Mild Distributions: A Simple and Universal Tool for Fourier and Time-Frequency Analysis

Hans G. Feichtinger,
Faculty of Mathematics, University Vienna& ARI (OEAW)
hans.feichtinger@univie.ac.at
www.nuhag.eu

University of Linz, AT, Sept. 1st, 2025

.

OEMG Workshop M06: Frames Multipliers

The Purpose of Talks and Publications

Typical (short) presentations describe a problem, motivate its importance and then go on to report on previous results and methods and then formulate the new results.
Thus at the end the audience has more detailed information about a specific problem.
A survey talk can try to establish connections between two concrete topics or point out surprising facts.
The overall idea of "Conceptual Harmonic Analysis" spans from concrete tasks in Gabor analysis to a wide range of problems treated in the context of coorbit theory, but also using methods related to decomposition spaces.

Mild Distributions, Modulation Spaces I

The by now *classical coorbit spaces* deal with general families of Banach spaces derived from (integrable) group representations, with wavelet analysis and Gabor theory (via the Heisenberg group) as special cases. One uses weighted (mixed norm) spaces to obtain atomic decompositions and Banach frames. Sometimes one can even get ONBs, but often (!Balian-Low) this is not possible.

The focus in this presentation is in a way on a very special situation, just three spaces, the Banach Gelfand Triple $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$, avoiding weights, quasi-Banach spaces, Frechet spaces etc., while still providing powerful tools for applied problems. Thus I would rather widen the scope of applications of this "simple and universal tool".

Let us recall that these three spaces are all Fourier invariant members of the family of *modulation spaces*.

Mild Distributions, Modulation Spaces II

Within the wider context mentioned above we are thus reducing our attention to a very limited setup:

1. We work in the context of time-frequency and Gabor analysis
2. From the coorbit point-of-view we have the Schrödinger representation of the (reduced) Heisenberg group;
3. From the viewpoint of decomposition spaces we deal with uniform coverings (thus BUPUs)
4. we avoid weight, and thus $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ and its dual are the minimal/maximal space in the given family
5. we focus on $\mathbb{R}^d$, although $\mathcal{G}$ LCA would be possibe.

Portfolio of Relevant Applications I

In the last years a wide range of settings has been evaluated, where the setting of the Banach Gelfand Triple is useful (or unavoidable). We start with the background in Gabor Analysis:

1. Atoms in $\boldsymbol{S}_0(\mathbb{R}^d)$ guarantee Bessel conditions for Gabor families and the existence of Gabor frames for sufficiently dense sets $\Lambda \subset \mathbb{R}^d \times \widehat{\mathbb{R}}^d$;
2. Dual Gabor atoms are also in $\boldsymbol{S}_0(\mathbb{R}^d)$, with continuous dependence of the dual Gabor atom on the lattice used;
3. For atoms $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ the reconstruction formula of f from $V_g f$ is valid in the sense of Riemann integrals ($f \in \boldsymbol{L}^2(\mathbb{R}^d)$);
4. *Fundamental Identity for Gabor Analysis* requires $\boldsymbol{S}_0(\mathbb{R}^d)$;
5. the study of Gabor multipliers and KNS-calculus;
6. new approach to Quantum Harmonic Analysis.

Portfolio of Relevant Applications II

But there are many more classical topics:

1. Classical Fourier Transform: summability kernels in $\boldsymbol{S}_0(\mathbb{R}^d)$;
2. Poisson's Formula is valid for $\boldsymbol{S}_0(\mathbb{R}^d)$;
3. The extended Fourier transform for $\boldsymbol{S}_0'(\mathbb{R}^d)$ (mild distributions) extends the theory of Argabright-Lamadrid, and is useful for recent studies of quasi-crystals;
4. Fractional Fourier transform, or more generally metaplectic operators leave $\boldsymbol{S}_0(\mathbb{R}^d)$ invariant (LCTs, quadratic Fourier transforms etc.), with consequences for PDE (e.g. Schrödinger equation) and pseudo-diff-ops;
5. Variants of the Wigner distribution, Weyl-calculus;
6. Generalized stochastic processes over LCA groups

Abstract: Mild Distributions I

The concept of mild distributions provides a flexible and universal framework that unifies classical distribution theory with modern techniques from Fourier and time-frequency analysis. From an applied perspective the Banach space of mild distributions is the set of all signals, whether they are discrete or continuous, period or non-periodic. Even an irregularly sampled signal can be viewed as a mild distribution, and the usual identification of periodic and discrete signals as finite sequences is best understood in this way.

From a mathematical point of view the space of mild distributions is the dual of a Banach algebra of test functions inside the Fourier algebra, meanwhile known as Feichtinger algebra $S_0(G)$. The rich collection of properties of this functor (assigning a Banach algebra of functions for each locally compact Abelian group) implies corresponding versatility of the space of $\boldsymbol{S_0'}(\mathbb{R}^d)$.

Abstract: Mild Distributions II

It can be also identified with the subspace of all tempered distributions having a bounded spectrogram (short time Fourier transform or STFT).
Thus every mild distribution has a (Fractional) Fourier transform.
Dirac combs are mild distributions and Poisson's formula shows that their Fourier transforms are Dirac combs over the orthogonal subgroup.
The presentation will give a short summary and display a small selection of the *many application areas* (in classical Fourier analysis or modern time-frequency analysis) of $\boldsymbol{S}_0'(\mathbb{R}^d)$. It will also highlight the role of weak*-convergence which corresponds to uniform convergence of the STFT (Short-Time Fourier Transform) over compact subsets of phase space.

