NuHAG Numerical Harmonic Analysis Group

Signals ARE mild distributions: An alternative approach to Fourier Analysis

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

LMU talk, held Friday 7.6.2024, 14:15

What is a signal? I

ChatGPT says:
A SIGNAL is any kind of measurable quantity that conveys information. It is a function that represents the variation of some physical quantity over time or space (!? temperature, audio-signals, optical signals?).

In physics, a FORCE is any interaction that, when unopposed, changes the motion of an object. It can cause an object with mass to change its velocity (to accelerate), which includes starting, stopping, or changing direction. Force can also cause objects to deform. The concept of force is fundamental in classical mechanics and is described by Newton's laws of motion.

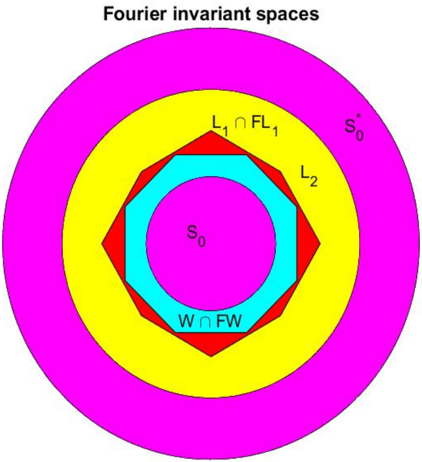
What is a signal? II

Since signal processing obviously means
handle signals: move them from input side to output
one should have a clear definition of **what a signal IS**. As a matter of fact this "ontological question" is rarely asked.

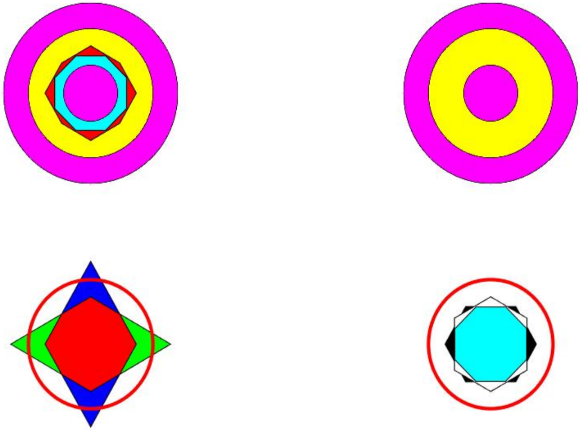
ChatGPT provides the following answer:

```
An ontological description of signals involves
defining what signals are in the most fundamental
terms, considering their nature, existence, and the
relationships they have within various contexts.
Here is a structured ontological description.
Signals are entities that carry information from one
point to another.  They are manifestations of some
underlying phenomena that convey data, commands,
or other forms of communication.
```

Fourier Invariant Spaces



Fourier Invariant Spaces II



Mathematical Analysis vs. Engineering Reality I

Fourier Analysis is an important branch of **Mathematical Analysis**. Once Fourier Analysis had been developed far enough with the help of L. Schwartz's theory of *tempered distributions* PDE became accessible to a Fourier-analytic treatment. The development which has been initiated by work of L. Hörmander has a long-lasting impact in the domain of PDEs, and the study of PDEs with variable coefficients has led to the development of the theory of *pseudo-differential operators*, typically described via the Kohn-Nirenberg calculus, i.e. via a *time-variant* transfer function containing all the relevant information about the operator.
Thus in the mathematical description we have functions or distributions describing domain and range of various operators (e.g. Fourier transforms), as well as different ways to describe operators (via their distributional kernel, a kind of "continuous matrix representation") or the KNS-symbol, and so on.

Mathematical Analysis vs. Engineering Reality II

Clearly *convolution* (between functions, or functions and distributions, or functions and operators in Quantum Harmonic Analysis) are important operations, and the (appropriate) Fourier transform transfers these (usually considered *cumbersome*) operations to simple pointwise multiplication.
In physics and in particular in engineering, especially in the *digital age* Fourier Analysis is at the basis of data transmissions and **signal processing** (think if data formats used from storage or streaming of audio or image content), and therefore also highly important.
Unfortunately the connections between the mathematical developments and the real-world applications are getting more and more vague and the members of one community have a hard time to even understand what the other side is doing, or needs (in terms of tools, etc.). It is true that first year students (typically in communication theory) learn about time-invariant systems.

