

Fourier Analysis using Mild Distributions

Hans G. Feichtinger, Univ. Vienna

`hans.feichtinger@univie.ac.at`

`https://nuhagphp.univie.ac.at/home/fei.php`

September 4th, 2025

AIMC56 CONFERENCE



Links to my Homepage

All my papers are found at

https://nuhagphp.univie.ac.at/home/feipub_db.php

and my talks are at www.nuhag.eu/feitalks (OR)

<https://nuhagphp.univie.ac.at/home/feitalks.php>

Maybe my alternative approach to Fourier analysis as provided by

www.nuhag.eu/ETH20 is found interesting (look for ETH EXTRA Lecture Notes 2024 in the list of publications).

Other lecture notes should be found at

<https://nuhagphp.univie.ac.at/home/skripten.php>

Feel free to contact me if you have access problems.

hans.feichtinger@univie.ac.at



General Views

To get started a few personal comments:

I consider myself an *application oriented harmonic analyst*. From my education I was starting mathematically with **abstract harmonic analysis (AHA)** under Hans Reiter, while I was also a teacher student in *physics*.

Over the years I have supervised quite a few PhD students and realized many joint projects with applied people from different areas, especially in the context of engineering applications.

While I consider myself an expert on function spaces the goal of most of my research is to make methods from Fourier Analysis useful for the applied community, to reveal mathematical structures to make dubious objects like the Dirac delta accessible, and so on. Sometimes it was necessary to develop new methods, sometime it is interesting to explore the usefulness of new tools to classical questions.



The Key Topics relevant for this talk

Short Historical Overview:

- 1 Fourier Series (1822)
- 2 Fourier Transforms via Lebesgue (ca. 1922)
- 3 Fourier Transform for tempered distributions (around 1950)
- 4 the Fast Fourier Transform (1965)
- 5 Gabor Analysis (1946, mathematics starting 1980)
- 6 Coherent states, phase space analysis (Folland 1989)
- 7 Wiener Amalgam spaces (HGFei, 1980)
- 8 $S_0(G)$ (Feichtinger algebra), *modulation spaces* (1983)
- 9 Wavelet Analysis (1986), painless expansions
- 10 Coorbit Theory (1989) (with K.Gröchenig)



A quick approach to Mild Distributions I

Throughout I will use the normalization of the Fourier Transform

$$\widehat{f}(s) = \int_{\mathbb{R}^d} f(t) e^{-2\pi i s t} dt = \int_{\mathbb{R}^d} f(t) \overline{\chi_s(t)} dt.$$

Using the Schwartz space $\mathcal{S}(\mathbb{R}^d)$ of *rapidly decreasing functions* the extended *Fourier transform* for the dual space, the space of *tempered distributions* is defined via $\widehat{\sigma}(f) = \sigma(\widehat{f})$, $f \in \mathcal{S}(\mathbb{R}^d)$. It is convenient to use the Gauss function $g_0(t) = e^{-\pi|t|^2}$ because it is Fourier invariant: $\mathcal{F}(g_0) = \widehat{g_0} = g_0 \in \mathcal{S}(\mathbb{R}^d)$.

Then we can define the *STFT (Short Time Fourier Transform)*

$$V_g \sigma(t, s) = \sigma(M_s T_t g_0),$$

with $T_t f(x) = f(x - t)$, $M_s f(t) = \chi_s(t) f(t)$. It can be shown that $V_g \sigma$ of at most polynomial growth for any $\sigma \in \mathcal{S}'(\mathbb{R}^d)$.



Setting the stage I

In this situation we can define the modulation space $(M^p(\mathbb{R}^d), \|\cdot\|_{M^p})$, for $1 \leq p \leq \infty$, by setting

$$M^p(\mathbb{R}^d) := \{\sigma \in \mathcal{S}'(\mathbb{R}^d) \mid V_g \sigma \in L^p(\mathbb{R}^{2d})\}$$

and endow each such space with the norm

$$\|\sigma\|_{M^p(\mathbb{R}^d)} := \|V_g \sigma\|_{L^p(\mathbb{R}^{2d})}$$

It is not difficult to show that each of these spaces is a [Banach spaces](#), and in fact also [Fourier invariant](#).

Furthermore one has the duality results

$$(M^p(\mathbb{R}^d), \|\cdot\|_{M^p})' = (M^q(\mathbb{R}^d), \|\cdot\|_{M^q}), \text{ for } 1 \leq p < \infty.$$



Feichtinger Algebra and Mild Distributions

We then have the following chain of dense embeddings:

$$\mathcal{S}(\mathbb{R}^d) \hookrightarrow \mathbf{M}^1(\mathbb{R}^d) \hookrightarrow \mathbf{M}^p(\mathbb{R}^d) \hookrightarrow \mathbf{M}^\infty(\mathbb{R}^d) \hookrightarrow \mathcal{S}'(\mathbb{R}^d)$$

and using my more “classical notation”

$(\mathbf{S}_0(\mathbb{R}^d), \|\cdot\|_{\mathbf{S}_0}) = (\mathbf{M}^1(\mathbb{R}^d), \|\cdot\|_{\mathbf{M}^1})$ (Feichtinger algebra)
and $(\mathbf{S}'_0(\mathbb{R}^d), \|\cdot\|_{\mathbf{S}'_0}) = (\mathbf{M}^\infty(\mathbb{R}^d), \|\cdot\|_{\mathbf{M}^\infty})$ (mild distributions) we
have the following “rigged Hilbert space” situation:

$$(\mathbf{S}_0(\mathbb{R}^d), \|\cdot\|_{\mathbf{S}_0}) \hookrightarrow = (\mathbf{L}^2(\mathbb{R}^d), \|\cdot\|_2) \hookrightarrow (\mathbf{S}'_0(\mathbb{R}^d), \|\cdot\|_{\mathbf{S}'_0}),$$

the last embedding being w^* -dense only.

We summarize this situation by calling

$(\mathbf{S}_0, \mathbf{L}^2, \mathbf{S}'_0)(\mathbb{R}^d)$ THE Banach Gelfand Triple.

Via *Wilson bases* it can be identified with $(\ell^1, \ell^2, \ell^\infty)$
(the *canonical Banach Gelfand Triple*).



Official Abstract Provided I

Taking the current viewpoint classical Fourier Analysis started from the study the orthogonal expansion of periodic functions in the Hilbert space $L^2(\mathbb{T})$, using the orthonormal system of pure frequencies (described by complex exponential functions). Moving on to the real line it was natural to view the Fourier transform, given by the usual integral formula, as a linear mapping defined on the Lebesgue space $L^1(\mathbb{R}^d)$, mapping into the space $C_0(\mathbb{R}^d)$ of continuous functions, vanishing at infinity (by the Riemann-Lebesgue Lemma). The convolution theorem then goes on and describes this mapping as an injective Banach algebra homomorphism, converting convolution into pointwise multiplication. This is also the basis for the study of summability methods which are used in order to verify the validity of Fourier inversion formulas.