Motivation I

At the beginning of the last century it was realized that one can build "auto-mobiles", i.e. vehicles which do not need horse power or steam engine in order to move autonomously. The construction principles for cars was developed and raised to high standards, streets and highways have been built, and individual mobility was supported: Every family has typically two cars in the garage, streets are crowded, and we all spend a lot of time "on the road" due to the fact the public transport and infra-structure for the charging of electric cars is still unsatisfactory.

It does not help too much to have nowadays the most efficient Diesel engines, powerful SUVs populating crowded streets in our towns, and we have to rethink mobility. New problems come up (last mile, public transport, etc.), but

new concepts help to make progress.

Motivation II

At the beginning of the last century the Lebesgue integral has been fully developed. Fourier Analysis, as one can learn it nowadays from mathematical books is based on the technology of Lebesgue integration, which allows to define the Lebesgue spaces such as $\left(\boldsymbol{L}^p(\mathbb{R}^d), \|\cdot\|_p\right)$. Integration theory is crucial for the definition of the Fourier transform and for the definition of the convolution. They are combined in of the convolution theorem:
The FT converts convolution into pointwise multiplication.
The concept of Hilbert spaces allowed to understand the classical Fourier series expansions as a simple orthonormal expansion in a separable Hilbert space. More involved is the description of the Fourier transform on the Hilbert space $\boldsymbol{L}^2(\mathbb{R}^d)$. It can still be described as a unitary automorphism of $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$, although the "building blocks of the FT", i.e. the pure frequencies $\chi_s(t) := exp(2\pi ist), \; s, t \in \mathbb{R}^d$ do not belong to $\mathcal{H}$.

Motivation III

When it comes to the study of operators (e.g. compact operators) it is also natural to first restrict the attention to operators which are bounded on $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$ and only then study more specialized classes of (often compact) operators, such as Hilbert-Schmidt operators or Schatten S_p-class operators (with singular values in ℓ^p).

Often Hilbert spaces allow to pursue nicely the analogy between the Euclidean situation, where orthogonal projection realize the best approximation of a given element from a closed subspaces. The sampling operator, in its most simple version the mapping

$$f \mapsto \sqcup\!\sqcup \cdot f := \sum_{k \in \mathbb{Z}^d} f(k)\delta_k,$$

is not defined on any of the spaces $\boldsymbol{L}^p(\mathbb{R}^d)$ nor does it map into such a space, it only maps into $\boldsymbol{S}_0'(\mathbb{R}^d)$ for bounded f.

Motivation IV

Another example of such an analogy is the

kernel theorem for Hilbert-Schmidt operators.

While any linear mapping on $\mathbb{R}^n$ can be described by matrix multiplication on column vectors with a unique $n \times n$-matrix $\boldsymbol{A}$, we have a continuous analogue:

HS-operators are exactly those (compact) linear mappings on $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$ which can be described as integral operators of the form

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

with $K \in \boldsymbol{L}^2(\mathbb{R}^{2d})$ (taking the role of continuous matrices). Also

$$\|T\|_{\mathcal{HS}} = \sqrt{\mathrm{trace}(TT^*)} = \|K\|_{\boldsymbol{L}^2(\mathbb{R}^{2d})}.$$

Unfortunately such operators are never invertible.

Motivation V

Despite the fact that engineers learn about the fact that one should use the Dirac Delta (?function?) for the description of time-invariant channels (linear systems, operators), or have to learn the classification of signals into continuous and discrete, periodic and non-periodic (and then decaying, say square integrable), one must say that the mathematical theory describing "signals" and thus the operations on them (e.g. signal processing tasks) is far from satisfactory from an applied viewpoint. What about even slightly time-varying signals, like heart-beat, or a piece of music (compressed using the MP3 algorithm), the compression of images or other practical tasks. Which of the mathematically well established function spaces are useful?

Are signals really elements of $\boldsymbol{L}^2(\mathbb{R}^d)$ (or finite energy signals)?

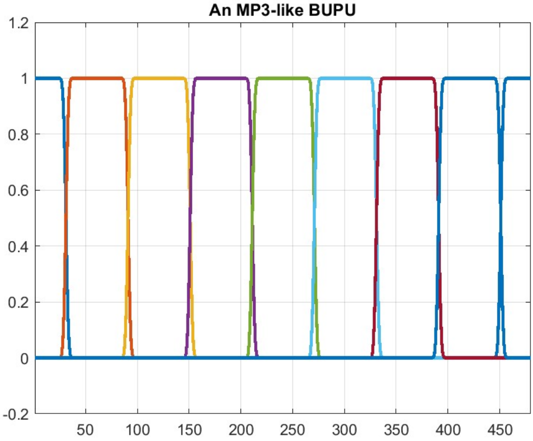


Figure: mp3bupu.jpg

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 *MP*3 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).

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 $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$.
Going to continuous variable the same principle applies, but now the Fourier coefficients are indexed by $\ell^1(\mathbb{Z})$. If we assume that a function should have locally an absolutely convergent Fourier series, i.e. belongs to Wiener's Algebra $\left(\boldsymbol{A}(\mathbb{T}), \|\cdot\|_{\boldsymbol{A}}\right)$, then one takes the ℓ^1-norm of the coefficients at each position.

Finally $(\mathbf{S}_0(\mathbb{R}^d), \|\cdot\|_{\mathbf{S}_0})$ can be described by taking the sum of all these $\mathbf{A}(\mathbb{T})$-norms over all the positions. In the context of the so-called *Wiener amalgam spaces* (as discussed by the author since 1980) we have $\mathbf{S}_0(\mathbb{R}^d) = \mathbf{W}(\mathcal{F}\mathbf{L}^1, \ell^1)(\mathbb{R}^d)$.

