C:\coursfei\ws2021\ETHExtra21A.tex

hgfei: VERSION of 01.01.2021

1 Recalling some basic facts from linear algebra

I assume that everybody is familiar with the basics of *linear algebra*, including matrix multiplication. The usual view on matrix multiplication is to think of $\mathbf{x} \mapsto \mathbf{y} := \boldsymbol{A} * \mathbf{x}$ as having a mapping which maps the *column vector* in $\mathbb{R}^n$ (or $\mathbb{C}^n$) into the vector $\mathbf{y} \in \mathbb{R}^m$ (or $\mathbb{C}^m$), assuming that $\boldsymbol{A}$ is an $m \times n$-matrix.
However there is an alternative viewpoint: if we consider $\boldsymbol{A}$ as a collection of n column vectors, $\mathbf{a}_1, .., \mathbf{a}_n \in \mathbb{C}^m$ and $\mathbf{x} \in \mathbb{C}^n$, then we can form the linear combination, and we observe that

$$\boxed{\boldsymbol{A} * \mathbf{x} = \textstyle\sum_{k=1}^n x_k \mathbf{a}_k \ (!)} \qquad (1)$$ matrmult1

In fact, this is equivalent to the (well known? but) easy to verify fact that for unit vectors $\mathbf{e}_k = [0, 0, ..1, ..0]$, with the 1 in coordinate number k one has

$$\boxed{\boldsymbol{A} * \mathbf{e}_k = \mathbf{a}_k, \quad 1 \leq k \leq n.} \qquad (2)$$ matrmult2

As we also know, matrix multiplication is defined in such a way that one can say that for a matrix $\boldsymbol{B} = [\mathbf{b}_1, ..., \mathbf{b}_k]$ of size $k \times m$ one has with $\boldsymbol{C} = \boldsymbol{B} * \boldsymbol{A}$ a matrix of size $k \times n$ (mapping $\mathbb{C}^n$ via $\mathbb{C}^m$ common dimension (!Domino Principle) into $\mathbb{C}^k$, thus mapping vectors in $\mathbb{C}^n$ into $\mathbb{C}^k$:

$$\boxed{\boldsymbol{C} = [\boldsymbol{B} * \mathbf{a}_1, .., \boldsymbol{B} * \mathbf{a}_n]} \qquad (3)$$ matrmult

In WORDS, for the most important special case of vectors in $\mathbb{C}^m$: if we form n linear combinations of the unit vectors in $\mathbb{C}^m$ (the columns of $\boldsymbol{A}$) and then use these vectors in order to produce linear combinations of these linear combinations (using the columns of $\boldsymbol{B}$ as coefficients) we can recompute them as linear combinations of the **original vectors**, with the coefficients represented by the columns of the product matrix $\boldsymbol{C}$!! So we have the motto:

Linear combinations of linear combinations of vectors are just (other) linear combinations

added 20.11.2020: Let us mention in passing that multiplication with $\boldsymbol{A}'$ (MATLAB: transpose conjugate, or transpose matrix for real matrices) acts on from $\mathbb{C}^m$ back to $\mathbb{C}^n$, but it can be viewed as a right matrix multiplication on row vectors in $\mathbb{C}^m$. As such it describes the adjoint action on the dual space of $\mathbb{C}^m$, with

$$\langle \mathbf{y}, \boldsymbol{A} * \mathbf{x} \rangle_{\mathbb{C}^m} = \langle \mathbf{y} * \boldsymbol{A}', \mathbf{x} \rangle_{\mathbb{C}^n}, \quad \mathbf{x} \in \mathbb{C}^n, \mathbf{y} \in \mathbb{C}^m. \qquad (4)$$ adjointmatrix

We also have (usual matrix multiplication):

$$\boxed{\boldsymbol{A}' * \mathbf{y} = (\langle \mathbf{y}, \mathbf{a}_k \rangle)_{k=1}^n, \quad \mathbf{y} \in \mathbb{C}^m.} \qquad (5)$$ scalprods

Then the question is: *Can we get the original n vectors back, as linear combinations of the new vectors?*, i.e. can we revert the procedure of taking linear combinations? Now, if these original vectors are linear independent, then there is only a *single representation* of each vector, namely as 1-times the vector itself, and all other vectors have to come with a zero, otherwise one would get that vector as linear combinations of some others, which is not possible (by the characterization of linear independence).

The special case m=n, linear mapping $\mathbb{R}^n \to \mathbb{R}^n$

Let us discuss an important special case: Assume we have n vectors in $\mathbb{C}^n$ and take n linear combinations with coefficients stored in an $n \times n$-matrix $\boldsymbol{A}$. So the first column $\mathbf{a}_1$ of $\boldsymbol{A}$ describes the coefficients of the first linear combination to be built, the second column contains the information about the second linear combination, and so on.
As we will see we are talking about the *inversion of a square matrix*. Note that the inverse of an element $\boldsymbol{A}$ in the matrix algebra $\mathcal{M}_{n,n}$ is another $n \times n$-matrix $\boldsymbol{B}$ with

$$\boldsymbol{B} * \boldsymbol{A} = \mathbf{Id}_n = \boldsymbol{A} * \boldsymbol{B}.$$

It is a standard observation that such a matrix $\boldsymbol{B}$ is *uniquely determined* if it exists.
More surprisingly, it is implied by one of the relations, namely

$$\boldsymbol{B} * \boldsymbol{A} = \mathbf{Id}_n \quad \text{or} \quad \mathbf{Id}_n = \boldsymbol{A} * \boldsymbol{B}. \tag{6}$$ invmatrnn

This fact belongs to the basic results of linear algebra. In the first case the columns of $\boldsymbol{A}$ must span $\mathbb{R}^n$ and hence (since $m = n$) they are a basis for $\mathbb{R}^n$, in the second case the columns of $\boldsymbol{A}$ span $\mathbb{R}^n$, and hence have to be linear independent. In each case we obtain that there exists some inverse and consequently a *left inverse* or a *right inverse* is already an inverse in the sense of (6).
For the case of the unit vector basis in $\mathbb{R}^3$, (to make it more concrete), namely $\mathbf{e}_1 = [1, 0, 0]; \mathbf{e}_2 = [0, 1, 0]$ and $\mathbf{e}_3 = [0, 0, 1]$ and a 3×3-matrix $\boldsymbol{A}$ this comes down to the question whether there is a 3×3-matrix $\boldsymbol{B}$ such that

$$\boldsymbol{B} * \boldsymbol{A} = \mathbf{1}_3$$

with $\mathbf{1}_3$ being the unit matrix of size 3:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{7}$$

It turns out that this is possible only if and only if $\boldsymbol{A}$ is an *invertible* matrix, and that then of course $\boldsymbol{B} = \boldsymbol{A}^{-1}$ is the matrix to be found.
Even more interesting is the case when the given vectors are a basis for some vector space (so they are linear independent and span the n-dimensional space). Then the new vectors, obtained as linear combinations of the given vectors, using an invertible (!) matrix $\boldsymbol{A}$, are *also a basis* of the same space.
Now assume that we want to have a vector in the given space, how can we get its coordinates in the new basis?? For now we will not pursue this path further, but we want to introduce a concrete and intuitive example:

lyshift2.jpg

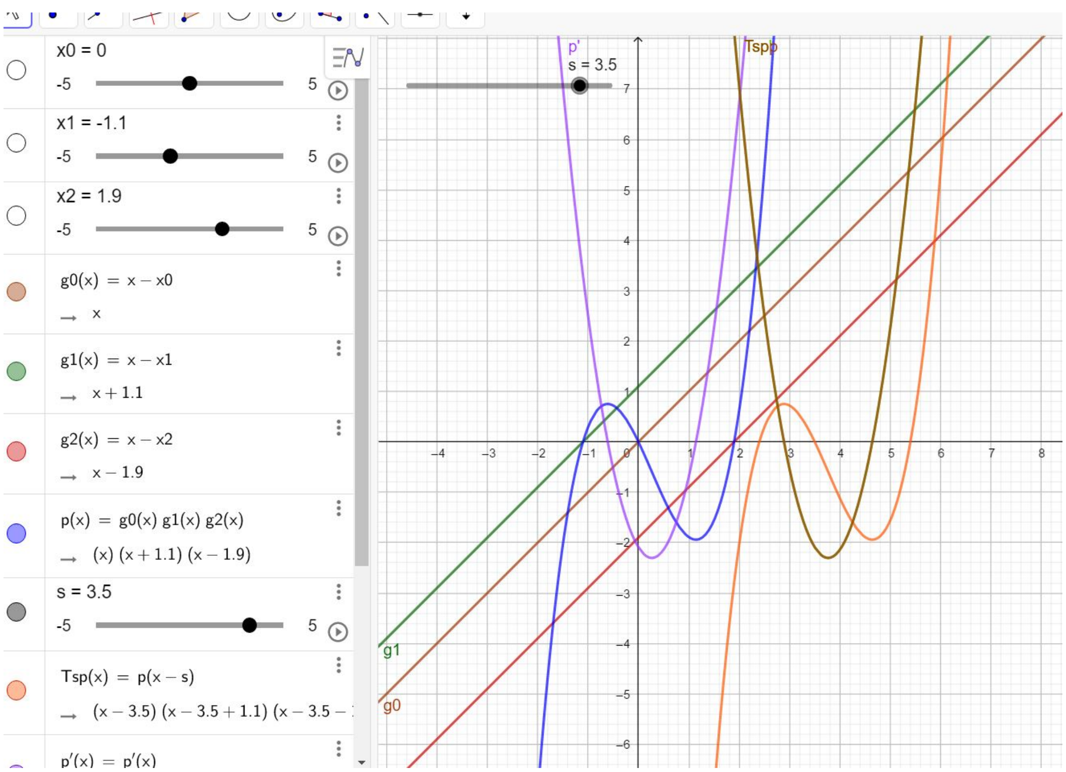

Abbildung 1: polyshift2.jpg

$\mathscr{P}_3(\mathbb{R})$, the space of cubic polynomial functions on the real line

(with real or complex coefficients, for illustration we will show of course pictures of such polynomials with real coefficients, using GEOGEBRA).It may be a good exercise for the students to do the analog story for quadratic polynomial functions, i.e. $\mathscr{P}_2(\mathbb{R})$.
There are many well known facts about such polynomials. First of all each of them has exactly 3 zeros in the complex plain, counting multiplicities. In fact, one zero x_1 has to be real (since a truly cubic polynomial tends to plus and minus infinity, as $x \to \infty$ and $x \to -\infty$, for a positive leading coefficient!). The rest (obtained by division) through $(x - x_1)$ is a quadratic polynomial.
It is clear that $\mathscr{P}_3(\mathbb{R})$ is a vector space (it is closed under addition and scalar multiplication, hence under the formation of linear combinations). It is four-dimensional, with the standard basis $\{1 = x_0, x, x^2, x^3\}$, the so called *monomial basis*, or if we take time t as a variable $\mathscr{M}_3 = \{1, t, t^2, t^3\}$, in contrast to the MATLAB convention for `polyval` which uses $\mathscr{M}^3 = \{x^3, x^2, x, 1\}$, hence `polyval([0,1,2,3],10) == 123`, and $p(x) = x^3 - x$ has the coefficient sequence $[1, 0, -1, 0]$ in this 'coordinate system'.
MATLAB suggestion:

```
for jj=1:10; plot(bas,polyval(rand(1,4)-0.5, bas));
figure(gcf); pause(1); end;
```

The space $\mathscr{P}_3(\mathbb{R})$ has many useful properties which allows to illustrate some of the more abstract key ideas which require basic ideas from linear functional analysis.

While vector analysis just works with column and row-vectors, and matrices and adjoint matrices, we will try to use $\mathscr{P}_3(\mathbb{R})$ as a *prototype* of a vector space which allows many interesting linear operators, it is translation invariant, and so on. We will think of $\mathscr{P}_3(\mathbb{R})$ (we could also use the infinite dimensional space $\mathscr{P}(\mathbb{R})$ of ALL polynomials, but for now let us stay with the 4-dimensional special case).
The first thing we are considering is the so-called *dual space*, we write $\mathscr{P}_3(\mathbb{R})'$ or $\mathscr{P}_3(\mathbb{R})^*$, the space of all (bounded) linear functionals on $\mathscr{P}_3(\mathbb{R})$. We can think of these functionals σ as linear measurements:

$$\sigma(\alpha p(x) + \beta q(x)) = \alpha\sigma(p(x)) + \beta\sigma(q(x)).$$

Now let us give simple examples:

1. The Dirac measure at $x_0 \in \mathbb{R}$: $\delta_{x_0} : p(x) \mapsto p(x_0)$, δ_0 or simple δ, the *Dirac Delta*.
2. another functional is $\sigma : p(x) \mapsto p'(y)$, we will write δ'_y.
3. for any pair $a < b$ the integral is such a function:
$$p(x) \mapsto \int_a^b p(x)dx.$$
4. one may also put weights and consider, for some Riemann integrable function k the functional
$$p(x) \mapsto \int_a^b p(x)k(x)dx.$$
5. as it turned out (ca. 100 years ago, searching for the most general integral, see the Riesz representation theory) one can use Riemann-Stieltjes integrals of the form
$$p(x) \mapsto \int_a^b p(x)dFx,$$
where F is a function of *bounded variation*, i.e. a difference of two monotonic functions, the increasing part of F and the decreasing part of F[1].

Clearly there are infinitely many such linear functionals on $\mathscr{P}_3(\mathbb{R})$, and any linear combination of such linear functionals is again a linear functional, i.e. we can form $\sigma = 3\sigma_1 + 5\sigma_2$ and so on. So we got another vector space, the space of linear functionals, we think of *linear measurements* on $\mathscr{P}_3(\mathbb{R})$.
But since with the choice of a basis we can identify $\mathscr{P}_3(\mathbb{R})$ with $\mathbb{R}^4$, using the basis $\mathscr{M}_3 = \{1, t, t^2, t^3\}$, or simply by the identity

$$\boxed{p(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3} \quad (8) \quad \boxed{\texttt{polymath1}}$$

or, since **MATLAB** does not allow the index/coordinate 0:

$$\boxed{p(x) = a_1 + a_2 x + a_3 x^2 + a_4 x^3} \quad (9) \quad \boxed{\texttt{polyvec1}}$$

or according to MATLAB conventions

$$\boxed{p(x) = b_1 x^3 + b_2 x^2 + b_2 x + b_4,} \quad (10) \quad \boxed{\texttt{polymatl}}$$

with $p(x) == polyval(\mathbf{b}, x)$ and $\mathbf{b} = flip(\mathbf{a})$ (reverse order)!

[1]Think of walking in the mountains

We also have **linear operators** on the vector space $\mathscr{P}_3(\mathbb{R})$, e.g.

1. translation operator : $T_a : p(x) \mapsto p(x-a)$;
2. dilation operators: $D_r : p(x) \mapsto p(rx)$;
3. derivative operator: $D : p(x) \mapsto p'(x)$

Matrix Representation of Shift operator

Right shift:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ -3 & 1 & 0 & 0 \\ 3 & -2 & 1 & 0 \\ -1 & 1 & -1 & 1 \end{pmatrix} \tag{11}$$

Left shift:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ 3 & 2 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \tag{12}$$

It is clear that the inverse operator for T_z is T_{-z} (the case $z = 1$ is displayed above), and correspondingly the matrix representing the inverse mapping (which "undoes" the effect of the forward mapping) is just the *matrix inverse* (the inverse matrix) of the matrix representing the operator.
In the same basis $\mathscr{M}^3$ the dilation operator with dilation factor $r > 0$ is just $diag(((ones(1,4)* r).(3:-1:0)))$. For $r = 2$ this looks as follows.

$$D2 = \begin{pmatrix} 8 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{13}$$

Numerical realization of the matrix product $D2 * H \neq H * D2$.
By recalling that the differential operator $p(x) \mapsto p'(x)$ maps x^k to kx^{k-1}, for $k \geq 1$ we have the following matrix representation

$$D = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \tag{14}$$

which commutes in fact with H and gives $H * D == D * H$

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 6 & 2 & 0 & 0 \\ 3 & 2 & 1 & 0 \end{pmatrix} \tag{15}$$

which confirms the observation made by the GEOGEBRA experiment (and of course analytic derivation, which we do not have to do here).

Let us recall some other basic fact from linear algebra:

$$\boxed{\text{Any linear mapping } T : \mathbb{R}^n \mapsto \mathbb{R}^m \text{ is represented by some } m \times n\text{-matrix } \boldsymbol{A}}$$

.

This $m \times n$-matrix is obtained by putting the images of the unit vectors $\mathbf{e}_k, 1 \leq k \leq n$ as the columns of $\boldsymbol{A}$, i.e. $\mathbf{a}_k := T(\mathbf{e}_k)$, $1 \leq k \leq n$.

If we consider a linear mapping from $\sigma : \mathbb{R}^n \mapsto \mathbb{R} = \mathbb{R}^1$ we call it a *linear functional* on $\mathbb{R}^n$, which consequently gives a uniquely determined row vector

$$\mathbf{b} := [b_1, .., b_n] := [\sigma(\mathbf{e}_1), ..., \sigma(\mathbf{e}_n)]$$

with the property

$$\sigma(\mathbf{x}) = \sigma\left(\sum_{k=1}^{n} x_k \mathbf{e}_k\right) = \sum_{k=1}^{n} x_k \sigma(\mathbf{e}_k) = \sum_{k=1}^{n} x_k b_k = \langle \mathbf{x}, \mathbf{b} \rangle. \tag{16}$$

linfunct1

For the case of complex numbers one would have to take the conjugate of $\sigma(\mathbf{e}_k)$, since $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{k=1}^{d} x_k \overline{y_k}$.

Using the basis $\mathscr{M}^3$ of monomials in decreasing order (MATLAB convention) and thus identifying $p(x)$ with its coefficient sequence according to equation (10) polymat1 the Dirac delta, defined by

$$\delta(p(x)) = \delta_0(p) = p(0) = b_4$$

hence the "representing" vector for this functional is just $[0, 0, 0, 1] = \mathbf{e}_4 \in \mathbb{R}^4$,because $\langle \mathbf{b}, \mathbf{e}_4 \rangle = b_4$.

Exercise 1. Compute the "representing" sequences $\mathbf{b} \in \mathbb{R}^n$ for one of the other linear functionals listed above, e.g., for $h > 0$ the local average functional over a symmetric interval of length h:

$$p(x) \mapsto \frac{1}{2h} \int_{-h}^{h} p(y) dy. \tag{17}$$

locaver

Exercise 2. In order to test that differentiation commutes with translations one *can* build the corresponding matrices. Let us just show the code for the differential operator, based on the well known fact that $D(x^k) = kx^{k-1}, k \geq 1$.

```
% DIFFOP.M  HGFei 24.09.2020, \ETH course
% Usage:  Dop = diffop(deg);
function  Dop = diffop(deg);
if nargin == 0; deg = 5, end;  % disp(diffop(deg)); end;
Dop =  zeros(deg+1);
for jj=1:deg; Dop(jj+1,jj) =  deg+1-jj; end;
```

Exercise 3. Determine the nullspace of the differential operator (describe it, what it the dimension). What is the *characteristic polynomial* (`poly(D)` gives the coefficients). Is this matrix orthogonally diagonalizable?

Exercise 4. Write a MATLAB program which allows to establish the matrix for the monomial basis for any degree and any shift parameter (positive or negative). Then test, whether the usual composition law, $T_a \circ T_b = T_{a+b}$ is valid in the matrix representation, and $(T_a)^{-1} = T_{-a}$.

ystem02b.jpg

Abbildung 2: system02b.jpg

Illustration of the linear system : $p(x) \mapsto T(p)(y) = q(y) = p'(y)/5 + p(y+4)$ applied to the polynomial function $p(x) = x(x-1)(x+1)$.
The graph to the left shows the output of the "linear system" and also indicates that shifting the input (from blue to green) results in a similar shift of the output signal (from red to brown).
We will demonstrate that it is enough to understand for such a system $T : p(x) \mapsto q(x)$ how the mapping can be described, knowing only the mapping (in fact linear functional)

$$\sigma = \sigma_T : \; p(x) \mapsto q(0) = [T(p)](0) \tag{18}$$ linfunct

and that the behavior of the whole system, which is **translation invariant** (we therefore write **TILS**: Translation Invariant Linear System[2]) can be described knowing only this functional. As we will see it is almost the impulse response of the system, but this has to be discussed separately.
In fact, we can describe the system as a "moving average", which comes from the fact that the functional listed in the exercise above as (locaver 17) will result in the system

$$p(x) \mapsto q(y) = \frac{1}{2h} \int_{y-h}^{y+h} p(x)dx. \tag{19}$$ moveaver00

The following THEOREM will serve as a motivation for a correct discussion of the term *impulse response* of a system. In particular it will make clear, that we have to deal with three types of objects: the signals (here $p(x) \in \mathscr{P}_3(\mathbb{R})$), the operators (or systems) on the space of signals (here and later on denoted by $T : \mathscr{P}_3(\mathbb{R}) \to \mathscr{P}_3(\mathbb{R})$), and certain linear

[2] As a synonym for linear operator (here on $\mathscr{P}_3(\mathbb{R})$), which *commutes with translations.*

functionals (here $p(x) \mapsto T(p)(0)$, later related, in combination with a flip-operator, to the impulse response).

polysys00 **Theorem 1.** *Assume that $T : \mathscr{P}_3(\mathbb{R}) \to \mathscr{P}_3(\mathbb{R})$ is a linear mapping which* **commutes with translations**[3]*, i.e. $T \circ T_z = T_z \circ T, \forall z \in \mathbb{R}$. Then there exists a (uniquely determined) linear functional $\sigma = \sigma_T \in \mathscr{P}_3(\mathbb{R})^*$ such that the output of the system can be described (as moving average, i.e.) by*

$$\boxed{T(p)(y) = q(y) = \sigma(T_{-y}p), \quad \forall y \in \mathbb{R}.} \tag{20}$$ polysysrep1

Engineers would probably write $q(y) = \sigma(p(. + y))$.

Proof. In order to relate the behavior of the *system* at 0 to the position at the given argument y we have to move y to 0. In fact

$$q(y) = q(0 + y) = [T_{-y}q](0).$$

By assumption (now comes the **crucial identity!** which will be part of examination):

$$\boxed{q(y) = [T_{-y}(T(p))](0) = [(T_{-y} \circ T)(p)](0) = [T \circ T_{-y}(p)](0) = [T(T_{-y}p)](0) = \sigma_T(T_{-y}p)} \tag{21}$$ crucial00

by the definition of the functional σ, and WE ARE DONE. □

Next we make the simple observation that the concatenation of TILS gives another TILS (!Exercise)

compTILS1 **Lemma 1.** *Given two translation invariant linear systems T_1, T_2 the concatenation of these two systems, in any order, give again translation invariant systems.*
Assume that T_1 is "represented by σ_1 and T_2 is "represented by σ_2 then there exists some $\sigma \in \mathscr{P}_3(\mathbb{R})^*$ such that

$$T_1(T_2(p))(y) = \sigma(T_{-y}p), \quad \forall y \in \mathbb{R}. \tag{22}$$ concatsigm

Proof. We leave the detailed proof to the reader. First one has to verify that the composition of linear operators gives again a linear operator. Then one has to (iteratively) verify that the composition of two translation invariant operators has again this property. □

WE ARE AT A CRITICAL POINT OF THIS EXAMPLE! We could go on and **define** the **convolution of two such functionals** by setting $\sigma_1 \star \sigma_2 := \sigma$, as in the lemma above. It would be clear that we have the associative law (coming from the associative law of concatenation of linear operators), so

$$\sigma_1 \star (\sigma_2 \star \sigma_3) = (\sigma_1 \star \sigma_2) \star \sigma_3$$

and also e.g. distributive laws, such as (we omit brackets!)

$$\sigma_1 \star (\alpha\sigma_2 + \beta\sigma_3) = \alpha\sigma_1 \star \sigma_2 + \beta\sigma_1 \star \sigma_3,$$

or other obvious properties, such as

$$\alpha(\sigma_1 \star \sigma_2) = (\alpha\sigma_1) \star \sigma_2 = \sigma_1 \star (\alpha\sigma_2).$$

What else could we ask for? [Please think ahead. 23.09, hgfei]

[3] We write $T \circ S$ for the *concatenation* of operators, i.e. $[T \circ S](f) = T(S(f))$.

We also have to provide a converse, i.e. in order to establish the fact that there is indeed a one-to-one correspondence between such functionals and TILS. We have to take essentially the following three steps:

1. Show that relation (20) polysysrep1 can be used as a definition for a TILS, i.e. given a functional it defines a linear operator which commutes with translations!

2. Show that every system arises in this way. Of course one expects that the functional which should be used (given the TILS) should be given by (just put $y = 0$ in (20) polysysrep1)

$$\sigma(p) = [Tp](0), \quad p \in \mathscr{P}_3(\mathbb{R}). \tag{23}$$

polysysrep2

3. Show that the representation is unique, or that different systems correspond to different functionals.

We will do this in the more applied context, but it is another good `Exercise` to fill in details! The hard part may be to check what is it that has to be *proved*.
Now let us look at some of the possible examples.

1. First we look at the identity operator. Then we get

$$\sigma_{Id}(p(x)) = (\mathrm{Id}(p))(0) = p(0) = \delta_0(p).$$

In other words: The Dirac delta (functional) is "representing" the identity operator, in fact, we have (to confirm)

$$(\mathrm{Id}(p))(y) = p(y) = \delta_0(T_{-y}p) = p(0 - (-y)) = p(y).$$

Another viewpoint is to move the Dirac delta around, but the only and most natural way to move a functional, i.e. to DEFINE $T_y\sigma$ is via the rule

$$\boxed{[T_y\sigma](p) := \sigma(T_{-y}f).} \tag{24}$$

shiftsigm1

In the standard engineering literature the above relationship is expressed as an *important fact*, called the **sifting property** of the Dirac Delta, often written as

$$\int_{-\infty}^{\infty} f(t)\delta(t-y)dt = f(y) \Leftrightarrow [T_y\delta_0](f) = \delta_0(T_{-y}f) = f(0 - (-y)) = f(y). \tag{25}$$

sifting02

In this sense we can also say, a linear operator T is a TILS if and only if

$$T_y \circ T \circ T_{-y} = T, \quad \forall y \in \mathbb{R}.$$

This corresponds to the fact that instead of moving the averaging (or pointwise evaluation process in the current case) to the point $y \in \mathbb{R}$ one moves the polynomial ($p(x)$ of rather $q(y)$) back to the position zero, then applies the functional, and then moves the information obtained in this way back to the position y.

OVERALL the translation invariance establishes some inherent *coherence* (or group invariance property) between the collection of point evaluations at the different points, applied to the output.

The example (signals modeled as cubic polynomials) allows us to go through the same process **in a uniform way**, covering the case of continuous or discrete, periodic or non-periodic (thus including the case of discrete and periodic, i.e. finite vectors)!!

In order to illustrate THIS approach to convolution (as a rule which allows to understand it as a way to combine two such moving averages by composing them). Later on we will speak about *filters* and also have to establish the fact that the order (!surprisingly at this point) does NOT matter, but this has to be verified.
In order to indicate the relationship between systems T and their corresponding functional σ we will occasionally use symbols such as T_σ (the system generated by the functional σ) or σ_T (the functional arising from the system). In this sense we have

$$T = T_{\sigma_T}, \quad \sigma = \sigma_{T_\sigma}.$$

A key point to this observation is the fact, that simple *translation operators* correspond to Dirac Deltas.
So let us first assume that we start from the functional $\sigma = \delta_{t_0}$. Then the corresponding TILS (and vice versa) can be computed by applying our general correspondence rule polysysrep1 (20). We obtain[4]

$$q(y) = Tp(y) = \delta_{t_0}(T_{-y}p) = p(t_0 - (-y)) = p(y + t_0) = T_{-t_0}p(y),$$

or described in an argument-free way

$$T_{\delta_{t_0}}(p) = T_{-t_0}(p), \quad t_0 \in \mathbb{R}, p \in \mathscr{P}_3(\mathbb{R}), \tag{26}$$ polysysdelto1

or even more compact: The system generated from δ_{t_0} is a shift operator:

$$\boxed{T_{\delta_{t_0}} = T_{-t_0}, \quad t_0 \in \mathbb{R}.} \tag{27}$$ polysysdelto

Next we look at the convolution of two point measures (Dirac measures).

convDirac1 **Lemma 2.** *For any two values t_1 and t_2 one has*

$$\delta_{t_1} \star \delta_{t_2} = \delta_{t_1+t_2}.$$

Proof. The system which corresponds to each of these functionals is a simple translation operator. Since $T_{-t_1} \circ T_{-t_2} = T_{-(t_1+t_2)}$ it is clear that the functional representing the composed operator has to be $\delta_{t_1+t_2}$ □

This observation has the immediate consequence that the convolution of point measures is commutative, i.e. does not depend on the order of the factors. Of course this property is also valid for two functionals which are both finite discrete measures, i.e. of the form $\sigma = \sum_{k=1}^{n} c_k \delta_{t_k}$.
Observe the analogy to matrix representation and matrix multiplication: Due to the linearity of a given mapping T one can reduce the description of a mapping T to the knowledge of a set of basis vectors, which is stored as columns in matrix format, we need n of them, from the target space $\mathbb{C}^m$, of course. Given two such operators, which can be composed, result in another operator, hence one needs an efficient way to compute the new matrix. This is how matrix multiplication is designed. For TILS we even reduce it further (essentially one row of that matrix), and need a new composition, called convolution. We will later discuss cyclic convolution of vectors (representing periodic discrete signals).

[4]Observe that $Tp(y)$ can only be interpreted as $[T(p)](y)$, because T applies only to polynomials, the same at the end of the line, $[T_{-t_0}(p)](y)$ is the only possible reading!.

Let us illustrate the situation for such linear systems from the MATLAB point of view, looking at the linear system:

$$p(x) \mapsto T(p)(y) = 2p'(x) + p(x-1) - p''(x)$$

We can ask, what is the row vector (i.e. which linear functional) that corresponds to this system, i.e. which linear mapping do we have from the coefficients of the input signal $p(x)$ to the output signal $q(y)$ (the linearity of T is clear!), and in addition we can ask, whether this system is invertible. In other words, how can we compute the coefficients of the input signal which is needed in order to obtain as the output given cubic polynomial. By just computing the output for the input vectors in the family $\mathscr{M}^3$, expressed in standard coordinates (again from highest to lowest exponent), we find that the matrix representing T with respect to the basis $\mathscr{M}^3$ is

$$TT = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ -3 & 2 & 1 & 0 \\ -1 & -1 & 1 & 1 \end{pmatrix} \tag{28}$$ reprTT

Since this is a *lower triangular* matrix with non-zero entries it is clear that $\det(TT) = 1 \neq 0$ (product of the diagonal elements). We may compute the inverse, and because of the fact that $1/\det(T) = 1$ is an integer it will be integer-valued, in fact we get

$$TT^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -3 & 1 & 0 & 0 \\ 9 & -2 & 1 & 0 \\ -11 & 3 & -1 & 1 \end{pmatrix} \tag{29}$$

Finally we have $[TT(x^k)(0) = [-1,-1,1,1]$ (which is of course just the last row of the matrix for TT), given the input $p(x)$ as a sequence of coefficients.
We have (see GEOGEBRA figure) the polynomial

$$p(x) = x(x-1)(x+1) = x^3 - x, \quad \text{or coefficients } [1,0,-1,0];$$

Applying the matrix TT to this column vector gives us (check with MATLAB) [1,3,-4,-2], i.e. the output is supposed to be of the form $q(y) = y^3 + 3y^2 - 4y - 2$ (of course one should also check it by ordinary differentiation).
Doing the same for the other system and plotting the result over $[-7,2]$, and then adjusting the height of the plot we obtain the following figure, confirming the correctness of computations done in a similar way!

ystemML2.pdf

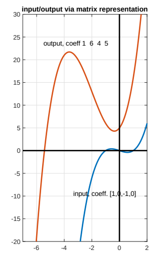

Abbildung 3: systemML2b.pdf: MATLAB experiment confirming GEOGEBRA observations: translation invariance of operation, directly computed within GEOGEBRA, using derivatives and shift operators

Concluding this section let us summarize our finding in the coordinate representation of both the operator TT and the corresponding linear function σ_{TT}

$$TT = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ -3 & 2 & 1 & 0 \\ -1 & -1 & 1 & 1 \end{pmatrix} \tag{30}$$ reprTT1

This is the matrix representation for TT given in (**??** represTT) with respect to the monomial basis (in MATLAB, i.e. in decreasing order). In other words, the columns contain the output $TT(p(x))$, for $p(x) = x^3$ in the first row and so on.
For the associated functional σ_{TT} we have to look at the value of the same output only at the position $y = 0$. Obviously this is just the constant term of $q(y)$, i.e. the *last coefficient*. In other words we have the row vector

$$\mathbf{b} = [-1, -1, 1, 1]$$

as the description of the linear functional.
In other words, for a general polynomial of the form $p(x) = a_1x^3 + a_2x^2 + a_3x + a_4$ the value

$$\sigma(p(x)) = \langle \mathbf{a}, \mathbf{b} \rangle = -a_1 - a_2 + a_3 + a_4.$$

So our claim is (FOR THE MOMENT LEFT AS AN EXERCISE, please send your suggestions or solutions or question to the lecturer!): One can reconstruct the full matrix from σ, i.e. it is possible to retrieve the full matrix from its last row (under the assumption that the matrix represents a translation invariant operator).

2 Transferring to non-periodic, continuous signals

OUTLOOK to the next steps: Motivated by the signal space $\mathscr{P}_3(\mathbb{R})$ we move on to a more realistic scenario, i.e. to $\big(\boldsymbol{C}_0(G), \|\cdot\|_\infty\big)$, for a locally compact Abelian group, like $(\mathbb{R}^d, +)$ (Euclidean space).
For the sequel we will basically provide a general approach to invariant systems which commute with translation. Although the approach works for any dimension, and also for discrete signals, we will put a focus on the case $\mathbb{R}$ (one-dimensional, audio-signals) and $\mathbb{R}^d$ with $d = 2, 3$, as it is used for image processing (or volume data).
We can give two formal definition of a BIBOS (bounded input - bounded output system). First the mathematically precise (and simple one). For a detailed discussion of the topic and relations to the published literature I refer to the very recent ARXIV article by M. Unser at EPFL (**un20 :** [89]).

BIBOS1 **Definition 1.** A BIBOS is a bounded linear operator $T : \big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big) \to \big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$ which commutes with translations, i.e. which satisfies

$$\exists\, C > 0 \text{ such that } \forall f \in \boldsymbol{C}_0(\mathbb{R}^d): \quad \|T(f)\|_\infty \leq C\|f\|_\infty \tag{31}$$ boundedness

and

$$T \circ T_z = T_z \circ T, \quad \forall z \in \mathbb{R}^d. \tag{32}$$ Tzcommut

For a practical verification that a system T is a BIBOS system I would prefer the following (more realistic, with respect to verification) assumption:

BIBOS2 **Definition 2.** A BIBOS is a bounded linear operator $T : \boldsymbol{C}_c(\mathbb{R}^d) \to \boldsymbol{C}_0(\mathbb{R}^d)$ which commutes with translations, i.e. which satisfies $T \circ T_z = T_z \circ T, \quad \forall z \in \mathbb{R}^d$. and satisfies a boundedness condition for $T(k)$:

$$\exists C > 0 \text{ such that } \forall k \in \boldsymbol{C}_c(\mathbb{R}^d): \quad \|T(k)\|_\infty \leq C\|k\|_\infty \tag{33}$$ boundedness2

and in addition:

$$\forall k \in \boldsymbol{C}_c(\mathbb{R}^d) \text{ and } \varepsilon > 0\ \exists R > 0 \text{ such that} \quad |x| \geq R \Rightarrow |T(k)(x)| \leq \varepsilon. \tag{34}$$ decayinf2

Note that one has for a general measure (as treated in courses on measure theory, see e.g. [Cohn, Donald, L.] Measure Theory [14]: $f \mapsto \int_{\mathbb{R}^d} f(x)d\mu(x)$ is only well defined on $\boldsymbol{C}_c(\mathbb{R}^d)$ and for any radius $R > 0$ there exists some constant $C_R > 0$ such that for any $f \in \boldsymbol{C}_c(\mathbb{R}^d)$ with $f(x) = 0$ for $|x| \geq R$ one has:

$$|\mu(f)| = \left|\int_{\mathbb{R}^d} f(x)d\mu(x)\right| \leq C_R\|f\|_\infty. \tag{35}$$ unbdmeas1

A typical example is the usual Riemann integral, with C_R being typically the volume of $B_R(0)$, the ball of radius R around zero.

3 Third week

First let us recall the one-to-one relationship between linear functionals (on $\mathscr{P}_3(\mathbb{R})$) and TILS on $\mathscr{P}_3(\mathbb{R})$, by the formula

$$\boxed{T(p)(y) = q(y) = \sigma(T_{-y}p), \quad \forall y \in \mathbb{R}.} \tag{36}$$

polysysrep1r

3.1 Norm of discrete measures

$$\|\sum_{k=1}^{n} c_n \delta_{x_i}\|_{M_b} = \sum_{k=1}^{n} |c_k|.$$

iscmeas0.pdf

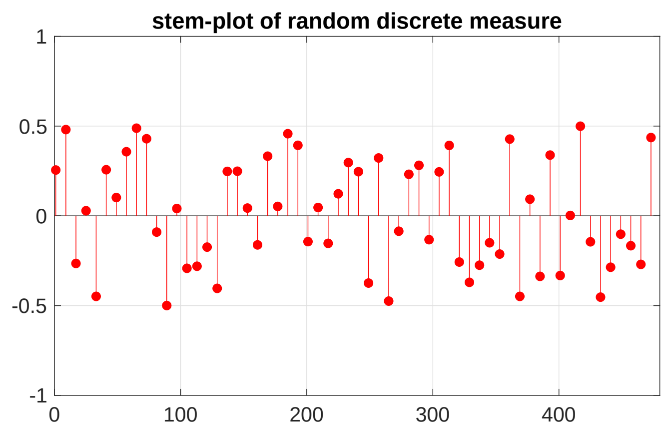

Abbildung 4: maxdiscmeas0.pdf

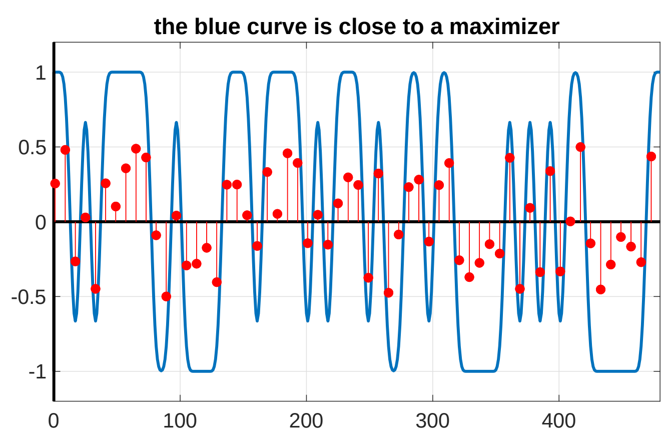

Abbildung 5: maxdiscmeas1.pdf

iscmeas2.pdf

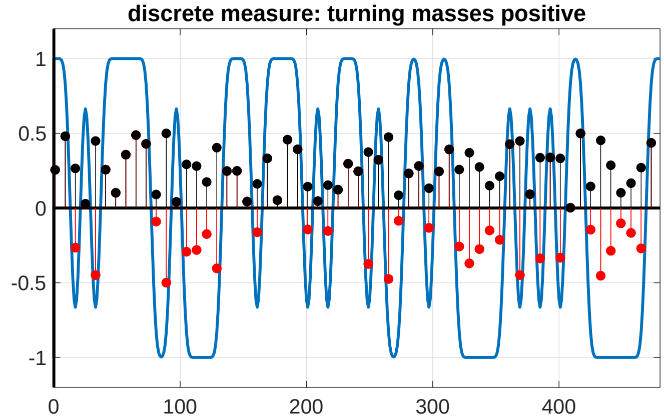

Abbildung 6: maxdiscmeas2.pdf

3.2 The continuous case

Claim: $\mu_k : f \mapsto \int_{\mathbb{R}^d} f(x)k(x)dx$ defines a bounded measure for each $k \in \boldsymbol{C}_c(\mathbb{R}^d)$:

$$\|\mu_k\|_{\boldsymbol{M}_b(\mathbb{R}^d)} = \int_{\mathbb{R}^d} |k(x)|dx = \|k\|_{\boldsymbol{L}^1}.$$

othsign1.pdf

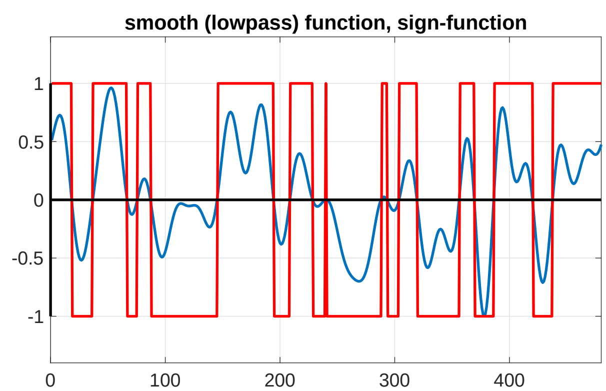

Abbildung 7: smoothsign1.pdf

$$\int_{\mathbb{R}^d} f(x)k(x)dx \equiv \int_{\mathbb{R}^d} |k(x)|dx =: \|k\|_{\boldsymbol{L}^1}$$

The idea: for $f = \operatorname{sign}(k)$ the above estimate provides equality, and thus for

Lispnorm1 **Lemma 3.** *For $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ one has:*

$$\|\mu_k\|_{\boldsymbol{M}} = \|k\|_{\boldsymbol{L}^1}, \quad k \in \boldsymbol{C}_c(\mathbb{R}^d). \tag{37}$$ LispinMb1

Proof. The proof of equality (37) relies on a standard strategy applied in many situations:
LispinMb1
i) First we show that one has for any given $k \in \boldsymbol{C}_c(\mathbb{R}^d)$:

$$\|\mu_k\|_{\boldsymbol{M}} \leq \|k\|_{\boldsymbol{L}^1}.$$

othsign2.pdf

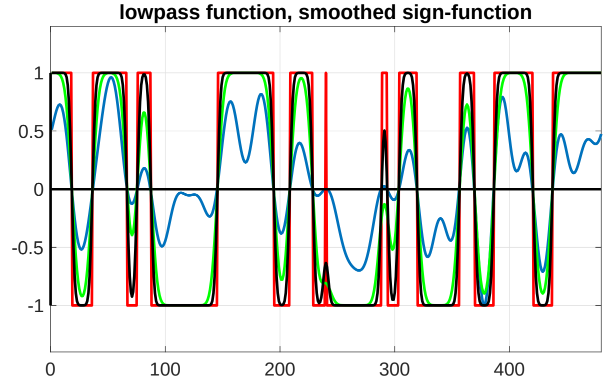

Abbildung 8: smoothsign2.pdf

This is an obvious consequence of the fact that one has for any $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ that $f \cdot k \in \boldsymbol{C}_c(\mathbb{R}^d)$, hence it is Riemann integrable, and

$$|\mu_k(f)| = |\int_{\mathbb{R}^d} f(x)k(x)dx| \leq \int_{\mathbb{R}^d} |f(x)||k(x)|dx \leq \|f\|_\infty \|k\|_1.$$

This implies that $\|k\|_1$ is an upper bound for $\|\mu_k\|_{\boldsymbol{M}_b(\mathbb{R}^d)}$.
ii) In the second step one has to show that it is actually the sup(remum), resp. that there is no smaller lower bound. In other words, one has to show that for any $\varepsilon > 0$ the number $\|k\|_1 - \varepsilon$ is not anymore an upper bound, or equivalently expressed: For any $\varepsilon > 0$ there exists some $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ (in fact in $\boldsymbol{C}_c(\mathbb{R}^d)$) with the restriction $\|f\|_\infty = 1$ such that one has

$$|\mu_k(f)| = |\int_{\mathbb{R}^d} f(x)k(x)dx| \geq \|k\|_1 - \varepsilon.$$

The choice of a suitable f (a smoothed version of the signum function for k) is illustrated by the plots.
The detailed explanation and the proof for complex-valued functions $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ is given as Proposition 6 (Functions to Measures), ca. p. 68, in the SKRIPT (HGFei) Instead if the sign-function one takes the phase function, but it requires some trick because this phase function may be discontinuous near the zeros of the function k.

□

3.3 Regular BUPUs

We will use a lot the so-called BUPUs, the Bounded Uniform Partitions of Unity: Families of bump functions which have small support, finite (limited) overlap, and which add up to the constant one (hence *partition of unity*!).

UPUETH01.pdf

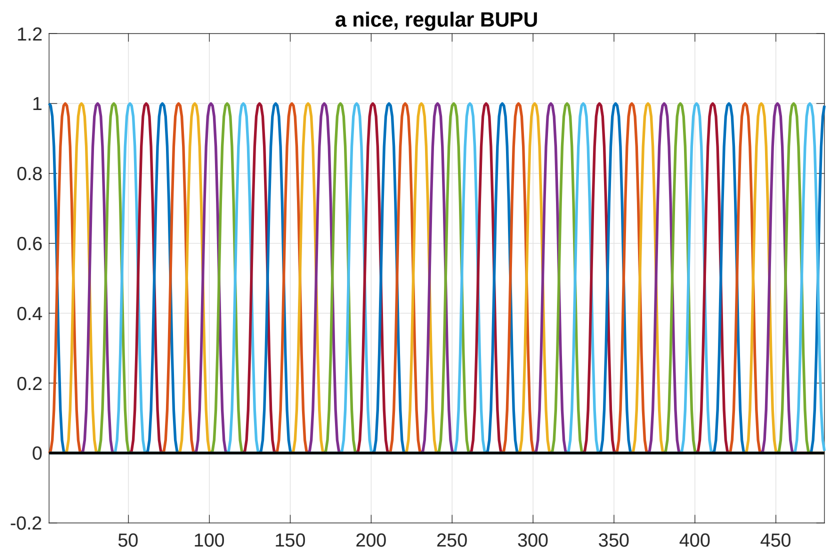

Abbildung 9: BUPUETH01.pdf

We use them in order to *cut a bounded measure into localized pieces* and we also use it in order to approximate function in $\boldsymbol{C}_0(\mathbb{R}^d)$ by e.g. piecewise polynomial functions, so-called *splines*. In computer graphics cubic splines are the most popular ones, because they have continuous second derivatives, i.e. the allow to represent smooth curves in a flexible way (so to say: roads which allow smooth driving, i.e. the steering wheel position, corresponding to the current radius of curvature, is changing only in a continuous way and without jumps!).
For the case $d \geq 2$ one can obtain BUPUs by taking tensor products of BUPUs (may be even the same), say $\Psi \otimes \Psi$ consists of the collection

$$(\psi_i \otimes \psi_{i'})_{(i,i') \in II}$$

with

$$(f \otimes g)(x, y) = f(x)g(y), \quad x, y \in \mathbb{R}^d,$$

and

$$II = \{(i, i') \,|\, i, i' \in I\}.$$

4 Functional analytic tools and arguments

For the sake of *repeated use* in similar, but different contexts, let us summarize some basic principles from functional analysis resp. some typical arguments used widely. Some of the statements below are well-known functional analytic principles (basic facts, so to say) and others are more like examples of best practice. They are supposed to present arguments, which are very useful and which can be applied in many situations.
First we recall that we are using often the fact that $\left(\boldsymbol{C}_0(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$ is a Banach space, i.e. every Cauchy sequence , in fact every Cauchy net $(f_n)_{n\geq 1}$ of $(f_\alpha)_{\alpha\in I}$ in $\boldsymbol{C}_0(\mathbb{R}^d)$ (with respect to the sup-norm) has a limit inside of the space $\boldsymbol{C}_0(\mathbb{R}^d)$.
The usual argumentation is

1. $(\left(\boldsymbol{C}_b(\mathbb{R}^d),\, \|\cdot\|_\infty\right))$ is a complete space, because *uniform limits of continuous functions have to be continuous* functions, and since Cauchy sequences are bounded in the norm sense, i.e. $\sup_n \|f_n\|_\infty < \infty$ the limit will be also a bounded function (this has to be verified first);

2. $\left(\boldsymbol{C}_0(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$ is a closed subspace, hence it is also complete. In other words, if one has $f_n \in \boldsymbol{C}_0(\mathbb{R}^d)$ for all $n \geq 1$ then also the limit (which exists in $\left(\boldsymbol{C}_b(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$, by (i)) must belong to $\left(\boldsymbol{C}_0(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$ (this argument has been worked out).

