What is a Continuous Orthonormal Basis? How Ideas from Physics can be put on a Sound Mathematical Ground via Mild Distributions

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

Int. Prague Sem. Function Spaces
ZOOM Talk held 16.12.2025

ABSTRACT provided for this talk

When it comes to the use of the family of Dirac measures or the real line of the Euclidean space $\mathbb{R}^d$ physicist and engineers often use weird formulas, involving divergent integrals, often manipulated in a formal way, based on the analogy to the finite discrete case, where such manipulations are justified within linear algebra.
The talk is going to describe some ongoing discussion which aims at a mathematically sound description of such formal manipulations in the context of mild distributions. Mild distributions form together with the Hilbert space $\boldsymbol{L}^2(\mathbb{R}^d)$ and the underlying Banach algebra $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ of test functions (the Feichtinger algebra) a so-called Banach Gelfand triple or rigged Hilbert space, comparable with the Schwartz setting.
In this setting the claim that the extended Fourier transform is just change of basis, moving from the Dirac basis to the (equivalent) Fourier basis, becomes well justified. —

Goals of this lecture

- Background from Linear Algebra.
- Recall the view on Dirac bases.
- Discuss how problematic they are.
- Remind the concept of Banach Gelfand Triples.
- Specifically $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$.
- The Fourier transform as a basic
  unitary Banach Gelfand Triple automorphism.
- Definition of *mild* (Dirac-like) bases.
- Examples and illustrations, e.g.
  Fractional Fourier Transforms as a change of bases!

Memories of our Linear Algebra Courses I

Finite dimensional vector spaces $\boldsymbol{V}$ over $\mathbb{C}$ (or $\mathbb{R}$) allow to form *linear combinations* and thus define the concept of generating families and linear independence. If both situations occur at the same time such a family of vectors is called a basis. The number of vectors in a basis is always the same, we call it the dimension of $\boldsymbol{V}$. There are very good and not so good bases, and commonly (for good reasons) orthonormal bases in $\mathbb{C}^n$ (now endowed with the Euclidean scalar product) are the best one! On the one hand it is easy to find the (unique) coefficients, just by taking scalar products. An ONB for $\mathbb{C}^n$ corresponds to a unitary matrix $\mathbf{U}$ with columns $(\mathbf{u}_k)_{k=1}^n$ and one has for all $\mathbf{x} \in \mathbb{C}^n$ TWO formulas:

$$\mathbf{x} = \sum_{k=1}^{n} \langle \mathbf{x}, \mathbf{u}_k \rangle \mathbf{u}_k = \sum_{k=1}^{n} P_k(\mathbf{x}) \quad \text{and} \quad \|\mathbf{x}\|_{\mathbb{C}^n} = \sqrt{\sum_{k=1}^{n} |\langle \mathbf{x}, \mathbf{u}_k \rangle|^2}. \tag{1}$$

In terms of matrices we have for these $n \times n$ square (!) matrices

$$\mathbf{U} * \mathbf{U}' = \mathrm{Id}_n \quad \text{and} \quad \mathbf{U}' * \mathbf{U} = \mathrm{Id}_n\,. \tag{2}$$

We may also remember the criterion for ONBs: A basis (e.g. a linear independent family of n vectors in $\mathbb{C}^n$) is an ONB if and only if this norm equality holds true (which at the end relies on the Pythagorean Principle.)

However one has now to emphasize the it does not characterize ONBs per se. In the redundant case (more then n vectors) this property leads to the concept of **tight frames**.

For the plane ($\boldsymbol{V} = \mathbb{R}^2$) the Mercedes-Benz trio (three vectors of equal length, with mutual angles of 120 degrees) are a perfect example of a tight frame of *redundancy* $3/2$.
In general, such systems of vectors are obtained by projecting an ONB of an ambient larger space into lower dimensional Hilbert space. In finite dimensions the redundancy is the quotient of the dimensions of the involved spaces.

Hilbert Spaces and (Tight) Frames I

Taking these simple concepts to infinite dimensional Hilbert spaces by introducing the usual concepts of a total (or spanning) set, or linear independence (via finite subfamilies) is not very efficient. Also the concept of (Schauder) basis is quite delicate but more of academic than practical relevance. These concepts allow "infinitely bad" systems which do not share the properties we are used to have in the finite dimensional situation.
Stable sets of generators allow to express any vector using coefficients whose norm (in the ℓ^2-sense) is comparable to the length. A Riesz basis is good if the condition number is not too big, i.e., if the length of vectors is comparable to the length of the linear combination generated. Such a *uniform control* of the quality allows to take infinite limits and get series expansions.
In the last decades (following the development of wavelet theory and Gabor expansions) it has become clear that the correct

analogy is to discuss frames and Riesz basic sequences in abstract Hilbert spaces.
A frame is called a tight frame if (up to normalization) it has the properties of norm reproduction (**??**).
Putting vectors in $\mathbb{C}^m$ as columns into some $m \times n$-matrix $\boldsymbol{A}$ allows to claim: A matrix $\boldsymbol{A}$ is of maximal rank $r = \min(m, n)$ if and only if it's columns are linear independent (case $n \leq m$) or generate all of $\mathbb{C}^m$ ($n \geq m$, frame case). Both properties occur (basis case) if and only if $\boldsymbol{A}$ is invertible. Then there is a dual system, a *biorthogonal Riesz sequence or dual frame*, described by pinv($\boldsymbol{A}'$), the pseudo-inverse of the transpose conjugate $\boldsymbol{A}'$.

