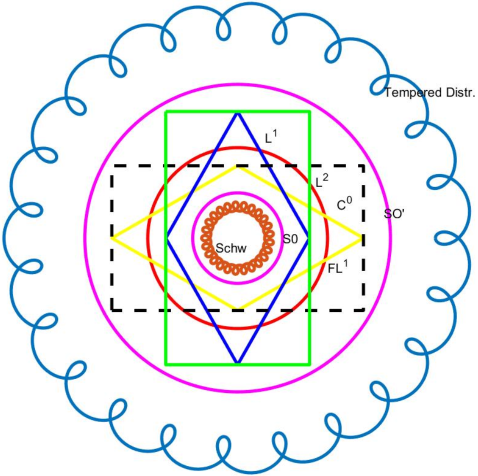
NuHAG Numerical Harmonic Analysis Group

Approximation of continuous problems in Fourier Analysis by finite dimensional ones: The setting of the Banach Gelfand Triple

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

Talk held at IWOTA 2019
Lisbon, July 22th, 2019

A Zoo of Banach Spaces for Fourier Analysis

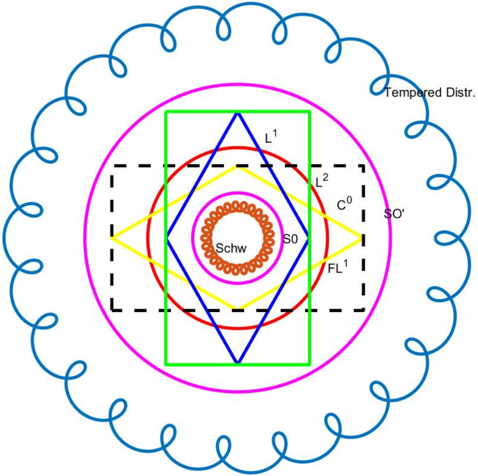


Approximation of continuous problems in Fourier Analysis by finite dimensional ones: The setting of the Banach Gelfand Triple

Hans G. Feichtinger

Talk held at IWOTA 2019
Lisbon, July 22th, 2019

A Zoo of Banach Spaces for Fourier Analysis



Many years of Computational Gabor Analysis I

The CONTENT of this talk is based on ca. 30 years of experience with experimental computational harmonic analyis.

The PURPOSE of this talk is to share some general ideas about the transition between continuous and corresponding discrete problems, to some extent concerning the connections between the theory of function spaces, the domain of real world problems and finally computational issues.

In the framework of Harmonic Analysis this has led me to the development of the idea of Conceptual Harmonic Analysis, which is more than just a synthesis of Abstract Harmonic Analysis (AHA) and Numerical Harmonic Analysis.

Many years of Computational Gabor Analysis II

As we will see, families of Banach spaces play a big role for a proper description of the setting, and among them the minimal families, sometimes called *rigged Hilbert spaces* or in our case the Banach Gelfand Triple $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$ will serve as a good exmpale for the illustration of our ideas.

The *families of Banach spaces* we have in mind are families of weighted $\boldsymbol{L}^p$-spaces for the reconstruction of band-limited functions (or functions from spline-type spaces, like shifted Gaussians), in wavelet theory this are the well-known Besov-Triebel-Lizorkin spaces, and for time-frequency analysis we are dealing with modulation spaces.

Various types of SAMPLING

Sampling Theory is a wide field which has many interesting aspects and challenges:

- Given data, can one approximate a regression manifold?
- How can on meet the challenge of large data sets, high dimensions or irregular sampling?
- One direction is the task to *approximate/reconstruct* from a given data set **(PASSIVE SAMPLING)**;
- If one can decide about the sampling strategy: How should one sample most efficiently? **(ACTIVE SAMPLING)**
- Of course efficiency depends on both the reconstruction method and the request (typically the choice of a function space): small error in *some norm, chosen by the user!*;

Irregular Sampling Algorithms

Let us just give a short summary, for the case of irregular sampling. Here one has typically the following situation:
Assume we are given *irregular samples* $(f(x_i))_{i \in I}$ in $\mathbb{R}^d$ of a function f which, as a tempered distribution has spectral support in a ball or cube Ω of known size.
Then good results concerning iterative algorithms will tell us: Given the size/diameter of Ω and the density of $(x_i)_{i \in I}$ a certain, say iterative algorithm requires so and so many interactions to guarantee that for any of the norms which apply to f (e.g. weighted $\boldsymbol{L}^p$-norms, with $1 \leq p \leq \infty$ and weights up to order 10) the relative reconstruction error after a fixed number of iterations with be at most 3%.
Via norm equivalences on the given space one can even estimate Sobolev norms for the involved functions.

What we have learned in this setting I

In the setting of irregular sampling we have - over the years - learnt a lot. Let us list just a few issues, which allow us to describe the situation (resp. the *quality of algorithms*).

- Algorithms have to be universal in the sense that e.g. band-limited functions may belong to various types of spaces, typically weighted L^p-spaces. When applying a given algorithm the knowledge of the user about the membership of the function from which data are drawn **must not be part of the algorithm**, i.e. it has to work automatically, with guaranteed recovery rates (depending on the amount and quality of information available).

What we have learned in this setting II

- As one does not have in real-world applications the exact date one has to be able to analyze the effect of perturbation of the assumptions: What if the function is only "approximately band-limited", what if the data provided are not take at exactly described positions but only within some margin (jitter error), what if instead of exact pointwise values small local averages are provided, and so on. Is the algorithm still convergent under these conditions? And is the limit close to the limit that one would expect?

What we have learned in this setting III

- When running numerical simulations one typically chooses a discrete setting. But then a couple of questions arise: Does the existing continuous theory guarantee that the algorithms, now applied to an analogue discrete situation also have to converge in the same way? Or do we need (!in effect) a separate discrete theory? And/but how does the computable discrete computation help to find solutions to the corresponding continuous problem? What kind of errors have to be taken into account.

The idea of CONCEPTUAL HARMONIC ANALYSIS

I am promoting since 2008 (first talk in this direction in Cambridge!) the idea of Conceptual Harmonic Analysis (CHA), an attempt to unify concepts from

Abstract and Computational Harmonic Analysis.

While engineers are satisfied with a simple replacement of a continuous Fourier transform by the corresponding FFT I would like to view things from a *Constructive Approximation* point if view.

- Given a numerical method that allows to determine the values of a given functions pointwise, how can one obtain information about its Fourier transform (via FFT methods).
- Given an Given some pseudo-differential operator (say via its Weyl-symbol or Kohn-Nirenberg symbol, or as a (singular) integral kernel, how can we reliably compute eigenvalues and eigenvectors, or solving corresponding linear equations?

