;
4. Combine the three spaces to the Banach Gelfand Triple $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$;
5. Show typical application situations, mostly in Fourier Analysis and Gabor Analysis resp. time-frequency analysis (TFA).

Comparison with the Number System I

The trio of "function spaces" can be compared with the trio of number systems (fields of numbers), namely the chain

$$\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C} \tag{12}$$

While there is a obvious distinction in their *appearance* it is also clear how to interpret each of these objects as a subset of the larger ones (e.g. rationals as periodic infinite decimal expressions), and all the computations which can be done at a lower level can be expanded in a natural *unique* way to the larger one.

The best example is multiplication and inversion, think of the number $1/\pi^2$, or the claim that $e^{2\pi i} = 1$. This is not as simple as forming the multiplicative inverse of $3/4$, which is $4/3$ (observe transition from actual to symbolic computation!).

A schematic description: the simplified setting

In our picture this simple means that the inner "kernel" is mapped into the "kernel", the Hilbert space to the Hilbert space, and at the outer level two types of continuity are valid (norm and w^*)!

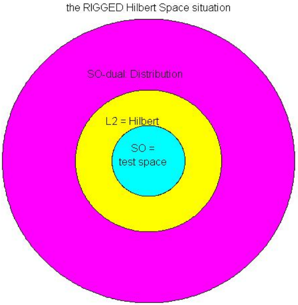


Figure: Compare the situation with $\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$

Comparison with the Number System II

There are many good reasons to **extend** the rational numbers (which already are a field and thus allow for quite a variety of operations) to the field of *real numbers*. It is lack of *completeness* which is the problem with $\mathbb{Q}$. It is easy to find a *Cauchy sequence* of rationals $q_n, n \geq 1$ with the property that $q_n^2 \to 2$ for $n \to \infty$. BUT there is no rational number q such that $q^2 = 2$!
The abstract way which allows to embed each *metric space* into a *complete metric space* (where every Cauchy-sequence has a limit) makes use of equivalence classes of Cauchy-sequences.
In the case of the rational number $\mathbb{Q}$ with the distance $d(q_1, q_2) = |q_1 - q_2|$ each such equivalence class contains (more or less) *exactly one* infinite decimal expression.

Comparison with Fourier Analysis for Engineers I

We will see that the use of certain **symbols**, specifically *integrals* within an engineering context is better understood at the "symbolic level", e.g. the *Fourier inversion formula*.
Let us give an example: Sometimes the validity of the Fourier inversion formula is justified by the (so-called) validity of the following formula

$$\int_{-\infty}^{\infty} e^{2\pi i s t} ds = \delta(t). \tag{13}$$

Such a claim is of course **very problematic to mathematicians** who try to take it literal and object to the existence of the integral on the left hand side as a Lebesgue integral (the best possible one), and the pointwise interpretation of the equality, because the "delta-function" should not be described pointwise.

Comparison with Fourier Analysis for Engineers II

Instead of just discarding the equation (13) as non-sense we can take it as a symbol, but we have to learn to read it properly.

Expressions of the form $\int_{-\infty}^{\infty} h(s)e^{2\pi i s t} ds$ are generally useful and allow us to regain g from its Fourier transform $h = \hat{g}$, given by $\hat{g}(s) = \int_{-\infty}^{\infty} g(t)e^{-2\pi i s t} dt$, at least for (good, i.e.) test functions. In this sense we can read (13) as the claim that $\mathcal{F}^{-1}(\mathbf{1}) = \delta_0$, the inverse Fourier transform of the function contant one is the Dirac delta (distribution or measures).

This sounds reasonable if we assume that the forward or inverse Fourier transform of objects like $\mathbf{1}$ of δ_0 **"exist"**[1], since the convolution theorem suggest that for test functions f one has

$$\delta_0 * f = f \Leftrightarrow \widehat{\delta_0} \cdot \widehat{f} \quad (\text{clearly } = \mathbf{1} \cdot \widehat{f}).$$

[1] Another problematic setting with the danger of drifting into philosophical discussions about the existence of objects!

Comparison with Fourier Analysis for Engineers III

So one of goals of this course will be to *build a bridge between engineering intuition and "symbolic manipulations" and strict mathematical description*, without going too deep into complicated mathematical theory (involving Lebesgue integration, which does not help here, or topological vector spaces, which are used as the foundation to the Schwartz theory of tempered distributions, indicating how they can be replaced to a large extent by Banach space arguments.

Coming back to (13) let us indicate our plan:

First we have to extend the domain of the forward and inverse Fourier transform from the space of test functions to a larger vector space of *generalized functions*. Then we have to show that δ_0 and $\mathbf{1}$ correspond to each other! Finally we can verify the validity of the convolution theorem in this more general context, justifying claim (13) in a different way.

Comparison with Fourier Analysis for Engineers IV

But the fact that the (generalized) inverse Fourier transform has the (necessary) property of bringing **1** back to δ_0 by itself does not guarantee that the classical Fourier inversion formula is giving a description of the inverse mapping to the Fourier transform, if we change (and specifically expand) the domain.
This is like saying, that it is obvious that we have

$$\pi \cdot \frac{1}{\pi^2} \cdot \pi = 1 \quad \text{in } \mathbb{R}.$$

Such a claim is trivial at the symbolic level, but would have to be a bit complicated if realized "numerically" (or constructively).
We can justify formula (13) later on also by verifying that the so-called w^*-w^*-continuity of the extended Fourier transform on $\boldsymbol{S}_0'(\mathbb{R}^d)$ enforces that (13) is not only valid but *characteristic* for the *inverse* (of the) Fourier transform!

What can we learn from the Number Systems

Multiplication and division (correctly interpreted as inversion of the multiplication for non-zero elements!), well defined on $\mathbb{Q}$, can be extended in a very natural way once we know a few things:

1. how to create the generalized objects from the given set of object (e.g. infinite decimal expression viewed as sequences of their approximations with finite precision);
2. how to embed the original structure into to new object, including the algebraic properties (e.g. multiplication, or Euclidean distances) in a *compatible way*![2]
3. show how new objects are approximated by old ones;
4. extend the structures and demonstrate that the extended structure is characterized by these natural properties.

[2]Like $(3/4)^2 = 9/16 = 0.5625 = 0.75^2$.

Papoulis comment on distribution theory

It is very interesting to read to introduction to the original version of A. Papoulis on *The Fourier Transforms and its Applications* (first published in 1962), one of the standard works for applied Fourier Analysis, specifically for Engineers, in the second half of the last century, see [11].
Note that at this time the theory of Schwartz distributions was still quite fresh, that Papoulis argues that it is a powerful but a theory which is too complicated for engineers. Note also that this book has been written shortly before the time the FFT was even invented (by Cooley and Tuckey, see [1]), which clearly has a deep impact on modern (computational) Fourier Analysis.
Papoulis writes in his preface:

Possible way to introduce Generalized Functions 1a

There are two different ways to introduce generalized functions.

The first one is through *equivalence classes of sequences of test functions*, while the second one uses *functional analytic* ideas, i.e. defines the space of *distributions* as a set of linear functionals on some topological vector space. This means one takes all *linear functionals* which respect the convergence (typically describe by families of seminorms on the vector space), i.e. which are continuous. We will follow this second approach, but with a simple Banach space approach, where continuity can be expressed simply by boundedness, the function σ has to satisfy $|\sigma(f)| \leq C\|f\|_{\boldsymbol{S}_0}$ for some $C > 0$ and all $f \in \boldsymbol{S}_0(\mathbb{R}^d)$.

The main advocate of the *sequential approach* is the was J. Ligthill, whose book [9] appeared first in 1958.

Benefits and Problems with the Sequential Approach 1b

It is clear that the sequential approach is modeled after the construction of the real numbers $\mathbb{R}$ from the rationals $\mathbb{Q}$, resp. by applying the general concept of *completion of metric spaces*.
Unfortunately (unlike one has the infinite decimal representation for $\mathbb{R}$) the general situation does not allow to work with a specific *representative* or a unique sequence of test functions, but one has to work effectively with equivalence classes of so-called *regular* (meaning "somehow convergent") sequences.
This makes the handling in this approach quite involved and even for simple (if not almost trivial statements) one has to work hard (or leave the details to the reader, so that she/he is left with a lot of work).