Mathematical Analysis vs. Engineering Reality III

They hear about different types of signals, discrete and continuous, periodic and non-periodic, and corresponding different so-called *Fourier transformations* (one for each setting!), with the FFT being the work-horse of many digital signal processing algorithms. Invariant systems are then easily understood as convolution operators by the *impulse response*, but for the continuous setting one has to invoke the mysterious Delta-function, which verifies the so-called *sifting property*

$$f(t) = \int_{-\infty}^{\infty} f(y)\delta(x-y)dy \quad (= f * \delta_0),$$

or the claim that

$$\delta(t) = \int_{-\infty}^{\infty} e^{2\pi i s t} ds \quad (= \mathcal{F}^{-1}(\mathbf{1})).$$

Submitted Abstract I

Since the presentation of the Segal algebra $\left(\boldsymbol{S}_0(G), \|\cdot\|_{\boldsymbol{S}_0}\right)$ (for LCA groups) in February 1979 this space, meanwhile known as "*Feichtinger's Algebra*", and its dual, more recently popularized under the name of "*mild distributions*" have been a tremendously useful tool for Harmonic Analysis. Our THESIS:
Mild distributions are a suitable model for signals as they are understood in the application domain.

In fact: *Signals* are described by *Measurements*, very much like a picture is not an $\boldsymbol{L}^2$-function on a square well defined almost everywhere, but rather something which allows to take digital images using fine sensors (with output proportional to intensity of light). **MP3** takes a digital audio recording and transforms it into a compact format (by forming specific linear combinations of basic measurements).

Submitted Abstract II

Trying to describe signals as elements of $\boldsymbol{S}_0'(\mathbb{R}^d)$ and measurements as test functions from $\boldsymbol{S}_0(\mathbb{R}^d)$ we can build up the Banach Gelfand Triple $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$ which should be taken as the basis for Time-Frequency and Gabor Analysis, but also for Classical Fourier Analysis or the theory of pseudo-differential operators.

The extended (fractional) Fourier transform leaving the space of mild distributions invariant supports the physicists intuition that signals can be either described in the time or in the frequency domain, and of course in many different ways as (bounded, continuous) functions on phase space.

The Plan for this Talk

Unlike technical talks (presenting theorems and proofs) or *survey talks* (which I like more), providing a summary over a certain territory, this is *perspective talk*, under the regime:

What if we would not have 200 years of Fourier Analysis and corresponding developments of analysis behind us?

Maybe we should assume that a novice student has basic understanding of complex numbers, exponential functions and principles of functional analysis. How could we describe the mathematics needed for applications in signal processing?
We will discuss "what signals are" and measurements, and how we think of them as possible mathematical objects, allowing conclusions, make predictions, develop algorithms.

Abstract versus Concrete I

As a warm-up exercise let us just think of those poor students, who are exposed to university courses in mathematics, physics, chemistry or engineering.
Usually it is thought that students in mathematics have to deal with abstract objects while the applied sciences learn more how to describe concrete things.

But this is (for ME) a misleading perspective!

Mathematics is emphasizing more the logical structure of things, while the emphasis in physics may be more on *models of the reality* that help to explain physical phenomena. Finally engineers want to make things work, build a steam engine or a mobile communication system, and for this purpose it is often a matter of efficiency to identify the relevant formulas and use them appropriately.

Abstract versus Concrete II

It is correct, that even real numbers are not "as concrete as fractions" of the form 3/4, but eventually analysis courses help us to understand that real and complex numbers are just objects which follow certain rules, in fact they form a FIELD. Thus the usual computations are possible, and even the definition and verification of the exponential law for complex numbers can be derived on such grounds (and thus can be used easily, once it is understood).

On the other hand physicists talk about "point masses", "point charges", which are convenient for the description of physical processes, and we do not ask "whether they exist" (in an ontological sense).

