NuHAG Numerical Harmonic Analysis Group

Convolutions, Fourier Transforms and Rigged Hilbert Spaces

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

University PECS, Seminar Talk
Sept. 27th, 2021

NEW ABSTRACT, provided 13.09.2021 I

The talk will indicate that there are many different motivations to introduce convolution, but also the Fourier transform, as a unitary transformation which diagonalizes convolutions (or: the Convolution Theorem demonstrates that the FT turns convolution into ordinary pointwise multiplication).

The theory of Rigged Hilbert Spaces, resp. the specific Banach Gelfand Triple, based on the Segal algebra S_0 comes in handy in order to describe the properties of the Fourier transform and to make the analogy between the discrete setting (this is based on the DFT, the discrete FT, which is implemented as the FFT, the Fast Fourier Transform).

As it will be explained, there are several new and simplified approaches to this Banach space of test functions and its dual, which is meanwhile known as the space of "mild distributions". It can be characterized as a kind of completion of the space of test functions, but also as a subspace of tempered distributions with bounded spectrogram.

Four periods

Four Periods of HGFei

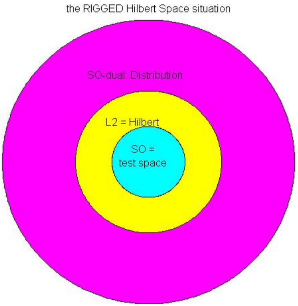
- 1969 - 1989: **Abstract Harmonic Analysis** (AHA);
- 1989 - 2015: Application Oriented Mathematics, using MATLAB, creating NuHAG;
- 2000 - now: Editor to JFAA: Journal of Fourier Analysis and Applications
- 2006 - now: **Conceptual Harmonic Analysis**: An attempt to recombine abstract and computational harmonic analysis, to reconciliate the mathematical and the engineering resp. physicists approach to Fourier Analysis.
- starting 2015: attempt to popularize the concepts of Banach frames, Banach Gelfand Triples, mild distributions.

Overview

- Recall different approaches to convolution
- Mathematicians vs. Engineers/Physicists
- Rigged Hilbert Spaces/Distributions
- Arguments/Techniques
- The Banach Gelfand Triple $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$.

The Banach Gelfand Triple

When we talk about ordinary MULTIPLICATION we refer to the chain of inclusions $\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$. Here we use



Multiplication

Let us compute in a slightly unconventional way:

4 5 6 * 1 3

. 12 15 18 which gives

4 17 21 18 ,

or by collecting the terms:

>> 4 thsd 12 hundr. 15 ten 18 unit = **5 9 2 8**

but also (flipping the order!)

6 5 4 * 3 1 = **18 21 17 4** gives **2 0 2 7 4**

Compare to $(4x^2 + 5x + 6) \cdot (1x + 3)$

respectively $(6x^2 + 5x + 4) \cdot (3x + 1)$,

noting that we have (for example)

$$456 = p(10) \quad \text{for} \quad p(x) = 4x^2 + 5x + 6.$$

Polynomials and DFT

There is also a nice connection to **elementary probability theory**: The polynomial $w(x) = (x + x^2 + x^3 + x^4 + x^5 + x^6)/6$ encodes exactly the information which should be known about a (fair) dice. Each possible value k which may appear as a possible outcome has the probability $1/6$, while these are exactly the coefficients of $w(x)$ at x^k, for $k = 1, ..., 6$!

The probability for a sum of *two independent dices* is encoded in the polynomial $w(x)^2$, which shows a linear increase of the coefficients, from $k = 2$ up to $k = 7$ and then it goes down up to $k = 12$.

The probability of obtaining a total score of 5 is just the sum of the probabilities of getting the pairs

$$(1,4), (2,3), (3,2) \quad \text{and} \quad (4,1),$$

which totals $4/36 = 1/9$, for example.

The Discrete Fourier Transform I

Obviously the probability distribution for 20 dices will give non-zero values for $k = 20$ up to $k = 120$, encoded in $w(x)^{20}$!
But how can we do such a computation efficiently?
Use the DFT (Discrete Fourier Transform), in the realization as FFT (Fast Fourier Transform), which can be interpreted as the transition from a sequence, viewed as a sequence of coefficients of a polynomial to the values of that polynomial *at the unit roots of order* N!
Correspondingly the IFFT routine recovers these coefficients from the values of the polynomial at exactly these points.
It is also clear that the values of $p \cdot q(x) := p(x)q(x)$ at the unit roots are just taken pointwise, while the ordinary multipication of polynomials requires a lot of shuffling (in order to get the Cauchy-Product, at the coefficient level).