Connection to Matrix Theory I

The concepts of frames and Riesz basis sequences has a clear analogy in the context of (rectangular) matrices. Any matrix is both a linear mapping $\mathbf{c} \mapsto \boldsymbol{A} * \mathbf{c}$, producing linear combinations of the columns of $\boldsymbol{A}$ via $\sum_{k=1}^{n} c_k \mathbf{a}_k$. Thus the range of the linear mapping is the column space of the matrix, and the columns are linear independent if and only of this mapping is INJECTIVE.

In contrast, the mapping is SURJECTIVE if and only if the columns generate all of $\mathbb{C}^m$ (i.e. they form a frame there).

Note that surjectivity of $\boldsymbol{A}$ is equivalent to injectivity if $\boldsymbol{A}'$ and vice versa, where $\mathbf{y} \mapsto \boldsymbol{A}' * \mathbf{y} = \langle \mathbf{y}, \mathbf{a}_k \rangle$ is the coefficient mapping.

A tight frame $(g_i)_{i\in I}$ instead establishes (again, suitable normalized, often called a Parseval frame) an isometric isomorphism to a closed subspace of ℓ^2 via the *coefficient mapping* $C : f \mapsto \langle f, g_i \rangle_{i\in I}$. In this case the adjoint mapping from ℓ^2 to $\mathcal{H}$ is the *reconstruction operator* $R = C^* : \mathbf{c} \mapsto \sum_{i\in I} c_i g_i$ on the range, respectively realizes the pseudo-inverse mapping, and $C \circ R$ represents the projection onto the range of C. Given $f \in \mathcal{H}$ the coefficients obtained by taking $\tilde{c} = (C \circ R \circ C)(f)$ are the minimal norm least squares (MNLSQ) coefficients for the representation of $f \in \mathcal{H}$, in analogy to the valid identity (resulting from SVD):

$$\mathrm{pinv}(\boldsymbol{A}) = \mathrm{pinv}(\boldsymbol{A}' * \boldsymbol{A}) * \boldsymbol{A}' = \boldsymbol{A}' * \mathrm{pinv}(\boldsymbol{A} * \boldsymbol{A}').$$

In the maximal rank case one of them is invertible!!
(the inverse frame operator or the inverse Gramian).

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

Time-shifts and Frequency shifts

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

–

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

with $x, \omega, t \in \mathbb{R}^d$, compatible with the Fourier transform

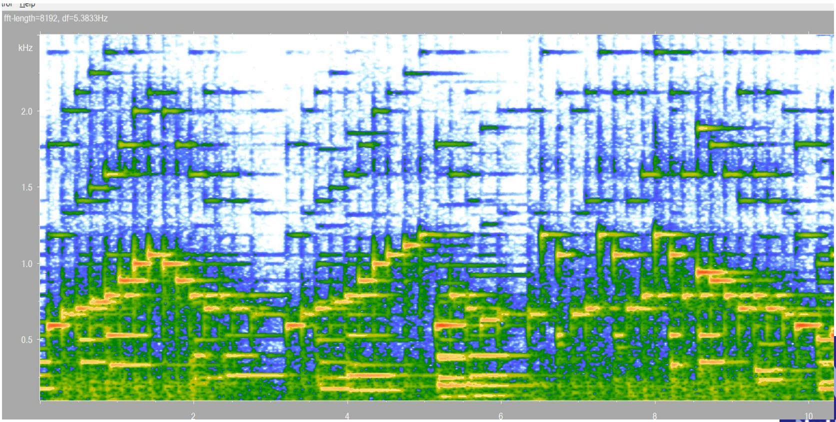
$$(T_x f)\widehat{\ } = M_{-x}\hat{f} \qquad (M_\omega f)\widehat{\ } = T_\omega \hat{f}$$

