NuHAG Numerical Harmonic Analysis Group

Convolutions and Fourier Transforms: Existence and Good Properties

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

DMV-OEMG Conference
Passau, Sept. 29th, 2021 (15:30)

Orientation

Given the large amount of material provide in these slides let me give a short summary of the purpose of this talk:

- Convolution is relevant and interesting, both for mathematical analysis and real world applications, e.g. for digital signal and image processing;
- Engineers/physicists and mathematicians view the field differently, (different worlds);
- There is a simple (!better!) way to introduce convolution for bounded measures (as a starting point);
- OUTLOOK: Based on the Segal algebra $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ ("Feichtinger's algebra") one can develop a theory of mild distributions and show that the FT turns convolution into pointwise multiplication, even in the most general case.

Convolutions everywhere I

If I reflect upon my work in the last decades, then it is clear that convolution plays a significant role there. It all started with Reiter's book on Harmonic Analysis, where the Lebesgue space $\left(\boldsymbol{L}^1(G), \|\cdot\|_1\right)$ resp. the weighted version, $\left(\boldsymbol{L}^1_w(G), \|\cdot\|_{1,w}\right)$ (so-called *Beurling algebras*) play an important role: they all for Banach algebras with respect to convolution! With a suitable norm they have to satisfy something like

$$\|g * f\|_1 \leq \|g\|_1 \|f\|_1, \quad f, g \in \boldsymbol{L}^1(G). \tag{1}$$

nor

There are of course many others, among them specifically the so-called *Segal algebras* $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$, which are in fact *Banach ideals* in $\left(\boldsymbol{L}^1(G), \|\cdot\|_1\right)$, which means that

$$\|g * f\|_{\boldsymbol{B}} \leq \|g\|_1 \|f\|_{\boldsymbol{B}}, \quad g \in \boldsymbol{L}^1(G), f \in \boldsymbol{B}. \tag{2}$$

nor

Convolutions everywhere II

Similar estimates hold for function space such as $(\boldsymbol{L}^p(G), \|\cdot\|_p)$, which are not contained in $\boldsymbol{L}^1(G)$. Following Y. Katznelson we call them *homogeneous Banach spaces*.

From the point of view of Abstract Harmonic Analysis (**AHA**) such estimates can be established for general LCA groups, and thus one has to adapt the concrete interpretation of symbol "$\star$" for convolution (!as well as the meaning of the corresponding Fourier Transform) to the concrete group, typically $\mathbb{T}$ (Fourier series), $G = \mathbb{R}^d$ (Fourier transforms, as used for PDE), finite groups for applications for time-series or for digital image processing.

Looking at the usefulness of convolution for engineering applications one finds the distinction between discrete and continuous, or periodic and non-periodic functions and correspondingly variants for the FT (e.g. the FFT for finite,

Convolutions everywhere III

i.e. discrete and periodic signals). What is in common to all these settings is the so-called **Convolution Theorem**

$$\mathcal{F}(f * g) = \mathcal{F}(f) \cdot \mathcal{F}(g), \quad (3)$$

con

and of course the usual rules for computation with an abstract multiplication (i.e. *bilinearity* and *associativity*).
Especially for the case $G = \mathbb{R}^d$, which is non-compact and carries "continuous", non-periodic functions (say in $\boldsymbol{L}^2(\mathbb{R}^d)$) many naive engineering terms, like the description of a function as a superposition of unit vectors (as it is possible in the discrete case) become *extremely vague* and very often "non-existent" integrals are used to juggle at a symbolic level and "derive" results, which are valid, but which should (and could) be derived on more solid mathematical grounds.

Convolutions everywhere IV

Often the Dirac measure δ_0 is used (sometimes in a magic way), and it's so-called sifting property is used to "represent a given function f" as a weighted integral of the form:

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

Surprisingly this "formula", which expresses nothing more than the trivial fact, that the evaluation of a (continuous) function f at x gives $f(x) \in \mathbb{C}$, can be turned into a tool which allows to "prove" meaningful claims!

Obviously there are big differences in the way how engineers (or physicists) or mathematicians are viewing things. All to often mathematicians cannot accept engineering proofs as more than a heuristic, making things plausible, and develop a more rigorous treatment of terms such as the Dirac measure δ_0.

