NuHAG Numerical Harmonic Analysis Group

The Extended Fourier Transform in the Context of Mild Distributions

Hans G. Feichtinger, Univ. Vienna and ARI
hans.feichtinger@univie.ac.at
www.nuhag.eu TALKS at www.nuhag.eu/feitalks

Harmonic and Spectral Analysis Conference
Recent Developments and Results
DEBRECEN, October 10th, 2025

Submitted Abstract I

In contrast to Abstract Harmonic Analysis which describes signals defined over locally compact groups and makes use of the corresponding Fourier decomposition the treatment of actual signals (audio-recordings, images, movies for streaming) rather views these signals as a collection of localized objects which are provided "up to some resolution". The MP3-coding scheme is a good example. It provides full information about a piece of music for the duration of a song, up to 20 kHz, by storing the relevant information for segments of length 512 samples (at the sampling rate of 44100, for HiFi recordings). This extension is different from the approach to a generalized Fourier transform using the Schwartz theory of tempered distributions.

Submitted Abstract II

The mathematical theory of mild distributions arose from investigations in time-frequency, specifically Gabor Analysis, i.e. the attempt to put D. Gabor's idea from 1946, claiming that any signal can be described as a double series of time-frequency shifted Gaussians, on solid mathematical grounds. As it turned out, the Segal algebra $S_O(^d)$ (the Feichtinger algebra) is a universal tool serving this purpose.

The talk will describe the setting, how it can be used efficiently for the mathematical description of basic problems in engineering. The thesis is simply that "mild distributions" (elements of $S_0^*(^d)$, the dual of S_0) are the perfect model for signals, with their evaluation on test functions in S_0 being the possible measurements. Another important feature of this setting: all relevant operators have a continuous matrix representation using mild distributions of 2 variables (the Kernel Theorem in this setting).

Submitted Abstract III

By using appropriate concepts of mild convergence (appearing naturally) one can turn heuristic arguments into more conceptual considerations and derive the form of special version of the Fourier transform (e.g. the classical variant for periodic functions, or the DFT/FFT variant for the case of finite signals) into natural variants of one general scheme (the Extended FT for mild distributions).
The goal of this presentation will be to shed some light on the treatment of classical concepts (like the spectrum), using this approach.

Fourier history of in a nut-shell

1. 1822: J.B.Fourier proposes: Every periodic function can be expanded into a Fourier series using only pure frequencies;
2. up to 1922: concept of functions developed, set theory, Lebesgue integration, $\left(\boldsymbol{L}^2(\mathbb{R}), \|\cdot\|_2\right)$;
3. first half of 20th century: Fourier transform for $\mathbb{R}^d$;
4. A. Weil: Fourier Analysis on Locally Compact Abelian Groups;
5. L. Schwartz: Theory of Tempered Distributions
6. Cooley-Tukey (1965): FFT, the Fast Fourier Transform
7. L. Hörmander: Fourier Analytic methods for PDE (Partial Differential Equations);

The Life of Fourier: 1768 - 1830

https://en.wikipedia.org/wiki/Joseph_Fourier

Classical Fourier Series

The classical approach to the theory of *FOURIER SERIES* appears in the following form: Looking at the partial sums of the (formally then infinite) Fourier series we expect them to approximate "any periodic function" in **some sense**[1]:

$$s_N(x) = \frac{a_0}{2} + \sum_{n=1}^{N} [a_n \cos(2\pi nx) + b_n \sin(2\pi nx)]. \tag{1}$$

Assuming this is possible it is not so hard to find out, using the properties of the building blocks ($\cos(x), \sin(x)$, addition rules, derivatives, integration) that one can expect for any $z \in \mathbb{R}$:

$$a_n = \int_z^{z+1} f(x)\cos(2\pi nx)dx, \quad b_n = \int_z^{z+1} f(x)\sin(2\pi nx)dx. \tag{2}$$

[1] For simplicity we assum period 1!

What are the Ingredients and Questions 1

In my course on Fourier series I was taught (like many classical talks) that the representation

$$f(x) \approx \frac{a_0}{2} + \sum_{n=1}^{\infty} [a_n \cos(2\pi nx) + b_n \sin(2\pi nx)]. \tag{3}$$

should be taken only as a "formal expression", which has to be formalized using various kinds of mysterious tricks!

But was does this mean?

What kind of concrete, mathematical questions should be asked? Why and how are summability methods saving the situation, and in which sense?

Until now Fourier series are seen as a mystery!

What are the Ingredients 2

First of all we have to note that by the time (1822!!, which was at the life-time of Carl Friedrich Gauss! [1777-1855]) the modern **concept of a function** was not available as it is now. Thanks to Leonhard Euler ([1707 - 1783]) the complex numbers and their connection to trigonometric functions had been known

$$e^{ix} = \cos(x) + i\sin(x), \quad i = \sqrt{-1}. \tag{4}$$

It was known what *polynomials* are and how to compute with them, and even to take "polynomial of infinite degree" (power series, with well defined regions of uniform convergence), hence **Taylor expansions** were known (going back to the English mathematician Brook Taylor [1685-1731]).

What are the Ingredients 3: the Integral

Of course the determination of the coefficients using **integrals** (over the period of the involved functions) is one of the corner-stones of the classical theory, raising some questions:

- What is the meaning of an integral *in the most general case*?
- What kind of functions can be integrated (over $[a, b]$?
- What can be said about the Fourier coefficients $(a_n)_{n\geq 0}$ or $(b_n)_{n\geq 1}$?

While the foundations of "Calculus" had been laid down by Isaac Newton [1642 - 1726] and Gottfried Wilhelm Leibniz [1646 - 1716] long before Fourier it was **Bernhard Riemann** [1826 - 1866] who gave a clean definition and showed that e.g. every *continuous function* can be integrated over any interval $[a, b]$. He showed that the Fourier coefficients tend to zero $(n \to \infty)$.