The Discrete Fourier Transform II

In order to realize this one has to start with a sequence of length $128 = 2^7$ (for such a length the FFT is particularly fast!) and put the values $1/6$ into this sequence at positions $2, 3, 4, 5, 6, 7$, noting that the *first position* for the description of a polynomial is occupied by the *constant term*.
The length is choosen ins such a way that we can be sure that $w(x)^{20}$ can be recovered, as this is a polynomial of order $6 \cdot 20 = 120 < 128$ (the next power of two).
So we have a two-line code:

```
a = zeros(1,128); a(2:7) = 1/6;
b = real(ifft(fft(a).^20));
```

Side Note on Multiplication of Integers I

The multiplication (hence powers) of long integers can be obtained in this way (even with thousands of digits) by observing that any integer is just the value of a particular polynomial (given by the digits of that number in the *decimal system* at $x = 10$. So we have for example (in MATLAB)

$$123 == \text{polyval}([1,2,3],10);$$

Note that further extension of this process, called multiplication to real numbers is obtained by approximation, and then even to $\mathbb{C}$, turning this set of *complex numbers* into a highly useful *field*, with quite similar computation rules, having a similar formalism, even if $1/\pi + i * 3\sqrt{5}$ looks like a strange object (for computation).

Why convolution?

While mathematical books describe a lot how convolution is defined and what kind of properties one has, often based on Lebesgue integration theory, they do NOT ask why convolution should be studied! In fact, it is often said, that the "complicated" convolution is turned into pointwise multiplication via the Fourier Transform which makes the Fourier transform so important [Convolution Theorem]:

$$\mathcal{F}(f \star g) = \mathcal{F}(f) \cdot \mathcal{F}(g), \quad f, g \in \boldsymbol{L}^1(\mathbb{R}^d).$$

In the course of Brad Osgood in Stanford (also on YouTube) he introduces the Fourier transform first and computes the function h which satisfies $\mathcal{F}(h) = \mathcal{F}(f) \cdot \mathcal{F}(g)$,
coming up with convolution in this way!

Vague and even mysterious explanations for TILS

In the engineering books the need of convolution and Fourier transforms is justified by the study of TILS:

Translation Invariant Linear Systems

There are different ways to explain the behaviour of TILS, From a mathematical point of view the goal must always be to "represent" the system as a convolution operator, where the "convolution kernel" is some function, or at least a kind of distribution. In the theory of multipliers in the 1960s alongside with bounded measures the terms *pseudo-measures* and *quasi-measures* appeared. The best way is to make use not only of measures, but of distributions!

But it turns out that *mild distributions* suffice!

TILS: Lüke/Ohm

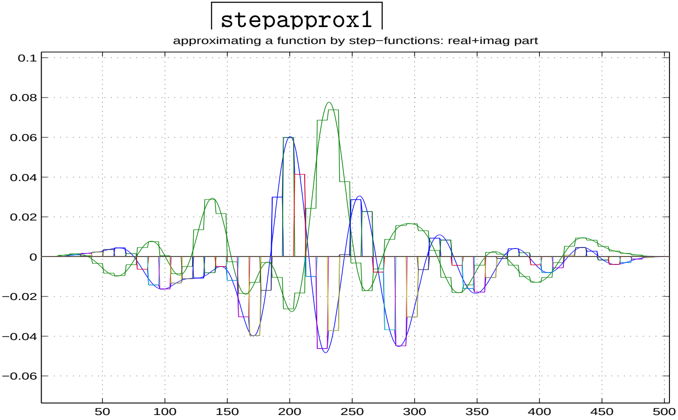


Abbildung: A typical illustration of an approximation to the input signal of a TILS, named T, preparing for the use of the Dirac impulse.

Questions arising from these pictures

- In which sense does the limit of the rectangular functions exist?
- What kind of argument is given for the transition to integrals? Do we collect (as I learned in the physics course) uncountably many infinitely small terms in order to get the integral?
- In which sense are these step function convergent to the input signal f, and how are the steps determined?
- What has to be assumed about the boundedness properties of the operator T? In other words, which kind of convergence of signals in the domain will guarantee corresponding (or different) convergence in the target domain?

Starting from systems

Let us first look at BIBOS systems, i.e. at systems which convert bounded input in bounded output. If one wants to avoid problems with integration technology and sets of measure zero it is reasonable to assume that the operator has $\left(C_b(\mathbb{R}^d), \|\cdot\|_\infty\right)$ as a domain and as target space.

Recall the **scandal in systems theory** observed by I. W. Sandberg.