The Setting: Mild Distributions

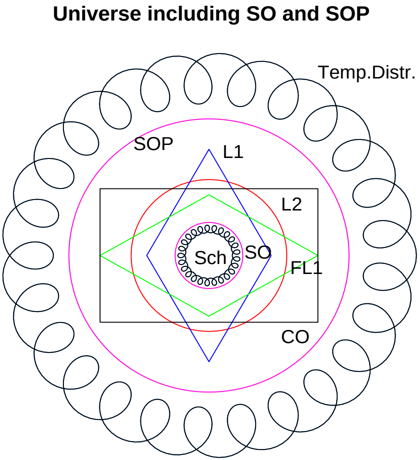
The setting for which I will describe a certain approach (based on the Banach Gelfand Triple (BGT) resp. rigged Hilbert space concept $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$ is related to topics such as

- Fourier transforms (approximation via FFT);
- Time-Frequency and Gabor Analysis (finite vs. continuous);
- Gabor multipliers, Anti-Wick operators,
- general linear operators between $\boldsymbol{L}^p$-spaces,
- finite dimensional approximation (in the w^*-sense);

Active Sampling and Abstract Harmonic Analysis

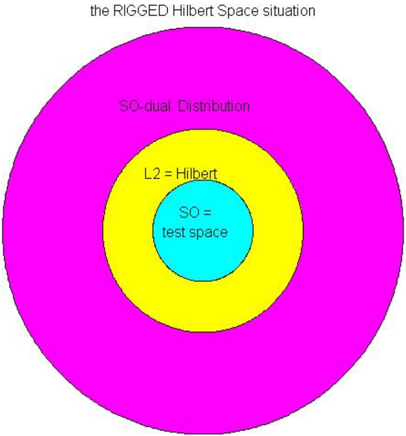
I want to discuss here a particular aspect of *active sampling*, based on ideas from AHA (Abstract Harmonic Analysis) and Computational Time-Frequency Analysis (TFA) or Gabor Analysis. AHA tells us, that we can to Fourier and Gabor analysis on any (!) LCA (locally compact Abelian) group G. In particular on $\mathbb{R}^d$, $\mathbb{Z}^d$ (or any other discrete lattice $\Lambda = \boldsymbol{A} * \mathbb{Z}^d \lhd \mathbb{R}^d$.
But we can also do it in the discrete and periodic setting (engineering terminology), resp. over *finite Abelian groups*, which is actually what is/can be done quite efficiently using mathematical software (MATLAB in my case, for 30 years now!).
AHA, furthermore tells us, how to describe the analogies between different groups: we have translations, dual groups, hence modulations, and so on. But how should one relate observations made over different groups to each other?

The Classical Setting of Test Functions & Distributions



A schematic description: THE Banach Gelfand Triple

Testfunctions $\subset$ Hilbert space $\subset$ Distributions, like $\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$!



The Key-players for Time-Frequency Analysis (TFA)

Time-shifts and Frequency shifts (II)

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

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

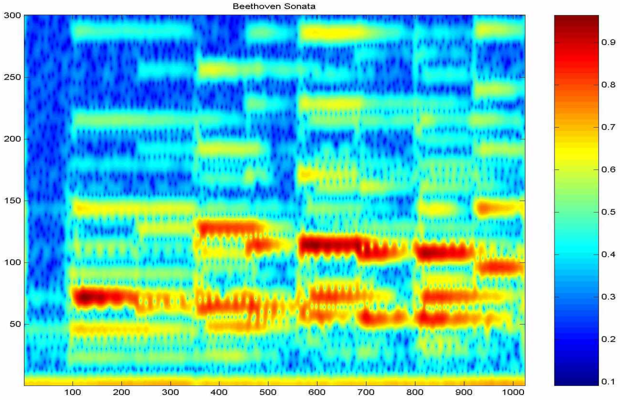
$$(T_x f)\widehat{\ } = M_{-x}\hat{f} \qquad (M_\omega f)\widehat{\ } = T_\omega \hat{f}$$

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 waterfall melody (Beethoven piano sonata) using the spectrogram: energy distribution in the TF = time-frequency plan:



Gabor Analysis: Modern Viewpoint I

The fact that the set $\{g_\lambda \mid \lambda \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d\}$ is a highly redundant family of vectors in the Hilbert space $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$ led to the suggestion of **D. Gabor in 1946** to select a "critical" family, namely $\lambda \in \Lambda_0 := \mathbb{Z} \times \mathbb{Z}$, hoping that in this way the properties of a (non-orthogonal) basis (modern view-point: a Riesz basis) could be achieved.
EVERY element $f \in \boldsymbol{L}^2(\mathbb{R})$ should be representable in a unique way, with coefficients representing the energy distribution within f in a unique way. Equivalently, recovery of $V_g(f)$ and hence f from samples at Λ_0 should be (stably) possible.

Gabor Analysis: Modern Viewpoint II

It follows from **Moyal's formula** (energy preservation):

$$\|V_g(f)\|_{\boldsymbol{L}^2(\mathbb{R}^d \times \widehat{\mathbb{R}}^d)} = \|g\|_2 \|f\|_2, \quad f, g \in \boldsymbol{L}^2. \tag{1}$$

This setting is well known under the name of coherent frames when $g = g_0$, the Gauss function. Its range is the *Fock space*. Approximate reconstruction results are due to Ferenc Weisz (appropriate fine Riemanian sums, even norm convergence in $(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2)$).
The possibility of having such *tight Gabor frames* is resulting from the continuous reconstruction formula, valid for arbitrary $\boldsymbol{L}^2$-atoms g. Writing again for $\lambda = (t, \omega)$ and $\pi(\lambda) = M_\omega T_t$, and furthermore $g_\lambda = \pi(\lambda) g$ we have in fact for any $g \in \boldsymbol{L}^2(\mathbb{R}^d)$ with $\|g\|_2 = 1$:

$$f = \int_{\mathbb{R}^d \times \widehat{\mathbb{R}}^d} \langle f, g_\lambda \rangle g_\lambda \, d\lambda.$$

Gabor Analysis: Modern Viewpoint III

Todays Rules of the Game

Choose a good *window* or *Gabor atom* (any $g \in \mathcal{S}(\mathbb{R}^d)$ will do) and try to find out, for which lattices $\Lambda \in \mathbb{R}^{2d}$ the signal f resp. its STFT (with that window) can recovered in a STABLE way from the samples, i.e. from the values $\langle f, \pi(\lambda)g \rangle$.
We speak of *tight Gabor frames* (g_λ) if we can even have the expansion (for some constant $A > 0$)

$$f = A \cdot \sum_{\lambda \in \Lambda} \langle f, g_\lambda \rangle g_\lambda, \quad \forall \quad f \in \mathbf{L}^2(\mathbb{R}^d).$$

Note that in general *tight frames* can be characterized as orthogonal projections of orthonormal bases of larger spaces!!!