What are the Ingredients 4: the Timeline

Another non-trivial part of the reasoning is the computation of the formula. In fact, it is only a necessary condition on the coefficients which can be easily obtained, using integrals

Isaac Newton [1642 - 1726]

Gottfried Wilhelm Leibniz [1646 - 1716]

AFTER FOURIER

Bernhard Riemann [1826 - 1866]

Karl Weierstrass [1815-1897]

Henri Leon Lebesgue [1875 - 1941]

Norbert Wiener [1894 - 1964]

Andre Weil [1906 1998]

What are the Ingredients 5: Perfect Integrals

By the beginning of the 20th century **Henri Leon Lebesgue** had developed his integral, and also given lectures on the application of this new techniques to trigonometric series.
He published a number of important papers between 1904 and 1907 in this direction.
From a modern (functional analytic) view-point his integral, which included the definition of the so-called *Lebesgue spaces* such as $(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1)$ or $(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2)$ (and of course later the $\boldsymbol{L}^p$-theory, duality etc.) opened the way to the field of (linear) **functional analysis**, which developed rapidly, the foundations being lead by e.g. David Hilbert [1862 - 1943], Friedrich Riesz [1880 - 1956] and Stefan Banach [1892 - 1945].

Convergence Issues: Pointwise

So let us return to the question of convergence: The key question being: **In which sense do the partial sums converge?**

In fact, it turned out that a more general problem appeared: What does convergence mean, and can one form classes of functions (nowadays Banach spaces or even topological vector spaces of such objects) such that one can guarantee convergence in those space in the corresponding norm (or topology).
The classical view-point was of course: Can one establish pointwise convergence (Dirichlet-conditions, J.P. Lejeune-Dirichlet [1805-1859])? Or uniform convergence at least for continuous functions (no, according to A.N.Kolmogorov [1903-1987], already in 1923 a found a counter-example and in 1926 he was able to prove that the Fourier series of an L^1-function can **diverge everywhere!**).

Convergence Issues: The idea of Summability

Of course one has to mention **Lipolt Fejer** [1880 - 1959] and a long list of names pursuing the problems related to **summability**.

The idea is to change the question from the question of convergence of the (partial sum) of the Fourier series to *the question of recovering a function from its Fourier coefficients*. For example, Fejer was suggesting to take (as a replacement for the ordinary partial sums) the *arithmetic means of the partial sums*.

Fejer's Theorem of 1900 states that for every continuous periodic function f the (now known as) Fejer means of the Fourier series converges uniformly to f.

Orthogonal Expansions, ONBs in Hilbert Spaces

The pure frequencies $\chi_n(x) := exp(2\pi inx),\ n \in \mathbb{Z}$ form a *complete orthonormal system* for the Hilbert space $\mathcal{H} = \left(\boldsymbol{L}^2(\mathbb{T}), \|\cdot\|_2\right)$:

$$f = \sum_{n\in\mathbb{Z}} \langle f, \chi_n\rangle \chi_n, \quad f \in \mathcal{H}, \tag{5}$$

with unconditional convergence in the $\boldsymbol{L}^2$-norm

$$\|f\|_2 := \sqrt{\int_0^1 |f(x)|^2}.$$

The *coefficients* $c_n := \langle f, \chi_n\rangle,\ n \in \mathbb{Z}$ are uniquely determined and satisfy Parseval's equality:

$$\|f\|_2 = \sqrt{\sum_{n\in\mathbb{Z}} |c_n|^2}. \tag{6}$$

Almost everywhere convergence, Lusin's Conjecture

The convergence issue, in the form of Lusin's conjecture about the convergence of Fourier series in the *pointwise almost everywhere sense* was provided by Lennart in his famous Acta Mathematica paper of 1966. He showed that for every $f \in \boldsymbol{L}^2(\mathbb{T})$ the Fourier series is *almost everywhere convergent*.

Lennart Carleson On convergence and growth of partial sums of Fourier series. *Acta Math.*, 116:135–157, 1966.
This result was of course the counterpoint to Kolmogorov's negative results in the $\boldsymbol{L}^1$-setting (Kolmogorov was a student of Lusin).

Engineering View-Point I

Engineers use the Fourier transform and the concept of convolution for various important applications:

1. Translation invariant systems are described either by their *impulse response* (convolution kernels) or by their *transfer function*, which is the Fourier transform of the impulse response, i.e.

$$T(f) = \sigma * f \quad \text{or} \quad T(f) = \mathcal{F}^{-1}(h \cdot \widehat{f}),$$

   obviously with $h = \mathcal{F}(\sigma)$.

2. Sampling is described as periodization on the Fourier transform side, hence complete recovery can be realized for band-limited functions, if the Nyquist criterion is satisfied, i.e. if *aliasing* is avoided.

Engineering View-Point II

The derivation of these two key principles (and others) is either done only heuristically, or by mystification of an object, known ad Dirac Delta (function? measure? distribution?).

Typically in all these cases one has to distinguish between discrete and continuous signals, but also between periodic and non-periodic, one- or multi-dimensional (e.g. images), because each setting requires the use of a different form of what is always called the Fourier transform and the inverse Fourier transform.

Often enough there is not much discussion whether a given signal "has a Fourier transform" or whether integrals make sense.

Striking Role of Lebesgue Integration I

It appears absolutely natural that the study of the concepts of a Fourier transform and the convolution of functions requires the "optimal theory of integration" (Lebesgue Theory, $\boldsymbol{L}^1(\mathbb{R}^d)$):

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

but also the pointwise (a.e.) definition of convolution

$$f * g(x) := \int_{\mathbb{R}^d} f(x-y) g(y) dy.$$

Combined to the Convolution Theorem we have $\widehat{f * g} = \widehat{f} \cdot \widehat{g}$.

Abstract Harmonic Analysis (AHA)

Abstract Harmonic Analysis goes on to define the Banach convolution algebra $\left(\boldsymbol{L}^1(G), \|\cdot\|_1\right)$, starting from the existence of a Haar measure on a LCA (locally compact, Abelian) group G, and goes on to define the Fourier transform as the concrete form of a Gelfand Transform. The dual group $\widehat{G}$ is identified with the domain, and Pontryagin's Theorem allows to identify the dual group of $\widehat{G}$ with the original group G.
This approach allows a unified viewpoint towards the different settings. The dual group of the torus group $\mathbb{T}$ is $\mathbb{Z}$, a discrete, non-compact group, the dual group (the group of pure frequencies, with pointwise multiplication) of $\mathbb{R}^d$ is $\widehat{\mathbb{R}^d}$, by identifying $s \in \mathbb{R}^d$ with the character $\chi_s(t) = exp(2\pi i\, st), t \in \mathbb{R}^d$.