Official Abstract Provided II

As it turned out, L^p -spaces are not so well suited for the study of the Fourier transform (only $L^2(\mathbb{R}^d)$ is invariant) and even less for the modern theory of Time-Frequency Analysis (TFA) or Gabor Analysis (a branch of TFA dealing with discrete representations). Instead, modulation spaces are better suited. The prototypes in this family are the Segal algebra $\mathcal{S}_0(\mathbb{R}^d)$, the so-called **Feichtinger algebra**, and its dual, meanwhile known as Banach space $\mathcal{S}'_0(\mathbb{R}^d)$ of *mild distributions*. Together with the Hilbert space $L^2(\mathbb{R}^d)$ they form **THE BGT: Banach Gelfand Triple** $(\mathcal{S}_0, L^2, \mathcal{S}'_0)(\mathbb{R}^d)$, which behaves in many ways like the triple $\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$ of $\mathbb{Q}$ (rational), $\mathbb{R}$ (real) and $(\mathbb{C}$: complex) numbers. Although these spaces can be defined and are useful in the context of general locally compact Abelian (LCA) groups we will discuss a (long) list of features making them extremely useful, for the study of both theoretical and applied (say in signal processing) problems.



Official Abstract Provided III

THE BGT allows for an elegant derivation of the basic principles of Fourier Analysis, but it is also a central tool for TFA and Gabor Analysis. It also provides a solid foundation of a mind-set called Conceptual Harmonic Analysis (CHA). Although it is pursued by the author now for several years, it still lacks popularity, competing with traditional approaches.

Overall CHA suggests to promote the integration of concepts from basic functional and harmonic analysis with numerical computations via efficient algorithms. It also supports the modelling real world problems (such as recovery of signals from sampling), thus avoiding purely heuristic “hand-waving” arguments.



Executive Summary I

We use the term *Mild Distributions* for the elements of the dual space for the Feichtinger algebra $(\mathbf{S}_0(\mathbb{R}^d), \|\cdot\|_{\mathbf{S}_0})$, also known as *modulation space* $(\mathbf{M}^\infty(\mathbb{R}^d), \|\cdot\|_{\mathbf{M}^\infty})$.

It is well known that many problems in Fourier and Time-Frequency Analysis can be well described using *THE Banach Gelfand Triple* $(\mathbf{S}_0, \mathbf{L}^2, \mathbf{S}'_0)(\mathbb{R}^d)$. The many good properties of $(\mathbf{S}_0(G), \|\cdot\|_{\mathbf{S}_0})$ allow to use it, typically as a replacement to the Schwartz space $\mathcal{S}(\mathbb{R}^d)$, especially in the context of LCA groups. It is used in the context of *Gabor Analysis* and appears implicitly in Classical Fourier Analysis, since summability kernels belong to $\mathbf{S}_0(\mathbb{R}^d)$.

It is the purpose of this talk to feature the usefulness of mild distributions, e.g., by emphasizing the problems of the traditional approach using the *Lebesgue spaces* $\mathbf{L}^p(\mathbb{R}^d)$.



Using the Fourier Transform I

One can look at *Fourier Analysis* or more generally *Time-Frequency Analysis* from many different angles.

The technical/mathematical viewpoint takes the Fourier transform as an integral transform and tries to extend it to the Hilbert space $(L^2(\mathbb{R}^d), \|\cdot\|_2)$ (a unitary automorphism, by Plancherel's Theorem), or to (tempered) distributions (via duality).

While engineers typically do similar things in a *discrete context* it is in the world of physics where “*continuous bases*” of Dirac measures $(\delta_x)_{x \in \mathbb{R}^d}$ are used for heuristic arguments.

Abstract Harmonic Analysis (AHA) classifies the different types of Fourier transforms as realizations of a general abstract principle (dual group of “pure frequencies”), with the DFT/FFT corresponding to the use of the cyclic group $\mathbb{Z}_N$ of order N .



Using the Fourier Transform II

Looking around in the real world and without formally defining what a “signal” IS we can say:

Very few signals in the real world are either periodic (in a strict sense), or well defined almost everywhere and square integrable (meaning belonging to $L^2(\mathbb{R}^d)$). Nevertheless we have nowadays very good **signal processing algorithms** which help us to record, transmit, modify signals in many ways, even in real time. But is the established mathematical theory helping us (except for heuristic considerations)?

Recall, that **TF-shifts** are defined for $\lambda = (t, s) \in \mathbb{R}^d \times \hat{\mathbb{R}}^d$:

$$\pi(\lambda)g(x) = [M_s T_t]g(x) = e^{2\pi i s x} g(x - t).$$



Basic Considerations concerning Mild Distributions I

Just recall the definition of $(\mathcal{S}'_0(\mathbb{R}^d), \|\cdot\|_{\mathcal{S}'_0})$. First we need the definition of the STFT of a tempered distribution with respect to the Gaussian window $g_0(t) = \exp(-\pi|t|^2)$: For $\sigma \in \mathcal{S}'(\mathbb{R}^d)$ set:

$$V_{g_0}\sigma(t, s) = \sigma(M_{-s}T_s g_0), \text{ with } \lambda = (t, s) \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d. \quad (1)$$

We can define $\mathcal{S}'_0(\mathbb{R}^d) = \mathcal{M}^\infty(\mathbb{R}^d)$ by setting

$$\mathcal{S}'_0(\mathbb{R}^d) := \{\sigma \in \mathcal{S}'(\mathbb{R}^d) \mid V_{g_0}\sigma \in \mathcal{C}_b(\mathbb{R}^{2d})\}, \quad \|\sigma\|_{\mathcal{S}'_0} := \|V_{g_0}\sigma\|_\infty < \infty. \quad (2)$$

We have many good properties, such as invariance under the Fourier transform (defined in many equivalent ways), e.g. via

$$\widehat{\sigma}(f) = \sigma(\widehat{f}), \quad f \in \mathcal{S}_0(\mathbb{R}^d).$$



Basic Operations on Mild Distributions

- 1 First: $(\mathbf{S}'_0(\mathbb{R}^d), \|\cdot\|_{\mathbf{S}'_0})$ is a Banach space with the dual norm;
- 2 TF-shifts $\pi(\lambda) = M_s T_t$ act isometrically, automorphisms (e.g. dilations) act boundedly on $(\mathbf{S}_0(\mathbb{R}^d), \|\cdot\|_{\mathbf{S}_0})$;
- 3 Every $\sigma \in \mathbf{S}'_0$ has a Fourier transform; it is an isometric mapping on $(\mathbf{S}'_0(\mathbb{R}^d), \|\cdot\|_{\mathbf{S}'_0})$, preserving mild convergence, and characterized by $\mathcal{F}(\chi_s) = \delta_s$;
- 4 Every $\sigma \in \mathbf{S}'_0$ is the mild limit of test functions, e.g. from $\mathbf{S}_0(\mathbb{R}^d)$, which may be obtained by *regularization*, or of (finite discrete) measures, hence the unique extension of the usual integral version of the FT;
- 5 Fractional FTs, even *metaplectic operators* act boundedly.