The next step is to discuss the *completeness* of $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b}) = \left(\boldsymbol{C}_0'(\mathbb{R}^d),\, \|\cdot\|_{\boldsymbol{C}_0'}\right)$. It is a special case of a more general principle, but let us give the proof in the more concrete context of $(\boldsymbol{B},\, \|\cdot\|_{\boldsymbol{B}}) = \left(\boldsymbol{C}_0(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$:

compldualsp1 **Lemma 4.** *Given a normed space $(\boldsymbol{B},\, \|\cdot\|_{\boldsymbol{B}})$, the dual space $(\boldsymbol{B}',\, \|\cdot\|_{\boldsymbol{B}'})$, i.e. the collection of all bounded linear functionals $\mu : f \mapsto \mu(f)$ is complete with respect to the dual norm, given by*

$$\|\mu\|_{\boldsymbol{B}'} = \sup_{\|f\|_{\boldsymbol{B}}\leq 1} |\mu(f)|. \tag{38}$$ dualnorm1

Even this very useful general fact can be further generalized (resp. represents a special case of the more general statement) to (choose $(\boldsymbol{B}^2,\, \|\cdot\|^{(2)})$ just $\mathbb{R}$ or $\mathbb{C}$ with the absolute value as a norm):

complopsp1 **Lemma 5.** *Let* $(\boldsymbol{B}^1,\, \|\cdot\|^{(1)})$ *be a normed space, and* $(\boldsymbol{B}^2,\, \|\cdot\|^{(2)})$ *be a Banach space, i.e. a complete normed space. Then* $\mathcal{L}(\boldsymbol{B}^1, \boldsymbol{B}^2)$*, endowed with the operator norm*

$$|||T||| = \sup_{f \in \boldsymbol{B}^1, \|f\|_{\boldsymbol{B}^1} \le 1} \|Tf\|_{\boldsymbol{B}^2}$$

is a Banach space.

Proof. Assume the $(T_n)_{n=1}^{\infty}$ is a Cauchy sequence in $\mathcal{L}(\boldsymbol{B}^1, \boldsymbol{B}^2)$ (endowed with the operator norm $|||\cdot|||$). Then clearly for any $f \in \boldsymbol{B}^1$ we have

$$\|T_n f - T_m f\|_{\boldsymbol{B}^2} = \|(T_n - T_m)f\|_{\boldsymbol{B}^2} \le |||T_n - T_m|||\, \|f\|_{\boldsymbol{B}^1}, \tag{39}$$

i.e. $(T_n f_n)_{n=1}^{\infty}$ is a Cauchy sequence in $(\boldsymbol{B}^2,\, \|\cdot\|^{(2)})$. By the completeness assumption the limit exists in $(\boldsymbol{B}^2,\, \|\cdot\|^{(2)})$, and we may call it $T_0(f)$, i.e.

$$T_0(f) = \lim_{n \to \infty} T_n f. \tag{40}$$

limTn

It is easy to check (left to the reader) that $f \mapsto T_0(f)$ is a linear mapping, so it remains to show that the limit operator is also a bounded operator, in fact we even have for any $n_0 \ge 1$:

$$|||T_0||| \le \sup_{n \ge n_0} |||T_n|||. \tag{41}$$

opnormest2

First, by choosing as $\varepsilon = 1$ we can find n_0 such that one has for $n, m \ge n_0$: $|||T_n - T_m||| \le 1$. Hence we have for any $f \in \boldsymbol{B}^1$ with $\|f\|_{\boldsymbol{B}^1} \le 1$:

$$\|T_n f\|_{\boldsymbol{B}^2} \le |||T_{n_0} - T_n|||\, \|f\|_{\boldsymbol{B}^1} + \|T_{n_0} f\|_{\boldsymbol{B}^2} \le 1 + |||T_{n_0}|||\,, \tag{42}$$

Toestim

and by taking the limit

$$\|T_0 f\|_{\boldsymbol{B}^2} = \lim_{n \to \infty} \|T_n f\|_{\boldsymbol{B}^2} \le 1 + |||T_{n_0}|||\,. \tag{43}$$

Toestim2

This implies immediately that the limit mapping is in fact a bounded linear operator (hence the symbol T_0 is well justified!) and the claimed operator norm estimation (41) opnormest2 is valid. □

5 TILS: measures representing operators

The key connection is the relation (slightly different from the "moving average approach for $\mathscr{P}_3(\mathbb{R})$), using the flip-operator, given by $f^{\checkmark}(x) = f(-x)$.

$$T_\mu f(y) = \mu(T_y f^{\checkmark}) \quad \text{for} \quad \mu_T(f) = T(f^{\checkmark})(0). \tag{44}$$ sysfuncrel

By this trick the correspondence between the operator T_z (right shift) and the discrete point measure δ_z (instead of $-z$!) arises.
We have to go through different steps, in particular we have to show that the linear operator T_μ defined by (sysfuncrel 44) actually defines a translation invariant and bounded operator on $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$.
Basically the linearity and the boundedness are easy to check, including the estimate

$$|||T_\mu|||_{\boldsymbol{C}_0(\mathbb{R}^d)} \leq \|\mu\|_{\boldsymbol{M}_b(\mathbb{R}^d)} .$$

What is more interesting is the verification that such a bounded and linear operator commutes with any translation operator T_z, $z \in \mathbb{R}^d$: Using the simple fact that

$$(T_z f)^{\checkmark} = T_{-z} f^{\checkmark}$$

we get:

$$[T_\mu(T_z f)](y) = \mu(T_y(T_{-z} f^{\checkmark})) = \mu(T_{y-z} f^{\checkmark}) = [T_\mu f](y-z) = [T_z(T_\mu f)](y). \tag{45}$$ Tsigcomm1

This being true for any $y \in \mathbb{R}^d$ we have $T_\mu(T_z f) = T_z(T_\mu f)$ for any $f \in \boldsymbol{C}_0(\mathbb{R}^d)$, or equivalently the desired commutativity for operators

$$T_\mu \circ T_z = T_z \circ T_\mu, \quad \forall z \in \mathbb{R}^d.$$

Next we have to show (once the boundedness is established) that for any input $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ $T_\mu(f)$ is not only a bounded, but also a continuous function, vanishing at infinity. Since $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ is also uniformly continuous, we can use the translation invariance of T_μ in order to derive the *uniform* continuity, due to the estimate

$$\|T_x(T_\mu f) - T_\mu f\|_\infty = \|T_\mu(T_x f - f)\|_\infty \leq |||T_\mu|||\, \|T_x f - f\|_\infty \to 0 \text{ for } |x| \to 0. \tag{46}$$ presCub

The final step, in order to guarantee the decay at infinity. We use the fact that compactly supported elements are dense in $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ as well as in $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$. More precisely we proceed in two steps.
First we prove that for *compactly supported* μ and f, i.e. $\mu = \sum_{i \in F} \mu\psi_i$ and $f = \sum_{j \in E} f\psi_j$, where E and F are finite sets, one finds that the resulting output signal $T_\mu(f)$ has compact support inside of $B_{2R}(0)$, because $T_\mu(f)(y) = 0$ for $|y| > 2R$. Thus altogether $T_\mu(f) \in \boldsymbol{C}_c(\mathbb{R}^d) \subset \boldsymbol{C}_0(\mathbb{R}^d)$. The general case is derived therefrom by a standard approximation argument! (a good exercise for the interested reader!).
So let us first observe that due to the properties of BUPUs for an finite set (here E and F) we have: There exists $R > 0$ such that for $|x| \geq R$ one has $\psi_i(x) = 0$ (because each of the bump functions in the BUPU have bounded, i.e. compact support). We can take the same radius for E and F.

Let us now look at the convolution product $T_{\mu\,\psi_i}(f\psi_j)$ and try to find a control on the region where this can be non-zero, for $i \in F$ and $j \in F$. By definition we have

$$T_{\mu\,\psi_i}(f\psi_j)(y) = \mu\,\psi_i(T_y(\psi_j f)^{\checkmark}) = \mu(\psi_i \cdot T_{-y}\psi_j{}^{\checkmark} \cdot T_y f^{\checkmark}) = \mu([\psi_i \cdot T_{-y}\psi_j{}^{\checkmark}] \cdot T_y f^{\checkmark}). \tag{47}$$

convpiec1

Now we note that for the case $|y| > 2R$ the two functions $\psi_i, i \in F$ and $T_{-y}\psi_j{}^{\checkmark}, j \in E$ cannot be non-zero at the same time, hence the pointwise product is zero and the resulting action of μ on the zero-function (multiplication of zero by $T_{-y}f^{\checkmark}$ gives zero again!) is zero. Combining these facts we see that $T_\mu f \in \boldsymbol{C}_{ub}(\mathbb{R}^d)$, with compact support, hence $T_\mu f \in \boldsymbol{C}_c(\mathbb{R}^d) \subset \boldsymbol{C}_0(\mathbb{R}^d)$.
The final step is taken by observing that we have (due to the absolute convergence, namely the condition $\sum_{i\in I} \|\mu\,\psi_i\|_{\boldsymbol{M}_b} = \|\mu\|_{\boldsymbol{M}_b}$):

$$\|\mu - \sum_{i\in F} \mu\,\psi_i\|_{\boldsymbol{M}_b} \to 0 \ \text{ for } F \nearrow I, \tag{48}$$

absconv

and similarly

$$\|f - \sum_{j\in E} f\psi_j\|_\infty \to 0 \text{ for } E \nearrow I. \tag{49}$$

convsumE

The detailed verification of this last condition is (for now) left to the reader (!Exercise).
In other words, the elements which satisfy the extra conditions just made (in order to come up with the compact support condition for $T_\mu f$) is verified for a *dense subspace* of elements (that linear combinations of elements with this condition share this condition is easy to check, but perhaps not trivial).
So we can create sequences of such measures with compact support, we may call them m_n, convergent to μ in $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$, and also sequences of functions f_n, convergent to f in $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$ (by choosing for $E = F$ e.g. all the elements in the BUPU which intersect $B_n(0)$, the ball of radius $n \in \mathbb{N}$ around 0 in $\mathbb{R}^d$).
Then (please try to work out the argument yourself) one has ...
$T_{\mu_n}(f_n)$ is a Cauchy sequence in $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$, and since this is a Banach space, the limit (which is taken using the uniform, hence pointwise convergence), which is of course $T_\mu(f)$ (why?) belongs to $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$ (and not just to $\boldsymbol{C}_{ub}(\mathbb{R}^d)$).
We do not do the detailed discussion of why we have that the (new) measure assigned to a given operator T_μ is just the original measure, and conversely, that the operator, assigned to σ_T, for a given translation invariant system is of course (in the sense of as expected!) the original TILS T (cf. original Lect.Notes).
In other words, we have that the mappings $T \mapsto \mu_T$ and $\mu \mapsto T_\mu$ are inverse to each other (and of course are both linear!).
Let us also observe that obviously one has

$$|\mu_T(f)| = |T(f^{\checkmark})(0)| \le |||T|||\ \|f^{\checkmark}\|_\infty = |||T|||\ \|f\|_\infty,$$

hence

$$\|\mu_T\|_{\boldsymbol{M}_b} \le |||T|||\ := |||T|||_{\boldsymbol{C}_0(\mathbb{R}^d)}\,. \tag{50}$$

muTest1

Recall that we also have the other estimate

$$|T_\mu f(y)| = |\mu(T_y f^{\checkmark})| \leq \|\mu\|_{\boldsymbol{M}_b} \|f\|_\infty, \quad \forall y \in \mathbb{R}^d,$$

hence

$$\|T_\mu f\|_\infty \leq \|\mu\|_{\boldsymbol{M}_b} \|f\|_\infty, \quad \forall f \in \boldsymbol{C}_0(\mathbb{R}^d),$$

or finally

$$|||T_\mu||| \leq \|\mu\|_{\boldsymbol{M}_b}, \quad \forall \mu \in \boldsymbol{M}_b(\mathbb{R}^d). \tag{51}$$

opnormest1

Observing that the two mappings are inverse to each other and both are non-expansive ((50) and (51) we come to the important conclusion that the mapping is an isometric one:

mulestf opnormest1

$$|||T_\mu||| = \|\mu\|_{\boldsymbol{M}_b}, \quad \forall \mu \in \boldsymbol{M}_b(\mathbb{R}^d). \tag{52}$$

isomTmu

We can summarize: The correspondence between TILS (described as T_μ) and linear functionals on $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$, named *bounded measures* is an isometric one. Operators T_{μ_n} are Cauchy sequences in $(\mathcal{L}_{\mathbb{R}^d}(\boldsymbol{C}_0(\mathbb{R}^d)), |||\cdot|||)$ if and only if the corresponding measures μ_n, $n \geq 1$ are a Cauchy sequence in $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$.

There was a break here

itofourA.pdf

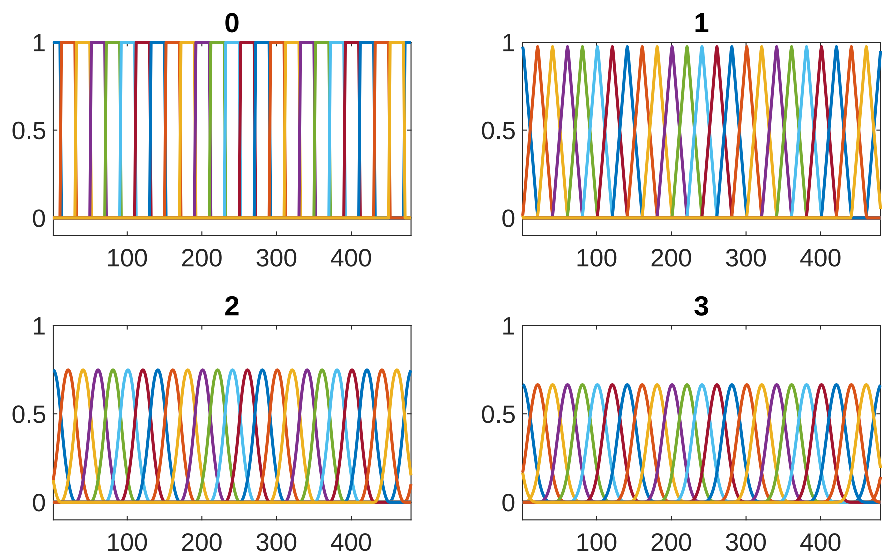

Abbildung 10: bupuitofourA.pdf

Let us recall:

$$D_\Psi \mu = \sum_{i \in I} \mu(\psi_i) \delta_{x_i}, \quad \text{with} \quad D_\Psi \mu(f) = \mu(\mathrm{Sp}_\Psi f), \quad f \in \boldsymbol{C}_0(\mathbb{R}^d),$$

hence

$$|D_\Psi \mu(f)| \leq \|\mu\|_{\boldsymbol{M}_b} \|\mathrm{Sp}_\Psi f\|_\infty \leq \|\mu\|_{\boldsymbol{M}_b} \|f\|_\infty,$$

or

$$\|D_\Psi \mu\|_{\boldsymbol{M}_b} \leq \|\mu\|_{\boldsymbol{M}_b}, \quad \forall \mu \in \boldsymbol{M}_b(\mathbb{R}^d),$$

pprfourA.pdf

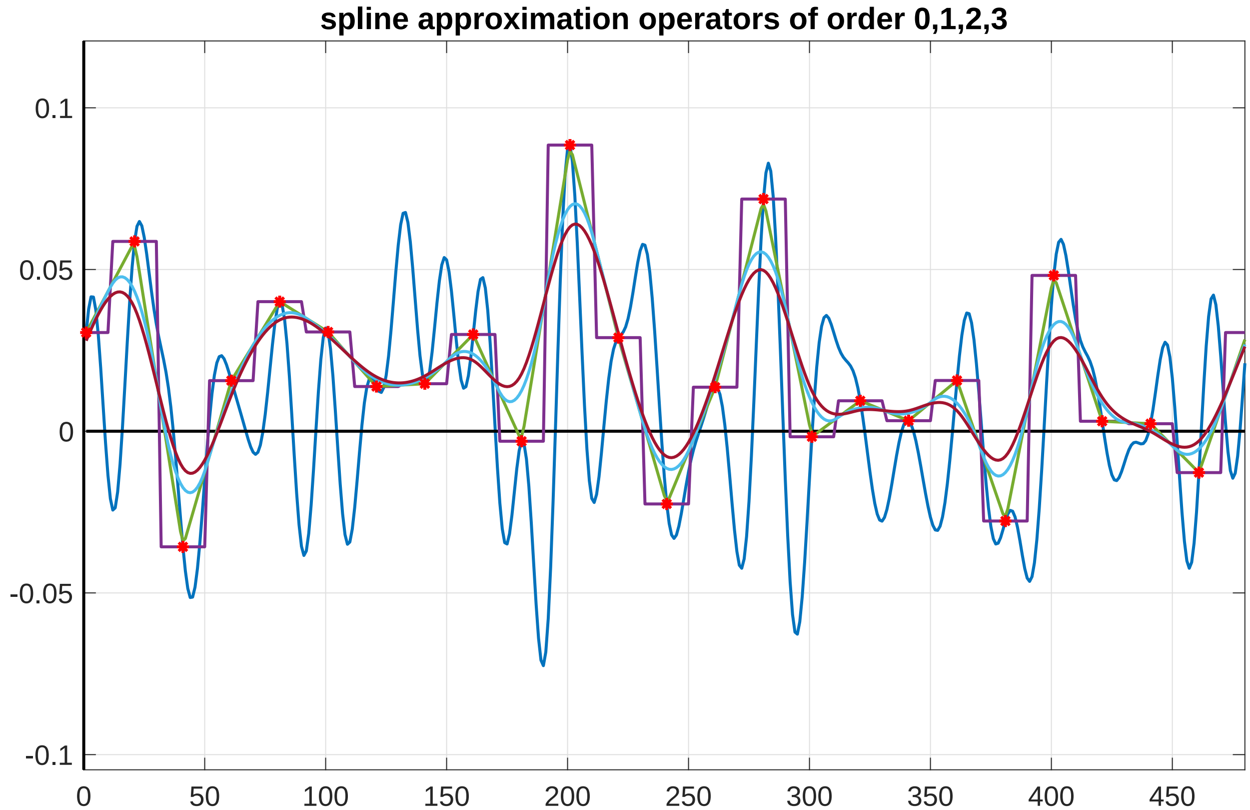

Abbildung 11: splinapprfourA.pdf:
Approximation of a given smooth, real-valued function by spline functions of order $1,2,3,4$ resp. degree $0,1,2,3$ using the (quasi-)interpolation operator Sp_Ψ.

in other words, the linear mappings $\mathrm{D}_\Psi : \mu \to D_\Psi \mu$ (with $|\Psi| \leq 1$ are all non-expansive (on $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b}))$.
Distinguish (!) between $(\mu, f) \to C_\mu(f)$ is a bilinear mapping from $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b}) \times \left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$ into $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$, and convolution with $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ (related to concatenation of operators) is using the same symbol.
Let us recall that the very natural relation

$$(\mu_1 \star \mu_2) \star f = \mu_1 \star (\mu_2 \star f) \tag{53}$$ assocconv00

is a consequence of the *definition* of the internal multiplication (called convolution) in $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$. The second symbol is always the application of the measure (the application of the convolution operator) generated by the measure to the left.
So far we have established the following facts:

MbconvBanalg **Theorem 2.** *The Banach space $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ is a Banach algebra with respect to the multiplication, given through convolution. δ_0 is the unit element in the multiplicative group of this Banach algebra. In particular we have submultiplicativity of the norm:*

$$\|\mu_1 \star \mu_2\|_{\boldsymbol{M}_b} \leq \|\mu_1\|_{\boldsymbol{M}_b} \cdot \|\mu_2\|_{\boldsymbol{M}_b}, \quad \mu_1, \mu_2 \in \boldsymbol{M}_b(\mathbb{R}^d). \tag{54}$$ convMbsubm

inappr02.pdf

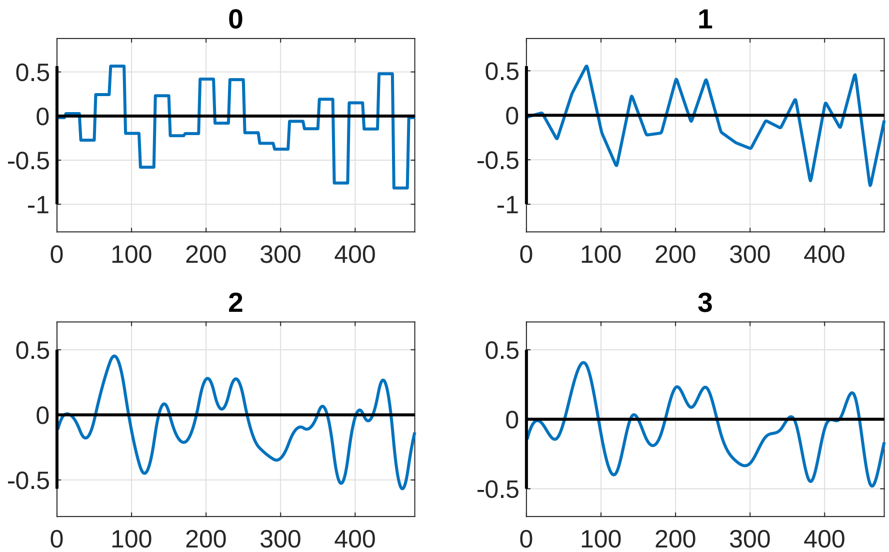

Abbildung 12: splinappr02.pdf:Approximation of a smooth signal by splines via Sp_Ψ .

Proof. It is clear that the set of all bounded linear operators on $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$ which commute with translations (TILS) is a Banach algebra with respect to composition. In fact, it is easy to check that it is closed inside of $\boldsymbol{L}(\boldsymbol{C}_0(\mathbb{R}^d))$, the Banach algebra of all bounded linear operators on $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$, endowed with the operator norm. Hence it is a Banach algebra (under composition of operators), endowed with the operator norm, which is of course submultiplicative. By transfer of structure we obtain (54). □ convMbsubm

$$\|\mu \star f\|_\infty \leq \|\mu\|_{\boldsymbol{M}_b} \|f\|_\infty, \quad \mu \in \boldsymbol{M}_b(\mathbb{R}^d), f \in \boldsymbol{C}_0(\mathbb{R}^d). \tag{55}$$ mufestim

$$\mu(f) = \mu \star f^{\checkmark}(0) \quad (= [\mu \star f^{\checkmark}](0)). \tag{56}$$ TILSmeas

5.1 Convolution and the commutativity question

It is clear that $D_\Psi \mu$ is w^*-convergent to $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ for $|\Psi| \to 0$, which means that

$$\lim_{|\Psi|\to 0} D_\Psi \mu(f) = \mu(f), \quad \forall f \in \boldsymbol{C}_0(\mathbb{R}^d). \tag{57}$$ Dpsiconv04

Upon replacing f by $T_y f^{\checkmark}$ one can see that

$$\lim_{|\Psi|\to 0} (D_\Psi \mu \star f - \mu \star f)(y) = 0, \quad \forall y \in \mathbb{R}^d. \tag{58}$$ Dpsiconv05

In fact, using the equicontinuity of the family $D_\Psi \mu \star f$ one can show that the convergence is uniform over compact sets. For details you should check with Lemma 8 of the script (labeled [unifcomp01], p. 18). In other words, for any $R > 0$ and $\varepsilon > 0$ there exists $\delta_0 > 0$ such that for $|\Psi| \leq \delta > 0$ on has for all y with $|y| \leq R$:

$$|D_\Psi \mu \star f(y) - \mu \star f(y)| \leq \varepsilon. \tag{59}$$ DpsimufR

But in fact much more is true (because essentially one has control about the behaviour of $D_\Psi\mu \star f$ at infinity (so fare we leave out details), see script.

DPsiconvappr **Proposition 1.** *For every $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ and $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ one has*

$$lim_{|\Psi|\to 0} \|D_\Psi\mu \star f - \mu \star f\|_\infty = 0. \tag{60}$$ convdiscr1

Moreover the limit is uniform for (bounded and) equicontinuous sets $M \subset \boldsymbol{C}_0(\mathbb{R}^d)$.

Remark 1. A verbal description of the above statement may be given as follows: Given any TILS T the output signal $T(f)$ (for input $f \in \boldsymbol{C}_0(\mathbb{R}^d)$) can be obtained by convolving f with the uniquely determined measure μ. It can be approximated by a finite sum of shift operators, with coefficients which depend on the concentration of mass (thinking of a probability measure μ), namely $\mu(\psi_i)$.

Proof. The idea of the proof is to make use of the fact that compactly supported measures are dense in $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ and that $\boldsymbol{C}_c(\mathbb{R}^d)$ is dense in $\Big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\Big)$, and that the pointwise convergence described by (58) [Dpsiconv05] is uniform over the compact set where the convolution product $\mu \star f$ is non-zero. □

Distinguish (!) between $(\mu, f) \to C_\mu(f)$ is a bilinear mapping from $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b}) \times \Big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\Big)$ into $\Big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\Big)$, and convolution with $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ (related to concatenation of operators) is using the same symbol.
Let us recall that the very natural relation

$$(\mu_1 \star \mu_2) \star f = \mu_1 \star (\mu_2 \star f) \tag{61}$$ assocconv00b

is a consequence of the *definition* of the internal multiplication (called convolution) in $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$. The second symbol is always the application of the measure (the application of the convolution operator) generated by the measure to the left.

DPsiconv04 **Lemma 6.** *Given $\mu, \nu \in \boldsymbol{M}_b(\mathbb{R}^d)$ we claim that one has for $f \in \boldsymbol{C}_0(\mathbb{R}^d)$, for $|\Psi| \to 0$:*

$$(D_\Psi\mu \star \mathrm{D}_\Psi\nu) \star f = D_\Psi\mu \star (\mathrm{D}_\Psi\nu \star f) \to (\mu \star \nu) \star f \quad in \Big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\Big). \tag{62}$$ DPsimunu

In particular one has w^*-convergence of

$$D_\Psi\mu \star \mathrm{D}_\Psi\nu \ \to^{w^*} \ \mu \star \nu, \quad for \ |\Psi| \to 0. \tag{63}$$ DPsimunuwst

Proof. The first statement follows from the uniform boundedness[5] of the family $(D_\Psi\mu)_{|\Psi|\le 1}$ in $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$, for any given $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$, since one has the estimate:

$$\|(D_\Psi\mu \star \mathrm{D}_\Psi\nu) \star f - (\mu \star \nu) \star f\|_\infty \le$$

using the commutativity of convolution (probably avoidable!)

$$\le \|D_\Psi\mu \star (\mathrm{D}_\Psi\nu \star f) - D_\Psi\mu \star (\nu \star f)\|_\infty + \|\nu \star (D_\Psi\mu \star f - \mu \star f)\|_\infty \le$$

[5] For an elementary proof of the *uniform boundedness principle* (related to the original proof of H. Hahn using a kind of gliding hump technique, is given in [86]).

$$\leq \|D_\Psi\mu\|_{\boldsymbol{M}_b}\|\mathrm{D}_\Psi\nu \star f - \ \nu \star f\|_\infty + \|\nu\|_{\boldsymbol{M}_b}\|D_\Psi\mu \star f - \mu \star f\|_\infty.$$

Since we have $\|D_\Psi\mu\|_{\boldsymbol{M}_b} \leq \|\mu\|_{\boldsymbol{M}_b}$ Prop. 1 can be invoked. The convergence described in (63) follows from the validity of (62), observing that for any $f \in \boldsymbol{C}_0(\mathbb{R}^d)$:

$$[D_\Psi\mu \star \mathrm{D}_\Psi\nu](f) = [(D_\Psi\mu \star \mathrm{D}_\Psi\nu) \star f^\checkmark](0) \to_{|\Psi|\to 0} ([\mu \star \nu] \star f^\checkmark)(0) = \mu \star \nu(f).$$

□

We have seen, that Dirac measures correspond to translation operators, which of course commute. Hence we have

commMdRd **Lemma 7.** *$(\boldsymbol{M}_d(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ is a closed,* **commutative** *subalgebra of $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ with respect to convolution.*

Proof. We know already that $(\boldsymbol{M}_d(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ is a closed subspace of $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$, hence in fact a Banach space with the norm inherited from $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$. Thus we only have to check that it is closed under convolution and that convolution is commutative in $\boldsymbol{M}_d(\mathbb{R}^d)$. Since we have seen that

$$\delta_x \star \delta_y = \delta_{x+y} = \delta_y \star \delta_x, \quad x, y \in \mathbb{R}^d,$$

it is clear that convolution of measures is commutative for finite discrete measures, and then also, by taking limits, for arbitrary discrete bounded measures (resp. infinite series of Dirac measures). The technical approximation argument relies on the submultiplicativity of the norm (54), see also Lemma 32 below. □

In order to derive commutativity of convolution (inside of $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$) we have to combine the observations above:

CommutConvM **Proposition 2.** *For any two bounded measures $\mu_1, \mu_2 \in \boldsymbol{M}_b(\mathbb{R}^d)$ one has*

$$\mu_1 \star \mu_2 = \mu_2 \star \mu_1. \tag{64}$$ muitcomm

Thus $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ is a commutative Banach algebra with respect to convolution, with identity δ_0, i.e.$\delta_0 \star \mu = \mu$, $\forall \mu \in \boldsymbol{M}_b(\mathbb{R}^d)$.

Proof. We will use an indirect argument. If one could find two bounded measures $\mu_1, \mu_2 \in \boldsymbol{M}_b(\mathbb{R}^d)$ such that

$$\mu_1 \star \mu_2 \neq \mu_2 \star \mu_1 \tag{65}$$ muitcommneg

one also would find some $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ with $\|f\|_\infty = 1$ and $\varepsilon > 0$ such that

$$|[\mu_1 \star \mu_2](f) - [\mu_2 \star \mu_1](f)| = \varepsilon_0 > 0. \tag{66}$$ muitcommnegf

But by applying (63)[6] we can choose sufficiently fine BUPUs such that we have both

$$|[\mu_1 \star \mu_2](f) - [\mathrm{D}_\Psi(\mu_1) \star \mathrm{D}_\Psi(\mu_2)](f)| < \varepsilon_0/4$$

and

$$|[\mu_2 \star \mu_1](f) - [\mathrm{D}_\Psi(\mu_2) \star \mathrm{D}_\Psi(\mu_1)](f)| < \varepsilon_0/4.$$

[6] We do not claim that $\mathrm{D}_\Psi(\mu_1 \star \mu_2) = \mathrm{D}_\Psi(\mu_1) \star \mathrm{D}_\Psi(\mu_2)$!

Now upon writing (according to Lemma commMdRd 7)

$$\mathrm{D}_\Psi(\mu_2) \star \mathrm{D}_\Psi(\mu_1) = \nu = \mathrm{D}_\Psi(\mu_1) \star \mathrm{D}_\Psi(\mu_2)$$

we arrive, by means of the triangular equation, at

$$|\mu_1 \star \mu_2(f) - \mu_2 \star \mu_1(f)| \leq |\mu_1 \star \mu_2(f) - \nu(f)| + |\nu(f) - \mu_2 \star \mu_1(f)| < \varepsilon_0/2,$$

in contradiction to the indirect assumption (muitcommnegf 66).
The action of δ_0 within $\boldsymbol{M}_b(\mathbb{R}^d)$ is of course just the expected one:

$$[\delta_0 \star \mu](f) = \mu \star [\delta_0 \star f^{\checkmark}](0) = \mu \star f^{\checkmark}(0) = \mu(f), \quad \forall \mu \in \boldsymbol{M}_b(\mathbb{R}^d).$$

□

5.2 Different variants of convolution

Given the commutativity of convolution we see a new problem. Given two functions $k_1, k_2 \in \boldsymbol{C}_c(\mathbb{R}^d)$, how can or should one define the convolution of these two functions? We cannot (so far!) convolve two general elements from $\boldsymbol{C}_0(\mathbb{R}^d)$, BUT, we can interpret one or both of the two functions as bounded measures. Thus we have so far the function $k_1, k_2 \in \boldsymbol{C}_c(\mathbb{R}^d) \subset \boldsymbol{C}_0(\mathbb{R}^d)$, the associated bounded measures $\mu_1 := \mu_{k_1}$ and $\mu_2 := \mu_{k_2}$ in $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$. This raises the questions about the potential ambiguity of convolutions. First we will check that the following functions coincide in $\boldsymbol{C}_0(\mathbb{R}^d)$:

$$\mu_{k_1} \star k_2 \quad \text{and} \quad \mu_{k_2} \star k_1 \quad \text{and} \quad k_1 \star_{pw} k_2 \tag{67}$$ convfunctequ

respectively the following measures:

$$\mu_{k_1} \star \mu_{k_2} \quad \text{and} \quad \mu_{k_1 \star_{pw} k_2} \quad \text{and} \quad \mu_{\mu_{k_2} \star k_1} \quad \text{and} \quad \mu_{k_1 \star_{pw} k_2} \tag{68}$$ convmeasequ

Of course we will not do each and any of them in great detail, but we go through the less obvious ones step by step!
Of course all these elements can be described as a continuous function with compact support, which can be described pointwise:

$$k_1 \star_{pw} k_2(y) = \int_{\mathbb{R}^d} k_1(y-z)k_2(z)dz = \int_{\mathbb{R}^d} k_2(y-z)k_1(y)dz = k_2 \star_{pw} k_1(y). \tag{69}$$ kiktconv

Formula (kiktconv 69) coincides with the interpretation as $C_{\mu_{k_2}}(k_1)$ since

$$C_{\mu_2}(k_1)(y) = \int_{\mathbb{R}^d} T_y({k_1}^{\checkmark})(z)k_2(z)dz = \int_{\mathbb{R}^d} {k_1}^{\checkmark}(z-y)k_2(z)dz = k_1 \star_{pw} k_2(y). \tag{70}$$ checkkikt

In order to convince ourselves that $\mu_{k_1 \star_{pw} k_2} = \mu_{k_1} \star \mu_{k_2}$ we only have to show that the action on any given $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ is the same (because by density the action of this functional coincides on all of $\boldsymbol{C}_0(\mathbb{R}^d)$). Given $k_1, k_2 \in \boldsymbol{C}_c(\mathbb{R}^d)$ and the corresponding measures $\mu_1 = \mu_{k_1}$ and $\mu_2 = \mu_{k_2}$ we can produce the convolution product (via composition) $\mu := \mu_1 \star \mu_2$, whereas of course one can check (fairly easily) that $k_1 \star_{pw} k_2 \in \boldsymbol{C}_c(\mathbb{R}^d)$ also defines a measure, namely $\mu_{k_1 \star_{pw} k_2}$.

Recall that for any $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ one has the connection: $\mu(k) = C_\mu(k^{\checkmark})(0)$.

$$\mu(k) = C_\mu k^{\checkmark}(0) = C_{\mu_1 \star \mu_2} k^{\checkmark}(0) = [C_{\mu_1}(C_{\mu_2} k^{\checkmark})](0) = \mu_1[(C_{\mu_2} k^{\checkmark})^{\checkmark}](0). \quad (71)$$ muk1muk2

Hence we have to compute that one has

$$C_{\mu_2} k^{\checkmark}(y) = \mu_2(T_y k) = \int_{\mathbb{R}^d} k(z-y)\, k_2(z) dz =_{z \to z+y} \int_{\mathbb{R}^d} k(z) k_2(z+y) dz \quad (72)$$ Cmutkchk

Note that we have to apply a flip operator to this function, changing the argument $k_2(x+y)$ to $k_2(x-y)$ in the formula below (for the outer flip operator) from (71):

$$\mu(k) = C_{\mu_1}\left[(C_{\mu_2} k^{\checkmark})^{\checkmark}\right](0) = \int_{\mathbb{R}^d} k_1(y) \left(\int_{\mathbb{R}^d} k(z) k_2(z-y) dz \right) dy = \quad (73)$$

$$=_{Fubini} \int_{\mathbb{R}^d} k(z) \left(\int_{\mathbb{R}^d} k_1(y) k_2(z-y) dy \right) dz = \mu_{k_1 \star_{pw} k_2}(k). \quad (74)$$

In this way (we will encounter various similar situations) we will be free to interpret convolutions whenever they exist according to one or the other method, with the extra benefit that *associativity* and *commutativity* of convolution are granted within this setting, as one might expect.

5.3 The Lebesgue space

Most mathematical books, in particular **re68:** [78] or the more recent edition **rest00:** [79], by my advisor Hans Reiter, or the books of Folland (**fo92:** [48], **fo94:** [49]) take Lebesgue integration theory as granted to start with the space $\left(\boldsymbol{L}^1(\mathbb{R}^d),\, \|\cdot\|_1\right)$ as the starting point for Fourier analysis (over $\mathscr{G} = \mathbb{R}^d$ or even more generally, starting from the Haar measure over a locally compact Abelian (LCA) group).
It is the goal of this subsection to *link our approach*, which starts with $\left(\boldsymbol{C}_0(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$ and its dual, called the space of *bounded measures* $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$, with the approach based on measure theory.
First of all let us emphasize that the terminology is well justified, because (different variants of) the *Riesz Representation Theorem* allow to identify these two objects. Books and courses on measure theory start from a σ-additive regular Borel measure, which allows to determine $\mu(A)$ for certain non-pathological sets (called μ-measurable), including all the open or compact sets (due to the Borel assumption). They go on to define expressions like

$$\int_{\mathscr{G}} f(x) d\mu(x)$$

first for measurable step functions (i.e. function of the form

$$f(x) = \sum_{k=1}^{n} c_k \mathbf{1}_{A_k}$$

which take constant values $c_k \in \mathbb{R}$ over measurable sets A_k, and then by taking suitable limits for the most general functions, called the μ-integrable functions.

The Lebesgue measure on $\mathbb{R}^d$ is built from the natural expression of volumes of parallelepipeds, i.e. sets of the form

$$Q = \prod_{k=1}^{d} [a_k, b_k], \quad \text{with} \quad \mathrm{Vol}(Q) = \prod_{k=1}^{d} (b_k - a_k).$$

As usual we write $\int_{\mathbb{R}^d} f(x)dx$ for the integral with respect to the Lebesgue measure, and of course we have for any $f \in \boldsymbol{C}_c(\mathbb{R}^d)$ that this integral exists and equals the usual Riemann integral (taken over a compact domain).
The space $\boldsymbol{\mathcal{L}}^1(\mathbb{R}^d)$ consists of all measurable functions on $\mathbb{R}^d$ which have finite absolute integral

$$\|f\|_1 := \int_{\mathbb{R}^d} |f(x)|dx < \infty.$$

Since this is only a *semi-norm*, meaning that $\|f\|_1 = 0$ only implies that $f(x) = 0$ almost everywhere, one resort to the definition of $\left(\boldsymbol{L}^1(\mathbb{R}^d), \, \|\cdot\|_1\right)$ as the vector space of *equivalence classes* of Lebesgue integrable functions.
The important facts are given on the following page:

1. $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ is a Banach space, containing $\boldsymbol{C}_c(\mathbb{R}^d)$ as dense subspace;

2. $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ is even a Banach algebra with respect to convolution, defined by the usual convolution integral,

$$f \star g(x) = \int_{\mathbb{R}^d} g(x-y)f(y)dy \tag{75}$$ convptwdef

which exists almost everywhere and defines another element (named $f \star g$) in $\boldsymbol{L}^1(\mathbb{R}^d)$ satisfying the expected submultiplicativity

$$\|f \star g\|_{\boldsymbol{L}^1} \leq \|f\|_{\boldsymbol{L}^1} \|g\|_{\boldsymbol{L}^1}, \quad f, g \in \boldsymbol{L}^1(\mathbb{R}^d). \tag{76}$$ Lisubmul1

3. On $\boldsymbol{L}^1(\mathbb{R}^d)$ the *Fourier transform* is well defined as

$$\hat{f}(\omega) = \int_{\mathbb{R}^d} f(t) \cdot e^{-2\pi i \omega \cdot t}\, dt \tag{77}$$ FT-def00

and defines a function $\widehat{f} \in \boldsymbol{C}_0(\mathbb{R}^d)$ (Riemann-Lebesgue Lemma), turning convolution into pointwise multiplication (Convolution Theorem)

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

4. Whenever $\widehat{f} \in \boldsymbol{L}^1(\mathbb{R}^d)$ the original function f can be recovered pointwise by the *inverse Fourier transform integral*

$$f(t) = \int_{\mathbb{R}^d} \hat{f}(\omega) \cdot e^{2\pi i t \cdot \omega}\, d\omega. \tag{78}$$ IFT-def00

The main facts used from Lebesgue integration theory are statements which guarantee the interchange of integration and (pointwise) limits of functions, such as the Lebesgue Dominated Convergence Theorem and the Fubini-Tonelli Theorem, justifying the change of integration orders for multiple integrals. Finally, the specific properties of the Lebesgue integral justify the change of variables (usually just translation, or affine transformation laws, bringing in determinants of the involved linear transformation, such as rotation, dilation and so on).
Let us just formulate the **Dominated Convergence Principle**:

domconvprinc **Lemma 8.** *Let $(f_n)_{n \geq 1}$ be a sequence in $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$, and assume that we have*

1. $|f_n(x)| \leq |g(x)|$ *almost everywhere (pointwise domination) for some $g \in \boldsymbol{L}^1(\mathbb{R}^d)$;*

2. $f_0(x) = \lim_{n\to\infty} f_n(x)$ *exists almost everywhere.*

Then $f_0 \in \boldsymbol{L}^1(\mathbb{R}^d)$ and $\lim_{n\to\infty} \|f_n - f_0\|_{\boldsymbol{L}^1} = 0$.

At a technical level it appears obvious that Fourier Analysis, as formulated above, cannot be done without Lebesgue integration, and that the *completeness* property of $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ is an essential building block for the Fourier transform.
On the other hand, the focus on integration techniques narrows the view and puts emphasize on the technicalities, and not on the content of Fourier Analysis, as this lecturer is seeing it! Comments on the dual of $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ (usually described as $\left(\boldsymbol{L}^\infty(\mathbb{R}^d), \|\cdot\|_\infty\right)$, to be given later!).

Intermezzo, just for 08.10.2020

Let us summarize the situation:
We have the **(pointwise) Banach algebra** $\left(\boldsymbol{C}_0(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$ as our starting point. It is also translation invariant and flip-invariant, containing $\boldsymbol{C}_c(\mathbb{R}^d)$ as a dense subspace. Hence most of the time we can reduce analytic statements by working with elements of this subspace only, all of which are Riemann integrable!
Justified by the Riesz Representation Theorem (see books on functional analysis) it is justified to just *call* the bounded linear functionals on $\left(\boldsymbol{C}_0(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$ **bounded measures**. The *dual space* (of all such linear functionals) is another Banach space, denoted by $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$.
Even without knowing much about $\left(\boldsymbol{C}_0(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$ and its dual $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ we can already establish the key connection between *translation invariant linear systems* T and linear functionals μ representing them, by the *isometric!* pairing

$$T_\mu f(y) = \mu(T_y f^{\checkmark}) \quad \text{for} \quad \mu_T(f) = T(f^{\checkmark})(0). \tag{79}$$

sysfuncrelCIT

There are two very specific subspaces of $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ which are very useful, namely the subspace of *finite, discrete measures* of the form $\mu = \sum_{k=1}^n c_k \delta_{x_k}$ (let us call them the "simple" measures) and the "nice" measures, which have a *density* $k \in \boldsymbol{C}_c(\mathbb{R}^d)$:

$$\mu(f) = \int_{\mathbb{R}^d} f(x)\ k(x)dx, \quad f \in \boldsymbol{C}_0(\mathbb{R}^d),$$

(the integral can be a Riemann integral, since the integrand is continuous).
While both of them generate (almost) disjoint (only $\mu = 0$ is common to them) closed subspaces of $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$, they are both w^*-dense in $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$, which is enough to get full information about the most general case by taking (suitable) limits of such (simple or nice) easier to understand objects within $\boldsymbol{M}_b(\mathbb{R}^d)$. In fact, the closure of the space of simple measures is the collection of bounded, discrete measures (or absolutely convergent sums of Dirac measures) resp. $\left(\boldsymbol{L}^1(\mathbb{R}^d),\, \|\cdot\|_1\right)$ (to be discussed in more detail below).
Since the collection of all TILS forms a closed (!commutative, as we have seen) subalgebra of $\mathcal{L}(\boldsymbol{C}_0(\mathbb{R}^d))$ (with the operator norm) it is natural to transfer the composition structure of that algebra of operators to the (after all isometrically isomorphic) space, namely $\boldsymbol{M}_b(\mathbb{R}^d)$.
Finally we have explained that the different ideas about convolution which we have introduced as the action of $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ on $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ by $f \mapsto C_m(f) = \mu \star f)$ respectively the internal composition (also called convolution), characterized by

$$C_{\mu_1 \star \mu_2} = C_{\mu_1} \circ C_{\mu_2}$$

are fully compatible and satisfy (in our context so far) the expected algebraic rules, such as *associativity, commutativity* or the *distributive law*.
We still have to get more information on convergence issues, in order to understand the connection to impulse response and transfer function, widely used in system theory.

5.4 Lebesgue space as completion

Based on these facts we can describe, or actually *introduce* the space $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ as a closed subspace of $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$.
Recall that we have already an isometric embedding of $\boldsymbol{C}_c(\mathbb{R}^d)$, endowed with the norm $\|\cdot\|_1$ into $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$, by assigning to $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ the measure

$$\mu_k : f \mapsto \int_{\mathbb{R}^d} f(x)k(x)dx, \quad f \in \boldsymbol{C}_0(\mathbb{R}^d),$$

i.e. with

$$\|k\|_{\boldsymbol{L}^1} = \|\mu_k\|_{\boldsymbol{M}_b(\mathbb{R}^d)}.$$

Liintro **Definition 3.** $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ is the closure of $\{\mu_k \mid k \in \boldsymbol{C}_c(\mathbb{R}^d)\}$ in $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$.

Few facts concerning $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right) \subset (\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$:

charLisp **Proposition 3.** *$\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ is a closed ideal in $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$. It coincides with the (closed) subspace of elements with continuous shift, given by*

$$\boldsymbol{M}_c(\mathbb{R}^d)(\mathbb{R}^d) := \{\mu \in \boldsymbol{M}_b(\mathbb{R}^d) \mid \|T_x\mu - \mu\|_{\boldsymbol{M}_b} \to 0 \textit{ for } x \to 0\}.$$

Being a closed subspace of a Banach space $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ is itself a Banach space, and hence absolutely convergent series are convergent.