Comments on the Approach by Duality, 2a

Aside from the fact that one has to make use of a few basic principles from the theory of Banach (and perhaps Hilbert) spaces the introduction of generalized functions, or perhaps better *distributions* is to *define them* as linear spaces of linear functionals. What is a bit less convenient at first sight is the necessity of embedding ordinary functions into which can be done using the Riemann integral (or Haar meausure), or more generally the Lebesgue integral for the most general examples of *regular distributions* (e.g. bounded measures with *density* in $\boldsymbol{L}^1(\mathbb{R}^d)$).
We define the distribution *induced by a function h* on $\mathbb{R}^d$ via

$$\sigma_h(f) = \int_{\mathbb{R}^d} f(x)\ h(x) dx, \tag{14}$$

i.e. by *integration of the argument* $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ *against h*.

Benefits of the Approach based on Duality, 2b

Aside from the fact that perhaps the view-point that signals *ARE IN SOME SENSE* linear functionals, which can be measured, without having necessarily a pointwise value (what about *room temperature* as a function of time and space coordinates: we can only measure some averages!) and pointwise defined functions are perhaps more of an idealization (compare to concrete linear functionals) one has several advantages from the duality approach.

First of all it is easy to verify *completeness* of the space of distributions. Secondly one has in addition to the norm convergence in the dual space also the so-called w^*-convergence.

We will see that with a couple of basic facts from *linear functional analysis* we can prove quite a few things (partially based on linear algebra considerations).

The key-players for time-frequency analysis

Time-shifts and Frequency shifts

$$T_x f(t) = f(t - x)$$

and $x, \omega, t \in \mathbb{R}^d$

$$M_\omega f(t) = e^{2\pi i \omega \cdot t} f(t) \,.$$

Behavior under Fourier transform

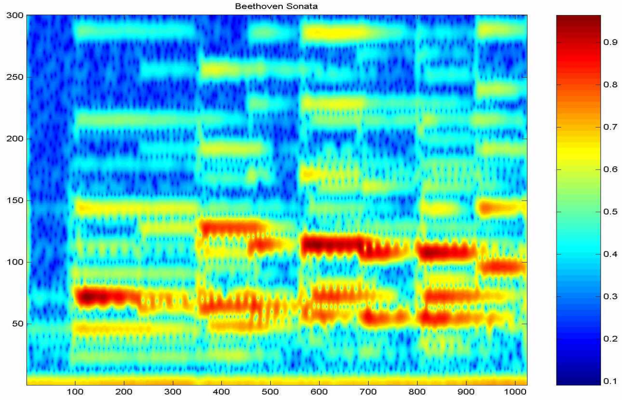
$$(T_x f)\widehat{\ } = M_{-x} \hat{f} \qquad (M_\omega f)\widehat{\ } = T_\omega \hat{f}$$

The Short-Time Fourier Transform

$$V_g f(\lambda) = \langle f, M_\omega T_t g \rangle = \langle f, \pi(\lambda) g \rangle = \langle f, g_\lambda \rangle, \;\; \lambda = (t, \omega);$$

A Typical Musical STFT

A typical waterfall melody (Beethoven piano sonata) depictured using the spectrogram, displaying the energy distribution in the TF = time-frequency plan:



A Banach Space of Test Functions (Fei 1979)

A function in $f \in \boldsymbol{L}^2(\mathbb{R}^d)$ is in the subspace $\boldsymbol{S}_0(\mathbb{R}^d)$ if for some non-zero g (called the "window") in the Schwartz space $\mathcal{S}(\mathbb{R}^d)$

$$\|f\|_{S_0} := \|V_g f\|_{\boldsymbol{L}^1} = \iint_{\mathbb{R}^d \times \widehat{\mathbb{R}}^d} |V_g f(x,\omega)| dx d\omega < \infty.$$

The space $(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0})$ is a Banach space, for any fixed, non-zero $g \in \boldsymbol{S}_0(\mathbb{R}^d))$, and different windows g define the same space and equivalent norms. Since $\boldsymbol{S}_0(\mathbb{R}^d)$ contains the Schwartz space $\mathcal{S}(\mathbb{R}^d)$, any Schwartz function is suitable, but also compactly supported functions having an integrable Fourier transform (such as a trapezoidal or triangular function) are suitable. It is convenient to use the Gaussian as a window.

Basic properties of $M^1 = S_0(\mathbb{R}^d)$

Lemma

Let $f \in \boldsymbol{S}_0(\mathbb{R}^d)$, then the following holds:

(1) $\pi(u,\eta)f \in \boldsymbol{S}_0(\mathbb{R}^d)$ *for* $(u,\eta) \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d$, *and* $\|\pi(u,\eta)f\|_{\boldsymbol{S}_0} = \|f\|_{\boldsymbol{S}_0}$.

(2) $\hat{f} \in \boldsymbol{S}_0(\mathbb{R}^d)$, *and* $\|\hat{f}\|_{\boldsymbol{S}_0} = \|f\|_{\boldsymbol{S}_0}$.

In fact, $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ is the smallest non-trivial Banach space with this property, and therefore contained in any of the $\boldsymbol{L}^p$-spaces (and their Fourier images), for $1 \leq p \leq \infty$, and dense for $p < \infty$. Later on we will make use of the fact that $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ coincides with the Wiener amalgam space $\boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1, \ell^1)(\mathbb{R}^d)$.e In fact it was introduced in this way by the author ([2], see [?]).

Various Function Spaces

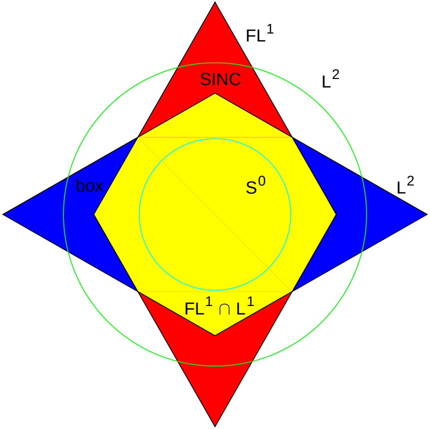


Figure: The usual Lebesgues space, the Fourier algebra, and the Segal algebra $\mathbf{S}_0(\mathbb{R}^d)$ inside all these spaces

BANACH GELFAND TRIPLES: a new category

Definition

A triple, consisting of a Banach space $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$, which is densely embedded into some Hilbert space $\mathcal{H}$, which in turn is contained in $\boldsymbol{B}'$ is called a Banach Gelfand triple.

Definition

If $(\boldsymbol{B}_1, \mathcal{H}_1, \boldsymbol{B}_1')$ and $(\boldsymbol{B}_2, \mathcal{H}_2, \boldsymbol{B}_2')$ are Gelfand triples then a linear operator T is called a [unitary] Gelfand triple isomorphism if

1. T is an isomorphism between $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$.
2. T is [unitary] isomorphism between $\mathcal{H}_1$ and $\mathcal{H}_2$.
3. T extends to a weak* isomorphism as well as a norm-to-norm continuous isomorphism between $\boldsymbol{B}_1'$ and $\boldsymbol{B}_2'$.

A schematic description: the simplified setting

In our picture this simple means that the inner "kernel" is mapped into the "kernel", the Hilbert space to the Hilbert space, and at the outer level two types of continuity are valid (norm and w^*)!

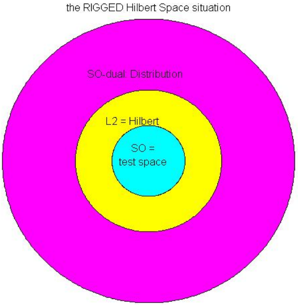


Figure: Compare the situation with $\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$

The prototypical examples over the torus

In principle every CONB (= *complete orthonormal basis*) $\Psi = (\psi_i)_{i \in I}$ for a given Hilbert space $\mathcal{H}$ can be used to establish such a unitary isomorphism, by choosing as $\boldsymbol{B}$ the space of elements within $\mathcal{H}$ which have an absolutely convergent expansion, i.e. satisfy $\sum_{i \in I} |\langle x, \psi_i \rangle| < \infty$.
For the case of the Fourier system as CONB for $\mathcal{H} = \boldsymbol{L}^2([0,1])$, i.e. the corresponding definition is already around since the times of N. Wiener: $\boldsymbol{A}(\mathbb{T})$, the space of absolutely continuous Fourier series. It is also not surprising in retrospect to see that the dual space $\boldsymbol{PM}(\mathbb{T}) = \boldsymbol{A}(\mathbb{T})'$ is space of *pseudo-measures*. One can extend the classical Fourier transform to this space, and in fact interpret this extended mapping, in conjunction with the classical Plancherel theorem as the first unitary Banach Gelfand triple isomorphism, between $(\boldsymbol{A}, \boldsymbol{L}^2, \boldsymbol{PM})(\mathbb{T})$ and $(\ell^1, \ell^2, \ell^\infty)(\mathbb{Z})$.