- I. W. Sandberg. A note on the convolution scandal. *Signal Processing Letters, IEEE*, 8(7) (2001) p.210–211.
- I. W. Sandberg The superposition scandal. Circuits Syst. Signal Process., 17/6, (1998) p.733-735.

Formulas in engineering books

In one of the books on Fourier Analysis for engineers, at an established US university, I found the following key arguments for the proof of the Fourier inversion theorem:

$$\int_{-\infty}^{\infty} e^{2\pi isx} ds = \delta(x) \quad \text{meaning?} \quad \mathcal{F}^{-1}\mathbf{1} = \delta_0?$$

combined with the so-called *sifting property of the Delta Dirac*

$$f(x) = \int_{-\infty}^{\infty} \delta(x-y) f(y) dy \quad \text{meaning?} \quad f = \delta_0 \star f.$$

If we write these relations in the mathematical style they are of course meaningful (in the distributional setting), which is not based on the fact that the integrals exist (as Lebesgue integrals, but expressing dualities).

TILS: Translation invariant linear systems

A translation invariant system is a linear operator which commutes with translation. In typical situations, i.e. when translation is isometric and continuous on the domain, meaning

$$\lim_{x \to 0} \| T_x f - f \|_{\boldsymbol{B}} \to 0 \quad \forall f \in \boldsymbol{B}.$$

such operators, if they are bounded, satisfy

$$T(g * f) = g * f, \quad \forall f \in \boldsymbol{B}.$$

Then they typical argument is

$$T(f) = T(\delta_0 * f) = T(\delta_0) * f = h * f$$

for the *impulse response* $h := T(\delta_0)$. But does δ_0 belong to the domain of T? And if this is the case, can one say that

$$\widehat{T(f)} = \widehat{h} \cdot \widehat{f}.$$

Taking a closer look

Taking at the closer look and assuming that T is defined on $(\boldsymbol{L}^1(\mathbb{R}), \|\cdot\|_1)$ or $(\boldsymbol{L}^2(\mathbb{R}), \|\cdot\|_2)$, so that δ_0 has to be seen as a (generalized) limit (not in norm) of compressed version of a given function in $\boldsymbol{L}^1(\mathbb{R})$ by dilation, e.g.

$$g_n(x) = 2^{-n} g(x/2^n), n \to \infty,$$

the claim is actually (with various limits involved):

$$T(f) = T(\lim_{n\to\infty} g_n * f) = \lim_{n\to\infty} T(g_n * f) =$$

$$= [\lim_{n\to\infty} T(g_n)] * f = T(\delta_0) * f.$$

No help from the side of mathematics

One might hope that mathematical books on Fourier analysis cover all these topics in a more appropriate fashion and explain all the technical details in order to show that the claims are correct, even if the explanation is just a heuristic one.
BUT CURRENT books in the field cover very little of these problems, but emphasize completely different aspects of Fourier analysis.
One may find details about the famous result of L. Carleson about the almost everywhere convergence of a Fourier series for any periodic function in $\boldsymbol{L}^2$, even if this is not of great *practical importance*. Moreover, very often one may not find a detailed proof of **Shannon's sampling theorem**, which obviously is very important for the understanding of signal processing (of band-limited functions, via regular samples, taken at Nyquist rate).

Questions arising from these pictures

- In which sense does the limit of the rectangular functions exist?
- What kind of argument is given for the transition to integrals? Do we collect (as I learned in the physics course) uncountably many infinitely small terms in order to get the integral?
- In which sense are these step function convergent to the input signal f, and how are the steps determined?
- What has to be assumed about the boundedness properties of the operator T? In other words, which kind of convergence of signals in the domain will guarantee corresponding (or different) convergence in the target domain?

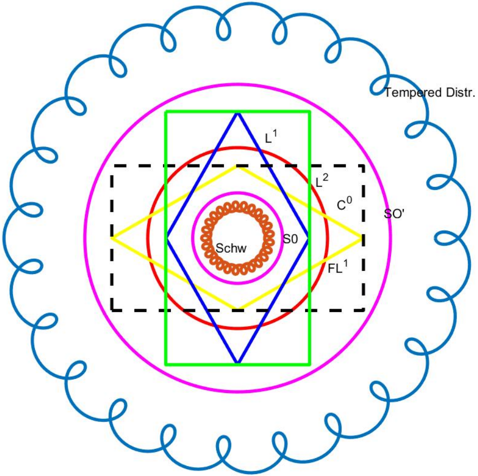
Starting from systems

Let us first look at BIBOS systems, i.e. at systems which convert bounded input in bounded output. If one wants to avoid problems with integration technology and sets of measure zero it is reasonable to assume that the operator has $\left(\boldsymbol{C_b}(\mathbb{R}^d), \|\cdot\|_\infty\right)$ as a domain and as target space.