Gabor Analysis: Modern Viewpoint IV

Another basic fact is that for each $g \in \mathcal{S}(\mathbb{R}^d)$ one can find, if Λ is dense enough (e.g. $a\mathbb{Z} \times b\mathbb{Z} \subset \mathbb{R}^d$ for $ab < 1$ in the Gaussian case) a *dual Gabor window* $\widetilde{g}$ such that one has at least

$$f = \sum_{\lambda \in \Lambda} \langle f, \widetilde{g}_\lambda \rangle g_\lambda = \sum_{\lambda \in \Lambda} \langle f, g_\lambda \rangle \widetilde{g}_\lambda \qquad (2)$$

$\widetilde{g}$ can be found as the solution of the (positive definite) linear system $S\widetilde{g} = g$, where $Sf = \sum_{\lambda \in \Lambda} \langle f, g_\lambda \rangle g_\lambda$, so using $\widetilde{g}$ instead of g for analysis or synthesis corrects for the deviation from the identity operator. An important fact is the commutation relation $S \circ \pi(\lambda) = \pi(\lambda) \circ S$, for all $\lambda \in \Lambda$.
Thus (2) is just $S \circ S^{-1} = Id = S^{-1} \circ S$ in disguise!).

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 $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$

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

The space $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ is a Banach space, for any fixed, non-zero $g \in \boldsymbol{S}_0(\mathbb{R}^d)$), and different windows g define the same space and equivalent norms. Since $\boldsymbol{S}_0(\mathbb{R}^d)$ contains the Schwartz space $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$, 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 $M^1 = S_0(\mathbb{R}^d)$

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

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 the $\boldsymbol{L}^p$-spaces (and their Fourier images).

Given the very interesting properties of the Banach algebra (both with respect to convolution and pointwise multiplication), such as the Fourier invariance, one obtains quite similar properties for the dual spaces, which we denote by $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$.
We find out that

1. $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$ is a Banach space of (tempered distributions), and for any non-zero (Schwartz) window uniform convergence of $V_g(f_n)$ corresponds to norm convergence in $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$;
2. $\boldsymbol{S}_0'(\mathbb{R}^d)$ contains any $\left(\boldsymbol{L}^p(\mathbb{R}^d), \|\cdot\|_p\right), 1 \leq p \leq \infty$;
3. The Fourier transform, which leaves $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ invariant extends in a unique, w^*-w^*--continuous way to an automorphism of $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$, given via

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

BANACH GELFAND TRIPLES: a new category

Definition

A triple, consisting of a Banach space $\boldsymbol{B}$, which is dense in 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 [a unitary operator resp.] an 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 $\boldsymbol{B}_2'$.

Absolutely Convergent Fourier Series

In his studies Norbert Wiener considered the Banach algebra $\left(\boldsymbol{A}(\mathbb{T}), \|\cdot\|_{\boldsymbol{A}}\right)$ of absolutely convergent Fourier series. It was one of the early **Banach algebras**, with *Wiener's inversion Theorem* being an important first example.

Later on it was natural to study the dual space, which of course contains the dual space of $\left(\boldsymbol{C}(\mathbb{T}), \|\cdot\|_{\infty}\right)$, which by the Riesz representation theorem can be identified with the bounded (regular Borel) measures on the torus it was natural to call these functions *pseudo-measures*.

Since $\boldsymbol{A}(\mathbb{T})$ can be identified with $\boldsymbol{L}^1(\mathbb{T})$ (viewed as subspaces of $\boldsymbol{L}^2(\mathbb{T})$ and $\ell^2(\mathbb{Z})$ respectively), it is natural to expect (and prove distributionally) that $\boldsymbol{PM}(\mathbb{T})$ is isomorphic to $\ell^\infty(\mathbb{Z})$ via the (extended) Fourier transform.

Preview and First Conclusions

1. Only for "good functions" one can hope to recover them from pointwise samples;
2. Not all the "generalized functions" (including elements from $\boldsymbol{L}^p(\mathbb{R}^d)$!) can be even sampled pointwise, but only after some applying some mollifier;
3. In cases like $L^\infty(\mathbb{R}^d)$ (or more concrete exponential functions, even the function $\mathbf{1}$) cannot be approximated uniformly (i.e. in norm) by test functions, one has to allow w^*-convergence.
4. since regularizers are well understood (there are also many choices) we will concentrate on the (*mutual!*) approximation of test functions and discrete sequences!

Fourier Transforms of Tempered Distributions

His construction *vastly extends the domain of the Fourier transform* and allows even polynomials to have a Fourier tranform. Among the objects which can now be treated are also the Dirac measures δ_x, as well as **Dirac combs** $\sqcup\!\sqcup = \sum_{k \in \mathbb{Z}^d} \delta_k$.
Poisson's formula, which expresses that one has for $f \in \boldsymbol{\mathcal{S}}(\mathbb{R}^d)$

$$\sum_{k \in \mathbb{Z}^d} f(k) = \sum_{n \in \mathbb{Z}^d} \widehat{f}(n), \tag{3}$$

can now be recast in the form

$$\widehat{\sqcup\!\sqcup} = \sqcup\!\sqcup.$$

Sampling and Periodization

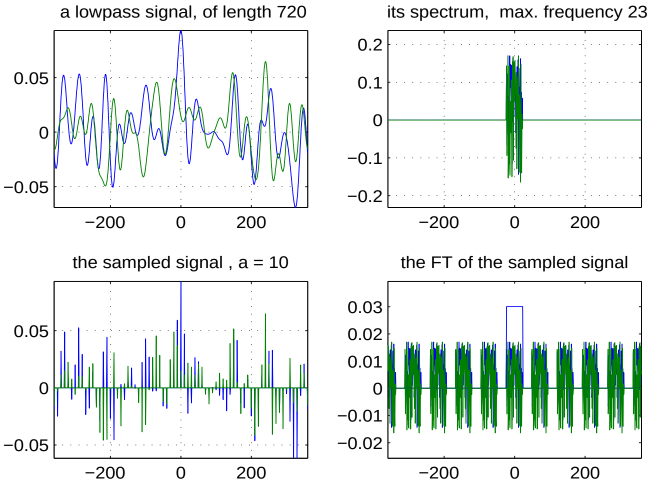
Typically one thinks of fine lattices in $\Lambda \lhd \mathbb{R}^d$ as being obtained by compression/dilation of the standard lattice, i.e. $\Lambda = \boldsymbol{A} * \mathbb{Z}^d$ ($\boldsymbol{A}$ non-singular).
If the function is smooth enough (Sobolev space) and the (regular) samples are taken fine enough (e.g. on a fine, hexagonal grid in the plane) one expects to get a very good approximation, e.g. by smooth (quasi)-interpolation of the data.
In the same way one expects that periodization of f does not produce a big error if the periodization lattice is course enough compared to the decay or concentration properties of f.
Combining the two methods (properly) one obtains a discrete and periodic signal, which in fact can be viewed naturally as a function over some finite Abelian group!