The Short-Time Fourier Transform

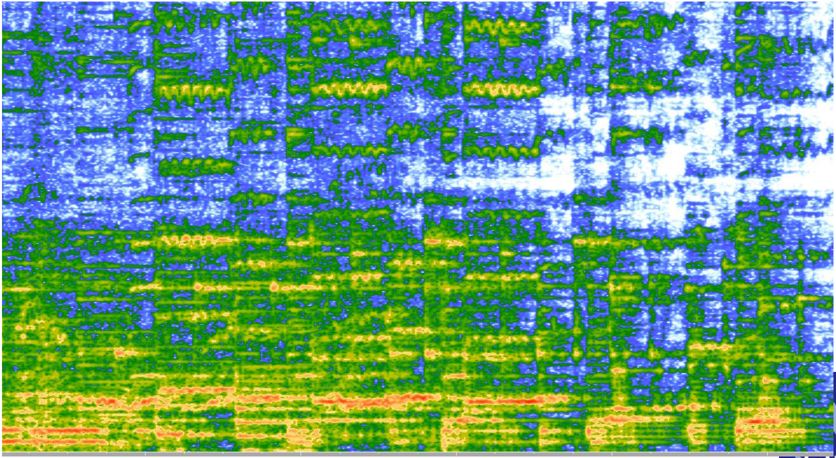
$$V_g f(\lambda) = \langle f, M_\omega T_t g\rangle = \langle f, \pi(\lambda) g\rangle = \langle f, g_\lambda\rangle,\ \ \lambda = (t, \omega);$$

A Typical Musical STFT

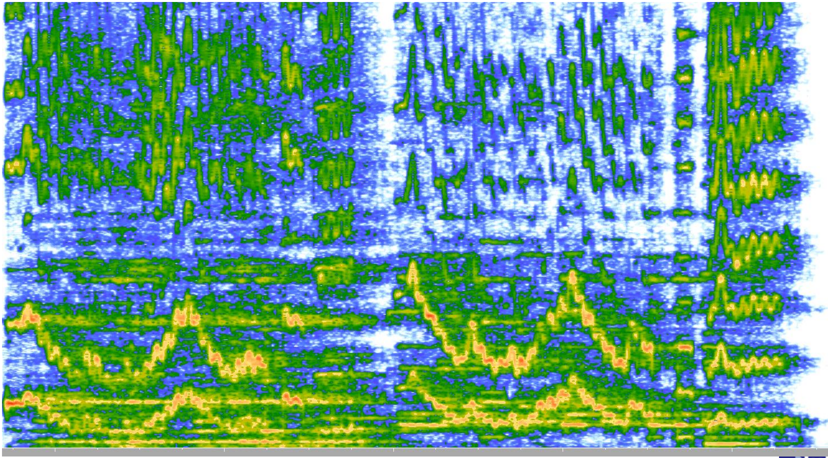
A typical piano spectrogram (Mozart), from recording



A Musical STFT: Brahms, Cello



A Musical STFT: Maria Callas



A 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.

Lemma

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

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

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

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

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

QRC The Three Layer Model

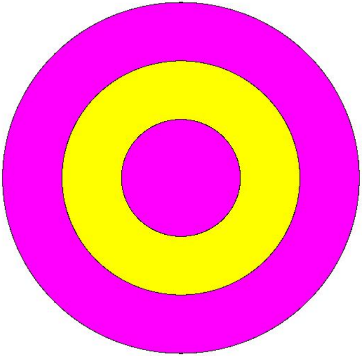


Abbildung: The THREE LAYER model, e.g. $\mathbb{Q} \hookrightarrow \mathbb{R} \hookrightarrow \mathbb{C}$
or $\ell^1 \hookrightarrow \ell^2 \hookrightarrow \ell^\infty$, or $\boldsymbol{S}_0(\mathbb{R}^d) \hookrightarrow \boldsymbol{L}^2(\mathbb{R}^d) \hookrightarrow \boldsymbol{S}_0'$

Tight Gabor Frames

Gaborian families are obtained from a fixed function, typically a symmetric, real-valued and somehow smooth function g (called the Gabor atom) by applying time-frequency shifts. Regular Gabor families arise if one takes all the TF-shifts along a lattice $\Lambda \lhd \mathbb{R}^d \times \widehat{\mathbb{R}}^d$. In such a case one can show that the dual (and also the canonical tight) frame is another Gabor frame, generated by the *dual Gabor atom* $\widetilde{g}$ (with respect to (g, Λ)) or the *canonical tight Gabor atom*. For our discussion we can thus restrict our attention to such tight Gabor atoms.
Often one takes $\Lambda = a\mathbb{Z}^d \times b\mathbb{Z}^d$. It is well known that for the choice of g as a (Fourier invariant) Gaussian $g_0(t) = \exp(-\pi|t|^2)$ it suffices to take $ab < 1$, and for $max(a, b) < 1/2$ (e.g.) the tight Gabor atom will be close to the original Gaussian.