Recall the **scandal in systems theory** observed by I. W. Sandberg.

- I. W. Sandberg. A note on the convolution scandal. *Signal Processing Letters, IEEE*, 8(7) (2001) p.210–211.
- I. W. Sandberg The superposition scandal. Circuits Syst. Signal Process., 17/6, (1998) p.733-735.

A Zoo of Banach Spaces for Fourier Analysis



Mathematicians approach I

We are used to learn that convolution is first well defined (even just using the Riemann integral, for $f, g \in \boldsymbol{C}_c(\mathbb{R}^d)$ (continuous complex functions with compact support):

$$[f \star g](x) = \int_{\mathbb{R}^d} f(x-y)g(y)dy, \quad x, y \in \mathbb{R}^d \qquad (1)$$

con

and then shows that

$$\boldsymbol{C}_c(\mathbb{R}^d) \star \boldsymbol{C}_c(\mathbb{R}^d) \subset \boldsymbol{C}_c(\mathbb{R}^d) \quad \text{with} \quad combinedwiththeestimate \|f \star g\|_1 \leq$$

where

$$\|f\|_1 := \int_{\mathbb{R}^d} |f(x)|dx, \quad f \in \boldsymbol{C}_c(\mathbb{R}^d).$$

Mathematicians approach II

This suggests to view $(C_c(\mathbb{R}^d), \|\cdot\|_1)$ as a normed algebra which allows to introduce $(L^1(\mathbb{R}^d), \|\cdot\|_1)$ as the (abstract) completion of this normed algebra. Consequently we see that $(L^1(\mathbb{R}^d), \|\cdot\|_1)$ is a Banach algebra with respect to convolution, which is obtained by extension of the "ordinary pointwise convolution" described above.

However, using *Lebesgue integration theory* (see H. Reiter's book) one can use *Fubini's Theorem* to establish the almost everywhere existence of the integral (1) and the norm estimate for the convolution product can be well established.

However, it is overall mis-leading to define convolution, or to claim that the "existence of convolution" could depend on the a.e. realization of the convolution integral!

Via bounded measures I

While the *abstract completion* via *equivalence classes of Cauchy sequences* is providing a (up to isomorphism unique) Banach algebra, which suffices for many proofs concerning $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$.

Alternatively *Lebesgue integration theory* allows to provide a "model case" for $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$, in the usual way, with the integral well defined for *equivalence classes of measurable functions* with the extended interpretation of the norm $\|\cdot\|_1$.

The functional analytic view-point starts from $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$, defined as the dual of $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$, and starts from the isometric embedding of $(\boldsymbol{C}_c(\mathbb{R}^d), \|\cdot\|_1)$ into $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$:

$$\mu_k : f \mapsto \int_{\mathbb{R}^d} f(x)\, k(x) dx \quad \text{with} \quad \|k\|_1 = \|\mu_k\|_{\boldsymbol{M}_b}.$$

TILS via Convolution I

But why are *engineers* interested in this mathematical operations: because it allows them to understand better time-invariant linear systems, which - mathematically speaking are *linear operators* T which commute with translations, i.e. which satisfy

$$T \circ T_z = T_z \circ T, \quad \forall z \in \mathbb{R}^d. \tag{2}$$

tra

But for a mathematician this is not a well-defined description, as it fails to provide the information of the domain (and target space)! So there is some freedom. So let us think of some translation invariant Banach space $(\boldsymbol{X}, \|\cdot\|_{\boldsymbol{X}})$, such that T is (in addition to (2)) also bounded on this space. The most natural choice might be (for engineers) the space $\left(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\right)$, and for mathematicians perhaps $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ or also $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$.

TILS via Convolution II

For the two last mentioned cases one has a full characterization (to be given in a moment), while the first case is related to what I. Sandberg has called the *scandal in system theory*.
For $(\boldsymbol{X}, \|\cdot\|_{\boldsymbol{X}}) = \left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ one has Wendel's Theorem, characterizing $\boldsymbol{H}_G(\boldsymbol{L}^1, \boldsymbol{L}^1)$ (isometrically) as the space of convolution operators using uniquely determined bounded measures $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$. In fact, Looking at the behaviour of $T(e_\alpha)$, where $(e_\alpha)_{\alpha\in I}$ is some *Dirac sequence* we observe that this family is bounded in $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$, hence in $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$, ans hence there is a subnet which is w^*-convergent in $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$. The limit, let us call it μ_0 turns out the be the correct measure giving $T(f) = \mu \star f$, given by the pointwise relation

$$T(f)(x) = \mu(T_x), \quad f \in \boldsymbol{C}_c(\mathbb{R}^d).$$

A few references I

[la71] R. Larsen.
An Introduction to the Theory of Multipliers.
Springer-Verlag, New York-Heidelberg, 1971.