The Fourier transform as BGT automorphism

The Fourier transform $\mathcal{F}$ on $\mathbb{R}^d$ has the following properties:

1. $\mathcal{F}$ is an isomorphism from $\boldsymbol{S}_0(\mathbb{R}^d)$ to $\boldsymbol{S}_0(\widehat{\mathbb{R}}^d)$,
2. $\mathcal{F}$ is a unitary map between $\boldsymbol{L}^2(\mathbb{R}^d)$ and $\boldsymbol{L}^2(\widehat{\mathbb{R}}^d)$,
3. $\mathcal{F}$ is a weak* (and norm-to-norm) continuous bijection from $\boldsymbol{S}_0'(\mathbb{R}^d)$ onto $\boldsymbol{S}_0'(\widehat{\mathbb{R}}^d)$.

Furthermore, we have that Parseval's formula

$$\langle f, g \rangle = \langle \widehat{f}, \widehat{g} \rangle \tag{15}$$

is valid for $(f, g) \in \boldsymbol{S}_0(\mathbb{R}^d) \times \boldsymbol{S}_0'(\mathbb{R}^d)$, and therefore on each level of the Gelfand triple $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$.

Some concrete computations (M.DeGosson: Wigner Transform)

For $\phi \in \mathcal{S}(\mathbb{R}^n)$ the short-time Fourier transform (STFT) V_ϕ with window ϕ is defined, for $\psi \in \mathcal{S}'(\mathbb{R}^n)$, by

$$V_\phi \psi(z) = \int_{\mathbb{R}^n} e^{-2\pi i p\cdot x'} \psi(x') \overline{\phi(x'-x)} \mathrm{d}x'. \qquad (16)$$

The STFT is related to a well-known object from quantum mechanics, the cross-Wigner transform $W(\psi, \phi)$, defined by

$$W(\psi,\phi)(z) = \left(\tfrac{1}{2\pi\hbar}\right)^n \int_{\mathbb{R}^n} e^{-\frac{i}{\hbar} p\cdot y} \psi(x+\tfrac{1}{2}y)\overline{\phi(x-\tfrac{1}{2}y)} \mathrm{d}y. \qquad (17)$$

In fact, a tedious but straightforward calculation shows that

$$W(\psi,\phi)(z) = \left(\tfrac{2}{\pi\hbar}\right)^{n/2} e^{\frac{2i}{\hbar} p\cdot x} V_{\phi^{\checkmark}_{\sqrt{2\pi\hbar}}} \psi_{\sqrt{2\pi\hbar}}\left(z\sqrt{\tfrac{2}{\pi\hbar}}\right) \qquad (18)$$

where $\psi_{\sqrt{2\pi\hbar}}(x) = \psi(x\sqrt{2\pi\hbar})$ and $\phi^{\checkmark}(x) = \phi(-x)$;

This formula can be reversed to yield:

$$V_\phi\psi(z) = \left(\tfrac{2}{\pi\hbar}\right)^{-n/2} e^{-i\pi p\cdot x} W(\psi_{1/\sqrt{2\pi\hbar}}, \phi^\vee_{1/\sqrt{2\pi\hbar}})(z\sqrt{\tfrac{\pi\hbar}{2}}). \quad (19)$$

In particular, taking $\psi = \phi$ one gets the following formula for the usual Wigner transform:

$$W\psi(z) = \left(\tfrac{2}{\pi\hbar}\right)^{n/2} e^{\frac{2i}{\hbar}p\cdot x} V_{\psi_1}(\psi_2)(z\sqrt{\tfrac{2}{\pi\hbar}}).$$

with $\psi_1 = \psi^\checkmark_{\sqrt{2\pi\hbar}}$ and $\psi_2 = \psi_{\sqrt{2\pi\hbar}}$.

Another reference is the book of K. Gröchenig [7], which contains (in the terminology used there) in Lemma 4.3.1 the following formula, using the convention $g^{\checkmark}(x) = g(-x)$:

$$W(f,g)(x,\omega) = 2^d e^{4\pi i x\omega} V_{g^{\checkmark}} f(2x, 2\omega). \tag{20}$$

Charly (in [7]) also provides the following *covariance property*

$$W(T_u M_\eta f) = Wf(x-u, \omega-\eta). \tag{21}$$

$$W(\hat{f}, \hat{g})(x,\omega) = W(f,g)(-\omega, x). \tag{22}$$

The Wigner approach to $\boldsymbol{S}_0(\mathbb{R}^d)$ is also used in [8].

Usefulness of $S_0(\mathbb{R}^d)$ in Fourier Analysis

Most consequences result form the following inclusion relations:

$$\boldsymbol{L}^1(\mathbb{R}^d) * \boldsymbol{S}_0(\mathbb{R}^d) \subseteq \boldsymbol{S}_0(\mathbb{R}^d); \tag{23}$$

$$\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) \cdot \boldsymbol{S}_0(\mathbb{R}^d) \subseteq \boldsymbol{S}_0(\mathbb{R}^d); \tag{24}$$

$$(\boldsymbol{S}_0'(\mathbb{R}^d) * \boldsymbol{S}_0(\mathbb{R}^d)) \cdot \boldsymbol{S}_0(\mathbb{R}^d) \subseteq \boldsymbol{S}_0(\mathbb{R}^d); \tag{25}$$

$$(\boldsymbol{S}_0'(\mathbb{R}^d) \cdot \boldsymbol{S}_0(\mathbb{R}^d)) * \boldsymbol{S}_0(\mathbb{R}^d) \subseteq \boldsymbol{S}_0(\mathbb{R}^d); \tag{26}$$

$$\boldsymbol{S}_0(\mathbb{R}^d) \widehat{\otimes} \boldsymbol{S}_0(\mathbb{R}^d) = \boldsymbol{S}_0(\mathbb{R}^{2d}). \tag{27}$$

1. $\boldsymbol{S}_0(\mathbb{R}^d)$ is a valid domain of Poisson's formula;
2. all the classical Fourier summability kernels are in $\boldsymbol{S}_0(\mathbb{R}^d)$;
3. the elements $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ are the natural building blocks for Gabor expansions;

The Banach Gelfand Triple

The Banach Gelfand Triple $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$ is for many applications in theoretical physics and engineering, but also for *Abstract Harmonic Analysis* a good replacement for the Schwartz Gelfand Triple $(\boldsymbol{\mathcal{S}}, \boldsymbol{L}^2, \boldsymbol{\mathcal{S}}')$.

Lemma

$$(\boldsymbol{S}_0' * \boldsymbol{S}_0) \cdot \boldsymbol{S}_0 \subseteq \boldsymbol{S}_0, \quad (\boldsymbol{S}_0' \cdot \boldsymbol{S}_0) * \boldsymbol{S}_0 \subseteq \boldsymbol{S}_0, \tag{28}$$

Clearly $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ is a Banach space and NOT a *nuclear Frechet* space, but still there is a kernel theorem!
The main exception are applications to PDE where $\boldsymbol{S}_0(\mathbb{R}^d)$ is not well suited, but there is a family of so-called *modulation spaces* which allows also to overcome this problem, and even go for the theory of *ultra-distributions*, putting weighted $\boldsymbol{L}^1$-norms on the STFT (see [7] for a first glimpse!).