Fourier Multipliers

In analogy to "system theory" the theory of (Fourier) multipliers was an important topic in the 60s, when the multipliers, the linear operators commuting with translations where studied on the Lebesgue spaces $\boldsymbol{L}^p(G)$. It was shown that they can be characterized as convolution operators by some *quasi-measures* and also as Fourier multipliers by some (other) quasi-measures, but unfortunately the space of quasi-measures has no global restrictions and this is to big to claim a Fourier transform relationship (in the spirit of impulse response versus transfer function).

Shannnon's Sampling Theorem

The exact recovery of band-limited functions in $\boldsymbol{L}^2(\mathbb{R}^d)$ with $\mathrm{spec}(f) = \mathrm{supp}(\widehat{f}) \subset \Omega$ from regular samples (i.e. along some discrete lattice) can be seen as the Fourier variant of the classical Fourier series expansion of $\widehat{f}$.
In fact, the Fourier coefficients of (a periodic version) of $\widehat{f}$ are exactly the samples of f, and the pure frequencies, restricted to the (say cubic) domain give the so-called SINC-function, yielding

$$f(t) = C_\alpha \sum_{k \in \mathbb{Z}^d} f(\alpha k) \, \mathrm{SINC}(t - \alpha k)$$

for any such band-limited f in $\boldsymbol{L}^p(\mathbb{R}^d)$, with $1 < p < \infty$.

Theory of Tempered Distributions

It is of course well known that the theory of *tempered distributions* as introduced by Laurent Schwartz allows to define the (extended) Fourier transform, not only for $f \in \boldsymbol{L}^p(\mathbb{R}^d)$ with $1 \leq p \leq 2$ (Hausdorff-Young), but for any locally integrable function of polynomial growth (say), or to partial derivatives of the Dirac delta. The define partial differential operators and are thus the key-players for PDE theory, now understood as multiplication operators by polynomials on the Fourier side (L. Hörmander, etc.). Schwartz-Bruhat theory imitates the situation over LCA groups. For experts this approach is still of limited scope and they replace the Schwartz space $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ of rapidly decreasing functions by even smaller (nuclear Frechet) spaces, in fact algebras, of test functions which allow to introduce spaces of ultra-distributions.

Downsides of Classical Approach I

The classical/mathematical approach to Fourier Analysis has a couple of downsides:

1. It is hard to use classical methods to analyze time-variant signals, such as music or heart-beat signals;
2. It is not possible to describe time-variant filters (focussing at different levels at different locations), or slowly-time variant signals as they appear in mobile communication (due to Doppler), or astronomy (change in atmosphere);
3. Engineers use numerical variants of existing algorithms without bothering too much about actual approximation of the underlying continuous model;
4. most of the existing theory is mathematically highly demanding (topological vector spaces, etc.) and thus difficult to use for applied scientists.

Downsides of Classical Approach I

The classical/mathematical approach to Fourier Analysis has a couple of downsides:

1. It is hard to use classical methods to analyze time-variant signals, such as music or heart-beat signals;
2. It is not possible to describe time-variant filters (focussing at different levels at different locations), or slowly-time variant signals as they appear in mobile communication (due to Doppler), or astronomy (change in atmosphere);
3. Engineers use numerical variants of existing algorithms without bothering too much about actual approximation of the underlying continuous model;
4. most of the existing theory is mathematically highly demanding (topological vector spaces, etc.) and thus difficult to use for applied scientists.

Mild Distributions: Key Facts I

1. The Feichtinger algebra $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ and its dual, the space of mild distributions are isometrically invariant under TF-shifts and the FT!, via $\widehat{\sigma}(f) = \sigma(\widehat{f})$;
2. All relevant function spaces (such as periodic or discrete signals or $\boldsymbol{L}^p$-spaces) are contained in $\boldsymbol{S}_0'(\mathbb{R}^d)$, in fact, it makes sense to view $\boldsymbol{S}_0'(\mathbb{R}^d)$ as the "space of signals";
3. The natural convergence in $\boldsymbol{S}_0'(\mathbb{R}^d)$ is the w^*-convergence, also called "mild convergence". It corresponds to uniform convergence of their STFT over compact subsets of $\mathbb{R}^d \times \widehat{\mathbb{R}}^d$;
4. Dirac combs (over lattices) belong to $\boldsymbol{S}_0'(\mathbb{R}^d)$;
5. Sequences from $\boldsymbol{S}_0(\mathbb{R}^d)$, or discrete (and/or periodic) signals allow to approximate any $\sigma \in \boldsymbol{S}_0'(\mathbb{R}^d)$.

Identifying Concrete Special Cases I

First of all let us recall that $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ is a Banach algebra of continuous and absolutely integrable functions on $\mathbb{R}^d$ which is dense in $\boldsymbol{S}_0(\mathbb{R}^d)$. In fact, it is a so-called *Segal algebra*, hence a Banach ideal in $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$, satisfying

$$\|g * f\|_{\boldsymbol{S}_0} \leq \|g\|_{\boldsymbol{L}^1} \|f\|_{\boldsymbol{S}_0}, \quad g \in \boldsymbol{L}^1, f \in \boldsymbol{S}_0.$$

Consequently the Fourier inversion formula applies directly to functions $f \in \boldsymbol{S}_0(\mathbb{R}^d)$, but due to the simply fact that $\|\widehat{f}\|_{\boldsymbol{C}_0} \leq \|f\|_{\boldsymbol{L}^1}$ (for $f \in \boldsymbol{S}_0(\mathbb{R}^d)$) it is clear that the FT can be extended to a bounded linear mapping from $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ onto $\left(\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_{\mathcal{F}\boldsymbol{L}^1}\right) \hookrightarrow \left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$.
Due to the Fourier invariance we also have

$$\|h \cdot f\|_{\boldsymbol{S}_0} \leq \|h\|_{\mathcal{F}\boldsymbol{L}^1} \|f\|_{\boldsymbol{S}_0}, \quad h \in \mathcal{F}\boldsymbol{L}^1, f \in \boldsymbol{S}_0.$$