Abstract versus Concrete III

Even more abstract seems to be the idea of a force. Nobody can SEE or TOUCH a force, such as the *gravitational force*. But we can easily let it act, but letting an object fall to the ground. In a similar way we understand *centrifugal forces* arising in a rotating object, and it appears quite concrete if we add such forces in order to model a carousel.

Having said this we realize that we do not have to say, WHAT a signal is (we should think of all possible signals, e.g. those recorded or measured by various XXX-meters), but rather how we can register or *measures* a signal and how to modify it (via some *signal processing* method.

So let me finally to COME TO THE POINT.

The case of Audio Signals

In order to further motivate our approach further let me define an *audio signal* to be something which we can record and which allows to compute an STFT, a short time Fourier transform (say with the help of the ARI programme STX!). What we need is a microphone and a basic understanding of the algorithm producing the STFT. Such a program allows to choose the *sampling rate* (so the microphone converts vibrations of the air into the digital world) and a *window length* (the program chooses a kind of Gabor window). Of course we expect that this output will depend linearly on the input, meaning that the superposition principle applies, so loud signals will produce large amplitudes. And of course there is (time-) shift invariance.

We can say, a (realistic) audio signal is something where the so computed output (STFT) is a bounded (discrete) sequence!

One more "abstract" mathematical model where we can learn is probability theory. How can we check whether some dataset is following a normal distribution? Say the conscription office records the size of 10.000 young men of age 18, and wants to make a test. For sure they are all found in the range of say 100 to 250 cm. What any person having been exposed to statistics would do is to try a *histogram*. Given the large number of probes one could make it rather fine, otherwise one might be satisfied with 10 or 20 bins. According to standard mathematical description we would think of the continuous probability distribution, with Gaussian kernel $g(t)$ and for a bin of form $I_n = [a_n, b_n)$ the height of the corresponding bar of the histogram would be just

$$h_n = \int_{a_n}^{b_n} g(t)dt.$$

But is it realistic that such a sharp classification is valid? Could we determine the size of a person within millimeters (or less)??

Consequently it is more realistic to take a certain uncertainty in the measuring process into account, we would say there is a narrow Gaussian (or whatever) distribution, because even measuring a person 10 times would give perhaps 10 different values...
So it is a more realistic model to replace the integral of $g(t)$ against the (sharp) indicator of $I = [a, b)$ by a smoothed version of the same, which - as is easily verified - turns the collection of indicator functions of the different bins into a BUPU, a *continuous partition of unity*, we write BUPU $\Psi = (\psi_n)_{n \in I}$.
It is an important property of measures (in fact even for bounded linear functionals on $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$) that the sum of the localized versions is the total sum (or just 100% in the case of a probability measure).

From here to the claimed model (inspired by STX, say) it is just one more step:
We do not only measure the local mass of a probability measure, respectively here the level of loudness, but we split the signal (which is discrete already) into frequency levels and determine the frequency content at each level. So we fix one such window (a continuous, finite length signal), such as a trapezoid function or Gaussian smoothed box-car function.
Then can try to request:
A signal is something which allows measurements, which are uniformly bounded over all the positions in time and frequency!

Taking this as a starting point we can apply functional analytic (combined with Fourier Analysis). In fact:

A Couple of Natural Questions

We are facing the following natural questions:

1. What is the influence of the choice of the window? Maybe one has for each different window another family of "signals"?
2. Is it necessary to know the boundedness for each offset in both time and frequency? Or is it enough to know it for a certain lattice (sampling in the time domain and the frequency domain). Could there be a blow-up "in between" those lattice points?
3. What about convergence of signals? When should we view signals close to each other? What does it mean that the CD allows to realize a perfect reconstruction of the concert in your home?