Tight Gabor Frames I

First let us fix one such (close to Gaussian, regular, right) Gabor frame $(g^{\blacktriangle}, a, b)$ and observe:
Every complex-valued sequence in $\ell^2(\mathbb{Z}^{2d})$ defines a signal/function in $\mathcal{H} = \left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$ via synthesis:
$f := \sum_{k,n} c_{k,n} g^{\blacktriangle}_{k,n}$, with $g^{\blacktriangle}_{k,n} := M_{nb} T_{ak} g^{\blacktriangle}$.
On the other hand the mapping $f \mapsto \langle f, g^{\blacktriangle}_{k,n} \rangle_{\boldsymbol{L}^2}$ defines a bounded linear mapping from $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$ into $\ell^2(\mathbb{Z}^{2d})$, mapping f to the MNLSQ-coefficients allowing a representation of $f \in \boldsymbol{L}^2(\mathbb{R}^d)$:

$$f = \sum_{(k,n)\in\mathbb{Z}^2} \langle f, g^{\blacktriangle}_{k,n} \rangle g^{\blacktriangle}_{k,n} \quad \text{and} \quad \|f\|^2_{\boldsymbol{L}^2} = \sum_{(k,n)\in\mathbb{Z}^2} |\langle f, g^{\blacktriangle}_{k,n} \rangle|^2.$$

NOTE that the MP3 compression scheme uses a similar representation for discrete (digital audio) signals over $\mathbb{Z}$.

Alternative Approach to Mild Distributions I

First we characterize $\boldsymbol{S}_0'(\mathbb{R}^d)$ inside of $\boldsymbol{\mathcal{S}}'(\mathbb{R}^d)$:

Theorem

A tempered distribution $\sigma \in \boldsymbol{\mathcal{S}}'(\mathbb{R}^d)$ belongs to $\boldsymbol{S}_0'(\mathbb{R}^d)$ (or equivalently: extends continuously from the dense subspace $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ of $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ to all of $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$) if and only if $V_g(\sigma)(\lambda) := \sigma(\pi(\lambda)g)$ is a bounded and continuous function on $\mathbb{R}^d \times \widehat{\mathbb{R}}^d$, and for any non-zero $g \in \boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ the expression $\|V_g(\sigma)\|_\infty$ defines an equivalent norm on $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$.
Moreover, w^*-convergence of a sequence $\sigma_n \to \sigma_0 \in \left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ is equivalent to the POINTWISE convergence of $V_g(\sigma_n)(\lambda)$ to $V_g(\sigma_0)(\lambda)$ (in fact, uniform convergence over compact subsets of $\mathbb{R}^d \times \widehat{\mathbb{R}}^d$!).

Since clearly $\left(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\right)$ is thus a subspace of $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$ (and hence any *almost periodic function*, etc.) the above characterization allows to derive two possible approaches to the space $\boldsymbol{S}_0'(\mathbb{R}^d)$ of mild distributions via a *sequential approach*. The objects of the extended domain for the Short-Time Fourier Transform are (equivalence classes of) so-called *mild Cauchy sequences* (in short **ECmiCS**).

Definition

Fix any non-zero $g \in \boldsymbol{C}_c(\mathbb{R}^d)$ with $\hat{g} \in \boldsymbol{L}^1(\mathbb{R}^d)$. A sequence $(h_n)_{n\geq 1}$ in $\boldsymbol{C}_b(\mathbb{R}^d)$ is called a *mild Cauchy sequence* if the sequence $\|V_g(h_n)\|_\infty$ is bounded and for every pair $(t,s) \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d$ we a Cauchy sequence: $(V_g(h_n)(s,t))_{n\geq 1}$.

The uniform continuity of the functions V_{g_n} implies that

$$\exists\, H(s,t) = \lim_{n\to\infty} V_g(h_n)(s,t) \tag{3}$$

defines a STFT for the "limit" $\|H\|_\infty \leq \sup_{n\geq 1} \|V_g(h_n)\|_\infty$.

Starting from these mild Cauchy sequences one can do the same thing as with the construction of $\mathbb{R}$ from $\mathbb{Q}$: One has to identify Cauchy sequences which "represent the same object", by determining **equivalence classes** of Cauchy sequences.
By identifying these equivalence classes with the limit of their STFTs we can thus claim that every such equivalence class has a bounded STFT and this determines a unique element in $\boldsymbol{S}_0'(\mathbb{R}^d)$.
IN SHORT: can be identified with with the space of all equivalence of mild Cauchy sequences! using

$$\boldsymbol{S}_0 \cdot (\boldsymbol{S}_0 * \boldsymbol{S}_0') \subset \boldsymbol{S}_0 \subset \boldsymbol{C}_b.$$

The Euclidean Situation

In the Euclidean Situation interesting objects are for example the Dirac trains supported by lattices $\Lambda \lhd \mathbb{R}^d$:

$$\sqcup\!\sqcup := \sum_{\lambda \in \Lambda} \delta_\lambda .$$

These (unbounded, discrete) measures belong to $\boldsymbol{S}_0'(\mathbb{R}^d)$ and thus have a distributional Fourier transform. In fact, the usual argument in distribution theory making use of Poisson's formula (now not only for $f \in \boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ but for any $f \in \boldsymbol{S}_0(\mathbb{R}^d)$!) gives ($\Lambda = \mathbb{Z}^d$):

$$\widehat{\sqcup\!\sqcup} = \sqcup\!\sqcup . \tag{4}$$

This fact allows to prove *Shannon's sampling Theorem* (based on intuitive arguments).