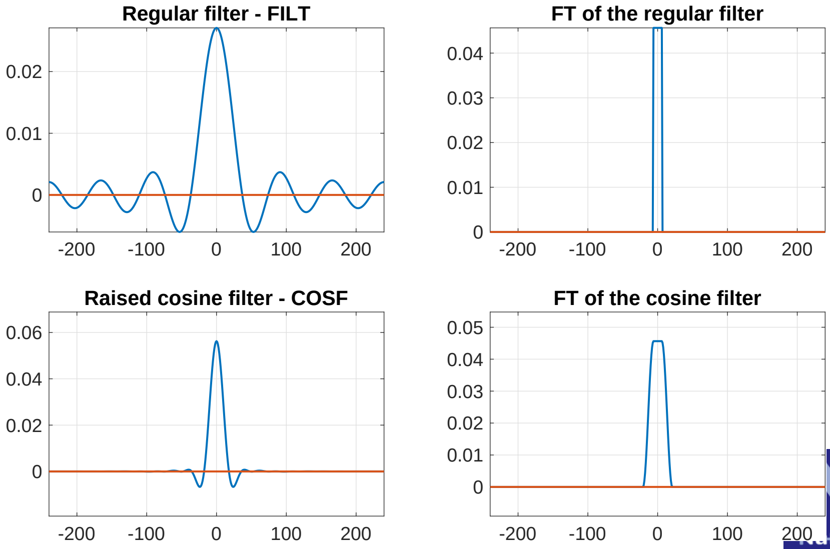
Sampling and Periodization on the FT side

This result is the key to prove **Shannon's Sampling Theorem** which is usually considered as the fundamental fact of digital signal processing (Claude Shannon: 1916 - 2001).
Shannon's Theorem says that one can have **perfect reconstruction** for band-limited functions.
If the so-called *Nyquist criterion* is satisfied (sampling distance small enough), i.e. $\mathrm{supp}(\hat{f}) \subset [-1/\alpha, 1/\alpha]$, then

$$f(t) = \sum_{k \in \mathbb{Z}^d} f(\alpha k) g(x - \alpha k), \quad x \in \mathbb{R}^d. \tag{4}$$

A Visual Proof of Shannon's Theorem




Regular filter - FILT
0.02
0.01
0
-200
-100
0
100
200
FT of the regular filter
0.04
0.03
0.02
0.01
0
-200
-100
0
100
200
Raised cosine filter - COSF
0.06
0.04
0.02
0
-200
-100
0
100
200
FT of the cosine filter
0.05
0.04
0.03
0.02
0.01
0
-200
-100
0
100
200