The dual space thus is $\mathbf{W}(\mathcal{F}\mathbf{L}^\infty, \ell^\infty)(\mathbb{R}^d)$, the space of *translation-bounded pseudo-measures*, which simply means that one takes the sup-norm over all the frequencies and positions. It is convenient to call the elements of the dual space mild distributions.

Using smooth BUPUs one can characterize this space as the space of tempered distributions having (local Fourier) coefficients in $\ell^\infty(\mathbb{Z}^{2d})$. Different BUPUs define the same space.

The most important property is: $\mathcal{F}(\mathbf{S}_0(\mathbb{R}^d)) = \mathbf{S}_0(\mathbb{R}^d)$ and by duality that the *extended FT* is an automorphism of $\mathbf{S}_0'(\mathbb{R}^d)$, also respecting w^*-convergence (which corresponds to coordinatewise convergence in this case). $\mathbf{S}_0(\mathbb{R}^d)$ is w^*-dense in $\mathbf{S}_0'(\mathbb{R}^d)$.

What are Numbers

Let us take as a *motivating example* for a development of concepts for signal models the number systems: If we want to do the standard computations the following FIELDS are of relevance. Identifying objects in the (smaller,) more simpler variant with specific elements in the larger environment we have a chain

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

Here we note that rational number allow typically perfect computations. The (multiplicative) inverse of $3/4$ is easily recognized to be $4/3$, since $\frac{3 \cdot 4}{4 \cdot 3} = 1$.
It looks(!) similar with expressions of the form $\pi^2/3$ and $3/\pi$, but there the explanation of the terms is more complicated.
Finally we identify real number $r \in \mathbb{R}$ with complex numbers of the form $z = a + i * b = r + i * 0$.

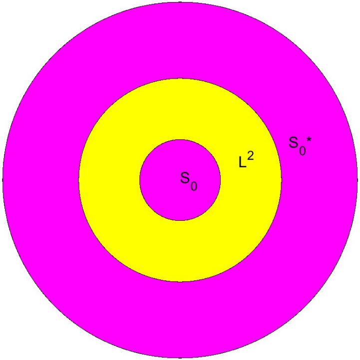
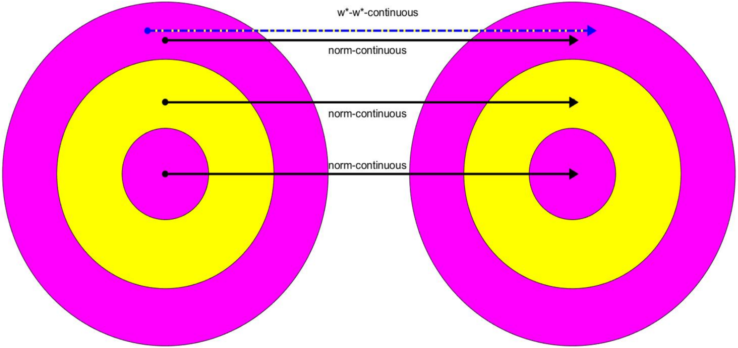


Figure: SOBGT00.jpg

Banach-Gelfand-Tripel-Homomorphisms



The Analogies to the Number system

The inner circle represents the "decent functions". $\boldsymbol{S}_0(\mathbb{R}^d)$ is a Banach space of absolutely Riemann integrable functions. Since it is Fourier invariant the Fourier Inversion Theorem is valid in a pointwise sense. Moreover, since the Dirac comb $\sqcup\!\sqcup = \sum_{k\in\mathbb{Z}^d}\delta_k$ belongs to the dual space Poisson's form is equivalent to the statement $\mathcal{F}(\sqcup\!\sqcup) = \sqcup\!\sqcup$. All the classical summability kernels are in $\boldsymbol{S}_0(\mathbb{R}^d)$, and $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)\cdot\boldsymbol{S}_0(\mathbb{R}^d)\subseteq\boldsymbol{S}_0(\mathbb{R}^d)\subset\boldsymbol{L}^1(\mathbb{R}^d)$.

On the other hand $\boldsymbol{S}_0(\mathbb{R}^d)$ is dense in $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$, and thus Plancherel's Theorem can be obtained by approximation, just using the isometric property of the FT on $\boldsymbol{S}_0(\mathbb{R}^d)$.

Finally, the rule $\widehat{\sigma}(f) = \sigma(\widehat{f})$, $f\in\boldsymbol{S}_0(\mathbb{R}^d)$, allows to extend the FT to all of $\boldsymbol{S}_0'(\mathbb{R}^d)$. It is uniquely determined by the fact that $\mathcal{F}(\chi_s) = \delta_x$, like a change of basis (as realized by the DFT)!

Shortcoming of Current Approaches

Although it appears natural to start from $(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1)$ (Lebesgue space) if one wants to define the Fourier transform or convolution (via integrals) there are limitations, in the sense that

- For $p > 2$ the Fourier transform my not be defined
- The spectrum of a function $f \in \boldsymbol{L}^\infty(\mathbb{R}^d)$ has to be defined very indirectly
- The "existence" of a FT or a convolution product may become problematic
- a theory of transformable (potentially unbounded) measures is assymmetric (Argabright-de Lamadrid).
- individual consideration for distributions may cause unexpected artefacts (non-associativity, and so on).
- convergence for tempered distributions is a rather limited, even $n^2\delta_n$ tends to zero distributionally!