Regularizations I

There are many ways to regularize a given mild distributions, such as a *Dirac comb* $\sqcup\sqcup = \sum_{k \in \mathbb{Z}^d} \delta_k$, which shows neither decay nor smoothness (decay in the frequency direction).

The typical methods to approximate (in the “mild sense”) a given $\sigma \in \mathbf{S}'_0$ is to convolve it with a Dirac sequence, normalized in $(L^1(\mathbb{R}^d), \|\cdot\|_1)$ but sitting inside of $\mathbf{S}_0(\mathbb{R}^d)$, hence its FT (a summability kernel) is a bounded approximate unit in the pointwise sense. Thus it is good to know that $\mathbf{S}_0(\mathbb{R}^d) * \mathbf{S}'_0(\mathbb{R}^d) \subset \mathbf{C}_b(\mathbb{R}^d)$ with corresponding norm estimates, but more importantly

$$(\mathbf{S}'_0(\mathbb{R}^d) * \mathbf{S}_0(\mathbb{R}^d)) \cdot \mathbf{S}_0(\mathbb{R}^d) \subseteq \mathbf{S}_0(\mathbb{R}^d)$$

$$(\mathbf{S}'_0(\mathbb{R}^d) \cdot \mathbf{S}_0(\mathbb{R}^d)) * \mathbf{S}_0(\mathbb{R}^d) \subseteq \mathbf{S}_0(\mathbb{R}^d).$$



Regularizations II

There is also a discrete alternative, which makes use of *Gabor frames* with windows $g \in \mathcal{S}(\mathbb{R}^d) \subset \mathcal{S}_0(\mathbb{R}^d)$. Think, for example of $g = g_0$, the standard Gaussian, with lattice $\Lambda = a\mathbb{Z}^d \times b\mathbb{Z}^d$, for $ab < 1$.

Then the dual window $\tilde{g}$ also belongs to $\mathcal{S}_0(\mathbb{R}^d)$ (by a result of Gröchenig and Leinert) and consequently one has:

$$\sigma \mapsto V_g \sigma|_{\Lambda} \in \ell^\infty(\Lambda)$$

and

$$\sigma = \sum_{\lambda \in \Lambda} V_g \sigma(\lambda) \pi(\lambda) \tilde{g},$$

is unconditionally mildly convergent.



Regularizations III

Obviously the finite partial sums belong to $\mathbf{S}_0(\mathbb{R}^d)$ (or even to $\mathcal{S}(\mathbb{R}^d)$) and so it is not surprising that the family of partial sums

$$\sigma_F = \sigma_{f,g,\Lambda,F} := \sum_{\lambda \in F} V_g \sigma(\lambda) \pi(\lambda) \tilde{g}$$

are mildly convergent to σ . In other words:

Given a compact domain $Q \subset \mathbb{R}^d \times \widehat{\mathbb{R}}^d$ and $\varepsilon > 0$ there exists some finite set F_0 (typically $\Lambda \cap (B_R(0) + Q)$, where R depends only on the concentration properties of g and $\tilde{g}$): such that one has for any finite set $F \supseteq F_0$:

$$|V_g \sigma(q) - V_g \sigma_F(q)| \leq \varepsilon, \quad \forall q \in Q.$$



The MP3-BUPU: 512 samples long

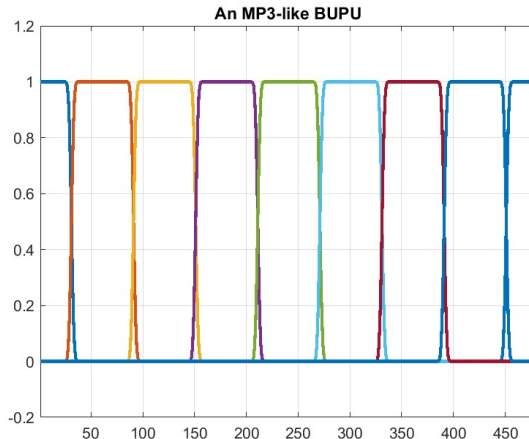


Figure: mp3bupu.jpg



Motivation from the MP3 coding method I

Next I will provide an “[a posteriori motivation](#)” for the construction, meaning that first there was the mathematical construction and only later it turned out that it is related to the MP3 coding scheme (see Peter Balazs PhD thesis from 2005). When it comes to the processing of music the best known compression algorithm for audio data is the *MP3 coding scheme*. It starts by cutting an audio signal into pieces of finite length, e.g. 30 seconds of 44100 samples corresponds to $30 * 4410 / 496$ slices of length $512 = 2^9$ ca. 2667 FFTs have to be performed. The slight overlap of the slices (by 16 samples) allows to ensure that the ends of each slice equals zero, so there is *NO JUMP* due to the cyclic prolongation inherent in the DFT/FFT.

In mathematical terms we apply a regular BUPU, a partition of unity which is generated by translation along the lattice $496\mathbb{Z}$, from a window of length 512 (a typical smooth plateau function).



Motivation from the MP3 coding method II

Given each position one thus obtains 512 Fourier coefficients, and the compression relies on the so-called *masking effect*, i.e. the effect that certain (low-amplitude) frequencies cannot be heard by human beings if they are close to more dominant ones. So they can be discarded without loss of information (from the point of view of perception).

Due to the fact that the DFT defines an isometry one can claim that the ℓ^2 -norm of each of the slices corresponds to the ℓ^2 -norm of the corresponding slice, and since these slices are almost orthogonal to each other the total ℓ^2 -norm (over all positions and all 512 frequencies) defines an equivalent norm on $(L^2(\mathbb{R}^d), \|\cdot\|_2)$.