A large variety of characterizations

There is a large variety of characterizations of $\left(\mathbf{S}_0(\mathbb{R}^d), \|\cdot\|_{\mathbf{S}_0}\right)$ and $(\mathbf{S}_0'(\mathbb{R}^d), \|\cdot\|_{\mathbf{S}_0'})$ (see e.g. [8]).
For example, a tempered distribution in $\mathcal{S}'(\mathbb{R}^d)$ belongs to $\mathbf{S}_0'(\mathbb{R}^d)$ if and only if its STFT (well defined for $g \in \mathcal{S}(\mathbb{R}^d)$!) is a bounded function. Norm convergence is equivalent to uniform convergence of spectrograms, while w^*-convergence (!very important) corresponds to *uniform convergence over compact sets* of the corresponding STFTs. It is again independent of the choice of the window, even any non-zero $g \in \mathbf{S}_0(\mathbb{R}^d)$ can be used here.
There are *atomic characterizations*, or characterizations via *Wiener amalgam spaces*, for example

$$\mathbf{S}_0(\mathbb{R}^d) = \mathbf{W}(\mathcal{F}\mathbf{L}^1, \ell^1)(\mathbb{R}^d).$$

The Space $S_0'(\mathbb{R}^d)$ of distributions

In this section we will show that the dual space $S_0'(\mathbb{R}^d)$ is a quite natural object, and that the Fourier transform can be extended in a unique and natural way to $S_0'(\mathbb{R}^d)$, using w^*-convergence. Since the space $\left(S_0(\mathbb{R}^d), \|\cdot\|_{S_0}\right)$ is *separable* the restriction to sequences is in fact well justified (as opposed to convergence of *nets* or *filters* in general topological vector spaces).
First of all we start with a trivial remark: A linear functional $\sigma : f \mapsto \sigma(f)$ from $\left(S_0(\mathbb{R}^d), \|\cdot\|_{S_0}\right)$ into $(\mathbb{C}, |\cdot|)$ is continuous if and only if it is bounded. In other words, it satisfies

$$\|f_n - f_0\|_{S_0} \to 0 \text{ for } n \to \infty \quad \Rightarrow \quad \sigma(f_n - f_0) \to 0 \text{ in } \mathbb{C}$$

if and only if there exists $C > 0$ such that

$$|\sigma(f)| \leq C\|f\|_{S_0} \quad \forall f \in S_0(\mathbb{R}^d). \tag{29}$$

The Space $S_0'(\mathbb{R}^d)$ of distributions

The minimal constant can be also characterized as

$$\|\sigma\|_{S_0'} = \sup_{f:\|f\|_{S_0}\leq 1} \{|\sigma(f)|\}. \tag{30}$$

Making use of the atomic characterization of $S_0(\mathbb{R}^d)$ one can show:

Theorem

For any nonzero $g \in S_0(\mathbb{R}^d)$ the S_0'-norm is equivalent to the sup-norm $\|V_g(\sigma)\|_\infty$, in other words: Norm convergence in $(S_0'(\mathbb{R}^d), \|\cdot\|_{S_0'})$ is the same as uniform convergence at the spectrogram level.

Again using Wiener amalgams suggests (correctly) to identify the dual space as $S_0' = W(\mathcal{F}L^1, \ell^1)' = W(\mathcal{F}L^\infty, \ell^\infty) \supset W(M, \ell^\infty)$.

w^*-convergence in $S_0'(\mathbb{R}^d)$

Thm. 8 suggests to look for a weaker concept of convergence compared to norm convergence, because it will never be possible to e.g. approximated a periodic function by compactly supported ones, even if the norm is a relatively weak norm compared to the ordinary $\boldsymbol{L}^p$-norms.

The answer is of course provided by the w^*-convergence.

Recall that $\sigma_0 = w^*\text{-}\lim_{n\to\infty}\sigma_n$ if and only if

$$\lim_{n\to\infty} \sigma_n(f) = \sigma_0(f), \quad \forall f \in \boldsymbol{S}_0. \tag{31}$$

As a first observation note that $\|\delta_{1/n} - \delta_0\|_{\boldsymbol{S}_0'} = 2$ while $\delta_0 = w^*\text{-}\lim_{n\to\infty}\delta_{x_n}$ for any sequence $x_n \to 0$ for $n \to \infty$.

Characterizing w^*-convergence and approximation

Note that a bounded sequence $(\sigma_n)_{n\geq 1}$ in $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$ is w^*-convergent if and only if convergence takes place on a dense or just total subset of $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$.
Thus for example is it enough to test the validity of (31) for any compactly supported function f with $\hat{f} \in \boldsymbol{L}^1(\mathbb{R}^d)$, or for all the band-limited functions in $f \in \boldsymbol{L}^1(\mathbb{R}^d)$.
The theory of Gabor frames on the other hand implies that it is enough to verify pointwise convergence of the STFT with respect to the Gaussian window $g_0(t) = e^{-\pi|t|^2}$ for all the lattice points of any fixed lattice Λ of the form $\Lambda = a\mathbb{Z}^d \times b\mathbb{Z}^d$ with $a \cdot b < 1$, i.e.

$$V_{g_0}(\sigma_n)(\lambda) \to V_{g_0}(\sigma_0)(\lambda) \quad \text{for } n \to \infty.$$

Comment: A closed subset of $\boldsymbol{S}_0(\mathbb{R}^d)$ is compact if and only if the convergence takes place uniformly in $\ell^1(\Lambda)$.

Interesting examples of w^*-convergence

It is often the w^*-convergence (sometimes appearing in disguise) which is used for handwaving arguments in Fourier Analysis.

1. One has $lim_{\alpha\to\infty} \sqcup\!\sqcup_\alpha = \delta_0$ (as is easily verified by applying it to compactly supported functions in $\boldsymbol{S}_0(\mathbb{R}^d)$);
2. The absolute Riemann-integrability of $f \in \boldsymbol{S}_0(\mathbb{R}^d) \subset \boldsymbol{W}(\boldsymbol{C}_0, \ell^1)(\mathbb{R}^d)$ implies that $lim_{\beta\to 0}\beta^d \sqcup\!\sqcup_\beta = \mathbf{1}$;
3. For any $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ one has

$$lim_{\alpha\to\infty} \sqcup\!\sqcup_\alpha * g = g.$$

4. The same relation on the Fourier transform (with $\beta = 1/\alpha$) is used to explain the form of the continuous Fourier transform (by letting the "period go to infinity").

How can we DEFINE Generalized Functions

The theory of *generalized functions* is clearly supposed to allow certain "objects" which are *beyond the scope of the usual concept of a pointwise well-defined functions* f (or $f(t)$ as engineers would write in order to emphasize the character of the domain of f).
The Dirac "function" $\delta(t)$ (engineering way of writing) is an example, and is usually described as the *limit of a sequency of box-functions*, with shrinking basis (to zero), and constant area 1.
In general there are two ways of defining linear spaces of generalized function, or we will call them "distributions"[3]

[3] Even if they are not distributions in the classical setting, e.g. because they are defined over LCA groups, without reference to differentiability.

Overall Perspective, Versatile Tools

When it comes to applications it is like real life: We would like to have the most universal and reliable tool at for a good price.
Translated into the scientific world: Even if *mathematicians* are willing to create the most *complicated and and fancy tools* these tools might not be used by other (more applied) scientist. If they are lucky they may receive great respect, but this does not mean that the applied scientist have the patience or skill or just willingness to learn and then use such a tools.
Of course *sometimes* only complicated tools do the job and one needs the top experts and *specialists* to tackle those few problems, but the daily life one should ideally have a good equipment helping the users to solve their problems themselves.

Regularization of Distributions I

We have claimed in the introduction that distributions can be approximated by test functions. In fact a good example is $\sqcup\!\sqcup = \sqcup\!\sqcup_{\mathbb{Z}^d} = \sum_{k \in \mathbb{Z}^d} \delta_k$, which is a well defined element of $\boldsymbol{S}_0'(\mathbb{R}^d)$ (it is even Fourier invariant, according to Poisson's formula!).
It cannot be viewed as a regular distributions coming from any possible test functions because it has *two defects*

- First of all it is not a continuous function, because it is a sum of Dirac measures;
- Secondly it does not have decay at infinity, since all the involved Dirac measures have the same coefficient 1.

Whenever we want to approximate (in fact in the w^*-sense) we have to improve both the t local and the *global* properties of the distribution!

Regularization of Distributions II

There are various ways to improve the local and global properties. Typically it is *convolution* by a test function which helps to improve the *local properties* while pointwise multiplication by a test function improves the decay at infinity, i.e. the *global properties*. Let us therefore recall the two version of the dilation operatot that will be useful for this purpose. One is the $\boldsymbol{L}^1$-norm preserving, where the index describes the shrinkage or expansion of the support, also *stretching* operator (for $\rho > 1$):

$$[\mathrm{St}_\rho g](x) = \rho^{-d} g(x/\rho), \quad \rho \neq 0, \tag{32}$$

and the value preserving *dilation operator*

$$[\mathrm{D}_\rho h](x) = h(\rho x), \quad \rho \neq 0. \tag{33}$$

Regularization of Distributions III

While the first is compatible with the structure of the Banach convolution algebra $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ the second is compatible with the pointwise structure of $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$ or $\left(\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_{\mathcal{F}\boldsymbol{L}^1}\right)$ (the Fourier algebra). We have in particular

$$\|\mathrm{St}_\rho f\|_1 = \|f\|_1 \quad \text{and} \quad \|\mathrm{D}_\rho h\|_\infty = \|h\|_\infty .$$

$$\mathrm{St}_\rho(g * f) = \mathrm{St}_\rho(g) * \mathrm{St}_\rho(f)$$

$$\mathrm{D}_\rho(h \cdot f) = \mathrm{D}_\rho(h) \cdot \mathrm{D}_\rho(f).$$

Of course $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ is invariant with respect to any automorphism of the underlying group, so in particular with respect to both of these (commutative) *dilation groups*, but of course not in the isometric sense (like St_ρ on $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ and D_ρ on $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$).