Proof. By definition $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ is a closed subspace of $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$.
In order to show that it is a closed ideal we can resort to the fact that for compactly supported measures μ and for $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ one has $\mu \star k \in \boldsymbol{C}_c(\mathbb{R}^d)$ (see the proof that $\boldsymbol{M}_b(\mathbb{R}^d) \star \boldsymbol{C}_0(\mathbb{R}^d) \subset \boldsymbol{C}_0(\mathbb{R}^d)$). By taking limits one finds that $\boldsymbol{M}_b(\mathbb{R}^d) \star \boldsymbol{L}^1(\mathbb{R}^d) \subseteq \boldsymbol{L}^1(\mathbb{R}^d)$.
In order to show that $\boldsymbol{L}^1(\mathbb{R}^d) \subset \boldsymbol{M}_{bc}(\mathbb{R}^d)$ (without resorting to the characterization of $\boldsymbol{L}^1(\mathbb{R}^d)$ as the space of Lebesgue integrable functions!) we observe that the uniform continuity of $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ implies that

$$\lim_{x \to 0} \|T_x k - k\|_{\boldsymbol{L}^1} = 0,$$

In fact, the family $(T_x k)_{|x| \le 1}$ has joint support in some compact set Q. Hence one has (writing $|Q|$ for the volume of Q):

$$\|T_x k - k\|_{\boldsymbol{L}^1} \le |Q| \|T_x k - k\|_\infty \to 0 \text{ for } x \to 0. \tag{80}$$ contshift02

That $\boldsymbol{M}_{bc}(\mathbb{R}^d)$ is a closed ideal of $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ (with respect to convolution) is quite easy to verify. The closedness is easy to check (left to the reader), and the ideal property follows from the fact that convolution commutes with translation (!also inside of $\boldsymbol{M}_b(\mathbb{R}^d)$, since

$$\delta_x \star (\nu \star \mu) = (\delta_x \star \nu \star) \star \mu, \quad x \in \mathbb{R}^d, \mu, \nu \in \boldsymbol{M}_b(\mathbb{R}^d),$$

hence one has for $\nu \in \boldsymbol{M}_{bc}(\mathbb{R}^d)$ and $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$:

$$\|T_x(\nu \star \mu) - \nu \star \mu\|_{\boldsymbol{M}_b} \le \|T_x\nu - \nu\|_{\boldsymbol{M}_b} \|\mu\|_{\boldsymbol{M}_b} \to 0 \text{ for} |x| \to 0.$$

The fact that any $\mu \in \boldsymbol{M}_{bc}(\mathbb{R}^d)$ can be approximated by functions $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ will establish the coincide with this space with $\boldsymbol{L}^1(\mathbb{R}^d)$.
The argument used will consists of two parts. For $k \in \boldsymbol{C}_c(\mathbb{R}^d) \subset \boldsymbol{L}^1(\mathbb{R}^d)$ we know (cf. above) that $\mu \star k \in \boldsymbol{L}^1(\mathbb{R}^d)$. But we will use the continuity of translation in $\boldsymbol{M}_{bc}(\mathbb{R}^d)$ in order to show that $\mu \star k$ (more correctly read as $\mu \star \mu_k$, via convolution within $\boldsymbol{M}_b(\mathbb{R}^d)$) approximates μ.
Given $\mu \in \boldsymbol{M}_{bc}(\mathbb{R}^d)$ and $\varepsilon > 0$ we can find some $\delta > 0$ such that $\|T_x\mu - \mu\|_{\boldsymbol{M}_b} < \varepsilon/2$ for $|x| \leq \delta$. For any $k \in \boldsymbol{C}_c(\mathbb{R}^d)$, non-negative, with $\text{supp}(k) \subset B_{\delta/2}(0)$ and any BUPU Ψ with $\Psi| < \delta/2$.
The continuity of translation within $\boldsymbol{M}_c(\mathbb{R}^d)$ implies that one has

$$\|\mu \star k - \mu\|_{\boldsymbol{M}_b} \leq \varepsilon$$

for any $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ with sufficiently small support and $\int_{\mathbb{R}^d} k(x)dx = 1$ implies that μ can be approximated by function $\mu \star k \in \boldsymbol{L}^1(\mathbb{R}^d)$, because we know already that it is closed inside $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$. Consequently $\boldsymbol{M}_c(\mathbb{R}^d) \subseteq \boldsymbol{L}^1(\mathbb{R}^d)$, and the proof is complete. □

CbLiinLi **Corollary 1.**

$$\boldsymbol{C}_b(\mathbb{R}^d) \cdot \boldsymbol{L}^1(\mathbb{R}^d) = \boldsymbol{L}^1(\mathbb{R}^d) \quad \textit{and} \quad \boldsymbol{C}_0(\mathbb{R}^d) \cdot \boldsymbol{L}^1(\mathbb{R}^d) = \boldsymbol{L}^1(\mathbb{R}^d)^{7}.$$

Proof. Either characterization of $\boldsymbol{L}^1(\mathbb{R}^d)$ can be used. E.g. $\mu h \in \boldsymbol{M}_c(\mathbb{R}^d)$ for $\mu \in \boldsymbol{M}_c(\mathbb{R}^d)$ and $h \in \boldsymbol{C}_{ub}(\mathbb{R}^d)$, due to the estimate

$$\|T_x(\mu h) - \mu h\|_{\boldsymbol{M}_b} \leq \|(T_x\mu - \mu) \cdot T_x h\|_{\boldsymbol{M}_b} + \|\mu \cdot (T_x h - h)\|_{\boldsymbol{M}_b}. \tag{81}$$ McCubMult

The second claim, i.e. the statement that every $f \in \boldsymbol{L}^1(\mathbb{R}^d)$ can be written in the form $f = f_1 \cdot h$, for some $f_1 \in \boldsymbol{L}^1(\mathbb{R}^d)$ and some $h \in \boldsymbol{C}_0(\mathbb{R}^d)$ will follow from the *Cohen-Hewitt Factorization Theorem* (see **hero70 :** [61]). As it is not needed now we postpone the proof. □

We can also establish a link between "our approach" and the realization of the *completion* as set of functions well defined almost everywhere.

Lisum1 **Lemma 9.** *The elements $f \in \boldsymbol{L}^1(\mathbb{R}^d)$ can be represented as those measures $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ which can be represented as absolutely convergent series of elements $k_n \in \boldsymbol{C}_c(\mathbb{R}^d)$, i.e.*

$$f = \sum_{n=1}^{\infty} k_n, \quad \textit{with} \quad \sum_{n=1}^{\infty} \|k_n\|_{\boldsymbol{L}^1} < \infty.$$

Proof. This claim is based on the general functional analytic fact which allows to characterize the closure of a subspace in a Banach space with the collection of absolutely convergent series. □

Remark 2. One of the basic facts concerning $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ (starting from the Lebesgue theory) is the fact that the condition $\sum_{n=1}^{\infty} \|k_n\|_{\boldsymbol{L}^1} < \infty$ implies that $\sum_{n=1}^{\infty} k_n(x)$ is absolutely convergent almost everywhere and the resulting function g is a representative

[7] In both cases we use the "complex product", i.e. $\boldsymbol{C}_0(\mathbb{R}^d) \cdot \boldsymbol{L}^1(\mathbb{R}^d) := \{hf \mid h \in \boldsymbol{C}_0(\mathbb{R}^d), f \in \boldsymbol{L}^1(\mathbb{R}^d)\}$

for the equivalence class which defines the measure $\mu_g \in \boldsymbol{L}^1(\mathbb{R}^d)$ given by (the Lebesgue integration, or via Riemann integrals for $\boldsymbol{C}_c(\mathbb{R}^d)$):

$$f \mapsto \int_{\mathbb{R}^d} f(x)g(x)dx,$$

also with the isometric embedding

$$\|\mu_g\|_{\boldsymbol{M}_b} = \|g\|_{\boldsymbol{L}^1}, \quad g \in \boldsymbol{L}^1(\mathbb{R}^d).$$

From the measure theoretic point of view $\boldsymbol{L}^1(\mathbb{R}^d)$ can be characterized as the subset of *absolutely continuous* measures, i.e. those measures $\nu \in \boldsymbol{M}_b(\mathbb{R}^d)$ with the property that Lebesgue null-sets are also null-sets with respect to ν. The direct direction of obvious, since for a set M of measure zero on has $\int_M g(x)dx = 0$. The converse is based on the Theorem of Radon-Nikodym.

intL1A **Lemma 10.** *The mapping $f \to \int_{\mathbb{R}^d} g(x)dx$ which is well defined on $\boldsymbol{C}_c(\mathbb{R}^d)$ and continuous on $(\boldsymbol{C}_c(\mathbb{R}^d), \|\cdot\|_1)$ extends in a unique way to a bounded linear functional on $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$*[8]

Proof. Any $f \in \boldsymbol{L}^1(\mathbb{R}^d)$ is the limit (in the norm of $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$) of a sequence $\mu_{k_n}, n \geq 1$ of "nice measures" induced by a family of Riemann integrable functions $k_n, n \geq 1$ in $\boldsymbol{C}_c(\mathbb{R}^d)$. Hence they are clearly Lebesgue integrable and

$$L - \int_{\mathbb{R}^d} k_n(x)dx = R - \int_{\mathbb{R}^d} k_n(x)dx.$$

(Lebesgue integral equals Riemann integral for $k \in \boldsymbol{C}_c(\mathbb{R}^d)$).
But due to the isometric property ($\|\mu_k\|_{\boldsymbol{M}_b} = \|(\|_{\boldsymbol{L}^1} k)$ for $k\ in \boldsymbol{C}_c(\mathbb{R}^d)$) these integrals have to form a Cauchy sequence, since

$$\left|\int_{\mathbb{R}^d} k_n(x)dx - \int_{\mathbb{R}^d} k_m(x)dx\right| \leq \|k_n - km\|_{\boldsymbol{L}^1} \to 0 \quad \text{for} \quad n, m \to \infty,$$

and so we can *define* (formally, without knowing whether the values $f(x)$ exist or not (if we do not want to use Lebesgue integration theory), by

$$\text{“} \int_{\mathbb{R}^d} f(x)dx'' = \int_{\mathbb{R}^d} f = lim_{n\to\infty} \int_{\mathbb{R}^d} k_n(x)dx.$$

Of course, for those who have learned in a course on measure theory that $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ is a Banach space it will be clear that $f = \boldsymbol{L}^1 - \lim_n k_n$ belongs to $\boldsymbol{L}^1(\mathbb{R}^d)$ and is Lebesgue integrable, and the formal limit makes also sense in a pointwise almost everywhere sense (even if it might be necessary to replace the Cauchy sequence (k_n) by some subsequence in order to achieve $f(x) = \lim_{n\to\infty} k_n(x)$ almost everywhere).

□

[8] While the integral is exactly defined in the sense of Lebesgue for $f \in \boldsymbol{L}^1(\mathbb{R}^d)$, viewed in the usual sense, it is just a symbol for the extended linear functional, if we write $\int_{\mathbb{R}^d} g(x)dx$ for general $g \in \boldsymbol{L}^1(\mathbb{R}^d)$. It should thus be interpreted, whenever necessary, as limit of $\int_{\mathbb{R}^d} g_n(x)dx$, for any sequence $(g_n)_{n\geq 1}$ approximating $g \in \boldsymbol{L}^1(\mathbb{R}^d)$ via elements $g_n \in \boldsymbol{C}_c(\mathbb{R}^d)$.

6 Extension of measures to $\big(\boldsymbol{C}_b(\mathbb{R}^d),\ \|\cdot\|_\infty\big)$

First we extend the action of bounded linear functionals $\mu \in (\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b}) = \big(\boldsymbol{C}_0'(\mathbb{R}^d),\ \|\cdot\|_{\boldsymbol{C}_0'}\big)$ to all of $\big(\boldsymbol{C}_b(\mathbb{R}^d),\ \|\cdot\|_\infty\big)$, by the simple observation that one has for any $h \in \boldsymbol{C}_b(\mathbb{R}^d)$ and any BUPU $(\psi_i)_{i\in I}$ by setting

$$\mu(h) = \sum_{i\in I} \mu(\psi_i h). \tag{82}$$

muCbact

This is defining a bounded and of course linear functional on $\big(\boldsymbol{C}_b(\mathbb{R}^d),\ \|\cdot\|_\infty\big)$, due to the formula $\|\mu\|_{\boldsymbol{M}_b} = \sum_{i\in I} \|\mu\,\psi_i\|_{\boldsymbol{M}_b}$, combined with the fact that there exists a family ψ_i^* in $\boldsymbol{C}_c(\mathbb{R}^d)$ with $\|\psi_i^*\|_\infty = 1$ and $\psi_i^* \cdot \psi_i = \psi_i$, allowing the estimate

$$|\mu(h)| = \sum_{i\in I} |\mu(\psi_i h)| = \sum_{i\in I} |(\mu\psi_i)(\psi_i^* h)| \le \sum_{i\in I} \|\mu\psi_i\|_{\boldsymbol{M}_b} \|\psi_i^* h\|_\infty \le \|h\|_\infty \sum_{i\in I} \|\mu\psi_i\|_{\boldsymbol{M}_b} \tag{83}$$

Mbactest1

hence

$$|\mu(h)| \le \|h\|_\infty \|\mu\|_{\boldsymbol{M}_b}, \quad h \in \boldsymbol{C}_b(\mathbb{R}^d), \mu \in \boldsymbol{M}_b(\mathbb{R}^d), \tag{84}$$

Mbactest2

i.e. the functional μ defined a priori only for $\big(\boldsymbol{C}_0(\mathbb{R}^d),\ \|\cdot\|_\infty\big)$ is now extended to the larger space $\big(\boldsymbol{C}_b(\mathbb{R}^d),\ \|\cdot\|_\infty\big)$, with the same norm estimate. Thus the norm of the linear functional $h \to \mu(h)$ on $\big(\boldsymbol{C}_b(\mathbb{R}^d),\ \|\cdot\|_\infty\big)$ equals the norm of the original functional on $\big(\boldsymbol{C}_0(\mathbb{R}^d),\ \|\cdot\|_\infty\big)$.

Remark 3. The above considerations imply the following harmless (but sometimes useful) estimate:

$$|\mu \cdot h| \le \|\mu\|_{\boldsymbol{M}_b} \|h\|_\infty, \quad \mu \in \boldsymbol{M}_b(\mathbb{R}^d), h \in \boldsymbol{C}_b(\mathbb{R}^d). \tag{85}$$

normmuh

The discussion above hides a little bit the question of independence of the description given above from the BUPU $(\psi_i)_{i\in I}$ used in formula (82). We will fix this problem by establishing an even more general fact:

MbCbindep1 **Lemma 11.** *For any $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ and any $h \in \boldsymbol{C}_b(\mathbb{R}^d)$ one has for any bounded approximate identity $(h_\beta)_{\beta\in J}$ in $\big(\boldsymbol{C}_0(\mathbb{R}^d),\ \|\cdot\|_\infty\big)$:*

$$\mu(h) = \lim_{\beta\in J} \mu(h_\beta h). \tag{86}$$

MbCblim1

Proof. First let us recall that we have $h_\beta h \in \boldsymbol{C}_0(\mathbb{R}^d) \cdot \boldsymbol{C}_b(\mathbb{R}^d) \subset \boldsymbol{C}_0(\mathbb{R}^d)$ and hence the expression $\mu(h_\beta h)$ is well defined, and we only have to show that for any such approximated identity (h_β) the limit exists and provides the same limit as used in definition (82), which by consequence is independent of the BUPU used.
Details to be given later .. □

The independence question is obtained by the observation that the partial sums $\sum_{i\in F} \psi_i$ (with $F \subset I$) form such a bounded approximation to the identity in $\big(\boldsymbol{C}_0(\mathbb{R}^d),\ \|\cdot\|_\infty\big)$.

Material for 08.10.2020 and later!

7 Impulse Response and Transfer Function

In engineering books one often finds: $T(\delta_0)$, the *impulse response* to the given system, fully characterizes the TILS T. Equivalently: The *transfer function* H, equivalently, describes the behaviour of T on the Fourier transform side. $Tf = \mathcal{F}^{-1}(H \cdot \widehat{f})$.
Of course such a description raises a couple of question:

- Is the unit impulse (δ_0 in our setting) in the domain of T, or how can T be extended to a domain which allows to view δ_0 as input; we will see that this is OK for BIBO systems, but not for arbitrary TILS on $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$.
- Even if the domain of T can be extended (from $\boldsymbol{C}_0(\mathbb{R}^d)$ to say $\boldsymbol{M}_b(\mathbb{R}^d)$, as we will see), in which sense is the usual argument valid, which pretends/claims that the impulse response is the limit (but in which sense?) of the output signals $T(b_n)$, of boxcar-functions b_n, with shrinking basis and unit area?
- What can we say about the convergence under the additional assumption that the impulse response belongs to $\boldsymbol{L}^1(\mathbb{R}^d)$, i.e. if it is an absolutely integrable, measurable function? We will see later that a Cauchy condition on $T(b_n)$ in $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ is equivalent to this additional assumption.
- Even if $h = T(\delta_0)$ exists (in whatever sense, as a function, or as a distribution, i.e. a generalized function), does it have a Fourier transform?
- In which sense is the description of $Tf = h \star f$ equivalent to $\mathcal{F}(Tf) = H \cdot \widehat{f}$?
  Of course one would expect that h (in the case of BIBOS systems it is $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$) has a Fourier transform and that $H = \mathcal{F}(h)$.
- Given a function H in the frequency domain (say a bounded function), can we define the operator $T(f) = \mathcal{F}^{-1}(H \cdot \mathcal{F}(f))$ and when is it a BIBOS system (or bounded on some other space, such as the Hilbert space $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$ (via Plancherel's Theorem). Is it justified to say that the description of TILS via impulse response or transfer functions are in fact *equivalent*?

For the discrete setting (periodic or non-periodic) the characterization of a TILS by the impulse response system is quite easy. In fact, the unit vector at the origin, or Dirac impulse, is an ordinary signal and all the other functions can be written as sums of a shifted Dirac measure.
However, in the case of continuous variable the Dirac measure does not belong to $\boldsymbol{L}^1(\mathbb{R}^d)$ anymore, even is it clear whether the operator T, which is a priori only defined (as bounded and linear operator) is well defined for the input signal δ_0! Consequently the "recipe" of calling $h := T(\delta_0)$ the *impulse response* may not be meaningful.
Hence our next goal is to actually show that T can be extended to a bounded operator on $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$, and thus $h \in T(\delta_0)$ is meaningful. In fact, we will show that $h = \sigma \in \boldsymbol{M}_b(\mathbb{R}^d)$, the "representing functional" for the operator T.

7.1 Coming back to the scandal story (Sandberg)

Since we are able to extend the action of a bounded measure, originally given on $\boldsymbol{C}_0(\mathbb{R}^d)$, we can also extend the TILS C_μ induced by the functional $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ to a TILS on $\left(\boldsymbol{C}_b(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$.

CbTILSch **Theorem 3.** *Any TILS on $\left(\boldsymbol{C}_0(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$ extends in a natural way to $\left(\boldsymbol{C}_b(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$. The extended operator, which we still denote by C_μ, has the extra property that for any bounded sequence $(h_n)_{n=1}^\infty$ which converges to some limit function $h_0 \in \boldsymbol{C}_b(\mathbb{R}^d)$, uniformly over compact sets the same is true for the output, i.e.*

$$\mu \star h_n(x) \to \mu \star h_0(x), \quad \textit{uniformly over compact sets in}\, \mathbb{R}^d. \tag{87}$$ convC0conv

Conversely, any TILS on $\left(\boldsymbol{C}_b(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$ with the extra property (87) is a convolution operator with some uniquely determined $\mu \in \boldsymbol{M}(\mathbb{R}^d)$ and the operator norm is equal to $\|\mu\|_{\boldsymbol{M}_b}$.

Although this claim looks quite natural it turns out that it is easier to provide a proof for the domain $\left(\boldsymbol{C}_{ub}(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$, more or less by the same argument as in the proof of the characterization of TILS on $\left(\boldsymbol{C}_0(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$ (just without the need to check for decay at infinity, of the output side).
However, if the input is only continuous and *uniformly continuous* we have to provide an extra argument in order to derive the continuity (at each point) as well. Of course we can use that for any compact subset the uniform convergence of continuous functions will be continuous.
Thus, without going into details, we just claim that this is valid from the fact that first it can be shown to be true for measures with compact support, and then, for general measures it follows by an approximation argument. The key observation (expressed via BUPUs) is contained in the following lemma:

convlocal1 **Lemma 12.** *Given any compact set $Q \subset \mathbb{R}^d$, and a measure $\mu withcompactsupport$, i.e. with $\mu = \mu \sum_{i\in F} \mu\psi_i$ one can find a finite set $E \subset I$ such that*

$$\mu \star f(x) = \mu \star \sum_{j\in E} f\psi_j$$

for any $f \in \boldsymbol{C}_b(\mathbb{R}^d)$.

Proof. In fact, as in one of the previous proofs, one finds that the support of $\psi_i \cap T_y(\psi_j{}^{\checkmark}) \neq 0$ for some $y \in Q$ only for finitely many indices $j \in I$. $\square$

Alternative argument (for the same claim). Assume that for a regular BUPU, with centers $k \in \mathbb{Z}^d$, one has

$$\mathrm{supp}(\psi_k) \subset B_1(k), k \in \mathbb{Z}^d.$$

Since only finitely many indices $i \in \mathbb{Z}^d$ are involved we have $\mathrm{supp}(\psi_i) \subset K$ for some compact set $K \subset \mathbb{R}^d$.
$\mathrm{supp}(\psi_j{}^{\checkmark}) \subset B_1(y-j) \subset Q + B_1(0)$. and $\mathrm{supp}(\psi_i) \subset B_1(i)$. The intersection can only be non-zero if $j \notin B_2(0) + Q + K$, but this is just true for a finite subset of indices.
CHECK: Maybe Q or $-Q$, K or $-K$?

7.2 The use of adjoint operators

The general idea will be to make use of the *adjoint operator*. This will be an important concept from functional analysis, especially for the treatment of distributions later on. Let us quickly recall a basic fact about *adjoint operators* on the dual of some Banach space.

defadjBan1 **Definition 4.** Given an operator $T \in \mathcal{L}(\boldsymbol{B})$ the *adjoint operator* T' is the bounded linear operator defined on $(\boldsymbol{B}', \|\cdot\|_{\boldsymbol{B}'})$, given by

$$T'\sigma(f) = \sigma(Tf), \quad f \in \boldsymbol{B}, \sigma \in \boldsymbol{B}'. \tag{88}$$ defadj03

propadj03 **Lemma 13.** *The operator T' is a bounded linear operator on $(\boldsymbol{B}', \|\cdot\|_{\boldsymbol{B}'})$ and*

$$|||T'|||_{\boldsymbol{B}'} = |||T|||_{\boldsymbol{B}}, \quad T \in \mathcal{L}(\boldsymbol{B}).$$

Proof. This is a standard result from function analysis, which is based on the Hahn-Banach Theorem. □

Note: We will show that the adjoint action (to the convolution operator $C_\nu : f \mapsto \nu \star f$, defined on $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$) on $\boldsymbol{M}_b(\mathbb{R}^d)$ (viewed as the dual of $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$, on which $\boldsymbol{M}_b(\mathbb{R}^d)$ acts through convolution operators), can be expressed by convolution with the flipped version of ν. Thus we have for any $\nu \in \boldsymbol{M}_b(\mathbb{R}^d)$:

$$|||C'_\nu|||_{\boldsymbol{M}_b(\mathbb{R}^d)} = |||C_\nu|||_{\boldsymbol{C}_0(\mathbb{R}^d)} = \|\nu\|_{\boldsymbol{M}_b} = \|\nu^\checkmark\|_{\boldsymbol{M}_b}.$$

The reasoning behind this additional viewpoint is the fact that adjoint operators are known to be w^*-w^*--continuous, i.e. they map w^*-convergent sequences in a dual space into w^*-convergent sequences, in our case this means

$$\mu_n(f) \to \mu_0(f) \;\forall f \in \boldsymbol{C}_0(\mathbb{R}^d) \Rightarrow T'\mu_n(h) \to T'\mu_0(h) \;\forall h \in \boldsymbol{C}_0(\mathbb{R}^d). \tag{89}$$ adjTwwst

We will see that the adjoint operator for an operator $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$ commuting with translation is also an operator (now on $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$) commuting with translation. So do we have *yet another convolution*??
Any adjoint operator is not only a bounded operator on $(\boldsymbol{B}', \|\cdot\|_{\boldsymbol{B}'})$, but it is also w^*-w^*--continuous, but this will be discussed later. In fact, any such w^*-w^*- operator S is the adjoint of some operator $T \in \mathcal{L}(\boldsymbol{B})$.
The abstract argument, showing that adjoint operators preserve w^*-w^*- convergence, is easy to verify (see (89) above).

Lemma 14. *Any adjoint operator preserves w^*-convergence.*

Proof. We work with the setting of (89). Given the w^*-convergent sequence (μ_n) in the dual space, $h \in \boldsymbol{C}_0(\mathbb{R}^d)$ (in the underlying Banach space), and $\varepsilon > 0$ we have to find an index n_0 such that

$$|T'(\mu_n)(h) - T'(\mu_0)(h)| < \varepsilon, \quad \forall n \geq n_0.$$

Rewriting the difference term which has to be estimated as $|\mu_n(T(h)) - \mu_0(T(h))|$ we find that we just have to apply the assumption of the specific choice $f := T(h)$ and choose n_0 appropriately. □

Next let us identify the adjoint action of the convolution operator [9] $C_\mu : f \mapsto \mu \star f$. In the discussion we will observe that the involution in $\left(\boldsymbol{C}_0(G), \|\cdot\|_\infty\right)$, given by $f \mapsto f^\checkmark$ (with $f^\checkmark(x) = f(-x)$) plays an important role. It can be also "extended" to measures, by the definition $\mu^\checkmark(f) = \mu(f^\checkmark)$. It is a good exercise to check yourself that $(\delta_x)^\checkmark = \delta_{-x}$, hence $\delta_0 = \delta_0{}^\checkmark$. Also the *consistency condition*

$$\mu_{k^\checkmark} = (\mu_k)^\checkmark, \quad k \in \boldsymbol{C}_c(\mathbb{R}^d),$$

confirm that this is the "right definition".
THE FOLLOWING CLAIM WILL APPEAR BELOW WITH A SHORT PROOF!

consistMbCG1 **Proposition 4.** *The adjoint operator of the convolution operator* $C_\mu : f \to \mu \star f$ *on* $\left(\boldsymbol{C}_0(G), \|\cdot\|_\infty\right)$ *is the (internal) convolution operator* $\nu \to \mu^\checkmark \star \nu$.

We could reformulate the claim of the proposition as the relationship

$$C'_\nu = C_{\nu^\checkmark}, \quad \nu \in \boldsymbol{M}_b(\mathbb{R}^d). \tag{90}$$ adjconvflip2

Note: it is trivial that the mapping $\mu \to \mu^\checkmark$ is not only norm-norm continuous on $(\boldsymbol{M}_b(G), \|\cdot\|_{\boldsymbol{M}_b})$ but also w^*-w^*- continuous, as it the adjoint mapping to $f \to f^\checkmark$, i.e. $\mu^\checkmark(f) = \mu(f^\checkmark)$, $f \in \boldsymbol{C}_0(G)$.

Proof. In both cases we know that it is enough to verify the relationship for the bounded, discrete measures, i.e. finally just for the Dirac measures. So let us write (for a moment) $\mu \star' \nu$ for the *adjoint action* of μ on ν, which is defined as usual by

$$[(C_\mu)'(\nu)](f) =_{\text{or}} [\mu \star' \nu](f) := \nu(\mu \star f) = \nu(C_\mu(f)). \tag{91}$$ adjconv02

We test it for $\mu = \delta_x$ and $\nu = \delta_y$. Then we have

$$\delta_x \star' \delta_y(f) = \delta_y(\delta_x \star f) = \delta_y(T_x f) = f(y-x) = \delta_{y-x}(f) = (\delta_{-x} \star \delta_y)(f), \quad \forall x, y \in \mathbb{R}^d, \tag{92}$$ convadj1

telling us (since obviously $\delta_x{}^\checkmark = \delta_{-x}$) that we can describe the adjoint action on $\boldsymbol{M}_b(\mathbb{R}^d) = \boldsymbol{C}'_0(\mathbb{R}^d)$ by convolution with the flipped measure:

$$\delta_x \star' \delta_y = \delta_x{}^\checkmark \star \delta_y \ . \tag{93}$$ adjconvDirac

$\square$

[9] The use of this symbol is now justified!

THIS PART IS NOT USED!

In order to derive the same claim for general measures we first extend it from individual Dirac measures to bounded, discrete measures, and then by approximation to general measures.
The fact that $D_\Psi\mu(f) = \mu(\mathrm{Sp}_\Psi(f)) \to \mu(f)$ for $|\Psi| \to 0$ implies that the bounded (and hence the finite) discrete measures are w^*-dense in $\boldsymbol{M}_b(\mathbb{R}^d)$ (in a constructive way).
Formally this means: Given $f_1, ..., f_n \in \boldsymbol{C}_0(\mathbb{R}^d)$ and $\varepsilon > 0$ there exists some $\delta > 0$ such that one has for $0 < \delta \leq \delta_0 > 0$:

$$|D_\Psi\mu(f_k) - \mu(f_k)| \leq \varepsilon, \quad \text{for} k = 1, ...n.$$

It will be helpful to work with regular BUPUs Ψ arising as shifts of $\psi = \psi_0$ which is supposed to be symmetric, i.e. satisfying $\psi = \psi^\checkmark$. Then one has:

$$\mathrm{D}_\Psi(\mu^\checkmark) = (D_\Psi\mu)^\checkmark, \quad \mu \in \boldsymbol{M}_b(\mathbb{R}^d). \tag{94}$$ Dpsicheck

We can use the fact that for $\lambda\in\Lambda$ one has $\psi_{-\lambda} = (\psi_\lambda)^\checkmark$, is $\psi = \psi^\checkmark$ (i.e. if the BUPU is made up by a symmetric function ψ).
In fact, we have for any regular BUPU $(\psi_\lambda)_{\lambda\in\Lambda}$: $\mu(\psi_{-\lambda}) = \mu([\psi_\lambda]^\checkmark) = \mu^\checkmark(\psi_\lambda), \lambda\in\Lambda$.

$$(D_\Psi\mu)^\checkmark(f) = D_\Psi\mu(f^\checkmark) = \mu(\mathrm{Sp}_\Psi(f^\checkmark)) = \sum_{\lambda\in\Lambda} f^\checkmark(\lambda)\mu(\psi_\lambda)$$

$$= \sum_{\lambda\in\Lambda} f(-\lambda)\mu(\psi_\lambda) =_{\lambda\to-\lambda} \sum_{\lambda\in\Lambda} f(\lambda)\mu(\psi_{-\lambda}) = \sum_{\lambda\in\Lambda} f(\lambda)\mu^\checkmark(\psi_\lambda) = \mathrm{D}_\Psi\mu^\checkmark(f).$$

By summing up over finitely many Dirac measures and then taking limits (in fact, using absolutely convergent series) one obtains: for any (finite or even bounded) discrete measure ν and any $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ one has:

$$C'_{D_\Psi\mu}(\nu) = D_\Psi\mu^\checkmark \star \nu, \tag{95}$$ DPsiadj1

In particular we have for $\mu, \nu \in \boldsymbol{M}_b(\mathbb{R}^d)$ and any BUPU Ψ:

$$C'_{D_\Psi\mu}(\mathrm{D}_\Psi\nu) = D_\Psi\mu^\checkmark \star (\mathrm{D}_\Psi\nu), \tag{96}$$ DPsiadj2

where $'\star'$ denotes internal convolution within $\boldsymbol{M}_b(\mathbb{R}^d)$.
Taking limits on both sides implies the claim made in proposition consistMbCG1 4. WE JUST HAVE TO BE CAREFUL in taking limits.
In fact, we have by the adjointness definition for $f \in \boldsymbol{C}_0(\mathbb{R}^d)$:

$$C'_\mu(\nu)(f) = \nu(C_\mu \star f) = \lim_{|\Psi|\to 0} \nu(C_{D_\Psi\mu} \star f).$$

On the other hand we find, using $\mu^\checkmark \star \nu = \nu \star \mu^\checkmark$ that

$$\|(\nu \star \mu^\checkmark) \star f - (\nu \star [D_\Psi\mu]^\checkmark) \star f\|_\infty \leq \|\nu\|_{\boldsymbol{M}_b} \|[\mu - D_\Psi\mu]^\checkmark f\|_\infty$$

Recall that we define $\mu^{\checkmark}$ via $\mu^{\checkmark}(f) = \mu(f^{\checkmark})$, and that $(\delta_x)^{\checkmark} = \delta_{-x}, x \in \mathbb{R}^d$.

convflipmf **Lemma 15.** *Given $\nu \in \boldsymbol{M}_b(\mathbb{R}^d)$ and $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ one has*

$$\nu^{\checkmark} \star f^{\checkmark} = (\nu \star f)^{\checkmark}. \tag{97}$$ flipconv

Proof. Using the fact that $(f^{\checkmark})^{\checkmark} = f$ and $(T_y f)^{\checkmark} = T_{-y} f^{\checkmark}$ we have:

$$[\nu^{\checkmark} \star f^{\checkmark}](y) = \nu^{\checkmark}(T_y f) = \nu((T_y f)^{\checkmark}) = \nu(T_{-y} f^{\checkmark}) = \nu \star f(-y) = (\nu \star f)^{\checkmark}(y).$$

□

Remark 4. Of course one can verify directly, writing out integrals and applying the substitution $z \to -z$ that one has of course also the pointwise relation

$$k_1{}^{\checkmark} \star_{pw} k_2{}^{\checkmark} = (k_1 \star_{pw} k_2)^{\checkmark}, \quad k_1, k_2 \in \boldsymbol{C}_c(\mathbb{R}^d). \tag{98}$$ convpwflip

adjflipconv **Proposition 5.** *Given $\nu \in \boldsymbol{M}_b(\mathbb{R}^d)$ the adjoint action for the convolution operator $C_\nu : f \mapsto \nu \star f$ coincides with the internal convolution by $\nu^{\checkmark}$, or in formulas*

$$C_\nu'(\mu) = \nu^{\checkmark} \star \mu, \quad \mu \in \boldsymbol{M}_b(\mathbb{R}^d). \tag{99}$$ adjfliprel

Proof. We have to show that the abstractly defined measure $C_\nu'(\mu)$ acts on any given $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ in the same way as the bounded measure $\nu^{\checkmark} \star \mu$. Thus let $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ be given. Then we have by Lemma convflipmf 15 above

$$[C_\nu'(\mu)](f) = \mu(\nu \star f) = [\mu \star (\nu \star f)^{\checkmark}](0) = [(\mu \star \nu^{\checkmark}) \star (f^{\checkmark})](0) = (\nu^{\checkmark} \star \mu)(f).$$

□

In fact, for practical applications the following formula which is equivalent to (adjfliprel 99) will be used (replacing $\nu^{\checkmark}$ by ν)

$$\mu(\nu^{\checkmark} \star f) = (\nu \star \mu)(f), \quad \forall f \in \boldsymbol{C}_0(\mathbb{R}^d). \tag{100}$$ adjflip

Using this formula it is easy to show that $C_\nu : \mu \mapsto \nu \star \mu$ (internal convolution) preserves w^* convergence:

convnuwst **Lemma 16.** *Assume that a bounded sequence $(\mu_n)_{n\geq 1}$ in $\boldsymbol{M}_b(\mathbb{R}^d)$ is w^*-convergent, with $\mu_0 = w^*\text{-}\lim_{n\to\infty} \mu_n$. Then for any $\nu \in \boldsymbol{M}_b(\mathbb{R}^d)$*

$$w^*\text{-}\lim_{n\to\infty} \nu \star \mu_n = \nu \star \mu_0.$$

Proof. By assumption we have $\lim_n \mu_n(f) = \mu_0(f)$, for any $f \in \boldsymbol{C}_0(\mathbb{R}^d)$. Using (adjflip 100) we see that

$$\nu \star \mu_n(f) = \mu_n(\nu^{\checkmark} \star f) \to \mu_0(\nu^{\checkmark} \star f) = \nu \star \mu_0(f), \tag{101}$$ useful01

because $\nu^{\checkmark} \star f$ is just another element of $\boldsymbol{C}_0(\mathbb{R}^d)$. □

syswst1 **Corollary 2.** *Let T be a BIBOS system. Given any sequence $(k_n)_{n\geq 1}$ in $\boldsymbol{C}_c(\mathbb{R}^d)$ (or even $\boldsymbol{L}^1(\mathbb{R}^d)$), which is bounded in $\boldsymbol{L}^1(\mathbb{R}^d)$, and such that $w^*\text{-}\lim_{n\to\infty} \mu_{k_n} = \delta_0$, then the measure μ representing T can be described as w^*-limit of the output sequence $T(k_n)$ (viewed now as a sequence in $\boldsymbol{M}_b(\mathbb{R}^d)$).*

Proof. We use that $T(k) = \mu \star k$ for $k \in \boldsymbol{C}_c(\mathbb{R}^d)$. The boundedness of (k_n) in $\boldsymbol{L}^1(\mathbb{R}^d)$ corresponds to the boundedness of the sequence (μ_{k_n}) in $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$. We also have seen that $T(k_n) = \mu \star \mu_{k_n}$ is w^*-convergent to $\mu \star \delta_0 = \mu$. So in fact we have that the sequence $T(k_n)$ (which consists of continuous and de facto absolutely Riemann integrable functions) is w^*-convergent to the impulse response $\mu = \mu \star \delta_0 = T(\delta_0)$. □

We will discuss such Dirac sequences at length later on. Just think of triangular functions on $\mathbb{R}$, compressed, and with constant area 1. We will use the operators $\mathrm{St}_\rho, \rho \to 0$ for this purpose.
We will show later on that general so-called *Dirac sequences*, which tend to δ_0, can be used.

As a consequence any BIBOs system of the form $T : f \mapsto \mu \star f$, with $\mu(f) = T(f^\checkmark)(0)$ in $\boldsymbol{M}_b(\mathbb{R}^d)$ with $\|\mu\|_{\boldsymbol{M}_b} = |||T|||_{\boldsymbol{C}_0(\mathbb{R}^d)}$ can also be considered as a bounded operator on $\boldsymbol{L}^1(\mathbb{R}^d)$ (or also $\boldsymbol{M}_b(\mathbb{R}^d)$), with the same norm:

$$|||C_\mu|||_{\boldsymbol{L}^1(\mathbb{R}^d)} = |||C_\mu|||_{\boldsymbol{M}_b(\mathbb{R}^d)} = \|\mu^\checkmark\|_{\boldsymbol{M}_b} = \|\mu\|_{\boldsymbol{M}_b}. \tag{102}$$

Liopnorm02

Since the flip-operator $\mu \to \mu^\checkmark$ is isometric and also satisfies

$$(\mu \star \nu)^\checkmark = \nu^\checkmark \star \mu^\checkmark, \tag{103}$$

flipconvMb

also the operator $k \mapsto \mu \star k$ can be viewed as the adjoint of $f \mapsto \mu^\checkmark \star f$ (on $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$) and is therefore a bounded and continuous operator on $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$, also commuting with any translation, and (thanks to the approximation by discrete convolutions) *with arbitrary convolutions!*. The operator, which acts on $\boldsymbol{L}^1(\mathbb{R}^d)$ is the adjoint of the mapping $C_{\mu^\checkmark}$.
In particular, one has $\mu = T(\delta_0)$, due to the simple relationship

$$T(\delta_0) = \mu \star \delta_0 = \mu. \tag{104}$$

impulsresp00

In the engineering literature the convergence is not discussed, i.e. the kind of limit that one takes by inserting a ***Dirac sequence*** **as input signals, and how the resulting output signals are convergent!** We will see that in the case of BIBO-systems one can guarantee that this results are convergent in the w^*-sense, resp. in the sense of strong operator convergence, see Proposition consistMbCG1 4 respectively Proposition [?] and the subsequent Lemma convnuwst 16, andin particular Corollary syswst1 2 below.

Another observation (checking for consistency).
We expect that translation can be extended from $\boldsymbol{C}_c(\mathbb{R}^d)$ not only to $\boldsymbol{C}_0(\mathbb{R}^d)$ and even $\boldsymbol{C}_b(\mathbb{R}^d)$ in the usual way, but also to $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$, due to the fact that we have

$$T_x f = \delta_x \star f$$

We also could try to approximate μ (in the w^*-sense) by a sequence (k_n) (more correctly (μ_{k_n})), by some bounded sequence in $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ (smoothed versions of μ!), and then expect that the shifted measures is approximated by the (equally) shifted nice function:

$$T_x \mu = w^*\text{-}\lim{}_{n\to\infty} \mu_{T_x k_n}.$$

In other words, the translation operator should be the unique w^*-continuous extension of the usual translation operator. This is possible, but it is more elegant to define translation by transposition, recalling that

$$C'_{\delta_{-x}}(\mu) = \delta_{-x}{}^{\checkmark} \star \mu = \delta_x \star \mu$$

suggesting to define directly (referring to (adjflip 100)):

$$T_x\mu(f) = (\delta_x \star \mu)(f) = \mu(\delta_{-x} f) = \mu(T_{-x} f),$$

This natural extension of operator defined for decent functions to *generalized functions* (here measures, later distributions) is called the **consistency principle**.

convwstcont2 **Lemma 17.** *The operator $T : \boldsymbol{C}_c(\mathbb{R}^d) \to \boldsymbol{M}_b(\mathbb{R}^d)$, given by $k \to \mu \star k$ resp. $\mu_k \to \mu \star \mu_k$ is continuous with respect to the norm in $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$, but it is also w^*-w^*--continuous, i.e. if*

$$\nu_n(f) \to \nu_0(f) \; for \; n \to \infty$$

for any $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ one also has

$$(\mu \star \nu_n)(f) \to (\mu \star \nu_0)(f).$$

Proof. We have for $n \to \infty$:

$$(\mu \star \nu_n)(f) = [(\mu \star \nu_n) \star f^{\checkmark}](0) = \mu \star (\nu_n \star f^{\checkmark})(0) \to \mu \star (\nu_0 \star f^{\checkmark})(0) = (\mu \star \nu_0)(f).$$

because obviously

$$\|\mu \star (\nu_n \star f^{\checkmark} - \nu_0 \star f^{\checkmark})\|_\infty \; \leq \|\mu\|_{\boldsymbol{M}_b} \|\nu_n \star f^{\checkmark} - \nu_0 \star f^{\checkmark}\|_\infty,$$

which tends to zero for $n \to \infty$. $\square$

8 Defining the Fourier Stieltjes transform

The usual Fourier transform is defined for $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ via

$$\hat{k}(s) = \int_{\mathbb{R}^d} k(t) \cdot e^{-2\pi i s \cdot t}\, dt = \int_{\mathbb{R}^d} \overline{\chi_s(t)}\, k(t) dt = \mu_k(\chi_{-s}), \tag{105}$$ FT-def2

with $\chi_s(t) = e^{2\pi i s \cdot t}$, the *pure frequency*. This suggest to define the Fourier for $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ by

$$\widehat{\mu}(s) = \mu(\chi_{-s}), \quad s \in \mathbb{R}^d.$$

In fact, the integral representation of the Fourier transform persists to be valid even $f \in \left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ (the Banach space of Lebesgue integrable functions on $\mathbb{R}^d$), but we prefer to avoid Lebesgue integrals. In fact, we plan to describe $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ in an alternative way (as a closed subspace, and introduce it as a closed ideal within $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$, with respect to convolution).
Similar considerations apply to the (so far "so-called"(!)) *inverse Fourier transform* which is of the form

$$\mathcal{F}^{-1}(h)(t) = \int_{\mathbb{R}^d} h(\omega) \cdot e^{2\pi i t \cdot \omega}\, d\omega = [\mathcal{F}(h)]^{\checkmark}(t) \tag{106}$$ IFT-def2

assuming that the integral exists. Of course we expect (but this has to be analytically derived) to have

$$\mathcal{F}^{-1}(\mathcal{F}(f)) = f \quad \text{and} \quad \mathcal{F}(\mathcal{F}^{-1}(h)) = h, \tag{107}$$ FourInvRel2

but due to the problem with integrability of $\widehat{f}$ (for general $f \in \boldsymbol{L}^1(\mathbb{R}^d)$) we have to clearly define the domain for which this Fourier inversion relation holds true.
Since $\widehat{f} \in \boldsymbol{C}_0(\mathbb{R}^d)$ for $f \in \boldsymbol{L}^1(\mathbb{R}^d)$ one may think of absolute (Riemann) integrability of $\widehat{f}$ as a sufficient condition for the validity of (107). We will discuss this in more detail later on.
In a sloppy form one can describe the **Convolution Theorem** as the statement that convolution is turned into pointwise multiplication under the Fourier transform respectively that pointwise multiplication is turned into convolution under the Fourier transform (or its inverse).
In order to be able to define the Fourier transform for measures and derive the *convolution theorem*, expressed in formulas as

$$\mathcal{F}(\mu_1 \star \mu_2) = \mathcal{F}(\mu_1) \cdot \mathcal{F}(\mu_2), \qquad \mu_1, \mu_2 \in \boldsymbol{M}_b(\mathbb{R}^d), \tag{108}$$ convFT00

or in the more usual form:

$$\widehat{\mu_1 \star \mu_2} = \widehat{\mu_1} \cdot \widehat{\mu_2}, \tag{109}$$ convFThat

convFTmeas **Theorem 4.** *The FT on $\boldsymbol{M}_b(\mathbb{R}^d)$ is injective, and turns convolution into pointwise multiplication, i.e. in fact, it is a homomorphism of Banach algebras. This also implies (once more) that convolution is commutative (because obviously pointwise multiplication is a commutative operation).*

The basis for the Fourier transform is the observation that the pure frequencies (resp. complex exponential functions $\chi_s(x) := exp(2\pi i s \cdot x)$ are (joint!) eigenvectors to general convolution operators. The strategy of proof is the following one:

1. First of all the *exponential law* guarantees that

$$(T_x\chi_s)(z) = \chi_s(z-x) = \chi_s(-x)\chi_s(z) = e^{-2\pi isx}\chi_s(z);$$

2. Then show a similar result for discrete measures;

3. Then show that the limit $D_\Psi\mu \to \mu$ is meaningful.

In fact, the Fourier (Stieltjes) transform will be defined on all of $\boldsymbol{M}_b(\mathbb{R}^d)$ and it will be shown that the convergence of $D_\Psi\mu$ to μ guarantees pointwise (in fact uniform convergence over compact sets of the frequency domain) to $\widehat{\mu}$.

First we observe that the Fourier transform of δ_x is a the pure frequency χ_x (which also appears as the eigenvalue of the shift operator T_x, which is the convolution operator by δ_x on $\boldsymbol{C}_b(\mathbb{R}^d)$, applied to the eigenvector χ_s).
Consequently (by forming absolutely convergent sums in $\left(\boldsymbol{C}_{ub}(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$) we find that also $\mathcal{F}(D_\Psi\mu) \in \boldsymbol{C}_{ub}(\mathbb{R}^d)$, for any BUPU Ψ.
We know that the discrete measures for a closed subspace of $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ hence we cannot expect to have norm convergence of $D_\Psi\mu$ to the (w^*)limit μ in $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$, which due the estimate $\|\widehat{\mu}\|_\infty \leq \|\mu\|_{\boldsymbol{M}_b}$ *would imply* that also $\widehat{\mu}$ (as the uniform limit, which it is not in fact!) would allow us to conclude easily that $\widehat{\mu} \in \boldsymbol{C}_{ub}(\mathbb{R}^d)$.
Note that we could however approximate μ first in the norm of $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ by a compactly supported measure $\nu = \mu \cdot p$, for some $p \in \boldsymbol{C}_c(\mathbb{R}^d)$ (e.g. $p = \sum_{i\in F}\psi_i$) and then apply the discretization operator to ν. Then the measures $\mathrm{D}_\Psi\nu$ would be finite discrete measures and their Fourier transforms would be linear combinations of pure frequencies, or so-called *trigonometric polynomials*. But still we would not have uniform convergence of these functions (i.e. norm convergence to $\hat{\nu}$ in $\left(\boldsymbol{C}_{ub}(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$.