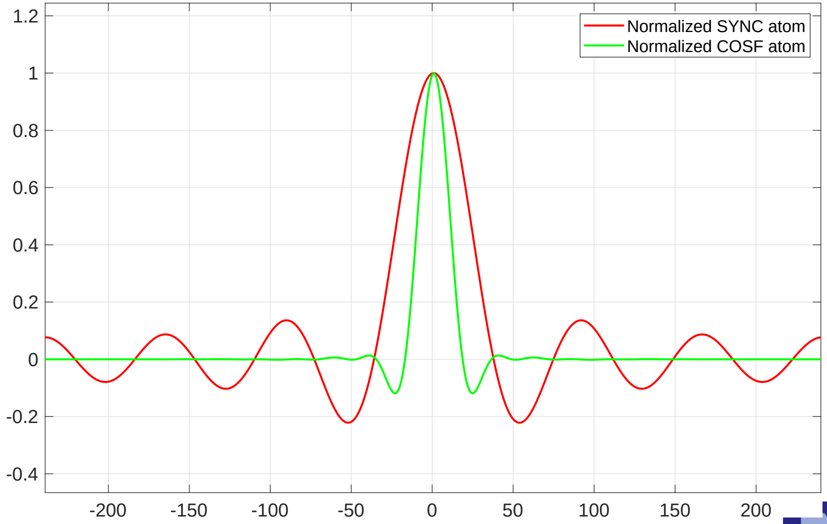


Abbildung: shannplot005.eps

It is obvious that the green graph is the better localized one. It is

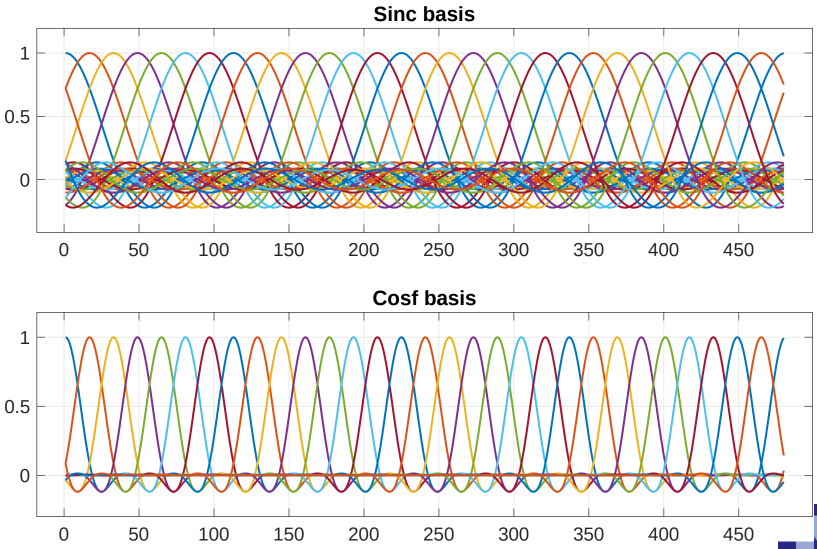


Abbildung: shannplot006.eps

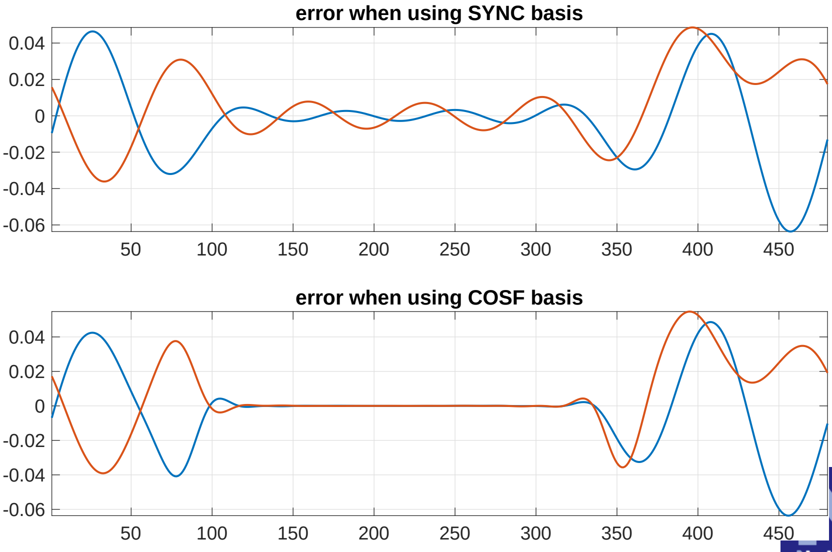


Abbildung: shannplot007.eps

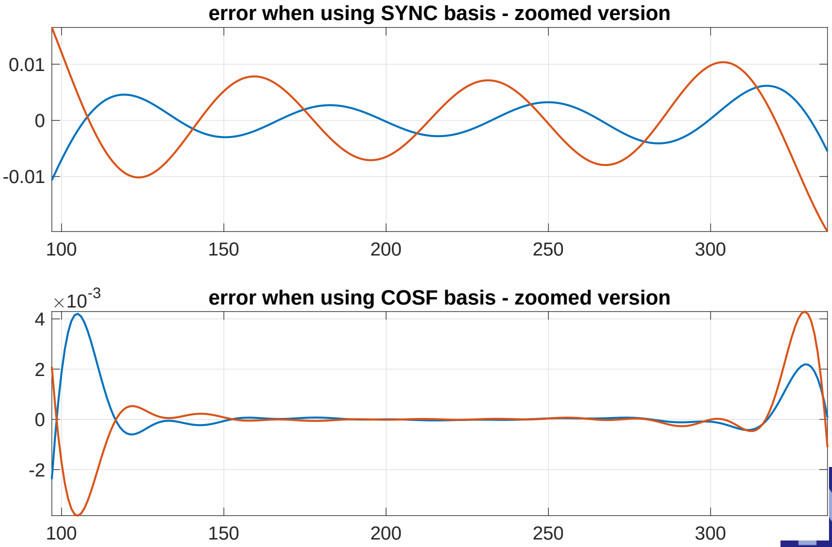


Abbildung: shannplot008.eps

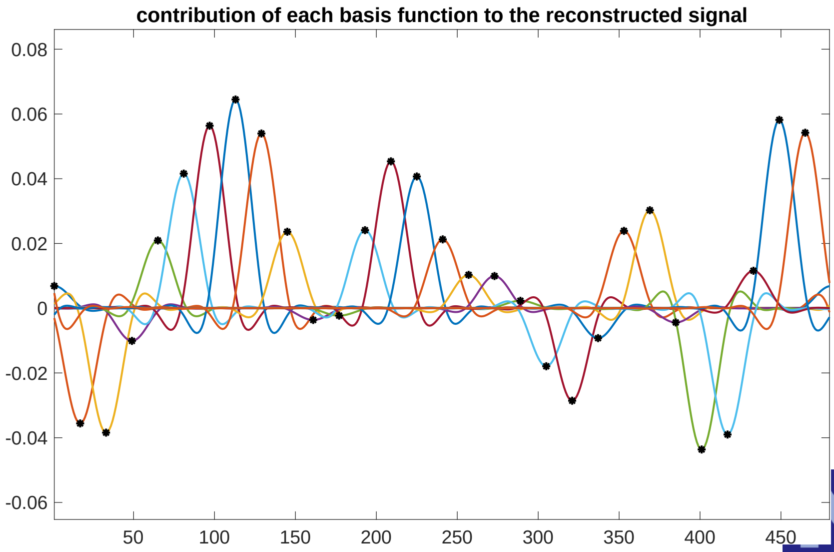


Abbildung: shannplot009.eps

How to use the Banach Gelfand Triple I

I TEND TO COMPARE THE SITUATION OF GELFAND TRIPLES WITH THE **number triple** $\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$
It is with the rational numbers where we do actually our **computations** (e.g. inversion: $3/4 \mapsto 4/3$). The real numbers have the advantage of being **complete**, but inversion is only taking place at the symbolic level: $1/\pi$ is only defined implicitly by its property: $\pi \cdot 1/\pi = 1$ (with "infinite decimal expressions"). Finally certain things (like exponential law) are better understood in the **complex domain**.

- Recall some concepts from linear algebra can be transferred to the setting of test functions in $(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0})$. Approximate recovery from regular samples (via quasi-interpolation or piecewise linear interpolation) are possible;

How to use the Banach Gelfand Triple II

- Many things (like Plancherel's) Theorem for the Hilbert space are obtained by taking limits of Cauchy sequences, resp. looking at $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$, defined via Lebesgue integration;
- Many important objects arising in signal analysis can only be understood well in the distributional setting of $\boldsymbol{S}_0'(\mathbb{R}^d)$, endowed with the w^*-convergence.
  For example, ⧢ has only a Fourier transform in the distributional sense (via Poisson formula), convolution with ⧢ turns into *periodization*, pointwise multiplication corresponds to *regular sampling*.

Now let us combine these two effects, by doing a coarse periodication (period $= n$) and a fine sampling (at rate $1/n$), providing discrete signals of length $N = n^2$.

Test functions and finite, discrete signals

Let us start with a few important observations:

- The periodic and discrete (unbounded) measures are exactly those which arise as periodic repetitions of a fixed finite sequence of the form $\sum_{k=0}^{N-1} a_k \delta_k$.
- The Fourier transform of such a sequence can be calculated directly using the FFT
- for any $f \in \boldsymbol{W}(\mathbb{R}^d)$ one has for $b = 1/a$:

$$\mathcal{F}[\sqcup\!\sqcup_{Na} * (\sqcup\!\sqcup^{a} \cdot f)] = \sqcup\!\sqcup^{Nb} \cdot (\sqcup\!\sqcup_{b} * \hat{f}) = \sqcup\!\sqcup_{b} * (\sqcup\!\sqcup^{Nb} \cdot \hat{f}) \quad (5)$$

Approximation by finite/discrete signals I

We shall define here $\sqcup\!\sqcup_a := \sum_k \delta_{ak}$ and $\sqcup\!\sqcup^a = \frac{1}{a}\sum_n \delta_{\frac{n}{a}}$. Then $\mathcal{F}\sqcup\!\sqcup_a = \sqcup\!\sqcup^a$. In fact, one has for $a = 1$ according to Poisson's formula $\mathcal{F}\sqcup\!\sqcup_1 = \sqcup\!\sqcup^1$, and the general formula follows from this by a standard dilation argument: Mass preserving compression St_ρ is converted into "value-preserving" dilation D_ρ on the Fourier transform side, and $\mathrm{D}_\rho\sqcup\!\sqcup^1 = \sqcup\!\sqcup^{1/\rho}$.

w^*-approximation

The w^*-approximation is quite crucial to describe the transition between different settings, e.g.

- periodic functions, with period going to infinity;
- sampling a continuous functions, with increasing density;
- approximation of a general (bounded and unbounded) measure by discrete measures (in fact in $\boldsymbol{W}(\boldsymbol{M},\ell^\infty)(\mathbb{R}^d) = \boldsymbol{W}(\boldsymbol{C}_0,\ell^1)(\mathbb{R}^d)')$.