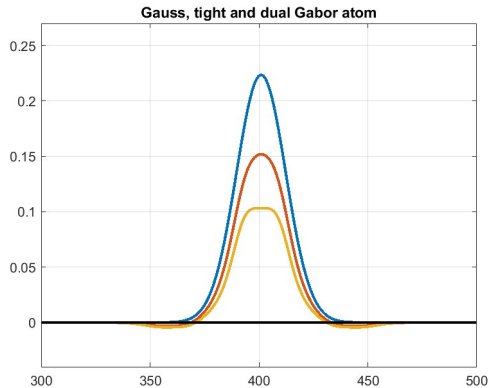
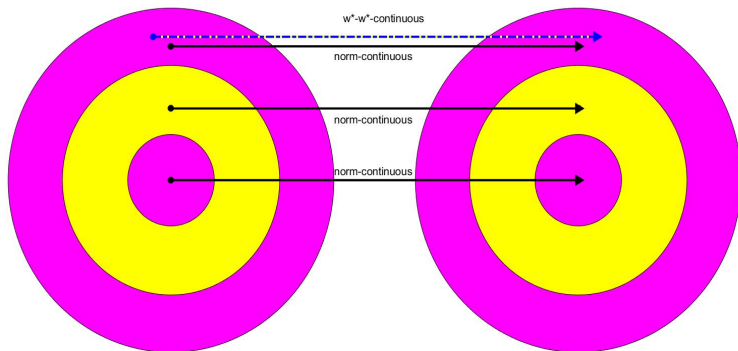


Figure: Gau8002tgtdu.jpg



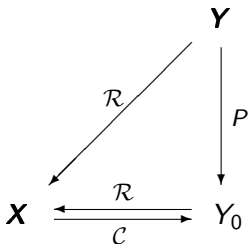
Banach-Gelfand-Tripel-Homomorphisms



DIAGRAMS I

Think of $\mathbf{X}$ as $(\mathbf{S}_0, L^2, \mathbf{S}'_0)$, and $\mathbf{Y} = (\ell^1, \ell^2, \ell^\infty)$:

BANACH (Gelfand Triple) FRAME case: $\mathcal{C}$ is injective, but not surjective, and $\mathcal{R}$ is a left inverse of $\mathcal{C}$. This implies that $P = \mathcal{C} \circ \mathcal{R}$ is a projection in $\mathbf{Y}$ onto the range Y_0 of $\mathcal{C}$ in $\mathbf{Y}$:



Modulation Spaces and Gabor Expansions I

The problem of Gabor expansions of “signals” is *equivalent* to the question of reconstructing a STFT $V_g \sigma$ from samples.

If the samples are taken from a lattice $\Lambda \triangleleft \mathbb{R}^d \times \widehat{\mathbb{R}}^d$ in phase space we call this a *regular Gabor expansions*. Often one takes $\Lambda = a\mathbb{Z}^d \times b\mathbb{Z}^d$. Otherwise we talk of *irregular Gabor families*.

Anyway, the usual consideration is that of a problem of **frame** in Hilbert space context, where the frame condition can be expressed by a simple pair of inequalities. In words:

The sampling energy is supposed to be equivalent to energy (L^2 -norm) of the signal.

But it turns out that at least for regular Gabor families generated by some $g \in \mathcal{S}_0(\mathbb{R}^d)$ that then one has in fact such a retract situation according to the above diagram.



Gabor Expansions of Mild signals I

Being defined as an adjoint mapping, which in fact extends the classical integral transform given by

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

which leaves $(\mathbf{S}_0(\mathbb{R}^d), \|\cdot\|_{\mathbf{S}_0}) \hookrightarrow (\mathbf{S}'_0(\mathbb{R}^d), \|\cdot\|_{\mathbf{S}'_0})$ isometrically invariant, we know that the (extended) FT is w^*-w^* -continuous, i.e. it preserves *mild convergence*.

But there is a more natural equivalent expression for this: A sequence (or bounded net) σ_α is **mildly convergent** to σ_0 if the corresponding STFTs are *uniformly convergent over compact subsets of $\mathbb{R}^d \times \widehat{\mathbb{R}}^d$* (phase space).



Gabor Expansions of Mild signals II

In the context of Gabor Analysis, let us choose on of the most simple cases, namely a Gaussian family of redundancy 2, with $a = 1/\sqrt{2} = b$. Then the elements of $(\mathbf{S}_0, \mathbf{L}^2, \mathbf{S}'_0)(\mathbb{R})$ can be characterized by the possibility of representing them as double series over the lattice, with coefficients in $(\ell^1, \ell^2, \ell^\infty)$.

Now there is a 'tight Gabor family' $(g_{n,k}) = (M_{bn}T_{ak}g_t)_{(n,k) \in \mathbb{Z}^2}$ (red shape below) which allows to write any $f \in (\mathbf{S}_0, \mathbf{L}^2, \mathbf{S}'_0)$ as

$$f = \sum_{(k,n) \in \mathbb{Z}^2} \langle f, g_{n,k} \rangle g_{n,k}. \quad (3)$$

The Fourier invariance of g_0 and the symmetry $a = b$ implies that also the tight version, i.e. this g is Fourier invariant, and thus

$$\mathcal{F}(g_{n,k}) = \mathcal{F}(M_{nb}T_{ka}g) = T_{nb}M_{-ka}\mathcal{F}(g) = e^{2\pi i knab}g_{-k,n}.$$



Gabor Expansions of Mild signals III

This implies that the Fourier transform can be simply applied (up to a well explained phase factor) as a simple permutation of Gabor atoms.

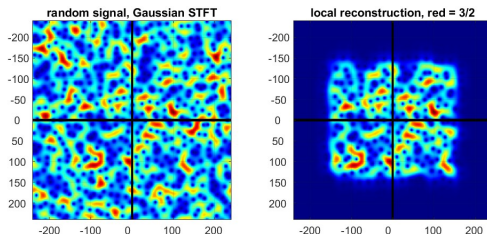


Figure: GabLocRec01.jpg



What are the USPs for the space of mild distributions?

A couple of bold claims!

- 1 It is large enough to contain (almost) everything
- 2 It is small and simple enough (a Banach space)
- 3 It can be defined on any LCA group, from scratch (ETH20)
- 4 It can serve as a *model for general SIGNALS*
- 5 It allows to describe *general OPERATORS*
- 6 It is easy/intuitively to use (no Lebesgue, little topology)
- 7 It suggests *structure preserving approximation schemes*
- 8 It covers discrete *and* continuous, periodic *and* non-periodic
- 9 It covers mostly the engineering or physicists viewpoint