8.1 Basic properties of the Fourier transform

basicFT **Theorem 5.** *(i) The Fourier (Stieltjes) transform maps the Banach convolution algebra* $(\boldsymbol{M}_b(\mathbb{R}^d), \star, \|\cdot\|_{\boldsymbol{M}_b})$ *non-expansively and injectively into the pointwise Banach algebra* $(\boldsymbol{C}_b(\mathbb{R}^d), \cdot, \|\cdot\|_\infty)$.
It restricts to a non-expansive, injective[10] *mapping from* $\left(\boldsymbol{L}^1(\mathbb{R}^d),\, \|\cdot\|_1\right) \to \left(\boldsymbol{C}_0(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$ *(with dense range). [Riemann-Lebesgue Lemma]*

There are additional (well-known) properties, such as $f^*(x) = \overline{f^\checkmark}$ is mapped into the conjugate of the Fourier transform, i.e. $\mathcal{F}(f^*) = \overline{\widehat{f}}$, and so on.
ADDED 27.10.2020

Proof. We are going to prove the injectivity of the F-ST transform $\mu \to \widehat{\mu}$, making use of (FRFST). Let us assume that $\widehat{\mu}(s) = 0$ for every $ats \in \mathbb{R}^d$. Then clearly based on (175) we have (by choosing $\nu = \mu_k$ for an arbitrary $k \in \boldsymbol{C}_c(\mathbb{R}^d)$). Let us concentrate on functions $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ with $\widehat{k} \in \boldsymbol{L}^1(\mathbb{R}^d)$, such as $k = \Delta$.

$$0 = \nu(\widehat{\mu}) = \mu(\widehat{\nu}) = \mu(\widehat{k}).$$

But since $\boldsymbol{C}_c(\mathbb{R}^d)$ is modulation and flip invariant we even have

$$0 = \mu(\widehat{M_u k^\checkmark}) = \mu(T_u(\widehat{k})^\checkmark) = \mu \star \widehat{k}$$

for any $k \in \boldsymbol{C}_c(\mathbb{R}^d)$. But $\boldsymbol{C}_c(\mathbb{R}^d)$ is even D_ρ invariant, and we can replace $\widehat{k}$ by $\mathrm{St}_\rho \widehat{k}$, which gives us for any $f \in \boldsymbol{C}_0(\mathbb{R}^d)$, using the (established) rules for convolution:

$$0 * f = (\mu \star \mathrm{St}_\rho \widehat{k}) * f = \mu * (\mathrm{St}_\rho \widehat{k} \star f).$$

choosing now k such that $\int_{\mathbb{R}^d} \widehat{k}(y) = 1$ (easily done by rescaling) we find that the limit of the argument to the right (for $\rho \to 0$) is just the given function f. Hence $\mu \star f = 0$ for any $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ and so finally $\mu = 0 \in \boldsymbol{M}_b(\mathbb{R}^d)$. [11] □

Given the **injectivity** of the Fourier (Stieltjes) transform it makes sense to look at the Fourier algebra $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$, defined as the range of $\boldsymbol{L}^1(\mathbb{R}^d)$ under the Fourier transform. Using Lebesgue's integration theory this would be simply

$$\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) := \{\widehat{f} \,|\, f \in \boldsymbol{L}^1(\mathbb{R}^d)\}$$

with the Fourier transform given as the usual Lebesgue integral. For us it is a subspace of all F-ST-transforms of specific measures. In any case, we know now that $\widehat{f} = 0$ implies that $f = 0$ (viewed as a measure, otherwise $f(x) = 0$ a.e.), and thus it makes sense to *endow* $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ with the norm

$$\|\widehat{f}\|_{\mathcal{F}\boldsymbol{L}^1} = \|f\|_{\boldsymbol{L}^1}, \quad f \in \boldsymbol{L}^1(\mathbb{R}^d). \tag{110}$$ FourAlgnorm

[10] Meaning: $\widehat{f} = 0$ implies that $f = 0$ in $\boldsymbol{L}^1(\mathbb{R}^d) \subset \boldsymbol{M}_b(\mathbb{R}^d)$, or classically $f(x) = 0$ almost everywhere.
[11] I GUESS THERE WILL BE EASIER ARGUMENTS, USING MORE ABSTRACT ARGUMENTS, BUT THIS IS STRICTLY BASED ON WHAT WE HAVE ALREADY ESTABLISHED. 27.10.2020, HGFEI

8.2 Outlook on the upcoming topics, as of 08.10.2020

We will discuss Dirac sequences and more generally bounded approximate units in Banach algebras a bit more closer. They play an important role in Banach algebras which do not have a unit element (with respect to the "multiplication"), such as $\left(\boldsymbol{C}_0(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$ and $\left(\boldsymbol{L}^1(\mathbb{R}^d),\, \|\cdot\|_1\right)$.
In order to efficiently discuss such objects we will introduce the concept of *Banach modules* over Banach algebras.
Statements like

$$\boldsymbol{L}^1(\mathbb{R}^d) \star \boldsymbol{L}^2(\mathbb{R}^d) \subset \boldsymbol{L}^2(\mathbb{R}^d)$$

combined with the corresponding norm-estimates

$$\|g \star f\|_2 \leq \|g\|_1 \|f\|_2, \quad g \in \boldsymbol{L}^1(\mathbb{R}^d), f \in \boldsymbol{L}^2(\mathbb{R}^d)$$

will be important special cases of considerations in this context.
Finally we are going to derive the *Plancherel Theorem*, showing that the Fourier transform $\mathcal{F}\colon \boldsymbol{L}^1 \cap \boldsymbol{L}^2 \to \boldsymbol{L}^2$ can be extended to a unitary automorphism ("energy preserving" $\left(\boldsymbol{L}^2(\mathbb{R}^d),\, \|\cdot\|_2\right) \to \left(\boldsymbol{L}^2(\mathbb{R}^d),\, \|\cdot\|_2\right)$) mapping.
Another direction will be to avoid the problem that there is no inclusion relation between $\boldsymbol{C}_0(\mathbb{R}^d)$ and its dual space $\boldsymbol{M}_b(\mathbb{R}^d)$ (just $\{\mu_k, k \in \boldsymbol{C}_c(\mathbb{R}^d)\}$ is a kind of common dense subspace of both of them, by replacing $\left(\boldsymbol{C}_0(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$ by the smaller *Wiener algebra* $\left(\boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d),\, \|\cdot\|_{\boldsymbol{W}}\right)$ and its dual, the space of translation-bounded measures $\left(\boldsymbol{W}(\boldsymbol{M}, \boldsymbol{\ell}^\infty)(\mathbb{R}^d),\, \|\cdot\|_{\boldsymbol{W}}\right)$.
This dual space contains also Dirac-combs $⧢_\alpha := \sum_{k \in \mathbb{Z}^d} \delta_{\alpha k}$ and has the additional advantage that there is a natural embedding of the space $\boldsymbol{W}(\mathbb{R}^d)$ into its dual (in the usual way):

$$\left(\boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d), \|\cdot|\boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d)\|\right) \hookrightarrow \left(\boldsymbol{W}(\boldsymbol{M}, \boldsymbol{\ell}^\infty)(\mathbb{R}^d), \|\cdot|\boldsymbol{W}(\boldsymbol{M}, \boldsymbol{\ell}^\infty)(\mathbb{R}^d)\|\right).$$

9 GEOGEBRA and MATLAB

```
M = [1,1;1,-2];  % a symmetric 2x2-matrix
[V,D] = eig(M)
V =
   -0.2898   -0.9571
    0.9571   -0.2898
D =
   -2.3028         0
         0    1.3028

edt(M) = -3
```

9.1 Circulant matrices

```
C = gallery('circul', 5)
C =
     1     2     3     4     5
     5     1     2     3     4
     4     5     1     2     3
     3     4     5     1     2
     2     3     4     5     1

>> convect(rand(1,5))  %self-made routine!
ans =
    0.8507    0.5606    0.9296    0.6967    0.5828
    0.5828    0.8507    0.5606    0.9296    0.6967
    0.6967    0.5828    0.8507    0.5606    0.9296
    0.9296    0.6967    0.5828    0.8507    0.5606
    0.5606    0.9296    0.6967    0.5828    0.8507


>> A = convmat(rand(1,m));
>> A = A+A';   % random real symmetric matrix!
>> [VA, DA] = eig(A);
```

9.2 Intermediate considerations

Let us recall that we have started to introduce $\left(\boldsymbol{C}_0(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$ and $\left(\boldsymbol{C}_b(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$ as pointwise Banach algebras, with respect to the sup-norm.
Then we have started to look at the dual space, which we call $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$, the space of bounded measures on $\mathbb{R}^d$. With the help of BUPUs one can show that $\|\mu\|_{\boldsymbol{M}_b} = \sum_{i\in I} \|\mu\, \psi_i\|_{\boldsymbol{M}_b}$, for any $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$. This implies that the compactly supported measures are norm-dense in $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$, namely those of the form $\sum_{i\in F} \mu\, \psi_i$, for some finite set $F \subset I$. In the same way any $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ can be approximated by compactly supported functions of the form $\sum_{i\in E} \psi_i f \in \boldsymbol{C}_c(\mathbb{R}^d)$, where E is some finite subset of I.
Among these bounded measures there are the discrete ones (bounded discrete measures are just absolutely convergent series of Dirac measures) and the "nice" ones (official terminology: absolutely continuous measures), arising from densities $g \in \boldsymbol{L}^1(\mathbb{R}^d) \supset \boldsymbol{C}_c(\mathbb{R}^d)$, in the form

$$f \mapsto \int_{\mathbb{R}^d} f(x)\, g(x) dx, \quad f \in \boldsymbol{C}_0(\mathbb{R}^d).$$

The identification of the TILSs with the convolution operators $f \mapsto \mu \star f$ gives an additional multiplicative structure within $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ (with the associativity coming for free), and the closed ideal $\left(\boldsymbol{L}^1(\mathbb{R}^d),\, \|\cdot\|_1\right)$ inside of $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$.
The discrete measures can be used to demonstrate that $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ is a *commutative* Banach algebra with respect to convolution, because

$$\delta_x \star \delta_y = \delta_{x+y} = \delta_y \star \delta_x.$$

This is *equivalent* to the composition law $T_x \circ T_y = T_{x+y}$ via the all important *identification of measures with impulse responses of a BIBOS system*.
That the operator $k \mapsto \mu \star k$ can be extended to the bounded measures follows from the w^*-continuity of the mapping $\nu \star k_n$ to $\nu \star \mu$ if $k_n(f) \to \mu(f)$, for some sequence $(k_n)_{n\geq 1}$ in $\boldsymbol{M}_b(\mathbb{R}^d)$.
The usual approach to Fourier Analysis (see most mathematical books, such as **rest00 :** [79], or Rudin, Hewlett-Ross, and many others) start with a definition of the Fourier transform and convolution via the integral transform, all defined on the Lebesgue space $\left(\boldsymbol{L}^1(\mathbb{R}^d),\, \|\cdot\|_1\right)$. This gives the impression that the Fourier transform IS an integral transform.
The truth is, that it can be *realized* as an integral transform, with many interesting properties, defined for the commutative Banach algebra $\left(\boldsymbol{L}^1(\mathbb{R}^d),\, \|\cdot\|_1\right)$ with bounded approximate units (the Dirac sequences), mapping the more involved convolution into pointwise multiplication. Hence, among others, it is very important to e.g. partially invert a TILS (at least for band-limited functions, as we will see).
Given such a starting point one can ask, whether e.g. a pure frequency χ_s "has a Fourier transform" (expecting a Dirac delta δ_s as answer). But for answering this question one has to resort to the (much more complicated) theory of distributions (in the sense of Laurent Schwartz). This is a great theory, especially for PDE (partial differential equations), but - **as we are going to demonstrate in this course** is by no means

necessary. Using ideas from the theory of function spaces (e.g. Wiener amalgams) and time-frequency analysis (e.g. Gabor analysis, spectrograms) a conceptually much more simple approach allows to deal with the "generalized Fourier transform". It will lead us to the consideration of the so-called *Banach Gelfand triple* $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$ as a natural domain for the Fourier transform. It will be compared to the triple $\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$ as a natural domain for the computation with "numbers".
The generalized Fourier transform can be viewed as the *natural extension* of the ordinary Fourier (and inverse Fourier) transform pair, given as integral transforms on $\boldsymbol{L}^1(\mathbb{R}^d) \cap \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$, e.g. by considering w^*-limits (such as Dirac sequences). As we will see, the generalized formula will provide us with a tool that allows to claim not only $\mathcal{F}(\delta_0) = \mathbf{1}$ but also the converse, namely $\mathcal{F}^{-1}(\mathbf{1}) = \delta_0$, often written as

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

So will concentrate on the properties of that integral transform (for us the restriction of the Fourier Stieltjes transform) to the subset $\boldsymbol{L}^1(\mathbb{R}^d) \subset \boldsymbol{M}_b(\mathbb{R}^d)$, with $\widehat{\mu}(s) = \mu(\chi_{-s}), s \in \mathbb{R}^d$.
Nevertheless it is important and useful to derive the properties of the (integral) Fourier transform in the $\boldsymbol{L}^1$-setting, especially the (pointwise) Fourier Inversion formula and the all important *Plancherel Theorem* (energy conservation law for the Fourier transform).

9.3 Locality of convolution

An important observation is now, that

$$\sum_{i \in F} \mu\, \psi_i \star \sum_{j \in E} f \psi_j$$

is not only in $\boldsymbol{C}_{ub}(\mathbb{R}^d)$, but it also has compact support.
In order to verify this let us recall that each of the functions $\psi_i, i \in I$ has compact support, hence for the finite set $F \subset I$ one can find R_1 such that $\psi_i(z) = 0$ for any $z \in \mathbb{R}^d$ with $|z| \geq R_1$.
In a similar way we find that for a finite set $E \subset I$ one finds a radius $R_2 > 0$ such that $\psi_j(u) = 0$ for $|u| > R_2$.
Now let us look (for fixed, arbitrary) indices $i \in F$ and $j \in E$ at the convolution product

$$[(\mu\psi_i) \star (f\psi_j)](y) = \mu\psi_i(T_y[f\psi_j]^{\checkmark}) = \mu([\psi_i \cdot T_y(\psi_j{}^{\checkmark})] \cdot T_y f^{\checkmark}). \qquad (111)$$ convlocal03

This convolution product has certainly only a chance to be non-zero if there exists some $z \in \mathbb{R}^d$ such that

$$[\psi_i \cdot T_y(\psi_j{}^{\checkmark})](z) \neq 0.$$

But this means that $\psi_i(z) \neq 0$, or $|z| \leq R_1$ and $\psi_j(y - z) \neq 0$, so $|y - z| \leq R_2$. But this gives

$$|y| \leq |z| + |y - z| \leq R_1 + R_2,$$

or expressed differently, the given convolution product (for any $i \in F$, $j \in E$) will vanish outside of the compact ball of radius $R_1 + R_2$ around zero.

Since $f \mapsto \mu \star f$ preserves uniform continuity it is clear that under the given circumstances one has

$$\mu \star f \in \boldsymbol{C}_c(\mathbb{R}^d) \subset \boldsymbol{L}^1(\mathbb{R}^d).$$

By taking limits on both objects ($F \nearrow I, E \nearrow I$) we find the $\mu \star f$ is in $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$. Recall that we define $\boldsymbol{L}^1(\mathbb{R}^d)$ as the closure of $\boldsymbol{C}_c(\mathbb{R}^d)$ (identified with a subspace of $\boldsymbol{M}_b(\mathbb{R}^d)$, via $k \longleftrightarrow \mu_k$) in $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$.
Among others we will make use of the following results:

DpsiconvLi **Proposition 6.** *For any $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ and $f \in \boldsymbol{L}^1(\mathbb{R}^d)$ one has*

$$\lim_{|\Psi| \to 0} \|D_\Psi \mu \star f - \mu \star f\|_{\boldsymbol{L}^1(\mathbb{R}^d)} = 0, \quad \forall f \in \boldsymbol{L}^1(\mathbb{R}^d). \tag{112}$$ DpsiconvL1A

Proof. OUTLINE of the argument: The statement is true for compactly supported μ and $f \in \boldsymbol{C}_c(\mathbb{R}^d)$. The important point is then that for any $\Psi = (\psi_i)_{i \in I}$ with $|\Psi| \le 1$ the same set $F \subset I$ can be used, hence one can conclude from uniform convergence

$$\lim_{|\Psi| \to 0} \|D_\Psi \mu \star k - \mu \star k\|_\infty = 0, \quad \forall k \in \boldsymbol{C}_c(\mathbb{R}^d) \tag{113}$$ DpsiconvLinfA

that one also has $\boldsymbol{L}^1(\mathbb{R}^d)$-convergence.
The general case can be obtained therefrom by approximation. $\square$

10 Fourier transforms and involutions

Recall the involutions $f^{\checkmark}(x) = f(-x)$, and $f^* = \overline{f^{\checkmark}}$

$$\widehat{f^{\checkmark}} = \widehat{f}^{\checkmark}, \quad \widehat{f^*} = \overline{\widehat{f}}.$$

$$[\mathcal{F}(f^{\checkmark})](s) = \int_{\mathbb{R}^d} f^{\checkmark}(x)\chi_{-s}(x)dx = \int_{\mathbb{R}^d} f(-x)\chi_s(-x)dx =$$

after applying the substitution $z = -x$ and using that $exp(2\pi isz) = exp(2\pi i(-s)(-z))$:

$$\int_{\mathbb{R}^d} f(z)\chi_s(z)dz = \int_{\mathbb{R}^d} f(z)\chi_{-s}(-z)dz = \widehat{f}(-z) = (\widehat{f})^{\checkmark}(z)$$

In a similar way one shows that $\mathcal{F}(f^*) = \overline{\widehat{f}}$ or $\widehat{f^*} = \overline{\widehat{f}}$. Indeed

$$\widehat{f^*}(s) = \int_{\mathbb{R}^d} f^*(x)\chi_{-s}(x)dx = \int_{\mathbb{R}^d} \overline{f(-x)} \cdot e^{-2\pi isx}dx =_{-x=z} \int_{\mathbb{R}^d} \overline{f(z)} \cdot \overline{e^{-2\pi isz}}dz = \overline{\widehat{f}(s)}.$$

In a similar way it is shown that

$$\mathcal{F}(\overline{f}) = (\widehat{f})^*,$$

e.g. by using above facts and the fact that $\overline{f} = (f^*)^{\checkmark}$.

We start from the **Fundamental Relation for the Fourier Transform** (on $\boldsymbol{L}^1(\mathbb{R}^d)$):

$$\int_{\mathbb{R}^d} f(t)\hat{g}(t)dt = \int_{\mathbb{R}^d} \widehat{f}(s)g(s)ds, \quad \text{for all} \quad f, g \in \boldsymbol{L}^1(\mathbb{R}^d). \tag{114}$$ L1FourL1B

This is easy to verify, because it just expresses the double integral

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(t)g(s)e^{-2\pi it\cdot s}dtds$$

can be realized in two different ways (but the value is the same, according to Fubini's Theorem), by changing the order of integration.

11 Even and odd functions or measures

evenodddef1 **Definition 5.** A function is called *even* if $f^{\checkmark} = f$. It is called *odd* if $f^{\checkmark} = -f$. Similar definitions apply to measures ($\mu^{\checkmark}(f) = \mu(f^{\checkmark})$).

Remark 5. The prototypical examples of even functions are of the form $\cos(2\pi sx)$ (the real part of the pure frequency), while $\sin(2\pi sx)$ (the imaginary parts) are odd functions.

multevenodd **Lemma 18.**

- *The pointwise product of two even or odd functions is even, while one obtains an odd function for the mixed case.*
- *Very much the same is the case for convolutions.*
- *In particular, one can state: The convolution product of two even or two odd functions is again an even functions.*

The proof if these facts is left as an exercise. Also note that for even function f one has

$$[\mu \star f](y) = \mu(T_y f), \quad y \in \mathbb{R}^d, \mu \in \boldsymbol{M}_b(\mathbb{R}^d).$$

This also applies to the pointwise convolution product:
The value of $k_1 \star k_2$ at y is the integral over $k_1 \cdot T_y k_2$, if k_2 is even.

12 BOX FOURIER equals the SINC function

Strictly speaking the *rectangular functions* or *boxcar* functions $\mathbf{1}_{[a,b]}$ are elements of $\boldsymbol{L}^1(\mathbb{R}^d)$, as **any Riemann integrable function with compact support**! Thus we can compute the Fourier transform using the Riemann integrals.
Since

$$\int_0^{1/2} \cos(2\pi s x) dx = \frac{1}{2\pi s} \sin(2\pi s x)\Big|_0^{1/2}$$

and hence

$$\mathcal{F}(\mathbf{1}_{[-1/2,1/2]})(s) = \mathbf{2} \int_0^{1/2} cos(2\pi s x) dx = 2\frac{1}{2\pi s} \sin(2\pi s 1/2) = \frac{\sin(\pi s)}{\pi s} =: \mathrm{SINC}(s).$$

boxSINC **Lemma 19.** *The Fourier transform of any integrable step function belongs to* $\boldsymbol{C}_0(\mathbb{R}^d)$.

Proof. Any integrable step function is the $\boldsymbol{L}^1$-limit of finite linear combinations of dilated and shifted rectangular functions, hence the Fourier transform is the uniform limit of their Fourier transforms.
Since the dilation of a boxcar function results in a simple dilation of the Fourier transform due to the formula $\mathcal{F}(\mathrm{St}_\rho(f)) = \mathrm{D}_\rho(\widehat{f})$ resp. $\mathcal{F}(\mathrm{D}_\rho f) = \mathrm{St}_\rho(\widehat{f}), \rho > 0$, and

$$|\mathcal{F}(T_x f)| = |M_{-s}\widehat{f}| = |\widehat{f}|$$

we see clearly that the Fourier transform of

$$\mathcal{F}\left(\sum_{k=1}^n c_k T_{x_k} \mathrm{D}_{\rho_k} \mathbf{1}_{[-1/2,1/2]}\right)$$

is exactly

$$\sum_{k=1}^n c_k M_{-x_k} \mathrm{St}_\rho SINC \in \boldsymbol{C}_0(\mathbb{R}^d).$$

Taking limits in $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$, hence uniform limits on the Fourier transform side, we arrive at the claim that $\widehat{f} \in \boldsymbol{C}_0(\mathbb{R}^d)$, known as the *Riemann-Lebesgue Lemma.* □

RiemLebsThm **Theorem 6.** *The Fourier transform, restricted to the Banach algebra* $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ *(with pointwise almost everywhere defined convolution) maps the Banach convolution algebra with involution* $f \mapsto f^*$ *into the pointwise Banach algebra* $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$ *with involution* $f \mapsto \overline{f}$. *The mapping is a continuous, injective algebra homomorphism and*

$$\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) := \mathcal{F}(\boldsymbol{L}^1(\mathbb{R}^d)) := \{\widehat{f} \mid f \in \boldsymbol{L}^1(\mathbb{R}^d)\}$$

is a dense subalgebra of $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$.
For $f \in (\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1)(\mathbb{R}^d)$ *one can recover (the continuous version of)* f *by the Inverse Fourier transform:*

$$f(t) = \int_{\mathbb{R}^d} \widehat{f}(s) e^{2\pi s \cdot t} ds, \quad t \in \mathbb{R}^d.$$

Remark 6. L1FourL1B Let us recall the fundamental relationship for the Fourier transform on $\boldsymbol{L}^1(\mathbb{R}^d)$ (114). It can be reinterpreted as

$$\mu_f(\widehat{g}) = \mu_g(\widehat{f}), \quad f, g \in \boldsymbol{L}^1(\mathbb{R}^d). \tag{115}$$ FundFourL1

So assuming that $\widehat{f}(s) = 0$ for any $s \in \mathbb{R}$ (i.e. that $\widehat{f} = 0$ (the zero function), then one has $\mu_f(\hat{g})$ for any $g \in \boldsymbol{L}^1(\mathbb{R}^d)$, or $\mu_f(h)$ for any $h \in \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$. But $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ is dense in $\boldsymbol{C}_0(\mathbb{R}^d)$, and thus for a sequence $(h_n)_{n \geq 1}$ which is uniformly convergent to a given function $h \in \boldsymbol{C}_0(\mathbb{R}^d)$ one has

$$\mu_f(h) = \lim_{n \to \infty} \mu_f(h_n) = 0,$$

or with other words: $\mu_f = 0$, but this implies that $f = 0$ in $\boldsymbol{L}^1(\mathbb{R}^d)$ (hence pointwise for $f \in \boldsymbol{C}_c(\mathbb{R}^d)$ with $\widehat{f} \in \boldsymbol{L}^1(\mathbb{R}^d)$).

Proof. Since we have $\|\widehat{f}\|_\infty \leq \|f\|_{\boldsymbol{L}^1}$ for any $f \in \boldsymbol{L}^1(\mathbb{R}^d)$ it is enough (due to the completeness of $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$) to show that $\hat{k} \in \boldsymbol{C}_0(\mathbb{R}^d)$ for any $k \in \boldsymbol{C}(\mathbb{R}^d)$.
So the first step is to recall the concept of Riemann integrability of $f \in \boldsymbol{C}_c(\mathbb{R}^d)$, which means that the (in fact uniformly) continuous step function is approximated (from above and from below) by step functions whose difference (in the $\boldsymbol{L}^1$-sense) shrinks, hence the difference between $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ and a suitable step function (with finitely values over cubes, in $\mathbb{R}^d$) can be made as small as we want. But for those functions we have explicit expressions for their Fourier transforms, namely suitable (modulated and dilated) SINC-functions.
The density of $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ in $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$ can be argued as follows: Let us first deal with the case $d = 1$: Any piecewise linear function with compact support, i.e. any linear combination of triangular functions, belongs to $\mathcal{F}\boldsymbol{L}^1(\mathbb{R})$, since $\Delta = \mathcal{F}(\mathrm{SINC}^2)$ (for even functions with $f \in \boldsymbol{L}^1(\mathbb{R})$ and $\widehat{f} \in \boldsymbol{L}^1(\mathbb{R}^d)$ one has $\mathcal{F}(\mathcal{F}(f)) = f^{\checkmark} = f$!). But any $h \in \boldsymbol{C}_0(\mathbb{R}^d)$ can be easily approximated by a piecewise linear function with compact support, i.e. by a finite linear combination of shifted triangular functions, which belongs to $\mathcal{F}\boldsymbol{L}^1(\mathbb{R})$.
For the multidimensional case one can take tensor products of such functions (i.e. coordinatewise products of such functions). For example

$$\mathrm{SINC}(x, y) = \mathrm{SINC}(x)\,\mathrm{SINC}(y), \quad x, y \in Rst$$

belongs obviously to $\boldsymbol{C}_0(\mathbb{R}^2)$, and so on.
The fact that the inclusion is a proper one case be obtained from the fact that it impossible (due to examples) to estimate the sup-norm of $\widehat{f}$ from below for functions $f \in \boldsymbol{L}^1(\mathbb{R}^d)$ with $\|f\|_{\boldsymbol{L}^1} = 1$. In other words, one can exhibit examples (see SCRIPT) of functions with $\boldsymbol{L}^1$-norm equal to one but with $\|f\|_\infty$ arbitrary small.

Such functions cannot be non-negative, because one has for $g \geq 0$:

$$\|g\|_{\boldsymbol{L}^1} = \int_{\mathbb{R}^d} g(x)dx = \hat{g}(0) = |\hat{g}(0)| \leq \|\hat{g}\|_\infty \leq \|g\|_{\boldsymbol{L}^1}.$$

The injectivity of $f \mapsto \widehat{f}$ has to be pointed out separately (in the context of the Fourier inversion theorem). $\square$

FTtriang **Lemma 20.** *The triangular function* $\Delta(x) = \max(\min(1-x, 1+x))$ *is the convolution product* $\mathbf{1}_{[-1/2,1/2]} \star \mathbf{1}_{[-1/2,1/2]}$. *Its Fourier transform is thus the non-negative function* $\widehat{\Delta} = \mathrm{SINC}^2$, *which integrable and vanishes exactly on* $\mathbb{Z} \setminus \{0\}$.

Proof. Since the boxcar function $\mathbf{1}_{[-1/2,1/2]}$ is even the resulting convolution square is even as well. It is supported on $[-1,1]$ and hence it is enough to compute the value over $[0,1]$, which is easy to do, since

$$\mathbf{1}_{[-1/2,1/2]} \star_{pw} \mathbf{1}_{[-1/2,1/2]}(y) = \int_y^1 1dx = 1-y.$$

$\square$

GauDistrFct1.jpg

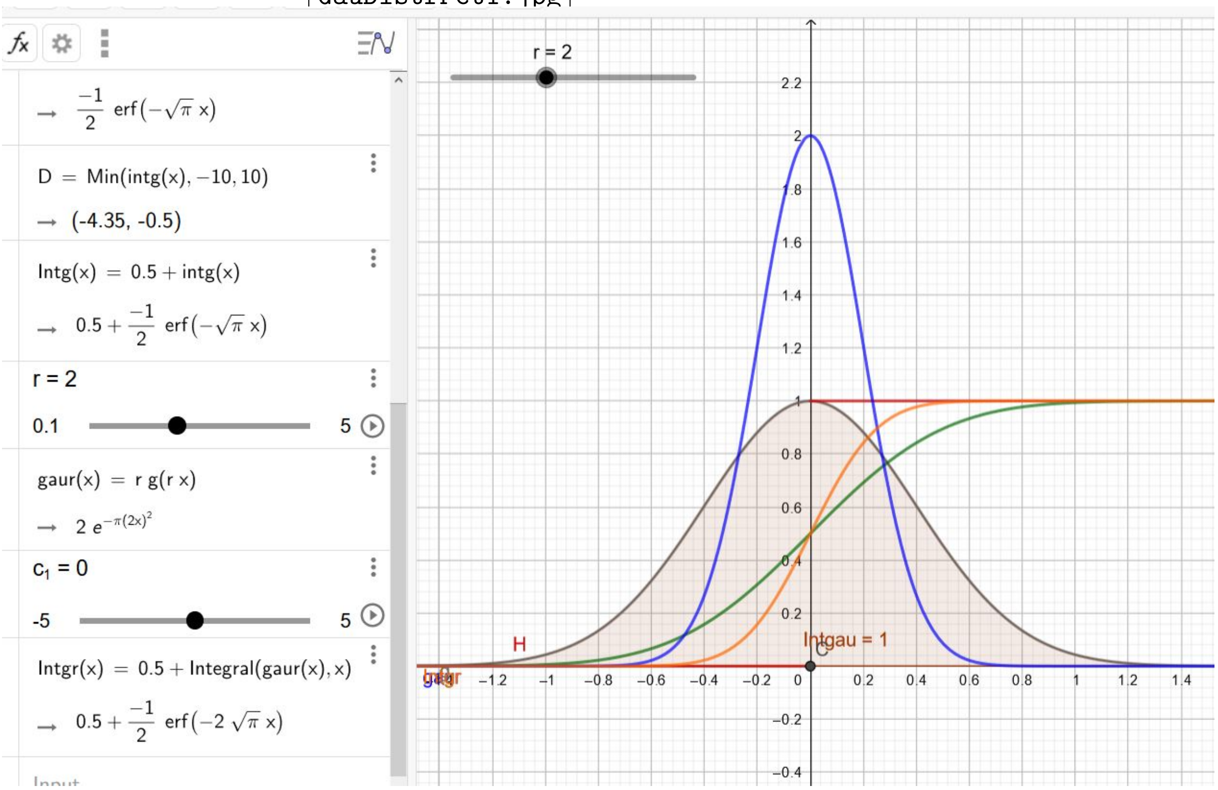

Abbildung 13: GauDistrFct1.jpg:
This is a plot showing Gauss functions with different compression factors, and their indefinite integrals, the so-called "DISTRIBUTION FUNCTION". As one can see, for any Dirac sequence (which is a sequence of functions all with integral 1, and shrinking support size, say) the corresponding integrated version tends the the *Heavyside function, or unit step function $H(x) = 1$ for $x \geq 0$ and zeros elsewhere. The convergence is valid for any point, except for the location of the step, i.e. for $x = 0$. Thus this concept of convergence is* **weaker** *than the [at that time standard] concept of pointwise convergence for any $x \in \mathbb{R}$. This is way the wording "weak" comes into play. Finally it was observed that one should talk about various concepts of weak convergence for measures. Considering the bounded linear functionals as elements of the dual space of $\boldsymbol{C}_0(\mathbb{R}^d)$ the introduced w^*-convergence concept is the most important and natural one.*

Recall that for $d = 1$ the name Fourier Stieltjes transforms arose in the context of Riemann-Stieltjes integrals, taken with respect to a distribution function $F = \int_{-\infty}^{x} 1\, d\mu(y)$, with Riemannian (Stieltjes) sums of the form

$$\sum_{k=1}^{r} f(\xi_k)(F(x_{k+1}) - F(x_k)) \tag{116}$$ RiemStisum

of a sequence $a = x_0 < x_1 \dots x_r = b$ and with $max_{k=1,..r}\, |x_{k+1} - x_k| \to 0$.
Those distribution functions are of *bounded variation* for any $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ and are *absolutely continuous* if and only if $\mu = \mu_g$ for some $g \in \boldsymbol{L}^1(\mathbb{R}^d)$. Such functions have the property that $F'(x) = g(x)$ a.e., i.e. F can be recovered from F'.

12.1 Illustration of Convergence in Distribution

https://de.wikipedia.org/wiki/Konvergenz_in_Verteilung

A sequence of distribution functions $(F_n)_{n\geq 1}$ is *weakly convergent* to a distribution function F if

$$\lim_{n\to\infty} F_n(t) = F(t), \quad \text{for all continuity points of } F, \tag{117}$$ convdistrib

OR

$$\lim_{n\to\infty} E(f(X_n)) = E(f(X)) \quad \text{for all } f \in \boldsymbol{C}_b(\mathbb{R}).$$

We can demonstrate the w^*-convergence of discrete measures by looking at the corresponding *distribution functions* $F_\mu(t) = \mu(\mathbf{1}_{[c,t]}$ (for some small, fixed value c).

Psimu01b.pdf

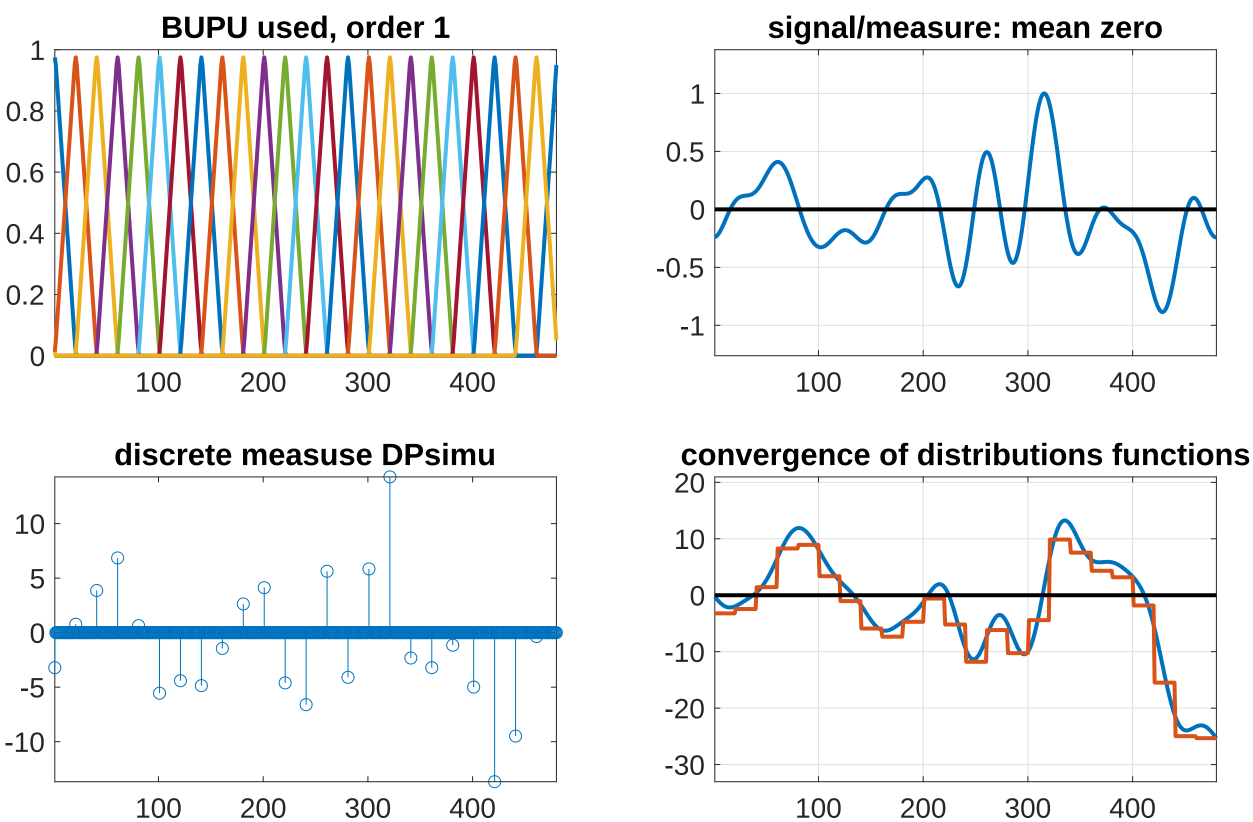

Abbildung 14: convDPsimu01b.pdf:
The (triangular) BUPU, the smooth signal/measure, the discretized measure illustrate by the STEM-command, and the corresponding distribution function, which has a jump of sice c_k at the position of the corresponding (positive or negative) Dirac measure.

dztofive.pdf

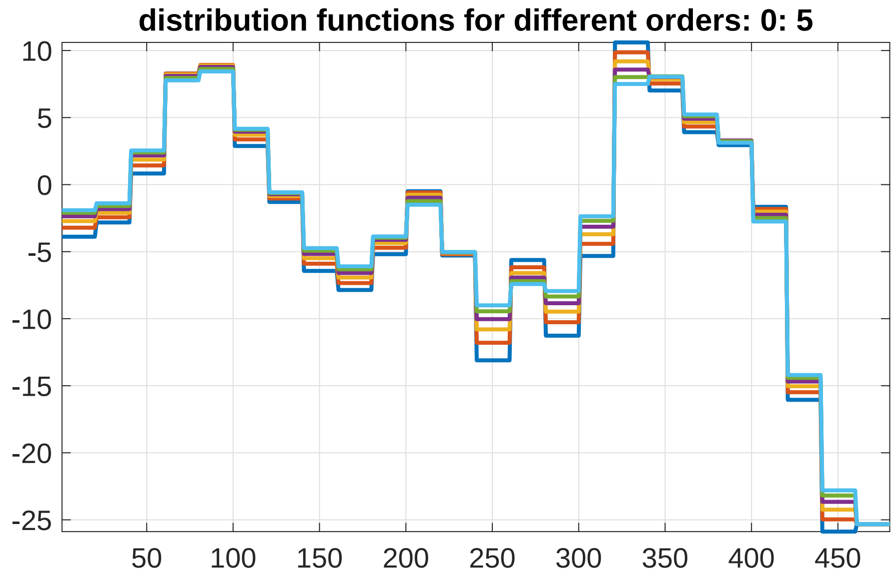

Abbildung 15: DPsimuordztofive.pdf:
Distribution function obtained by B-spline BUPUs of different order: the higher the order, the more smearing effect is obtained, because the supports of the B-splines are increasing with the order.

13 Convolution Demo

It is clear that the boxcar-function (as we use it) is an even function. The same is true for dilated versions (Exercise: the dilation operators both preserve the property of a function to be even or odd), and thus also the convolution square or the convolution product of two such symmetric indicator functions will be symmetric. Since the support of $\mathbf{1}_{[-1/2,1/2]}$ is exactly $[-1/2, 1/2]$ it is clear that the convolution square will vanish for $|x| > 1$. Since it is even one only has to determine the values for $0 < x < 1$, which is obviously (visualize the product of one box function against a shifted box function) is just $1 - x$ (!).

vargap1c.pdf

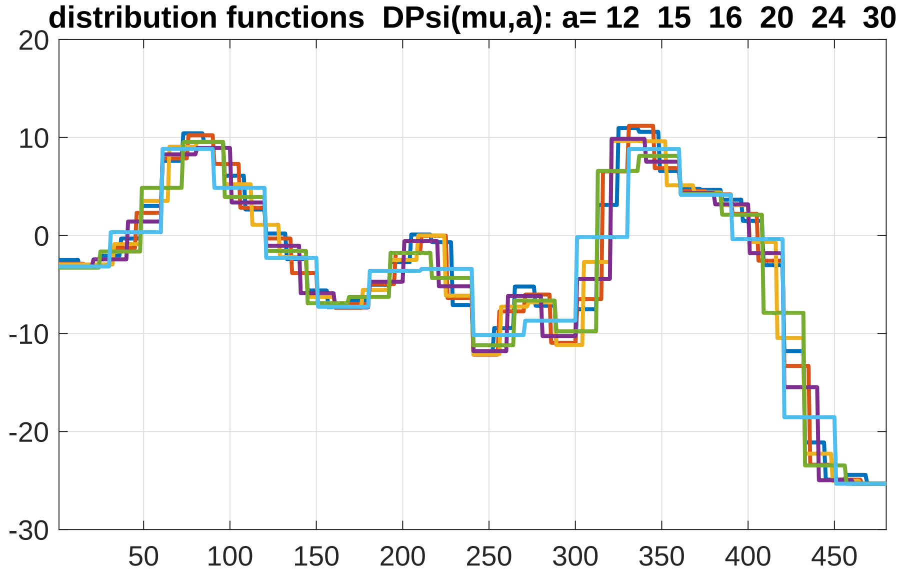

Abbildung 16: DPsimuvargap1c.pdf:
Distribution function of discrete measures, for finer and finer partitions of unity, so that the approximation of the resulting step-functions are better and better!

14 Approximate units for convolution

vdemo01a.pdf

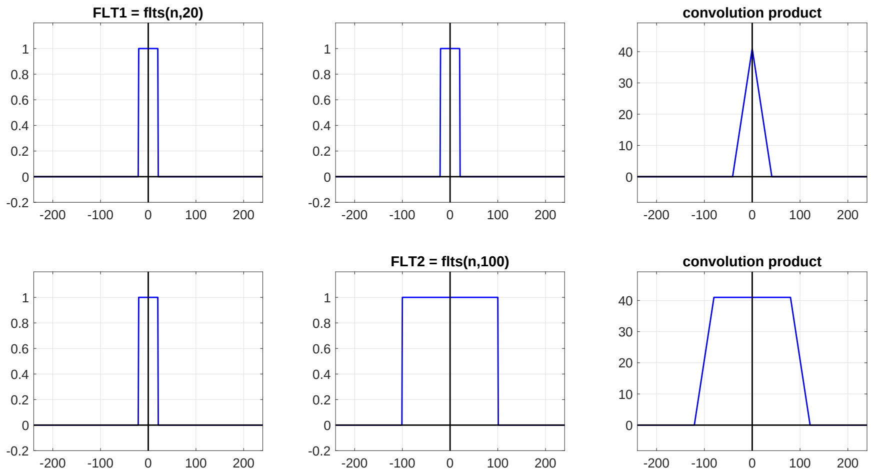

Abbildung 17: convdemo01a.pdf

vdemo02a.pdf

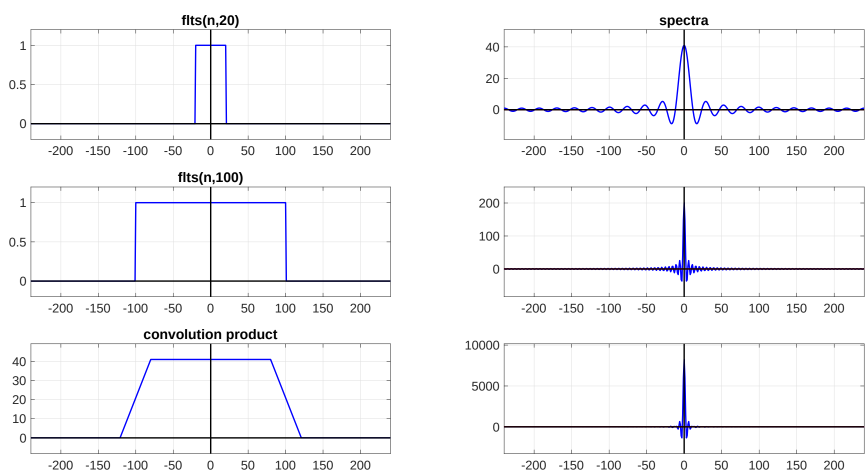

Abbildung 18: convdemo02a.pdf

vpower01.pdf

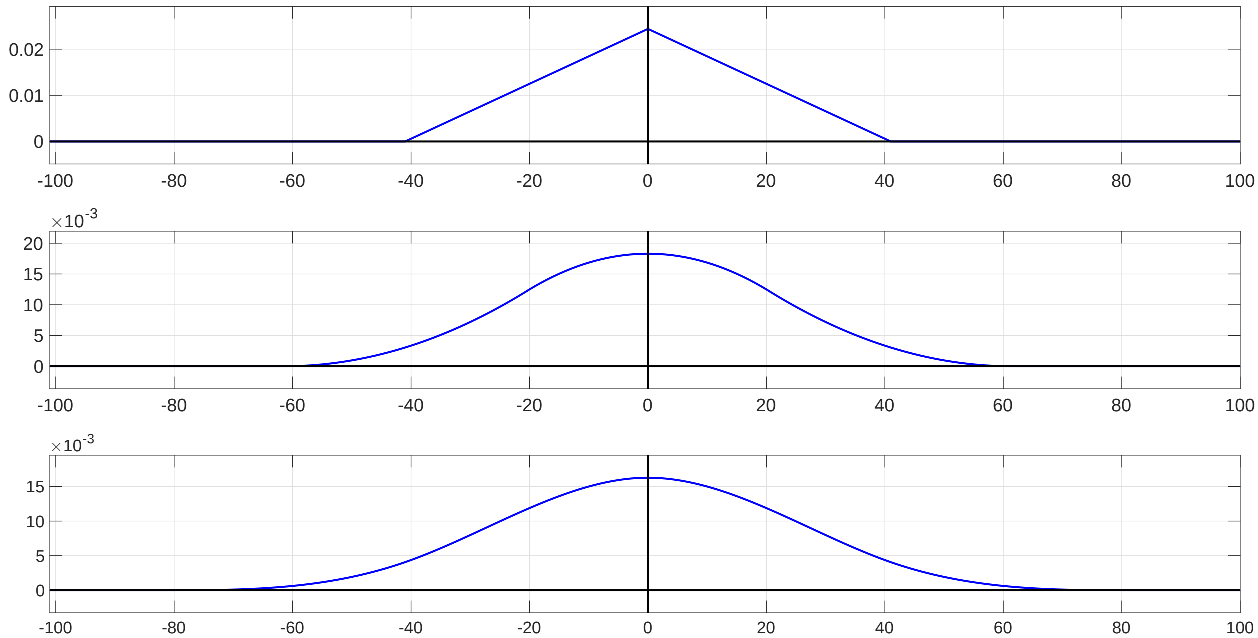

Abbildung 19: convpower01.pdf
Convolution powers of the boxcar function: 2,3,4 factors

umhist01.jpg

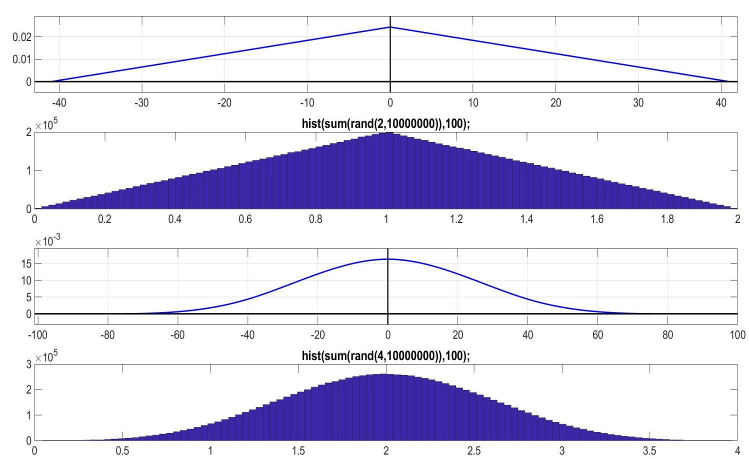

Abbildung 20: sumhist01.jpg:behaviour of histograms compared to B-splines'

dpwlin12.eps

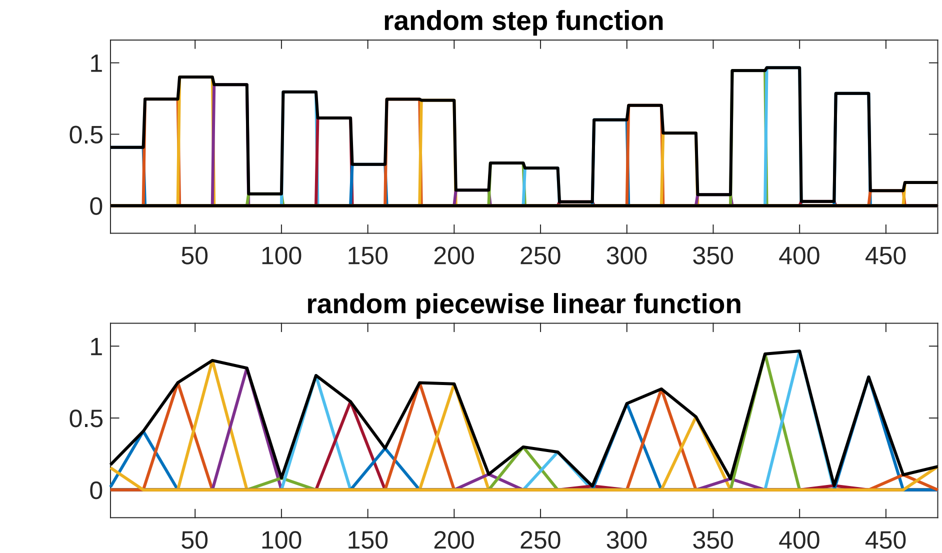

Abbildung 21: randpwlin12.eps

lin12fft.eps

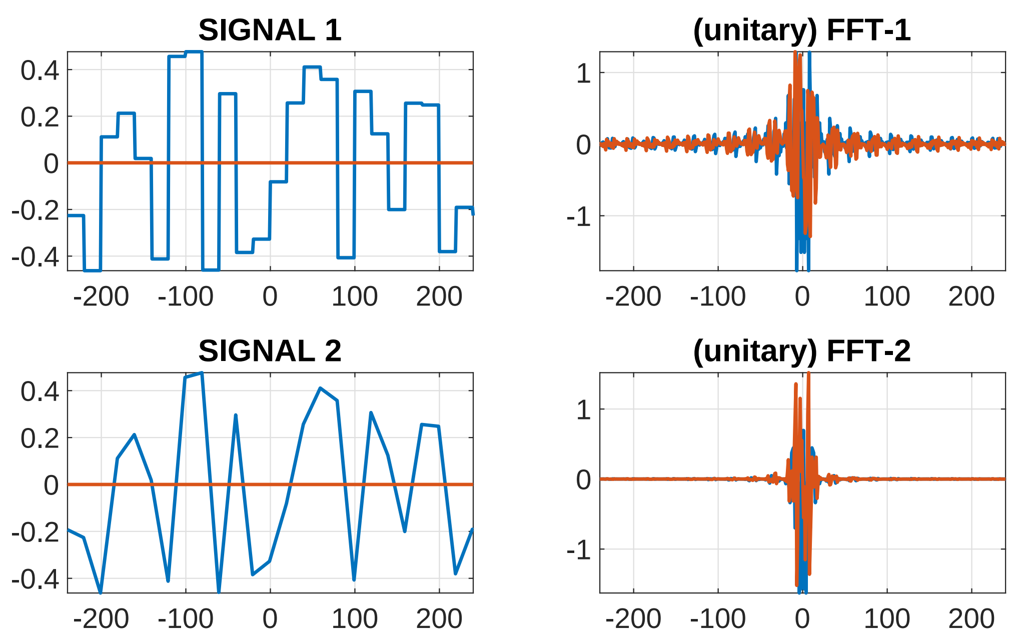

Abbildung 22: randpwlin12fft.eps
random step function and random piecewise linear functions and their DFTs. ATTENTION: for better vizualization the mean value has been subtracted, resp. the First Fourier coefficient has been set to zero!