[un20] M. Unser.
A note on BIBO stability.
IEEE Trans. Sign. Proc., 68:5904–5913, 2020.

[sa98] I. W. Sandberg.
The superposition scandal.
Circuits Syst. Signal Process., 17(6):733–735, 1998.

[fe17] H. G. Feichtinger.
A novel mathematical approach to the theory of translation invariant linear systems.
In I. Pesenson, Q. Le Gia, A. Mayeli, H. Mhaskar, and D. Zhou, editors, *Recent Applications of Harmonic Analysis to Function Spaces, Differential Equations, and Data Science.*, Applied and Numerical Harmonic Analysis, pages 483–516. Birkhäuser, Cham, 2017.

Impulse Response and Transfer Function I

The above mathematical observation is *intuitively supported* by the reported observation that $T(e_\alpha)$ is taking a limit, which (taken the justification of the following term as granted) is called the *impulse response* of the system T, since

$$T(\delta_0) = T(\lim_{\alpha\to\infty} e_\alpha) = \lim_{\alpha\to\infty}(T(e_\alpha)) = \mu_0! \qquad (3)$$

imp

As it turns out, this formula can be justified under suitable conditions, but not by choosing $(\boldsymbol{X}, \|\cdot\|_{\boldsymbol{X}}) = \big(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\big)$ (as has been observed by I. Sandberg). BUT the main reason is the fact that this space is NOT separable, and that an abstract operator my vanish on $\boldsymbol{C}_0(\mathbb{R}^d)$ but still be non-trivial (a so-called Banach limit).
The situation is quite different for $(\boldsymbol{X}, \|\cdot\|_{\boldsymbol{X}}) = \big(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\big)$.
Here we can make use of Plancherel's Theorem.

Impulse Response and Transfer Function II

ult

Theorem

Any bounded linear operator $T : (\mathbf{L}^2(\mathbb{R}^d), \|\cdot\|_2) \to (\mathbf{L}^2(\mathbb{R}^d), \|\cdot\|_2)$ *can be characterized via a pointwise multiplication on the Fourier transform side, by some (uniquely determined)* $h \in (\mathbf{L}^\infty(\mathbb{R}^d), \|\cdot\|_\infty)$, *namely via*

$$\mathcal{F}(T(f)) = \widehat{f} \cdot h, \quad f \in \mathbf{L}^2(\mathbb{R}^d). \tag{4}$$

Pla

In fact, this identification is isometric, i.e. satisfies

$$\|h\|_\infty = \|T\|_{\mathbf{L}^2(\mathbb{R}^d)} .$$

This pointwise multiplier on the Fourier transform is called *transfer function* of the system T. For the case $T(f) = \mu \star f$ with $\mu \in \mathbf{M}_b(\mathbb{R}^d)$ we have of course $h = \widehat{\mu}$.

Conflict of Concepts I

There is a clear conflict of these concepts, if one realizes that the (distributional) Fourier transform of the *chirp function* $h(s) = exp(2\pi is^2) \in \boldsymbol{C}_b(\mathbb{R}^d)$ is just the function itself, i.e. $\hat{h} = h$. Hence it is obvious that pointwise multiplication by h (on the Fourier transform side) is a valid transfer function describing a decent TILS on $\boldsymbol{L}^2(\mathbb{R}^d)$.
But $\mathcal{F}^{-1}(\mathbf{1}_{[-1/2,1/2]}) = \text{SINC} \in \boldsymbol{L}^2(\mathbb{R}^d) \setminus \boldsymbol{L}^1(\mathbb{R}^d)$ does not allow to do the convolution product (using the Lebesgue integral) AT ANY POINT $x \in \mathbb{R}^d$!

As a matter of fact (cf. my ETH notes) one can and should question the usual heuristic derivation of the *"representation theorem"* for TILSs as given in many engineering books. But we skip this here.

