NuHAG Numerical Harmonic Analysis Group

Distribution Theory based on Time-Frequency Analysis

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

September 24th, 2013
OEMG Congress 2013, INNSBRUCK

Distribution Theory for Signal Processing

Distribution theory is typically associated with the treatment of PDE, and the Schwartz space $\mathcal{S}(\mathbb{R}^d)$ is well suited for this purpose. But it is only a topological vector space (Frechet space) and not so easy to explain to engineers.
The fact that there is no differentiability structure on general LCA groups as well as the fact that many engineering applications (characterizations of translation invariant systems, sampling and periodization, Shannon's sampling theorem) call an alternative theory of generalized functions suitable for a solid mathematical tool to prove results relevant for applications.
It is the purpose of this talk to present such a setting, based on considerations coming up in time-frequency analysis. In fact, both for $\mathbb{R}^d$ or general LCA groups this approach is more convenient than the corresponding Schwartz-Bruhat theory.

The Description is given inside of $\mathcal{S}'(\mathbb{R}^d)$

Although it is not too difficult to present the theory envisaged in the context of LCA groups, based on elementary properties of the Banach algebra $\mathbf{L}^1(G)$ (with respect to convolution) and the Fourier transform (and a bit of Wiener amalgam space theory), but also from scratch, starting from elementary concepts, it is much easier to describe the whole theory within the (well established) realm of tempered distributions.

In other words we rather refer her to a space of distributions which is contained in $\mathcal{S}'(\mathbb{R}^d)$, but contains practically all the objects of relevance for signal processing applications, including all the $\mathbf{L}^p$-spaces, as a well as some important unbounded measures.

BUT THE WHOLE THEORY CAN BE BUILT, WITH SLIGHTLY MORE EFFORT, WITHOUT REFERENCE TO $\mathcal{S}'(\mathbb{R}^d)$.

This is like discussing the field $\mathbb{Q}(\sqrt{2})$ either as a subfield of $\mathbb{R}$ or as an extension of $\mathbb{Q}$.

What is time-frequency analysis

Time-frequency analysis is a part of analysis, which studies functions or (tempered) distributions via their **spectrograms**, i.e. their *localized* or *sliding-window Fourier transforms*, and judge linear operators by their effect on spectrograms.
Based on this method one defines a family of Banach spaces of functions (resp. distributions), called **modulation spaces** (introduced around 1983). The three most important ones are the Segal algebra $\mathbf{S}_0(\mathbb{R}^d)$, the Hilbert space $\mathbf{L}^2(\mathbb{R}^d)$ and the dual space $\mathbf{S}_0'(\mathbb{R}^d)$, and together they form a so-called **Banach Gelfand triple**, quite analogously to the triple $(\mathcal{S}, \mathbf{L}^2, \mathcal{S}')(\mathbb{R}^d)$.
Gabor Analysis is the branch of TF-analysis dealing with sampled spectrograms resp. atomic decompositions of distributions using TF-shifted copies of a *Gabor atom*.

The key-players for time-frequency analysis

Time-shifts and Frequency shifts

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

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

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

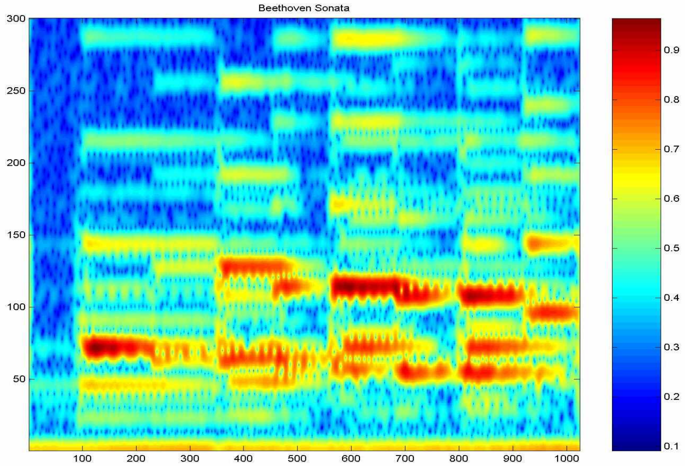
Behavior under Fourier transform

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

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



Some Heuristic Claims in Fourier Analysis

From Fourier Series to Fourier Transforms

How can we explain the form of the (forward and inverse) Fourier transform, given the well-known Fourier series expansion formulas for periodic, square integrable functions?

The usual trick is to considers for a square integrable function with compact support its p-periodic versions, for $p \to \infty$.

In the limit Riemannian sums turn to the integral of the form

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

which makes sense for $f \in \mathbf{L}^1(\mathbb{R}^d)$, with inversion formula

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

which is a bit more problematic (except for $\hat{f} \in \mathbf{L}^1(\mathbb{R}^d)$).