Identifying Concrete Special Cases II

It is a well-known fact that $f \in \boldsymbol{L}^1(\mathbb{R}^d)$ does NOT imply $\widehat{f} \in \boldsymbol{L}^1(\mathbb{R}^d)$, and thus there are many cases which do not allow to apply the Fourier inversion formula directly, even (or especially) if one makes use of Lebesgue integration theory. Thus one has to apply *summability methods*. Choosing some function $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d} g(x)dx = \widehat{g}(0) = 1$ one defines $h = \widehat{g}$, thus with $h(0) = 1$ and multiplies $\widehat{f}$ by $\mathrm{D}_\rho h$ (with $\mathrm{D}_\rho h(y) = h(\rho y)$, for $\rho \to 0$. Since $\mathrm{D}_\rho h = \mathcal{F}()$, with $\mathrm{St}_\rho g(x) = \rho^{-d} g(x/\rho)$ forming a Dirac sequence one can recover f from $\widehat{f}$ since

$$f = \lim_{\rho \to 0} \mathrm{St}_\rho g * f \text{ in } \left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right).$$

It has been verified by F. Weisz (ELTE Univ., Budapest) that **all** *the classical summability kernels* actually belong to $\boldsymbol{S}_0(\mathbb{R}^d)$, thus providing a kind of universal argument for their usefulness.

Dirac Combs I

Another important setting are discrete, periodic signals. Usually one has to use the DFT (Discrete Fourier transform), realizable by the FFT (Fast FT) in order to compute the FT in this setting. But since such signals can be viewed as finite linear combinations of Dirac combs. Noting furthermore that they can be obtained by transformation of the standard Dirac comb $\sqcup\!\sqcup = \sum_{k\in\mathbb{Z}^d} \delta_k$ under the (extended) Fourier transform, which is (aside from the Gaussian) the most important Fourier invariant signals. In order to verify $\mathcal{F}(\sqcup\!\sqcup) = \sqcup\!\sqcup$ one "has to verify" *Poisson's formula*:

$$\sum_{k\in\mathbb{Z}^d} \widehat{f}(k) = \sqcup\!\sqcup(\widehat{f}) = \sqcup\!\sqcup(f) = \sum_{k\in\mathbb{Z}^d} f(k), \quad f \in \boldsymbol{S}_0(\mathbb{R}^d).$$

We mention that the same result is much weaker if one only requires the invariance in the sense of tempered distributions, i.e. Poisson's formula for $f \in \mathcal{S}(\mathbb{R}^d)$.

Regular Partitions of Unity in $\mathcal{F}L^1(\mathbb{R})$

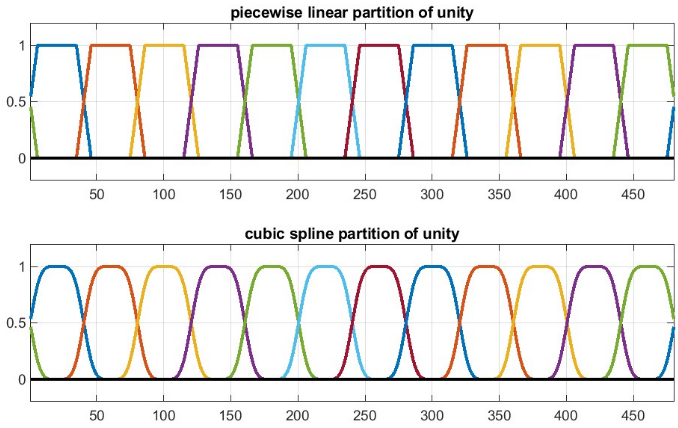


Abbildung: BUPUPLNCUB25A.jpg

Periodic Functions I

The classical Fourier transform for *periodic functions* should actually be seen as the Fourier transform applied to functions which have been obtained by *periodization*! of some compactly supported function (e.g. by localizing the periodic function with the help of a piecewise linear partition of unity); In this way we find that the functions in Wiener's algebra of absolutely convergent Fourier series arise by periodization of functions from $\boldsymbol{S}_0(\mathbb{R}^d)$. Such functions of the form $p = f * ⧢_\Lambda$ for some lattice Λ, with $⧢_\Lambda = \sum_{\lambda\in\Lambda} \delta_\lambda \ \delta_\lambda$, and by the convolution theorem:

$$\mathcal{F}(p) = \mathcal{F}(⧢_\Lambda * f) = \mathcal{F}(⧢_\Lambda) \cdot \widehat{f} = ⧢_{\Lambda^\perp} \cdot f = \sum_{\lambda^\perp\in\Lambda^\perp} \widehat{f}(\lambda^\perp)\delta_{\lambda^\perp}.$$

Of course, these coefficients do not depend on the particular choice of f (as long as $p = f * ⧢_\Lambda$) and the coefficients appearing are just the classical Fourier coefficients.

Letting the Period go to Infinity

The transition from the classical Theory of Fourier Expansions for periodic functions to the continuous integral formula is often motivated heuristically by *letting the period go to infinity* In the sense of mild convergence one has in fact

$$\lim_{\tau \to \infty} ⧢_{\tau \mathbb{Z}^d} * f = f, \quad f \in \boldsymbol{S}_0(\mathbb{R}^d)$$

and thus these periodic version tend to f in the sense of mild distributions. Recalling the dilation property of the FT we have $\mathcal{F}(⧢_{\tau \mathbb{Z}^d}) = \tau^{-d} ⧢_{\mathbb{Z}^d/\tau}$. This corresponds to mild convergence of the discrete measures (in fact uniformly bounded measures) on the FT side:

$$\widehat{f} \cdot \tau^{-d} ⧢_{\mathbb{Z}^d/\tau} \to \widehat{f}, \quad f \in \boldsymbol{S}_0(\mathbb{R}^d).$$