P1: Simplicity I

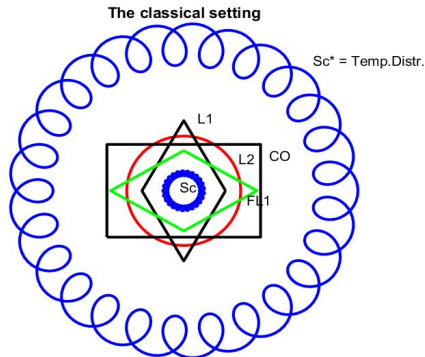


Figure: Classical spaces and tempered distributions

Classical spaces and Fourier

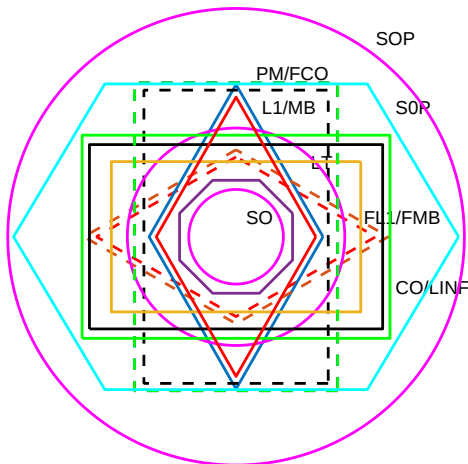


Figure: Classical spaces and Fourier

TOPIC P1: SIMPLICITY

The Banach Gelfand Triple (S_0, L^2, S_0^*)

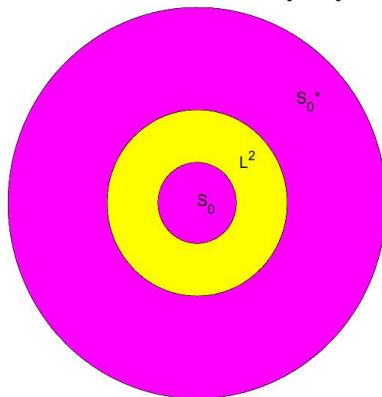


Figure: THE (simple) Banach Gelfand Triple

P2: Historical Remarks, Schwartz-Bruhat I

The space $(\mathbf{S}'_0(G), \|\cdot\|_{\mathbf{S}'_0})$ of *translation bounded quasi-measures* has been introduced shortly after the paper on the “New Segal Algebra” $(\mathbf{S}_0(G), \|\cdot\|_{\mathbf{S}_0})$ (which goes now by the name of Feichtinger’s algebra) in [?]. It appears in the Compt. Rend. Note [?]. This paper from 1980 has as many/few citations as the recent survey article [?] by Mads Jakobsen, which is a sign that recognition came rather late.

In many earlier talks I have tried to emphasize the possible role of the so-called **Banach Gelfand Triple** $(\mathbf{S}_0, L^2, \mathbf{S}'_0)$, especially for the Euclidean case, as a universal tool for many questions arising naturally in Fourier Analysis and Time-Frequency Analysis, in particular in the context of Gabor Analysis.

This construction (also called a “rigged Hilbert space” setting) is an important corner stone for the development of the idea of “**Conceptual Harmonic Analysis**”, because very often



P2: Historical Remarks, Schwartz-Bruhat II

Banach Gelfand Triple Morphisms (BGT-morphisms) allow to widen the view on *unitary mappings*. The prototypical example is Plancherel's Theorem. It can be obtained by observing that $(\mathbf{S}_0(\mathbb{R}^d), \|\cdot\|_{\mathbf{S}_0})$ is a space of continuous, Riemann integrable functions which is invariant under the Fourier transform. Thus the Fourier Inversion Theorem applies pointwise (even with Riemann integrals), and by observing the L^2 -isometric property one naturally extends the integral transform to a unitary automorphism of $(L^2(\mathbb{R}^d), \|\cdot\|_2)$. But finally the (unique, w^* - w^* -continuous) extension to all of $\mathbf{S}'_0(\mathbb{R}^d)$ allows to observe that it just maps pure frequencies into Dirac Deltas, similar to the fact that the DFT maps pure frequencies on $\mathbb{Z}_N$ into corresponding unit vectors in the frequency domain.



P2: Historical Remarks, Schwartz-Bruhat III

So this space is around for 45 years, and I have been using it a lot for various applications. But still, its usefulness outside of time-frequency analysis (where it is usually described as the modulation space) is not so clear within a wider community. Therefore it is the purpose of this talk to provide a list of facts and properties of $\mathcal{S}'_0(G)$ which make it so useful. We will also explain the effect of fundamental properties of $(\mathcal{S}_0(G), \|\cdot\|_{\mathcal{S}_0})$ for the various descriptions of $\mathcal{S}'_0(G)$, including the characterization of w^* -convergent sequences (or better bounded nets). As a final introductory remark let me note that (for a similar reason) I decided to call the elements of $\mathcal{S}'_0(\mathbb{R}^d)$ “**mild distributions**” (because they behave better than general tempered distributions), see [?] or [?].



P2: Mild versus Tempered Distributions I

Although it is clear that the Schwartz space $\mathcal{S}(\mathbb{R}^d)$ and its (topological) dual, the space $\mathcal{S}'(\mathbb{R}^d)$ of *tempered distributions* not only Fourier invariant, but also closed under differentiation, and thus are the ideal tool for finding solutions of PDEs, it is quite cumbersome to work with countable families of (semi-)norms instead of just one simple norm as in the case of $(\mathcal{S}_0(G), \|\cdot\|_{\mathcal{S}_0})$ or $(\mathcal{S}'_0(G), \|\cdot\|_{\mathcal{S}'_0})$. However, for general LCA groups the so-called *Schwartz-Bruhat space* is really cumbersome and may not have much to do with differentiation (see [?]), which has been used in the famous Acta paper by Andre Weil ([?]) about the metaplectic group (including Fractional Fourier transforms as a compact subgroup of unitary operators).

It was then H. Reiter who picked up this property in his revision of Weil's work in the Lecture notes [?].



P3: Extending the FT to all $(L^p(\mathbb{R}^d), \|\cdot\|_p)$!

We have the following sandwich situation for $1 \leq p \leq \infty$:

$$(\mathbf{S}_0(G), \|\cdot\|_{\mathbf{S}_0}) \hookrightarrow (L^p(G), \|\cdot\|_p) \hookrightarrow (\mathbf{S}'_0(G), \|\cdot\|_{\mathbf{S}'_0}), \quad ,$$

and thus it is no problem to define the FT for any such function. This is in contrast to the limitations arising with the classical L^p -spaces, where only the *Hausdorff-Young* estimate is valid (with the usual convention $1/q + 1/p = 1$):

$$\mathcal{F}(L^p(G)) \hookrightarrow L^q(\widehat{G}), \quad 1 \leq p \leq 2.$$

But, for example, with the observation that $\mathcal{F}(L^p(\mathbb{R}^d)) \hookrightarrow \mathbf{S}'_0(\mathbb{R}^d)$ we know much more than just the membership of $\widehat{f} \in \mathbf{S}'(\mathbb{R}^d)$. Of course, for $p > 2$ $\mathcal{F}L^p(\mathbb{R}^d)$ will contain true distributions which are not represented by locally integrable functions!