Banach Gelfand Triples appear to be the correct structure in order to imitate situations like those encountered by the inclusion of the number systems $\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$.

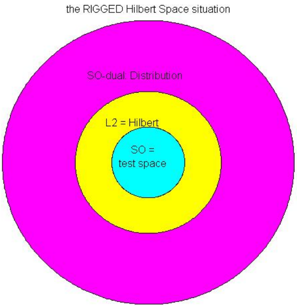


Figure: Three layers

The "inner layer" is where the actual computations are done, the focus in mathematical analysis is all to often with the (yellow) Hilbert spaces (taking the role of $\mathbb{R}$, more complete with respect to a scalar product, more symmetric, because it allows to be identify the dual, via the Riesz representation Theorem, very much like matrix theory is working, with row and column vectors), and the outside world where things sometimes can be explained, and with completeness in an even more general sense (distributional convergence). In other words, we do not assume anymore that $\sigma_n(f)$ is convergent for all $f \in \mathcal{H}$ (the completion of the test functions in $\mathcal{H}$), but *only for* elements f in the core space!
What we are going to suggest/present is THE Banach Gelfand Triple

$$(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$$

consisting of the *Feichtinger algebra* $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$, the Hilbert space $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$ and the dual space $\left(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'}\right)$, known as space of *mild distributions*. Note that these spaces can be defined without great difficulties on any LCA group $\mathcal{G}$ and that it satisfies many desirable *functorial properties*, see the early work of V. Losert (lo83-1).

For $\mathbb{R}^d$ the most elegant way (which is describe in gr01 or ja18) is to define it by the integrability (actually in the sense of an infinite Riemann integral over $\mathbb{R}^{2d}$ if you want) of the STFT

$$V_{g_0}(f)(x,y) := \int_{\mathbb{R}^d} f(y)g(y-x)e^{-2\pi isy}\,dy$$

and the corresponding norm

$$\|f\|_{\boldsymbol{S}_0} := \int_{\mathbb{R}^{2d}} |V_{g_0}(f)(x,y)|dxdy < \infty.$$

From a practical point of view one can argue that one has the following list of good properties of $\boldsymbol{S}_0(\mathbb{R}^d)$.

Theorem

1. $\boldsymbol{S}_0(\mathbb{R}^d) \hookrightarrow \left(\boldsymbol{W}(\boldsymbol{C}_0, \ell^1)(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{W}}\right) \hookrightarrow \boldsymbol{L}^1(\mathbb{R}^d) \cap \boldsymbol{C}_0(\mathbb{R}^d)$;
2. $\mathcal{F}(\boldsymbol{S}_0(\mathbb{R}^d)) = \boldsymbol{S}_0(\mathbb{R}^d)$ *(isometrically);*
3. *Isometrically invariant under TF-shifts*

$$\|\pi(\lambda)(f)\|_{\boldsymbol{S}_0} = \|M_s T_t f\|_{\boldsymbol{S}_0} = \|f\|_{\boldsymbol{S}_0}, \quad \forall (t,s) \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d.$$

4. $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ *is an essential double module (convolution and multiplication)*

$$\boldsymbol{L}^1(\mathbb{R}^d) * \boldsymbol{S}_0(\mathbb{R}^d) \subseteq \boldsymbol{S}_0(\mathbb{R}^d) \quad \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) \cdot \boldsymbol{S}_0(\mathbb{R}^d) \subseteq \boldsymbol{S}_0(\mathbb{R}^d),$$

in fact a Banach ideal and hence a double Banach algebra.

5. *Tensor product property* $\boldsymbol{S}_0(\mathbb{R}^d) \widehat{\otimes} \boldsymbol{S}_0(\mathbb{R}^d) \approx \boldsymbol{S}_0(\mathbb{R}^{2d})$ *which implies the Kernel Theorem.*
6. *Restriction property: For* $H \lhd G$: $R_H(\boldsymbol{S}_0(G)) = \boldsymbol{S}_0(H)$.

1. $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ has various equivalent descriptions, e.g.
   - as *Wiener amalgam space* $\boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1, \ell^1)(\mathbb{R}^d)$;
   - via *atomic decompositions* of the form

$$f = \sum_{i \in I} c_i \pi(\lambda_i) g \text{ with } (c_i)_{i \in I} \in \ell^1(I).$$

2. $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ is invariant under group automorphism;
3. $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ is invariant under the *metaplectic group*, and thus under the *Fractional Fourier transform* as well as the multiplication with *chirp signals*: $t \mapsto exp(-i\alpha t^2)$, for $\alpha \geq 0$.

In addition $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ is quite universally useful in Classical Fourier Analysis and of course for *Time-Frequency Analysis* and *Gabor Analysis*, and as I am going to show also for QHA: Quantum Harmonic Analysis. In short, it is easier to handle than the Schwartz-Bruhat space or even the Schwartz space $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$, and since $\boldsymbol{\mathcal{S}}(\mathbb{R}^d) \hookrightarrow \left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ it is (much) bigger.