Giving a Meaning to Terms used in Physics I

While it is difficult to explain how discrete measures can converge to a continuous function in $\boldsymbol{L}^1(\mathbb{R}^d)$, even in the sense of bounded measures, it is plausible that one has mild (and in this case in fact equivalently *vague*) convergence. At a technical level it boils down to the observation that Riemann sums converge to the Riemann integral whenever on applies them to a function $h = f \cdot g$, with f, g and hence $h \in \boldsymbol{S}_0(\mathbb{R}^d)$.
In a similar way one can give a meaning to formulas used by physicists, such as the *resolution of identity* via pure frequencies

$$\mathbf{Id} = \int_{\mathbb{R}^d} |\chi_s\rangle\langle\chi_s| ds$$

Giving a Meaning to Terms used in Physics II

(using the Bra-Ket notation of Dirac) or the *sifting property of the Dirac Delta* used by engineers:

$$f = \int_{\mathbb{R}^d} f(x)\delta_x dx, \quad f \in \boldsymbol{S}_0(\mathbb{R}^d)$$

or also (in the sense of mild limits):

$$\mathbf{Id} = \int_{\mathbb{R}^d} |\delta_x\rangle\langle\delta_x| \; dx,$$

which is used in order to describe translation invariant linear systems, even if δ_0 is not in the domain of this operators (say defined on $\boldsymbol{L}^2(\mathbb{R}^d)$).

The key-players for time-frequency analysis I

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

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

Some important properties of the STFT I

Starting from the Hilbert space $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$ with scalar product

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

one can find that for fixed g the linear mapping $f \mapsto V_g f$ is isometric from $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$ into $\left(\boldsymbol{L}^2(\mathbb{R}^{2d}), \|\cdot\|_2\right)$. But also

$$\|V_g f\|_\infty = \max_{\lambda\in\mathbb{R}^d\times\widehat{\mathbb{R}}^d} |V_g f(\lambda)| \leq \|f\|_2\|g\|_2$$

by the Cauchy-Schwarz inequality, since $\|\pi(\lambda)g\|_2 = \|g\|_2$, for any $\lambda \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d$.
There is also orthogonality for the windows, since we have

$$V_g f(\lambda) = \langle f, \pi(\lambda)g\rangle = \langle \pi(\lambda)^* f, g\rangle,$$

Some important properties of the STFT II

which gives essentially symmetry between f and g:

$$V_g f(t,s) = e^{2\pi i t s} V_{\widehat{g}} \widehat{f}(s,-t).$$

The next property is sometimes called the *covariance property* of the STFT.

Lemma (3.1.3)

Whenever $V_g f$ is defined, we have

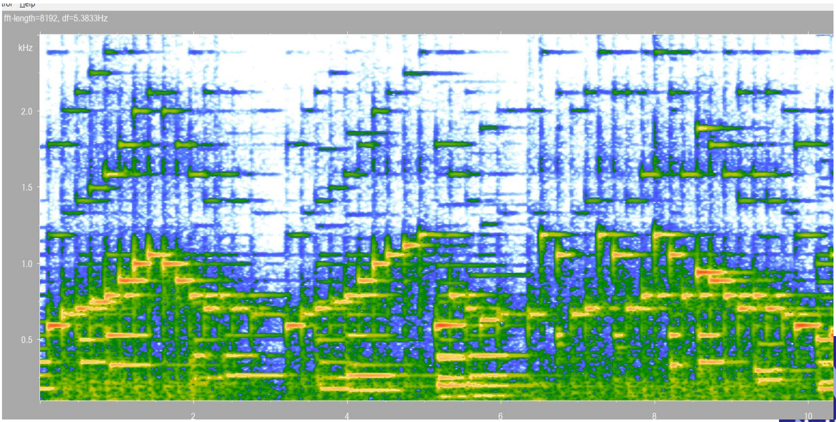
$$V_g(T_u M_\eta f)(x,\omega) = e^{-2\pi i u \cdot \omega} V_g f(x-u, \omega-\eta) \quad (3.14)$$

for $x, u, \omega, \eta \in \mathbb{R}^d$. In particular,

$$|V_g(T_u M_\eta f)(x,\omega)| = |V_g f(x-u,\omega-\eta)| .$$

A Typical Musical STFT

A typical piano spectrogram (Mozart), from recording



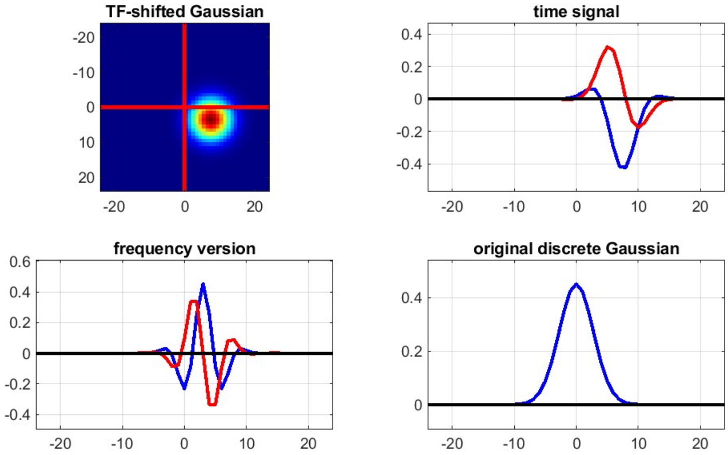
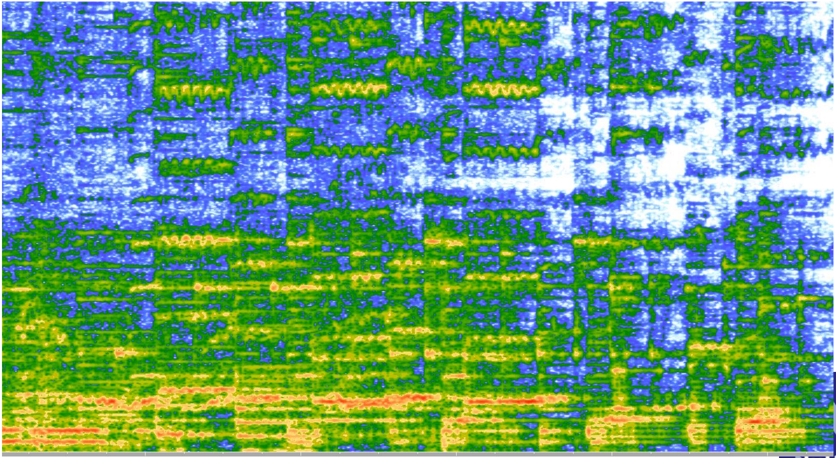
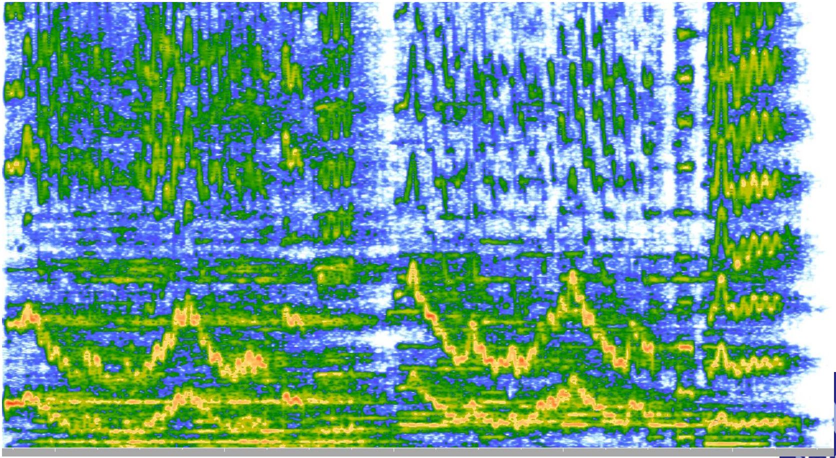