Basic Functional Analysis I

The following lemma helps to answer to the *first question*:

Lemma

Given a bounded subset $\Phi = (\phi_j)_{j \in J} \subset \boldsymbol{C}_b(\mathbb{R}^d)$ *there is a smallest Banach space* $\boldsymbol{B}_\Phi \hookrightarrow \big(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\big)$, *containing* $M \subset \boldsymbol{B}$ *is a bounded subset, namely*

$$\boldsymbol{B}_\Phi := \{f = \sum_{n=1}^{\infty} c_n \phi_n, (c_n) \in \ell^1(\mathbb{N})\}. \tag{1}$$

Each of these representations is called admissible with $\mathbf{c} = (c_n)_{n=1}^{\infty}$.

$$\|f\|_{\boldsymbol{B}_\Phi} := \inf_{adm.repr.off} \{\|\mathbf{c}\|_{\ell^1} = \sum_{n=1}^{\infty} |c_n|\} < \infty. \tag{2}$$

Basic Functional Analysis II

The particular choice $\Phi = \{M_s T_t g_0, t, s \in \mathbb{R}^d\}$ the resulting space is just Feichtinger's algebra $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ and mild distributions can be defined to be the elements of the dual of this space, which is characterized by the atomic representation with TF-shifted Gaussians.

However, Gabor Analysis tells us that there are many other ways to describe this space. We do not provide the technical details but state the facts. Most of them are derived using sometimes basic, sometimes deep results from Gabor Analysis.

Basic Functional Analysis III

- If one replaces the Gauss function $g_0(t) = \exp(-\pi|t|^2)$ by any other nonzero element from $\boldsymbol{S}_0(\mathbb{R}^d)$ one obtains the same space.
- For any $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ there exists some $\delta > 0$ such that for any δ-dense lattice Λ $\Phi = \{\pi(\lambda)g | \lambda \in \Lambda\}$ is a valid set of *atoms* generating the same space $\boldsymbol{S}_0(\mathbb{R}^d)$.
- $\mathcal{S}(\mathbb{R}^d) \hookrightarrow (\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0})$ as a dense subspace.

Convergence of Signals by Measurement I

It is plausible to call two signals (informally) "*very similar* of one has observed very little difference during a lot of measurements. In a more formal way one can characterize this type of convergence (let us consider for simplicity ordinary sequences, and not nets) as follows

Definition

A sequence $(\sigma_n)_{n\in\mathbb{N}}$ of mild distributions is mildly convergent with limite σ_0 if one has

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

There is also a corresponding notion of *mild Cauchy sequences*, which opens the way to define mild distributions as limits of *equivalence classes of mild Cauchy sequences* (ECMiS).

Convergence of Signals by Measurement II

Consequently norm convergence of $(\sigma_n)_{n\in\mathbb{N}}$ corresponds to *uniform convergence* of $V_{g_0}\sigma_n$ over $\mathbb{R}^d \times \widehat{\mathbb{R}}^d$. Again, due to the atomic decomposition property of $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ different non-zero atoms from $\boldsymbol{S}_0(\mathbb{R}^d)$ define equivalent norms.
Sometimes this concept of convergence is too strong. Among others the embedding $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right) \hookrightarrow \left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$ implies that any bounded measure defines an element of $\boldsymbol{S}_0'(\mathbb{R}^d)$, in particular the Dirac measures $\delta_x : f \mapsto f(x)$. But since $\|\delta_x - \delta_y\|_{\boldsymbol{S}_0'} = 2$ for $x \neq y$ one does not have norm convergence of δ_{x_n} to δ_0 if $x_n \neq 0$ but $x_n \to 0$ for $n \to \infty$.
But *mild convergence* (or w^*-convergence in $\boldsymbol{S}_0'(\mathbb{R}^d)$) corresponds to *uniform convergence over compact subsets* $K \subset \mathbb{R}^d \times \widehat{\mathbb{R}}^d$.