Some Heuristic Claims in Fourier Analysis II

Fourier Transforms versus FFT

Whenever the Fourier transform of some decent function f has to be computed the FFT (Fast Fourier transform) is applied, mostly by analogy. Integrals are then replaced by sums, the function is (approximately) represented by its sampling values, at least if they are taken over a sufficiently long interval at a sufficiently high rate. But all these vague statements, can one guarantee (resp. under which conditions) that the FFT applied to such a sequence has anything to do with the samples of the Fourier transform (understood as integral transform) over a suitable lattice?
In the strict sense such heuristics to be wrong, in the strict mathematical sense, but on the other hand they are still reasonable from a practical point of view (model error?).

Some Heuristic Claims in Fourier Analysis III

Fourier Analysis and Fourier Synthesis

There is a number of problems with the usual interpretation, viewing the signal f as a (continuous) superposition of pure frequencies with amplitudes $\hat{f}(s)$ for the pure frequency

$$\chi_s(t) := e^{i\pi s \cdot t},$$

where $s \cdot t$ stands for the usual scalar product on $\mathbb{R}^d$.

Considering then Riemannian sums corresponding to this integral, or truncated sums, one quickly sees that they are all trigonometric polynomials (or series), which are periodic, and hence it does *not makes sense* to measure the difference between such "approximating sums" and their limit $f \in \mathbf{L}^2(\mathbb{R}^d)$.

Some Heuristic Claims in Fourier Analysis IV

From the engineering point of view once can describe a TILS (translation invariant linear systems resp. operators) are just convolution operators. This comes from the following *heuristic* consideration, using $T_x f(y) = f(y-x)$, noting that $T_x(\delta_0) = \delta_x$, referring to the so-called *sifting-property* of the Dirac:

$$f = \int_G f(x)\, \delta_x dx = \int_G T_x(\delta_0) f(x) dx;$$

hence with $\mu := T(\delta_0)$, the **impulse response** we have:

$$Tf = \int_G T(T_x(\delta_0)) f(x) dx = \int_G T_x(\mu) f(x) dx = f * \mu,$$

so at the Fourier transform side we get the **transfer function**

$$\widehat{Tf} = \hat{\mu} \cdot \hat{f}.$$

Axiomatic Setting

We will need two kinds of objects for our purpose. Once a space of "good functions", which are easy to handle, but still large enough to contain many functions, for which e.g. Poisson's formula is valid. It should be invariant under the usual Fourier transform, but also maybe a Banach algebra under pointwise multiplication or change of coordinates.

Correspondingly the dual space, containing now a sufficiently large collection of objects needed for signal processing applications, should not be too small, which in turn implies that the space of test functions should not be too large.

Just think of $\mathbf{C}_0(\mathbb{R}^d)$ with the sup-norm, which is simple to use, and has the bounded measures as a dual space, but is by far not Fourier invariant.

The key-players for time-frequency analysis II

The typical "window" used in order to determine the spectrogram is the Gauss function $g(t) = e^{-\pi|t|^2}, t \in \mathbb{R}^d$.
The classical theory of tempered distributions allows to find out that for each $\sigma \in \boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ the STFT $V_g(\sigma)$ is of at most polynomial growth, i.e. $V_g(\sigma) = \mathcal{O}(1 + |(x,\omega)|)^s$ for some positive s, while $f \in \boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ if and only if $V_g(f)$ is decaying faster than any polynomial, or $V_g(f) = \mathcal{O}|(x,\omega)|^r$ for all $r \in \mathbb{R}$.
Moreover, since $\|g\|_2 = 1$ one can prove that

$$\|V_g(f)\|_{\mathbf{L}^2(\mathbb{R}^d \times \widehat{\mathbb{R}}^d)} = \|f\|_{\mathbf{L}^2(\mathbb{R}^d)},$$

and in particular $V_g(f) \in \mathbf{L}^2(\mathbb{R}^d \times \widehat{\mathbb{R}}^d)$ if and only if $f \in \mathbf{L}^2(\mathbb{R}^d)$.
For fixed s resp. r the Shubin Classes $\mathbf{Q}_s(\mathbb{R}^d)$ correspond to weighted $\mathbf{L}^2$-spaces, with radial symmetric weights $v_s(x,\omega) := (1 + |(x,\omega)|)^s$.

A Banach Space of Test Functions (Fei 1979)

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

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

The space $\left(\mathbf{S}_0(\mathbb{R}^d), \|\cdot\|_{\mathbf{S}_0}\right)$ is a Banach space, for any fixed, non-zero $g \in \mathbf{S}_0(\mathbb{R}^d)$), and different windows g define the same space and equivalent norms. Since $\mathbf{S}_0(\mathbb{R}^d)$ 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 $\mathbf{M}^1 = \mathbf{S}_0(\mathbb{R}^d)$

Lemma

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

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

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

In fact, $\left(\mathbf{S}_0(\mathbb{R}^d), \|\cdot\|_{\mathbf{S}_0}\right)$ is the smallest non-trivial Banach space with this property, and therefore contained in any of the $\mathbf{L}^p$-spaces (and their Fourier images).

There are many other independent characterization of this space, spread out in the literature since 1980, e.g. atomic decompositions using ℓ^1-coefficients, or as $\mathbf{W}(\mathcal{F}\mathbf{L}^1, \ell^1) = \mathbf{M}^0_{1,1}(\mathbb{R}^d)$.