Sampling and Periodization on the FT side I

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

$$f(t) = \sum_{k \in \mathbb{Z}^d} f(\alpha k) g(x - \alpha k), \quad x \in \mathbb{R}^d. \tag{5}$$

For more general sampling lattices the support size conditions have to avoid overlapping of the periodization of the spectrum with respect to the orthogonal lattice $\Lambda^{\perp}$.

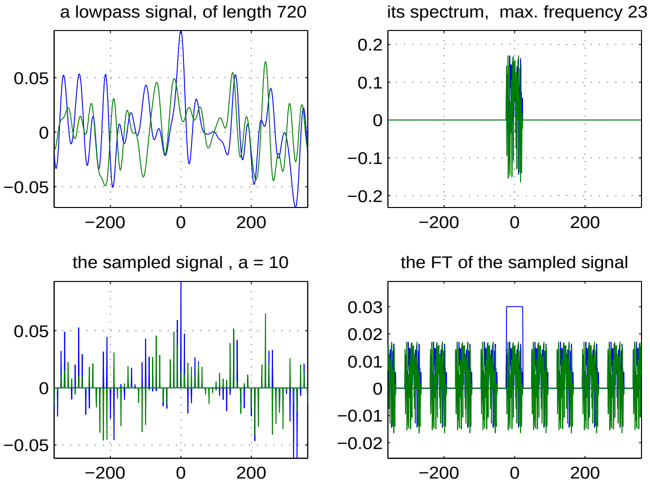
Recalling that the atom $g \in \boldsymbol{L}^1(\mathbb{R})$ is chosen in such a way that $\widehat{g}(s) \equiv 1$ on $\Omega = \mathrm{supp}(\widehat{f})$ we have $f = f * g$, and since the pointwise evaluation of the convolution can be realized as inner product and since in our case $\widehat{g}$ is real-valued, thus $g(x) = g) - x)$, we have

$$f(\alpha k) = f * g(\alpha k) = \langle f, T_{\alpha k} g \rangle$$

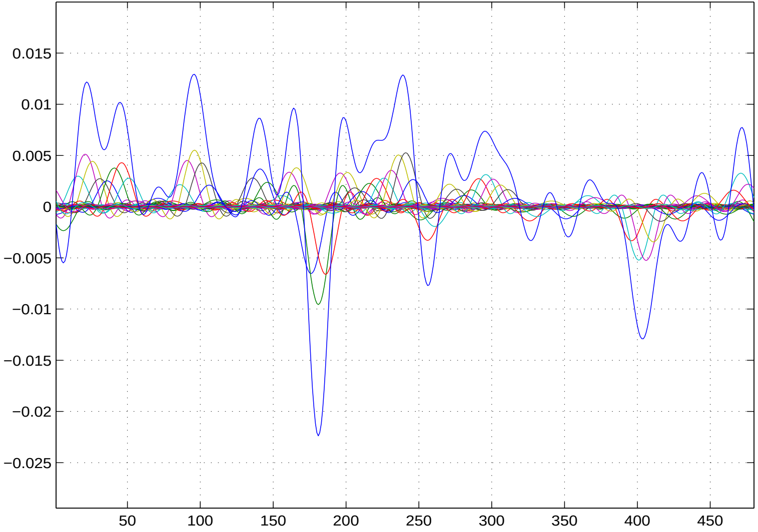
which implies the following version of the Shannon Theorem, valid for the full range of $p \in [1, \infty]$, at the cost of a bit oversampling:

$$f = \sum_{\lambda \in \Lambda} \langle f, T_\lambda g \rangle T_\lambda g, \quad \mathrm{supp}(\widehat{f}) \subseteq \Omega. \tag{6}$$

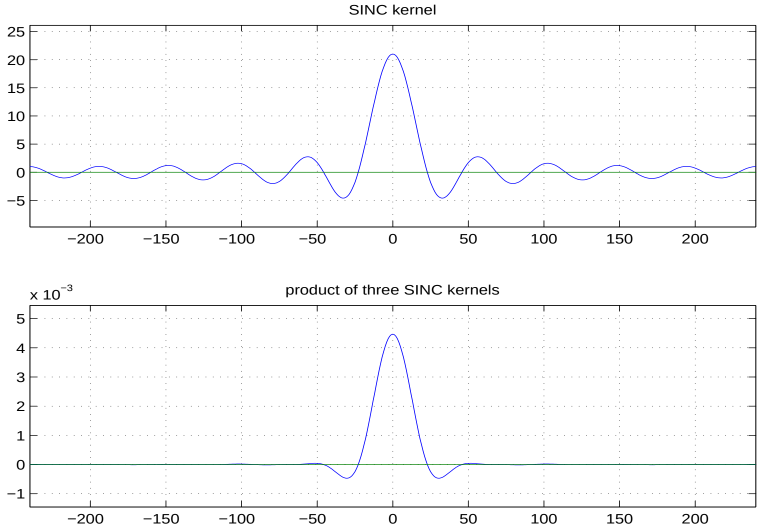
A Visual Proof of Shannon's Theorem



Non-locality of the Classical Shannon Theorem



Non-locality of the Classical Shannon Theorem



Decomposition of Identity I

In any of these case one could say that the tight frame elements $g_j, j \in J$ constitute a decomposition of the identity as a sum of projection operators. In fact, writing

$$P_j(f) = \langle f, g_j \rangle g_j$$

we can characterize tight frames by the fact that

$$\mathrm{Id}_{\mathcal{H}} = \sum_{j \in J} P_j.$$

Of course we assume typically that the family of atoms have equal norm, often they are normalized as with $\|g_j\|_{\mathcal{H}} = 1$. This suggests to consider the "continuous ONBs" as the integral over the rank one operators!