Theorem

1. $\mathcal{F}(\boldsymbol{S}_0'(\mathbb{R}^d)) = \boldsymbol{S}_0'(\mathbb{R}^d)$ *via* $\widehat{\sigma}(f) := \sigma(\widehat{f}), f \in \boldsymbol{S}_0'$.
2. *Identification of TLIS:* $\boldsymbol{H}_{\mathcal{G}}(\boldsymbol{S}_0, \boldsymbol{S}_0') \approx \boldsymbol{S}_0'(G)$ *(as convolutions of the form)* $T(f) = \sigma * f$;
3. *Kernel Theorem:* $\mathcal{B} := \mathcal{L}(\boldsymbol{S}_0, \boldsymbol{S}_0') \approx \boldsymbol{S}_0'(\mathbb{R}^{2d})$ *Inner Kernel Theorem reads:* $\mathcal{L}(\boldsymbol{S}_0', \boldsymbol{S}_0) \approx \boldsymbol{S}_0(\mathbb{R}^{2d})$.
4. *Regularization via product-convolution or convolution-product operators:* $(\boldsymbol{S}_0' * \boldsymbol{S}_0) \cdot \boldsymbol{S}_0 \subseteq \boldsymbol{S}_0, \; (\boldsymbol{S}_0' \cdot \boldsymbol{S}_0) * \boldsymbol{S}_0 \subseteq \boldsymbol{S}_0$
5. *The finite, discrete measures or trig. pols. are* w^*-*dense.*
6. $H \lhd G \to \boldsymbol{S}_0(H) \hookrightarrow \boldsymbol{S}_0(G)$ *via* $\iota_H(\sigma)(f) = \sigma(R_H f), f \in \boldsymbol{S}_0(G)$. *Moreover the range characterizes* $\{\tau \in \boldsymbol{S}_0(G) \,|\, \mathrm{supp}(\tau) \subset H\}$.

Theorem

1. $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'}) = (\boldsymbol{M}^\infty(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}^\infty})$, *with* $V_g(\sigma)$ *and* $\|\sigma\|_{\boldsymbol{S}_0'} = \|V_g(\sigma)\|_\infty$, *hence norm convergence corresponds to uniform convergence on pahse space. Also* w^*-*convergence is uniform convergence over compact subsets of phase space.*
2. $(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}) \hookrightarrow (\boldsymbol{L}^p(\mathbb{R}^d), \|\cdot\|_p) \hookrightarrow (\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$, *with density for* $1 \leq p < \infty$, *and* w^*-*density in* $\boldsymbol{S}_0'$. *Hence, facts valid for* $\boldsymbol{S}_0$ *can be extended to* $\boldsymbol{S}_0'$ *via* w^*-*limits.*
3. *Periodic elements* $(T_h\sigma = \sigma, h \in H)$ *correspond exactly to those with* $\tau = \mathcal{F}(\sigma)$ *having* $\mathrm{supp}(\tau) \subseteq H^\perp$.
4. *The (unique) spreading representation* $T = \int_{\mathbb{R}^d \times \widehat{\mathbb{R}}^d} F(\lambda)\pi(\lambda)d\lambda$, $F \in \boldsymbol{S}_0(\mathbb{R}^d \times \widehat{\mathbb{R}}^d)$ *for* $T \in \mathcal{B}$ *extends to the isomorphism* $T \leftrightarrow \eta(T)$ $\eta : \mathcal{B} \approx \mathcal{L}(\boldsymbol{S}_0, \boldsymbol{S}_0')$, *uniquely determined by the correspondence with*
$$\eta(\pi(\lambda)) = \delta_\lambda, \lambda \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d.$$

Some conventions

Scalar product in $\mathcal{HS}$:

$$\langle T, S\rangle_{\mathcal{HS}} = \mathrm{trace}(T * S^*)$$

In feko98 the notation

$$\alpha(\lambda)(T) = [\pi \otimes \pi^*(\lambda)](T) = \pi(\lambda) \circ T \circ \pi(\lambda)^*, \quad \lambda \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d,$$

and the covariance of the KNS-symbol is decisive:

$$\sigma(\pi \otimes \pi^*(\lambda)(T)) = T_\lambda(\sigma(T)), \quad T \in \mathcal{L}(\boldsymbol{S}_0, \boldsymbol{S}_0'), \lambda \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d.$$

Kernel Theorems I

The so-called Kernel Theorem for $\boldsymbol{S}_0$-spaces allows to establish a number of further unitary BGTr-isomorphism. It involves certain types of operators, their (integral) kernels, but also their representation as *pseudo-differential* operators, via the Weyl or Kohn-Nirenberg symbol, or (important for applications in mobile communication) their spreading distribution.
These situations allow to make use of the general principles. In order to understand the transformation one can start from the core spaces, where the analogy with the finite dimensional case is valid in a very natural sense, e.g. if one has $K(x,y) \in \boldsymbol{S}_0(\mathbb{R}^{2d})$, then

$$K(x,y) = T(\delta_y)(x) = \delta_x(T(\delta_y)),$$

in analogy to the matrices

$$a_{n,k} = [T(\mathbf{e}_k)]_n = \langle \mathbf{e}_n, T(\mathbf{e}_k) \rangle_{\mathbb{C}}^m.$$

Kernel Theorems II

The Hilbert space case of the well-known characterization

Theorem

There is a unitary BGTr isomorphism between $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^{2d})$ *and* $\left(\mathcal{N}_{w*}(\boldsymbol{S}_0', \boldsymbol{S}_0), \mathcal{HS}, \mathcal{L}(\boldsymbol{S}_0, \boldsymbol{S}_0')\right)(\mathbb{R}^d)$ *which is a unitary mapping between* $\left(\boldsymbol{L}^2(\mathbb{R}^{2d}), \|\cdot\|_2\right)$ *and* $\left(\mathcal{HS}, \|\cdot\|_{\mathcal{HS}}\right)$*, with*

$$\langle T_1, T_2\rangle_{\mathcal{HS}} = \mathrm{trace}(T_1 \circ T_2^*).$$

Alternative BGTr descriptions use the *Kohn-Nirenberg* symbol or the spreading representation : $T \in \mathcal{N}_{w*}(\boldsymbol{S}_0', \boldsymbol{S}_0)$ iff

$$T = \int_{\mathbb{R}^d \times \widehat{\mathbb{R}}^d} H(\lambda)\pi(\lambda)d\lambda, \quad H \in \boldsymbol{S}_0(\mathbb{R}^{2d}).$$

Kernel Theorems III

[feko98]: H. G. Feichtinger and W. Kozek.
Quantization of TF lattice-invariant operators on elementary LCA groups.
Gabor Book 1998.