Convolutions everywhere V

It is obviously not just a strange "ordinary function", since it should satisfy $\int_{-\infty}^{\infty} \delta_0(x)dx = 1$, e.g. in the context of distribution theory, specifically in the context the Schwartz space $\mathcal{S}'(\mathbb{R}^d)$ of *tempered distributions*.
On the other hand, if it comes to the "correct treatment" of the convolution it appears natural to invoke the "best available integral", namely the Lebesgue integral, which has been established as a standard tool for mathematical analysis during the last 100 years.
However, it is clear that we can understand the action of δ_x (often written as $\delta(. - x)$) by taking a so-called Dirac sequence at $t = 0$ and shift it to the position x.

Basics and Motivation

Let us first give a quick overview over the plan for this talk: Convolutions have been in the focus of my work since my PhD thesis (in 1974), but the reflection concerning the motivation for such studies came much later. We will start with

- Questions about the appearance of convolutions;
- The way how they appear naturally in applications;
- The way how convolution is treated mathematically;
- Provide a synthesis of view-points by allowing a glimpse on the idea of Conceptual Harmonic Analysis (CHA).

The goal of CHA is to bridge the gap between the different worlds (of physics, engineering, mathematics and mathematics, taking computational and functional analytic aspects into account, providing a *unified approach*).

Signals and Systems for Engineers

If we browse engineering books or courses on "Signals and Systems" we can learn about different types of signals:

- discrete or continuous;
- periodic or non-periodic;
- one-dimensional, multi-dimensional (images);
- corresponding concepts of convolution and
- different forms of Fourier transforms adapted to the different settings, e.g. the DFT/FFT for the (i.e. finite) setting!
- (translation invariant) systems are described as convolution operators which are turned into pointwise multiplications operators by the FT (filters).

Wide range of applications: MP3, JPEG, dig. signal processing,...!

TILS: Lüke/Ohm

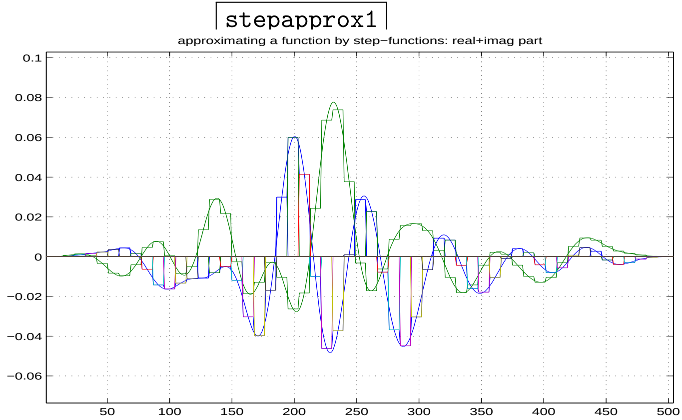


Abbildung: A typical illustration of an approximation to the input of a TILS T, preparing for the use of the Dirac impulse.

The explanation of this plot I

1. The picture shows (real and imaginary part of a) function, which is approximated by step functions;
2. These step functions are obtained by compressing a boxcar function $\mathbf{1}_{[-1/2,1/2]}$ and shifting it to the correct position;
3. Obviously the step functions get closer to the original function as the spacing gets more and more narrow ($h \to 0$);
4. On the other had the compressed rectangular function, all assumed to satisfy $\int_{-\infty}^{\infty} b(x)dx = 1$ tend, in the limit, to δ_0, the Dirac "function", thus justifying the rule

$$\int_{-\infty}^{\infty} \delta(y)dy = 1.$$

One must say, that an attempt to make the statement found in this context mathematically solid claims is a challenging task!

Questions arising from these pictures

- In which sense does the limit of the rectangular functions exist? Maybe the symbol δ or δ_0 is just a *phantom*?
- 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, e.g. uniformly, or in the $\boldsymbol{L}^1$-sense, and how are the steps determined (samples, local averages)?
- 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.

Abstract Harmonic Analysis I

In a nut-shall Abstract Harmonic Analysis (AHA) provides a unified view-point on many of the above mentioned concepts by analyzing the underlying common structure in a unified way.

Starting from a locally compact Abelian (LCA) group G the task is to analyse objects defined over such a group (functions, distributions, signals), typically belonging to translation invariant Banach spaces (or topological spaces of such objects). The Fourier transform provides a way to *decompose* or *approximate* these functions by "superposition" of pure frequencies, called *characters* in this general setting.
These characters (continuous functions from G into the torus groups $\mathbb{U}$ satisfying the exponential law), in turn, form a group, the so-called *dual group* $\widehat{G}$, and the Fourier transform maps objects on G into objects on $\widehat{G}$.