Physicists Notation I

Physicists (and sometimes engineers) use this model to describe the Fourier transform in a similar way, calling the family of Dirac measures $(\delta_t)_{t\in\mathbb{R}^d}$ a continuous orthonormal basis. The alternative (kind of complementary basis) is the Fourier basis, consisting of the pure frequencies $(\chi_s)_{s\in\mathbb{R}^d}$, given as complex exponential functions $\chi_s(t) = \exp(2\pi its)$, with $t, s \in \mathbb{R}^d$.
The main purpose of this talk is to explain that the setting of mild distributions will allow us to give a mathematically justified approach to such considerations.
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:

$$f = \int_{-\infty}^{\infty} f(t)\delta(x-t)dt = \int_{-\infty}^{\infty} \langle f, \delta_t\rangle \delta_t dt,$$

The Fourier Inversion Theorem I

Another, more tricky case, which will direct us towards a better understanding of the situation is the Fourier inversion theorem. Looking at the Fourier transform and its inversion formula we find that we have in a compact description:

$$\widehat{f}(s) = \langle f, \chi_s \rangle, \quad f \in \boldsymbol{L}^1(\mathbb{R}^d),$$

with $\chi_s(t) = \exp(2\pi i s t)$, the pure frequencies. and

$$f(t) = \int_{\mathbb{R}^d} \widehat{f}(s)\chi_s(t)dt, \quad \widehat{f} \in \boldsymbol{L}^1(\mathbb{R}^d),$$

or again in compact form writing $Q_s(f) = \langle f, \chi_s \rangle \chi_s$:

$$\mathrm{Id} = \int_{\mathbb{R}^d} Q_s ds. \quad (!!???)$$

THE Banach Gelfand Triple (BGT) $(\boldsymbol{S_0}, \boldsymbol{L^2}, \boldsymbol{S_0'})(\mathbb{R}^d)$

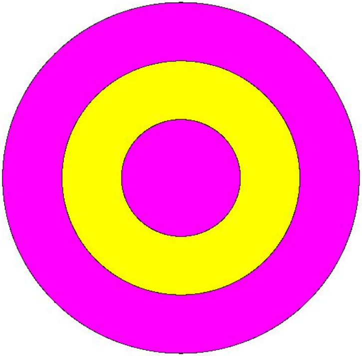


Abbildung: $\boldsymbol{S_0}(\mathbb{R}^d) \hookrightarrow \boldsymbol{L}^2(\mathbb{R}^d) \hookrightarrow \boldsymbol{S_0'}$

Analogy to Wiener's Algebra (Abs. Conv. FS) I

This last formula is quite similar to the expansion of a periodic function in $\left(\boldsymbol{L}^2(\mathbb{T}), \|\cdot\|_2\right)$ with respect to is Fourier coefficients. Engineers call these building blocks (characters) the "pure frequencies", typically of the form $\chi_k(t) = \exp(2\pi i k t)$. The form an ONB for $\left(\boldsymbol{L}^2(\mathbb{T}), \|\cdot\|_2\right)$ and thus we have for any $f \in \mathcal{H} = \left(\boldsymbol{L}^2(\mathbb{T}), \|\cdot\|_2\right)$:

$$f = \sum_{k \in \mathbb{Z}} \langle f, \chi_k \rangle \chi_k = \sum_{k \in \mathbb{Z}} \widehat{f}(k) \chi_k \quad \text{and} \quad \|f\|_{\boldsymbol{L}^2}^2 = \sum_{k \in \mathbb{Z}} |\widehat{f}(k)|^2,$$

with unconditional norm convergence and Fourier coefficients in $\ell^2(\mathbf{Z}_N)$ (by Parseval's identity).

A similar statement, bringing tight frames into the picture in a classical setting would be the representation of band-limited signals in Shannon's Theorem in the oversampling regime.