P3: Extending the FT to all $(L^p(\mathbb{R}^d), \|\cdot\|_p)$

Among the spaces discussed above we have $p = \infty$ as an important special case. Obviously we have a continuous and dense embedding of $(\mathcal{S}_0(G), \|\cdot\|_{\mathcal{S}_0})$ into $(L^1(G), \|\cdot\|_1)$ (it is a true **Segal algebra in Reiter's sense**), and thus by duality we have

$$(L^\infty(\mathbb{R}^d), \|\cdot\|_\infty) \hookrightarrow (\mathcal{S}'_0(\mathbb{R}^d), \|\cdot\|_{\mathcal{S}'_0}).$$

Hence any $h \in L^\infty(\mathbb{R}^d)$ has a Fourier transform in $\mathcal{S}'_0(\mathbb{R}^d)$, and any such mild distribution has a support, defined in the usual way, as one is used from the classical theory of distributions.

The idea to call $\text{supp}(\widehat{h})$ the “**spectrum of**” h (we write $\text{spec}(h)$) can be justified by demonstrating that this natural viewpoint is compatible with Reiter's definition of the spectrum, using closed ideals of $(L^1(\mathbb{R}^d), \|\cdot\|_1)$ (cf. Reiter's Ideal Theorem for Segal algebras, see [?] and [?]).



P4: Translation-bounded measures

Of course there are mild distributions which are not regular, i.e. not represented by (locally) integrable functions, but true, potentially unbounded measures.

Obvious examples are Dirac measures $\delta_x : f \mapsto f(x)$, but also Dirac combs (more or less Haar measures on lattices $\Lambda \triangleleft \mathbb{R}^d$):

$$\sqcup \sqcup \Lambda := \sum_{\lambda \in \Lambda} \delta_\lambda.$$

Poisson's formula (valid for $f \in \mathcal{S}_0(\mathbb{R}^d)$) implies

$$\mathcal{F}(\sqcup \sqcup) = \sqcup \sqcup, \quad \text{and} \quad \mathcal{F}(\sqcup \sqcup \Lambda) = C_\Lambda \sqcup \sqcup \Lambda^\perp,$$

where $\Lambda^\perp$ is the *orthogonal group* to Λ in $\widehat{\mathbb{R}^d} = \mathbb{R}^d$.

I am currently engaged in a project with colleagues active in *quasi-crystals* who find this form of FT quite convenient.



P5: Sampling Theory, Shannon's Theorem I

One of the basic principles of digital signal processing is the fact that a band-limited signal can be recovered well from discrete (regular) samples if only the *Nyquist condition* is satisfied. At the basis (also for the finite discrete setting) is Poisson's formula, giving us

$$\mathcal{F}(\sqcup) = \sqcup.$$

By applying a dilation one obtains a similar statement for general lattices of the form $\sqcup_a := \sqcup_{a\mathbb{Z}^d}$, namely (up to normalization)

$$\mathcal{F}\sqcup_a = C_a \sqcup_b, \quad b = 1/a.$$

and obviously one has (by the extended convolution theorem):

$$\mathcal{F}(a \cdot f) = C_a(\widehat{f} * \sqcup_b), \quad f \in \mathbf{S}_0(\mathbb{R}^d).$$



P5: Sampling Theory, Shannon's Theorem II

Since a bandlimited function $f \in L^1(\mathbb{R}^d)$ belongs to $\mathbf{S}_0(\mathbb{R}^d)$ we can state that the information in the sampled signal $\sum_{k \in \mathbb{Z}^d} f(ak)\delta_{ak}$ is the same as in the b -periodic version of $\hat{f}$. So if the periodized spectrum is well separated it allows to recover $\hat{f}$ by a simple convolution on the time side (resulting in Shannon's reconstruction formula).



P6: Translation invariant Systems and Convolution

In the work on multipliers between L^p -spaces Gaudry has introduced the space of *quasi-measures*, which (according to Cowling, see [?]) can be identified with $\mathcal{FL}_{loc}^\infty$.

As one can find in the book of Larsen about Multipliers (see [?]) any operator between (reflexive) L^p -spaces can be represented as a convolution operator with a (uniquely determined) quasi-measure, but since there are no global restrictions in this space one cannot extend the Fourier transform to all the quasi-measures. Thus it is not clear, whether such a “multiplier”, understood as Fourier multiplier, can be described by pointwise multiplication with some “transfer function” on the Fourier side. WE HAVE:

Multipliers T from $(\mathcal{S}_0(G), \|\cdot\|_{\mathcal{S}_0})$ to $(\mathcal{S}'_0(G), \|\cdot\|_{\mathcal{S}'_0})$ are exactly convolution operators by unique elements $\sigma \in \mathcal{S}'_0(G)$, (*impulse response*) or as a pointwise multiplier on the Fourier side by $\hat{\sigma} \in \mathcal{S}'_0(\mathbb{R}^d)$ (which takes the role of a *transfer function*).



P7: All the classical cases are just Special Cases I

It is clear that periodization provides a *periodic* function, and pointwise multiplication of $h \in \mathbf{C}_b(\mathbb{R}^d)$ with a Dirac comb $\sqcup \sqcup \Lambda$ defines a weighted Dirac comb (with bounded coefficients), which belong to $\mathbf{S}'_0(\mathbb{R}^d)$ and satisfy $\text{supp}(h \cdot \sqcup \sqcup \Lambda) \subseteq \Lambda$.

For the case of tempered distributions this **does not mean that we have a distribution which is already defined on the subgroup!**

Similar von continuous subgroups like the x -axis in $\mathbb{R}^2$. But - this is meant by claiming that $\mathbf{S}'_0(\mathbb{R}^d)$ is not too big - we have

- Any $\sigma \in \mathbf{S}'_0$ with $\text{supp}(\sigma) \subseteq H \triangleleft \mathbb{R}^d$ arises from a unique $\tau \in \mathbf{S}'_0(H)$ via $\sigma(f) = \tau(f|_H)$, $f \in \mathbf{S}_0(\mathbb{R}^d)$.
- The above situation is equivalent to the claim that $\hat{\sigma}$ is a $H^\perp$ periodic distribution in $\mathbf{S}'_0(\mathbb{R}^d)$.

These facts ARE IN CONTRAST with tempered distributions!