Abstract Harmonic Analysis II

Using a natural topology on $\widehat{G}$ it is a LCA group, with the dual group being naturally isomorphic to G (Pontryagin's Thm.).
It is well known, that $\widehat{G}$ is discrete if and only if G is compact, or that the members of the dual group form an ONB for $\ell^2(G)$ for any finite Abelian group, so in particular $\#G = \#\widehat{G}$ and the (discrete) Fourier transform on G maps finite vectors, viewed as elements of $\ell^2(G)$, is a unitary change of bases to $\ell^2(\widehat{G})$.

For the standard case (vectors of finite length, "pixel"-images viewed as matrices) this change of bases is implemented efficiently via the FFT/FFT2 algorithms, the Fast Fourier Transform (or mathematically just the DFT). The issue here is efficiency of the computation of the DFT by suitable algorithms ($>$ Computational Harmonic Analysis).

How does convolution appear?

In both worlds convolution plays a crucial role.

1. In the applications the main motivation for the study of convolutions appears to be the characterization of TILS (translation invariant systems) as convolution operators, which are turned into pointwise multipliers, BUT often the vague description is not put on solid groups (seen from the mathematical view-point).
2. In Mathematical Analysis Lebesgue's integration theory allows to study in great detail the Banach algebra $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ and the property of the Fourier transform as a mapping between different function spaces.
3. The theory of Tempered Distributions (L. Schwartz) is a sound basis for the treatment of PDEs.

A list of problems not touched in this way

There is not enough time to discuss a (potentially long) list of short-comings that these relatively isolated bodies of knowledge in the different disciplines brings along!

1. The way how Fourier Analysis is developing (e.g.) in engineering and mathematics is rather divergent than convergent (new algorithms for mobile communication versus sophisticated function spaces, say);
2. AHA is just stressing the analogy between the different groups, while ignoring the fact that one would like to *approximate* the FT of a given (smooth and decaying) function, making use of FFT-based methods.
3. Time-frequency methods are developed for both the continuous and the discrete setting, again little is done to justify the transition, e.g. compute $\widehat{f}$ via FFT of samples.

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)

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

Discrete Convolutions I

We see in the discrete setting:

1. Discrete convolution is more or less multiplication of polynomials (via Cauchy Products), i.e. at the coefficient level;
2. The DFT/FFT can be used to realize this
3. Discrete systems allow the unit vector $\mathbf{e}_0$ as input, and therefore $T(\mathbf{x})$ can be written as convolution product by the "impulse response" $T(\mathbf{e}_0)$ (recall $\mathcal{F}(\mathbf{e}_0) = \mathbf{1}$), since

$$T(\mathbf{x}) = T(\mathrm{Id}(\mathbf{x})) = T(\mathbf{e}_0 \star \mathbf{x}) = T(\mathbf{e}_0) \star \mathbf{x}.$$

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 (as an illustration) is just the sum of the probabilities of obtaining the pairs

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

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

Discrete Convolution, Summary

So summarizing shortly the findings concerning discrete convolutions:

1. Discrete convolution corresponds to the Cauchy Product, describing the (pointwise) multiplication of polynomials at the coefficient level;
2. The DFT/FFT (resp. FFT2, etc.) allows to efficiently implement such convolutions, which are at the basis of digital signal processing;
3. There are nice connections to the multiplication of (long) integers or *elementary probability*.

When does the convolution $f \star g$ exist?

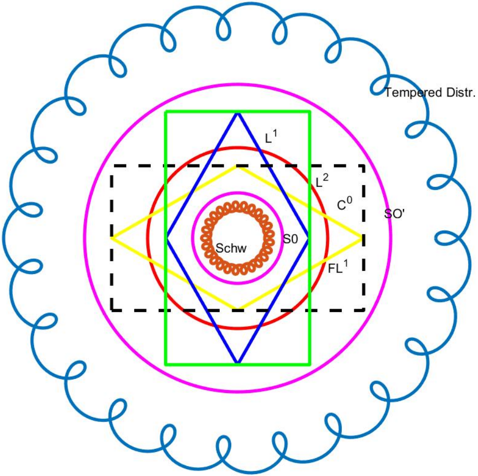
Given two functions $f, g \in \boldsymbol{C}_c(\mathbb{R}^d)$ (compactly supported, continuous, complex-valued functions on $\mathbb{R}^d$) it is well known that the classical way to define their convolution product is via a pointwise integral formula:

$$[g \star f](x) := \int_{\mathbb{R}^d} f(x-y)g(y)dy, \quad x, y \in \mathbb{R}^d. \tag{4}$$