Regularization of Distributions IV

The approximation of distributions requires the application of both of these *regularizes*, in any order.
So let us look at the Product-Convolution (short: PC) operator

$$\sigma \mapsto \mathrm{St}_\rho g * (\mathrm{D}_\rho h \cdot \sigma), \quad \text{for } \rho \to 0.$$

Here $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ should satisfy $\int_{\mathbb{R}^d} g(x)dx = 1$, while $h \in \boldsymbol{S}_0(\mathbb{R}^d)$ is has to satisfy the condition $h(0) = 1$. Since $\widehat{\mathrm{St}_\rho f} = \mathrm{D}_\rho \widehat{f}$ one could be the (inverse) FT of the other. In a similar way one has Convolution-Product (CP) operators of the form

$$\sigma \mapsto \mathrm{D}_\rho h \cdot (\mathrm{St}_\rho g * \sigma), \quad \text{for } \rho \to 0,$$

with the same conditions on g and h in $\boldsymbol{S}_0(\mathbb{R}^d)$.

Regularization of Distributions V

So we finally just have to verify that these operators map in fact (for fixed) $\rho \neq 0$ the space $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$ into $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$, even in the sense that a w^*-convergent and bounded sequence (or net) in $\boldsymbol{S}_0'(\mathbb{R}^d)$ with $w^*\text{-}\lim \sigma_n = \sigma_0$ is mapped into a norm convergent sequence within $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$.
Note that these operators act uniformly bounded (w.r.t. ρ) on each of the spaces $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$, $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$ and $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$!!
A similar behaviour (we call it a *regularizing sequence* for the Banach Gelfand Triple $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d))$ can be verified for the partial sum operator for a Gabor expansion, with Gabor atom $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ and canonical dual (or minimal ℓ^2-norm coefficients) $\widetilde{g}$ (which also automatically belongs to $\boldsymbol{S}_0(\mathbb{R}^d)$).

Typical Applied Questions

Let us thus list a couple of questions where Fourier Analysis has to play a role and doing it properly may be appreciated even by very applied persons from the engineering or physics community:

1. The theory of Translation Invariant Systems works with *convolution* by the *impulse response* or alternatively with the *transfer function* (i.e. multiplication on the Fourier transform side);
2. The Shannon-Sampling Theorem, allowing to reconstruct band-limited functions from regular samples at or above the Nyquist rate; it is based on *Poisson's Formula*;
3. *Generalized Stochastic Processes* can be seen as a combination of distribution theory with classical stochastic processes. They can be modelled as linear operators from $\boldsymbol{S}_0(\mathbb{R}^d)$ to some Hilbert space $\mathcal{H}$ (of random variables).

Sampling and Periodization on the FT side

The convolution theorem can then be used to show that sampling corresponds to periodization on the Fourier transform side, with the interpretaton that

$$\sqcup\!\sqcup \cdot f = \sum_{k \in \mathbb{Z}^d} f(k)\delta_k, \quad f \in \mathcal{S}(\mathbb{R}^d).$$

In fact, we have

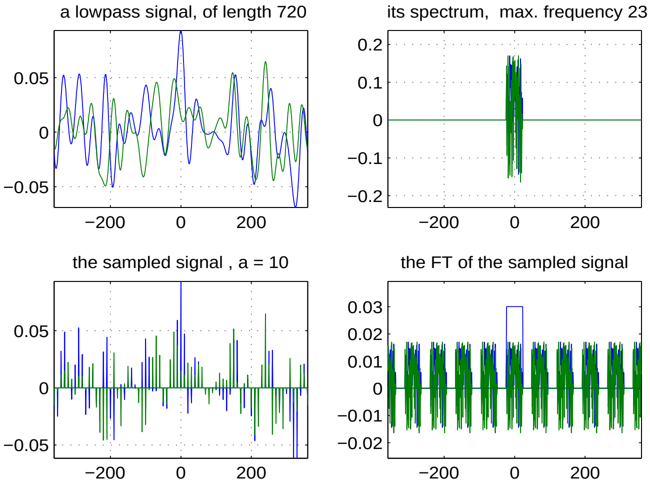
$$\widehat{\sqcup\!\sqcup \cdot f} = \widehat{\sqcup\!\sqcup} * \widehat{f} = \sqcup\!\sqcup * \widehat{f}.$$

This result is the key to prove **Shannon's Sampling Theorem** which is usually considered as the fundamental fact of digital signal processing (Claude Shannon: 1916 - 2001).

Recovery from Regular Samples: Shannon's Theorem

If we try to recover a real function $f : \mathbb{R} \to \mathbb{R}$ from samples, i.e. from a sequence of values $(f(x_n))_{n \in I}$, where I is a finite or (countable) infinite set, we cannot expect perfect reconstruction. In the setting of $\left(\boldsymbol{L}^2(\mathbb{R}), \|\cdot\|_2\right)$ any sequence constitutes only set of measure zero, so knowing the sampling values provides *zero information* without side-information.
On the other hand it is clear the for a *(uniformly) continuous* function, so e.g. a continuous function supported on $[-K, K]$ for some $K > 0$ piecewise linear interpolation (this is what MATLAB does automatically when we use the PLOT-routine) is providing a good (in the uniform sense) approximation to the given function f as long as the maximal distance between the sampling points around the interval $[-K, K]$ is small enough.
Shannon's Theorem says that one can have **perfect reconstruction** for band-limited functions.

A Visual Proof of Shannon's Theorem



Shannon's Sampling Theorem

It is kind of clear from this picture that one can recover the spectrogram of the original function by isolating the central copy of the periodized version of $\widehat{f}$ by multiplying with some function $\hat{g}$, with g such that $\hat{g}(x) = 1$ on spec(f) and $\hat{g}(x) = 0$ at the shifted copies of $\widehat{f}$. This is of course only possible if these shifted copies of spec(f) do not overlap, resp. if the sampling is dense enough (and correspondingly the periodization of $\widehat{f}$ is a coarse one). This conditions is known as the *Nyquist criterion*. If it is satisfied, or supp(f) $\subset [-1/\alpha, 1/\alpha]$, then

$$f(t) = \sum_{k \in \mathbb{Z}^d} f(\alpha k) T_{\alpha k} g(x), \quad x \in \mathbb{R}^d.$$

Proof an extension of the Shannon Sampling Theorem

Although the Hilbert space is very nice we will often encounter non-perfect situations, in the following respect:

- the sampled function may not belong to $\boldsymbol{L}^2(\mathbb{R}^d)$ but rather some $\boldsymbol{L}^p(\mathbb{R}^d)$, or in some weighted $\boldsymbol{L}^p$-space;
- the function might not be strictly band-limited, but only approximately, with "small tails" on the Fourier transform side, e.g. $f \in \mathcal{H}_s(\mathbb{R}^d)$, some Sobolev space;
- the samples might *not be regular*, either due to *jitter error*, or generically *irregular sampled*, perhaps with some outliers, so that one has to perform *scattered data approximation* of the underlying function f from the data $(f(x_i))$.

In all these cases we should have suitable mathematical tools and algorithms in order to analytically study the problem. As we will see *Wiener amalgam spaces* are a highly appropriate tool.

Band-limited functions in $L^p(\mathbb{R}^d)$

Let us begin with the case of band-limited functions in $\boldsymbol{L}^p(\mathbb{R}^d)$, for some $p \in [1, \infty]$. The first question is, *what does it mean for the Fourier transform to be zero outside some cube* Q_0? Especially for $p > 2$ where the Hausdorff-Young inequality (implying that $\mathcal{F}\boldsymbol{L}^p(\mathbb{R}^d) \subset \boldsymbol{L}^q(\mathbb{R}^d)$, for $1/p + 1/q = 1$)

Since $\boldsymbol{L}^p(\mathbb{R}^d) \hookrightarrow \boldsymbol{S}_0'(\mathbb{R}^d)$ it is clear that the Fourier transform exists in the sense of $\boldsymbol{S}_0'(\mathbb{R}^d)$ and hence we assume that $\operatorname{supp}(\widehat{f}) \subseteq Q_0$.