Sufficient conditions for $\mathbf{M}^1 = \mathbf{S}_0(\mathbb{R}^d)$

There is a long list of sufficient conditions for a (continuous and integrable) function to belong to $\mathbf{S}_0(\mathbb{R}^d)$. For simplicity let us discuss the case $d = 1$:

- Assume f is compactly supported with integrable Fourier transform;
- Assume $f \in \mathbf{L}^1(\mathbb{R}^d)$ is bandlimited, i.e. having bounded spectrum (supp($\hat{f}$));
- Assume that $f \in \mathcal{S}(\mathbb{R}^d)$, or just that $f, f' \in \mathbf{L}^1(\mathbb{R})$;
- Assume that f is piecewise linear with a finite number of nodes;
- Assume that $f \in H^s(\mathbb{R}^d)$ and also $f \cdot (1 + |t|)^s \in \mathbf{L}^2(\mathbb{R})$ for some $s > 1$, then $f \in \mathbf{S}_0(\mathbb{R})$.

Making use of $\mathbf{S}_0(\mathbb{R}^d)$

The space of test functions is relatively large, and consists of continuous, absolutely Riemann-integrable functions. It turns out to be very suitablein a variety or situations, even in the context of classical Fourier analysis. For example, it is the optimal domain for **Poisson's formula** :

Lemma

$$\sum_{n\in\mathbb{Z}^d} f(n) = \sum_{k\in\mathbb{Z}^d} \hat{f}(k), \quad \forall f \in \mathbf{S}_0(\mathbb{R}^d).$$

Since the Segal algebra is invariant under general automorphisms the general version of Poisson's formula, involving a lattice $\Lambda = \mathbf{A} * \mathbb{Z}^d$ with $det(\mathbf{A}) \neq 0$ and its orthogonal lattice $\Lambda^\perp$ is valid as well. Equivalently, with $\sqcup\!\sqcup_\Lambda := \sum_{\lambda\in\Lambda} \delta_\lambda$

$$\mathcal{F}(\sqcup\!\sqcup_\Lambda) = C_\Lambda \sqcup\!\sqcup_{\Lambda^\perp} .$$

Making use of $\mathbf{S}_0(\mathbb{R}^d)$

The space $\mathbf{S}_0(\mathbb{R}^d)$ is also suitable in order to "catch" the information in a function by taking samples.

Lemma

For any lattice $\Lambda \lhd \mathbb{R}^d$ the mapping $f \mapsto (f(\lambda))_{\lambda\in\Lambda}$ is continuous from $\mathbf{S}_0(\mathbb{R}^d)$ to $\ell^1(\Lambda)$. Conversely, for given $f \in \mathbf{S}_0(\mathbb{R}^d)$ one can guarantee that for any sufficiently dense lattice Λ any adapted partition of with $\sum_{\lambda\in\Lambda} \psi(x-\lambda) \equiv 1$ on can guarantee that the quasi-interpolant $Q_\Psi(f) = \sum_{\lambda\in\Lambda} f(\lambda) T_\lambda \psi$, will give a good approximation of f in the $\mathbf{S}_0$-sense, hence in any of the $\mathbf{L}^p$-norms.

This applies in particular to piecewise linear interpolation (this is how MATLAB displays function values!) A similar statement is valid for periodization.

Making use of $\mathbf{S}_0(\mathbb{R}^d)$

A combination of the last two observations together with the fact that the Fourier transform (in the distributional sense) of a discrete and periodic signal can be computed *exactly* via the FFT allowed Norbert Kaiblinger to provide a proof that the Fourier transform of $f \in \mathbf{S}_0(\mathbb{R})$ can be computed approximately by applying the FFT to a sampled and periodized version of f and then applying piecewise linear interpolation on the Fourier transform side on a localized version of the resulting periodic and discrete signal.

There are also similar statements applicable in the context of Gabor analysis. In this way continuous Gabor problems can be (approximately, in the $\mathbf{S}_0$-sense) solved using finite Gabor models.

Making use of $\mathbf{S}_0(\mathbb{R}^d)$: Summability Kernels

Finally let us mention that it appears that most (in fact all the ones inspected so far) of the classical summability kernels belong to $\mathbf{S}_0(\mathbb{R}^d)$. It is just the “bad” box-function which does not belong to $\mathbf{S}_0(\mathbb{R}^d)$ (because it is not continuous) as well as its Fourier transform, the so-called SINC-function (which is in fact not integrable).
This statement includes of course the Fejer, De la Vallee-Poussin or Weierstrass kernels, and many others (as investigated by Ferenc Weisz).