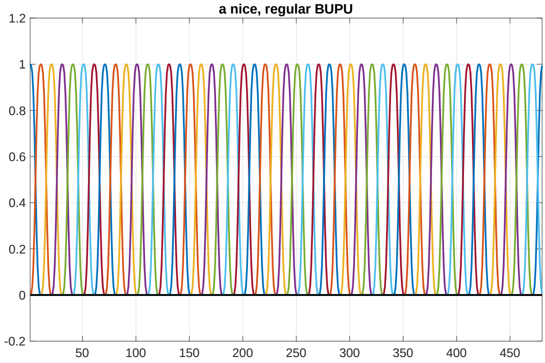
con

Observing that $(\boldsymbol{C}_c(\mathbb{R}^d), \|\cdot\|_1)$ turn out to be a commutative, normed Banach algebra with respect to convolution

A Zoo of Banach Spaces for Fourier Analysis



Let us consider the space of *cubic splines* in $\left(\boldsymbol{L}^2(\mathbb{R}), \|\cdot\|_2\right)$, then it is known to be generated by the shifted versions of the cubic B-spline. which form a BUPU,a Bounded Uniform Partitions of Unity and a Riesz basis for their closed linear span:



BUPUETH01.pdf

Orthogonal Projection

The key-properties of cubic splines in $(\boldsymbol{L}^2(\mathbb{R}), \|\cdot\|_2)$ can be described as follows:

Lemma

The family $(T_n\phi)_{n\in\mathbb{Z}}$ forms a Riesz basis for the closed linear subspace $\boldsymbol{V}_\phi$ of $(\boldsymbol{L}^2(\mathbb{R}), \|\cdot\|_2)$ constituted by the cubic splines in $\boldsymbol{L}^2(\mathbb{R})$. In fact, there is a biorthogonal family $(T_n\tilde{\phi})_{n\in\mathbb{Z}}$, which allows to describe the orthogonal projection onto $\boldsymbol{V}_\phi$ as

$$P_\phi(f) = \sum_{n\in\mathbb{Z}} f * \tilde{\phi}(n) T_n\phi.$$

There is also another function $\varphi \in \boldsymbol{V}_\phi$ such that $(T_n\varphi)_{n\in\mathbb{Z}}$ constitutes an ONB for $\boldsymbol{V}_\phi$.
BUT what happens on $\boldsymbol{L}^p(\mathbb{R})$? Is P_ϕ still continuous there?

There are different uses of BUPUS

Using the BUPUs we can define two operators, one on $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$ and the other on the dual space, $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$. We use the abstract notation $\Psi = (\psi_i)_{i\in I}$:

t02

Definition

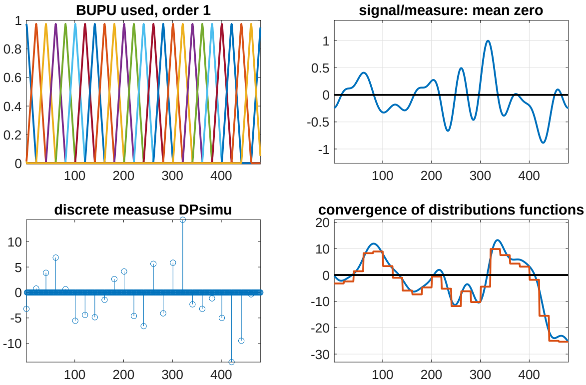
$$\mathrm{Sp}_\Psi f := \sum_{i\in I} f(x_i)\psi_i, \quad f \in \boldsymbol{C}_0(\mathbb{R}^d). \text{ QUASI-INTERPOLATION}$$

Psi

Definition

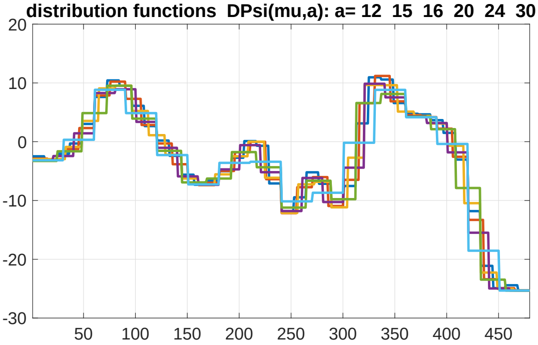
$$D_\Psi \mu := \sum_{i\in I} \mu(\psi_i)\delta_{x_i}, \quad \mu \in \boldsymbol{M}_b(\mathbb{R}^d). \text{ DISCRETIZATION}$$