[cofelu08]: E. Cordero, H. G. Feichtinger, and F. Luef.
Banach Gelfand triples for Gabor analysis.
In *Pseudo-differential Operators*, volume 1949 of *Lect. Notes Math.*, pages 1–33. Springer, Berlin, 2008.

[feja22] H. G. Feichtinger and M. S. Jakobsen.
The inner kernel theorem for a certain Segal algebra.
Monatsh. f. Math., 198(4):675–715, 2022.

Wilson Bases

For the case $\mathcal{G} = \mathbb{R}^d$ one can derive the kernel theorem also from the description of operators mapping ℓ^1 to ℓ^∞ or vice versa (in a w^*–to-norm continuous way).
The key is the fact, that local Fourier basis, but in particular the so-called Wilson bases are suitable for modulation spaces. In our situation we can formulate the following:

Theorem

Any Wilson ONB (obtained by a smart pairwise recombination of the elements of a tight Gabor frame of redundancy 2) establishes a unitary BGTr between $(\mathbf{S}_0, \mathbf{L}^2, \mathbf{S}_0')(\mathbb{R}^d)$ *and* $(\ell^1, \ell^2, \ell^\infty)$.

H. G. Feichtinger, K. Gröchenig, and D. F. Walnut.
Wilson bases and modulation spaces.
Math. Nachr., 155:7–17, 1992.

LINEAR ALGEBRA: Gilbert Strang's FOUR SPACES

Let us now take a LINEAR ALGEBRA POINT OF VIEW!
We recall the *standard linear algebra situation*. We view a given $m \times n$ matrix $\mathbf{A}$ either as a collection of *column* or as a collection of *row vectors*, generating $Col(\mathbf{A})$ and $Row(\mathbf{A})$. We have:

$$\textbf{row-rank}(\mathbf{A}) = \textbf{column-rank}(\mathbf{A})$$

Each *homogeneous linear system of equations* can be expressed in the form of scalar products[1] we find that

$$\mathbf{Null}(\mathbf{A}) = \quad \mathbf{Rowspace}(\mathbf{A})^{\perp}$$

and of course (by reasons of symmetry) for $\mathbf{A}' := conj(A^t)$:

$$\mathbf{Null}(\mathbf{A}') = \quad \mathbf{Colspace}(\mathbf{A})^{\perp}$$

[1] Think of $3x + 4y + 5z = 0$ is just another way to say that the vector $\mathbf{x} = [x, y, z]$ satisfies $\langle \mathbf{x}, [3, 4, 5] \rangle = 0$.

Geometric interpretation of matrix multiplication

Since *clearly* the restriction of the linear mapping $x \mapsto \mathbf{A} * x$ is injective we get an isomorphism $\tilde{T}$ between $Row(\mathbf{A})$ and $Col(\mathbf{A})$.

$$
\begin{tikzcd}
\mathbb{R}^n \arrow[d, "\mathbf{P}_{\mathbf{Row}}"'] \arrow[dr, "T : x \to \mathbf{A} * x"] & \\
Row(\mathbf{A}) \arrow[r, "\tilde{T} = T_{|row(\mathbf{A})}", "\mathcal{R}"'] & Col(\mathbf{A}) \subseteq \mathbb{R}^m
\end{tikzcd}
$$

Geometric interpretation of matrix multiplication

$$
\begin{tikzcd}
Null(\mathbf{A}) \subseteq \mathbb{R}^n \arrow[d, "\mathbf{P}_{\mathbf{Row}}"'] \arrow[drr, "T"] & & \mathbb{R}^m \supseteq Null(\mathbf{A}') \arrow[d, "\mathbf{P}_{\mathbf{Col}}"] \arrow[dll, "T'"'] \\
Row(\mathbf{A}) \arrow[rr, shift left, "\widetilde{T} = T_{|row(\mathbf{A})}"] & & Col(\mathbf{A}) \subseteq \mathbb{R}^m \arrow[ll, shift left, "inv(\widetilde{T})"]
\end{tikzcd}
$$

$$T = \widetilde{T} \circ P_{Row}, \quad pinv(T) = inv(\widetilde{T}) \circ P_{Col}.$$

Four spaces and the SVD

The **SVD** (the so-called Singular Value Decomposition) of a matrix, described in the MATLAB helpful as a way to write A as

$$A = U * S * V'$$

, where the columns of U form an ON-Basis in $\mathbb{R}^m$ and the columns of V form an ON-basis for $\mathbb{R}^n$, and S is a (rectangular) diagonal matrix containing the non-negative *singular values* (σ_k) of A. We have $\sigma_1 \geq \sigma_2 \ldots \sigma_r > 0$, for $r = rank(A)$, while $\sigma_s = 0$ for $s > r$. In standard description we have for A and $pinv(A) = A^+$:

$$A * x = \sum_{k=1}^{r} \sigma_k \langle x, v_k \rangle u_k, \quad A^+ * y = \sum_{k=1}^{r} \frac{1}{\sigma_k} \langle y, u_k \rangle v_k.$$

Generally known facts in this situation

The Four Spaces are well known from LINEAR ALGEBRA, e.g. in the **dimension formulas**:

ROW-Rank of $\mathbf{A}$ *equals COLUMN-Rank of* $\mathbf{A}$*.*

The defect (i.e. the dimension of the Null-space of $\mathbf{A}$*) plus the dimension of the range space of* $\mathbf{A}$ *(i.e. the column space of* $\mathbf{A}$*) equals the dimension of the domain space* $\mathbb{R}^n$*.* Or in terms of linear, homogeneous equations: The dimension of set of all solution to the homogeneous linear equations equals the number of variables minus the dimension of the column space of $\mathbf{A}$.