Atomic Decompositions for $\mathbf{S}_0(\mathbb{R}^d)$

Another interesting characterization of $\mathbf{S}_0(\mathbb{R}^d)$ is in terms of TF-shifted copies of Gauss functions along a lattice, with absolutely convergent sums:

Theorem

Denote by g_0 the (multi-dimensional) Gauss function $g_0(t) := e^{-\pi|t|^2}$. Then for any pair (a,b) of lattice constants with $ab < 1$ there exists a linear coefficient mapping C from $\mathbf{S}_0(\mathbb{R}^d)$ into $\ell^1(\mathbb{Z}^{2d})$ such that such that

$$f = \sum_{\lambda \in a\mathbb{Z}^d \times b\mathbb{Z}^d} C(f)(\lambda)\pi(\lambda)g_0, \quad \forall f \in \mathbf{S}_0(\mathbb{R}^d).$$

SO(Rd) in Gabor Analysis

Of course there would be a very long list of properties where the properties of windows from $\mathbf{S}_0(\mathbb{R}^d)$ are very valuable. We just mention a few ones:

- Any $g \in \mathbf{S}_0(\mathbb{R}^d)$ defines a Gabor family with the Bessel property, for any lattice $\Lambda \lhd \mathbb{R}^d \times \widehat{\mathbb{R}}^d$;
- The dual Gabor window depends continuously on both g and Λ, if everything is measured in the $\mathbf{S}_0$-norm (joint work with N. Kaiblinger);
- windows from $\mathbf{S}_0(\mathbb{R}^d)$ are optimal for proving the Janssen representation, or the definition of Gabor multipliers (joint work with F. Luef);
- $\mathbf{S}_0(G)$ is the correct setting for questions of non-commutative geometry (Franz Luef);

Basic properties of $\mathbf{M}^\infty(\mathbb{R}^d) = \mathbf{S}_0'(\mathbb{R}^d)$

The dual space of $\left(\mathbf{S}_0(\mathbb{R}^d), \|\cdot\|_{\mathbf{S}_0}\right)$, i.e. $\mathbf{S}_0'(\mathbb{R}^d)$ is the *largest* (reasonable) Banach space of distributions (resp. local pseudo-measures) which is isometrically invariant under all time-frequency shifts. As an amalgam space one has

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

the space of translation bounded quasi-measures, however it is much better to think of it as the modulation space $\mathbf{M}^\infty(\mathbb{R}^d)$, i.e. the space of all tempered distributions on $\mathbb{R}^d$ with bounded Short-time Fourier transform (for an arbitrary $0 \neq g \in \mathbf{S}_0(\mathbb{R}^d)$). Consequently norm convergence in $\mathbf{S}_0'(\mathbb{R}^d)$ is just uniform convergence of the STFT, while certain atomic characterizations of $\left(\mathbf{S}_0(\mathbb{R}^d), \|\cdot\|_{\mathbf{S}_0}\right)$ imply that w^*-convergence is in fact equivalent to locally uniform convergence of the STFT. – Hifi recordings!

The Dual Space and Fourier Transforms

Together with the space of test functions it is also of course the dual space which receives attention in this case. The Fourier transform extends in the usual way, via

$$\hat{\sigma}(f) := \sigma(\hat{f})$$

. We have already seen that

$$\mathcal{F}(\sqcup\!\sqcup_\Lambda) = C_\Lambda \sqcup\!\sqcup_{\Lambda^\perp} .$$

Since it is also true that the extended (forward or inverse) Fourier transform interchange the role of convolution and pointwise multiplication this formula can be used to proof that sampling (i.e. pointwise multiplication with $\sqcup\!\sqcup_\Lambda$ is turned into periodization of $\hat{f}$ with respect to the orthogonal lattice $\Lambda^\perp$. This can be taken as the basis for a simple proof of Shannon's sampling theorem (full reconstruction under Nyquist density assumption).

The Dual Space and Fourier Transforms

Distributions in $\mathbf{S}_0'(\mathbb{R}^d)$ are "not so bad", therefore it is easy to get rid of their lack of smoothness respectively decay at infinity by applying a product convolution operator.

Lemma

We have

$$\mathbf{S}_0 * (\mathbf{S}_0' \cdot \mathbf{S}_0) \subseteq \mathbf{S}_0$$

as well as

$$\mathbf{S}_0 \cdot (\mathbf{S}_0' * \mathbf{S}_0) \subseteq \mathbf{S}_0$$

Of course these statements can be used to regularize, i.e. approximate σ by convolving with Dirac-like compressed Gaussians and multiplying with slowly decaying (dilated) versions of some summability kernel, say.

Relevance of the w^*-approximation

The above statement about regularization indicates in which sense the of $\mathbf{S}_0'(\mathbb{R}^d)$ can be approximated by elements from the much smaller space $\mathbf{S}_0(\mathbb{R}^d)$ of test functions: It is *NOT in the sense of the norm!* but only in the so-called w^*-convergence. Let us recall the formal definition:

Definition

A net of distributions $(\sigma_\alpha)_{\alpha \in I}$ is w^*-convergent to a limit σ_0 if

$$lim_\alpha \sigma_\alpha(f) = \sigma_0(f) \quad \forall f \in \mathbf{S}_0(\mathbb{R}^d).$$

It is (only) in this sense that convolution with a Dirac sequence followed by multiplication with a summability kernel (or vice versa) allow to approximate a distribution by test functions. Most of the vague statements described in the introductory section can be resolved if one takes the w^* concept of convergence.

Relevance of the w^*-approximation

There is a long series of other situations where one has (only) w^*-convergence!