In fact, a simple operation like "piecewise linear interpolation" for functions in $\left(\boldsymbol{C}_b(\mathbb{R}), \|\cdot\|_\infty\right)$ can be seen as an operator of the form Sp_Ψ, using the triangular BUPU, and many numerical integration methods are just providing closed formulas for the integrals of such approximations.



simu01b.pdf

Abbildung: The (triangular) BUPU, the smooth signal/measure, the discretized measure illustrate by the STEM-command, and the corresponding distribution function, which has a jump of sice c_k at the position of the corresponding (positive or negative) Dirac measure.



uvargap1c.pdf

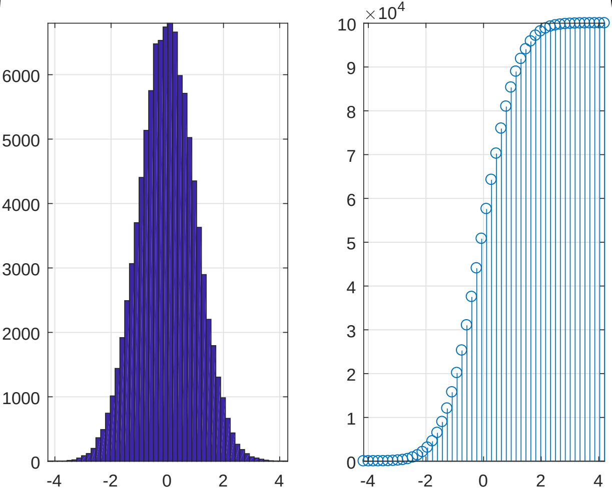


Abbildung: w^*-convergence is ptw. convergence of distribution fcts.

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):

$$[f \star g](x) = \int_{\mathbb{R}^d} f(x-y)g(y)dy, \quad x, y \in \mathbb{R}^d \tag{5}$$

con

and then shows that

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

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 $(\boldsymbol{C}_c(\mathbb{R}^d), \|\cdot\|_1)$ as a normed algebra which allows to introduce $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ as the (abstract) completion of this normed algebra. Consequently we see that $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ 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 (5) and the norm estimate for the convolution product can be well established.

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

Ways to complete the normed algebra 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

$$T \circ T_z = T_z \circ T, \quad \forall z \in \mathbb{R}^d. \tag{6}$$

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 (6)) 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 $(\boldsymbol{X}, \|\cdot\|_{\boldsymbol{X}}) = \big(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\big)$ one has Wendel's Theorem, characterizing (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 $\big(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\big)$, 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

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

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

$$T(\delta_0) = T(\lim_{\alpha\to\infty} e_\alpha) = \lim_{\alpha\to\infty} (T(e_\alpha)) = \mu_0! \tag{7}$$

imp

As it turns out, this formula can be justified under suitable conditions, but not by choosing $(\boldsymbol{X}, \|\cdot\|_{\boldsymbol{X}}) = (\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty)$ (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 $(\boldsymbol{X}, \|\cdot\|_{\boldsymbol{X}}) = (\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2)$. Here we can make use of Plancherel's Theorem.

Impulse Response and Transfer Function II

ult

Theorem

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

$$\mathcal{F}(T(f)) = \widehat{f} \cdot h, \quad f \in \mathbf{L}^2(\mathbb{R}^d). \tag{8}$$

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* $h(s) = exp(2\pi i s^2) \in \boldsymbol{C}_b(\mathbb{R}^d)$ 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 $\mathcal{F}^{-1}(\mathbf{1}_{[-1/2,1/2]}) = \mathrm{SINC} \in \boldsymbol{L}^2(\mathbb{R}^d) \setminus \boldsymbol{L}^1(\mathbb{R}^d)$ 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.

NEW WAY: 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 $\left(M_b(\mathbb{R}^d), \|\cdot\|_{M_b}\right)$ for $\left(C_0'(\mathbb{R}^d), \|\cdot\|_{C_0'}\right)$. Recall that $\mu_0 = w^*\text{-}\lim_{\alpha\to\infty} \mu_\alpha$ if and only for any $f \in C_0(\mathbb{R}^d)$: $\mu_\alpha(f) \to \mu_0(f)$ (w^*—-convergence of measures);
E.g. $\mu = w^*\text{-}\lim_{|\Psi|\to 0} D_\Psi \mu.$

NEW WAY: 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

$$\|T_x f\|_\infty = \|f\|_\infty, \quad f \in C_0(\mathbb{R}^d), x \in \mathbb{R}^d, \tag{9}$$

COt

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

NEW WAY: 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* $(M_b(\mathbb{R}^d), \|\cdot\|_{M_b})$.

$$(\mathcal{H}_{\mathbb{R}^d}(C_0(\mathbb{R}^d)), |||\cdot|||) \cong (M_b(\mathbb{R}^d), \|\cdot\|_{M_b}).$$

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.