SINC2dil.jp

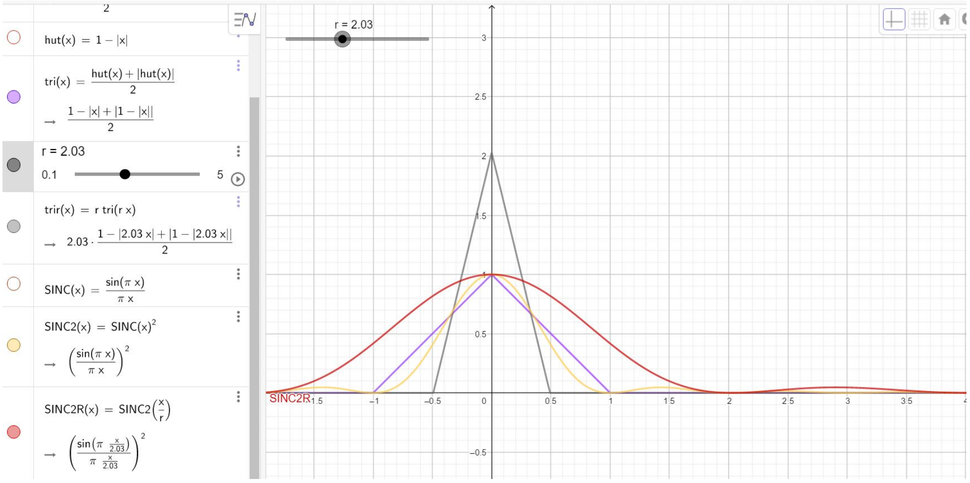

Abbildung 23: SINC2dil.jpThe triangular function and its Fourier transform: $SINC^2$.

15 Preparing for the Riemann-Lebesgue Lemma

In order to show that $\mathcal{F}(\boldsymbol{L}^1(\mathbb{R}^d)) \subset \boldsymbol{C}_0(\mathbb{R}^d)$ it will be sufficient to show that for a dense subspace of $\boldsymbol{L}^1(\mathbb{R}^d)$ one has control over the norm (via $\|\widehat{\mu}\|_\infty \leq \|\mu\|_{\boldsymbol{M}_b}$ respectively $\|\widehat{f}\|_\infty \leq \|f\|_{\boldsymbol{L}^1}, f \in \boldsymbol{L}^1(\mathbb{R}^d)$.

There are several different ways to show that $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) = \{\widehat{f} \,|\, f \in \boldsymbol{L}^1(\mathbb{R}^d)\} \subset \boldsymbol{C}_0(\mathbb{R}^d)$, in fact, it is a dense, but proper subspace of $\boldsymbol{C}_0(\mathbb{R}^d)$.

One of the classical methods to verify it is by resorting to step functions.

Let us first recall the situation for $d = 1$, i.e. the functions on the real line. We define the box-function as the real-valued symmetric function with looks like a square, namely First we claim that one has

$$\mathcal{F}(\text{box})(s) = \text{SINC}(s) = \sin(\pi s)/(\pi s), \quad s \in \mathbb{R}. \tag{118}$$ FTboxSINC

SINC FOUR JPG missing here! 03.12.2020

Using the standard rules for the Fourier transform with respect to dilation (producing narrow boxcar functions) and translation (building step functions) one can find that any step function over $\alpha\mathbb{Z}$ on $\mathbb{R}^d$ (with finitely non-zero values) has a Fourier transform which is a finite linear combination of stretched (dilated) and modulation (out of translation) SINC-functions, so it must belong to $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$.

Since we have the obvious estimate

$$\|\widehat{f}\|_\infty \leq \|f\|_{\boldsymbol{L}^1}, \quad f \in \boldsymbol{L}^1(\mathbb{R}),$$

and since obviously any $f \in \boldsymbol{C}_c(\mathbb{R})$ can be approximated in the $\|\cdot\|_1$-norm by such step functions (this is more or less by the construction of the Riemann integral!) we can conclude (based on the completeness of $\left(\boldsymbol{C}_0(\mathbb{R}), \|\cdot\|_\infty\right)$, that also the limit (which is of course $\widehat{f}$ must belong to $\boldsymbol{C}_0(\mathbb{R})$. >> $\mathbb{R}^d$, $d \geq 2$.

The non-integrability of the SINC-function is illustrated by the following plot (p.85 of SCRIPT) rInvThm1.jpg

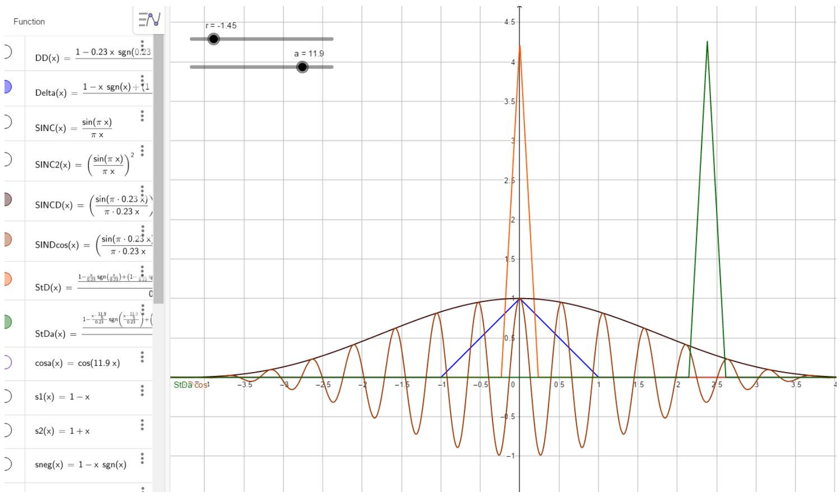

Abbildung 24: FourInvThm1.jpg: The ingredients of the Fourier inversion theorem: The shifted Dirac sequence goes to the modulated $SINC^2$ function.

15.1 The Riemann-Lebesgue Lemma for $\mathbb{R}^d$, $d \geq 2$

We just indicate that this can be done using tensor products, i.e. functions which are products of functions of each variable, following the general principle (first done for $d = 2$):

$$(f \otimes g)(x, y) = f(x)g(y), \quad x, y \in \mathbb{R}. \tag{119}$$ tensproddef02

Using MATLAB the tensor product is easy to realize, it is just the creation of a "multiplication table", i.e. an $n \times n$-matrix from two vectors of length n.
In order to make it independent from the given format of the input vectors $\mathbf{x}, \mathbf{y} \in \mathbb{C}$ we suggest to create the tensor product $\mathbf{z} := \mathbf{x} \otimes \mathbf{y}$ via

```
z = x(:) * y(:).';
```

Due to the exponential law one has (for pure frequencies well defined for different dimensions) one has a **separation of the variables**:

$$\chi_s = \chi_{s_1} \otimes \chi_{s_2}, \quad \forall s = (s_1, s_2) \in \mathbb{R}^d \tag{120}$$ sepvar00

Proof. The claim follows from the obvious splitting of the scalar product: for any $t = (t_1, t_2) \in \mathbb{R}^d$ (e.g. $t_1 \in \mathbb{R}^k$ and $t_2 \in \mathbb{R}^l$ with $d = k + l$) one has

$$\langle t, s \rangle = \langle t_1, s_1 \rangle + \langle t_2, s_2 \rangle$$

and consequently

$$exp(2\pi i \langle t, s \rangle) = exp(2\pi i \langle t_1, s_1 \rangle) \cdot exp(2\pi i \langle t_2, s_2 \rangle). \tag{121}$$ expsplit

$\square$

tensprodC0 **Lemma 21.**

$$\boldsymbol{C}_0(\mathbb{R}^k) \otimes \boldsymbol{C}_0(\mathbb{R}^l) \subset \boldsymbol{C}_0(\mathbb{R}^d). \tag{122}$$ C0tens12

INCplot4.pdf

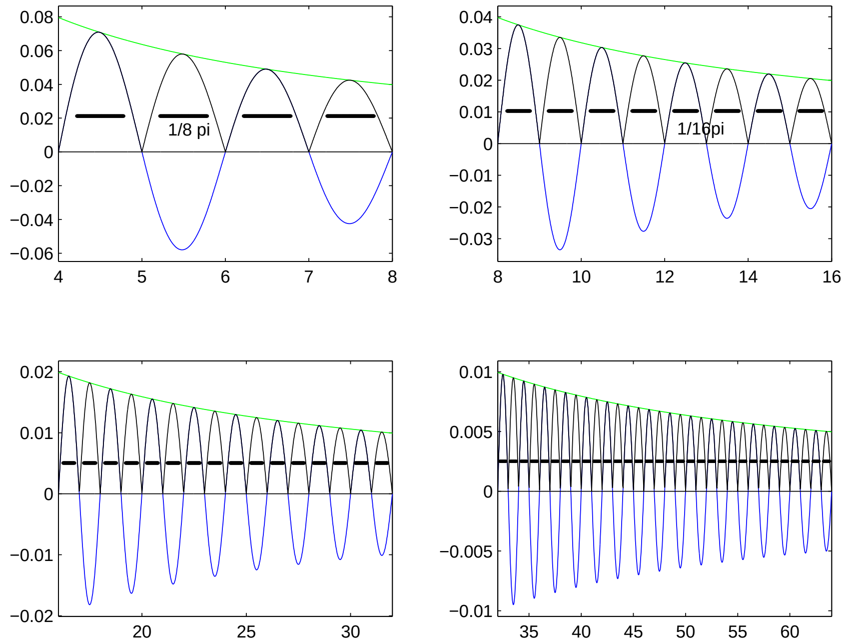

Abbildung 25: SINCplot4.pdf

This is a very simple exercise (!). (Exercise)

Remark 7. One can generate a Banach space, denoted by $\boldsymbol{C}_0\widehat{\otimes}\boldsymbol{C}_0$, which is a proper subspace of $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$. The dual space is called the space of *bimeasures*, which properly contains $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$.

FTsplit12 **Lemma 22.**

$$\mathcal{F}(f \otimes g) = \mathcal{F}(f) \otimes \mathcal{F}(g), \forall f, g \in \boldsymbol{L}^1. \tag{123}$$ FTtens

stepfctRd **Lemma 23.** *Step functions on $\mathbb{R}^d$ are obtained by sums of dilated translates of tensor products of boxcar functions. Hence their Fourier transform belongs to $\boldsymbol{C}_0(\mathbb{R}^d)$, and thus the proof of the Riemann Lebesgue Lemma remains valid.*

16 Approximate units for convolution

See GEOGEBRA classroom notes at `www.nuhag.eu/ETH20`

rInvThm1.png

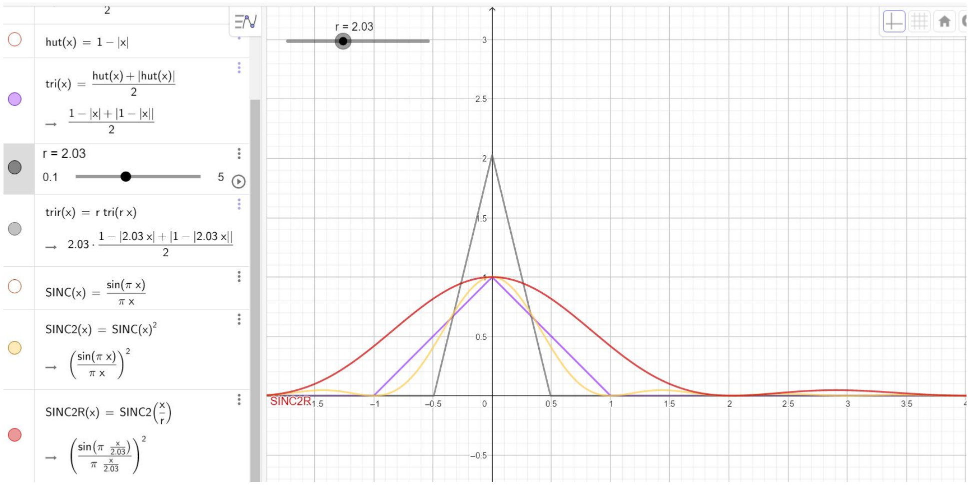

Abbildung 26: FourInvThm1.png
The triangular function and its Fourier transform: $SINC^2$. The compressed triangular function (compression by the factor 2, hence support $= [-1,1]/2 = [-1/2,1/2]$, and maximum $= 2$ at 0) has the red dilated version of the SINC-function as the Fourier transform, which in turn has its zeros at $2,4,6$ (and of course $-2,-4...$).

rInvThm1.png

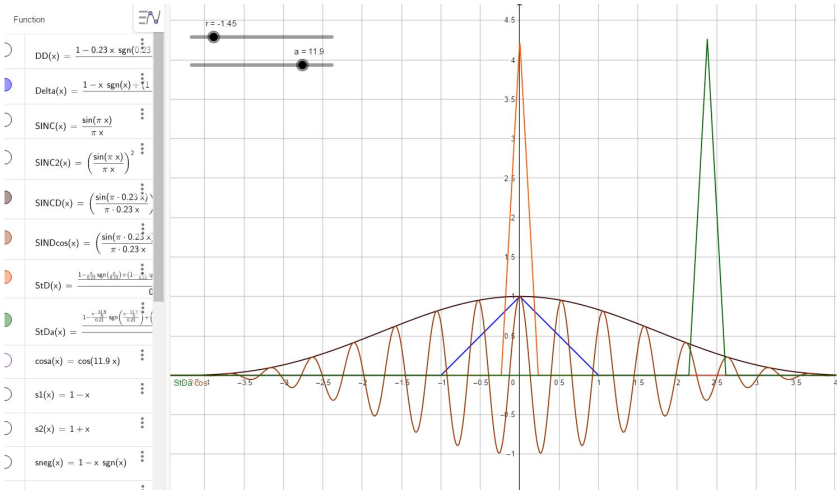

Abbildung 27: FourInvThm1.png: The ingredients of the Fourier inversion theorem: The shifted Dirac sequence goes to the modulated $SINC^2$ function.

The technical details of this part of the explanation are found in the Lecture Notes (Script file)

16.1 Discrete convolution, p.51

We have seen that the convolution $\mu := \mu_1 \star \mu_2$ is a well defined bounded measure for any pair of bounded measures μ_1, μ_2. We also have observed that

$$\delta_x \star \delta_y = \delta_{x+y}, \quad x, y \in \mathbb{R}^d.$$

The same hold true - by the same proof, for discrete measures over any discrete group, such as $\mathscr{G} = \mathbb{Z}$, or $\alpha\mathbb{Z}$ or $\Lambda = \boldsymbol{A} \star \mathbb{Z}^d \lhd \mathbb{R}^d$[12], but we will not make use of this abstract point of view.
Since the general bounded discrete measures are just limits (in the sense of the norm of $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$!) of finite discrete measures let us next take a look at the convolution of two finite discrete measures, say

$$\mu_1 = \sum_{k=1}^{n} a_k \delta_{x_k} \quad \text{and} \quad \mu_2 = \sum_{j=1}^{m} b_j \delta_{y_j},$$

we get a finite sum of discrete measures of the form

$$\delta_z = \delta_{x_k} \star \delta_{y_j}, 1 \leq k \leq n, 1 \leq j \leq m.$$

$$\mu = \sum_{l=1}^{r} c_l \delta_{z_l}, \quad \text{with } c_l = \sum_{(k,j): z_l = x_k + y_j} a_k b_j. \tag{124}$$ discrconv1

For the case that the sequence $(x_k)_{k=1}^n$ is an arithmetic progression, in the simplest case $x_k = k - 1$, for $1 \leq k \leq n$ and $y_j = j - 1$, for $1 \leq k \leq n$ we get:

$$\mu = \sum_{l=0}^{n+m-1} c_l \delta_l, \quad \text{with } c_l = \sum_{(k,j):(l-1)=(k-1)+(j-1)} a_k b_j. \tag{125}$$ discrconv2

With a slight change of notation, by putting $\alpha_{k-1} = a_k$ and $\beta_{j-1} = b_k$ and $\gamma_{l-1} = c_l$, and summing up to n (instead of n), one can describe the discrete convolution for measures supported on the semi-group $\mathbb{N}_0 = \{0, 1, 2, ...\}$, namely $\mu_1 = \sum_{k=0}^{n} \alpha_k \delta_k$ and $\mu_2 = \sum_{j=0}^{m} \beta_j \delta_j$ as

$$\mu = \left(\sum_{k=0}^{n} \alpha_k \delta_k \right) \star \left(\sum_{j=0}^{m} \beta_j \delta_j \right) = \sum_{l=0}^{n+m-1} \gamma_l \delta_l, \quad \text{with } \gamma_l = \sum_{(k,j): l = k+j} \alpha_k \beta_j. \tag{126}$$ discrconv3

But this is exactly the **Cauchy product**, i.e. the formula for the coefficients of the product polynomial $r(x) = p(x) \cdot q(x)$, for

$$p(x) = \sum_{k=0}^{n} \alpha_k x^k \quad \text{and} \quad q(x) = \sum_{j=1}^{m} \beta_j x^j.$$

[12] We use the symbol $\lhd$ to indicate that Λ is a *subgroup* of $(\mathbb{R}^d, +)$, in this case a lattice.

Exercise 5. With $p(x)$ and $q(x)$ given as above, or *degree* n and m respectively, we know that the resulting product polynomial $r(x) = p(x)q(x), x \in \mathbb{R}$ is a polynomial of degree $m + n$. Thus it can be described by $1 + n + m$ coefficients (index set is going from 0 to $m + n$). If we "multiply out" we find that:

The discrete convolution corresponds to pointwise multiplication of polynomials.

In other words, while we are doing a pointwise multiplication of polynomials (more correctly: of the polynomial function $x \to p(x)$ over $\mathbb{R}$ or $\mathbb{C}$) we have a corresponding operation (obviously commutative) at the level of coefficients, mapping the pair of coefficients (α,β, indexed by $\mathbb{Z}$) to γ.

Formally one typically collects the term in a systematic way:

$$\gamma_l = \sum_{k=0}^{l} \alpha_k \beta_{l-j} = \sum_{j=0}^{l} \beta_j \alpha_{l-j}, \quad 0 \leq l \leq r = m + n. \tag{127}$$ Cauchyprod2

Another observation is this one:

Exercise 6. Note that δ_0 (corresponding to $p(x) == 1$) is the unit element for convolution among the discrete measures, any OTHER such measure (once more: supported on $\mathbb{N}_0$) does not have an inverse otherwise.

Recall that a polynomial of the form $p(x) = \sum_{k=0}^{n} \alpha_k x^k$ with non-vanishing leading term x^n (i.e. with $\alpha_n \neq 0$) is said to be of *degree* n resp. of order $n + 1$ (meaning: it has potentially $n + 1$ non-zero coefficients).

The degree is of course good in order to find that it multiplies (the degree of $r(x) = p(x)q(x)$ is $r = n + m$!), while the order tells us how many (different) points we need in order to determine a polynomial. In fact, for any collection of $n + 1$ different points in the complex plane (specifically on the real line) it suffices to know the polynomial at those points in order to recover the polynomial completely!!!

One of the many ways to verify this is to make use of the idea of the Lagrange interpolation polynomials, for a given sequence $z_0, z_1, ... z_n \in \mathbb{C}$. We may look, for some fixed $k \in \{0, 1, ...n\}$ for a polynomial of degree k with $L_k(x_j) = \delta_{k,j}$ (Kronecker delta: equal to 1 for $k = j$ and zero for $k \neq j$).

The first step is to make sure that one has a polynomial which vanishes at all the points listed, except for $z = z_k$, so we just take

$$p_k(x) = \prod_{j:j \neq k} (z - z_j).$$

Then $p_k(x)$ has n linear factors, hence it is of degree n, but it cannot of one more zero at $z = z_k$, so we can take $L_k(x) = p_k(z)/p_k(z_k)$.

Given the values of the given polynomial at all those points it is now clear that

$$q(x) = \sum_{k=0}^{n} p(z_k) L_k(x), \quad x \in \mathbb{C} \tag{128}$$ Lagrint02

is a polynomial with $p(z_k) = q(z_k), 1 \leq k \leq n$. Hence the difference polynomial $p - q$ has $n + 1$ zeros, which implies that $p(x) = q(x)$, i.e. the coefficients are known.

In this whole story the so-called Vandermonde matrix and its (non-vanishing determinant, for $z_k \neq z_j$, pairwise) plays an important role.

Remark 8. There is also a very interesting and useful connection to probability theory, more precisely to discrete probability measures. Assume you through a coin, which tells you to take a step to the left or to the right (both with probability 1/2) along $\mathbb{Z}$, starting from 0. What is the probability of coming back to the range $\{-1, 0, 1\}$ after 7 steps? Or you trough a dice and you want to know what the probability is to reach more than 30 point after 8 times throwing the dice (or by summing the outcome of 8 independent dices? I plan to come back to such questions later on. It will also be a way to demonstrate the validity of the **Central Limit Theorem** in an *experimental setting*.

Maybe we end by providing a relatively simple invertibility criterion (valid for general Banach algebras). Recall, that δ_0 is the unit-element in the Banach algebra $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ (with convolution), since $\mu \star \delta_0 = \mu$ for all $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$. An element is called *invertible in $\boldsymbol{M}_b(\mathbb{R}^d)$* if there exists some $\nu \in \boldsymbol{M}_b(\mathbb{R}^d)$ such that $\mu \star \nu = \delta_0$.
Let us insert a relatively simple invertibility criterion (valid for general Banach algebras) at this place:

Mbinvert **Lemma 24.** *Assume that $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ has the property that for some $\rho > 0$:*

$$\|\mu - \rho\delta_0\|_{\boldsymbol{M}_b} < \rho.$$

Then μ is invertible inside of $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$, i.e. there exists some $\nu \in \boldsymbol{M}_b(\mathbb{R}^d)$, with the extra property $\|\nu\|_{\boldsymbol{M}_b} \leq \rho \cdot (1 - \|\rho^{-1}\mu - \delta_0\|_{\boldsymbol{M}_b})^{-1}$, such that $\mu \star \nu = \delta_0 = \nu \star \mu$.

Proof. First of all let us observe that upon rescaling (replacing μ by $\rho^{-1}\mu$) we may assume that $\|\rho^{-1}\mu - \delta_0\|_{\boldsymbol{M}_b} = \gamma < 1$. Let us write $\mu_1 := \rho^{-1}\mu$.
The key argument here is the simple convergence of a geometric series, using the submultiplicativity of the norm on $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ (see formula (convMbsubm 54)). In particular, assume that $\kappa \in \boldsymbol{M}_b(\mathbb{R}^d)$ satisfies $\|\kappa\|_{\boldsymbol{M}_b} \leq \gamma < 1$, then by the general estimate for convolution powers ($k^{*n} = k \star k \star ...k$, n-times) of κ we have:

$$\|\kappa^{r*}\|_{\boldsymbol{M}_b} \leq \|\kappa\|^r_{\boldsymbol{M}_b} \leq \gamma^r, \quad r \in \mathbb{N}. \tag{129}$$ powerMbnorm

By our assumption the measure $\kappa = \delta_0 - \mu_1$ satisfies this assumption, hence $\nu_1 := \sum_{k=0}^{\infty} \kappa^{k*}$ is an absolutely convergent series[13] in $\boldsymbol{M}_b(\mathbb{R}^d)$, with

$$\|\nu_1\|_{\boldsymbol{M}_b} \leq \sum_{k=0}^{\infty} \gamma^k = \frac{1}{1-\gamma}.$$

We claim that $\mu_1 \star \nu_1 = \delta_0 = \nu_1 \star \mu_1$. In fact $\mu_1 = \kappa - \delta_0$ and consequently

$$\mu_1 \star \nu_1 = (\kappa - \delta_0) \star \nu_1 = \sum_{k=1}^{\infty} \kappa^{*k} - \sum_{k=0}^{\infty} \kappa^{*k} = \kappa^{0*} = \delta_0.$$

Since convolution is commutative the inverse to the original measure μ is then given as $\nu := \rho\nu_1$, hence

$$\|\nu\|_{\boldsymbol{M}_b} \leq \frac{\rho}{1-\gamma}.$$

□

[13]With the convention: $\kappa^{0*} = \delta_0$!

Remark 9. We have to make a few simple observations:

- Assume that $\mu \in \boldsymbol{M}_d(\mathbb{R}^d)$ is a discrete measure. Then the condition is only satisfied if the coefficient of the Dirac measure at zero is sufficiently large. In fact, one can reduce that discussion to the case $\mu = \delta_0 + \sum_{k=1}^{\infty} c_k \delta_{x_k}$ with $\sum_{k=1}^{\infty} |c_k| < 1$.
- If the discrete measure is supported by some lattice $\Lambda \lhd \mathbb{R}^d$, then the inverse measure is supported by the same lattice, because the family of bounded discrete measures supported on some lattice Λ is a closed subalgebra of $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ (and if such a measure has an inverse in $\boldsymbol{M}_b(\mathbb{R}^d)$ then it also belongs to the subalgebra (as the proof shows).
- It is clear that invertibility of μ in the measure algebra implies invertibility of the continuous function $\widehat{\mu} \in \boldsymbol{C}_b(\mathbb{R}^d)$. This implies among others that for some $\eta > 0$ one has $|\widehat{\mu}(s)| \geq \eta$, for all $s \in \mathbb{R}^d$. Of course the inverse can be directly determined by inverting the Fourier transform.
- Wiener's inversion theorem is a kind of converse to the last statement.

16.2 Discrete Fourier Transform, FFT

One way to look at the **Discrete Fourier Transform** is to say, that it realizes the *discrete convolution* for polynomials, one the degree is high enough (to recover the coefficients of a polynomial from its values over $\mathbb{U}_N$, the (multiplicative, but commutative) group of unit roots of order N, which is a discrete subgroup of the unitary or *torus* group:

$$\mathbb{T} = \mathbb{U} := \{z \in \mathbb{C} \,|\, |z| = 1\},$$

which has $1 \in \mathbb{C}$ as its multiplicative neutral element, and with $z^{-1} = \overline{z}$ (complex conjugate) since we have

$$\overline{z}z = z\overline{z} = |z|^2 = 1^2 = 1, \quad z \in \mathbb{U}.$$

For $N \geq 1$ the unit roots of order N form a group with N elements and *generator*

$$\omega_N = e^{2\pi i/N} = \cos(2\pi/N) + i * \sin(2\pi/N).$$

For teaching purposes the choice $N = 8$ is quite convenient, because it is easy to draw the unit roots of order 8 on the blackboard. Here is how it looks. They are created by the following MATLAB command (see MATLAB Expo Talk):

```
u = exp(2*pi*i*(0:1/N:(N-1)/N));
```

In this way the discrete Fourier transform can be viewed as the linear mapping which assigns to each $\alpha = (\alpha_k)_{k=0}^{N-1}$ (with of course practically speaking $n << N-1$)

```
aa = rand(1,5);  u5 = uproots(5);
compnorm ( polygala(fliplr(aa),conj(u5)) , fft(aa));
comparing the evaluation of a polynomial at the unit roots of order N
 with the values of the FFT routine, which is a linear routine
% quotient of norms: norm(x)/norm(y) = 1
% difference of normalized versions  = 3.3197e-16
```

ots8dem2.eps

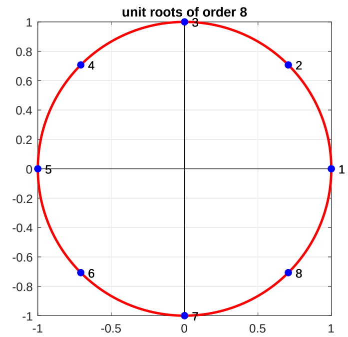

Abbildung 28: uroots8dem2.eps: ω_N is at position 2, or $360/8 = 45$ degrees.

Here we are working with *row vectors aa*, hence the flip operator has to be used in the form of `fliplr`. The unit roots of order 5 come in their natural order, i.e. as $\omega_5^k, k = 0, 1, 2, 3, 4$.
Since the Fourier transform has expression of the form $e^{-2\pi isx}$ they have to be taken in the opposite (namely clock-wise sense), still starting from $1 = e^0$. This is way we have to evaluate the polynomial at $conj(u5)$.
The justification for the `fliplr` command comes from the fact that we have the convention that `polyval([1,2,3],10) == 123` i.e. to describe the polynomials in *this context*, starting with the leading term!
So the summary is: Given a sequence $(a_k)_{k=1}^N \in \mathbb{C}^n$ the DFT (Discrete Fourier Transform) corresponds to the samples of a polynomial at the unit roots of oder N.
NOTE: The connection to the continuous Fourier transform will be discussed later on! It is an interesting subject.
Next steps: What can we say about the matrix $F = fft(eye(N))$ representing this linear transform? It has a number of interesting features:

1. It is a Vandermonde matrix
2. Up to the normalizing factor $\sqrt{N}$ it is a unitary matrix;
3. All the entries are complex numbers of absolute value 1;
4. The matrix is equal to its transpose matrix, hence the inverse matrix if just the $1/N$ times the conjugate, or $\mathrm{inv}(F) = F'/N$.

Note that this scheme correspond exactly to the multiplication of natural numbers described in the decimal system! After all, we have

```
polyval([1,2,3],10) == 123 = 1*10^2 + 2*10^1 + 3*10^0.
```

Exercise: Compute 11111^2, or 1111111^2? Do you see any pattern?

16.3 Towards impulse response

The fact, that we have a concrete realization of any TILS T on $\boldsymbol{C}_0(\mathbb{R}^d)$ allows us to extend its action to all of $\left(\boldsymbol{C}_b(\mathbb{R}^d),\,\|\cdot\|_\infty\right)$ by the same pairing (let us write H instead of T now):

$$Hf(y) = \mu(T_y f^{\checkmark}), f \in \boldsymbol{C}_b(\mathbb{R}^d) \quad \mu(f) = H(f^{\checkmark})(0), f \in \boldsymbol{C}_0(\mathbb{R}^d).$$

In contrast to the setting discussed earlier there is a problem when one tries to recover the system from the measure defined by $\mu(f) = H(f^{\checkmark})(0)$ because this might be the zero functional, even if the operator (abstractly defined on $\boldsymbol{C}_{ub}(\mathbb{R}^d)$ or $\boldsymbol{C}_b(\mathbb{R}^d)$ and commuting with translations) is not the zero operator. This observation, which is a consequence of the *Hahn-Banach Theorem* in functional analysis, has been explained in several papers by I. Sandberg, who talks about the *scandal in system theory*. See **sa98 :** [82], **sa02-1 :** [83],**sa94 :** [81], and several other papers in a similar spirit by the same author.
The key observation is the fact that in a situation as we have it here, namely a chain of closed, proper subspaces (all Banach spaces with respect to the sup-norm, namely

$$\left(\boldsymbol{C}_0(\mathbb{R}^d),\,\|\cdot\|_\infty\right) \hookrightarrow \left(\boldsymbol{C}_{ub}(\mathbb{R}^d),\,\|\cdot\|_\infty\right) \hookrightarrow CbRdN \tag{130}$$ propC0CubCb

there are always non-zero, bounded linear functionals which vanish on the smaller subspace, e.g. on $\boldsymbol{C}_{ub}(\mathbb{R}^d)$, but vanishing on $\boldsymbol{C}_0(\mathbb{R}^d)$.
The existence of such linear functionals is grounded (in the mathematical theory) on the Hausdorff maximal principle which is equivalent to the *axiom of choice*, i.e. it is highly *non-constructive*. The alleged systems H would not be observable within finite time, because their reaction to input signal $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ would be the zero signal, so they cannot be really observed.
As a consequence, the *scandal* is a mild scandal, and has more to to with an inappropriate description of the situation, the *mathematical modelling of BIBOS systems*. As we have tried to point out, the choice of $\left(\boldsymbol{C}_0(\mathbb{R}^d),\,\|\cdot\|_\infty\right)$ as the domain of such systems as a starting point, followed by an extension of that system (as just done) is a better way compared to a direct modelling of BIBOS systems as operators acting on $\left(\boldsymbol{C}_b(\mathbb{R}^d),\,\|\cdot\|_\infty\right)$ or (as it is done in **un20 :** [89] on $\left(\boldsymbol{L}^\infty(\mathbb{R}^d),\,\|\cdot\|_\infty\right)$, which requires Lebesgue integration theory and causes further problems, because the point evaluation is not well defined for general elements $h \in \boldsymbol{L}^\infty(\mathbb{R}^d)$!).
See [31] for more details. See also the recent article by M. Unser [89], which also puts forward (mathematical) short-comings in the discussion of linear translation invariant systems in the engineering literature.
It is clear that the same rules apply as for the action of $C_\mu : f \mapsto \mu \star f$ defined on $\boldsymbol{C}_0(\mathbb{R}^d)$, i.e. such systems commute with translations, and this implies that the systems preserve uniform continuity, i.e.

$$\boldsymbol{M}_b(\mathbb{R}^d) \star \boldsymbol{C}_b(\mathbb{R}^d) \subseteq \boldsymbol{C}_b(\mathbb{R}^d), \quad \boldsymbol{M}_b(\mathbb{R}^d) \star \boldsymbol{C}_{ub}(\mathbb{R}^d) \subseteq \boldsymbol{C}_{ub}(\mathbb{R}^d). \tag{131}$$ MbCbCub

16.4 Exponential law and characters

Due to the *exponential law* one has

$$\chi_s{}^{\checkmark}(z) = \chi_{-s}(z) \quad \text{and} \quad T_z(\chi_{-s}{}^{\checkmark}) = \chi_s(z) \cdot \chi_{-s},$$

and therefore

$$\mu \star \chi_s(z) = C_\mu \chi_s(z) = \mu(T_z(\chi_{-s}{}^{\checkmark})) = \mu(\chi_{-s}) \cdot \chi_s(z) = \widehat{\mu}(s)\chi_s(z), \tag{132}$$ FourConvEig

or expressed more compactly: χ_s is an *eigenvector* for C_μ with *eigenvalue* $\hat{\mu}(s)$:

$$C_\mu(\chi_s) = \widehat{\mu}(s) \cdot \chi_s. \tag{133}$$ eigenCmu1

$$\mathcal{F}(\delta_y) = \chi_{-y}, \quad \text{but?} \quad \mathcal{F}^{-1}(\chi_{-y}) = \delta_y? \tag{134}$$ FourDirac2b

17 The dual of $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right) : \left(\boldsymbol{L}^\infty(\mathbb{R}^d), \|\cdot\|_\infty\right)$

LinfRdLeb

We want to discuss shortly the dual space of $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$. Abstractly speaking it is of course just the space of all bounded linear functionals on the Banach space $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$.

In books on functional analysis one finds that it is the space usually denoted by $\left(\boldsymbol{L}^\infty(\mathbb{R}^d), \|\cdot\|_\infty\right)$, the space of all *essentially bounded, measurable functions*. More precisely, the set of equivalence classes of such functions, according the to usual identification

$$h_1 \equiv h_2 \quad \Leftrightarrow \quad h_1(x) = h_2(x) \text{ a.e.}. \tag{135}$$ equivLinf

Recall that a complex-valued measurable function is *essentially bounded* if there exists $C > 0$ such that

$$|h(x)| \le C \text{ a.e.}. \tag{136}$$ essbounddef

The natural (semi-)norm on

$$\mathcal{L}^\infty(\mathbb{R}^d) := \{h \,|\, h \text{ is essentially bounded}\} \tag{137}$$ Linfdef

is given by

$$\|h\|_\infty := \inf\{C \,|\, |h(x)| \le C \text{ a.e.}\}, \tag{138}$$ Linfnormdef

turns $\boldsymbol{L}^\infty(\mathbb{R}^d)$ (the space of equivalence classes, as described above) into a normed space. The classical approach using Lebesgue integration allows to show that one can identify $\left(\boldsymbol{L}^\infty(\mathbb{R}^d), \|\cdot\|_\infty\right)$ with $\boldsymbol{L}^{1'}(\mathbb{R}^d)$ (the dual space for $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$).

LinfdualLi **Lemma 25.** *There is a natural identification of $\left(\boldsymbol{L}^\infty(\mathbb{R}^d), \|\cdot\|_\infty\right)$ with the dual space of $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$, such that every $\sigma \in \boldsymbol{L}^{1'}(\mathbb{R}^d)$ is determined by a unique (up to equivalence) $h \in \mathcal{L}^\infty(\mathbb{R}^d)$ such that*

$$\sigma(f) = \int_{\mathbb{R}^d} f(x)h(x)dx. \tag{139}$$ Lidualrepr

This identification is isometric, i.e.

$$\|h\|_{\boldsymbol{L}^\infty(\mathbb{R}^d)} = \sup_{\|f\|_1 \le 1} \left| \int_{\mathbb{R}^d} f(x)h(x)dx \right| = \sup_{k \in \boldsymbol{C}_c(\mathbb{R}^d) : \|k\|_1 \le 1} \left| \int_{\mathbb{R}^d} k(x)h(x)dx \right| \tag{140}$$ Lidualiso01

17.1 Multipliers and related facts 15.10.2020

The first step can be shown without reference to the use of measurable functions.

CbinLinfRd **Lemma 26.** *$\left(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\right)$ is a closed subspace of $\boldsymbol{L}^\infty(\mathbb{R}^d)$. Moreover, $\boldsymbol{C}_c(\mathbb{R}^d)$ and hence $\boldsymbol{C}_0(\mathbb{R}^d)$ are w^*-dense subspaces of $\boldsymbol{L}^\infty(\mathbb{R}^d)$.*

Proof. First we have to show that any $h \in \boldsymbol{C}_b(\mathbb{R}^d)$ defines a bounded linear function on $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$, by checking the estimate (Lidualiso01 140) above, by simply testing it on $h \in \boldsymbol{C}_b(\mathbb{R}^d)$ and $f \in \boldsymbol{C}_c(\mathbb{R}^d)$ (using the Riemann integral).
The result is again

$$|\sigma_h(f)| \leq \|h\|_\infty \|f\|_{\boldsymbol{L}^1} \quad h \in \boldsymbol{C}_b(\mathbb{R}^d), f \in \boldsymbol{L}^1(\mathbb{R}^d),$$

or simply

$$\|\sigma_h\|_{\boldsymbol{L}^\infty(\mathbb{R}^d)} \leq \|h\|_\infty, \quad h \in \boldsymbol{C}_b(\mathbb{R}^d).$$

In order to show equality one has to find for $\varepsilon > 0$ some $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ with $\|k\|_{\boldsymbol{L}^1} \leq 1$ but

$$|\sigma_h(k)| \geq \|h\|_\infty - \varepsilon.$$

First let us assume that $h(x) \geq 0$ for all $x \in \mathbb{R}^d$. Then the continuity of $h \in \boldsymbol{C}_b(\mathbb{R}^d)$ implies that we can find some small ball $B_r(x_0)$ around some point $x_0 \in \mathbb{R}^d$ such that

$$|h(x)| \geq \|h\|_\infty - \varepsilon, \quad \forall x \in B_r(x_0).$$

Then for any $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ with $k(x) \geq 0$ and $\|k\|_{\boldsymbol{L}^1} = 1$ and with compact support inside of $B_r(x_0)$ one has, using the non-negativity of all the involved terms:

$$\int_{\mathbb{R}^d} h(x)k(x)dx \geq \min_{z \in B_r(x_0)} h(x) \int_{\mathbb{R}^d} k(x)dx \geq \|h\|_\infty - \varepsilon.$$

For the general, complex-valued case we have to invoke the phase function. Since we may assume that k has small support inside of some open ball $B_s(x_0)$ (with $s < r$) we can find some other function ϕ supported in $B_r(x_0)$ with $\phi \cdot k = k$ (i.e. $\phi(x) = 1$ on $\mathrm{supp}(k)$).
Since the continuous function $h \in \boldsymbol{C}_b(\mathbb{R}^d)$ is bounded away from zero on $\mathrm{supp}(\phi)$ the function $\varphi(x) := |h(x)|\phi(x)/h(x)$ is a continuous function, and we work with $k \cdot \varphi$ instead of k in the estimate from below. In fact, we still have $\mathrm{supp}(k \cdot \varphi) \subseteq B_r(x_0)$ and also $\|k \cdot \varphi\|_{\boldsymbol{L}^1} = 1$ (since the phase factor has absolute value 1). Thus

$$\sigma_h(k \cdot \varphi) = \int_{\mathbb{R}^d} h(x)\varphi(x)k(x)dx = \int_{\mathbb{R}^d} |h(x)|k(x)dx,$$

so that the rest of the estimate can be used unchanged.
Finally we have to argue why $\boldsymbol{C}_c(\mathbb{R}^d)$ (a subspace of $\boldsymbol{C}_b(\mathbb{R}^d) \subset \boldsymbol{L}^\infty(\mathbb{R}^d)$, via $h \mapsto \sigma_h$) is a w^*-dense subspace of $\boldsymbol{L}^\infty(\mathbb{R}^d)$. In other words we have to show that any $f \in \boldsymbol{L}^\infty(\mathbb{R}^d)$ (**OR should we use** σ**?**) can be approximated in the w^*-sense by a suitable sequence $(k_n)_{n\geq 1}$ from $\boldsymbol{C}_c(\mathbb{R}^d)$. REST to be done! $\square$

LinfLiconv **Lemma 27.**

$$\boldsymbol{L}^\infty(\mathbb{R}^d) \star \boldsymbol{L}^1(\mathbb{R}^d) \subseteq \boldsymbol{C}_{ub}(\mathbb{R}^d). \tag{141}$$ LinfLiCub

Proof. The convolution of $\sigma \in \boldsymbol{L}^\infty(\mathbb{R}^d)$ with $f \in \boldsymbol{L}^1(\mathbb{R}^d)$ can be viewed as a pointwise convolution product of the form

$$\sigma \star f(y) = \sigma(T_y f^\checkmark), \quad f \in \boldsymbol{L}^1(\mathbb{R}^d).$$

Consequently one has

$$|[\sigma \star f](z) - [\sigma \star f](z-y)| \leq |\sigma(T_z f^\checkmark - T_{z-y} f^\checkmark)| \leq \|\sigma\|_{\boldsymbol{L}^\infty(\mathbb{R}^d)} \|T_z(f^\checkmark - T_y f^\checkmark)\|_{\boldsymbol{L}^1(\mathbb{R}^d)}.$$

Since the $\boldsymbol{L}^1$-norm is translation and flip-invariant we have

$$\|T_z(f^\checkmark - T_y f^\checkmark)\|_{\boldsymbol{L}^1(\mathbb{R}^d)} = \|f - T_{-y} f\|_{\boldsymbol{L}^1(\mathbb{R}^d)} \to 0 \text{ for } y \to 0.$$

□

Lemma 28.

$$\boldsymbol{M}_b(\mathbb{R}^d) \star \boldsymbol{L}^\infty(\mathbb{R}^d) \subseteq \boldsymbol{L}^\infty(\mathbb{R}^d), \tag{142}$$ MbLinfconv

with the corresponding norm estimate

$$\|\mu \star f\|_{\boldsymbol{L}^\infty(\mathbb{R}^d)} \leq \|\mu\|_{\boldsymbol{M}_b} \|f\|_{\boldsymbol{L}^\infty(\mathbb{R}^d)}, \quad \mu \in \boldsymbol{M}_b(\mathbb{R}^d), f \in \boldsymbol{L}^\infty(\mathbb{R}^d). \tag{143}$$ MbLinfconvequ

Proof. We start from the fact that $\boldsymbol{M}_b(\mathbb{R}^d) \star \boldsymbol{L}^1(\mathbb{R}^d) \subseteq \boldsymbol{L}^1(\mathbb{R}^d)$, i.e. from the fact that $\boldsymbol{L}^1(\mathbb{R}^d)$ is a closed ideal in the Banach convolution algebra $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$.
Hence we have (by the general principles of functional analysis) some abstract action on the dual space, which can be expected to be the convolution by $\mu^\checkmark$. Hence let us consider the mapping $f \mapsto C'_\mu(f)$ given by the action on $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ (by definition a dense subspace of $\boldsymbol{L}^1(\mathbb{R}^d)$):

$$C'_{\mu^\checkmark}(f)(k) = f(\mu^\checkmark \star k), \quad f \in \boldsymbol{L}^\infty(\mathbb{R}^d), k \in \boldsymbol{C}_c(\mathbb{R}^d).$$

We know that $\left(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\right)$ is a closed subspace of $\left(\boldsymbol{L}^\infty(\mathbb{R}^d), \|\cdot\|_\infty\right)$.... (also $\boldsymbol{C}_{ub}(\mathbb{R}^d)$)

□

Linfxxx **Lemma 29.** *We will prove later on:*

$$\mathcal{H}_{\mathbb{R}^d}(\boldsymbol{L}^1(\mathbb{R}^d), \boldsymbol{C}_{ub}(\mathbb{R}^d)) = \mathcal{H}_{\mathbb{R}^d}(\boldsymbol{L}^1(\mathbb{R}^d), \boldsymbol{C}_b(\mathbb{R}^d)) \approx \boldsymbol{L}^\infty(\mathbb{R}^d).$$

$$\mathcal{H}_{\mathbb{R}^d}(\boldsymbol{L}^1(\mathbb{R}^d), \boldsymbol{C}_0(\mathbb{R}^d)) \subset \boldsymbol{L}^\infty(\mathbb{R}^d) \quad \textit{as a closed, proper subspace.}$$

17.2 Pure frequencies are eigen-vectors for TILS

For any operator $Tf = \mu \star f$ the output for a given pure frequency χ_s is just

$$T(\chi_s) = \mu(\chi_{-s})\chi_s = \widehat{\mu}(s)\chi_s \tag{144}$$

chareigenTILS

But in reality one can only produce a localized (time-limited) version of a pure frequency, and then one will find the

lconvpf1.pdf

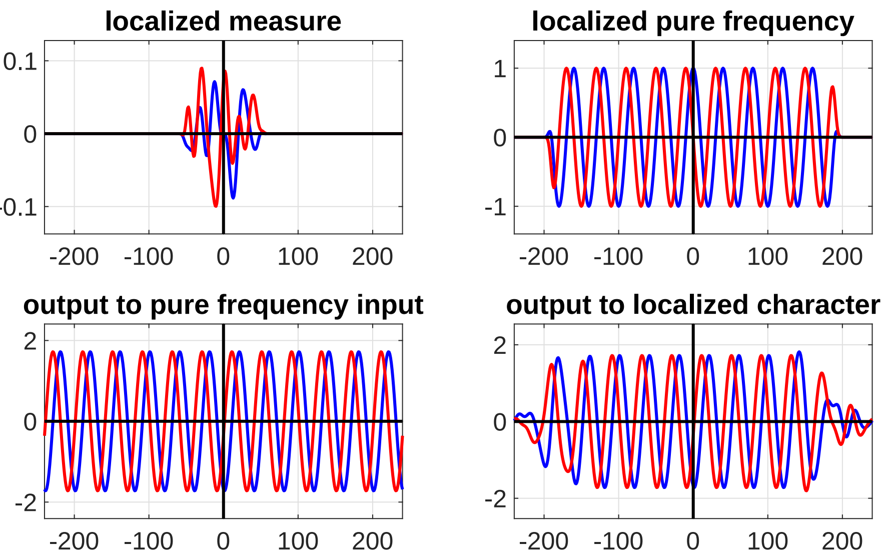

Abbildung 29: localconvpf1.pdf We see the localized convolution kernel or measure μ in the left upper window. Below we see the output of a pure frequency under convolution (done via FFT) with a particular pure frequency χ_s. The right, upper window shows the same (input) pure frequency, but restricted to a compact domain. Note that the cosine-part is evenly arranged (blue) around the center, while the sine-part is in read. Note the phase shift in the output side! Finally the right lower plot show the output of the localized pure frequency under the same system. Locally we have the same phase shift, so if you compare local pieces of the input (localized pure frequency) with the equally localized output in order to determine the phase shift, resp. $\widehat{\mu}(s)$

inputoutp2.pdf

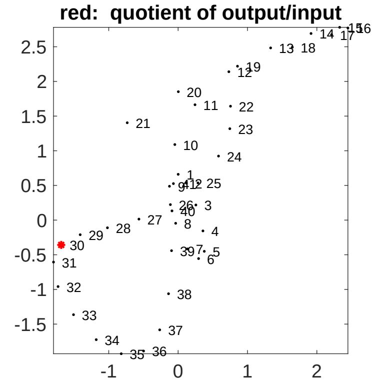

Abbildung 30: inputoutp2.pdf This plot shows the central Fourier coefficients (low frequencies) and marked in read the quotient (phase shift) obtained as illustrated by picture localconvpf1.pdf 17.2. It shows up at exactly the right position and with a highly precise value. The needed length of the input signal for a good determination of $\widehat{\mu}(s)$ depends on the size of the support of μ.