- for $x_n \to x_0$ one has $w^*-\lim \delta_{x_n} = \delta_{x_0}$;
- (finite or infinite) Riemannian sums, viewed as linear functionals on $\mathbf{S}_0(\mathbb{R}^d)$ converge in the w^*-sense to the Lebesgue (resp. Riemann) integral;
- mass-preserving dilation converges to δ_0 by stretching of the lattice;

The Kernel Theorem

One of the important statements about the Schwartz space is the so-called **kernel theorem**. The existence of if this theorem, that allows to "represent" linear operators using a kernel = (nucleus) is the reason why $\mathcal{S}(\mathbb{R}^d)$ is called a *nuclear Frechet-space*. But:

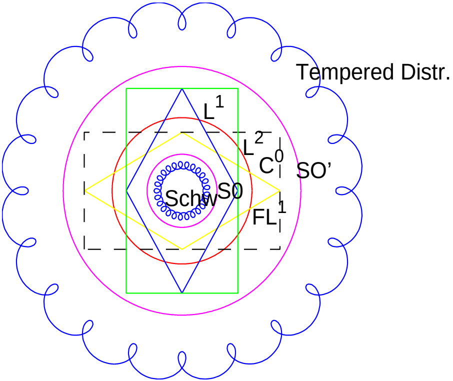
Theorem

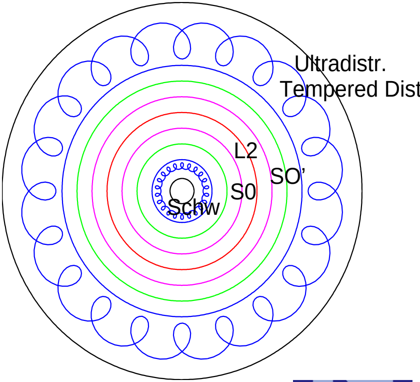
Every linear operator from $\mathbf{S}_0(\mathbb{R}^d)$ *to* $\mathbf{S}_0'(\mathbb{R}^d)$ *can be characterized by unique distribution from* $\sigma \in \mathbf{S}_0'(\mathbb{R}^{2d})$ *in the sense that*

$$\sigma(f \otimes g) = Tf(g), \quad \forall f, g \in \mathbf{S}_0(\mathbb{R}^d).$$

If σ was even a test function $K(.,.) \in \mathbf{S}_0(\mathbb{R}^{2d})$ one would say that $Tf(x) = \int_{\mathbb{R}^d} K(x,y)f(y)dy$, and $K(x,y) = \delta_y(T(\delta_k))$, the perfect analogue of matrix representation of linear mappings on $\mathbb{R}^n$: $T(x) = A * x$, with $a_{n,k} = \langle T(e_k), e_n \rangle$.

ANALYSIS: Spaces used to describe the Fourier Transform





The Fourier transform as a prototypical example

We all know that the (classical) Fourier transform is well defined on the space $\mathbf{L}^1(\mathbb{R}^d)$ of Lebesgue integrable functions via

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

The inverse Fourier transform then has the form

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

But is $\mathbf{L}^1(\mathbb{R}^d)$ the right setting. For pointwise computation by sampling it is too large, also for Fourier inversion, while on the other hand it is too small to show that pure frequencies χ_s are mapped the corresponding Dirac measure δ_s!

The idea of Banach Gelfand Triples

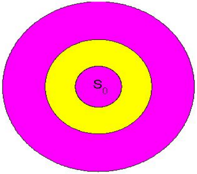
- that Hilbert spaces are themselves a too narrow concept and should be replaced Banach Gelfand Triples, ideally isomorphic to the canonical ones $(\ell^1, \ell^2, \ell^\infty)$;
- Demonstrate by examples (Fourier transform, kernel theorem) that this viewpoint brings us very close to the finite-dimensional setting!
- We could go on and show that the usual generalizations of linear algebra concepts to the Hilbert space case (namely linear independence and totality) are inappropriate in many cases and should be replaced by frame and Riesz basis, in fact by commutative diagrams in the category of BGTRs.

Banach Gelfand Triples and Rigged Hilbert space

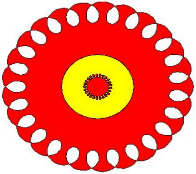
The next term to be introduced are Banach Gelfand Triples.
There exists already and established terminology concerning triples of spaces, such as the Schwartz triple consisting of the spaces $(\mathcal{S}, \mathbf{L}^2, \mathcal{S}')(\mathbb{R}^d)$, or triples of weighted Hilbert spaces, such as $(\mathbf{L}^2_w, \mathbf{L}^2, \mathbf{L}^2_{1/w})$, where $w(t) = (1 + |t|^2)^{s/2}$ for some $s > 0$, which is - via the Fourier transform isomorphic to another ("Hilbertian") Gelfand Triple of the form $(\mathcal{H}_s, \mathbf{L}^2, \mathcal{H}_s{}')$, with a Sobolev space and its dual space being used e.g. in order to describe the behaviour of elliptic partial differential operators.
The point to be made is that suitable Banach spaces, in fact imitating the prototypical Banach Gelfand triple $(\ell^1, \ell^2, \ell^\infty)$ allows to obtain a surprisingly large number of results resembling the finite dimensional situation.