The SVD also shows, that the *isomorphism* $\widetilde{T}$ *between the Row-space and the Column-space* can be described by a diagonal matrix, if suitable orthonormal bases are used.

Consequences of the SVD

We can describe the quality of the isomorphism $\widetilde{T}$ by looking at its condition number, which is σ_1/σ_r, the so-called **Kato-condition number** of T.
It is not surprising that for **normal matrices** with $A' * A = A * A'$ one can even have diagonalization, i.e. one can choose $U = V$, using to following simple argument:

$$Null(A) =_{always} \; Null(A' * A) = Null(A * A') = Null(A').$$

The most interesting cases appear if a matrix has maximal rank, i.e. if $rank(\mathbf{A}) = min(m, n)$, or equivalently if one of the two Null-spaces is trivial. Then we have either linear independent columns of **A** (injectivity of $T >>$ RIESZ BASIS for subspaces) or the columns of **A** span all of $\mathbb{R}^m$ (i.e. surjectivity, resp. $Null(A') = \{0\}$): $>>$ FRAME SETTING!

Geometric interpretation: linear independent set > R.B.

$$
\begin{tikzcd}
 & & \mathbb{R}^m \supseteq Null(\mathbf{A}') \arrow[dll, "T'"'] \arrow[d, "\mathbf{P}_{\mathbf{Col}}"] \\
Row(\mathbf{A}) = \mathbb{R}^n \arrow[rr, shift left, "\widetilde{T} = T_{|row(\mathbf{A})}"] & & Col(\mathbf{A}) \subseteq \mathbb{R}^m \arrow[ll, shift left, "inv(\widetilde{T}) = pinv(\mathbf{A})"]
\end{tikzcd}
$$

Geometric interpretation: generating set > FRAME

$$\begin{tikzcd}
Null(\mathbf{A}) \subseteq \mathbb{R}^n \arrow[d, "\mathbf{P}_{\mathbf{Row}}"'] \arrow[dr, "T"] & \\
Row(\mathbf{A}) \arrow[r, shift left, "\widetilde{T} = T_{|row(\mathbf{A})}"] & Col(\mathbf{A}) = \mathbb{R}^m \arrow[l, shift left, "inv(\widetilde{T}) = \mathbf{A}'"]
\end{tikzcd}$$

The frame diagram for Hilbert spaces:

Here $C : f \mapsto (\langle f, f_i \rangle)$ is the *analysis operator* and $\mathcal{R}$ the reconstruction operator (cf. the concept of Banach frames).

$$
\begin{tikzcd}
 & \ell^2(I) \arrow[dl, "\mathcal{R}"'] \arrow[d, "\mathbf{P}"] \\
\mathcal{H} \arrow[r, shift right, "\mathcal{C}"'] & \mathcal{C}(\mathcal{H}) \arrow[l, shift right, "\mathcal{R}"']
\end{tikzcd}
$$

This diagram is **fully equivalent** to the usual frame inequalities.

The frame diagram for Banach Gelfand Triples

$C : f \mapsto (\langle f, f_i \rangle)$ is the *analysis operator* and $\mathcal{R}$ the reconstruction operator. The typical situation for Gabor frames with $\mathcal{C}(f) = V_g f|_\Lambda$, for some nonzero atom $g \in \boldsymbol{S}_0(\mathbb{R}^d)$, and $\Lambda \lhd \mathbb{R}^d \times \widehat{\mathbb{R}}^d$, for example.

$$
\begin{tikzcd}
 & (\ell^1, \ell^2, \ell^\infty) \arrow[dl, "\mathcal{R}"'] \arrow[d, "\mathbf{P}"] \\
(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0') \arrow[r, shift right, "\mathcal{C}"'] & \mathcal{C}((\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')) \arrow[l, shift right, "\mathcal{R}"']
\end{tikzcd}
$$

The Kirkwood-Dirac Principle and SVD

There is a quite extensive recent literature on the Kirkwood-Dirac distribution (mostly motivated by applications in Quantum Theory). I prefer to call it the Kirkwood (1933)-Dirac (1945) principle. By analogy to the linear algebra case one can relate it to the simple linear algebra principle, that it might be good to represent a linear mapping using two different bases. The SVD is giving us a strong motivation: Any linear mapping between finite dimensional space can be represented as a diagonal matrix, using suitable ONBs in the domain and target space!
Some of the representations useful for Gabor analysis or the study of time-variant operators are exactly of this form, e.g. the Kohn-Nirenberg symbol, which can be described as a *time-variant transfer function*. It can also be interpreted as the description of the operator from the Fourier basis to the Dirac basis, and is thus a special case of the K-D-principle.