Mild convergence and w^*-convergence

It is easy to recognize that the described convergence is just the w^*-convergence in a dual Banach space, which is correctly described by nets (generalized sequences). However, since $(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0})$ is a *separable* Banach space, it is enough to work with sequences. Consequently the pointwise convergence of functionals (meaning application to any given $f \in \boldsymbol{S}_0(\mathbb{R}^d)$) implies norm boundedness of the sequence $(\sigma_n)_{n\in\mathbb{N}}$ in $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$. Recall that the dual norm on $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$ is just

$$\|\sigma\|_{\boldsymbol{S}_0'(\mathbb{R}^d)} := \sup_{\|f\|_{\boldsymbol{S}_0} \leq 1} |\sigma(f)|.$$

As a matter of fact we have

$$\|\sigma\|_{\boldsymbol{S}_0'(\mathbb{R}^d)} = \sup_{(t,s)\in\mathbb{R}^d\times\widehat{\mathbb{R}}^d} |V_{g_0}\sigma(t,s)|.$$

Basic Operations on Mild Distributions

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

Regularizations I

There are many ways to regularize a given mild distributions, such as a *Dirac comb* $\sqcup\!\sqcup = \sum_{k \in \mathbb{Z}^d} \delta_k$, which shows neither decay nor smoothness (decay in the frequency direction).
The typical methods to approximate (in the "mild sense") a given $\sigma \in \boldsymbol{S}_0'$ is to convolve it with a Dirac sequence, normalized in $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ but sitting inside of $\boldsymbol{S}_0(\mathbb{R}^d)$, hence its FT (a summability kernel) is a bounded approximate unit in the pointwise sense. Thus it is good to know that $\boldsymbol{S}_0(\mathbb{R}^d) * \boldsymbol{S}_0'(\mathbb{R}^d) \subset \boldsymbol{C}_b(\mathbb{R}^d)$ with corresponding norm estimates, but more importantly

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

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

Regularizations II

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

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

$$\sigma \mapsto V_{g_0}\sigma|_\Lambda \in \ell^\infty(\Lambda)$$

and

$$\sigma = \sum_{\lambda \in \Lambda} V_{g_0}\sigma(\lambda)\pi(\lambda)\widetilde{g},$$

is unconditionally mildly convergent.

Regularizations III

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

$$\sigma_F = \sigma_{f,g,\Lambda,F} := \sum_{\lambda \in F} V_{g_0}\sigma(\lambda)\pi(\lambda)\widetilde{g}$$

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

$$|V_{g_0}\sigma(q) - V_g\sigma_F(q)| \leq \varepsilon, \quad \forall q \in Q.$$

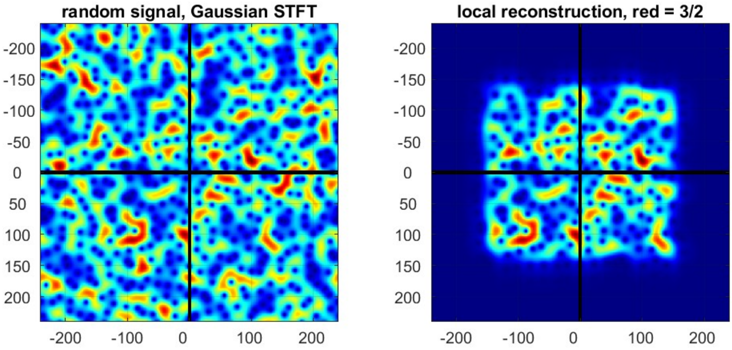


Figure: GabLocRec01.jpg

Mild Distributions and Operators I

While Hilbert spaces such as $\ell^2(\mathbb{N})$ are quite popular, the operators on Hilbert spaces are not so nice, although they form a C^*-algebra. Given an infinite matrix $\mathbf{A}$ it is not so easy to verify whether it defines a bounded operator on $\ell^2(\mathbb{N})$ (resp. whether the action on the finite sequences extends to all of ℓ^2). But most proofs of boundedness (say of pseudo-differential operators, or Gabor multipliers) start from the $\mathbf{L}^2$-case and then go on to discuss refined estimates on other Banach spaces.