Different Gelfand Triples

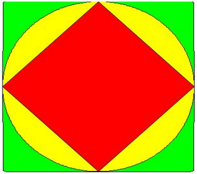
Fei-BGTr



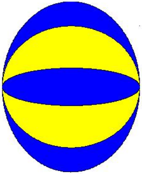
Schwartz GTr



l^1, l^2, l^∞



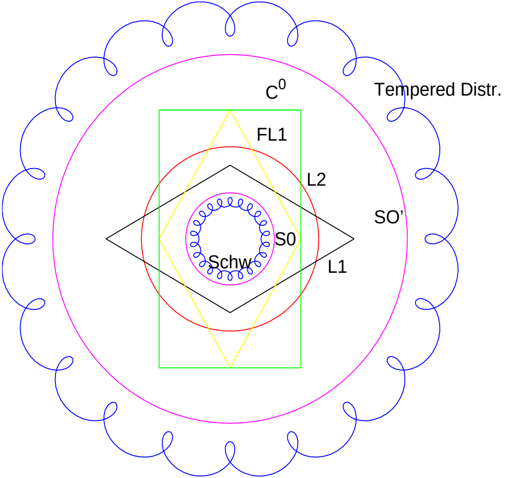
Sobolev GTr



A Classical Example related to Fourier Series

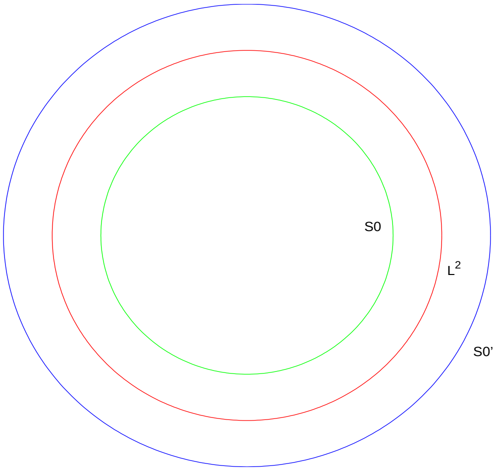
There is a well known and classical example related to the more general setting I want to describe, which - as so many things - go back to N. Wiener. He introduced (within $\mathbf{L}^2(\mathbb{U})$) the space $\left(\mathbf{A}(\mathbb{U}), \|\cdot\|_{\mathbf{A}}\right)$ of absolutely convergent Fourier series. Of course this space sits inside of $\left(\mathbf{L}^2(\mathbb{U}), \|\cdot\|_2\right)$ as a dense subspace, with the norm $\|f\|_{\mathbf{A}} := \sum_{n\in\mathbb{Z}} |\widehat{f}(n)|$.
Later on the discussion about Fourier series and generalized functions led (as I believe naturally) to the concept of pseudo-measures, which are either the elements of the dual of $\left(\mathbf{A}(\mathbb{U}), \|\cdot\|_{\mathbf{A}}\right)$, or the (generalized) inverse Fourier transforms of bounded sequences, i.e. $\mathcal{F}^{-1}(\ell^\infty(\mathbb{Z}))$.
In other words, this extended view on the Fourier analysis operator $\mathcal{C} : f \mapsto (\widehat{f}(n)_{n\in\mathbb{Z}})$ on the BGT $(\mathbf{A}, \mathbf{L}^2, \mathbf{PM})$ into $(\ell^1, \ell^2, \ell^\infty)$ is the prototype of what we will call a BGT-isomorphism.

A schematic description of the situation



The visualization of a Banach Gelfand Triple

The S_0 Gelfand triple



ANALYSIS: Calculating with all kind of numbers

We teach in our courses that there is a huge variety of *NUMBERS*, but for our daily life rationals, reals and complex numbers suffice. The most beautiful equation

$$e^{2\pi i} = 1.$$

It uses the exponential function, with a (purely) imaginary exponent to get a nice result, more appealing than (the equivalent)

$$cos(2\pi) + i * sin(2\pi) = 1 \quad \text{in} \quad \mathbb{C}.$$

But actual computation are done for rational numbers only!! Recall

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

The Banach space $(\mathbf{S}_0(\mathbb{R}^d), \|\cdot\|_{\mathbf{S}_0})$

Without differentiability there is a *minimal, Fourier and isometrically translation invariant Banach space* (called $(\mathbf{S}_0(\mathbb{R}^d), \|\cdot\|_{\mathbf{S}_0})$ or $(\mathbf{M}^1(\mathbb{R}^d), \|\cdot\|_{\mathbf{M}^1})$), which will serve our purpose. Its dual space $(\mathbf{S}_0'(\mathbb{R}^d), \|\cdot\|_{\mathbf{S}_0'})$ is correspondingly the largest among all Fourier invariant and isometrically translation invariant "objects" (in fact so-called local pseudo-measures or quasimeasures, orginally introduced in order to describe translation invariant systems as convolution operators).