Modelling via $\left(C_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$ I

1. We start with the pointwise Banach algebra $\left(C_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$, of continuous, complex-valued (hence bounded) functions vanishing at infinity, endowed with the sup-norm (written as $\|f\|_\infty$ for $f \in C_b(\mathbb{R}^d)$). It contains $C_c(\mathbb{R}^d)$ as a dense subspace, in fact it coincides with the closure of $C_c(\mathbb{R}^d)$ in $\left(C_b(\mathbb{R}^d), \|\cdot\|_\infty\right)$!
2. By (our, justified) definition the bounded linear functionals are called bounded measures, and we use the symbol $(M_b(\mathbb{R}^d), \|\cdot\|_{M_b})$ for $\left(C_0'(\mathbb{R}^d), \|\cdot\|_{C_0'}\right)$.

Modelling via $(C_0(\mathbb{R}^d), \|\cdot\|_\infty)$ II

3. Using the simple fact that $\|\delta_x\|_{M_b} = 1$ one observes that the closed linear span of the *Dirac functionals* $\delta_x(f) := f(x)$, for $f \in C_0(\mathbb{R}^d)$, are the discrete [bounded] measures, can be characterized as absolutely convergent series of the form $\nu = \sum_{k=1}^\infty c_k \delta_{x_k}$, with $\sum_{k=1}^\infty |c_k| < \infty$. We use the symbol $M_d(\mathbb{R}^d)$ for this closed subspace of $(M_b(\mathbb{R}^d), \|\cdot\|_{M_b})$.

Note that translations acts in a natural way on e.g. $C_b(\mathbb{R}^d)$, given by $T_x(f)(z) = T(z-x)$. The fact, these translations form a commutative group of isometric operators, i.e. satisfy

$$\|T_x f\|_\infty = \|f\|_\infty, \quad f \in C_0(\mathbb{R}^d), x \in \mathbb{R}^d, \tag{5}$$

COt

An object of great interest is the subalgebra of all bounded linear operators which commute with the shift operators! We use the symbol $\mathcal{H}_G(C_0(G))$.

Modelling via $(C_0(\mathbb{R}^d), \|\cdot\|_\infty)$ III

v00

Theorem

There is an (natural) isometric identification between translation invariant, linear systems on $(C_0(\mathbb{R}^d), \|\cdot\|_\infty)$ *(i.e. bounded linear mappings on* $(C_0(\mathbb{R}^d), \|\cdot\|_\infty)$ *commuting with translations, and the space* $(M_b(\mathbb{R}^d), \|\cdot\|_{M_b})$.
In fact, every such operator $T \in$ *is a convolution operator by a uniquely determined bounded measure* μ, *we write* $C_\mu(f)$ *or (later)* $\mu * f$, *for* $f \in C_0(\mathbb{R}^d)$ *and* $\mu \in M_b(\mathbb{R}^d)$.

The "non-trivial" part is of course to show that $C_\mu(f)$ is not only bounded and (in fact uniformly) continuous, but also still tending to zero at infinity.

The key-players for time-frequency analysis

Time-shifts and Frequency shifts

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

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

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

Behavior under Fourier transform

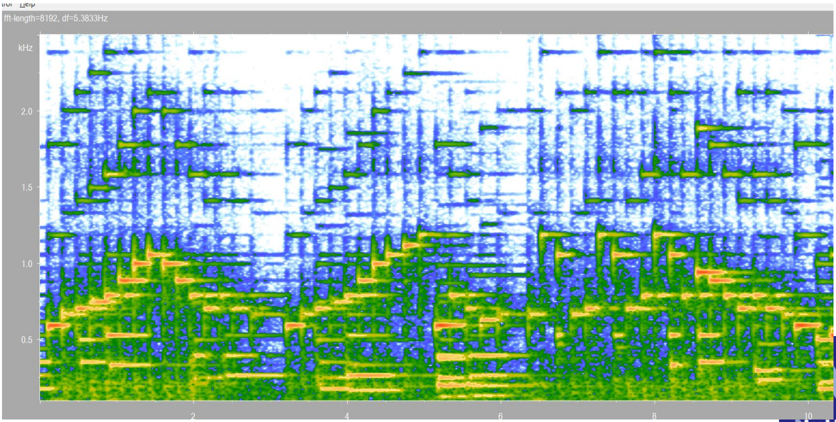
$$(T_x f)\widehat{\ } = M_{-x} \hat{f} \qquad (M_\omega f)\widehat{\ } = T_\omega \hat{f}$$