If we want to cover the case $p = 1$ we should avoid the SINC function (not in $\boldsymbol{L}^1(\mathbb{R}^d)$!) but rather choose some function h in $\boldsymbol{S}_0(\mathbb{R}^d)$ with $\widehat{h}(q) \equiv 1$ on Q_0 and $\widehat{h}(q + k/\alpha) = 0$ for $k \in \mathbb{Z}^d \setminus \{0\}$, for example some plateau-type function. Then

$$(\sqcup\!\sqcup_{1/\alpha} * \widehat{f}) \cdot h = \widehat{f},$$

or by the the inverse Fourier transform, for $g = C_\alpha \mathcal{F}^{-1} h \in \boldsymbol{S}_0(\mathbb{R}^d)$:

$$f = (\sqcup\!\sqcup_\alpha \cdot f) * g.$$

Band-limited functions in $L^p(\mathbb{R}^d)$ II

But are all the infinite sums convergent, and are the limits (of their partical sums) convergent in the given space $\left(L^p(\mathbb{R}^d), \|\cdot\|_p\right)$? After all, the choice of g resp. h does *not depend* on the parameter $p \in [1, \infty]$, but only on α and Q_0 (as long as on has $Q_0 \cap k/\alpha + Q_0 = \emptyset$ for $k \in \mathbb{Z}^d, k \neq 0$).
First of all we see that $h \cdot \widehat{f} = \widehat{f}$ for obvious reasons, or equivalently $h * g = g$ for some $g \in S_0(\mathbb{R}^d)$. Since we assume that $f \in L^p = W(L^1, \ell^p)$ this implies that one actually has

$$f = f * g \in W(L^1, \ell^p) * W(C_0, \ell^1) \subset W(C_0, \ell^p)(\mathbb{R}^d). \quad (34)$$

Consequently the samples belong to $\ell^p(\mathbb{Z}^d)$, but it is better to argue that $\sqcup\!\sqcup_\alpha \in W(M, \ell^\infty)$ and hence

$$\sqcup\!\sqcup_\alpha \cdot f \in W(M, \ell^\infty) \cdot W(C_0, \ell^p) \subset W(M, \ell^p)(\mathbb{R}^d). \quad (35)$$

Band-limited functions in $L^p(\mathbb{R}^d)$ III

Finally we prove the convergence of the Shannon sampling series:

$$(\sqcup\!\sqcup_\alpha \cdot f) * g = \left(\sum_{k \in \mathbb{Z}^d} f(\alpha k) \delta_{\alpha k} \right) * g = \sum_{k \in \mathbb{Z}^d} f(\alpha k) T_{\alpha k} g. \quad (36)$$

Since $\sqcup\!\sqcup_\alpha \cdot f \in \boldsymbol{W}(\boldsymbol{M}, \ell^p)$ the convergence in $\boldsymbol{W}(\boldsymbol{C}_0, \ell^p)(\mathbb{R}^d)$, and hence in $\left(\boldsymbol{L}^p(\mathbb{R}^d), \|\cdot\|_p\right)$ and uniformly follows from

$$\boldsymbol{W}(\boldsymbol{M}, \ell^p) * \boldsymbol{S}_0 \subset \boldsymbol{W}(\boldsymbol{M}, \ell^p) * \boldsymbol{W}(\boldsymbol{C}_0, \ell^1) \subset \boldsymbol{W}(\boldsymbol{C}_0, \ell^p)(\mathbb{R}^d). \quad (37)$$

For $p = \infty$ minor modifications may apply (if $f \notin \boldsymbol{C}_0(\mathbb{R}^d)$).

Band-limited functions in $L^p(\mathbb{R}^d)$ IV

We cannot go into details about the irregular case, but at least we mention that instead of an orthonormal basis of shifte SINC-functions one has a **frame** of shifted SINC functions describing the situation, since

$$f(t_i) = f * SINC(t_i) = \langle f, T_{t_i} SINC \rangle, \quad i \in I.$$

For the case of irregular samples $(f(t_i))$ of a band-limited function in $\boldsymbol{L}^p(\mathbb{R}^d)$ (with high enough density, depending only on Q_0!) one can write a Shannon-like series of the form $Af = \sum_{i \in I} w_i T_{t_i} g$ for well chosen *adaptive weights* (see [5]) and then goes on the prove

$$\|Af - f\|_{\boldsymbol{W}(\boldsymbol{C}_0, \ell^p)} \leq \gamma \cdot \|f\|_{\boldsymbol{W}(\boldsymbol{C}_0, \ell^p)}, \quad \text{for some } \gamma < 1$$

and for all Q_0 band-limited functions in $\boldsymbol{L}^p(\mathbb{R}^d)$, so that Banach's fix point theorem can be applied to do the rest ([4]).

Overall perspective

In this section we will explain how the setting of the Banach Gelfand Triple $(\boldsymbol{S_0}, \boldsymbol{L}^2, \boldsymbol{S_0'})(\mathbb{R}^d)$ can be used to formulate a number of important general principles, most of which actually extend to the setting of LCA (locally compact Abelian) groups, even if they do not have arbitary fine lattices.

In many cases this setting allows for *compact formulations* of natural statements, combined with a *unified principle of proof*!

The Fourier transform will be the prototypical example, the kernel theorem for linear operators the other one, but there are many more of these statements, also in connection with the theory of *Banach frames* and *Riesz projection bases*.

Fourier Transform as Banach Gelfand Triple Automorphism

First of all we can describe the Fourier transform on $\mathbb{R}^d$ as a *unitary Banach Gelfand Triple automorphism* of $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$, meaning that it is

- well defined (and isometric) on $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$;
- extending to a *unitary automorphism of* $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$;
- with a unique w^*-w^*-extension to $\boldsymbol{S}_0'(\mathbb{R}^d)$.

As you see the classical Lebesgue space (aside from the Hilbert space $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$) do not play an important role now, because we see the Fourier transform in a wider context than just being an *integral transform*. Only the view that the Fourier transform should be an *integral transform* suggest to choose $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ as a domain, but this is not good enough to find out that the Fourier transform of a pure frequency is just a Dirac.

BANACH GELFAND TRIPLES: a new category

Definition

A triple, consisting of a Banach space $\boldsymbol{B}$, which is dense in some Hilbert space $\mathcal{H}$, which then is naturally embedded (and w^*dense in) $\boldsymbol{B}'$ is called a Banach Gelfand triple.

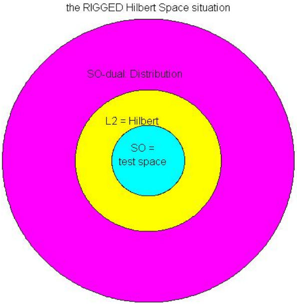
Definition

If $(\boldsymbol{B}_1, \mathcal{H}_1, \boldsymbol{B}_1')$ and $(\boldsymbol{B}_2, \mathcal{H}_2, \boldsymbol{B}_2')$ are Gelfand triples then a linear operator T is called a [unitary] Gelfand triple isomorphism if

1. T is an isomorphism between $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$.
2. T is [a unitary operator resp.] an isomorphism between $\mathcal{H}_1$ and $\mathcal{H}_2$.
3. T extends to a weak* isomorphism as well as a norm-to-norm continuous isomorphism between $\boldsymbol{B}_1'$ and $\boldsymbol{B}_2'$.

A schematic description: the simplified setting

In our picture this simple means that the inner "kernel" is mapped into the "kernel", the Hilbert space to the Hilbert space, and at the outer level two types of continuity are valid (norm and w^*)!



The prototypical examples over the torus

In principle every CONB (= *complete orthonormal basis*) $\Psi = (\psi_i)_{i \in I}$ for a given Hilbert space $\mathcal{H}$ can be used to establish such a unitary isomorphism, by choosing as $\boldsymbol{B}$ the space of elements within $\mathcal{H}$ which have an absolutely convergent expansion, i.e. satisfy $\sum_{i \in I} |\langle x, \psi_i \rangle| < \infty$.
For the case of the Fourier system as CONB for $\mathcal{H} = \boldsymbol{L}^2([0,1])$, i.e. the corresponding definition is already around since the times of N. Wiener: $\boldsymbol{A}(\mathbb{T})$, the space of absolutely continuous Fourier series. It is also not surprising in retrospect to see that the dual space $\boldsymbol{PM}(\mathbb{T}) = \boldsymbol{A}(\mathbb{T})'$, known as the space of *pseudo-measures*, appears. One can extend the classical Fourier transform to this space, and in fact interpret this extended mapping, in conjunction with the classical Plancherel theorem as the first unitary Banach Gelfand triple isomorphism, namely between $(\boldsymbol{A}, \boldsymbol{L}^2, \boldsymbol{PM})(\mathbb{T})$ and $(\boldsymbol{\ell}^1, \boldsymbol{\ell}^2, \boldsymbol{\ell}^\infty)(\mathbb{Z})$.