Although there is a rich zoo of Banach spaces around (one can choose such a family, the so-called Shubin classes - to intersect in the Schwartz class and their union is corresondingly $\mathcal{S}'(\mathbb{R}^d)$), we will restrict ourselves to Banach Gelfand Triples, mostly related to $(\mathbf{S}_0, \mathbf{L}^2, \mathbf{S}_0')(\mathbb{R}^d)$.

BANACH GELFAND TRIPLES: a new category

Definition

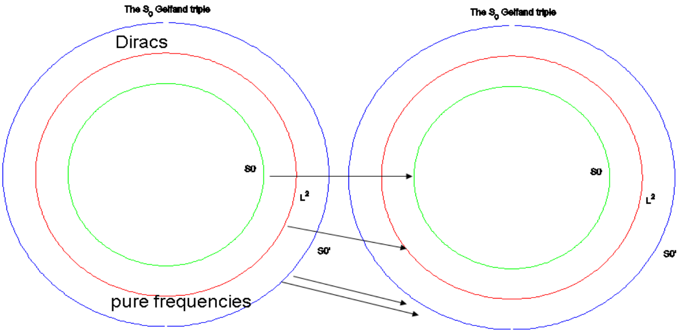
A triple, consisting of a Banach space $\mathbf{B}$, which is dense in some Hilbert space $\mathcal{H}$, which in turn is contained in $\mathbf{B}'$ is called a Banach Gelfand triple.

Definition

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

1. A is an isomorphism between $\mathbf{B}_1$ and $\mathbf{B}_2$.
2. A is [a unitary operator resp.] an isomorphism between $\mathcal{H}_1$ and $\mathcal{H}_2$.
3. A extends to a weak* isomorphism as well as a norm-to-norm continuous isomorphism between $\mathbf{B}_1'$ and $\mathbf{B}_2'$.

Gelfand triple mapping



Banach Gelfand Triples, ctc.

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 **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 $\mathcal{H} = \mathbf{L}^2([0,1])$, i.e. the corresponding definition is already around since the times of N. Wiener: $\mathbf{A}(\mathbb{U})$, the space of absolutely continuous Fourier series. It is also not surprising in retrospect to see that the dual space $\mathbf{PM}(\mathbb{U}) = \mathbf{A}(\mathbb{U})'$ 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 $(\mathbf{A}, \mathbf{L}^2, \mathbf{PM})(\mathbb{U})$ and $(\ell^1, \ell^2, \ell^\infty)(\mathbb{Z})$.

The BGT $(\mathbf{S}_0, \mathbf{L}^2, \mathbf{S}_0')$ and Wilson Bases

Among the many different orthonormal bases the wavelet bases turn out to be exactly the ones which are well suited to characterize the distributions by their membership in the classical Besov-Triebel-Lizorkin spaces.

For the analogue situation (using the modulation operator instead of the dilation, resp. the Heisenberg group instead of the "ax+b"-group) on finds that *local Fourier bases* resp. the so-called Wilson-bases are the right tool. They are formed from tight Gabor frames of redundancy 2 by a particular way of combining complex exponential functions (using Euler's formula) to cos and sin functions in order to build a Wilson ONB for $\mathbf{L}^2(\mathbb{R}^d)$.

In this way another BGT-isomorphism between $(\mathbf{S}_0, \mathbf{L}^2, \mathbf{S}_0')$ and $(\ell^1, \ell^2, \ell^\infty)$ is given, for each concrete Wilson basis.

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

Furthermore, we have that Parseval's formula

$$\langle f, g \rangle = \langle \widehat{f}, \widehat{g} \rangle \qquad (3)$$

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

The w^*- topology: a natural alternative

It is not difficult to show, that the norms of $(\mathbf{S}_0, \mathbf{L}^2, \mathbf{S}_0')(\mathbb{R}^d)$ correspond to norm convergence in $(\mathbf{L}^1, \mathbf{L}^2, \mathbf{L}^\infty)(\mathbb{R}^{2d})$.
The FOURIER transform, viewed as a BGT-automorphism is uniquely determined by the fact that it maps pure frequencies onto the corresponding point measures δ_ω.
This is a typical case, where we can see, that the w^*-continuity plays a role, and where the fact that $\delta_x \in \mathbf{S}_0'(\mathbb{R}^d)$ as well as $\chi_s \in \mathbf{S}_0'(\mathbb{R}^d)$ are important.
In the STFT-domain the w^*-convergence has a particular meaning: a sequence σ_n is w^*-convergent to σ_0 if $V_g(\sigma_n)(\lambda) \to V_g(\sigma_0)(\lambda)$ uniformly over compact subsets of the TF-plane (for one or any $g \in \mathbf{S}_0(\mathbb{R}^d)$).