Abbildung: g48TFshifts.jpg

A Musical STFT: Brahms, Cello



A Musical STFT: Maria Callas



A Banach Space of Test Functions (Fei 1979) I

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

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

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

A Banach Space of Test Functions (Fei 1979) II

Since one has for any pair $f, g \in \boldsymbol{L}^2(\mathbb{R}^d)$

$$\|V_g f\|_\infty \le \|f\|_2 \|g\|_2,$$

as a simple consequence of the Cauchy-Schwarz inequality, this is stronger then the corresponding norm in $\boldsymbol{L}^2(\mathbb{R}^{2d})$. In fact one has

$$\|V_g f\|_2 = \|f\|_2 \|g\|_2, \quad f, g \in \boldsymbol{L}^2(\mathbb{R}^d).$$

This implies that the range of V_g is a closed, invariant subspace of $\boldsymbol{L}^2(\mathbb{R}^d)$, and the projection operator is (twisted convolution operator), mapping $\left(\boldsymbol{L}^2(\mathbb{R}^{2d}), \|\cdot\|_2\right)$ onto $V_g(\boldsymbol{L}^2(\mathbb{R}^d))$. If $g \in \boldsymbol{S}_0(\mathbb{R}^d)$, then the convolution kernel is in $\boldsymbol{L}^1(\mathbb{R}^{2d})$. Assuming $\|g\|_2 = 1$ we have the *reconstruction formula*:

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

which can be approximated in $\boldsymbol{L}^2$ by Riemannian sums.

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

Lemma

Let $f \in S_0(\mathbb{R}^d)$, then the following holds:

1. $S_0(\mathbb{R}^d) \hookrightarrow L^1 \cap C_0(\mathbb{R}^d)$ *(dense embedding);*
2. $\left(S_0(\mathbb{R}^d), \|\cdot\|_{S_0}\right)$ *is a Banach algebra under both pointwise multiplication and convolution;*

(1) $\pi(u,\eta)f \in S_0(\mathbb{R}^d)$ *for* $(u,\eta) \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d$*, and* $\|\pi(u,\eta)f\|_{S_0} = \|f\|_{S_0}$.

(2) $\hat{f} \in S_0(\mathbb{R}^d)$*, and* $\|\widehat{f}\|_{S_0} = \|f\|_{S_0}$.

In fact, $\left(S_0(\mathbb{R}^d), \|\cdot\|_{S_0}\right)$ is the smallest non-trivial Banach space with this property, and therefore contained in any of the L^p-spaces (and their Fourier images), for $1 \le p \le \infty$.

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

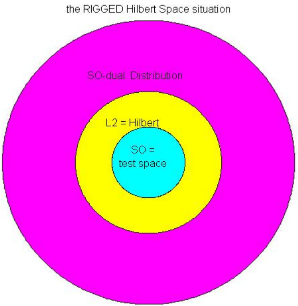


Abbildung: Three layers

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 *Feichtinger's 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 $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$,

known as space of *mild distributions*. Note that these spaces can be defined without great difficulties on any LCA group 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

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

and the corresponding norm

$$\|f\|_{S_0} := \int_{\mathbb{R}^{2d}} |V_{g_0}(f)(x,y)|dxdy < \infty.$$

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* $\boldsymbol{S}_0(\mathbb{R}^d)\hat{\otimes}\boldsymbol{S}_0(\mathbb{R}^d) \approx \boldsymbol{S}_0(\mathbb{R}^{2d})$ *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: $\boldsymbol{H}_G(\boldsymbol{S}_0, \boldsymbol{S}_0') \approx \boldsymbol{S}_0'(G)$ (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 $\{\tau \in \boldsymbol{S}_0(G) \mid \operatorname{supp}(\tau) \subset H\}$.

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 $\mathcal{H}S$:

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

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:

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

Observables and States I

The current status of the development suggest to consider to BE mild distributions (!synonymously). Signal spaces but also measurement have the structure of a linear space and of course one expects that the measurement process is a **bilinear mapping**. We propose to model this process by the pairing defined on $\boldsymbol{S}_0'(\mathbb{R}^d) \times \boldsymbol{S}_0(\mathbb{R}^d)$ (viewed as a Banach space), with measurements of signals being ugiven by

$$(\sigma, f) \to \sigma(f), \quad \sigma \in \boldsymbol{S}_0'(\mathbb{R}^d), f \in \boldsymbol{S}_0'(\mathbb{R}^d).$$

In analogy to the representation of linear mappings on $\mathbb{R}^n$ via $n \times n$-matrices $\boldsymbol{A}$ one can describe operators via "mild distributions on the product space". In fact, there is a *Kernel Theorem* describing a one-to-one correspondence between the (most general) bounded linear operators $T \in \mathcal{L}(\boldsymbol{S}_0, \boldsymbol{S}_0')(\mathbb{R}^d)$ and the corresonding *kernels* $\sigma = \sigma_T \in \boldsymbol{S}_0'(\mathbb{R}^{2d})$.