18 Approximate units for convolution

Definition 6. A bounded family of functions $(g_\rho)_{0<\rho\leq 1}$ is called a bounded approximate unit for the Banach convolution algebra $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ (or a *Dirac sequences*) if

$$\sup_{0<\rho\leq 1} \|g_\rho\|_{\boldsymbol{L}^1} < \infty, \tag{145}$$ Linobd01

and

$$\lim_{\rho\to 0} \|g_\rho \star f - f\|_{\boldsymbol{L}^1} = 0, \quad \forall f \in \boldsymbol{L}^1(\mathbb{R}^d). \tag{146}$$ apprconvLi

By a simple approximation argument one can show, that in the presence of condition (Linobd01 145) condition (apprconvLi 146) is equivalent to the validity of (apprconvLi 146) for just $f \in \boldsymbol{C}_c(\mathbb{R}^d)$ (It is a good exercise to verify this claim).

Note: In the book of Kanwal [68] the SINC function (which is NOT in $\boldsymbol{L}^1(\mathbb{R})$) is chosen as a Dirac sequence, but this is somewhat problematic. We will see that the compressed SINC function is convergent to the Dirac measure in the since of the dual space $\boldsymbol{S}_0'$ (but it is not w^*-convergent in $\boldsymbol{M}_b(\mathbb{R}^d)$, because even the approximating elements do not belong to this space).

Hence we will first prove the following result:

Diracgrho **Lemma 30.** *Given any* $g \in \boldsymbol{C}_c(\mathbb{R}^d)$ **with** $\int_{\mathbb{R}^d} g(x)dx = 1$ *and* $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ *we have*

$$\lim_{\rho\to 0} \|\mathrm{St}_\rho g \star k - k\|_\infty = 0. \tag{147}$$ Strhoconvunif

Since all the involved functions have common compact support (for $|\rho| \leq 1$) this implies convergence in $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$*, and finally by approximation for all of* $\boldsymbol{L}^1(\mathbb{R}^d)$*:*

$$\lim_{\rho\to 0} \|\mathrm{St}_\rho g \star f - f\|_{\boldsymbol{L}^1} = 0, \quad f \in \boldsymbol{L}^1(\mathbb{R}^d). \tag{148}$$ StrhoconvLinc

tLinFLi1.jpg

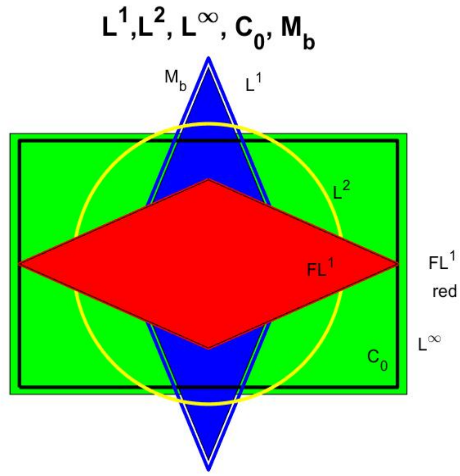

Abbildung 31: LiLtLinFLi1.jpg: blue rhomb represents $\boldsymbol{L}^1(\mathbb{R}^d)$, the green rectangle represents $\boldsymbol{L}^\infty(\mathbb{R}^d)$ resp. $\boldsymbol{C}_b(\mathbb{R}^d)$, with $\boldsymbol{C}_0(\mathbb{R}^d)$ inside (marked as a black line). Rotation of the blue rhomb (corresponding to taking the Fourier transform) gives us the red symbol for $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$. That it is contained inside of $\boldsymbol{C}_0(\mathbb{R}^d)$ is the statement of the Riemann-Lebesgue Lemma. The yellow circle represents $\boldsymbol{L}^2(\mathbb{R}^d)$. Note that $\boldsymbol{L}^1 \cap \boldsymbol{L}^\infty(\mathbb{R}^d) \subset \boldsymbol{L}^2(\mathbb{R}^d)$!

Proof. We can restrict our attention to the case $\rho \in (0,1]$ and assume without loss of generality that $\|k\|_{\boldsymbol{L}^1} = 1$. Since $\mathrm{supp}(g) \subset B_R(0)$ for some $R > 0$ the family $(\mathrm{St}_\rho g)_{\rho \in (0,1]}$ has joint support in $B_R(0)$. Following [78] (Ch.2.1.4) we can write $\mathrm{St}_\rho g \star k - k$ as pointwise integral compared to $f(x) = \int_{\mathbb{R}^d} \mathrm{St}_\rho g(y) \cdot f(x) dy$, taking into account that (by the transformation formula for integrals) this integral equals $\int_{\mathbb{R}^d} g(y) dy = 1$ for any choice of ρ. Thus we obtain the pointwise estimate

$$|[\mathrm{St}_\rho g \star k](x) - k(x)| \leq \int_{\mathbb{R}^d} |k(x-y) - k(x)||[\mathrm{St}_\rho g](y)| dy$$

Note that $\mathrm{supp}(\mathrm{St}_\rho g) = \rho\,\mathrm{supp}(g) \subseteq \rho B_R(0) = B_{\rho R}(0)$, can be made arbitrarily small, and we continue with

$$|[\mathrm{St}_\rho g \star k](x) - k(x)| \leq \int_{B_{\rho R}(0)} |k(x-y) - k(x)||[\mathrm{St}_\rho g](y)| dy$$

Using the uniform continuity of $f \in \boldsymbol{C}_c(\mathbb{R}^d)$ we can estimate the differences $|k(x-y) - k(x)|$ for $y \in \rho R \leq \delta$, and get thus an estimate of the form

$$\|\mathrm{St}_\rho g \star k - k\|_\infty \leq \varepsilon \int_{\mathbb{R}^d} |\mathrm{St}_\rho g(y)| d(y) = \varepsilon \|g\|_{\boldsymbol{L}^1}.$$

Finally we make use of the compact support of f: Assuming that $\mathrm{supp}(k) \subseteq B_K(0)$ for some $K > 0$ (or $k(x) = 0$ for $|x| \leq K$) we find that the pointwise estimate is valid on $B_{R+K}(0)$ in the pointwise sense, and the expression equals zero outside of this compact set. In this situation uniform convergence implies $\boldsymbol{L}^1$-convergence.
Finally, once the result is valid for the dense subspace $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ one can approximate it and obtain (148) is valid for every $f \in \boldsymbol{L}^1(\mathbb{R}^d)$. StrhoconvLino □

rinvsp5B.jpg

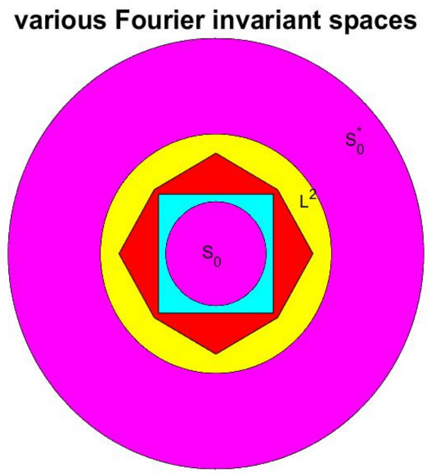

Abbildung 32: Fourinvsp5B.jpg: Within the universe of $\boldsymbol{S}_0'(\mathbb{R}^d)$ we have the Hilbert space $\boldsymbol{L}^2(\mathbb{R}^d)$ (yellow). It contains $\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1$, the space to which the Fourier inversion applies (in red). Inside of it (turquoise square) we have $\boldsymbol{W}(\mathbb{R}^d) \cap \mathcal{F}(\boldsymbol{W}(\mathbb{R}^d))$. This is the domain for the Poisson summation formula, or the space of continuous Riemann integrable functions to which the Fourier inversion via Riemann integrals applies. Finally we have inside the magenta space of test functions called $\boldsymbol{S}_0(\mathbb{R}^d)$.

Remark 10. If we do not want to normalize the compressed function g to have $\hat{g}(0) = 1$ the approximation will be to $\widehat{g}(0)f$ (instead of f).

Based on the proof of Lemma (30) we come up with the full statement, which looks a bit complicated, but is based on very simple and natural approximation arguments:

odapprunit00 **Lemma 31.** *For any $g \in \boldsymbol{L}^1(\mathbb{R}^d)$ with $\hat{g}(0) = \int_{\mathbb{R}^d} g(y)dy = 1$ the family $(\mathrm{St}_\rho g)_{\rho \to 0}$ forms a bounded approximate unit for $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$.*

Proof. Clearly the proof will go by reduction to the case that $g \in \boldsymbol{C}_c(\mathbb{R}^d)$, using the density of $\boldsymbol{C}_c(\mathbb{R}^d)$ in $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ (by definition!).
Our goal is to find, given $f \in \boldsymbol{L}^1(\mathbb{R}^d)$ and $\varepsilon > 0$, some parameter ρ_0 such that $\rho \in (0, \rho_0]$ implies for the given $g \in \boldsymbol{L}^1(\mathbb{R}^d)$:

$$\|\mathrm{St}_\rho g \star f - f\|_{\boldsymbol{L}^1} \leq \varepsilon. \tag{149}$$ stroappr1

Without loss of generality we may assume $\|f\|_{\boldsymbol{L}^1} = 1$.
Recall our observation that $g \mapsto \widehat{g}(0) = \int_{\mathbb{R}^d} g(y)dy$ is a continuous mapping from $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ into $\mathbb{C}$. Hence, by the density of $\boldsymbol{C}_c(\mathbb{R}^d)$ we can guarantee that for any $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ with $\|g - k\|_{\boldsymbol{L}^1} \leq \eta$ (for whatever $\eta > 0$, to be chosen later on) one has

$$|\hat{k}(0) - 1| = |\hat{k}(0) - \hat{g}(0)| \leq \|k - g\|_{\boldsymbol{L}^1} \leq \eta. \tag{150}$$ intesteta1

We will use now Lemma 30 to show, that one has for ρ well chosen

$$\|\mathrm{St}_\rho k \star f - \hat{k}(0)f\|_{\boldsymbol{L}^1} \leq \varepsilon/3. \tag{151}$$ estim009

FourInvWWRnew.jpg

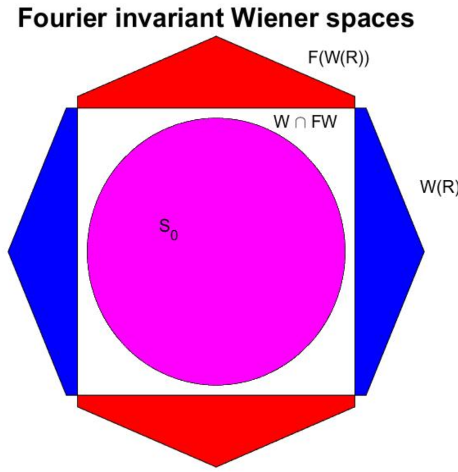

Abbildung 33: FourInvWWRnew.jpg: Showing the "small" Fourier invariant spaces, namely the (here) white square, the intersection of $\boldsymbol{W}(\mathbb{R}^d)$ with $\mathcal{F}(\boldsymbol{W}(\mathbb{R}^d))$, shown as blue resp. red object. Obviously this is Fourier invariant. But inside of this space we have $\boldsymbol{S}_0(\mathbb{R}^d) = \mathcal{F}(\boldsymbol{S}_0(\mathbb{R}^d))$, which is also invariant under the Fractional FT (see later).

Finally we observe that one we have the following estimate, *independent* of ρ!:

$$\|\mathrm{St}_\rho k \star f - \mathrm{St}_\rho g \star f\|_{\boldsymbol{L}^1} \leq \|\mathrm{St}_\rho (k-g)\|_{\boldsymbol{L}^1} \|f\|_{\boldsymbol{L}^1} = \|k-g\|_{\boldsymbol{L}^1} \|f\|_{\boldsymbol{L}^1} \leq \eta. \tag{152}$$ estim010

Using the triangle inequality we come up with the overall estimate

$$\|\mathrm{St}_\rho g \star f - f\|_{\boldsymbol{L}^1} \leq \|\mathrm{St}_\rho (g-k) \star f\|_{\boldsymbol{L}^1} + \|\mathrm{St}_\rho k \star f - \hat{k}(0) f\|_{\boldsymbol{L}^1} + |\hat{k}(0) - 1| \|f\|_{\boldsymbol{L}^1} \tag{153}$$ estim011

and combining (estim011 153), (estim010 152) and (intesteta1 150) we obtain:

$$\|\mathrm{St}_\rho g \star f - f\|_{\boldsymbol{L}^1} \leq \eta + \varepsilon/3 + \eta < \varepsilon. \tag{154}$$ estim011b

Hence one can obtain the required estimate by *first* choosing $\eta = \varepsilon/3$, then picking $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ such that $\|g-k\|_{\boldsymbol{L}^1} < \eta$ and then, for fixed (auxiliary function) k one chooses ρ_0 appropriately, so that the overall estimate (stroappr1 149) is achieved for any $\rho \in (0, \rho_0]$. $\square$

Let us give an application of formula (Stroconv7 172) below respectively of Lemma (bdapprunit00 31) above (once we know that $\mathcal{F}(\mathrm{SINC}^2) = \Delta$, due to the Fourier inversion theorem for $(\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1)(\mathbb{R}^d)$).

Libandl **Corollary 3.** *The band-limited function in $\boldsymbol{L}^1(\mathbb{R}^d)$, i.e.*

$$\boldsymbol{L}^1_{bdl}(\mathbb{R}^d) := \{f \in \boldsymbol{L}^1(\mathbb{R}^d) \mid \mathrm{supp}(\hat{f}) \ \mathit{compact}\}$$

is a dense subspace of $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$.

Proof. Due to the Fourier inversion theorem we know that $\mathcal{F}(\mathrm{SINC}^2) = \Delta^{\checkmark} = \Delta$, and thus $g = \mathrm{SINC}^2$ can be used to form a bounded approximate identity of the form $(\mathrm{St}_\rho g)_{\rho \to 0}$. Hence we can approximate any $f \in \boldsymbol{L}^1(\mathbb{R}^d)$ by convolution products of the form $\mathrm{St}_\rho g \star f$. But such function are band-limited, because we have $\mathrm{supp}(\Delta) \subset B_1(0)$ (in the multi-dimensional setting we also have compact support), and by the convolution theorem

$$\mathrm{supp}(\mathcal{F}(\mathrm{St}_\rho g \star f)) = \mathrm{supp}(\mathrm{D}_\rho \Delta \cdot \widehat{f}) \subset 1/\rho \cdot B_1(0) = B_{1/\rho}(0).$$

$\square$

Note: Although compactly supported elements are dense in $\big(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\big)$, and also band-limited elements are dense, the only function $f \in \boldsymbol{C}_c(\mathbb{R}^d)$ which is also band-limited is the zero-function (because $f \in \boldsymbol{C}_c(\mathbb{R}^d)$ implies that $\widehat{f}$ is an analytic function, and a non-zero analytic function cannot vanish on a whole interval, leave alone outside a given, compact set).
In reducing the Fourier transform to compact support it is important to be careful and to *avoid sharp truncation*, because this corresponds to multiplication by a boxcar function, resp. convolution by a (not integrable) SINC-function! For the case of the boxcar function (step-function in $\boldsymbol{L}^1(\mathbb{R})$) the negative effect is known as the *Gibbs phenomenon* (see WIKIPEDIA for details).

ecBump1.eps

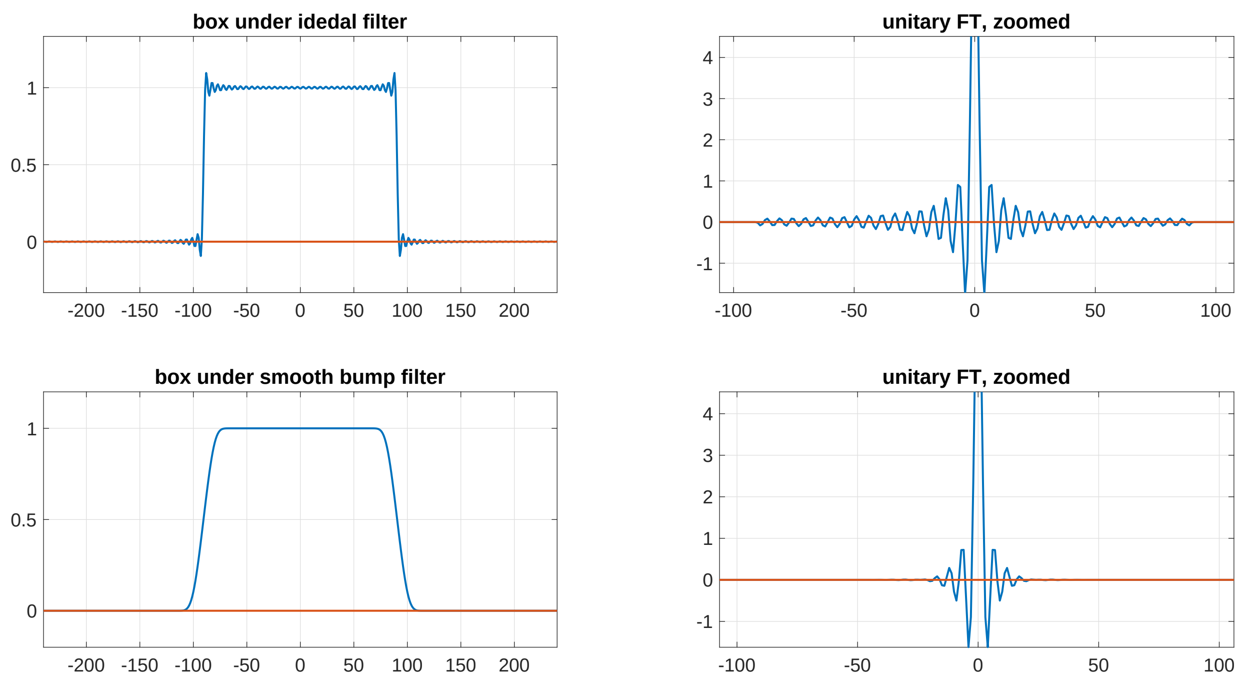

Abbildung 34: FiltRecBump1.eps

NOTE: It may be difficult to find the optimal balance, because in the above situation the choice of ρ depends on k. If one chooses a function which is approximating g very well it might be a not-so-smooth function k, and in such a case a choice of ρ_0 results which might be suboptimal (with respect to the final claim (stroappr1 149).

The book of Kanwal ([68]) also provides the following example (snapshot of a GEOGEBRA experiment):

lDirac02.jpg

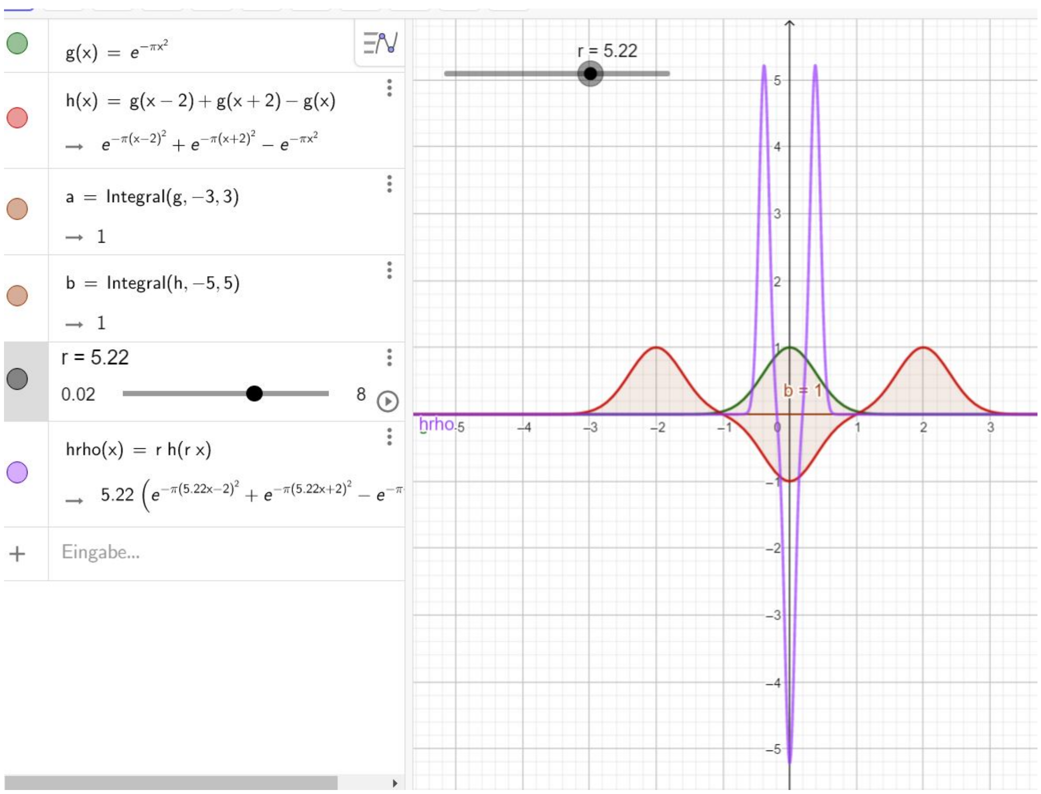

Abbildung 35: KanwalDirac02.jpg This GEOGEBRA SNAPSHOT exhibits a sum of three Gaussians, two positive ones and a negative one, with total mass $2 - 1 = 1$. The corresponding compressed versions of $h = T_{-2}g_0 - g_0 + T_2g_0$ have the property that $Strhoh(0) \to -\infty$ for $\rho \to 0$. A compressed version (member of the Dirac sequence) is displayed in violet color.

19 Some abstract bilinear considerations

bilinconv **Lemma 32.** *Assume the M is a bilinear mapping $\boldsymbol{B}^1 \times \boldsymbol{B}^2 \to \boldsymbol{B}$ which is bounded, in the following sense: There exists $C > 0$ such that*

$$\|M(b_1, b_2)\|_{\boldsymbol{B}} \leq C\|b_1\|_{\boldsymbol{B}^1}\|b_1\|_{\boldsymbol{B}^2}, \quad b_1 \in \boldsymbol{B}^1, b_2 \in \boldsymbol{B}^2. \tag{155}$$ bilinbd

Then M is a continuous mapping, meaning that

$$f_0 = \boldsymbol{B}^1 - \lim_{n\to\infty} f_n, \quad \textit{and} \quad g_0 = \boldsymbol{B}^2 - \lim_{n\to\infty} g_n$$

implies that

$$M(f_0, g_0) = \boldsymbol{B} - \lim_{n\to\infty} M(f_n, g_n).$$

The proof is based on the same argument that is used to prove that multiplication or addition are continuous on $\mathbb{R}$. Given a' close to a and b' close to b we know how to show that $|ab - a'b'|$ is small.

Proof. Given $\varepsilon > 0$ we show that for a suitable choice of n_0 we can guarantee that

$$\|M(f_n, g_n) - M(f_0, g_0)\|_{\boldsymbol{B}} \leq \varepsilon, \quad \forall n \geq n_0. \tag{156}$$ Mfngnest1

We first observe that convergence of the sequence g_n implies boundedness in $(\boldsymbol{B}^1, \|\cdot\|^{(1)})$. More precisely, there exists n_1 such that for $n \geq n_1$ on has

$$\|g_n\|_{\boldsymbol{B}^1} \leq \|g_n - g_0\|_{\boldsymbol{B}^1} + \|g_0\|_{\boldsymbol{B}^1} \leq \|g_0\|_{\boldsymbol{B}^1} + 1 := C_1 < \infty. \tag{157}$$ gnbound1

The key estimate comes by a application of the triangle inequality, by inserting the mixed term $M(f_0, g_n)$. Thus

$$\|M(f_n, g_n) - M(f_0, g_0)\|_{\boldsymbol{B}} \leq \|M(f_n, g_n) - M(f_0, g_n)\|_{\boldsymbol{B}} + \|M(f_0, g_n) - M(f_0, g_0)\|_{\boldsymbol{B}} \tag{158}$$ Mfngnest2

The first term in (158) can be estimated (first for fixed n)

$$\|M(f_n, g_n) - M(f_0, g_n)\|_{\boldsymbol{B}} \leq C\|f_n - f_0\|_{\boldsymbol{B}^1}\|g_n\|_{\boldsymbol{B}^2} \leq C_1 C\|f_n - f_0\|_{\boldsymbol{B}^1}. \tag{159}$$ Mfngnest3

The second term in (159) can be estimated as

$$\|M(f_0, g_n) - M(f_0, g_0)\|_{\boldsymbol{B}} \leq C\|f_0\|_{\boldsymbol{B}^1}\|g_n - g_0\|_{\boldsymbol{B}^2}. \tag{160}$$ Mfngnest4

This is suffices to choose $n_0 \geq n_1$ such that

$$\|f_n - f_0\|_{\boldsymbol{B}^1} < \varepsilon/(2CC_1) \quad \text{and} \quad \|g_n - g_0\|_{\boldsymbol{B}^2} \leq \varepsilon/(2C). \tag{161}$$ Mfngnest5

Combining the last three estimates we have verified that (156) is valid for $n \geq n_0$. $\square$

absconv0 **Proposition 7.** *Assume that M is a bounded, bilinear mapping as in Lemma [?]. Assume that $\boldsymbol{C}_1$ and $\boldsymbol{C}_2$ are dense subspaces of $(\boldsymbol{B}^1, \|\cdot\|^{(1)})$ and $(\boldsymbol{B}^2, \|\cdot\|^{(2)})$ respectively. Let T_α be a bounded net of operators on $(\boldsymbol{B}^1, \|\cdot\|^{(1)})$, i.e. assume that*

$$|||T|||_{\boldsymbol{B}^1} \leq B, \quad \text{i.e.} \quad \|T_\alpha(f)\|_{\boldsymbol{B}^1} \leq B\|f\|_{\boldsymbol{B}^1}, \quad \forall f \in \boldsymbol{B}^1, \tag{162}$$ absconv2

for some positive constant $B > 0$. Assume also that T_0 is a fixed operator[14] *on $(\boldsymbol{B}^1, \|\cdot\|^{(1)})$, also with $|||T_0|||_{\boldsymbol{B}^1} \leq B$. Then the following implication is true:*
Assume that one has verified that

$$\boldsymbol{B} - \lim_\alpha M(T_\alpha k, f) = M(T_0 k, f), \quad \forall k \in \boldsymbol{C}_1, f \in \boldsymbol{C}_2, \tag{163}$$ convdens1

then one can conclude that this limit relation is valid on $\boldsymbol{B}^1 \times \boldsymbol{B}^2$, i.e. that

$$\boldsymbol{B} - \lim_\alpha M(T_\alpha g, h) = M(T_0 g, h), \quad \forall g \in \boldsymbol{B}^1, h \in \boldsymbol{B}^2, \tag{164}$$ convdens

First note that the proof of Lemma bilinconv 32 is no only valid for sequences but also for *nets*. In fact it provides an error estimate for the output, given some deviations at the input.

Proof. Given a general input pair $g \in \boldsymbol{B}^1$ and $h \in \boldsymbol{B}^2$ we have to obtain, for a given $\varepsilon > 0$, an estimate of the form

$$\|M(T_\alpha g, h) - M(T_0 g, h)\|_{\boldsymbol{B}} \leq \varepsilon. \tag{165}$$ convdens0

As usual we are comparing the term to be estimated to terms for which relation (convdens1 163) can be used. By inserting two intermediate terms we split $M(T_\alpha g, h) - M(T_0 g, h)$ as:

$$[M(T_\alpha g, h) - M(T_\alpha k, f)] + [M(T_\alpha k, f) - M(T_0 k, f)] + [M(T_0 k, f) - M(T_0 g, h)]. \tag{166}$$ convdens2

we can estimate the expression (convdens0 165) by estimating the three terms separately. Let us first take care of the first one (the third can be done by exactly the same estimate, just setting $\alpha = 0$):
Note that we have the following situation: If we choose $k \in \boldsymbol{C}_1$ appropriately, with $\|g - k\|_{\boldsymbol{B}^1} < \eta$, then

$$\|T_\alpha g - T_\alpha k\|_{\boldsymbol{B}^1} = \|T_\alpha(g - k)\|_{\boldsymbol{B}^1} \leq |||T_\alpha|||_{\boldsymbol{B}^1} \|g - k\|_{\boldsymbol{B}^1} \leq B\eta,$$

and similarly $\|h - f\|_{\boldsymbol{B}^2}$ can be made arbitrarily small by a suitable choice of $f \in \boldsymbol{C}_2 \subset \boldsymbol{B}^2$. we can apply Lemma bilinbd 155 in order to obtain

$$\|M(T_\alpha g, h) - M(T_\alpha k, f)\|_{\boldsymbol{B}} < \varepsilon/3, \tag{167}$$ convdens3

and of course (as said, by the same argument) also the third term can be estimated as

$$\|M(T_0 g, h) - M(T_0 k, f)\|_{\boldsymbol{B}} < \varepsilon/3. \tag{168}$$ convdens3b

Now given this choice of $k \in \boldsymbol{C}_1$ and $f \in \boldsymbol{C}_2$ we can apply the assumption, i.e. the middle term can be estimated as well, according to the assumptions. There exists some α_0 such that $\alpha \succeq \alpha_0$ implies

$$\|M(T_\alpha k, f) - M(T_0 k, f)\|_{\boldsymbol{B}} < \varepsilon/3. \tag{169}$$ convdens5

A simple combination of the last three estimates implies the required estimate (convdens0 165). □

[14] Typically the w^*-limit of the operators T_α.

19.1 Applications of the abstract principle

First let us mention a situation, which we have already encountered for the discussion of convolution. How have we demonstrated that $\boldsymbol{M}_b(\mathbb{R}^d) \star \boldsymbol{C}_0(\mathbb{R}^d) \subseteq \boldsymbol{C}_0(\mathbb{R}^d)$?
We can say that this corresponds to case

$$M(\mu, f) = \mu \star f$$

which maps $\boldsymbol{M}_b(\mathbb{R}^d) \times \boldsymbol{C}_{ub}(\mathbb{R}^d)$ into $\boldsymbol{C}_{ub}(\mathbb{R}^d)$ with the estimate

$$\|\mu \star f\|_\infty \leq \|\mu\|_{\boldsymbol{M}_b} \|f\|_\infty .$$

Compare with the formulas (48) and (49), in the proof that any bounded measures describes a bounded linear operator from $\boldsymbol{C}_0(\mathbb{R}^d)$ into $\boldsymbol{C}_0(\mathbb{R}^d)$ (and not just $\boldsymbol{C}_{ub}(\mathbb{R}^d)$).
The proof was using the fact that

$$\boldsymbol{M}_{comp} = \{\mu = \sum_{i \in F} \mu\, \psi_i \text{ for some finite set } F\} \tag{170}$$

is a dense subspace of $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ and $\boldsymbol{C}_c(\mathbb{R}^d)$ is a dense subspace of $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$.
The explicit demonstration that also $\mu \star k$ has compact support (vanishes outside of some ball of radius $R > 0$), i.e. that

$$\boldsymbol{M}_{comp} \star \boldsymbol{C}_c(\mathbb{R}^d) \subset \boldsymbol{C}_c(\mathbb{R}^d) \subset \boldsymbol{C}_0(\mathbb{R}^d)$$

was the decisive step, the rest was just approximation in the spirit of Lemma (32).
Let us now demonstrate a number of concrete special cases of the abstract principle put forward in Lemma 7. In fact, the formulation of the abstract arguments above came from the observation that the arguments used in quite different situations show a formal similarity. One verifies a convergence statement for some special cases and goes on to extend it to the general case, using estimates for convolution or multiplication operators, combined with approximation arguments.

1. The first choice will be the setting of $(\boldsymbol{B}^1, \|\cdot\|^{(1)}) = (\boldsymbol{B}^2, \|\cdot\|^{(2)}) = \big(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\big)$, with $\boldsymbol{C}_1 = \boldsymbol{C}_2 = \boldsymbol{C}_c(\mathbb{R}^d)$, and operators St_ρ, for $\rho \to 0$, and the limit operator $T_0 = \widehat{\mu}(0)\delta_0$ (a rank one operator from $\boldsymbol{L}^1(\mathbb{R}^d)$ to $\boldsymbol{M}_b(\mathbb{R}^d)$)[15]

   We have seen that the operators St_ρ are isometric on $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$, since they are the adjoint of the operators D_ρ (which are value preserving, hence isometric on $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$). Thus we have $B = 1$ in Lemma 7.

   The bilinear mapping M is given as $M(\mu, f) = \mu \star f$, for $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ and $f \in \boldsymbol{C}_0(\mathbb{R}^d)$. Since we have $\|\mu \star f\|_\infty \leq \|\mu\|_{\boldsymbol{M}_b} \|f\|_\infty$ we have $C = 1$.

   For compactly supported measure and $f \in \boldsymbol{C}_c(\mathbb{R}^d)$ one can use the uniform continuity of f in order to derive the validity of the assumption (163) in this case.

   The conclusion of Proposition 7 is then: For $g \in \boldsymbol{L}^1(\mathbb{R}^d), \mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ one has

$$\|\mathrm{St}_\rho \mu \star f - \widehat{\mu}(0) f\|_{\boldsymbol{L}^1} \to 0 \text{ for } \rho \to 0. \tag{171}$$

[15] A slight deviation from described setting, since $\delta_0 \notin \boldsymbol{L}^1(\mathbb{R}^d)$.

Restricted to functions $g \in \boldsymbol{L}^1(\mathbb{R}^d)$ (by normalization) with $\int_{\mathbb{R}^d} g(x)dx = \hat{g}(0) = 1$ we show that compression of $\boldsymbol{L}^1(\mathbb{R}^d)$ functions produces bounded, approximate units for $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ (with convolution):

$$\lim_{\rho \to 0} \|\mathrm{St}_\rho g \star f - f\|_{\boldsymbol{L}^1} = 0, \tag{172}$$ Stroconv7

i.e. the family $(\mathrm{St}_\rho g)_{\rho \to 0}$ forms a bounded approximate unit for $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$.

2. The second choice is the pair $(\boldsymbol{B}^1, \|\cdot\|^{(1)}) = (\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ and $(\boldsymbol{B}^2, \|\cdot\|^{(2)}) = \left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$, again with $M(\mu, f) = \mu \star f$, hence $C = 1$. The operators we choose are the discretization operators D_Ψ. Since we know that $\|D_\Psi \mu\|_{\boldsymbol{M}_b} \leq \|\mu\|_{\boldsymbol{M}_b}$ for all $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ we also can choose $B = 1$. The limiting operator will be the identity operator on $\boldsymbol{M}_b(\mathbb{R}^d)$.

   As $\boldsymbol{C}_1$ we use the subspace of compactly supported measures in $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$, and $\boldsymbol{C}_c(\mathbb{R}^d)$ as dense subspace $\boldsymbol{C}_2$ of $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$.

   The validity of the convergence

$$\lim_{|\Psi| \to 0} \|D_\Psi \mu \star k - \mu \star k\|_\infty = 0$$

   for compactly supported measures and $f \in \boldsymbol{C}_c(\mathbb{R}^d)$ is easy to check, because all the involved functions $D_\Psi \mu \star k$ have *joint compact support*, and thus can be reduced to the pointwise convergence[16]

$$D_\Psi \mu \star f(y) = D_\Psi \mu(T_y k^\checkmark) \to \mu(T_y k^\checkmark) = \mu \star k(y)$$

3. Exchanging the space $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$ by $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ one comes up with the claim: For any $f \in \boldsymbol{L}^1(\mathbb{R}^d)$ and $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ one has:

$$\|D_\Psi \mu \star f - \mu \star f\|_{\boldsymbol{L}^1} = 0, \quad f \in \boldsymbol{L}^1(\mathbb{R}^d). \tag{173}$$ Lidiscr1

More such applications will be given below.

19.2 weak* convergence of bounded measures

Just for illustration of the same kind of arguments we state the following result:

Dpsiwst01 **Lemma 33.** *A bounded sequence $(\mu_n)_{n \geq 1}$ in $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ is w^*-convergent to $\mu_0 \in \boldsymbol{M}_b(\mathbb{R}^d)$ if and only if one has for every BUPU Ψ:*

$$D_\Psi \mu_n(f) \to \mathrm{D}_\Psi \mu_0(f), \quad \forall f \in \boldsymbol{C}_c(\mathbb{R}^d). \tag{174}$$ Dpsiwstconv

Proof. Assume that $\mu_0 = w^*\text{-}\lim_{n \to \infty} \mu_n$. Then one has for any BUPU Ψ;

$$D_\Psi \mu_n(f) = \mu_n(\mathrm{Sp}_\Psi(f)) \to_{n \to \infty} \mu_0(\mathrm{Sp}_\Psi(f)) = D_\Psi \mu_0(f)$$

[16] The derivation of *uniform convergence over compact sets* should be given in more details, to come!

by the assumption (just because $\mathrm{Sp}_\Psi(f) \in \boldsymbol{C}_0(\mathbb{R}^d)$).
For the converse observe that we have for any BUPU $\|D_\Psi \mu\|_{\boldsymbol{M}_b} \leq \|\mu\|_{\boldsymbol{M}_b}$. Let us assume that $\sup_{n\geq 1} \|\mu_n\|_{\boldsymbol{M}_b} \leq C_M < \infty$. Of course this implies that also $\|\mu_0\|_{\boldsymbol{M}_b} \leq C_M$ (this will be used later on, below).
By consequence one can extend the validity of (Dpsiwstconv 174) to $f \in \boldsymbol{C}_0(\mathbb{R}^d)$.

For general $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ one has for $n \geq n_0$ (suitably chosen)

$$|\mu_n(f) - \mu_0(f)| \leq |\mu_n(f) - \mu_n(\mathrm{Sp}_\Psi f)| + |\mu_n(\mathrm{Sp}_\Psi f) - \mu_0(\mathrm{Sp}_\Psi f)| + |\mu_0(\mathrm{Sp}_\Psi f) - \mu_0(f)| < \varepsilon.$$

In fact, first we can guarantee, for given f that for any BUPU Ψ with $|\Psi| \leq \delta_0$ (chosen properly) on has

$$\|f - \mathrm{Sp}_\Psi f\|_\infty \leq \varepsilon/(3C_M).$$

Then the first term (and similarly the last term) in the estimate above can be controlled

$$|\mu_n(f - \mathrm{Sp}_\Psi f)| \leq \|\mu_n\|_{\boldsymbol{M}_b} \|f - \mathrm{Sp}_\Psi f\|_\infty < \varepsilon/3.$$

The middle term follows from the identity

$$D_\Psi \mu(f) = \mu(\mathrm{Sp}_\Psi f), \quad f \in \boldsymbol{C}_0(\mathbb{R}^d),$$

which is the middle term, which tends to zero by choosing n appropriately. $\square$

The **Fundamental Relationship for the Fourier Stieltjes transform**, namely

$$\mu(\hat{\nu}) = \nu(\hat{\mu}), \quad \mu, \nu \in \boldsymbol{M}_b(\mathbb{R}^d), \tag{175}$$ FRFST

can also be derived in the same way. We choose the mapping $M(\mu, \nu) = \mu(\mathcal{F}(\nu))$, which is bilinear, well defined and bounded, thanks to the estimate

$$|\mu(\hat{\nu})| \leq \|\mu\|_{\boldsymbol{M}_b} \|\hat{\nu}\|_\infty \leq \|\mu\|_{\boldsymbol{M}_b} \|\nu\|_{\boldsymbol{M}_b}, \quad \mu, \nu \in \boldsymbol{M}_b(\mathbb{R}^d). \tag{176}$$ FRFSTest1

As a dense subspace we can choose the compactly supported measures. So let us verify it for compactly supported measures. As it is clear that it is valid for discrete measures (please check it directly yourself for $\mu = \delta_x$ and $\nu = \delta_y$):

$$\mu(\hat{\nu}) = \delta_x(\chi_{-y}) = e^{-2\pi i y x} = e^{-2\pi i x y} = \nu(\hat{\mu}). \tag{177}$$ FRFSTDirac

Hence we have (by taking absolutely convergent sums of this type):

$$D_\Psi \mu \left(\mathcal{F}(\mathrm{D}_\Psi \nu)\right) = \mathrm{D}_\Psi \nu \left(\mathcal{F}(D_\Psi \mu)\right). \tag{178}$$ FRFSTdisc1

In order to go to the limit we take the intermediate step, namely

$$D_\Psi \mu \left(\mathcal{F}(\nu)\right) = \nu \left(\mathcal{F}(D_\Psi \mu)\right), \tag{179}$$ FRFSTdisc2

which can be verified by verifying that for both sides the transition from $\mathrm{D}_\Psi \nu$ to ν can be justified, at least for case that $\nu \in \boldsymbol{M}_{comp}$ (and then by taking limits for general $\nu \in \boldsymbol{M}_b(\mathbb{R}^d)$).
In a second step the transition from (FRFSTdisc2 179) to the general formula is obtained in a similar way, again assuming first that $\mu \in \boldsymbol{M}_{comp}$ and then again by taking norm limits.
Details will be worked out later (hopefully, hgfei).
The last item is helpful in order to verify that the Fourier Stieltjes transform is injective (without making use of the Fourier inversion theorem, see script for details), but this question can be settled in many different ways.

Actually, similar results are valid for $\left(\boldsymbol{L}^p(\mathbb{R}^d),\, \|\cdot\|_p\right)$, for $1 \leq p < \infty$ and even for so-called *homogeneous Banach spaces* in the sense of Y. Katznelson (see **ka76:** [69]).
NOTE: The abstract principle described above makes it necessary to understand which spaces are contained in other space, and which ones are *dense in each other*.
Another abstract principle which plays an important role here is the following:

densdual **Lemma 34.** *Assume that* $(\boldsymbol{B}^1,\, \|\cdot\|^{(1)})$ *is a dense subspace of* $(\boldsymbol{B}^2,\, \|\cdot\|^{(2)})$*. Then there is a continuous embedding of* $\left(\boldsymbol{B}_2', \|\cdot\|_{\boldsymbol{B}_2'}\right)$ *into* $\left(\boldsymbol{B}_1', \|\cdot\|_{\boldsymbol{B}_1'}\right)$.

Proof. First of all it is clear that any bounded linear functional σ on $(\boldsymbol{B}^2,\, \|\cdot\|^{(2)})$ will act as a bounded linear functional on $\boldsymbol{B}^1$ by restriction. In fact, by the assumption about a continuous embedding we have

$$\|f\|_{\boldsymbol{B}^2} \leq C\|f\|_{\boldsymbol{B}^1}, \quad \forall f \in \boldsymbol{B}^1. \tag{180}$$ contemb01

Consequently we have

$$|\sigma(f)| \leq \|\sigma\|_{\boldsymbol{B}_2'} \|f\|_{\boldsymbol{B}^2} \leq [C\|\sigma\|_{\boldsymbol{B}_2'}]\|f\|_{\boldsymbol{B}^1}, \tag{181}$$ contemb02

and thus we have

$$\|\sigma\|_{\boldsymbol{B}_1'} \leq C\|\sigma\|_{\boldsymbol{B}_2'}. \tag{182}$$ contemb03

The role of the *density* of $\boldsymbol{B}^1$ in $(\boldsymbol{B}^2,\, \|\cdot\|^{(2)})$ is to ensure that the restriction mapping is injective. In fact, it is clear, for any two linear functionals σ_1 and σ_2 as above, if they coincide on $\boldsymbol{B}^1$ they are equal, because for any sequence $f_n, n \geq 1$ with $\lim_{n\to\infty}\|f_n - f\|_{\boldsymbol{B}^2}$ one has

$$\sigma_1(f) = \lim_{n\to\infty} \sigma_1(f_n) = \lim_{n\to\infty} \sigma_2(f_n) = \sigma_2(f).$$

□

Remark 11. The example of $\left(\boldsymbol{\ell}^1,\, \|\cdot\|_1\right) \hookrightarrow \left(\boldsymbol{\ell}^2,\, \|\cdot\|_2\right)$, with dual spaces being of the form $\left(\boldsymbol{\ell}^2,\, \|\cdot\|_2\right) \hookrightarrow \left(\ell^\infty,\, \|\cdot\|_\infty\right)$ show that one cannot expect density of $\boldsymbol{B}_2'$ in $\left(\boldsymbol{B}_1', \|\cdot\|_{\boldsymbol{B}_1'}\right)$ *in general.* However, it is true that $\boldsymbol{B}_2'$ is w^*-dense in $\left(\boldsymbol{B}_1', \|\cdot\|_{\boldsymbol{B}_1'}\right)$.

This principle will be applied later by to the situation of $\boldsymbol{S}_0(\mathbb{R}) \hookrightarrow \boldsymbol{W}(\mathbb{R}^d)$, or $\boldsymbol{S}_0(\mathbb{R}^d) \hookrightarrow \boldsymbol{L}^2(\mathbb{R}^d)$, and so on.
By definition $\boldsymbol{L}^1(\mathbb{R}^d)$ is a subalgebra (even a closed ideal) inside the Banach convolution algebra $\boldsymbol{M}_b(\mathbb{R}^d)$. The Fourier transform maps $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ into $\left(\boldsymbol{C}_b(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$ in a non-expansive way. The image of $\left(\boldsymbol{L}^1(\mathbb{R}^d),\, \|\cdot\|_1\right)$ will be denoted by $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ and is called the *Fourier algebra*

FourAlgDef00 **Definition 7.**

$$\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) := \{F = \mathcal{F}(\mu_f) \,|\, f \in \boldsymbol{L}^1(\mathbb{R}^d)\}. \tag{183}$$ FLidef00

Since $\left(\boldsymbol{L}^1(\mathbb{R}^d),\, \|\cdot\|_1\right)$ is a Banach algebra with respect to convolution the Convolution Theorem and the Riemann Lebesgue Theorem imply that $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ is a pointwise algebra, dense in $\left(\boldsymbol{C}_0(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$, because it contains all the linear combinations of triangular functions (for example).

LiFLidens **Lemma 35.** *$\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1$ is a dense subspace of both $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$.*

Proof. We start by observing the triangular function Δ belongs the $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ and to $\boldsymbol{C}_c(\mathbb{R}^d) \subset \boldsymbol{L}^1(\mathbb{R}^d)$, as well as all the dilated and shifted versions, hence all linear combinations.
In order to approximate $f \in \boldsymbol{L}^1(\mathbb{R}^d)$ it is enough to approximate a nearby $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ (in the $\|\cdot\|_1$-sense). But here we know that for the BUPU consisting of triangular functions (for $\mathbb{R}$, and tensor products for $d \geq 2$) we have uniform approximation, i.e.

$$\|\mathrm{Sp}_\Psi k - k\|_\infty \to 0 \quad \text{for } \operatorname{diam}(\Psi) \to 0. \tag{184}$$ splinappr

Since for any $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ the piecewise linear interpolation (i.e. the sum of shifted triangular functions) is a *finite* sum we have achieved the required approximation, noting that due to the joint compact support of the involved functions $\mathrm{Sp}_\Psi k$, $|\Psi| \leq 1$, uniform convergence implies of course $\boldsymbol{L}^1$-convergence. □

We will also see other arguments showing the validity of this claim in a different way. In fact, we will see that $\boldsymbol{S}_0(\mathbb{R}^d)$ and hence also $\boldsymbol{W}(\mathbb{R}^d) \cap \mathcal{F}\boldsymbol{W}(\mathbb{R}^d)$ are dense subspace of $\boldsymbol{L}^1(\mathbb{R}^d)$ and $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$, and in fact even of $(\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1)(\mathbb{R}^d)$ with its natural norm:

$$\|f\|_{\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1} = \|f\|_{\boldsymbol{L}^1} + \|f\|_{\mathcal{F}\boldsymbol{L}^1}.$$

This is a non-trivial result derived in [43]. Also, by the argument used above $(\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1)(\mathbb{R}^d)$ is dense in $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$.