The Short-Time Fourier Transform

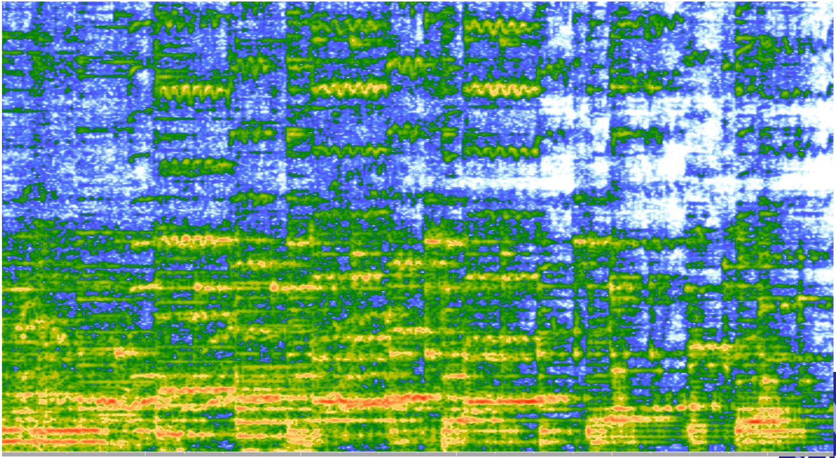
$$V_g f(\lambda) = \langle f, M_\omega T_t g \rangle = \langle f, \pi(\lambda) g \rangle = \langle f, g_\lambda \rangle, \;\; \lambda = (t, \omega);$$

A Typical Musical STFT

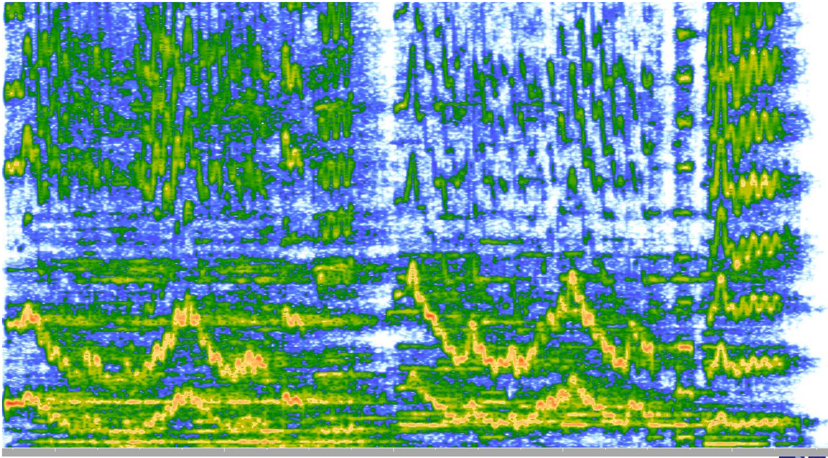
A typical piano spectrogram (Mozart), from recording



A Musical STFT: Brahms, Cello

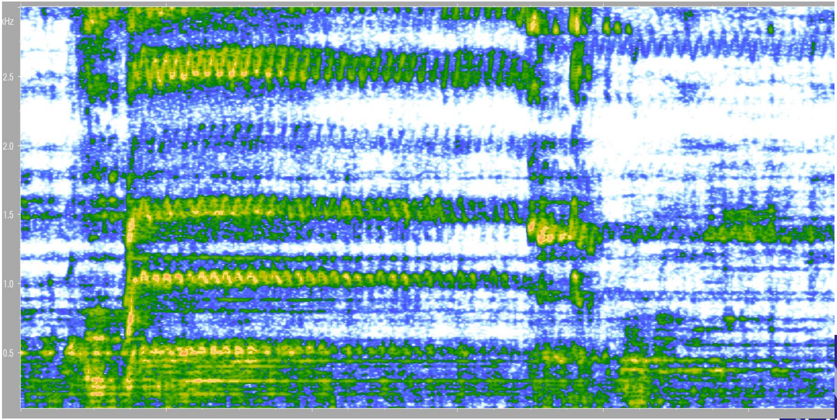


A Musical STFT: Maria Callas



A Musical STFT: Tenor: VINCERA!

Obtained via STX Software from ARI (Austrian Acad. Sci.)



A Banach Space of Test Functions (Fei 1979)

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

$$\|f\|_{\mathbf{S}_0} := \|V_g f\|_{\mathbf{L}^1} = \iint_{\mathbb{R}^d \times \widehat{\mathbb{R}}^d} |V_g f(x,\omega)| dx d\omega < \infty.$$

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

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

cSo

Lemma

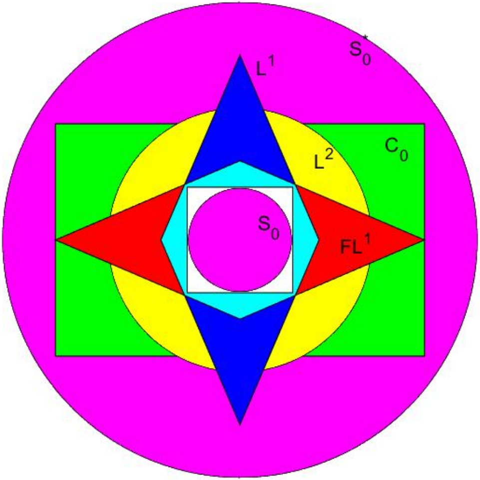
Let $f \in S_0(\mathbb{R}^d)$, then the following holds:

(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* $\|\hat{f}\|_{S_0} = \|f\|_{S_0}$.