Kernel Theorem for general operators in $\mathcal{L}(\mathbf{S}_0, \mathbf{S}_0')$

Theorem

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

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

Kernel Theorem II: Hilbert Schmidt Operators

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 $\mathbf{L}^2(\mathbb{R}^d)$ are just those compact operators which arise as integral operators with $\mathbf{L}^2(\mathbb{R}^{2d})$-kernels. The complete picture can be best expressed by a unitary Gelfand triple isomorphism. First the innermost shell:

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 $\mathbf{L}^2(\mathbb{R}^{2d})$ on the kernels. An operator T has a kernel in $K \in \mathbf{S}_0(\mathbb{R}^{2d})$ if and only if the T maps $\mathbf{S}_0'(\mathbb{R}^d)$ into $\mathbf{S}_0(\mathbb{R}^d)$, boundedly, but continuously also from w^*-topology into the norm topology of $\mathbf{S}_0(\mathbb{R}^d)$.

Kernel Theorem III

Remark: Note that for such regularizing kernels in $K \in \mathbf{S}_0(\mathbb{R}^{2d})$ the usual identification. Recall that the entry of a matrix $a_{n,k}$ is the coordinate number n of the image of the $n-$th unit vector under that action of the matrix $A = (a_{n,k})$:

$$k(x, y) = T(\delta_y)(x) = \delta_x(T(\delta_y)).$$

Note that $\delta_y \in \mathbf{S}_0'(\mathbb{R}^d)$ implies that $K(\delta_y) \in \mathbf{S}_0(\mathbb{R}^d)$ by the regularizing properties of K, hence the pointwise evaluation makes sense.

With this understanding the kernel theorem provides a (unitary) isomorphism between the Gelfand triple (of kernels) $(\mathbf{S}_0, \mathbf{L}^2, \mathbf{S}_0')(\mathbb{R}^{2d})$ into the Gelfand triple of operator spaces

$$(\mathcal{L}(\mathbf{S}_0', \mathbf{S}_0), \mathcal{HS}, \mathcal{L}(\mathbf{S}_0, \mathbf{S}_0')).$$

AN IMPORTANT TECHNICAL warning!!

How should we realize these various BGT-mappings?
Recall: How can we check numerically that $e^{2\pi i} = 1$??
Note: we can only do our computations (e.g. multiplication, division etc.) properly in the rational domain $\mathbb{Q}$, we get to $\mathbb{R}$ by approximation, and then to the complex numbers applying "the correct rules" (for pairs of real numbers).
In the BGT context it means: All the (partial) Fourier transforms, integrals etc. only have to be meaningful at the $\mathbf{S}_0$-level!!! (no Lebesgue even!), typically isometric in the $\mathbf{L}^2$-sense, and extend by duality considerations to $\mathbf{S}_0'$ when necessary, using w^*-continuity!
The Fourier transform is a good example (think of Fourier inversion and summability methods), similar arguments apply to the transition from the integral kernel of a linear mapping to its Kohn-Nirenberg symbol., e.g..

The Spreading Representation

The kernel theorem corresponds of course to the fact that every linear mapping T from $\mathbb{C}^n$ to $\mathbb{C}^n$ can be represented by a uniquely determined matrix $\mathbf{A}$, whose columns $\mathbf{a}_k$ are the images $T(\vec{e}_k)$. When we identify $\mathbb{C}^N$ with $\ell^2(\mathbf{Z}_N)$ (as it is suitable when interpreting the FFT as a unitary mapping on $\mathbb{C}^N$) there is another way to represent every linear mapping: we have exactly N cyclic shift operators and (via the FFT) the same number of frequency shifts, so we have exactly N^2 TF-shifts on $\ell^2(\mathbf{Z}_N)$. They even form an orthonormal system with respect to the Frobenius scal.prod.:

$$\langle \mathbf{A}, \mathbf{B} \rangle_{Frob} := \sum_{k,j} a_{k,j} \bar{b}_{k,j} = trace(A * B')$$

This relationship is called the spreading representation of the linear mapping T resp. of the matrix $\mathbf{A}$. It is a kind of operator version of the Fourier transform.

The unitary spreading BGT-isomorphism

Theorem

There is a natural (unitary) Banach Gelfand triple isomorphism, called the spreading mapping, which assigns to operators T from $(\mathbf{B}, \mathcal{H}, \mathbf{B}')$ the function or distribution $\eta(T) \in (\mathbf{S}_0, \mathbf{L}^2, \mathbf{S}_0')(\mathbb{R}^{2d})$. It is uniquely determined by the fact that $T = \pi(\lambda) = M_\omega T_t$ corresponds to $\delta_{t,\omega}$.

Via the symplectic Fourier transform, which is of course another unitary BGT-automorphism of $(\mathbf{S}_0, \mathbf{L}^2, \mathbf{S}_0')(\mathbb{R}^{2d})$ we arrive at the Kohn-Nirenberg calculus for pseudo-differential operators. In other words, the mapping $T \mapsto \sigma_T = \mathcal{F}_{symp}\eta(T)$ is another unitary BGT isomorphism (onto $(\mathbf{S}_0, \mathbf{L}^2, \mathbf{S}_0')(\mathbb{R}^{2d})$, again).

Selection of bibliographic items, see www.nuhag.eu

A copy of this talk can be downloaded from the talk-server at **www.nuhag.eu** ($>$ DB+tools), as well as corresponding literature.

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

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

N. Kaiblinger.
Approximation of the Fourier transform and the dual Gabor window.
J. Fourier Anal. Appl., 11(1):25–42, 2005.