Observables and States II

This identification can be seen as a starting point for the study of (many, quite general) operators T. Among the various descriptions of the corresponding signals σ_T of (phase space) "two variables", also known as observables, some are more suitable for the study of physical measurements than others. In particular, the *Weyl-Wigner calculus*, but also the pair of *Kohn-Nirenberg* or spreading function representations are the most important ones. They appear in the study of slowly time-variant systems or *pseudo-differential operators*.

If we think of the Hilbert-Schmidt version of the kernel theorem and observe that any $\sigma \in \boldsymbol{L}^2(\mathbb{R}^{2d}) \subset \boldsymbol{S}_0'(\mathbb{R}^{2d})$ defines a (compact) Hilbert-Schmidt operator. At this level the kernel theorem becomes a unitary mapping, if we use the scalar product for $T, S \in \mathcal{HS}$ by the formula $\langle T, S \rangle_{\mathcal{HS}} := \text{trace}(TS^*)$.

Observables and States III

Finally we point to the so-called *Inner Kernel Theorem*, which describes the operators T with kernel $\sigma_T \in \boldsymbol{S}_0(\mathbb{R}^{2d})$. These operators from a Banach space of operators which are all trace class (but in fact more restrictive). These operators have the property that they map bounded, w^*-convergent sequences in $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$ into norm convergent sequences in $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$, i.e. they have a *regularizing property*.

Typical examples are product-convolution operators of the form $f \mapsto (f * g) \cdot h$, with $g, h \in \boldsymbol{S}_0(\mathbb{R}^d)$, or finite partial sums for Gabor frames arising from (g, Λ), with $\Lambda \lhd \mathbb{R}^d \times \widehat{\mathbb{R}}^d, g \in \boldsymbol{S}_0(\mathbb{R}^d)$. Obviously the duality between those nuclear/regularizing operators and the general operators, which can be realized by the trace at the level of Hilbert-Schmidt operators, can be extended to this pairing, including unbounded operators on $\boldsymbol{L}^2(\mathbb{R}^d)$.

Observables and States IV

The Wigner function of a given function $f \in \mathbf{L}^2(\mathbb{R}^d)$ is a *real-valued function* on phase space $\mathbb{R}^d \times \widehat{\mathbb{R}}^d$. If f is normalized with $\|f\|_2 = 1$ then the scalar product between a bonded operator T on $\mathbf{L}^2(\mathbb{R}^d)$ and the corresponding projection operator on $\mathbf{L}^2(\mathbb{R}^d)$, given by $h \mapsto P_f(g) = \langle h, f \rangle_{\mathbf{L}^2} f$, is given by trace$(T \circ P_f)$, or equivalently the ordinary $\mathbf{L}^2(\mathbb{R}^{2d})$-scalar product of their Weyl symbols. Since the Wigner function of f (in $\mathbf{C}_0(\mathbb{R}^{2d})$) is just the Weyl symbol of the projection operator *Moyal's Identity* implies

$$\mathrm{trace}(T \circ P_f) = \int_{\mathbb{R}^d \times \widehat{\mathbb{R}}^d} \mathrm{Weyl}(T)(x,\xi) \cdot \mathrm{Wig}(f)(x,\xi) dx d\xi$$

Specifically of f is an eigenvector of $T = T^*$ one has:

$$\langle T, P_f \rangle_{\mathcal{HS}} = \langle f | T | f \rangle_{\mathbf{L}^2} = \langle f, \lambda f \rangle_{\mathbf{L}^2} = \lambda \|f\|^2_{\mathbf{L}^2} = \lambda.$$

Observables and States V

A Feichtinger state (according to Maurice de Gosson) arises from an element $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ which is normalized in $\boldsymbol{L}^2(\mathbb{R}^d)$ and since the space of self-adjoint operators obtained as absolutely convergent sums of such operators is just the space of operators with kernels in $\boldsymbol{S}_0(\mathbb{R}^{2d})$ (as described by the inner kernel theorem) the duality can be extended to all the operators in $\mathcal{L}(\boldsymbol{S}_0, \boldsymbol{S}_0')(\mathbb{R}^d)$. The so-called Weyl-Heisenberg operators $\widehat{T}_z$ with $z = (x, \xi)$ then just correspond to the Dirac measures δ_z on phase space. This can be considered as one of the starting points of Quantum Theory. There is also a raising area of analysis, called Quantum Harmonic Analysis, which goes back to a paper by R. Werner (1984). It introduces a convolution of operators and a (operator) Fourier transform which turns this convolution into multiplication....

Comments on Observables from Literature I

In quantum mechanics, an **observable** is a measurable physical quantity such as position, momentum, or energy. Each observable corresponds to a Hermitian (self-adjoint) operator acting on a Hilbert space of quantum states. The eigenvalues of this operator represent the possible outcomes of a measurement, and the state collapses into the corresponding eigenvector upon measurement. Let $\hat{A}$ be the Hermitian operator associated with an observable A. For a state $|\psi\rangle$, the expectation value of A is

$$\langle A \rangle = \langle \psi | \hat{A} | \psi \rangle$$

Eigenvalues $\{a_i\}$ and eigenvectors $\{|\phi_i\rangle\}$ satisfy

$$\hat{A}|\phi_i\rangle = a_i|\phi_i\rangle$$

The Bra-Ket Notation I

In quantum mechanics, the expression $\langle\phi|A|\phi\rangle$ represents the *expectation value* of the observable associated with the operator $\hat{A}$ when the system is in the state $|\phi\rangle$. Detailed reasoning:

- The ket $|\phi\rangle$ is a vector representing the quantum state.
- The bra $\langle\phi|$ is the Hermitian conjugate (dual vector) of $|\phi\rangle$.
- The operator $\hat{A}$ acts on $|\phi\rangle$ to produce a new vector $\hat{A}|\phi\rangle$.
- The inner product $\langle\phi|\hat{A}|\phi\rangle$ is a complex number obtained by combining $\langle\phi|$ with $\hat{A}|\phi\rangle$.
- Since $\hat{A}$ is Hermitian (self-adjoint), $\langle\phi|\hat{A}|\phi\rangle$ is real;
- Physically, if many measurements of the observable A are performed on identically prepared systems in state $|\phi\rangle$, their average outcome converges to $\langle\phi|\hat{A}|\phi\rangle$.