NEW WAY: Modelling via $(C_0(\mathbb{R}^d), \|\cdot\|_\infty)$ IV

LIS

Theorem

[Characterization of LTISs on $C_0(\mathbb{R}^d)$]
There is a natural isometric isomorphism between the Banach space $\mathcal{H}_{\mathbb{R}^d}(C_0(\mathbb{R}^d))$, endowed with the operator norm, and $(M(\mathbb{R}^d), \|\cdot\|_M)$, the dual of $(C_0(\mathbb{R}^d), \|\cdot\|_\infty)$, by means of the following pair of mappings:

1. *Given a bounded measure $\mu \in M(\mathbb{R}^d)$ we define the operator C_μ (to be called convolution operator with convolution kernel μ later on) via:*
$$C_\mu f(x) = \mu(T_x f^{\checkmark}). \tag{10}$$
2. *Conversely we define $T \in \mathcal{H}_{\mathbb{R}^d}(C_0(\mathbb{R}^d))$ the linear functional $\mu = \mu_T$ by*
$$\mu_T(f) = [Tf^{\checkmark}](0). \tag{11}$$

The claim is that both of these mappings: $C : \mu \mapsto C_\mu$ and the

con

imp

There are various benefits

The new approach has a number of benefits:

1. First of all it connects the engineering approach to TILS with the mathematical concepts for convolution in the measure algebra $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$, viewed as dual space to $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$.
2. The setting allows to formulate the general principles up to the validity of the Convolution Theorem

$$\mathcal{F}(\mu_1 \star \mu_2) = \mathcal{F}\mu_1 \cdot \mathcal{F}\mu_2, \quad \mu_1, \mu_2 \in \boldsymbol{M}_b(\mathbb{R}^d).$$

3. The approach works for general LCA groups!
4. Using the Banach Gelfand Triple $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(G)$ the expected analogy between "impulse response" and "transfer function" can be presented in a rigorous way!

Approximation by discrete convolution

The w^*-approximation of μ by the discrete measures $D_\Psi\mu$ (taking the role of finer and finer histograms) suggests to define the convolution of two measures as the limit of the product of their discretized version (similar to the product $\pi \cdot \sqrt{5}!$).
This is can in fact be justified:

v07

Lemma

Given $\mu, \nu \in \boldsymbol{M}_b$ one has for anz $f \in \boldsymbol{C}_0$ and $|\Psi| \to 0$:

$$(D_\Psi\mu * \mathrm{D}_\Psi\nu) * f = D_\Psi\mu * (\mathrm{D}_\Psi\nu * f) \to (\mu * \nu) * f \quad \text{in} \left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right). \tag{17}$$

DPs

In particular one has w^*-convergence of

$$D_\Psi\mu * \mathrm{D}_\Psi\nu \to^{w^*} \mu * \nu, \quad \text{for } |\Psi| \to 0. \tag{18}$$

DPs

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

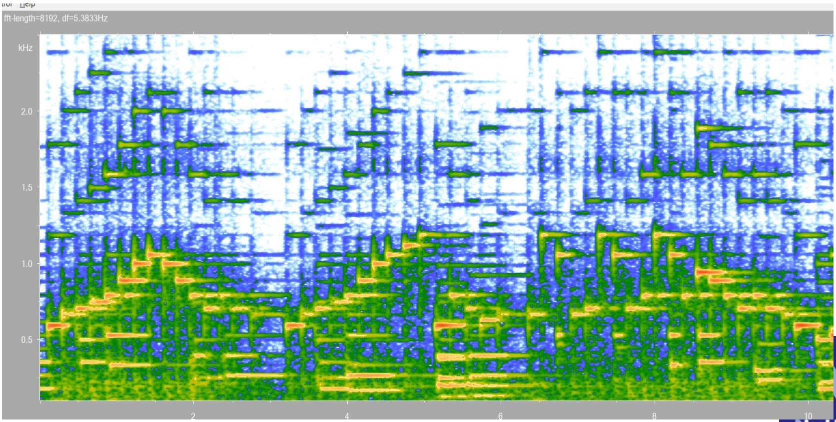
$$(T_x f)\widehat{\ } = M_{-x}\hat{f} \qquad (M_\omega f)\widehat{\ } = T_\omega \hat{f}$$