There is just one Fourier Transform (for each dimension) I

The reservoir $\mathcal{S}'_0(\mathbb{R}^d)$ contains all the typical cases, where usually for each type a specialized Fourier transform seems to be in place:

- 1 First the “continuous non-periodic” case:
 $(L^p(\mathbb{R}^d), \|\cdot\|_p) \hookrightarrow (\mathcal{S}'_0(\mathbb{R}^d), \|\cdot\|_{\mathcal{S}'_0})$ for $1 \leq p \leq 2$, hence the extended Fourier transform contains the **Plancherel transform** and the Hausdorff-Young setting.
- 2 **Periodic functions** (with whatever period) in L^p (locally) belong to $(\mathcal{S}'_0(\mathbb{R}^d), \|\cdot\|_{\mathcal{S}'_0})$ and the corresponding FT is concentrated on the dual lattice (classical **Fourier Series**).
- 3 **Periodic and discrete** signals belong to $\mathcal{S}'_0(\mathbb{R}^d)$ and their FT can be obtained by the usual **DFT/FFT** algorithm.
- 4 Similar claims apply for **discrete signals** (DTFT) or *almost periodic functions*.



There is just one Fourier Transform (for each dimension) II

More importantly, we can (easily) show that the **heuristic transitions** performed usually correspond to “**mild convergence**” (i.e. w^* -convergence in $\mathcal{S}'_0(\mathbb{R}^d)$), such as:

- ① lattice constants which are convergent, including to infinity, i.e. to the non-periodic case (FT for $L^1(\mathbb{R}^d)$);
- ② approximation of discrete FTs (via Dirac sequences) from the continuous case;
- ③ approximating the discrete and periodic case from non-periodic discrete version by periodization
- ④ approximation of continuous densities of bounded measures by discrete measures (using a simple discretization procedure), whose inverse is smoothing using BUPUs (such as piecewise linear interpolation), and so on.



Q1: The Kernel Theorem I

Clearly a linear mapping T from $\mathbb{C}^n$ to $\mathbb{C}^m$ have a matrix representation: $T(\mathbf{x}) = \mathbf{A} * \mathbf{x}$, where the entries are of the form

$$a_{j,k} = \langle T(\mathbf{e}_k), \mathbf{e}_j \rangle, 1 \leq k \leq m, 1 \leq j \leq n.$$

Hence one can expect that the continuous version allows to write at least (certain integral) operators as

$$T(f) = \int_{\mathbb{R}^d} K(x, y) f(y) dy, \quad f \in L^2(\mathbb{R}^d).$$

It turns out, that for $K \in \mathbf{S}_0(\mathbb{R}^{2d})$ these operators map $\mathbf{S}'_0(\mathbb{R}^d)$ into $\mathbf{S}_0(\mathbb{R}^d)$ in a w^* -to-norm continuous fashion and *vice versa*. Moreover in analogy to the discrete case one has

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



Q1: The Kernel Theorem II

Extending to the Hilbert space setting one finds that kernels in $L^2(\mathbb{R}^{2d})$ give rise to the well-known Hilbert Schmidt operators. In fact this is a unitary mapping, using the fact

$$\|K\|_{L^2} = \|T\|_{\mathcal{HS}} := \text{trace}(T \circ T^*).$$

The outer layer describes the most general operator. The correspondence identifies $\mathbf{S}'_0(\mathbb{R}^{2d})$ with the space of all bounded linear operators from $(\mathbf{S}_0(\mathbb{R}^d), \|\cdot\|_{\mathbf{S}_0})$ to $(\mathbf{S}'_0(\mathbb{R}^d), \|\cdot\|_{\mathbf{S}'_0})$. In this setting one can even describe multiplication or convolution operators, in particular the identity operator, which corresponds to the distribution $F \mapsto \int_{\mathbb{R}^d} F(x, x) dx$, $F \in \mathbf{S}_0(\mathbb{R}^{2d})$, which is well defined since the restriction of $F \in \mathbf{S}_0(\mathbb{R}^{2d})$ to the diagonal is in $\mathbf{S}_0(\mathbb{R}^d)$ and hence integrable.



Q1: The Kernel Theorem III

If one tries to rewrite the functional (representing the identity operator) in the usual way (or observing that of course the identity operator is an operator which commutes with translation, and thus has to be a convolution operator, with the usual Dirac measure $\delta_0 : f \mapsto f(0)$) we have

$$f(t) = \int_{\mathbb{R}^{2d}} \int_{\mathbb{R}^d} K(x, y) f(y) dy, \quad f \in \mathbf{S}_0(\mathbb{R}^d),$$

which is only possible if one has in each row

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

(this is more or less the transition from the *Kronecker delta* describing the unit-matrix to the Dirac delta, and is another way of expressing the “sifting property” of δ_0 .)



Q1: The Kernel Theorem IV

The composition law for matrices is the unique way of combining information about two linear mappings which can be composed (Domino rule) into a new matrix scheme, via standard matrix multiplication rules: $\mathbf{C} = \mathbf{A} * \mathbf{B}$. Thus one expects for the composition of operators a similar composition law for their kernels, something like

$$K(x, y) = \int_{\mathbb{R}^d} K_1(x, z) K_2(z, y) dz, \quad x, y \in \mathbb{R}^d.$$

If one make use of the kernel for the Fourier transform, i.e. $K_2(z, y) = \exp(-2\pi i \langle y, z \rangle)$ and $K_1(x, z) = \exp(2\pi i \langle x, z \rangle)$, then, even if the integrals do not make sense anymore in the Lebesgue sense, it still suggest to claim that the resulting product operator is the identity operator, which gives a meaning to formulas appearing in engineering books on the Fourier transform.



Q2: Spreading, Weyl, KNS-Calculus I

It is easy to explain the concept of the *spreading function* corresponding to a given $N \times N$ matrix $\mathbf{A}$, by decomposing it into (N cyclic) side-diagonals and then applying a Fourier transform on them. The KNS symbol $\kappa(\mathbf{A})$ (Kohn-Nirenberg) then arises as the (!*symplectic*) FT of the spreading function. It satisfies the following important *covariance property*:

$$\kappa(\pi(\lambda) \circ \mathbf{A} \circ \pi(\lambda)^*) = T_\lambda \kappa(\mathbf{A}).$$

This makes these descriptions suitable tools for the description of e.g. *Anti-Wick operators* (STFT-multipliers).



Q3: Mild distributions as upper symbols I

There are many good reasons to restrict the windows for the STFT to the space $\mathbf{S}_0(\mathbb{R}^d)$. Then one can assure that any $\sigma \in \mathbf{S}'_0$ has a bounded and continuous STFT, denoted by $V_g \sigma$.