The w^*- topology: approximation strategies

- How to approximate general distributions by test functions: Regularization procedures via product convolution operators, $h_\alpha(g_\beta * \sigma) \to \sigma$ or TF-localization operators: multiply the STFT with a 2D-summability kernel before resynthesis (e.g. partial sums for Hermite expansion);
- how to approximate an $\boldsymbol{L}^1$-Fourier transform by test functions: and classical summability
- how to approximate a test function by a finite disrete sequence using quasi-interpolation (N. Kaiblinger): $Q_\Psi f(x) = \sum_i f(x_i)\psi_i(x)$.

Smooth functions, sampling, discrete measures

The transition between smooth functions and discrete (regular) sequences (over a regular grid) are obtained by *sampling*, or equivalently by multiplying with a *fine Dirac comb*.
This is in fact equivalent with a periodization on the Fourier transform side.
But how can one express the fact that *dense sampling* contains a good deal of information provided by the smooth function? Of course one can (and should) think of recovery from the samples, ideally by *quasi-interpolation*.

Usually we try to approximate (in the w^*-sense) a distribution $\sigma \in \boldsymbol{S}_0'(\mathbb{R}^d)$ by a product-convolution operator or a convolution product operator, i.e.we use the following relations:

$$\sigma = w^*\text{-}\lim_{\rho \to 0} [(\sigma * \mathrm{St}_\rho g_0) \cdot \mathrm{D}_\rho g_0], \tag{6}$$

$$\sigma = w^*\text{-}\lim_{\rho \to 0} [(\sigma \cdot \mathrm{D}_\rho g_0) * \mathrm{St}_\rho g_0], \tag{7}$$

based on convolution and pointwise multiplier results for Wiener amalgam:

$$\boldsymbol{S}_0(\mathbb{R}^d) \cdot (\boldsymbol{S}_0'(\mathbb{R}^d) * \boldsymbol{S}_0(\mathbb{R}^d)) \subseteq \boldsymbol{S}_0(\mathbb{R}^d), \tag{8}$$

$$\boldsymbol{S}_0(\mathbb{R}^d) * (\boldsymbol{S}_0'(\mathbb{R}^d) \cdot \boldsymbol{S}_0(\mathbb{R}^d)) \subseteq \boldsymbol{S}_0(\mathbb{R}^d). \tag{9}$$

Proof.

The key arguments for both of these regularization procedures, be it convolution followed by pointwise multiplication (a CP or product-convolution operator), or correspondingly PC operators, are based on the pointwise and convolutive behavior of generalized Wiener amalgam spaces, such as the relation

$$\boldsymbol{S}_0(\mathbb{R}^d)*\boldsymbol{S}_0'(\mathbb{R}^d) = \boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1,\ell^1)*\boldsymbol{W}(\mathcal{F}\boldsymbol{L}^\infty,\ell^\infty) \subseteq \boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1,\ell^\infty).$$

□

The simple setting of Banach spaces (resp. dual spaces) allows also to work with a very simple, alternative way to introduce the dual space, namely to define

1. weak Cauchy-sequences $(h_n)_{n\geq 1}$ of test functions from $\boldsymbol{S}_0(\mathbb{R}^d)$, in the sense that $(\langle h, f_n \rangle)_{n\geq 1}$ is a Cauchy sequence in $\mathbb{C}$ for any $h \in \boldsymbol{S}_0(\mathbb{R}^d)$;
2. then define equivalence classes of such sequences;
3. then define the norm of such a class (it is finite by Banach-Steinhaus).

It is then not difficult to show that this is just another way (maybe more intuitive to engineers) to describe $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$!

How should we sample a continuous function?

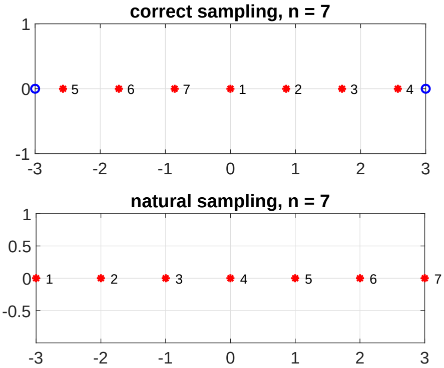


Abbildung: samppatt17A.pdf TEST777

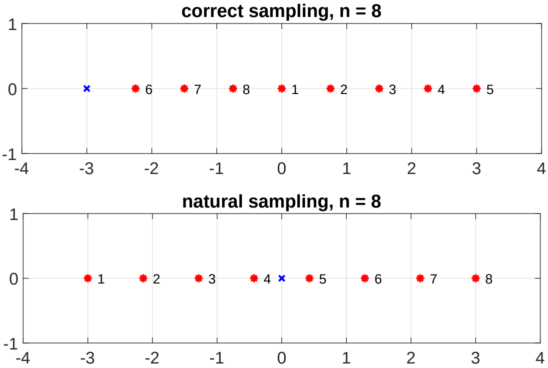


Abbildung: samppatt17B.eps

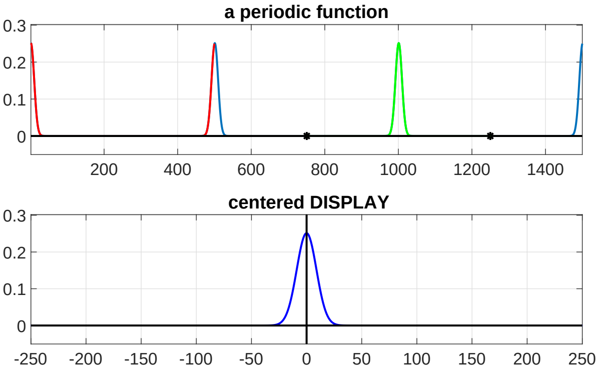


Abbildung: smpvaldem3A1.eps

Comparing signals of length n to $4n$

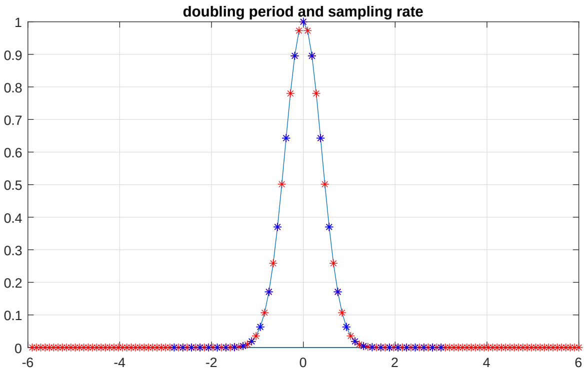


Abbildung: smpvaldem3A1.eps

Another form of Cauchy condition

While samples taken over differen periods and with different sampling rates are not immediately comparable to each other (except via interpolation/quasi-interpolation and resampling) the choice of particular sequences allow immediate comparison.
For example one may start to plot a Gauss function over $[-3, 3]$ with 144 samples taken at a rate of $1/6$.
Then this discrete (periodized) Gauss function will be invariant under the (unitary normalized) FFT.
taking now signal lengths of the form $4^k * 144$, say $n = 576$ and $n = 2304$ and then $n = 9216$ or $n = 36864$ one can then check whether discretely computed functions (e.g. Hermite functions up to some order) are *compatible*, which is in fact the case.