The Short-Time Fourier Transform

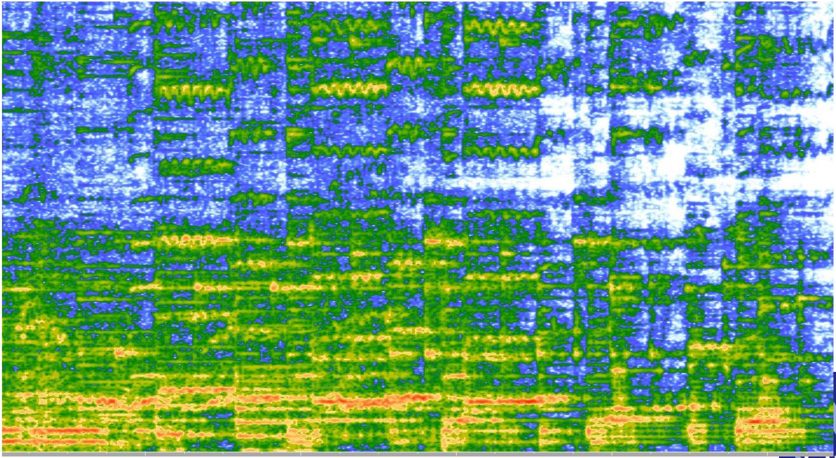
$$V_g f(\lambda) = \langle f, M_\omega T_t g\rangle = \langle f, \pi(\lambda) g\rangle = \langle f, g_\lambda\rangle, \;\; \lambda = (t, \omega);$$

A Typical Musical STFT

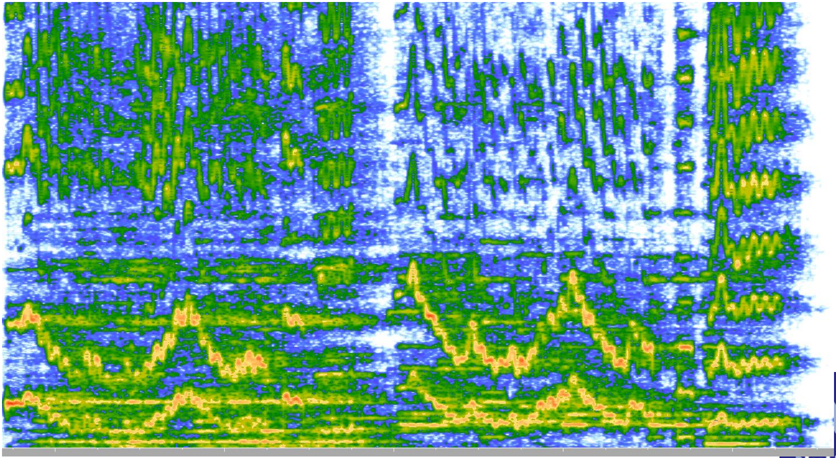
A typical piano spectrogram (Mozart), from recording



A Musical STFT: Brahms, Cello

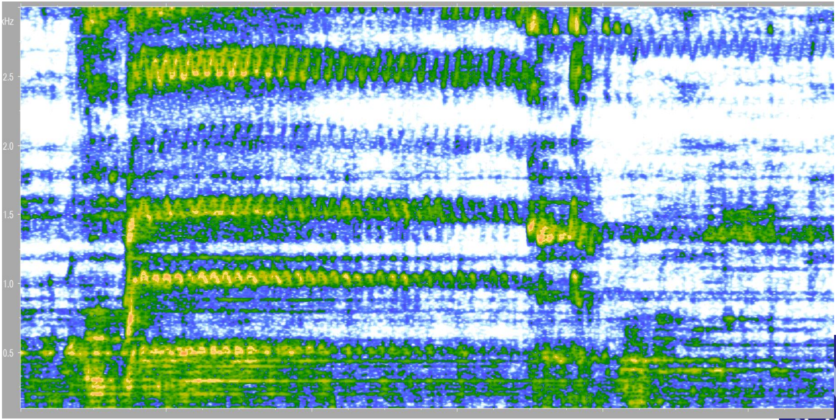


A Musical STFT: Maria Callas



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

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

The space $(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0})$ 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.