Thus $\langle\phi|A|\phi\rangle$ gives the *expected (mean) measurement value* of the observable/signal A in the quantum state $|\phi\rangle$.

Advantages of the Mild Context I

Although the kernel theorem allows to deal with operators which are not just bounded on $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$ but even preserve the structure of the BGT $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$, this level of generality is only of minor importance in this context.

Most of the time we are interested in self-adjoint (Hermitean) BGT-morphisms, i.e. linear operators which are self-adjoint at the Hilbert space level, but also map $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ into itself. In such a case the self-adjointness implies that they also map $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$ into itself (boundedly and w^*-w^*-continuously).

This has the big advantage that one may hope to find true eigenvectors, let as call them perhaps eigen-signals inside of $\boldsymbol{S}_0'(\mathbb{R}^d)$. The prototypical example is the commutative family of translation operators with pure frequencies χ_s as eigenvectors.

Summary on Mild Distributions I

The setting of THE Banach Gelfand Triple $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$ allows to describe not only phenomena arising in the context of time-frequency and Gabor Analysis (where they are almost indispensable tools) but they also help to make vague and heuristic transitions used especially in applied courses (for engineers and physicist) mathematically correct in the sense of a well-defined limit in the sense of mild convergence.

Compared to the well-established theory of tempered distributions it is much easier from a technical point of view, and since $\boldsymbol{S}_0'(\mathbb{R}^d) \hookrightarrow \boldsymbol{\mathcal{S}}'(\mathbb{R}^d)$ as a relatively small subspace statements which can be formulated in the context of mild distributions (such as $\mathcal{F}(\sqcup\!\sqcup) = \sqcup\!\sqcup$) are stronger! because it expresses the validity of Poisson's formula for either just $f \in \boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ or in $\boldsymbol{S}_0(\mathbb{R}^d)$. In addition it allows to handle the *Fractional Fourier transform.*

Links and Refs

Let us provide a few links:

ETH Course Notes are found at: www.nuhag.eu/ETH20

At the page www.nuhag.eu/feitalks:
you find my talks and also recorded talks from some courses, e.g.
FROM LINEAR ALGEBRA TO GELFAND TRIPLES.
Various Lecture Notes can be downloaded from
https://www.univie.ac.at/nuhag-php/home/skripten.php
The public version of my publications can be provided upon request, via hans.feichtinger@univie.ac.at !!

This talk provides some thought related to the CONCEPTUAL HARMONIC ANALYSIS idea.

References concerning Mild Distributions I

[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.

[fe24-5] H. G. Feichtinger.
Fourier analysis via mild distributions: group theoretical aspects.
volume Trends in Mathematics by Birkhäuser., pages 1–6, 2024.

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

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

References concerning the Banach Gelfand Triple I

[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.

[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.

[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.

[fe09] H. G. Feichtinger.
Banach Gelfand triples for applications in physics and engineering.
In *AIP Conference 2009*, volume 1146 of *AIP Conf. Proc.*, pages 189–228. Amer. Inst. Phys., 2009.

References concerning the Banach Gelfand Triple II

[fe18-3] H. G. Feichtinger.
Banach Gelfand Triples and some Applications in Harmonic Analysis.
In J. Feuto and M. Essoh, editors, *Proc. Conf. Harmonic Analysis (Abidjan, May 2018)*, pages 1–21, 2018.

[fe24] H. G. Feichtinger.
Sampling via the Banach Gelfand Triple.
In Stephen D. Casey, Maurice Dodson, Paulo J.S.G. Ferreira, and Ahmed Zayed, editors, *Sampling, Approximation, and Signal Analysis Harmonic Analysis in the Spirit of J. Rowland Higgins*, Appl. Num. Harm. Anal., pages 211–242. Cham: Springer International Publishing, 2024.

[fe24-7] H. G. Feichtinger.
THE Banach Gelfand triple and its role in classical Fourier analysis and operator theory.
Internat. Tbilisi Analysis & PDE Seminar, pages 67–75, 2024.

Fractional Fourier Transforms I

The *fractional Fourier transform* of order $\alpha \in \mathbb{R}$ is defined as:

$$\mathcal{F}_\alpha[f](u) = e^{-i\pi \operatorname{sgn}(\sin\alpha)/4} \int_{-\infty}^{\infty} f(t) K_\alpha(t,u)\, dt,$$

where the kernel $K_\alpha(t,u)$ is given by:

$$K_\alpha(t,u) = \begin{cases} \sqrt{1 - i\cot\alpha}\, e^{i\pi(t^2\cot\alpha - 2tu\csc\alpha + u^2\cot\alpha)}, & \alpha \neq n\pi, \\ \delta(t-u), & \alpha = 2n\pi, \\ \delta(t+u), & \alpha = (2n+1)\pi. \end{cases}$$

Using chirp multipliers and the ordinary Fourier transform, the Fractional Fourier Transform can be implemented as follows, using chirp multipliers. Recall that $\csc(\theta) = 1/\sin(\theta)$ and $\cot(\theta) = \cos(\theta)/\sin(\theta)$.

Fractional Fourier Transforms II

1. Multiply $f(t)$ by $e^{-i\pi t^2 \cot(\alpha)/2}$ to obtain:
$$f_1(t) = f(t) e^{-i\pi t^2 \cot(\alpha)/2}.$$
2. Compute the Fourier transform:
$$f_2(u) = \mathcal{F}[f_1(t)](u).$$
3. Multiply $f_2(u)$ by $e^{-i\pi u^2 \cot(\alpha)/2}$:
$$f_3(u) = f_2(u) \cdot e^{-i\pi u^2 \cot(\alpha)/2}.$$
4. Scale the result by $\sqrt{|\csc(\alpha)|}$ and include the phase factor:
$$\mathcal{F}_\alpha[f](u) = \sqrt{|\csc(\alpha)|} e^{-i\pi \mathrm{sgn}(\sin\alpha)/4} \cdot f_3(u).$$