Look at Fourier multipliers on $\left(\mathbf{L}^p(\mathbb{R}^d), \|\cdot\|_p\right)$. Any such Fourier multiplier $f \mapsto \mathcal{F}^{-1}(m \cdot \widehat{f})$ is also bounded on the dual space $\left(\mathbf{L}^q(\mathbb{R}^d), \|\cdot\|_q\right)$ and thus - by complex interpolation - also on $\left(\mathbf{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$). Since the pointwise multiplier space of $\left(\mathbf{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$ is just $\left(\mathbf{L}^\infty(\mathbb{R}^d), \|\cdot\|_\infty\right)$ this approach is OK.

Kernel Theorem I

For more general situation the **Kernel Theorem** is very useful:

Theorem

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

Formally sometimes one writes by "abuse of language"

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

defining the action of the functional $Kf \in \mathbf{S}_0'(\mathbb{R}^d)$ via

$$Kf(g) = \int_{\mathbb{R}^d} Kf(x)g(x)dx = \int_{\mathbb{R}^{2d}} k(x,y)g(x)f(y)dxdy.$$

Kernel Theorem II

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

Again the complete picture can again be best expressed by a unitary Gelfand triple isomorphism.

On the one hand we have kernels $k(x,y)$ in $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^{2d})$, or also spreading functions $\eta(K)$ in $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d \times \widehat{\mathbb{R}}^d)$ (also KNS-symbols), and on the other hand the operator BGTr $(\mathcal{N}_{w*}(\boldsymbol{S}_0', \boldsymbol{S}_0), \mathcal{HS}, \mathcal{L}(\boldsymbol{S}_0, \boldsymbol{S}_0'))$.

Kernel Theorem III

We first describe the innermost shell:

Theorem

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

Kernel Theorem IV

Remark: Note that for *regularizing* kernels in $\boldsymbol{S}_0(\mathbb{R}^{2d})$ the expected identification hold true:

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

$\delta_y \in \boldsymbol{S}_0'(\mathbb{R}^d)$ implies that $K(\delta_y) \in \boldsymbol{S}_0(\mathbb{R}^d)$ by the regularizing properties of K, hence pointwise evaluation makes sense.
With this understanding our claim is that the kernel theorem provides a (unitary) isomorphism between the Gelfand triple (of kernels) $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^{2d})$ into the Gelfand triple of operator spaces $(\mathcal{N}_{w*}(\boldsymbol{S}_0', \boldsymbol{S}_0), \mathcal{HS}, \mathcal{L}(\boldsymbol{S}_0, \boldsymbol{S}_0'))$. At the Hilbert-Schmidt $(\mathcal{HS}\langle \cdot, \rangle_{\mathcal{HS}})$ is given by

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

Distributional kernels are unavoidable

It sounds a bit strange that one has to resort to 'generalized functions' if one is just trying to describe harmless operators on a simple Hilbert space such as $\left(\boldsymbol{L}^2(\mathbb{R}), \|\cdot\|_2\right)$. But let us consider a simple multiplication operator $f \mapsto h \cdot f$, for some $h \in \boldsymbol{C}_b(\mathbb{R})$. By analogy with the discrete (= matrix) case one would expect that the "kernel" of such an operator is concentrated on the diagonal. But in the continuous case the diagonal $\{(x,x) \,|\, x \in \mathbb{R}^d\}$ is a set of measure zero in $\mathbb{R}^{2d}$! and thus it is impossible to find a kernel $k(x,y)$ which is even locally integrable, leave alone in $\boldsymbol{L}^2(\mathbb{R}^{2d})$.