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).

Various Function Spaces



BANACH GELFAND TRIPLES: a new category

Definition

A triple, consisting of a Banach space $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$, which is densely embedded into some Hilbert space $\mathcal{H}$, which in turn is contained in $\boldsymbol{B}'$ is called a Banach Gelfand triple.

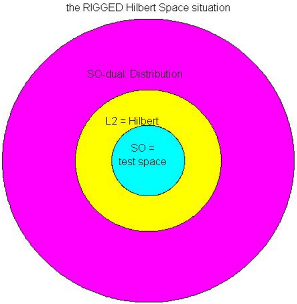
Definition

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

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

A schematic description: the simplified setting

In our picture this simple means that the inner “kernel” is mapped into the "kernel", the Hilbert space to the Hilbert space, and at the outer level two types of continuity are valid (norm and w^*)!



The prototypical examples over the torus

In principle every CONB (= *complete orthonormal basis*) $\Psi = (\psi_i)_{i \in I}$ for a given Hilbert space $\mathcal{H}$ can be used to establish such a unitary isomorphism, by choosing as $\boldsymbol{B}$ the space of elements within $\mathcal{H}$ which have an absolutely convergent expansion, i.e. satisfy $\sum_{i \in I} |\langle x, \psi_i \rangle| < \infty$.
For the case of the Fourier system as CONB for $\mathcal{H} = \boldsymbol{L}^2([0,1])$, i.e. the corresponding definition is already around since the times of N. Wiener: $\boldsymbol{A}(\mathbb{T})$, the space of absolutely continuous Fourier series. It is also not surprising in retrospect to see that the dual space $\boldsymbol{PM}(\mathbb{T}) = \boldsymbol{A}(\mathbb{T})'$ is space of *pseudo-measures*. One can extend the classical Fourier transform to this space, and in fact interpret this extended mapping, in conjunction with the classical Plancherel theorem as the first unitary Banach Gelfand triple isomorphism, between $(\boldsymbol{A}, \boldsymbol{L}^2, \boldsymbol{PM})(\mathbb{T})$ and $(\ell^1, \ell^2, \ell^\infty)(\mathbb{Z})$.

The Fourier transform as BGT automorphism

The Fourier transform $\mathcal{F}$ on $\mathbb{R}^d$ has the following properties:

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

Furthermore, we have that Parseval's formula

$$\langle f, g \rangle = \langle \widehat{f}, \widehat{g} \rangle \quad (6)$$

par

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

The general distributional case I

Those who know my work will not be surprised to see finally the use of the *Segal algebra* $(S_0(\mathbb{R}^d), \|\cdot\|_{S_0})$:

onv

Theorem

Let T be a TILS from $(S_0(\mathbb{R}^d), \|\cdot\|_{S_0})$ to $(S_0'(\mathbb{R}^d), \|\cdot\|_{S_0'})$. Then there exists a uniquely determine $\sigma \in S_0'$ such that one has for $f \in S_0(\mathbb{R}^d)$:

$$[\sigma \star f](z) = \sigma(T_z), \quad f \in S_0(\mathbb{R}^d).$$

Moreover, the operator norm of T and the S_0'-norm of σ are equivalent.

The general distributional case II

In this setting it is also clear that any such convolution kernel σ has a (generalized Fourier transform) via the standard convention

$$\widehat{\sigma}(f) := \sigma(\widehat{f}), \quad \sigma \in \mathbf{S}_0', f \in \mathbf{S}_0.$$

and that one may call σ the impulse response and $\widehat{\sigma}$ the transfer function of the given system/operators T.

This result applies also for e.g.
$T : \left(\mathbf{L}^p(\mathbb{R}^d), \|\cdot\|_p\right) \to \left(\mathbf{L}^q(\mathbb{R}^d), \|\cdot\|_q\right)$, for $1 \leq p, q < \infty$,
because $\mathbf{S}_0(\mathbb{R}^d) \hookrightarrow \mathbf{L}^p(\mathbb{R}^d) \hookrightarrow \mathbf{S}_0'(\mathbb{R}^d)$.