Exercise 7. Check yourself the details of the last statements, mostly by making a list of the partial claims involved in the proof.

The Fourier transform in the context of $(\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1)(\mathbb{R}^d)$ puts things into a symmetric context with respect to the time and frequency variables (on "equal footing"). In fact, both $\boldsymbol{L}^1(\mathbb{R}^d)$ and (hence) $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ are flip-invariant, and since the inverse Fourier transform is more or less the forward Fourier transform combined with the flip operator (before and afterwards) we have:

ptw2conv **Lemma 36.**

- *$(\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1)(\mathbb{R}^d)$ is a Fourier invariant subspace of $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ as well as $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$, and for the norm*

$$\|f\| := \|f\|_{\boldsymbol{L}^1} + \|\widehat{f}\|_{\boldsymbol{L}^1}$$

  the Fourier transform and the flip operator are isometric, i.e. $\|\widehat{f}\| = \|f\| = \|f^\checkmark\|$.

- *The Fourier transform converts pointwise multiplication (within $(\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1)(\mathbb{R}^d)$) to convolution.*

The proof is left to the reader as an exercise.

19.3 The Fourier inversion theorem

Fourinv3 **Lemma 37.** *1. For any $f \in \boldsymbol{L}^1(\mathbb{R}^d)$ and $g \in \boldsymbol{C}_0(\mathbb{R}^d)$ with we have*

$$\int_{\mathbb{R}^d} \mathrm{D}_\rho g(x) f(x) dx \to_{\rho \to 0} \; g(0) \int_{\mathbb{R}^d} f(x) dx. \tag{185}$$ dilconv03

In particular we have $\tau \in \boldsymbol{C}_0(\mathbb{R}^d)$ with $\tau(0) = 1$ and $\phi \in \boldsymbol{L}^1(\mathbb{R}^d)$:

$$\int_{\mathbb{R}^d} [M_t \mathrm{D}_\rho \tau(x)] \phi(x) dx = \int_{\mathbb{R}^d} \mathrm{D}_\rho \tau(x) \chi_t(x) \phi(x) dx \to_{\rho \to 0} \; \int_{\mathbb{R}^d} \phi(x) e^{2\pi i t x} dx. \tag{186}$$ dilconv03b

2. For any $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ and $h \in \boldsymbol{L}^1(\mathbb{R}^d)$ one has for any $t \in \mathbb{R}^d$:

$$\int_{\mathbb{R}^d} [T_t \mathrm{St}_\rho h](x) f(x) dx \to_{\rho \to 0} \; \hat{h}(0) \delta_t(f) = \hat{h}(0) f(t). \tag{187}$$ dilconv04

Proof. (ii) We will derive both statement based on the abstract principle just developed. Let us start with the first one. We will choose $\boldsymbol{B}^1 = \big(\boldsymbol{C}_0(\mathbb{R}^d), \, \|\cdot\|_\infty\big)$, $\boldsymbol{B}^2 = \big(\boldsymbol{L}^1, \, \|\cdot\|_1\big)$ and $(\boldsymbol{B}, \, \|\cdot\|_{\boldsymbol{B}}) = \mathbb{C}$ (with Euclidean distance), and simple duality pairing $M(f,g) = \int_{\mathbb{R}^d} f(x) g(x) dx$, thus with $C = 1$. As a dense subspace we can choose $\boldsymbol{C}_c(\mathbb{R}^d)$ in both cases, i.e. we only have to verify (dilconv03 185) for $f, g \in \boldsymbol{C}_c(\mathbb{R}^d)$.
The assumption (convdens1 163) is then verified, because for $f \in \boldsymbol{C}_c(\mathbb{R}^d)$, with $\mathrm{supp}(f) \subset B_R(0)$ one can use the continuity of $g \in \boldsymbol{C}_0(\mathbb{R}^d)$ at zero in order to guarantee that one has for some suitable chosen $\delta > 0$:

$$|g(0) - g(y)| \leq \varepsilon, \quad \forall |y| \leq \delta.$$

But this means that for $\rho \leq \delta / R$ one has

$$|g(0) - \mathrm{D}_\rho g(z)| = |g(0) - g(\rho z)| \leq \varepsilon, \quad \forall |z| \leq R,$$

whenever $0 < \rho R \leq \delta$, i.e. whenever $\rho \leq \rho_0 := \delta / R$. Hence

$$\left| \int_{\mathbb{R}^d} \mathrm{D}_\rho g(x) f(x) dx - \int_{\mathbb{R}^d} g(0) f(x) dx \right| \leq \varepsilon \int_{\mathbb{R}^d} |f(x)| dx \leq \varepsilon \|f\|_{\boldsymbol{L}^1}. \tag{188}$$ intedilapp1

Formula dilconv03b 186 is just a special case, arising by replacing $g \in \boldsymbol{L}^1(\mathbb{R}^d)$ by $\chi_s g \in \boldsymbol{L}^1(\mathbb{R}^d)$.
The second case is done by similar arguments.
HGFei comment 27.10.: I have made some changes, and will provide more details later on. See LECTURE NOTES!

□

Proof of the Fourier Inversion Theorem

We want to verify, that the Fourier inversion theorem is valid if both $f, \widehat{f} \in \boldsymbol{L}^1(\mathbb{R}^d)$:

$$f(t) = \int_{\mathbb{R}^d} \widehat{f}(s) e^{2\pi i t s} ds, \quad t \in \mathbb{R}^d. \tag{189}$$ Fourinv00

We will apply the *Fundamental Relationship for the Fourier Transform* (FRFT)

$$\int_{\mathbb{R}^d} \hat{g}(x) f(x) dx = \int_{\mathbb{R}^d} g(s) \widehat{f}(s) ds \quad \text{(FRFT)} \tag{190}$$ FRFLiRd

to suitable pairs, consisting of the given continuous and integrable function f and some well chosen functions g, namely $g = M_t \mathrm{D}_\rho \Delta \in \boldsymbol{L}^1(\mathbb{R}^d)$ with

$$\widehat{g} = \mathcal{F}(M_t \mathrm{D}_\rho \Delta) = T_t \mathrm{St}_\rho \mathcal{F}(\Delta) = T_t \mathrm{St}_\rho \, \mathrm{SINC}^2 \in \boldsymbol{L}^1(\mathbb{R}^d).$$

The right hand side of (FRFT) is exactly the left hand side of (186) [dilconv03b] for the choice $\tau = \Delta \in \boldsymbol{C}_0(\mathbb{R}^d)$ with $\tau(0) = 1$ and $\phi = \widehat{f} \in \boldsymbol{L}^1(\mathbb{R}^d)$. Taking the limit for $\rho \to 0$ one obtains as limit the right hand side of (189) [Fourinv00].
The left hand side of (FRFT) becomes in this situation just the left hand side of (187) [dilconv04], with $h = \mathrm{SINC}^2 \in \boldsymbol{L}^1(\mathbb{R}^d)$, and thus converges (for $\rho \to 0$) to $\gamma \cdot f(t)$, with

$$\gamma := \widehat{h}(0) = [\mathcal{F}(\mathrm{SINC}^2)](0) = \| \, \mathrm{SINC}^2 \, \|_{\boldsymbol{L}^1} > 0.$$

Since, by the (FRFT) these two expressions are equal and the limit is justified by Lemma 37 [Fourinv3] on each side (IF $f, \widehat{f} \in \boldsymbol{L}^1$) we arrive at the following formula:

$$\gamma(\mathcal{F}^{-1}(\widehat{f}))(t) = \gamma \cdot f(t) = \int_{\mathbb{R}^d} \widehat{f}(s) e^{2\pi i t s} ds = \int_{\mathbb{R}^d} \widehat{f}(s) e^{-2\pi i t(-s)} ds = \mathcal{F}(\widehat{f})^{\checkmark}(\mathbf{t}) \tag{191}$$ Fourinvgam

whenever $f, \widehat{f} \in \boldsymbol{L}^1(\mathbb{R}^d)$.

Remark 12. This also implies that the condition $\widehat{f} \in \boldsymbol{L}^1(\mathbb{R}^d)$ implies that $f \in \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ (!), justifying previous notations and showing that the inversion theorem holds for $f \in (\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1)(\mathbb{R}^d)$.
It also implies implies that the Fourier transform is injective over $(\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1)(\mathbb{R}^d)$, because (due to the linearity) $\widehat{f} = 0$ implies $f = 0$ (and so continuous functions in $(\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1)(\mathbb{R}^d)$ having the same Fourier transform would have to take the same values).
Moreover we have the very useful formula

$$\mathcal{F}^{-1}(\widehat{f}) = \gamma^{-1} [\mathcal{F}(\widehat{f})]^{\checkmark}, \quad f \in (\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1)(\mathbb{R}^d). \tag{192}$$ IFTform

The observations made so far makes sense also for other normalizations of the Fourier transform, e.g. with the exponential functions appearing without 2π in the exponent (because for them (FRF) is equally valid).
It is now an important consequence of the normalization which we have chosen, to take the term 2π into the definition of the pure frequencies χ_s which allows to find out that $\gamma = 1$, because the classical Gauss function $g_0(x) = exp(-\pi x^2)$ satisfies two important properties (not verified in detail *here*):

$$\mathcal{F}(g_0) = g_0 \quad \text{and} \quad g_0(0) = \widehat{g_0}(0) = \int_{\mathbb{R}^d} g_0(x) dx = 1. \tag{193}$$ Gaussprop0

FTFTflip **Corollary 4.** *For $f \in (\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1)(\mathbb{R}^d)$ one has*

$$\mathcal{F}(\mathcal{F}(f)) = f^{\checkmark}, \quad \textit{and consequently} \quad \mathcal{F}^4(f) = f.$$

Remark 13. We could have used g_0 also instead of the pair Δ, SINC^2 for the proof of the Fourier Inversion Theorem, but we found it more natural to reduce the role of this very specific property of the Gauss function to the determination of the constant γ. We do so also having in mind that such a very special function like the Gaussian does not exist in general LCA groups, while the arguments given up to this point are valid in full generality, if properly adapted.

For details see [85] (p.138,139), using an elementary differential equation which is Fourier invariant. Prop.1.8 (p.140) is the (FRFT) (for Schwartz functions). From this (p.143) the Plancherel Theorem (isometric in the $\boldsymbol{L}^2$-norm) on $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ is developed (and then the heat equation is derived).
A similar path is taken by Folland in his book [48], Chap.7.2 (p. 213), also using the Gauss function. But he does a lot of convolution first.
[Duoandikoetxea, Javier] Fourier Analysis, is another modern book. [18], also using (FRFT) and the Gauss function (Thm.1.13), doing over $\mathbb{R}^n$ right away.
Grafakos does the same in Chap.2, showing the Fourier invariance of the Gauss function $g(x) = exp(-x^2)$, with $\int_{-\infty}^{\infty} g(x)dx = \sqrt{\pi}$. (I guess [57]).

20 Plancherel's Theorem

20.1 Towards Plancherel's Theorem II

The *boxcar-function* $\mathbf{1}_{[-1/2,1/2]}$ has the obvious property that it coincides with its pointwise product.
We also have the convolution theorem : Convolution Theorem convFT00 108 which has implied that $\hat{\Delta} = \mathrm{SINC}^2$. We know that $\Delta(0) = 1$, but we have to determine

$$[\mathcal{F}(\mathrm{SINC}^2)](0) = \int_{-\infty}^{\infty} \mathrm{SINC}^2(x)dx = \mathrm{SINC} \star \mathrm{SINC}(0) =_? \mathrm{SINC}(0) = 1.$$

Note that both Δ and SINC^2 are even and (non-negative) real-valued functions. Since $\mathcal{F}(\Delta) = \mathrm{SINC}^2$ it follows that $\mathcal{F}^{-1}(\mathrm{SINC}^2) = \gamma\Delta$ for some $\gamma > 0$, based on the fundamental relation for the Fourier transform.

20.2 Different Normalizations of the FT

It is remarkable to mention that this (FRF) is also valid for the alternative version of the Fourier transform, we write $\mathcal{F}_{alt}$, given by

$$\mathcal{F}_{alt} f(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt$$

which has different normalizations for Plancherel's Theorem, namely

$$\|f\|_{\boldsymbol{L}^2}^2 = \frac{1}{2\pi} \|\widehat{f}\|_{\boldsymbol{L}^2}^2,$$

or another normalization of the inverse Fourier transform (see [90]), namely

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \widehat{f}(\omega) e^{-it\omega} d\omega.$$

Finally it is not valid anymore that convolution and multiplication just change their role. In fact, pointwise multiplication goes into convolution, again up to a scalar factor of the form 2π.

$$\mathcal{F}_{alt}(f \cdot g) = \frac{1}{2\pi} (\mathcal{F}_{alt} f \star \mathcal{F}_{alt} g).$$

Finally (Thm.7.7) in [90] the (Shannon) Sampling Theorem takes a less convenient form. For general band-limited $f \in \boldsymbol{L}^1(\mathbb{R}^d)$ takes the form, for a function $f \in \boldsymbol{L}^1(\mathbb{R}^d)$ with $\widehat{f}(\omega) = 0$ for $|\omega| \geq c$:

$$f(t) = \sum_{n \in \mathbb{Z}} f\left(\frac{n\pi}{c}\right) \frac{\sin(ct - n\pi)}{(ct - n\pi)}, \tag{194}$$ VretSamp1

with (claimed only) *uniform convergence*. In fact it is uniform and absolute (in $\big(\boldsymbol{C}_0(\mathbb{R}), \|\cdot\|_\infty\big)$), because effectively $f \in \boldsymbol{S}_0(\mathbb{R})$ in this case. Thus the Nyquist rate for $[-1/2, 1/2]$ is 2π, or sampling at $\mathbb{Z}$ ($c = \pi$) allows bandwidth π.

20.3 The advantage of having 2π in the exponent

Let us try to find out what the value of γ is!
The usual way found in all the books I have checked is via the Gauss function $g_0(t) = exp(-\pi|t|^2), t \in \mathbb{R}^d$, namely the two properties of the Gauss functions which play a role here

$$\mathcal{F}(g_0) = g_0, \quad g_0(0) = 1, \quad \int_{\mathbb{R}^d} g_0(s) ds = 1. \tag{195}$$ GauProps00

Then we have the perfectly symmetric situation. Plancherel's Theorem is becoming an *isometry*, the inverse Fourier transform (for $(\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1)(\mathbb{R}^d)$) is just the forward Fourier transform combined with the flip operator (corresponding to the transition from χ_{-s} to χ_s, and **pointwise multiplication and convolution simply change their roles.**
Of course one can (finding formulas in different books) transfer results about the Fourier transform as used in these notes to the alternative one or vice versa combining the integral transforms with dilation operators. The starting point could be the obvious transition formula

$$\mathcal{F} = \mathrm{D}_{2\pi} \circ \mathcal{F}_{alt}, \quad \text{respectively} \quad \mathcal{F}_{alt} = \mathrm{D}_{1/2\pi} \circ \mathcal{F}. \tag{196}$$ FTtoFTalt

Finally let us mention that some authors have the normalization of pure frequencies without the factor 2π, but prefer to have the inversion formula and the forward Fourier transform to be more similar. This implies to use (citation from [10], p.164)

$$\widehat{f}(v) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-iuv} f(u) du, \tag{197}$$ ButzNesFT

which gives the symmetric version of the Fourier transform, and also the isometric property for the Plancherel's transform, but still the Nyquist rate in the Shannon Sampling

Theorem does not have the same nice normalization. Also now/then both the convolution theorem AND the conversion of pointwise products to convolutions would have to bear the factor $\sqrt{2\pi}$, which is not nice.
Let us also mention here that Butzer/Nessel do the Fourier Stieltjes transform for $\mathbb{R}$ in Chap.4.3 and 5.3 of [10], p.219:

$$\check{\mu}(v) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ivu} d\mu(u). \tag{198}$$ FStBUNE71

They also do Poisson's formula (p.188,201,232,239), but the word SAMPLING does NOT yet appear in this book, hence the persistent problem with the Nyquist rate would still show up in that context. Also the convolution theorem does not show up in the book [10]. There was never a multi-dimensional version of that book!
Of course the original reason for choosing $exp(it)$ is Euler's formula:
$e^{it} = \cos(t) + i\sin(t)$, which *looks* more convenient.

20.4 Convolution operators on the Fourier algebra

By the Riemann-Lebesgue Theorem we know that the Fourier algebra $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ with the natural norm $\|\widehat{f}\|_{\mathcal{F}\boldsymbol{L}^1} = \|f\|_{\boldsymbol{L}^1}, f \in \boldsymbol{L}^1(\mathbb{R}^d)$ is a dense subalgebra of $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$. But it is also (by transfer of structure) a Banach space, even a Banach algebra.
In fact, any Cauchy sequence $(\widehat{f}_n)_{n\geq 1}$ in $\big(\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_{\mathcal{F}\boldsymbol{L}^1}\big)$ corresponds to a Cauchy sequence in $\big(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\big)$, namely $(f_n)_{n\geq 1}$. By the density we may replace this given sequence by an equivalent one with elements in the dense subspace $(\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1)(\mathbb{R}^d)$. Thus we may assume without loss of generality that the Fourier inversion theorem establishes a connection between f_n and $\widehat{f}_n$.
By the completeness of $\big(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\big)$ this sequence has a limit, let us call it f_0, i.e. $\lim_{n\to\infty}\|f_n - f_0\|_{\boldsymbol{L}^1} = 0$, but this is fully equivalent (thanks to the Fourier inversion theorem) to the claim that $\lim_{n\to\infty}\|f_n - f_0\|_{\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)} = 0$.
The submultiplicativity of the norm follows also in a trivial way, using the convolution theorem (now in the $\boldsymbol{L}^1(\mathbb{R}^d)$ setting):

$$\|\widehat{f} \cdot \widehat{g}\|_{\mathcal{F}\boldsymbol{L}^1} = \|f \star g\|_{\boldsymbol{L}^1} \leq \|f\|_{\boldsymbol{L}^1}\|g\|_{\boldsymbol{L}^1} = \|\widehat{f}\|_{\mathcal{F}\boldsymbol{L}^1}\,\|\widehat{g}\|_{\mathcal{F}\boldsymbol{L}^1}\,. \tag{199}$$ FLiBanalg1

Thus $\boldsymbol{M}_b(\mathbb{R}^d) \star \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) \subset \boldsymbol{M}_b(\mathbb{R}^d) \star \boldsymbol{C}_0(\mathbb{R}^d) \subseteq \boldsymbol{C}_0(\mathbb{R}^d)$, i.e. the action of convolution by bounded measures is well defined, but, does it also preserve the $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$-structure? The answer is provided by the following proposition.

MbconvFLi **Proposition 8.** *The Banach space* $\big(\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_{\mathcal{F}\boldsymbol{L}^1}\big)$ *is a homogeneous Banach space, and any convolution operator* C_μ *with* $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ *leaves* $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ *invariant, i.e. we have*

$$\boldsymbol{M}_b(\mathbb{R}^d) \star \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) \subset \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d), \quad \textit{hence } \boldsymbol{L}^1(\mathbb{R}^d) \star \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) \subseteq \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d), \tag{200}$$ MbconvFLi1

with the corresponding norm estimate

$$\|\mu \star h\|_{\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)} \leq \|\mu\|_{\boldsymbol{M}_b}\|h\|_{\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)}\,, \quad \mu \in \boldsymbol{M}_b(\mathbb{R}^d), h \in \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d). \tag{201}$$ MbconvFLiest

Let us look at the dense subspace $(\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1)(\mathbb{R}^d)$ of $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ (and of $\boldsymbol{C}_0(\mathbb{R}^d)$). Then we know that the convolution acts boundedly with respect to the $\boldsymbol{L}^1$-norm or the sup-norm, but it is also bounded with respect to the norm in the Fourier algebra, for simple reasons (and then extended to all of $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$):

$$\|\mu \star \widehat{f}\|_{\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)} = \|\widehat{\mu} \cdot f\|_{\boldsymbol{L}^1} \leq \|\widehat{\mu}\|_\infty \|f\|_{\boldsymbol{L}^1} \leq \|\mu\|_{\boldsymbol{M}_b} \|\widehat{f}\|_{\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)}. \tag{202}$$

MbconfFliest1

20.5 Towards the Fourier Plancherel Theorem

We want to derive from this that the Fourier transform (at least defined on $\boldsymbol{C}_c(\mathbb{R}^d)$) is isometric in the sense of the $\boldsymbol{L}^2(\mathbb{R}^d)$-norm, given by

$$\|f\|_{\boldsymbol{L}^2(\mathbb{R}^d)} := \left(\int_{\mathbb{R}^d} |f(x)|^2 dx\right)^{1/2},$$

or expressed by the natural scalar product

$$\langle f, g \rangle_{\boldsymbol{L}^2} := \int_{\mathbb{R}^d} f(x)\overline{g(x)}dx, \tag{203}$$

scalproddef

by setting as usual

$$\|f\|_{\boldsymbol{L}^2(\mathbb{R}^d)} := \sqrt{\langle f, f \rangle_{\boldsymbol{L}^2}}. \tag{204}$$

Hilbnorm

It is a consequence of the fact that

$$(f, g) \mapsto \langle f, g \rangle_{\boldsymbol{L}^2}$$

is a sesquilinear form, i.e. linear in the first variable and anti-linear in the second one[17], that one has the **Cauchy Schwartz inequality**

$$|\langle f, g \rangle_{\boldsymbol{L}^2}| \leq \|f\|_{\boldsymbol{L}^2} \|g\|_{\boldsymbol{L}^2}, \tag{205}$$

CSinequal

which in turn can be used to derive the triangle inequality for the $\boldsymbol{L}^2$-norm:

$$\|f + g\|_{\boldsymbol{L}^2} \leq \|f\|_{\boldsymbol{L}^2} + \|g\|_{\boldsymbol{L}^2}. \tag{206}$$

triangLtsp

Note let us choose $\widehat{g} = \overline{h} \in (\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1)(\mathbb{R}^d)$, so that the left hand side of (FRF) gives just $\langle f, h \rangle_{\boldsymbol{L}^2}$. Then we have

$$g = \mathcal{F}^{-1}(\overline{h}) = [\mathcal{F}(\overline{h})]^{\checkmark} = [(\widehat{h})^*]^{\checkmark} = \overline{\widehat{h}}, \tag{207}$$

LtisomFTsubs1

or in other words, the right hand side turns into

$$\int_{\mathbb{R}^d} \widehat{f}(s)\overline{\widehat{h}(s)}ds = \langle \widehat{f}, \widehat{h} \rangle_{\boldsymbol{L}^2}. \tag{208}$$

LtisomFT1

Setting now $f = h$ one obtains the important isometric property of the Fourier transform (with respect to the $\boldsymbol{L}^2(\mathbb{R}^d)$-norm), namely

$$\|f\|_{\boldsymbol{L}^2(\mathbb{R}^d)} = \|\widehat{f}\|_{\boldsymbol{L}^2(\mathbb{R}^d)}, \quad \forall f \in (\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1)(\mathbb{R}^d). \tag{209}$$

LtisomFT

we arrive at the fact that the Fourier transform is also unitary (i.e. isometric in the $\boldsymbol{L}^2(\mathbb{R})$-sense, by setting $\hat{g}(t) = \overline{h(t)}$).
Later on we will show that (new by 29.10.)

$$\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) = \boldsymbol{L}^2(\mathbb{R}^d) \star \boldsymbol{L}^2(\mathbb{R}^d) = \{f_1 \star f_2, f_1, f_2 \in \boldsymbol{L}^2(\mathbb{R}^d)\},$$

which originates from the idea that for $f \geq 0$ the function $\sqrt{f}$ belongs to $\boldsymbol{L}^2(\mathbb{R}^d)$.

[17] Physicists would choose the opposite convention.

21 Plancherel's Theorem

21.1 The energy preservation property of $\mathcal{F}$

One of the most important consequences of the Fourier inversion theorem, which establishes (among others) the symmetry between forward and inverse Fourier transform, thus contributing to the viewpoint, that the "time"-domain and the "frequency"-domain are equivalent.
While the restriction to $(\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1)(\mathbb{R}^d)$ ensure the existence of the involved integrals (as Lebesgue integrals) it does not even help us to find that $\mathcal{F}^{-1}(\mathrm{SINC}) = \mathbf{1}_{[-1/2,1/2]}$, because $\mathrm{SINC} \notin \boldsymbol{L}^1(\mathbb{R})$! We will see that Plancherel's Theorem, which extends the Fourier transform to $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$ (establishing an isometric automorphism!) allows us to come back from $\mathrm{SINC} \in \boldsymbol{L}^2(\mathbb{R})$ to $\mathbf{1}_{[-1/2,1/2]}$. Of course, this still does not allow us to invert the Fourier Stieltjes transform, not even to recover every $f \in \boldsymbol{L}^1(\mathbb{R}^d)$ from $\widehat{f}$, because $\boldsymbol{L}^1(\mathbb{R}^d) \cap \boldsymbol{L}^2(\mathbb{R}^d)$ is a proper subspace of $\boldsymbol{L}^1(\mathbb{R}^d)$. In the extreme case (but most naturally) we need to find that $\mathcal{F}^{-1}(\mathbf{1}) = \delta_0$, as the converse to the obvious relationship $\mathcal{F}(\delta_0) = \mathbf{1}$, which appears in engineering books often as the (questionable) formula

$$\mathcal{F}^{-1}(\mathbf{1})(t) = \int_{-\infty}^{\infty} \mathbf{1}(s) e^{2\pi i s t} ds = \int_{-\infty}^{\infty} e^{2\pi i s t} ds =_{???} \delta_0(t)??$$

We will see in due course in which sense such a symbolic manipulation is justified!
The step from the Fourier inversion theorem (with $\gamma = 1$) to Plancherel's theorem is based on the following result:

LtIsometr **Proposition 9.** *For $f \in (\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1)(\mathbb{R}^d)$ one has*

$$\langle f, \overline{f} \rangle_{\boldsymbol{L}^2} = \int_{\mathbb{R}^d} |f(x)|^2 dx = \int_{\mathbb{R}^d} |\widehat{f}(s)|^2 ds. \tag{210}$$ LTismetr1

Proof. The obvious strategy to go back to the (FRF) for $\boldsymbol{L}^1(\mathbb{R}^d)$, which reads:

$$\int_{\mathbb{R}^d} f(t)\widehat{g}(t) dt = \int_{\mathbb{R}^d} g(s)\widehat{f}(s) ds, \quad \text{for all} \quad f, g \in \boldsymbol{L}^1(\mathbb{R}^d). \tag{211}$$ L1FourL1

Now let us choose $\hat{g} = \overline{h} \in (\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1)(\mathbb{R}^d)$, so that the left hand side of (FRF) gives just $\langle f, h \rangle_{\boldsymbol{L}^2}$. Using the algebraic properties of $\mathcal{F}$ we obtain:

$$g = \mathcal{F}^{-1}(\overline{h}) = [\mathcal{F}(\overline{h})]^{\checkmark} = [(\widehat{h})^*]^{\checkmark} = \overline{\widehat{h}}, \tag{212}$$ LtisomFTsubs

or in other words, the right hand side turns into

$$\int_{\mathbb{R}^d} \widehat{f}(s)\overline{\widehat{h}(s)} ds = \langle \widehat{f}, \hat{h} \rangle_{\boldsymbol{L}^2}. \tag{213}$$ LtisomFT2

Setting now $f = h$ and taking square roots on both sides one obtains the important isometric property of the Fourier transform (with respect to the $\boldsymbol{L}^2(\mathbb{R}^d)$-norm), namely

$$\|f\|_{\boldsymbol{L}^2(\mathbb{R}^d)} = \|\widehat{f}\|_{\boldsymbol{L}^2(\mathbb{R}^d)}, \quad \forall f \in (\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1)(\mathbb{R}^d). \tag{214}$$ LtisomFT

$\square$

The full version of Plancherel's Theorem can be derived from this isometric property, using the density of $(\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1)(\mathbb{R}^d)$ in $\left(\boldsymbol{L}^2(\mathbb{R}^d),\, \|\cdot\|_2\right)$. We will provide the technical details later[18] on.

Plancherel00 **Theorem 7** (Plancherel's Theorem)**.** *The isometric (forward and inverse) Fourier transform on $(\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1)(\mathbb{R}^d)$ extends to a unitary automorphism of $\left(\boldsymbol{L}^2(\mathbb{R}^d),\, \|\cdot\|_2\right)$. This means that it defines a bijective mapping from $\left(\boldsymbol{L}^2(\mathbb{R}^d),\, \|\cdot\|_2\right)$ into itself, which preserves scalar products and in particular is isometric. Thus we have*

$$\langle f, g\rangle_{\boldsymbol{L}^2} = \langle \widehat{f}, \widehat{g}\rangle_{\boldsymbol{L}^2}, \quad f, g \in \boldsymbol{L}^2(\mathbb{R}^d). \tag{215}$$ PlanchScalP

Proof. First one extends the forward (and the inverse) mappings in natural way to all of $\boldsymbol{L}^2$ *ASSUMING FOR NOW* the density of $(\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1)(\mathbb{R}^d)$ in $\boldsymbol{L}^2(\mathbb{R}^d)$, to all of $\left(\boldsymbol{L}^2(\mathbb{R}^d),\, \|\cdot\|_2\right)$. Then we make use that both sides, namely the sesquilinear forms defining the scalar product (in the $\boldsymbol{L}^2$-sense) on the time or the frequency side are continuous, and therefore the transition can be made based on the bilinear mapping proposition described earlier. $\square$

Remark 14. This remarkable statement allows, among others, to give a meaning to the (correct) claim that $\mathcal{F}^{-1}(\mathrm{SINC}) = \mathbf{1}_{[-1/2,1/2]}$.
It also implies that two (say band-limited) functions $f, g \in \boldsymbol{L}^2(\mathbb{R}^d)$ with disjoint spectrum, i.e. with $\mathrm{supp}(\widehat{f}) \cap \mathrm{supp}(\widehat{g}) = \emptyset$ are orthogonal, because

$$\langle f, g\rangle_{\boldsymbol{L}^2} = \int_{\mathbb{R}^d} \widehat{f}(s)\overline{\widehat{g}(s)}dx = 0.$$

[18] In the current context the role of $\left(\boldsymbol{L}^2(\mathbb{R}^d),\, \|\cdot\|_2\right)$ as a Banach space of (classes of) measurable functions will be a minor one, compared to the existence of a scalar product on any of the decent subspaces, such as $(\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1)(\mathbb{R}^d)$ or later $\boldsymbol{S}_0(\mathbb{R}^d)$, which is preserved by the (forward and inverse) Fourier transform as explained.

22 Wiener and Feichtinger algebra: 25.10.2020

This is an attempt to derive certain properties of Wiener amalgam spaces which will turn out to be extremely useful for the rest of these notes. The general Wiener amalgams are denoted by $\boldsymbol{W}(\boldsymbol{L}^p, \boldsymbol{\ell}^q)(\mathbb{R}^d)$, or more generally $\boldsymbol{W}(\boldsymbol{B}, \boldsymbol{C})$, with a "local component" $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ and a *global component* $(\boldsymbol{C}, \|\cdot\|_{\boldsymbol{C}})$ and they consist of those (generalized) functions which belong locally to $\boldsymbol{B}$, and whose global behavior of the $\boldsymbol{B}$-norm is described by $\boldsymbol{C}$.
Instead of giving a formal definition for the general case here, let us take restrict our attention first to those space which can be already formed now, with a global $\boldsymbol{\ell}^1$-behavior (or equivalently, as we will see, a global $\boldsymbol{L}^1$-behaviour). There are quite different ways to describe such amalgams, using so-called *continuous* norms (then the symbol $\boldsymbol{W}(\boldsymbol{L}^p, \boldsymbol{L}^q)(\mathbb{R}^d)$ is justified), or the *discrete* version of the norm, which makes use of BUPUs. There are also *atomic characterizations* of these spaces and so on.
For us the following situation will be useful (unifying the properties of $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$ and $\left(\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_{\mathcal{F}\boldsymbol{L}^1}\right)$):

homFuncAlg **Definition 8.** A Banach algebra of continuous functions (with respect to pointwise multiplication) is called a *homogeneous Banach algebra* of functions if one has:

1. $\boldsymbol{A} \cap \boldsymbol{C}_c(\mathbb{R}^d) \hookrightarrow (\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}}) \hookrightarrow \left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$ as dense subspaces;
2. Translation is isometric on $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$, i.e. $\|T_x h\|_{\boldsymbol{A}} = \|h\|_{\boldsymbol{A}}, h \in \boldsymbol{A}, x \in \mathbb{R}^d$;
3. Translation is continuous on $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$, i.e.
$$\lim_{x\to 0} \|T_x h - h\|_{\boldsymbol{A}} = 0, \quad \forall h \in \boldsymbol{A}.$$
4. $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ has bounded approximate units; (?used?)
5. There exists a regular BUPU $(\psi_k)_{k\in\mathbb{Z}^d}$ in $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$, i.e. some $\psi \in \boldsymbol{A} \cap \boldsymbol{C}_c(\mathbb{R}^d)$ such that
$$\sum_{k\in\mathbb{Z}^d} \psi_k(x) = \sum_{k\in\mathbb{Z}^d} \psi(x-k) \equiv 1. \qquad (216) \quad \text{BUPUinAa}$$

As a consequence of 2. and 3. above one has $\boldsymbol{M}_b(\mathbb{R}^d) \star \boldsymbol{A} \subset \boldsymbol{A}$, we will need

$$\boldsymbol{L}^1(\mathbb{R}^d) \star \boldsymbol{A} \subseteq \boldsymbol{A}, \quad \text{with } \|g \star f\|_{\boldsymbol{A}} \le \|g\|_{\boldsymbol{L}^1} \|f\|_{\boldsymbol{A}}, \quad g \in \boldsymbol{L}^1(\mathbb{R}^d), f \in \boldsymbol{A}. \qquad (217) \quad \text{LiConvMod1}$$

For reasons of formal reference let us state the following lemma.

FLiprops **Lemma 38.** $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ *is homogeneous Banach algebra in the sense of definition* (homFuncAlg 8).

Proof. We will not go into details here. Just let us recall, that $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ is a commutative Banach algebra with respect to convolution. The Dirac families $(\mathrm{St}_\rho g)_{\rho\to 0}$ (for $g \in \boldsymbol{L}^1(\mathbb{R}^d)$ with $\hat{g}(0) = 1$) form bounded approximate units. The existence of band-limited approximate units implies that $\boldsymbol{C}_c(\mathbb{R}^d) \cap \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ is a dense subspace (even a dense ideal) within $\left(\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_{\mathcal{F}\boldsymbol{L}^1}\right)$.

Remark 15. As we have seen the mapping $k \to \int_{\mathbb{R}^d} k(x)dx$ is continuous on $\boldsymbol{C}_c(\mathbb{R}^d)$, endowed with the $\boldsymbol{L}^1$-norm (which coincides with the norm induced from $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ via the identification of k with μ_k). Hence using this "extended integral" one could describe the Fourier algebra simple as the image of $\boldsymbol{L}^1(\mathbb{R}^d)$ under the usual integral transform. This is, by the way, the approach taken whenever the Lebesgue integral is available. The argument then relies on the fact that $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ (now in the Lebesgue sense!) is a Banach space with respect to the $\|\cdot\|_1$, containing $\boldsymbol{C}_c(\mathbb{R}^d)$ as a dense subspace.

We have seen that $\Delta = \mathcal{F}(\mathrm{SINC}^2) \in \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$, hence we can have (many different) BUPUs in $\left(\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_{\mathcal{F}\boldsymbol{L}^1}\right)$ (e.g. $\psi = \Delta \otimes .. \otimes \Delta$ in the multi-dimensional case).
It is obvious that one has

$$\|T_s\widehat{f}\|_{\mathcal{F}\boldsymbol{L}^1} = \|M_s f\|_{\boldsymbol{L}^1} = \|f\|_{\boldsymbol{L}^1} = \|\widehat{f}\|_{\mathcal{F}\boldsymbol{L}^1}, \quad f \in \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d). \tag{218}$$ FLitrans1

It is an easy exercise to verify that $s_n \to s_0$ implies uniform convergence of the characters χ_{s_n} to χ_{s_0}, i.e. that for any $R > 0$ and $\varepsilon > 0$ there exists $\delta > 0$ such that

$$|s_n - s_0| \le \delta \quad \Rightarrow \quad |\chi_{s_n}(x) - \chi_{s_0}(x)| = |\chi_{s_n - s_0}(x) - 1| \le \varepsilon, \quad \forall |x| \le R,$$

because

$$exp(2\pi i\langle s_n - s_0, x\rangle) = exp(iy)$$

for some $y \in \mathbb{R}^d$ with

$$|y| \le \|s_n - s_0\|\|x\| \le \delta R.$$

Since the exponential function is continuous and satisfies $e^0 = 1$ we can assure that for $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ one has $\|M_s k - k\|_{\boldsymbol{L}^1} \to 0$ for $s \to 0$. By approximation the same is true for $f \in \boldsymbol{L}^1(\mathbb{R}^d)$, and this implies (by the formula $T_s\widehat{f} = \mathcal{F}(M_s f)$) that one has continuous shift in the Fourier algebra. □

Remark 16. Observe that property (LiConvMod1 217), stating that $\boldsymbol{L}^1(\mathbb{R}^d) \star \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) \subseteq \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$, has been already been verified in Proposition MbconvFLi 8.

We are going to introduce $\boldsymbol{W}(\boldsymbol{A},\boldsymbol{\ell}^1)$ resp. $\left(\boldsymbol{W}(\boldsymbol{A},\boldsymbol{\ell}^1), \|\cdot\|_{\boldsymbol{W}(\boldsymbol{A},\boldsymbol{\ell}^1)}\right)$ in the following way. The two prototypical examples (which we want to describe in a unified way) are given by $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}}) = \left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$ or $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}}) = \left(\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_{\mathcal{F}\boldsymbol{L}^1}\right)$. The resulting space will be called Wiener algebra, denoted by $\boldsymbol{W}(\boldsymbol{C}_0,\boldsymbol{\ell}^1)(\mathbb{R}^d)$, and $\boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1,\boldsymbol{\ell}^1)(\mathbb{R}^d)$ for the second case, also known as *modulation space* $(\boldsymbol{M}^1(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}^1})$ or as **Segal algebra** $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$, also called Feichtinger algebra in the literature.

WAlidef00 **Definition 9.** Let $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ be a homogeneous Banach algebra as defined in (8). Then we set homFuncAlg

$$\boldsymbol{W}(\boldsymbol{A},\boldsymbol{\ell}^1)(\mathbb{R}^d) := \left\{ f \in \boldsymbol{A} \,|\, \sum_{k\in\mathbb{Z}^d} \|f\psi_k\|_{\boldsymbol{A}} < \infty \right\}. \tag{219}$$ WAlidef

It is clear that the proposed expression actually defines a norm on $\boldsymbol{W}(\boldsymbol{A},\boldsymbol{\ell}^1)(\mathbb{R}^d)$ and that the embedding of $\left(\boldsymbol{W}(\boldsymbol{A},\boldsymbol{\ell}^1), \|\cdot\|_{\boldsymbol{W}(\boldsymbol{A},\boldsymbol{\ell}^1)}\right)$ into $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ is a continuous one. In the proof of the various statements we will make use of the following technical result (which also implies the equivalence of the norms introduced by different BUPUs).
As a preparation for the relevant estimates let us provide an alternative, so-called *atomic characterization* of $\boldsymbol{W}(\boldsymbol{A},\boldsymbol{L}^1)(\mathbb{R}^d)$.

WAliatomic **Proposition 10.** *For any fixed $R > 0$ the following characterization can be given: A function $h \in \boldsymbol{A}$ belongs to $\boldsymbol{W}(\boldsymbol{A},\boldsymbol{\ell}^1)$ if and only if h can be written as an absolutely convergent sum of so-called atoms: there exists a sequence $(h_k)_{k=1}^\infty$ with $\sum_{k=1}^\infty \|h_k\|_{\boldsymbol{A}} < \infty$ and $\mathrm{supp}(h_k) \subseteq B_R(x_k)$ for some sequence of points $(x_k)_{k=1}^\infty$. Any such representation of $h \in \boldsymbol{A}$ is called an admissible atomic representation of $h = \sum_{k=1}^\infty h_k$.*

Proof. First of all it is clear that any $f \in \boldsymbol{W}(\boldsymbol{A},\boldsymbol{\ell}^1)(\mathbb{R}^d)$ has such a representation. Just recall that the BUPU is formed from a function $\psi \in \boldsymbol{C}_c(\mathbb{R}^d)$, hence $\mathrm{supp}(\psi) \subseteq B_r(0)$ for some $r > 0$. This implies that the absolutely convergent series representation $f = \sum_{k\in\mathbb{Z}^d} f\psi_k$ satisfies the required side condition, namely $\mathrm{supp}(f\psi_k) \subset B_r(k)$.
For the converse we observe first that it is enough to find a *uniform estimate for all the atoms* h_k, because the rest can obtained by summation over the atoms:

$$\|h\|_{\boldsymbol{W}(\boldsymbol{A},\boldsymbol{\ell}^1)} \leq \sum_{k=1}^\infty \|h_k\|_{\boldsymbol{W}(\boldsymbol{A},\boldsymbol{\ell}^1)}\,.$$

In preparation for such a uniform estimate let us observe that for any radius $R > 0$ the number of non-zero terms $k \in \mathbb{Z}^d$ such that

$$B_R(x) \cap \mathrm{supp}(\psi_k) \subseteq B_R(x) \cap B_r(k) \neq \emptyset$$

is uniformly bounded. In fact the intersection is nonzero (if and) only if $k \in B_{R+r}(x)$. A simple estimate for the number of non-zero terms can be obtained as follows: for $\gamma = 1/2$ the open balls $B_\gamma(k)$ are disjoint. If $k \in B_{R+r}(x)$ then of course $B_\gamma(k) \subset B_{R+r+1}$. Thus the total area of all these small, disjoint balls, all having the same fixed volume, namely $\alpha := \mathrm{Vol}(B_\gamma(0))$. Write $\beta = \mathrm{Vol}(B_{R+r+1}(0)) = \mathrm{Vol}(B_{R+r+1}(x))$ for any $x \in \mathbb{R}^d$. Hence the count of points involved allows the following estimate:

$$\#\{k \,|\, k \in B_{R+r}(x)\} \leq \beta/\alpha.$$

The situation can be illustrated (for $d = 2$) by Fig./Abb. 22. The yellow ball represents the ball $B_{R+r}(0)$, while the larger ball/circle of area α is represented by the blue circle surrounding the yellow one.

Remark 17. The short verbal description of the situation is the following one: Given $R > 0$ there exists some constant $C_R > 0$ such that for any $h \in \boldsymbol{A}$ with $\mathrm{supp}(h) \subseteq B_R(z)$ for some $z \in \mathbb{R}^d$ one has: $h \in \boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)(\mathbb{R}^d)$ and

$$\|h\|_{\boldsymbol{W}(\boldsymbol{A},\boldsymbol{\ell}^1)(\mathbb{R}^d)} \leq C_R \|h\|_{\boldsymbol{A}}.$$

coverballs1.jpg

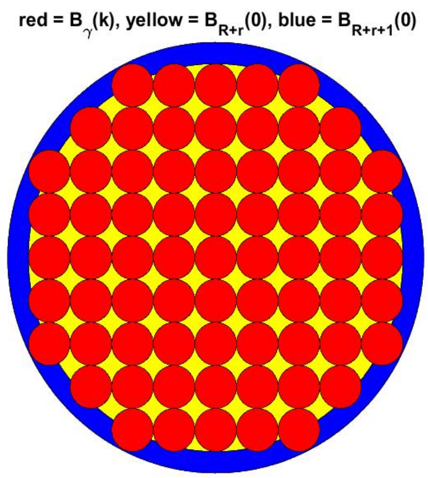

Abbildung 36: coverballs1.jpg

Recalling that $\|\psi_k\|_{\boldsymbol{A}} = \|\psi\|_{\boldsymbol{A}}$ for any $k \in \mathbb{Z}^d$ and the fact that $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ is a (pointwise) Banach algebra, we can estimate the $\boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)$-norm of such an atom via

$$\|h_k\|_{\boldsymbol{W}(\boldsymbol{A},\boldsymbol{\ell}^1)} \leq \sum_{j:\psi_j \cdot h_k \neq 0} \|h_k \psi_j\|_{\boldsymbol{A}} \leq (\beta/\alpha)\|\psi\|_{\boldsymbol{A}} \|h_k\|_{\boldsymbol{A}}. \tag{220}$$

WAliEstim1

$$\Rightarrow \quad \|h\|_{\boldsymbol{W}(\boldsymbol{A},\boldsymbol{\ell}^1)} \leq \sum_{k=1}^{\infty} \|h_k\|_{\boldsymbol{W}(\boldsymbol{A},\boldsymbol{\ell}^1)} \leq [\beta/\alpha]\|\psi\|_{\boldsymbol{A}} \sum_{k=1}^{\infty} \|h_k\|_{\boldsymbol{A}}.$$

□

This result immediately implies

WAliequiv **Corollary 5.** *Two BUPUs Φ and Φ define the same subspace ($\boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)(\mathbb{R}^d)$) of $\boldsymbol{A}$ with equivalent norms.*

The proof also reveals that the natural *atomic norm*, namely the expression

$$\|h\|_{atom} := \inf \left\{ \sum_{k=1}^{\infty} \|h_k\|_{\boldsymbol{A}} \;\middle|\; h = \sum_{k=1}^{\infty} h_k \text{ admissible} \right\} \tag{221}$$

atomicnorms

the infimum taken over all *admissible representation*, defines an equivalent norm on $\boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)(\mathbb{R}^d)$, with the equivalence constants only depending on ψ (resp. Ψ) and the chosen Radius R.

WAliprops1 **Proposition 11.** *The space $\boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)$, endowed with the natural norm*

$$\|f\|_{\boldsymbol{W}(\boldsymbol{A},\boldsymbol{\ell}^1)} := \sum_{k\in\mathbb{Z}^d} \|f\psi_k\|_{\boldsymbol{A}} < \infty \tag{222}$$ WAlinormdef

has the following properties:

1. $\Big(\boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1), \|\cdot\|_{\boldsymbol{W}(\boldsymbol{A},\boldsymbol{\ell}^1)}\Big)$ *is a Banach space, continuously embedded into* $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ *and also into* $\Big(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\Big)$.
2. $\boldsymbol{A} \cap \boldsymbol{C}_c(\mathbb{R}^d) \hookrightarrow \Big(\boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1), \|\cdot\|_{\boldsymbol{W}(\boldsymbol{A},\boldsymbol{\ell}^1)}\Big)$ *as a dense subspace;*
3. *The family of translation operators is uniformly bounded;*
4. *Translation is continuous in* $\Big(\boldsymbol{W}(\boldsymbol{A}, \boldsymbol{L}^1), \|\cdot\|_{\boldsymbol{W}(\boldsymbol{A},\boldsymbol{L}^1)}\Big)$, *i.e.*
$$\lim_{x\to 0} \|T_x f - f\|_{\boldsymbol{W}(\boldsymbol{A},\boldsymbol{\ell}^1)(\mathbb{R}^d)} = 0, \quad f \in \boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)(\mathbb{R}^d).$$
5. $\boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)(\mathbb{R}^d)$ *is also a Banach algebra with respect to pointwise multiplication.*
6. *If* $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ *is dilation invariant (either for* D_ρ *or* St_ρ*, for all* $\rho > 0$*), then the same is true for* $\boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)(\mathbb{R}^d)$.
7. *The same is true for affine transformation of the argument, i.e. for* $f \to \gamma^*(f)$, *with* $\gamma^* f(x) = f(\gamma(x))$ *for* $\gamma(x) = \boldsymbol{A} \star \mathbf{x}$, *some non-singular* $d \times d$*-matrix* $\boldsymbol{A}$.