Papers by Jens Fischer I

J. V. Fischer.
On the duality of regular and local functions.
Mathematics, 5(41), 2017.

J. V. Fischer.
There is only one Fourier Transform.
ResearchGate, doi 10.13140/rg. 2.2, 30950:335, 2017.

J. V. Fischer.
Four particular cases of the Fourier transform.
Mathematics, 12(6):335, 2018.

J. V. Fischer and R. L. Stens.
On the reversibility of discretization.
Mathematics, 8(4):619, 2020.

The Gabor Frame Operator I

It is clear from the introduction that $\boldsymbol{S}_0'(\mathbb{R}^d)$ is quite convenient for the analysis of operators in Gabor Analysis. Let us start with the question, whether a given function $g \in \boldsymbol{L}^2(\mathbb{R}^d)$ defines a Gabor frame with respect to a lattice $\Lambda \lhd \mathbb{R}^d \times \widehat{\mathbb{R}}^d$.
Then even the boundedness on $\boldsymbol{L}^2(\mathbb{R}^d)$ may be a problem (one has to require that the family $(\pi(\lambda)g)_{\lambda\in\Lambda}$ is a *Bessel family*. But even better we can observe that for any $g \in \boldsymbol{L}^2(\mathbb{R}^d)$ the mapping $f \mapsto V_g f|_\Lambda$ defines a bounded linear operator from $(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0})$ into $\ell^2(\Lambda)$, and the synthesis mapping $(c_\lambda)_{\lambda\in\Lambda} \mapsto \sum_{lainLa} c_\lambda \pi(\lambda) g$ defines a bounded linear mapping from $\ell^2(\Lambda)$ into $\boldsymbol{S}_0'(\mathbb{R}^d)$. Thus $S = S_{g,\Lambda} : f \mapsto \sum_{\lambda\in\Lambda} V_g f(\lambda)\pi(\lambda) g$ defines a bounded linear operator from $(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0})$ to $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$. Any such operator has a *spreading representation* $\eta(S) \in \boldsymbol{S}_0'(\mathbb{R}^d \times \widehat{\mathbb{R}}^d)$.

The Gabor Frame Operator II

It is important that $\boldsymbol{S}_0'(\mathbb{R}^{2d})$ is much smaller than $\boldsymbol{\mathcal{S}}'(\mathbb{R}^{2d})$. For tempered distributions supported by a discrete set one may have not only Dirac measures, but also (partial) derivatives of any order! It is one of the important properties of $S_{g,\Lambda}$ that such operators commute with TF-shifts from Λ. For the kernel of $S_{g,\Lambda}$ this means certain invariance properties, best described by Λ periodicity of its Kohn-Nirenberg symbol $\kappa(S)$, also in $\boldsymbol{S}_0'(\mathbb{R}^{2d})$, which then translate (via the symplectic Fourier transform) into support conditions of $\eta(S)$.

The important conclusion is then $\eta(S)$ is a weighted Dirac comb, and in fact (up to normalizations) the weights are just the samples of the STFT $V_g g$, taken at the adjoint lattice Λ° (commutator of Λ via TF-shifts). Note that for $\Lambda = a\mathbb{Z}^d \times b\mathbb{Z}^d$ the adjoint lattice is given by $\Lambda^\circ = \frac{1}{b}\mathbb{Z}^d \times \frac{1}{a}\mathbb{Z}^d$.

Generalized Stochastic Processes I

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

Thus it is natural to ask what a "stochastic signal" could be.

Recall that a (continuous) function assigns to each point t a value $f(t)$. A *generalized function* σ assigns to each test function k (average of point masses) some value $\sigma(k)$. A *stochastic process* $\rho(t)$ assigns to each value t a probability measure, ideally a member of some abstract Hilbert space $\mathcal{H}$ (defined in a probabilistic manner). So a *generalized stochastic signal* (**GSP**) could be

a bounded linear operator $\rho : \left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right) \to (\mathcal{H}, \|\cdot\|_{\mathcal{H}})$.

Generalized Stochastic Processes II

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

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

For every GSP we have a autocorrelation, another mild distribution $\sigma_\rho \in \boldsymbol{S}_0'(\mathbb{R}^{2d})$ with

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

As expected the autocorrelation of $\widehat{\rho}$ is just $\mathcal{F}(\sigma_\rho)$.
Wide sense stationarity means:

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

Or in the case of generalized stochastic processes

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

THANKS you for your attention

Maybe you visit www.nuhag.eu
and checkout other talks and related material. hgfei

HGFei publications:
https://nuhagphp.univie.ac.at/home/feipub$_d$*b.php*
www.nuhag.eu/talks : many talks on related subjects
www.nuhag.eu/ETH20 :
NOTE: Talks require access code:
"visitor" and "nuhagtalks"
The spectrograms shown have been produced with the help of STX (from ARI Vienna).