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

$$\langle f, g \rangle = \langle \widehat{f}, \widehat{g} \rangle \qquad (10)$$

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

Discrete Hermite functions I

One of the delicate question (arising e.g. in the discussion about finite quantum theory) is the question concerning a discrete analogue of *Hermite Functions*

One can use the these functions to define e.g. *Fractional Fourier transforms*. These (unitary) operators form a natural group under composition, isomorphic to the torus, and correspond (up to phase factors) to the rotation of the STFT (with Gaussian window, so in the *Fock space*) by a given angle.

Of course the ordinary Fourier transform corresponds to a rotation by 90 degrees, hence two times the FT corresponds to the flip operator (rotation by 180 degrees), and four times gives the identity operator.

Discrete Hermite functions II

The easiest description of the FrFT is to describe it as a *Hermite multiplier*, with a (suitable) pure frequency (over the index set $\mathbb{N}_0$). Recall, that $h_0 = g_0$, the Gauss function.
The Hermite functions form an ONB for $\left(\boldsymbol{L}^2(\mathbb{R}), \|\cdot\|_2\right)$ and are eigen-vectors for the Fourier tranform with eigenvalues $(-i)^n, n \geq 0$.
Consequently the projection from $\mathcal{H} = \boldsymbol{L}^2(\mathbb{R})$ onto the subspace of Fourier invariant elements can be realized by just projecting on the linear span of the Hermite functions of order $k = 0 mod(4)$.

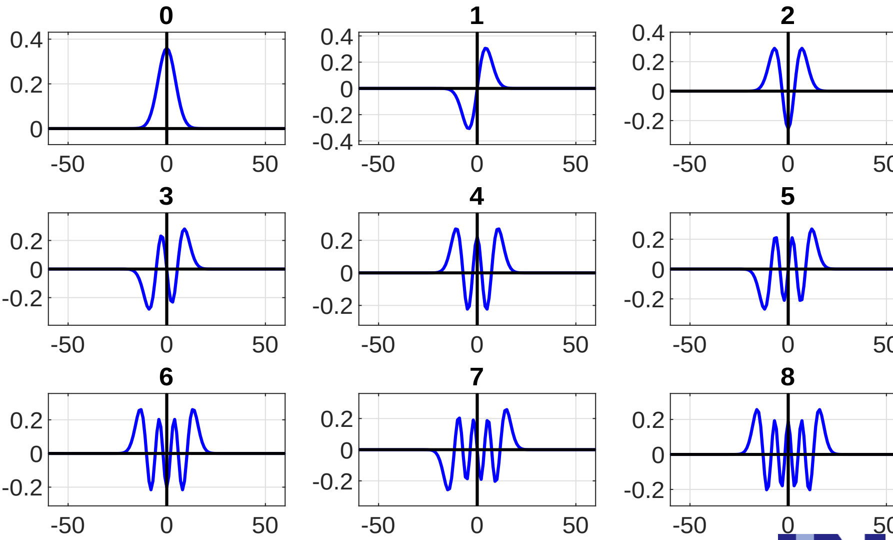


Abbildung: herm9of12.pdf

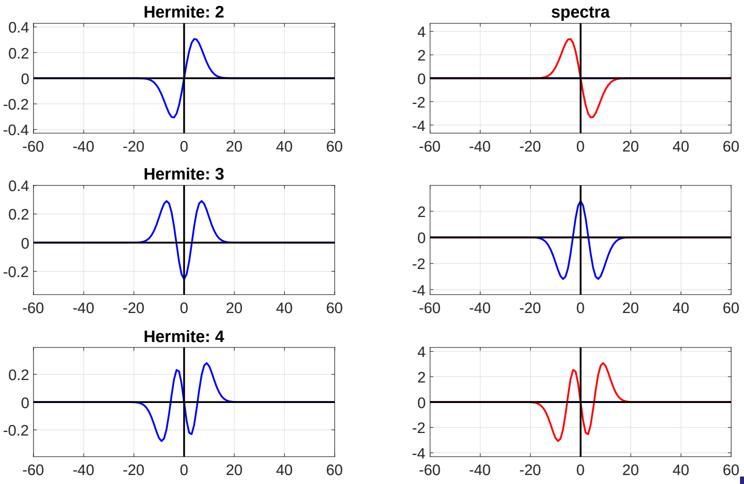


Abbildung: herm9of12B.eps

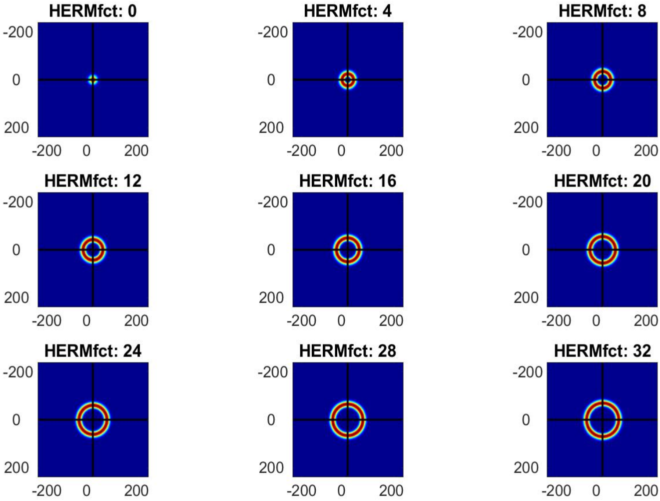
But how should these functions be defined or computed?
One can try to take the samples of the original Hermite functions, obtained by three-term recursion. By sampling the orthogonality may get lost for higher frequencies. One can also use the fact that the form an orthonormal system for the *Harmonic Oscillator*, i.e. for the operator

$$H(\psi) = \Delta(\psi) + x^2 \cdot \psi, \quad \psi \in \mathcal{S}(\mathbb{R}^d), \tag{11}$$

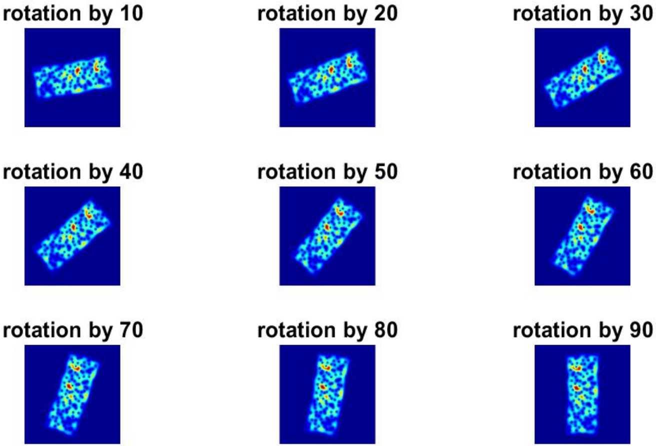
which is an unbounded, but self-adjoint operator.
Often this operator is replaced by a discrete variant of the Laplace operator (second order differences) and then the corresponding ONB for $\mathbb{C}^N$ are computed.
In fact, they form an ONB for $\mathbb{C}^N$ (by the spectral theorem for symmetric matrices), but only for lower order they are well suited as a replacement for the continuous Hermite functions.