Anti-Wick operators are based on the consideration that (for a normalized window $g \in \mathbf{S}_0(\mathbb{R}^d)$ with $\|g\|_2 = 1$) the inverse for the *isometric* embedding $f \mapsto V_g f$ (from $(\mathbf{L}^2(\mathbb{R}^d), \|\cdot\|_2)$ into $(\mathbf{L}^2(\mathbb{R}^{2d}), \|\cdot\|_2)$) is just the adjoint mapping V_g^* , of the form

$$V_g^*(F) = \int_{\mathbb{R}^d \times \widehat{\mathbb{R}}^d} F(\lambda) \pi(\lambda) g \, d\lambda,$$

for $F \in (\mathbf{L}^2(\mathbb{R}^d), \|\cdot\|_2)$. Thus we have the reconstruction formula

$$f = V_g^*(V_g f) = \int_{\mathbb{R}^d \times \widehat{\mathbb{R}}^d} V_g f(\lambda) \pi(\lambda) g \, d\lambda.$$



Q3: Mild distributions as upper symbols II

For $g \in \mathbf{S}_0(\mathbb{R}^d)$ this integral can even be approximated by corresponding Riemannian sums (according to F. Weisz). Thus it is natural to define for bounded functions $m \in \mathbf{C}_b(\mathbb{R}^d \times \widehat{\mathbb{R}}^d)$ an (Anti-Wick) operator of the form

$$\mathbf{A}_m(f) = V_g^*(V_g f) = \int_{\mathbb{R}^d \times \widehat{\mathbb{R}}^d} m(\lambda) V_g f(\lambda) \pi(\lambda) g \, d\lambda.$$

This obviously defines a bounded linear operator on $(\mathbf{L}^2(\mathbb{R}^d), \|\cdot\|_2)$, with

$$\|\mathbf{A}_m\|_{\mathbf{L}^2(\mathbb{R}^d)} \leq C \|m\|_\infty,$$

but it defines in fact a BGT-homomorphism, thus acting boundedly also on $(\mathbf{S}_0(\mathbb{R}^d), \|\cdot\|_{\mathbf{S}_0})$ or $(\mathbf{S}'_0(\mathbb{R}^d), \|\cdot\|_{\mathbf{S}'_0})$.



Q3: The “natural choice of upper symbols I

As it turns out, the correlation between neighboring elements from the family $\pi(\lambda)g, \lambda \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d$ introduces some smearing of the (upper) symbol. One can show that L^2 -symbols give Hilbert Schmidt operators, but the converse is not true, it is enough that a smeared version of m belongs to $L^2(\mathbb{R}^d \times \widehat{\mathbb{R}}^d)$.

With the idea that a “smeared version” of the upper symbol is supposed to be in $C_b(\mathbb{R}^{2d})$ we may (correctly) expect that ANY SIGNAL, i.e. any $m \in \mathcal{S}'_0(\mathbb{R}^{2d})$ defines a BGTr homomorphism, with control of the norms through $\|m\|_{\mathcal{S}'_0}$ (at all three levels!).

The key for a verification is the following observation:

The restriction of $f \mapsto V_g f$, which acts isometrically on $(L^2(\mathbb{R}^d), \|\cdot\|_2)$, is also bounded from $(\mathcal{S}_0(\mathbb{R}^d), \|\cdot\|_{\mathcal{S}_0})$, and not just into $(L^1(\mathbb{R}^d), \|\cdot\|_1)$. In fact, one has $V_g f \in \mathcal{S}_0(\mathbb{R}^{2d})$ if and only if f and g belong to $\mathcal{S}_0(\mathbb{R}^d)$.



Q4: Generalized Stochastic Processes I

As I learned from many discussions with engineers one should sometimes take a probabilistic viewpoint. A signal is not “a true signal” but just a *stochastic realization* of a process. So to say, even “measuring the same thing several times” does not mean that one gets exactly the same values.

Thus it is natural to ask what a “stochastic signal” could be. Recall that a (continuous) function assigns to each point t a value $f(t)$. A *generalized function* σ assigns to each test function k (average of point masses) some value $\sigma(k)$. A *stochastic process* $\rho(t)$ assigns to each value t a probability measure, ideally a member of some abstract Hilbert space $\mathcal{H}$ (defined in a probabilistic manner). So a *generalized stochastic signal* (**GSP**) is a bounded linear operator $\rho : (\mathbf{S}_0(\mathbb{R}^d), \|\cdot\|_{\mathbf{S}_0}) \rightarrow (\mathcal{H}, \|\cdot\|_{\mathcal{H}})$ (see work with W.Hörmann, 1989). [?].



Q4: Generalized Stochastic Processes II

For each GSP there is a Fourier transform, the associated *spectral process* and a *spectral representation* (inverse Fourier transform), by the usual rule

$$\hat{\rho}(f) = \rho(\hat{f}), \quad f \in \mathbf{S}_0(\mathbb{R}^d).$$

For every GSP we have a *autocorrelation* $\sigma_\rho \in \mathbf{S}'_0(\mathbb{R}^{2d})$ with

$$\sigma_\rho(f \otimes g) = \langle f, g \rangle_{\mathcal{H}}, \quad f, g \in \mathbf{S}_0(\mathbb{R}^d).$$

As expected the autocorrelation of $\hat{\rho}$ is just $\mathcal{F}_{\mathbb{R}^{2d}}(\sigma_\rho)$.

Wide sense stationarity means:

$$\langle \rho(x), \rho(y) \rangle_{\mathcal{H}} = \langle \rho(x+h), \rho(y+h) \rangle_{\mathcal{H}}, \quad x, y, h \in \mathbb{R}^d.$$

Or in the case of generalized stochastic processes

$$\langle \rho(f), \rho(g) \rangle_{\mathcal{H}} = \langle \rho(T_h f), \rho(T_h(g)) \rangle_{\mathcal{H}}, \quad h \in \mathbb{R}^d, f, g \in \mathbf{S}_0.$$



Where do MILD DISTRIBUTIONS appear naturally

We list topics according to relevance for engineering students!

- 1 Signals ARE mild distributions (talk at LMU, 7.6.2024)
- 2 Mild distributions describe *impulse response* and *transfer function*. for TILS (*translation invariant linear systems*).
- 3 This approach allows a natural approach to convolution and the Fourier transform (convolution theorem), see [?].
- 4 The interpretation of the FT as *unitary BGTr automorphism* unifies different FFT variants and reveals connections.
- 5 This gives easy approach to *Shannon Theorem*.
- 6 The *kernel theorem* (also KNS symbols, spreading function) opens a natural approach to “continuous matrix theory”.
- 7 Such considerations help in *Quantum Harmonic Analysis*.





Thank you for your attention!

All relevant papers are downloadable from my homepage

https://nuhagphp.univie.ac.at/home/feipub_db.php

Course notes (including YouTube Links) are found at

www.nuhag.eu/ETH20

Many other talks on related subjects can be downloaded from

<https://nuhagphp.univie.ac.at/home/feitalks.php>

User: visitor PWD: nuhagtalks