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

cSo

Lemma

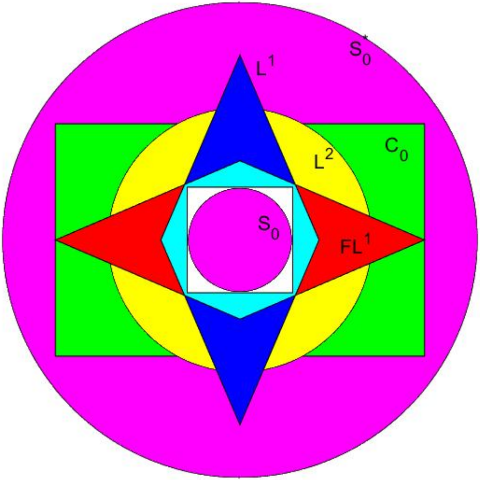
Let $f \in \boldsymbol{S}_0(\mathbb{R}^d)$, then the following holds:

(1) $\pi(u,\eta)f \in \boldsymbol{S}_0(\mathbb{R}^d)$ *for* $(u,\eta) \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d$, *and* $\|\pi(u,\eta)f\|_{\boldsymbol{S}_0} = \|f\|_{\boldsymbol{S}_0}$.

(2) $\hat{f} \in \boldsymbol{S}_0(\mathbb{R}^d)$, *and* $\|\hat{f}\|_{\boldsymbol{S}_0} = \|f\|_{\boldsymbol{S}_0}$.

In fact, $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ is the smallest non-trivial Banach space with this property, and therefore contained in any of the $\boldsymbol{L}^p$-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.

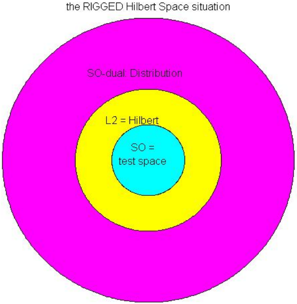
Definition

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

$$\langle f, g \rangle = \langle \widehat{f}, \widehat{g} \rangle \tag{19}$$

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* $\left(\mathbf{S}_0(\mathbb{R}^d), \|\cdot\|_{\mathbf{S}_0}\right)$:

onv

Theorem

Let T be a TILS from $\left(\mathbf{S}_0(\mathbb{R}^d), \|\cdot\|_{\mathbf{S}_0}\right)$ to $(\mathbf{S}_0'(\mathbb{R}^d), \|\cdot\|_{\mathbf{S}_0'})$. Then there exists a uniquely determine $\sigma \in \mathbf{S}_0'$ such that one has for $f \in \mathbf{S}_0(\mathbb{R}^d)$:

$$[\sigma \star f](z) = \sigma(T_z), \quad f \in \mathbf{S}_0(\mathbb{R}^d).$$

Moreover, the operator norm of T and the $\mathbf{S}_0'$-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

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

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

Iterated limits

net

Lemma

Assume that $(T_\alpha)_{\alpha \in I}$ *and* $(S_\beta)_{\beta \in J}$ *are two* **bounded** *(!) nets of operators in* $\mathcal{L}(\boldsymbol{V})$*, strongly convergent* T_0*, and* S_0 *resp. i.e.*

$$T_o(\mathbf{v}) = \lim_\alpha T_\alpha(\mathbf{v}) \quad \forall\, \mathbf{v} \in \boldsymbol{V} \quad \textit{and} \quad S_0(\mathbf{w}) = \lim_\beta S_\beta(\mathbf{w}) \quad \forall\, \mathbf{w} \in \boldsymbol{V}.$$

Then the net $(T_\alpha \circ S_\beta)_{(\alpha,\beta)}$ *is strongly convergent to* $T_0 \circ S_0$*:*

$$T_o[S_0(\mathbf{v})] = [T_0 \circ S_0](\mathbf{v}) = \lim_{\alpha,\beta} [T_\alpha \circ S_\beta](\mathbf{v}). \tag{20}$$

dou

In detail: Given $\mathbf{v} \in \boldsymbol{V}, \varepsilon > 0$ *there exists a pair of indices* $(\alpha_0, \beta_0) \in I \times J$ *such that for* $\alpha \succeq \alpha_0$ *in* I *and* $\beta \succeq \beta_0$ *in* J

$$\| T_0(S_0(\mathbf{v})) - T_\alpha(S_\beta(\mathbf{v}))\| \leq \varepsilon. \tag{21}$$

dou

THANKS!

www.nuhag.eu/bibtex: all papers
www.nuhag.eu/talks: many talks on related subjects
NOTE: **Talks require access code:**
"visitor" and "nuhagtalks"

At the page www.nuhag.eu/ETH20 you find many links to Lecture Notes and YouTube Lectures.
Various Lecture Notes can be downloaded from
https://www.univie.ac.at/nuhag-php/home/skripten.php

Feel free to REQUEST more information from the author.