Dirac and Fourier basis

Physicists and engineers use a lot the "continuous Dirac basis" and call it a *continuous orthonormal basis*. While it seems to the intuitive natural extension of the concept of the discrete case it has at least two drawbacks:

1. In which sense can one consider δ_x to be normalized!?
2. Is the system $(\delta_x)_{x \in \mathbb{R}^d}$ really minimal, as one may expect from a basis; wouldn't the family $(\delta_x)_{x \in}$ be equally suited to identify and reconstruct continuous test functions?

Similar things can be said about the Fourier basis $(\chi_s)_{s \in \mathbb{R}^d}$. Here the different normalizations of the FT in the literature are an indication that there is no optimal/natural normalization (can be discussed..). At least, there is always a normalization that turns the FT into a unitary automorphism of $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$ (Plancherel's Theorem).

Mild or Dirac-like Bases I

The wish to extend the K-D-principle to the continous (non-compact) setting, e.g. to $\mathbb{R}^d$, cf. next page.
This suggested terminology relates well to the use of the Dirac basis in physics and other application areas. Strictly speaking it would be more correct to call it a *tight, mild frame*, because a basis is supposed to be minimal, while a frame is a generating system which allows some redundancy. All the *relevant* properties of the would be valid also for the family (for illustration only) $(\delta_q)_{q\in\mathbb{Q}}$ as well, since $\mathbb{Q}$ is a dense subset of $\mathbb{R}$. In a way, one can consider the *tight Dirac frame* as the limit of Gabor frames of the form $(\mathrm{St}_\rho g_0, \rho a, b/\rho)$, using compressed Gaussians as atoms, for $\rho \to 0$, starting from any pair (a, b) with $ab < 1$. Thus the Dirac basis consists of in infinitely fine grid in time and an infinitely sparse grid in frequency, and thus the concept of redundancy does not make much sense anymore for this limiting case.

Mild or Dirac-like Bases II

Definition

We call a family $(\tau_x)_{x\in\mathbb{R}^d}$ a *mild basis* for $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$ if there exists some BGT-automorphism α such that $\tau_x = \alpha(\delta_x)$ for all $x \in \mathbb{R}^d$.
We call it a *mild orthonormal basis* (for short **MOB**) for $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$ if there exists some unitary BGT-automorphism with this property.
BETTER TERMINOLOGY (as of 1.9.2025): We call it a *unitary mild basis* (for short, a **UMB**) for $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$ if there exists some unitary BGT-automorphism with this property.
Among those MOBs the ones arising by the application of a unitary operator from the *metaplectic group* $\mathrm{Mp}(n)$ deserve special attention. We call them *metaplectic UMBs*, for short **MUMBs**.

Mild or Dirac-like Bases III

We might call such mild bases also *Dirac-like bases* for $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$. For the case of a Fractional Fourier transform $\mathcal{F}_\alpha$ (they form a compact subgroup inside of) we might call the corresponding Dirac-like basis a *fractional MOB*.

Using such a setup we can talk about the representation of a given linear operator $T \in \mathcal{L}(\boldsymbol{S}_0, \boldsymbol{S}_0')$ with respect to two chosen say UMBs. The case studied in two recent papers relates to the so-called Rihaczek distribution or the Kohn-Nirenberg symbol, using the Fourier basis (domain) and the Dirac basis in the target space (cf. definition of pseudo-differential operators).

M. Spriet:
Characterizing the Kirkwood-Dirac positivity on second countable LCA groups.
arXiv preprint arXiv:2507.23628, 2025.

Some References I

[gr01]: K. Gröchenig.
Foundations of Time-Frequency Analysis.
Appl. Numer. Harmon. Anal. Birkhäuser, Boston, MA, 2001.

[fest98] H. G. Feichtinger and T. Strohmer.
Gabor Analysis and Algorithms. Theory and Applications.
Birkhäuser, Boston, 1998.

[fezi98]: H. G. Feichtinger and G. Zimmermann.
A Banach space of test functions for Gabor analysis.
In H. G. Feichtinger and T. Strohmer, editors, *Gabor Analysis and Algorithms: Theory and Applications*, Applied and Numerical Harmonic Analysis, pages 123–170. Birkhäuser Boston, 1998.

[feko98]: H. G. Feichtinger and W. Kozek.
Quantization of TF lattice-invariant operators on elementary LCA groups.
In H. G. Feichtinger and T. Strohmer, editors, *Gabor Analysis and Algorithms: Theory and Applications*, Appl. Numer. Harmon. Anal., pages 233–266. Birkhäuser, Boston, MA, 1998.

Some References II

[cofelu08]: E. Cordero, H. G. Feichtinger, and F. Luef.
Banach Gelfand triples for Gabor analysis.
In *Pseudo-differential Operators*, volume 1949 of *Lect. Notes Math.*, pages 1–33. Springer, Berlin, 2008.

[fe19] H. G. Feichtinger.
Classical Fourier Analysis via mild distributions.
MESA, Non-linear Studies, 26(4):783–804, 2019.

[fe20-1]: H. G. Feichtinger.
A sequential approach to mild distributions.
Axioms, 9(1):1–25, 2020.

[fe24-6] H. G. Feichtinger.
Modelling Signals as Mild Distributions (Abdul Siddiqi Memorial Volume), pages 1–26.
Birkhäuser/Springer, Cham, 2024.

[fe25]: H. G. Feichtinger.
An axiomatic approach to mild distributions.
Axioms, pages 1–23, 2025.

THANKS to the audience

THANKS you for your attention

All my papers are found at

https://nuhagphp.univie.ac.at/home/feipub$_d$*b.php*

and my talks are at

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

NOTE: Talks require access code:

"visitor" and "nuhagtalks"