Wiener's Algebra of Abs. Conv. FS

In this setting N. Wiener was introducing the subspace corresponding to $\ell^1(\mathbb{Z}_N)$, namely $\left(\boldsymbol{A}(\mathbb{T}), \|\cdot\|_{\boldsymbol{A}}\right)$, the space of *absolutely convergent Fourier series*, with
$\|f\|_{\boldsymbol{A}} = \|(\widehat{f_k})\|_{\ell^1} = \sum_{k\in\mathbb{Z}} |\widehat{f}(k)|$.
Since such Fourier series converge uniformly and absolutely in $\left(\boldsymbol{L}^2(\mathbb{T}), \|\cdot\|_2\right)$ the resulting functions are continuous and pointwise represented by their Fourier series. Moreover we have $\left(\boldsymbol{A}(\mathbb{T}), \|\cdot\|_{\boldsymbol{A}}\right) \hookrightarrow \left(\boldsymbol{L}^2(\mathbb{T}), \|\cdot\|_2\right)$, as a dense subspace. By duality, we have $\left(\boldsymbol{L}^2(\mathbb{T}), \|\cdot\|_2\right) \hookrightarrow \boldsymbol{PM}(\mathbb{T})$, the dual of $\left(\boldsymbol{A}(\mathbb{T}), \|\cdot\|_{\boldsymbol{A}}\right)$, the space of pseudo-measures on the torus.
Via duality (or a minimal form of distribution theory) one can identify $\boldsymbol{PM}(\mathbb{T})$ with $\ell^\infty(\mathbb{Z})$ via the extended Fourier transform.

Tight Gabor Frames I

First let us fix one such (close to Gaussian, regular, right) Gabor frame $(g^{\blacktriangle}, a, b)$ and observe:
Every complex-valued sequence in $\ell^2(\mathbb{Z}^{2d})$ defines a signal/function in $\mathcal{H} = \left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$ via synthesis:
$f := \sum_{k,n} c_{k,n} g^{\blacktriangle}_{k,n}$, with $g^{\blacktriangle}_{k,n} := M_{nb} T_{ak} g^{\blacktriangle}$.
On the other hand the mapping $f \mapsto \langle f, g^{\blacktriangle}_{k,n}\rangle_{\boldsymbol{L}^2}$ defines a bounded linear mapping from $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$ into $\ell^2(\mathbb{Z}^{2d})$, mapping f to the MNLSQ-coefficients allowing a representation of $f \in \boldsymbol{L}^2(\mathbb{R}^d)$:

$$f = \sum_{(k,n)\in\mathbb{Z}^2} \langle f, g^{\blacktriangle}_{k,n}\rangle g^{\blacktriangle}_{k,n} \quad \text{and} \quad \|f\|^2_{\boldsymbol{L}^2} = \sum_{(k,n)\in\mathbb{Z}^2} |\langle f, g^{\blacktriangle}_{k,n}\rangle|^2.$$

NOTE that the MP3 compression scheme uses a similar representation for discrete (digital audio) signals over $\mathbb{Z}$.

Going to the Gabor Setting I

Going by analogy (kind of reverting the historical development of concepts), based on our current knowledge about Gabor expansions we could ask:

- What about functions with Gabor coefficients in $\ell^1(\mathbb{Z}^{2d})$?
- Is it reasonable to look at sums with coefficients in $\ell^\infty(\mathbb{Z}^{2d})$?
- Aside from the problems arising using a fixed (tight) Gabor family, what happens if one takes another lattice, another atom?
- Are there chances for computationally approximating the (infinitely many) coefficients in the right norm?
- How to make use of such representation (Gabor multipliers, time-variant filtering, etc.)?

Some Answers I

- The functions which arise with coefficients in $\ell^1(\mathbb{Z}^{2d})$ form a Banach space, even a Banach algebra, which is even invariant under the Fourier transform!
- The "distributions" which can be characterized with bounded coefficients can be characterized as the dual space, called the space of mild distributions.
- These spaces do not depend on the choice of the Gabor atom (e.g. any Schwartz window) and the corresponding (quotient) norms are equivalent.
- One can view the elements in the dual space as "signals" $\sigma \in \boldsymbol{S}_0'(\mathbb{R}^d)$, and the evaluation on test functions $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ as "measurements": $\sigma \mapsto \sigma(f), \quad f \in \boldsymbol{S}_0(\mathbb{R}^d)$.
- Natural convergence of sequences in $\boldsymbol{S}_0'(\mathbb{R}^d)$ is mild convergence: $\lim_{n\to\infty} \sigma_n(f) = \sigma_0(f), \quad \forall f \in \boldsymbol{S}_0(\mathbb{R}^d)$.

The Kernel Theorem I

This concept appears to allow a natural extension of the setting of $(\boldsymbol{A}^2, \boldsymbol{L}^2, \boldsymbol{PM})(\mathbb{T})$ to the Euclidean case.
Mild convergence corresponds to w^*-convergence, which in that case is equivalent to coordinate-wise convergence on the Fourier side.
There is also a kernel theorem, which allows to identify linear operators from $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ to $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$ with "kernels" $K(x,y)$ in $\boldsymbol{S}_0'(\mathbb{R}^{2d})$.
If $K(x,y)$ is a bounded and continuous function on $\mathbb{R}^{2d}$, then we have a bounded linear operator T_K, given by

$$Tf(x) = \int_{\mathbb{R}^d} K(x,y)f(y)dy, \quad f \in \boldsymbol{S}_0(\mathbb{R}^d).$$

In general one has $\langle Tf, g\rangle = \langle K, f \otimes g\rangle$.