The KERNEL THEOREM for $\mathcal{S}(\mathbb{R}^d)$

The *kernel theorem* for the Schwartz space can be read as follows:

Theorem

For every continuous linear mapping T from $\mathcal{S}(\mathbb{R}^d)$ into $\mathcal{S}'(\mathbb{R}^d)$ there exists a unique tempered distribution $\sigma \in \mathcal{S}'(\mathbb{R}^{2d})$ such that

$$T(f)(g) = \sigma(f \otimes g), \quad f, g \in \mathcal{S}(\mathbb{R}^d). \tag{38}$$

Conversely, any such $\sigma \in \mathcal{S}'(\mathbb{R}^{2d})$ induces a (unique) operator T such that (38) holds.

The proof of this theorem is based on the fact that $\mathcal{S}(\mathbb{R}^d)$ is a *nuclear Frechet space*, i.e. has the topology generated by a sequence of semi-norms, can be described by a metric which turns $\mathcal{S}(\mathbb{R}^d)$ into a complete metric space.

The KERNEL THEOREM for S_0 I

Tensor products are also most suitable in order to describe the set of all operators with certain mapping properties. The backbone of the corresponding theorems are the *kernel-theorem* which reads as follows (!! despite the fact that $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ is NOT a *nuclear Frechet space*)

One of the corner stones for the kernel theorem is: One of the most important properties of $\boldsymbol{S}_0(\mathbb{R}^d)$ (leading to a characterization given by V. Losert, [10]) is the tensor-product factorization:

Lemma

$$\boldsymbol{S}_0(\mathbb{R}^k)\hat{\otimes}\boldsymbol{S}_0(\mathbb{R}^n) \cong \boldsymbol{S}_0(\mathbb{R}^{k+n}), \tag{39}$$

with equivalence of the corresponding norms.

The KERNEL THEOREM for S_0 II

The **Kernel Theorem** for general operators in $\mathcal{L}(\boldsymbol{S}_0, \boldsymbol{S}_0')$:

Theorem

If K is a bounded operator from $\boldsymbol{S}_0(\mathbb{R}^d)$ to $\boldsymbol{S}_0'(\mathbb{R}^d)$, then there exists a unique kernel $k \in \boldsymbol{S}_0'(\mathbb{R}^{2d})$ such that $\langle Kf, g\rangle = \langle k, g \otimes f\rangle$ for $f, g \in \boldsymbol{S}_0(\mathbb{R}^d)$, where $g \otimes f(x,y) = g(x)f(y)$.

Formally sometimes one writes by "abuse of language"

$$Kf(x) = \int_{\mathbb{R}^d} k(x,y)f(y)dy$$

with the understanding that one can define the action of the functional $Kf \in \boldsymbol{S}_0'(\mathbb{R}^d)$ as

$$Kf(g) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} k(x,y)f(y)dy\, g(x)dx = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} k(x,y)g(x)f(y)dxdy.$$

The KERNEL THEOREM for S_0 III

This result is the "outer shell" of the Gelfand triple isomorphism. The "middle = Hilbert" shell which corresponds to the well-known result that Hilbert Schmidt operators on $\boldsymbol{L}^2(\mathbb{R}^d)$ are just those compact operators which arise as integral operators with $\boldsymbol{L}^2(\mathbb{R}^{2d})$-kernels.

The KERNEL THEOREM for S_0 IV

Theorem

The classical kernel theorem for Hilbert Schmidt operators is unitary at the Hilbert spaces level, with $\langle T, S\rangle_{\mathcal{HS}} = trace(T * S')$ as scalar product on $\mathcal{HS}$ and the usual Hilbert space structure on $\boldsymbol{L}^2(\mathbb{R}^{2d})$ on the kernels.
Moreover, such an operator has a kernel in $\boldsymbol{S}_0(\mathbb{R}^{2d})$ if and only if the corresponding operator K maps $\boldsymbol{S}_0'(\mathbb{R}^d)$ into $\boldsymbol{S}_0(\mathbb{R}^d)$, but not only in a bounded way, but also continuously from w^*-topology into the norm topology of $\boldsymbol{S}_0(\mathbb{R}^d)$.

In analogy to the matrix case, where the entries of the matrix

$$a_{k,,j} = T(\mathbf{e}_j)_k = \langle T(\mathbf{e}_j), \mathbf{e}_k\rangle$$

we have for $K \in \boldsymbol{S}_0$ the continuous version of this principle:

$$K(x,y) = \delta_x(T(\delta_y)), \quad x, y \in \mathbb{R}^d.$$

The Kernel Theorem as a BGT isomorphism

The different version of the kernel theorem for operators between $\mathbf{S}_0$ and $\mathbf{S}_0'$ can be summarized using the terminology of Banach Gelfand Triples (BGTR) as follows.

Theorem

There is a unique Banach Gelfand Triple isomorphism between the Banach Gelfand triple of kernels $(\mathbf{S}_0, \mathbf{L}^2, \mathbf{S}_0')(\mathbb{R}^{2d})$ *and the operator Gelfand triple around the Hilbert space* $\mathcal{HS}$ *of Hilbert Schmidt operators, namely* $(\mathcal{L}(\mathbf{S}_0', \mathbf{S}_0), \mathcal{HS}, \mathcal{L}(\mathbf{S}_0, \mathbf{S}_0'))$*, where the first set is understood as the* w^* *to norm continuous operators from* $\mathbf{S}_0'(\mathbb{R}^d)$ *to* $\mathbf{S}_0(\mathbb{R}^d)$*, the so-called regularizing operators.*

Spreading function and Kohn-Nirenberg symbol

1. For $\sigma \in \boldsymbol{S}_0'(\mathbb{R}^d)$ the *pseudodifferential operator* with *Kohn-Nirenberg symbol* σ is given by:

$$T_\sigma f(x) = \int_{\mathbb{R}^d} \sigma(x,\omega)\hat{f}(\omega)e^{2\pi i x\cdot\omega}d\omega$$

The formula for the integral kernel $K(x,y)$ is obtained

$$\begin{aligned} T_\sigma f(x) &= \int_{\mathbb{R}^d}\Big(\int_{\mathbb{R}^d}\sigma(x,\omega)e^{-2\pi i(y-x)\cdot\omega}d\omega\Big)f(y)dy \\ &= \int_{\mathbb{R}^d} k(x,y)f(y)dy. \end{aligned}$$

2. The *spreading representation* of T_σ arises from

$$T_\sigma f(x) = \iint_{\mathbb{R}^{2d}} \widehat{\sigma}(\eta,u)M_\eta T_{-u}f(x)dud\eta.$$

$\widehat{\sigma}$ is called the spreading function of T_σ.

Further details concerning Kohn-Nirenberg symbol

(courtesy of Goetz Pfander (Eichstätt):)

- *Symmetric coordinate transform:* $\mathcal{T}_s F(x,y) = F(x+\frac{y}{2}, x-\frac{y}{2})$
- *Anti-symmetric coordinate transform:* $\mathcal{T}_a F(x,y) = F(x, y-x)$
- *Reflection:* $\mathcal{I}_2 F(x,y) = F(x,-y)$
- *partial Fourier transform in the first variable:* $\mathcal{F}_1$
- *partial Fourier transform in the second variable:* $\mathcal{F}_2$

The kernel $K(x,y)$ can be described as follows:

$$K(x,y) = \mathcal{F}_2\sigma(\eta, y-x) = \mathcal{F}_1^{-1}\widehat{\sigma}(x, y-x)$$
$$= \int_{\mathbb{R}^d} \widehat{\sigma}(\eta, y-x) \cdot e^{2\pi i \eta \cdot x} d\eta.$$

Kohn-Nirenberg symbol and spreading function II

operator H	$Hf(x)$
$\updownarrow$	$=$
kernel κ_H	$\int \kappa_H(x,s)f(s)\,ds$
$\updownarrow$	$=$
Kohn–Nirenberg symbol σ_H	$\int \sigma_H(x,\omega)\widehat{f}(\omega)e^{2\pi i x\cdot\omega}\,d\omega$
$\updownarrow$	$=$
time–varying impulse response h_H	$\int h_H(t,x)f(x-t)\,dt$
$\updownarrow$	$=$
spreading function η_H	$\int\int \eta_H(t,\nu)f(x-t)e^{2\pi i x\cdot\nu}\,dt\,d\nu$
	$=$
	$\int\int \eta_H(t,\nu)M_\nu T_t f(x),dt\,d\nu\,,$

Spreading representation and commutation relations

The description of operators through the spreading function and allows to understand a number of commutation relations.
If an operator is a limit (in the strong operator topology) of translation operators it is just a convolution operator with some $\tau \in \boldsymbol{S}_0'(\mathbb{R}^d)$, resp. its spreading representation is just an element concentrated on the *time axis* (more or less representing $\hat{\tau}$, the "individual frequency contributions".
Similarly, multiplication operators require just the use of modulation operators, so their spreading function is concentrated in the frequency axis of the TF-plane.
Finally typical *Gabor frame operators* arising from a family of Gabor atoms (g_λ), where $\lambda \in \Lambda$, some lattice within $\mathbb{R}^d \times \widehat{\mathbb{R}}^d$ typically commute with TF-shift operators, one can say that they are obtained by periodizing the projection operator $f \mapsto \langle f, g \rangle g$ along the lattice.