As we found, a much more natural approach comes from another property (found e.g. in work of Ingrid Daubechies, and earlier H. Landau) on *localization operators*: The Hermite functions are also joint eigenvalues for all the STFT-multipliers (Anti-Wick operators) with radial symmetric weight. Whenever on uses a radial weight tending to zero at infinity one has a compact and self-adjoint operator!

Such operators can be realized in a discrete setting and provide actually a very useful common (!) family of discrete eigenvectors for all these Anti-Wick operators (which therefore commute!). According to N. Vasilevskii the only commutative Banach algebras of Anti-Wick operators are the shift invariant ones (e.g. time-invariant ones), or the ones which commute with fractional Fourier transforms and are consequently Hermite multipliers.


HERMfct: 0
HERMfct: 4
HERMfct: 8
HERMfct: 12
HERMfct: 16
HERMfct: 20
HERMfct: 24
HERMfct: 28
HERMfct: 32
-200
0
200
-200 0 200

A localized signal under fractional FT



frametitle: bibliography I

H. G. Feichtinger and G. Zimmermann.
A Banach space of test functions for Gabor analysis.
In H. G. Feichtinger and T. Strohmer, editors, *Gabor Analysis and Algorithms: Theory and Applications*, 1998.

H. G. Feichtinger and N. Kaiblinger.
Quasi-interpolation in the Fourier algebra.
J. Approx. Theory, 144(1):103–118, 2007.

H. G. Feichtinger and N. Kaiblinger.
Varying the time-frequency lattice of Gabor frames.
Trans. Amer. Math. Soc., 356(5):2001–2023, 2004.

H. G. Feichtinger and F. Weisz.
Inversion formulas for the short-time Fourier transform.
J. Geom. Anal., 16(3):507–521, 2006.

H. G. Feichtinger.
Choosing Function Spaces in Harmonic Analysis,
Vol. 4 of The February Fourier Talks at the Norbert Wiener Center, 2015.

frametitle: bibliography II

H. G. Feichtinger.
Elements of Postmodern Harmonic Analysis.
In *Operator-related Function Theory and Time-Frequency Analysis. The Abel Symposium 2012, Oslo, Norway, August 20–24, 2012*, pages 77–105. Cham: Springer, 2015.

H. G. Feichtinger.
Modulation Spaces: Looking Back and Ahead.
Sampl. Theory Signal Image Process., 5(2):109–140, 2006.

H. G. Feichtinger and M. S. Jakobsen.
Distribution theory by Riemann integrals. Arxiv, 2018.

M. S. Jakobsen and H. G. Feichtinger.
The inner kernel theorem for a certain Segal algebra. 2018.

M. S. Jakobsen.
On a (no longer) New Segal Algebra: A Review of the Feichtinger Algebra.
J. Fourier Anal. Appl., 24(6):1579–1660, 2018.

Book References

K. Gröchenig: Foundations of Time-Frequency Analysis, 2001.
H.G. Feichtinger and T. Strohmer: Gabor Analysis, 1998.
H.G. Feichtinger and T. Strohmer: Advances in Gabor Analysis, 2003. both with Birkhäuser.
G. Folland: Harmonic Analysis in Phase Space, 1989.
I. Daubechies: Ten Lectures on Wavelets, SIAM, 1992.
G. Plonka, D. Potts, G. Steidl, and M. Tasche. Numerical Fourier Analysis. Springer, 2018.

Some further books in the field are *in preparation*, e.g. on modulation spaces and pseudo-differential operators (Benyi/Okoudjou, Cordero/Rodino).

See also www.nuhag.eu/talks.

Conclusion and OUTLOOK I

1. For many branches of mathematical analysis and also application areas I see a strong need for more extensive numerical work, which requires constructive and realizable approaches with guaranteed convergence resp. precision.
   This means typically:
   Given some linear operator T from $(\boldsymbol{B}^1, \|\cdot\|^{(1)})$ to $(\boldsymbol{B}^2, \|\cdot\|^{(2)})$ and $\varepsilon > 0$ determine $\mathbf{y} = T(\mathbf{x})$, given $\mathbf{x}$, up to ε, or compute $\mathbf{a}$ approximately, given $T(\mathbf{a}) = \mathbf{b}$.
2. Of course, once such computational approaches have been established one should be able to answer questions of optimality (minimal resources in order to achieve this goal, sensitivity to model assumptions, computational costs, error analysis, scalability).

Conclusion and OUTLOOK II

3. In signal analysis (pseudo-differential operators, Fourier transform, etc.) one typically combines discretization with FFT-based methods. Hence the transition from say $\mathbb{R}^n$ to a suitable finite group (via sampling) and conversely (via quasi-interpolation) have to be investigated properly.
4. The Banach Gelfand Triple $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$ and more generally modulation spaces appear to be an appropriate framework, due to its Fourier invariance and the existence of a *kernel theorem*.
5. Much work is needed in order to turn the idea of Conceptual Harmonic Analysis (marriage of Abstract and Harmonic Analysis with Computational Harmonic Analysis) with life and turn it into a foundation of modern Fourier and Time-Frequency Analysis.

Conclusion and OUTLOOK III

Let us finally collect a few more recommendations:

- Good discretizations are structure preserving!
  For example: a good version of a discrete Gauss function should be (up to the normalization) be FFT invariant!
- Good discretization should show as many properties already for relatively small dimensions as close as possible to the continuous situation;
- Only then one can hope that numerical experiments and computations are supporting the theory and vice versa!
- Maybe the best estimates are not obtained for the classical function spaces such as $\left(\boldsymbol{L}^p(\mathbb{R}^d), \|\cdot\|_p\right)$, but maybe via modulation spaces $\left(\boldsymbol{M}^s_{p,q}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}^s_{p,q}}\right)$.

Thanks for your attention

maybe you visit www.nuhag.eu
and checkout other talks at
www.nuhag.eu/talks (!password)
and related material. hgfei