The Fourier Inversion Theorem

Let us now consider the identity $\mathrm{Id} = \int_{\mathbb{R}^d} Q_s ds$.
We are going to show that this is valid on $\boldsymbol{S}_0(\mathbb{R}^d)$.
For $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ the mapping $s \mapsto \langle f, \chi_s \rangle$ is well defined, and $\widehat{f} \in \boldsymbol{S}_0(\mathbb{R}^d) \subseteq \boldsymbol{C}_0(\mathbb{R}^d)$.
Of course one would like to interpret the integral as a limit of Riemannian sums. Note that these Riemannian sums are of the form $\sum_{k \in F} c_k \widehat{f}(s_k) \chi_{s_k}$ i.e. this are trigonometric polynomials (which do not belong to $\boldsymbol{S}_0(\mathbb{R}^d)$, nor to $\boldsymbol{L}^2(\mathbb{R}^d)$.) But in the sense of mild convergence one just has to look at the action on test functions $g \in \boldsymbol{S}_0(\mathbb{R}^d)$, so we have to look at terms $\lim_F \sum c_k f(s_k) g(s_k) = \int_{\mathbb{R}^d} f(t) g(t) dt = \langle \mathrm{Id}_{\boldsymbol{S}_0(\mathbb{R}^d)} f, g \rangle$, or

$$\langle \left(\int_{\mathbb{R}^d} Q_s ds \right) f, g \rangle = \int_{\mathbb{R}^d} \langle f, \chi_s \rangle \langle \chi_s, g \rangle = \langle f, g \rangle, \quad f, g \in \boldsymbol{S}_0(\mathbb{R}^d).$$

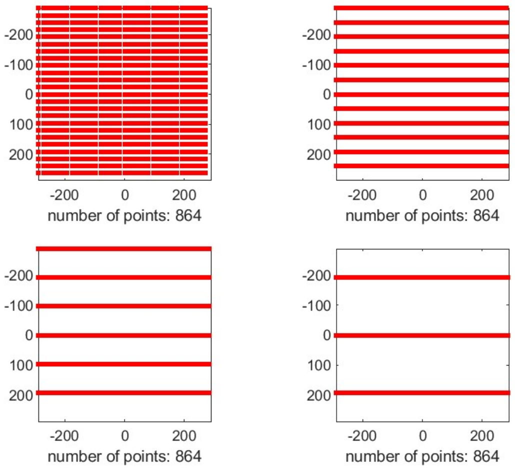


Abbildung: Four lattices with equal redundancy

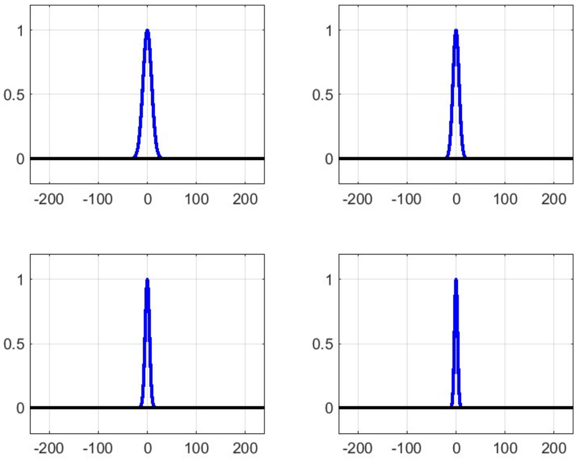


Abbildung: Corresponding better localized Gabor atoms

SUMMARY: We can think of the Dirac family as the "continuous limit" of a collection of tight Gabor lattices obtained by compression of the time axis.
While physicists would say that the Dirac basis consists of infinitely small Gabor atoms which are aligned along an infinitely fine (hence continuous) lattice we can say: For the test against any finite set of test functions the family of compressed version of any tight Gabor frame will have shrinking lattice constant $a \to 0$, and correspondingly $b \to \infty$, so that finally the approximation of a given function in $\boldsymbol{S}_0(\mathbb{R}^d)$ is of a given quality once the sampling rate is high enough (compare to ideas in wavelet theory, scale space approximation!).

Why is "continuous ONB" a bad notation!?

While it is clear that the family of Dirac measures $(\delta_x)_{x \in \mathbb{R}^d}$ is a family of mild distributions (in fact, of course a bounded family of bounded measures) it *may be OK* to argue that δ_x is orthogonal to δ_y, of one tries to compute the scalar product formally by taking limits of smoothed versions.

But δ_0 cannot be normalized, and that the scalar product of δ_0 by itself equals (by usual methods $+\infty$ does not indicate that we have an orthonormality condition.

Finally, there is some built in redundancy, which one would not expect to occur for a true "basis"!

Links and Refs

Let us provide a few links:

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

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

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