The symplectic Fourier transform

The *symplectic Fourier transform* connects the Kohn-Nirenberg symbol with the spreading function, i.e.

$$\mathcal{F}_s(\sigma(T)) = \eta(T) \quad \text{resp.} \quad \mathcal{F}_s(\eta(T)) = \sigma(T). \tag{40}$$

$$(\mathcal{F}_{symp} f)(k,l) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x,y) e^{-2\pi i (k\cdot y - l\cdot x)}; \quad f \in \boldsymbol{S}_0(\mathbb{R}^d \times \widehat{\mathbb{R}}^d). \tag{41}$$

It is completely characterized by its action on elementary tensors:

$$\mathcal{F}_{symp}(f \otimes \hat{g}) = g \otimes \hat{f}, \quad f, g \in \boldsymbol{S}_0(\mathbb{R}^d), \tag{42}$$

and extends from there in a unique way to a $w^* - w^*$ continuous mapping from $\boldsymbol{S}_0'(\mathbb{R}^{2d})$ to $\boldsymbol{S}_0'(\mathbb{R}^{2d})$, also ${\mathcal{F}_s}^2 = Id$.

Understanding the Janssen representation

The spreading representation of operators has properties very similar to the ordinary Fourier expansion for functions! Periodization at one side corresponds to sampling on the transform side, if we understand "translation" either at the level of ordinary translation of the Kohn-Nirenberg symbol (which is the *symplectic Fourier transform* of the spreading function), OR by conjugation of an operator by the corresponding TF-shifts.

In other words: for any given operator T and $\lambda \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d$ we can **define** [recall $\pi(x,\omega) = M_\omega T_x$ for $\lambda = (x,\omega)$]

$$\pi \otimes \pi^*(T) = \pi(\lambda) \circ T \circ \pi(\lambda)^*, \tag{43}$$

providing the important *covariance property* for KNS:

$$\sigma[\pi \otimes \pi^*(\lambda)(T)] = T_\lambda[\sigma(T)], \quad \lambda \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d. \tag{44}$$

Periodization goes over to sampling

If we have a "nice operator" T_0 we can form its periodic version $\sum_{\lambda\in\Lambda} \pi\otimes\pi^*(\lambda)(T_0)$ and it is still a well defined operator from $\boldsymbol{S}_0(\mathbb{R}^d)$ to $\boldsymbol{S}_0'(\mathbb{R}^d)$. Its KNS is just the Λ-periodization of T_0. Consequently its spreading function is obtained by sampling of $\eta(T)\in\boldsymbol{S}_0(\mathbb{R}^d\times\widehat{\mathbb{R}}^d)$, over the *adjoint lattice* Λ° and obtain in this case an ℓ^1-sequence.
The adjoint lattice Λ° can be characterized by the fact that

$$\mathcal{F}_s(\sqcup\!\sqcup_\Lambda) = C_\Lambda \sqcup\!\sqcup_{\Lambda^\circ}. \tag{45}$$

For the projection on the Gabor atom $P_g: f\mapsto\langle f,g\rangle g$ the spreading functions is essentially

$$[\eta(P_g)](\lambda) = Vg(g)(\lambda) = \langle g,\pi(\lambda)g\rangle, \quad \lambda\in\mathbb{R}^d\times\widehat{\mathbb{R}}^d.$$

Janssen representation II

An important insight concerning the connection between the Gabor atom g, the TF-lattice $\Lambda \lhd \mathbb{R}^d \times \widehat{\mathbb{R}}^d$ and the quality of the resulting Gabor frame resp. Gabor Riesz basis (e.g.condition number) clearly comes from the *Janssen representation* of the *Gabor frame operator* for any $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ with $\|g\|_2 = 1$:

$$S_{g,\Lambda}(f) = \sum_{\lambda \in \Lambda} P_{g_\lambda}(f) = \sum_{\lambda \in \Lambda} \pi \otimes \pi^*(\lambda)[P_g]. \qquad (46)$$

The periodization principle gives the Janssen representation

$$S_{g,\Lambda} = \eta^{-1}[\eta(S_{g,\Lambda})] = C_\Lambda \sum_{\lambda^\circ \in \Lambda^\circ} V_g(g))(\lambda^\circ)\pi(\lambda^\circ), \qquad (47)$$

as an absolutely convergent sum of TF-shifts from Λ°.

Fourier Standard Spaces of Operators

The kernel theorem allows to identify many spaces of linear operators (with different forms of continuity) with suitable FouSSs over $\mathbb{R}^{2d}$.
For example, there are the so-called *Schatten classes* of operators on the Hilbert space $\boldsymbol{L}^2(\mathbb{R}^d)$ which are compact operators with singular values in ℓ^p, for $1 \leq p < \infty$. These spaces are *operator ideals* within $\mathcal{L}(\mathcal{H})$, i.e. they are Banach spaces, continuously embedded into the space of compact operators over the Hilbert space $\mathcal{H}$, as well as two-sided Banach ideals, i.e. whenever one has an operator T in such a space, and two bounded operators S_1, S_2 on $\mathcal{H}$, then $S_1 \circ T \circ S_2$ also belongs to that *operator ideal* and the operator ideal norm is bounded by the operator ideal norm of T multiplied with the operator norms of S_1 and S_2.

J. Cooley and J. W. Tukey.
An algorithm for the machine calculation of complex Fourier series.
Math. Comp., 19:297–301, 1965.

H. G. Feichtinger.
On a new Segal algebra.
Monatsh. Math., 92:269–289, 1981.

H. G. Feichtinger.
Banach convolution algebras of Wiener type.
In *Proc. Conf. on Functions, Series, Operators, Budapest 1980*, volume 35 of *Colloq. Math. Soc. Janos Bolyai*, pages 509–524. North-Holland, Amsterdam, Eds. B. Sz.-Nagy and J. Szabados. edition, 1983.

H. G. Feichtinger and K. Gröchenig.
Iterative reconstruction of multivariate band-limited functions from irregular sampling values.

SIAM J. Math. Anal., 23(1):244–261, 1992.

H. G. Feichtinger, K. Gröchenig, and T. Strohmer.
Efficient numerical methods in non-uniform sampling theory.
Numer. Math., 69(4):423–440, 1995.

H. G. Feichtinger and T. Strohmer.
Gabor Analysis and Algorithms. Theory and Applications.
Birkhäuser, Boston, 1998.

K. Gröchenig.
Foundations of Time-Frequency Analysis.
Appl. Numer. Harmon. Anal. Birkhäuser, Boston, MA, 2001.

M. S. Jakobsen.
On a (no longer) New Segal Algebra: A Review of the Feichtinger Algebra.
J. Fourier Anal. Appl., pages 1 – 82, 2018.

M. J. Lighthill.

Introduction to Fourier analysis and Generalised Functions.
Cambridge Monographs on Mechanics and Applied Mathematics. Cambridge University Press, New York, 1958.

V. Losert.
A characterization of the minimal strongly character invariant Segal algebra.
Ann. Inst. Fourier (Grenoble), 30:129–139, 1980.

A. Papoulis.
The Fourier Integral and Its Applications.
McGraw-Hill Book Co., Inc., 1962.

H. Reiter and J. D. Stegeman.
Classical Harmonic Analysis and Locally Compact Groups. 2nd ed.
Clarendon Press, Oxford, 2000.