Proof. By the assumption of Def. homFuncAlg 8 we have for the atoms

$$\|h^n\|_{\boldsymbol{L}^1} \le C_R \|h^n\|_\infty \le C_R C_A \|h^n\|_{\boldsymbol{A}},$$

hence absolute convergence of the series in the Banach space $\Big(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\Big)$, as well as a continuous embedding of $\boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)(\mathbb{R}^d)$ into $\Big(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\Big)$.
For the completeness one has to show that absolutely convergent series with $\sum_{n=1}^\infty h^n$ in $\Big(\boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{W}(\boldsymbol{A},\boldsymbol{\ell}^1)(\mathbb{R}^d)}\Big)$ are convergent. Since the condition

$$\sum_{n=1}^\infty \|h^n\|_{\boldsymbol{W}(\boldsymbol{A},\boldsymbol{\ell}^1)} < \infty$$

implies absolute convergence in $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$, $h = \sum_{n=1}^\infty h^n$ is well defined in $\boldsymbol{A}$, and

$$h\psi_k = \sum_{n=1}^\infty h^n \psi_k, \quad k \in \mathbb{Z}^d,$$

and (due the allowed permutation of summation) one has

$$\sum_{k\in\mathbb{Z}^d} \|h\psi_k\|_{\boldsymbol{A}} = \sum_{n=1}^\infty \sum_{k\in\mathbb{Z}^d} \|h^n\psi_k\|_{\boldsymbol{A}} < \infty,$$

i.e. $h \in \boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)(\mathbb{R}^d)$ and $\|h\|_{\boldsymbol{W}(\boldsymbol{A},\boldsymbol{\ell}^1)} \le \sum_{n=1}^\infty \|h^n\|_{\boldsymbol{W}(\boldsymbol{A},\boldsymbol{\ell}^1)(\mathbb{R}^d)} < \infty$.

lgauBUP1.eps

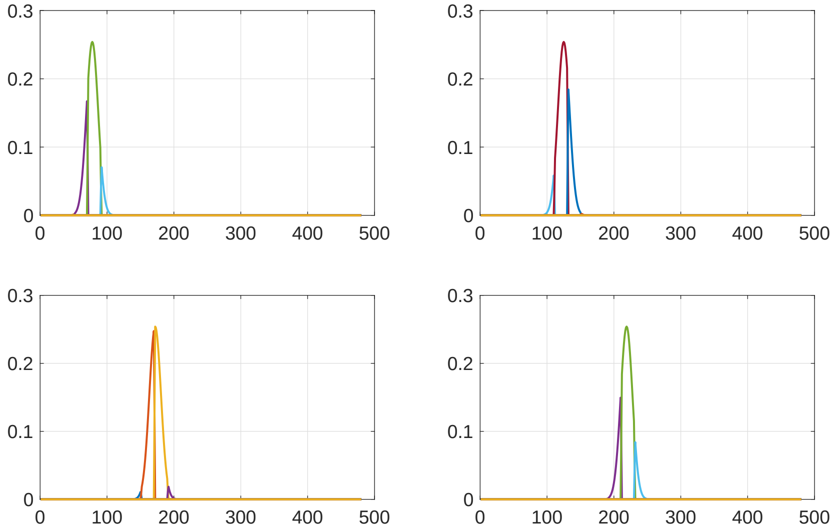

Abbildung 37: trlgauBUP1.eps:
This picture show a few typical situations, who a shifted version of a BUMP function is split into different pieces which are then measured separately. Depending on the configuration the resulting norm, computed by means of absolute sums of the norms of the different pieces (e.g. in the sup-norm or in the Fourier algebra) is not exactly the same for different shift parameters applied to the function whose norm in an amalgam space $\boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)(\mathbb{R})$ is measured.

22.1 Lack of isometric translation

In order to verify the uniform boundedness of the translation operators one just have to observe that the atomic norm is strictly isometric. In fact, for any atomic decomposition of $f \in \boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$ the function $T_x f$ has an equivalent representation, just with the support of the atoms being shifted (but preserving the norm exactly).
As for the density of the compactly supported elements it is also clear from the atomic characterization. For any finite partial sum of atoms the reminder (which has an arbitrary small total sum of norms) will have a corresponding atomic representation, hence with decreasing (atomic) norm.
For finite partial sums translation is continuous, again based on the atomic decomposition. Assume that $\operatorname{supp} h_k \subseteq B_r(x_k)$ implies for $|y| \leq 1$ that $T_y h_k \subseteq B_{r+1}(x_k)$. Thus the differences $(T_y h_n - h_n)$ constitute an atomic representation with respect to an atomic decomposition of size $R = r + 1$.
Finally, also the dilation invariance is easily obtained from the atomic representation. The dilated atoms will be supported on balls of a new radius $B_{\rho R}(z_k)$ for suitable centers z_k, $k \geq 1$, and thus for each fixed parameter $\rho > 0$ the boundedness of the dilation operator can be verified.[19]
The continuous embedding into $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ follows from the fact that finite partial sums belong to $\boldsymbol{A} \cap \boldsymbol{C}_c(\mathbb{R}^d)$, and thus to $\boldsymbol{L}^1(\mathbb{R}^d)$. Due to the finite size of the support

[19] Since there is a change of geometry the operator norm may be influenced by this change of diameters.

ntransl1.eps

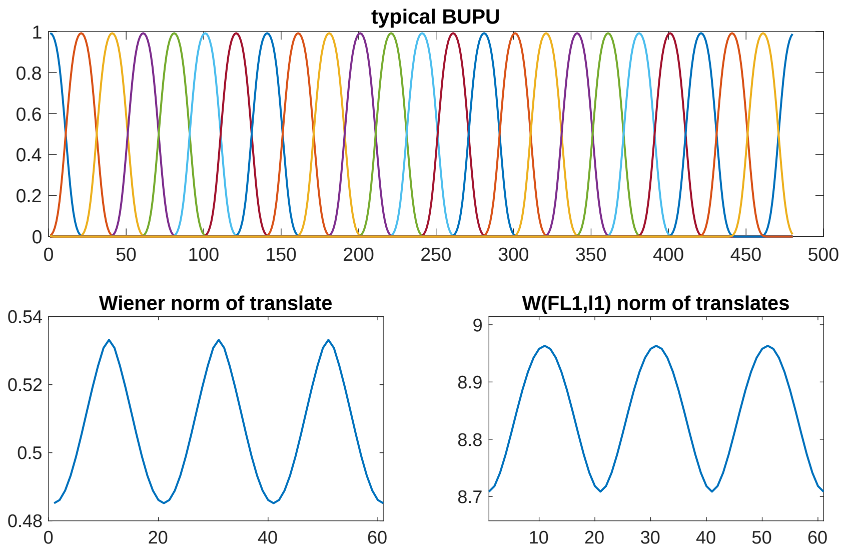

Abbildung 38: Wnormtransl1.eps:
Wiener amalgam norms of a collection of shifted copies of a standard Gaussian, showing that the norm is NOT strictly translation invariant, even for a decent bump functions. For other case with several bumps involved at different positions the behaviour may depend even more strongly on the BUPU used and the actual translation parameter. However, there are uniform bounds from above and below!

one can estimate the $\boldsymbol{L}^1$-norm (which is computable via Riemann integrals) as

$$\|h\|_{\boldsymbol{L}^1} \leq \sum_{k\in\mathbb{Z}^d} \|f\psi_k\|_{\boldsymbol{L}^1} \leq \sum_{k\in\mathbb{Z}^d} \|f\psi_k\|_\infty \mathrm{Vol}(B_r(k)) \leq C \sum_{k\in\mathbb{Z}^d} \|f\psi_k\|_{\boldsymbol{A}}$$

if $\mathrm{supp}(\psi) \subset B_r(0)$, since we have $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}}) \hookrightarrow \big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$.
Finally let us observe that the pointwise multiplicative structure is easy to verify, because clearly the product of two atoms can be viewed as an atom itself, since $\mathrm{supp}(h_1h_2) \subseteq \mathrm{supp}(h_1)\,\mathrm{supp}(h_2)$. $\square$

MbWali **Proposition 12.** *There exists $C > 0$ (depending only on the BUPU etc.) such that:*

$$\boldsymbol{M}_b(\mathbb{R}^d) \star \boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1) \subseteq \boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1), \quad \textit{with} \quad \|\mu \star f\|_{\boldsymbol{W}(\boldsymbol{A},\boldsymbol{\ell}^1)} \leq C\|\mu\|_{\boldsymbol{M}_b}\|f\|_{\boldsymbol{W}(\boldsymbol{A},\boldsymbol{\ell}^1)}. \quad (223)$$ MbWAliest

The proof of these results goes back to the techniques already developed in **fe83:** [24], and uses the fact that $\boldsymbol{M}_b(\mathbb{R}^d) = \boldsymbol{W}(\boldsymbol{M}_b, \boldsymbol{\ell}^1)$ according to Thm. measabsdecomp. (Thm. 4 of current version of Script, 25.10.2020).

Proof. The proof is essentially a combination of the fact (see Prop. homFuncAlg 8) that

$$\boldsymbol{M}_b(\mathbb{R}^d) \star \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) \subseteq \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$$

and the fact that both $\boldsymbol{M}_b(\mathbb{R}^d)$ and $\boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)$ can be described via absolutely convergent sums of localized pieces. Since the compactly supported elements are dense in both space it is enough to find an estimate under the assumption that $\mu = \sum_{i\in F} \mu\,\psi_i$ and $f = \sum_{j\in E} f\psi_j$ in $\boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)(\mathbb{R}^d)$.

We have for some fixed regular BUPU Ψ:

$$\|\mu\|_{\boldsymbol{M}_b} = \sum_{j\in\mathbb{Z}^d} \|\mu\psi_j\|_{\boldsymbol{M}_b} \quad \text{and} \quad \|f\|_{\boldsymbol{W}(\boldsymbol{A},\boldsymbol{\ell}^1)} = \sum_{k=1}^{\infty} \|f\psi_k\|_{\boldsymbol{A}}$$

(actually a finite sum). Hence (cf. (111) above!)

convMbWAli
$$\mu \star f = \sum_{i\in F}\sum_{j\in E}[(\mu\psi_i)\star(f\psi_j)]. \tag{224}$$

By relabelling the finite index sets of pairs by the natural numbers $l = 1, ..L$ and writing $h_k = (\mu\psi_i)\star(f\psi_j)$ for the corresponding term, we have the required norm estimate

$$\|h_l\|_{\boldsymbol{A}} \leq \|\mu\psi_i\|_{\boldsymbol{M}_b}\|f\psi_j\|_{\boldsymbol{A}}, \quad 1 \leq l \leq L$$

and furthermore

$$\mathrm{supp}(h_l) \subseteq B_r(i) + B_r(j) \subseteq B_{2r}(i+j).$$

Combining these observations allows to estimate the atomic norm of $\mu\star f$ in $\boldsymbol{W}(\boldsymbol{A},\boldsymbol{\ell}^1)(\mathbb{R}^d)$, thus completing the proof (by invoking Prop. 10).

DiracWAli **Lemma 39.** *Assume that one has for any $f \in \boldsymbol{A}$ and $g \in \boldsymbol{L}^1(\mathbb{R}^d)$ with $\hat{g}(0) = 1$*

DiracA1
$$\|\mathrm{St}_\rho g \star f - f\|_{\boldsymbol{A}} \to 0 \text{ for } \rho \to 0, \tag{225}$$

then the same relationship also holds for $\boldsymbol{W}(\boldsymbol{A},\boldsymbol{\ell}^1)(\mathbb{R}^d)$, i.e.

StrhoWAli
$$\lim_{\rho\to 0} \|\mathrm{St}_\rho g \star f - f\|_{\boldsymbol{W}(\boldsymbol{A},\boldsymbol{\ell}^1)} = 0. \tag{226}$$

In particular, the band-limited elements are dense in $\boldsymbol{W}(\boldsymbol{A},\boldsymbol{\ell}^1)(\mathbb{R}^d)$.

WduSOP1 **Lemma 40.** *Since $\boldsymbol{S}_0(\mathbb{R}^d)$ is a dense subspace of $\boldsymbol{W}(\mathbb{R}^d) = \big(\boldsymbol{W}(\boldsymbol{C}_0,\boldsymbol{\ell}^1)(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{W}}\big)$, the dual spaces are continuously embedded into each other, i.e. for some $C > 0$:*

WduSOP2
$$\|\mu\|_{\boldsymbol{S}_0'} \leq C\|\mu\|_{\boldsymbol{W}(\mathbb{R}^d)^*}, \quad \mu \in \boldsymbol{W}(\mathbb{R}^d)^*. \tag{227}$$

Proof. Due to the atomic characterization of $\boldsymbol{W}(\mathbb{R}^d) = \boldsymbol{W}(\boldsymbol{C}_0,\boldsymbol{\ell}^1)(\mathbb{R}^d)$ (cf. Lemma 10) it is enough to approximate atoms by compactly supported elements of $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$. In fact, using a sufficiently fine regular BUPU with generator $\psi \in \mathcal{F}\boldsymbol{L}^1 \cap \boldsymbol{C}_c(\mathbb{R}^d) \subset \boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1,\boldsymbol{\ell}^1)(\mathbb{R}^d)$ we obtain a good (uniform) approximation of the form $\mathrm{Sp}_\Psi f$, i.e. by a finite sum of terms from $\boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1,\boldsymbol{\ell}^1)(\mathbb{R}^d)$.
As a consequence of the dense embedding the abstract principle Lemma 34 is applicable and provides a proof for the claim. □

In general, the fact that the embedding of $\boldsymbol{S}_0(\mathbb{R}^d)$ in $\boldsymbol{W}(\mathbb{R}^d)$ is a proper one, implies that also the dual spaces are properly embedded into each other, i.e. there exists $\sigma \in \boldsymbol{S}_0'(\mathbb{R}^d)$ which cannot be (boundedly) extended to all of $\big(\boldsymbol{W}(\boldsymbol{C}_0,\boldsymbol{\ell}^1)(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{W}}\big)$. However, if $\sigma \in \boldsymbol{S}_0'(\mathbb{R}^d)$ is *non-negative*, i.e. if $f \geq 0 \Rightarrow \sigma(f) \geq 0$ then such an extension if always possible.

SOPposit **Lemma 41.** *Let σ be a non-negative, bounded linear functional on $\boldsymbol{S}_0(\mathbb{R}^d)$. Then is extends in a unique way to a linear functional on $\left(\boldsymbol{W}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{W}}\right)$.*

Proof. Given the positivity (resp. non-negativity) of σ it will be enough to show that there exists some $C_0 > 0$ such that for any $h \in \boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d)$ some $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ with $\|f\|_{\boldsymbol{S}_0(\mathbb{R}^d)} \leq C_0\|h\|_{\boldsymbol{W}(\mathbb{R}^d)}$.
In fact, we can start from any fixed function $\phi \in \mathcal{F}\boldsymbol{L}^1 \cap \boldsymbol{C}_c(\mathbb{R}^d)$ with $\phi(y) = 1$ over $B_r(0)$, and thus $T_x\phi(z) = 1$ over $B_r(x)$.
Since $\operatorname{supp}(f\psi_k) \subseteq B_r(k), k \in \mathbb{Z}^d$ we have

$$|f\psi_k(x)| \leq \|f\psi_k\|_\infty T_k\phi(x), \quad x \in \mathbb{R}^d.$$

Using using the positivity of σ, applied to the non-negative function

$$h(x) := \sum_{k\in\mathbb{Z}^d} \|f\psi_k\|_\infty |T_k\phi(x)| - |f(x)| \geq 0$$

one has (using the continuity of σ on $\left(\boldsymbol{W}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{W}}\right)$):

$$|\sigma(f)| \leq \sum_{k\in\mathbb{Z}^d} \|f\psi_k\|_\infty |\sigma(T_k\phi)| \leq [\|\sigma\|_{\boldsymbol{S}_0'} \|\phi\|_{\boldsymbol{S}_0}]\|f\|_{\boldsymbol{W}(\boldsymbol{C}_0,\boldsymbol{\ell}^1)(\mathbb{R}^d)}.$$

□

In order to verify the additional action of the Dirac sequence $\mathrm{St}_\rho g$, with $\rho \to 0$, let us recall that - in the same way as for $\boldsymbol{L}^1(\mathbb{R}^d)$, one first can verify the result for $g \in \boldsymbol{C}_c(\mathbb{R}^d)$. Since it is compressed one can assume that $\operatorname{supp}(\mathrm{St}_\rho g) \subset B_1(0)$ for $\rho < \rho_0$. Hence for the dense subspace of finite sums of atoms the terms $\mathrm{St}_\rho g \star h_n$ are also supported in balls $B_{r+1}(x_k)$. Since the we have (by assumption) convergence in $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ the result is clear in this case (convergence in the atomic norm).
The general case follows there by the same principle as in the proof for $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$. Since there exist functions g in $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ which are band-limited and satisfy $\hat{g}(0) = 1$ (e.g. $g = \mathrm{SINC}^2$, with $\hat{g} = \Delta$) the band-limited approximation can be achieved by convolution products of the form $\mathrm{St}_\rho \star f$ (whose Fourier transform is contained in $\rho^{-1}\operatorname{supp}(\hat{g})$). □

The two most important special cases of this construction appear by the choice $\boldsymbol{A} = \boldsymbol{C}_0(\mathbb{R}^d)$ respectively $\boldsymbol{A} = \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$, both with their natural norms.
The resulting space $\left(\boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{W}}\right)$ is called **Wiener's algebra**, and appears prominently in H. Reiter's book [78] (and hence in [79]), while the space $\boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1, \boldsymbol{\ell}^1)(\mathbb{R}^d)$ has been introduced as the "New Segal Algebra" $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ in [23] by the author, and has been baptised **Feichtinger algebra** in [79].
Wiener's algebra is the largest among the spaces $\boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)(\mathbb{R}^d)$, because we have assumed from the very beginning that $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}}) \hookrightarrow \left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$.
Any of the space $\boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)(\mathbb{R}^d)$ is a *Segal algebra*, i.e. a dense Banach ideal inside of $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ (with respect to convolution), but also a pointwise Banach algebra (see [78] for the definition of a Segal algebra).

Both of them are invariant under dilations and (by choosing $\rho = -1$) under the flip operator as well as the other involutions discussed so far $f \to \overline{f}$ and $f \to f^*$.

$\boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)$ is not contained in $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$, and hence cannot be Fourier invariant, but one may consider $\boldsymbol{W}(\mathbb{R}^d) \cap \mathcal{F}\boldsymbol{W}(\mathbb{R}^d)$. We will see later that this is a natural domain for *Poisson's summation formula*. Moreover, one can derive the validity of the Fourier inversion formula in the pointwise sense using only Riemann integrals, but we avoid the more technical discussion of this question here.
Compared to the spaces considered so far, namely $\boldsymbol{C}_0(\mathbb{R}^d)$ or $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$, which cannot be embedded into their dual spaces, the new spaces, as any Banach space $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}}) \hookrightarrow \boldsymbol{L}^1 \cap \boldsymbol{C}_0(\mathbb{R}^d)$ (with natural norm $\|f\| := \|f\|_{\boldsymbol{L}^1} + \|f\|_\infty$) are naturally embedded into their dual space, in the usual way of turning nice functions into functionals:

$$\mu_f(h) = \int_{\mathbb{R}^d} h(x) f(x) dx, \quad f, h \in \boldsymbol{B}. \tag{228}$$ BembBdual

In fact, if one has $\|f\|_{\boldsymbol{L}^1} \leq C_1 \|f\|_{\boldsymbol{B}}$ and $\|f\|_\infty \leq C_2 \|f\|_{\boldsymbol{B}}, \forall f \in \boldsymbol{B}$, then

$$|\mu_f(h)| \leq \int_{\mathbb{R}^d} |h(x)||f(x)| dx \leq \|h\|_\infty \|f\|_{\boldsymbol{L}^1} \leq [C_1 C_2 \|f\|_{\boldsymbol{B}}] \|h\|_{\boldsymbol{B}}.$$

or

$$\|\mu_f\|_{\boldsymbol{B}'} \leq C_1 C_2 \|f\|_{\boldsymbol{B}},$$

showing that the embedding $f \to \mu_f$ (the injectivity of this mapping is obvious!) is a continuous mapping from $\boldsymbol{B}$ to $(\boldsymbol{B}', \|\cdot\|_{\boldsymbol{B}'})$.

22.2 The Fourier Invariance of $\boldsymbol{S}_0(\mathbb{R}^d)$

The next and very important result is the Fourier invariance of $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ (resp. $\boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1, \boldsymbol{\ell}^1)(\mathbb{R}^d)$).
As a preparation let us note that by the Fourier inversion theorem (and the fact that Δ is an even function) we have seen that $\Delta \in \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$, and since it has compact support it also belongs to $\boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1, \boldsymbol{\ell}^1)(\mathbb{R}^d)$ (only finitely many terms of the given BUPU will intersect in a non-trivial way with $\mathrm{supp}(\Delta)$.) This is the argument why $\boldsymbol{A} \cap \boldsymbol{C}_c(\mathbb{R}^d) \subset \boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)(\mathbb{R}^d)$.

S0Fourinv **Theorem 8.** *The space $\boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1, \boldsymbol{\ell}^1)(\mathbb{R}^d)$ is invariant under the Fourier transform. Hence it is continuously embedded into $\boldsymbol{W}(\mathbb{R}^d) \cap \mathcal{F}\boldsymbol{W}(\mathbb{R}^d)$.*

Proof. Since we have seen that $\boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1, \boldsymbol{\ell}^1)(\mathbb{R}^d)$ is continuously embedded into $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ it is clear that it is also embedded into the Fourier invariant subspace of all those continuous functions which belong to $\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1$.
What we need for the proof of the theorem (using the atomic characterization, with atoms concentrated in balls $B_R(x_k)$, $k \geq 1$), is a collection of plateau functions $p \in \mathcal{F}(\boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1, \boldsymbol{\ell}^1)(\mathbb{R}^d))$ with $p(x) = 1$ on $\mathrm{supp}(\psi)$.
This is easily achieved, using a few properties of $\boldsymbol{S}_0(\mathbb{R}^d) := \boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1, \boldsymbol{\ell}^1)(\mathbb{R}^d)$. First we recall that there are non-trivial functions $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ with $Q = \mathrm{supp}(\widehat{f})$ compact. Taking any band-limited function in $\boldsymbol{S}_0(\mathbb{R}^d)$ we can convolve it with $f^* \in \boldsymbol{S}_0(\mathbb{R}^d)$, and

thus obtain the function $\widehat{f}\hat{f}^* = |\widehat{f}|^2$, which is compactly supported and non-negative (of course non-zero), in $\mathcal{F}(\boldsymbol{S}_0(\mathbb{R}^d))$. By dilation we may assume that there also bump functions of this kind in $\mathcal{F}(\boldsymbol{S}_0(\mathbb{R}^d))$ with arbitrary small support.
Since $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) \cdot \boldsymbol{S}_0(\mathbb{R}^d) \subset \boldsymbol{S}_0(\mathbb{R}^d)$ we also have

$$\boldsymbol{L}^1(\mathbb{R}^d) \star \mathcal{F}(\boldsymbol{S}_0(\mathbb{R}^d)) \subset \mathcal{F}(\boldsymbol{S}_0(\mathbb{R}^d)),$$

which in turn implies that we can convolve these bump functions with indicator functions $\mathbf{1}_Q$ of sufficiently large sets, so that we can create within $\mathcal{F}(\boldsymbol{S}_0(\mathbb{R}^d)) = \mathcal{F}^{-1}(\boldsymbol{S}_0(\mathbb{R}^d))$ also plateau functions p such that $p(x)\psi(x) = \psi(x), x \in \mathbb{R}^d$.
We will need the fact that $\mathcal{F}(p) \in \boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1, \boldsymbol{\ell}^1)(\mathbb{R}^d)$, or equivalently $p \in \mathcal{F}(\boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1, \boldsymbol{\ell}^1)(\mathbb{R}^d))$ (because the space is flip-invariant). In fact, we have

$$f\psi_k = f\psi_k \cdot T_k p, \quad k \in \mathbb{Z}^d. \tag{229}$$ platrepro1

By consequence we have for any $k \geq 1$:

$$\mathcal{F}(f\psi_k) = \mathcal{F}(f\psi_k) \star \mathcal{F}(p), \quad k \in \mathbb{Z}^d \tag{230}$$ platFTrepr1

$$\|\mathcal{F}(f\psi_k)\|_{\boldsymbol{S}_0(\mathbb{R}^d)} \leq \|\mathcal{F}(f\psi_k)\|_{\boldsymbol{L}^1} \|\mathcal{F}(p)\|_{\boldsymbol{S}_0(\mathbb{R}^d)} = \|\mathcal{F}(p)\|_{\boldsymbol{S}_0(\mathbb{R}^d)} \cdot \|f\psi_k\|_{\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)}$$

and adding up over the pieces

$$\|\widehat{f}\|_{\boldsymbol{S}_0} \leq \sum_{k\in\mathbb{Z}^d} \|\mathcal{F}(f\psi_k)\|_{\boldsymbol{S}_0(\mathbb{R}^d)} \leq \sum_{k\in\mathbb{Z}^d} \|f\psi_k\|_{\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)} \|\mathcal{F}(p)\|_{\boldsymbol{S}_0(\mathbb{R}^d)} = \|\mathcal{F}(p)\|_{\boldsymbol{S}_0(\mathbb{R}^d)} \|f\|_{\boldsymbol{S}_0},$$

demonstrating that the Fourier transform (and consequently the inverse Fourier transform) is a bounded automorphism of $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$. □

SOPdef2 **Definition 10.** We just define the dual space $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$ as the Banach space of all **mild distributions**. The extended Fourier transform can be defined on $\boldsymbol{S}_0'(\mathbb{R}^d)$ via

$$\widehat{\sigma}(f) = \sigma(\widehat{f}), \quad f \in \boldsymbol{S}_0(\mathbb{R}^d), \sigma \in \boldsymbol{S}_0'(\mathbb{R}^d). \tag{231}$$ SOPFTdef02

Let us recall the of course finite discrete measures are in the dual of $\boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)(\mathbb{R}^d)$ (for any choice of $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$), but also so-called **Dirac combs** *over lattices* $\Lambda = \boldsymbol{A} \star \mathbb{Z}^d$, for some non-singular $d \times d$-matrix $\boldsymbol{A}$.
In fact, we have seen for any ball $B_R(x)$ (independent of $x \in \mathbb{R}^d$) the number of points of a lattice is limited. Let us write

$$C_R := \#\{k \in \mathbb{Z}^d, k \in B_R(x_k)\}.$$

We can refer to Fig. 22 coverballs1.jpg (for the case $\Lambda = \mathbb{Z}^d$, for the purpose of illustration). The general situation is described by the picture presented in Fig. 22.2 periodOmega2a.pdf.

DcombSOP **Lemma 42.** *For any $\Lambda \lhd \mathbb{R}^d$ the Dirac comb $\delta_\Lambda := ⊔⊔_\Lambda := \sum_{\lambda\in\Lambda} \delta_\lambda$ belongs to $\boldsymbol{W}(\mathbb{R}^d)^*$, the dual of Wiener's algebra $\left(\boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{W}}\right)$, hence the linear mapping*

$$f \mapsto \sum_{\lambda\in\Lambda} f(\lambda) \tag{232}$$ Dcombdef

defines a mild distribution, i.e.$⊔⊔_\Lambda \in \boldsymbol{S}_0'$.

dOmega2a.pdf

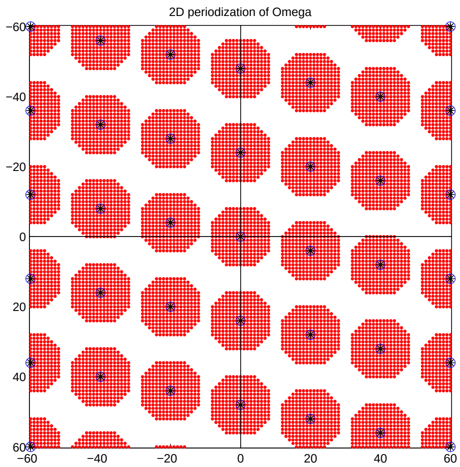

Abbildung 39: periodOmega2a.pdf: disjoint subsets of the fundamental domain along some lattice Λ.

Proof. It will be enough to observe that we have for any atom $h_k \in \boldsymbol{A} \cap \boldsymbol{C}_c(\mathbb{R}^d)$:

$$\left| \sum_{\lambda \in \Lambda} \delta_\lambda(h_k) \right| \leq \sum_{\lambda \in B_R(x_k)} |h_k(\lambda)| \leq C_R \|h_k\|_\infty \leq C_R C_2 \|h_k\|_{\boldsymbol{A}},$$

hence we have for $f = \sum_{k=1}^\infty h_k$:

$$\left| \sum_{\lambda \in \Lambda} \delta_\lambda(h_k) \right| \leq C_R C_2 \sum_{k=1}^\infty \|h_k\|_{\boldsymbol{A}},$$

and then by summing over all atoms, we have for $C_4 > 0$ (involving the equivalence constants comparing different norms)

$$\left| \sum_{\lambda \in \Lambda} \delta_\lambda(h_k) \right| \leq C_4 \|f\|_{atom}.$$

$\square$

oxshift.eps

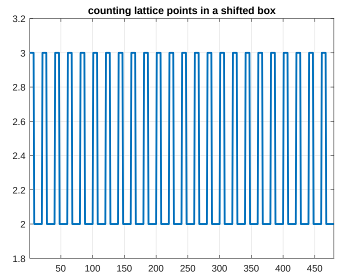

Abbildung 40: countboxshift.eps:
Counting the number of lattice points of the form $\alpha\mathbb{Z}$ inside of a given interval of fixed length, as a function of the initial position, e.g. $\#\{j \,|\, \alpha j \in [x-R, x+r]\}, r \in \mathbb{Z}$.

Pworld1.pdf

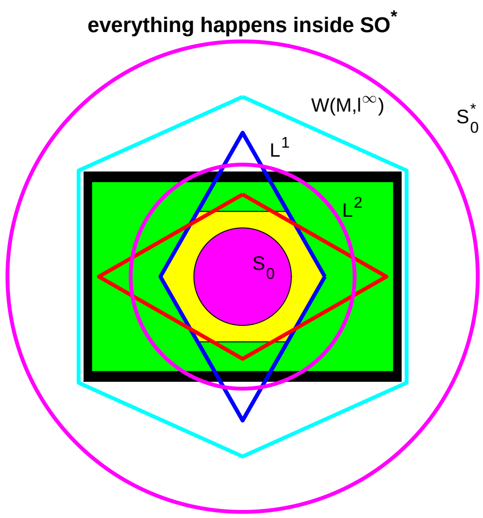

Abbildung 41: SOSPworld1.pdf: The universe inside of $\boldsymbol{S}_0'(\mathbb{R}^d)$:
yellow: Wiener's algebra, NOT contained in $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$. Light blue hexagon to the outside: dual space: $\boldsymbol{W}(\boldsymbol{M}, \boldsymbol{\ell}^\infty)(\mathbb{R}^d)$. Locally comparable to $\boldsymbol{M}_b(\mathbb{R}^d)$, globally like $\boldsymbol{\ell}^\infty$. Everything happens inside of $\boldsymbol{S}_0'(\mathbb{R}^d)$ (magenta big circle).

22.3 Equivalent norms on Wiener amalgams

The use of BUPUs $\Psi = (\psi_i)_{i \in I}$ in $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ is the most convenient way to define Wiener amalgam spaces, because they deliver a so-called *atomic decompositions* $f = \sum_{i \in I} f \psi_i$ right away, i.e. a linear procedure to obtain an atomic representation from the given signal. But in order to measure the membership of a given function in $\boldsymbol{A}$ (in our case) other simple methods can be employed. For example, it would be enough to control the expression $\sum_{k \in \mathbb{Z}^d} \|f T_k \phi\|_{\boldsymbol{A}}$ for some $\phi \in \boldsymbol{A} \cap \boldsymbol{C}_c(\mathbb{R}^d)$ as long as there is some $h \in \boldsymbol{A}$ and some $\psi \in \boldsymbol{A}$ forming a BUPU $\Psi = (\psi_i)_{i \in I}$. In fact, we have

$$\sum_{k \in \mathbb{Z}^d} \|f \psi_k\|_{\boldsymbol{A}} = \sum_{k \in \mathbb{Z}^d} \|(f T_k \phi) \cdot (T_k h)\|_{\boldsymbol{A}} \le \|h\|_{\boldsymbol{A}} \sum_{k \in \mathbb{Z}^d} \|(f \cdot T_k \phi)\|_{\boldsymbol{A}} < \infty. \qquad (233)$$ controlBUPU1

Conversely, the assumed compactness of the support of ϕ implies that

$$\phi = \sum_{l \in F} \phi \cdot \psi_l$$

and consequently, by the same argument as above, with changing roles, one has:

$$\sum_{j \in \mathbb{Z}^d} \|f \cdot T_j \phi \cdot T_j \left(\sum_{l \in F} \psi_l \right) \|_{\boldsymbol{A}} \leq \|\phi\|_{\boldsymbol{A}} \sum_{l \in F} \sum_{j \in \mathbb{Z}^d} \|f \cdot T_{j+l} \psi\|_{\boldsymbol{A}} \leq \#(F) \cdot \|f\|_{\boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)} \cdot$$

Another alternative for the norm, which is overcoming the difficulty of lack of strict (meaning isometric) translation invariance of the norm for a fixed BUPU (or the alternative problem of obtaining atomic decompositions in a linear way), is given by the use of the so-called *continuous control function (CCF)*.
Instead of measuring f by moving the function ψ along some lattice ($\mathbb{Z}^d$ or more general does not matter) one defines this continuous control function by moving ψ (in fact any non-zero bump function $\phi \in \boldsymbol{A} \cap \boldsymbol{C}_c(\mathbb{R}^d)$) continuously, and measuring constantly. Let us provide the definition of such CCFs for the general case of a *localizable* space $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$.

$$\kappa(f)(z) = \kappa(f, \phi, \boldsymbol{B})(z) := \|f \cdot T_z \phi\|_{\boldsymbol{B}}. \tag{234}$$ CCSdef01

In our setting (A **detailed proof** will be given in Lemma CSnorm2 72 below) the alternative characterization can be formulated as follows (details will be provided in section DiscContS0norm 45 below).

CCSnorm1 **Lemma 43.** *For each non-zero $\phi \in \boldsymbol{A} \cap \boldsymbol{C}_c(\mathbb{R}^d)$ the expression*[20]

$$\|f \mid \boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)\|_{\phi, cont} := \int_{\mathbb{R}^d} \kappa(f, \phi, \boldsymbol{A})(z) dz. \tag{235}$$ CCSnorm

defines an equivalent norm on $\boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)(\mathbb{R}^d)$, and $f \in \boldsymbol{A}$ belongs to $\boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)(\mathbb{R}^d)$ if and only if this expression is finite.

Remark 18. The fact that the continuous norm is applied to κ is the reason why these space are also denoted was $\boldsymbol{W}(\boldsymbol{A}, \boldsymbol{L}^1)(\mathbb{R}^d)$. However, since there are no inclusion relations between e.g. $\boldsymbol{L}^1(\mathbb{R}^d)$ and $\boldsymbol{L}^2(\mathbb{R}^d)$ and only the global properties of $\boldsymbol{L}^p$-spaces play a role in the definition of amalgam spaces, we prefer to use the discrete components which are natural to use. They also give an intuitive clue concerning (dense) embeddings, e.g. $\boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1) \hookrightarrow \boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^2)$.

Note: The space of the form $\boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)(\mathbb{R}^d)$ have been introduced as minimal invariant spaces in [22]. The key paper concerning the Segal algebra $\left(\boldsymbol{S}_0(G), \|\cdot\|_{\boldsymbol{S}_0}\right)$ is [23], see also [65] for a summary of some 13 characterizations of $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$.
Turning this abstract principle for $\boldsymbol{A} = \mathcal{F}\boldsymbol{L}^1$ into a concrete condition one has for any band-limited function $g \in \boldsymbol{L}^1(\mathbb{R}^d)$ (this $\phi := \widehat{g}^{\checkmark} \in \boldsymbol{A} \cap \boldsymbol{C}_c(\mathbb{R}^d)$): The following expression is another equivalent norm on $\boldsymbol{S}_0(\mathbb{R}^d)$, seen as a subspace of $(\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1)(\mathbb{R}^d)$:

$$\|f\|_{g,1,1} = \int_{\mathbb{R}^d} \|M_s g \star \widehat{f}\|_{\boldsymbol{L}^1} \, ds < \infty, \tag{236}$$ S0modul11FT

[20] The integral appearing here applies to a non-negative, continuous functions, and can be understood in the Riemannian sense!

because $\|T_s\phi \cdot f\|_{\mathcal{F}\boldsymbol{L}^1} = \|M_s g \star \widehat{f}\|_{\boldsymbol{L}^1}$. But we know already that $\boldsymbol{S}_0(\mathbb{R}^d)$ is Fourier invariant, hence one can equally well use the following norm:

$$\|f\|_{g,1,1} = \int_{\mathbb{R}^d} \|M_s g \star f\|_{\boldsymbol{L}^1}\, ds < \infty. \tag{237}$$ S0modul11

We will see later that the following is a norm equivalent to the $\boldsymbol{L}^2$-norm:

$$\|f\|_{g,2,2} = \left(\int_{\mathbb{R}^d} \|M_s g \star f\|_{\boldsymbol{L}^2}^2\, ds\right)^{1/2} < \infty. \tag{238}$$ S0modul22

More general norms (with $p, 1 \in [1, \infty]$) define modulation spaces, see [27], or [25]. For more details and proofs see Lemma CCSnorm2 ?? below.

JPUGaus1.jpg

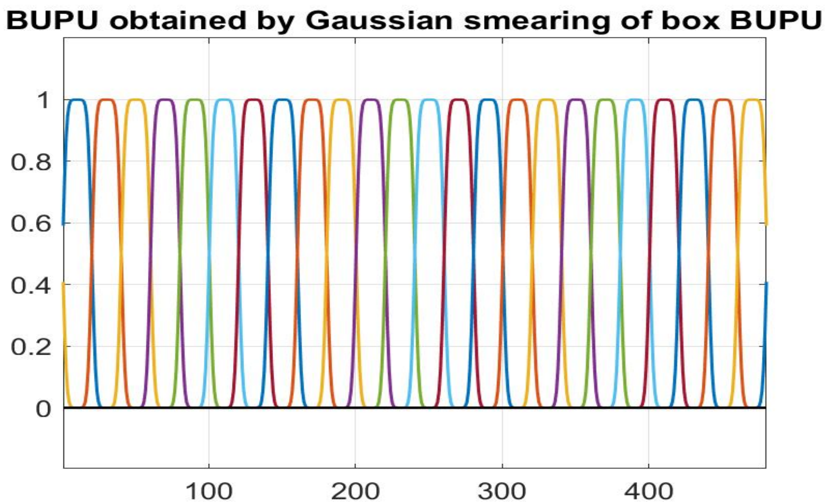

Abbildung 42: BUPUGaus1.jpg: smoothed bins as they are used for the creation of *histograms*.

22.4 Wiener's algebra

This is a kind of repetition on a more simple case, namely $\boldsymbol{A} = \boldsymbol{C}_0$. Recall how BUPUs look like:
We observe that one has for any $f \in \boldsymbol{C}_0(\mathbb{R}^d)$:

$$f = \sum_{i \in I} f\psi_i \quad \text{in } \Big(\boldsymbol{C}_0(\mathbb{R}^d), \, \|\cdot\|_\infty\Big),$$

meaning that the family of finite partial sums $\sum_{i \in F} f\psi_i$ converges to f in the sup-norm sense, or in full detail:
For any given $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ and $\varepsilon > 0$ one can find a finite set $F_0 \subset I$ such that one can guarantee for any finite set $F \supseteq F_0$ that

$$\Big\| f - \sum_{i \in F} f\psi_i \Big\|_\infty = \Big\| f(1 - \sum_{i \in F} \psi_i) \Big\|_\infty \leq \varepsilon. \tag{239}$$ C0convpart1

Just recall that this is done easily by observing that (by definition of $\boldsymbol{C}_0(\mathbb{R}^d)$) the set

$$Q := Q(f, \varepsilon) := \{x \,|\, |f(x)| \geq \varepsilon\}$$

is a bounded, in fact compact subset of $\mathbb{R}^d$. Hence only finitely many of the members of $\Psi = (\psi_i)_{i \in I}$ satisfy $\psi_i(q) \neq 0$ for some $q \in Q$. Let us call this set F_0. Then one has for $F \supseteq F_0$ and $q \in Q$ (since only terms which vanish over Q are omitted)

$$1 = \sum_{i \in I} \psi_i(q) = \sum_{i \in F} \psi_i(q) = \sum_{i \in F_0} \psi_i(q).$$

Consequently $f(q) - (\sum_{i \in F} f\psi_i)(q) = 0$ for all $q \in Q$. But for $x \notin Q$ one has

$$|f(x)(1 - \sum_{i \in F} \psi_i(x))| \leq |f(x)| \leq \varepsilon,$$

thus by combining the cases $q \in Q$ and $x \notin Q$ we arrive at the required claim (239). COconvpart1
Thus we have shown that the sum $\sum_{i\in F} f\psi_i$ is *unconditionally convergent* (think of summation over $\mathbb{Z}^d$, you can take iterated sums of any order etc.), but it is in general not absolutely convergent, i.e. we will not have $\sum_{i\in I} \|f\psi_i\|_\infty < \infty$. This (which is also more or less a reason whey $\boldsymbol{C}_0(\mathbb{R}^d)$ is NOT INCLUDED in $\boldsymbol{M}_b(\mathbb{R}^d)$) suggest to specifically consider the subspace of $\boldsymbol{C}_0(\mathbb{R}^d)$ which consists of all those functions f **for which one has absolute convergence.** We will call that space *Wiener algebra*[21].

WRd-def00 **Definition 11.**

$$\boldsymbol{W}(\mathbb{R}^d) := \boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d) := \{f \in \boldsymbol{C}_0(\mathbb{R}^d) \,|\, \sum_{i\in I} \|f\psi_i\|_\infty < \infty.\} \tag{240}$$ WRddef01

Adding $\boldsymbol{M}_b(\mathbb{R}^d)$ to the graph. The local behaviour of $\boldsymbol{W}(\mathbb{R}^d)$ is the same as that of $\boldsymbol{C}_0(\mathbb{R}^d)$ or $\boldsymbol{C}_b(\mathbb{R}^d)$, while the global behaviour (left/right) is the same as that of $\boldsymbol{L}^1(\mathbb{R}^d) = \boldsymbol{W}(\boldsymbol{L}^1, \boldsymbol{\ell}^1)(\mathbb{R}^d)$.

ener002b.jpg

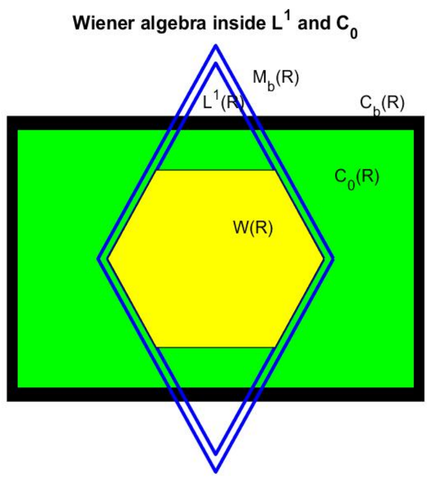

Abbildung 43: Wiener002b.jpg

Of course this raises the question, whether the definition depends on the particular choice of the BUPU Ψ (of course not, see the lemma below) and whether the natural expression of size, namely

$$\|f\|_{\boldsymbol{W}(\boldsymbol{C}_0,\boldsymbol{\ell}^1)} := \sum_{i\in I} \|f\psi_i\|_\infty \tag{241}$$ WC01inorm1

turns $\left(\boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d), \, \|\cdot\|_{\boldsymbol{W}}\right)$ into a Banach space.
First we verify the independence of the particular BUPU and then we collect the basic properties of the normed space $\left(\boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d), \, \|\cdot\|_{\boldsymbol{W}}\right)$.

WC0lr-props1 **Proposition 13.** *The following statements concerning* $\left(\boldsymbol{W}(\mathbb{R}^d), \, \|\cdot\|_{\boldsymbol{W}}\right)$ *are valid:*

[21] There are many objects with this name, but this is established terminology within the area of Time-Frequency analysis and function space theory.

WRLiFLi1.jpg

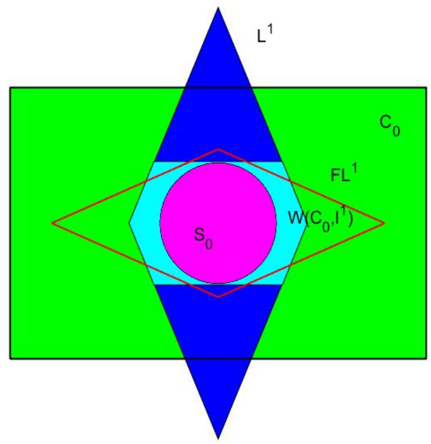

Abbildung 44: SOWRLiFLi1.jpg

1. *$\big(\boldsymbol{W}(\mathbb{R}^d),\, \|\cdot\|_{\boldsymbol{W}}\big)$ is a Banach space.*

2. *$\big(\boldsymbol{W}(\mathbb{R}^d),\, \|\cdot\|_{\boldsymbol{W}}\big)$ is continuously embedded into $\big(\boldsymbol{L}^1(\mathbb{R}^d),\, \|\cdot\|_1\big)$ as well as into $\big(\boldsymbol{C}_0(\mathbb{R}^d),\, \|\cdot\|_\infty\big)$, hence $\boldsymbol{W}(\mathbb{R}^d) \hookrightarrow \boldsymbol{L}^1 \cap \boldsymbol{C}_0$ (as a dense subspace).*

3. *$\boldsymbol{C}_c(\mathbb{R}^d)$ is a dense subspace of $\big(\boldsymbol{W}(\mathbb{R}^d),\, \|\cdot\|_{\boldsymbol{W}}\big)$, in fact*

$$\sum_{i\in F} f\psi_i \to f \;\; in \; \big(\boldsymbol{W}(\mathbb{R}^d),\, \|\cdot\|_{\boldsymbol{W}}\big).$$

4. *Translations and modulation operators act uniformly bounded on $\big(\boldsymbol{W}(\mathbb{R}^d),\, \|\cdot\|_{\boldsymbol{W}}\big)$, i.e. there exists $C_1 > 0$ such that*

$$\|T_x f\|_{\boldsymbol{W}(\mathbb{R}^d)} \leq C_1 \|f\|_{\boldsymbol{W}(\mathbb{R}^d)}, \quad \forall f \in \boldsymbol{W}(\mathbb{R}^d), x \in \mathbb{R}^d;$$

For the case of the continuous control function translations would be isometric.

5. *Translation is continuous within $\boldsymbol{W}(\mathbb{R}^d)$, i.e.*

$$\lim_{x\to 0} \|T_x f - f\|_{\boldsymbol{W}(\mathbb{R}^d)} = 0, \quad \forall f \in \boldsymbol{W}(\mathbb{R}^d).$$

6. *$\big(\boldsymbol{W}(\mathbb{R}^d),\, \|\cdot\|_{\boldsymbol{W}}\big)$ is an ideal in $\boldsymbol{M}_b(\mathbb{R}^d)$, with*

$$\|\mu \star f\|_{\leq} \|\mu\|_{\boldsymbol{M}_b} \|f\|_{\boldsymbol{W}(\mathbb{R}^d)}, \quad \mu \in \boldsymbol{M}_b(\mathbb{R}^d), f \in \boldsymbol{W}(\mathbb{R}^d). \tag{242}$$ MbconvWRd

7. *For any measure $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ and $f \in \boldsymbol{W}(\mathbb{R}^d)$ one has*

$$\lim_{|\Psi|\to 0} \|D_\Psi \mu \star f - \mu \star f\|_{\boldsymbol{W}(\mathbb{R}^d)} = 0. \tag{243}$$ DPsiWRd1

22.5 More general Wiener amalgam spaces

The use of the symbol $\boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)(\mathbb{R}^d)$ already indicates that one may have other spaces, to be denoted $\boldsymbol{W}(\boldsymbol{B}, \boldsymbol{\ell}^q)$, using the sequence space $\boldsymbol{\ell}^q(\mathbb{Z}^d)$, for $1 \leq q \leq \infty$, using the defining norm (for $q < \infty$, and $\|c_k\|_\infty$ for $q = \infty$):

$$\|(c_k)_{k\in\mathbb{Z}^d}\|_{\ell^q} = \left(\sum_{k\in\mathbb{Z}^d} |c_k|^q\right)^{1/q}. \tag{244}$$

lqnorm

As a *local component* $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ any Banach space of (generalized) functions can be used as long as $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ acts pointwise on $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ with an estimate of the form: For some $C = C_{A,B} < \infty$ one has:

$$\|h \cdot f\|_{\boldsymbol{B}} \leq C\|h\|_{\boldsymbol{A}}\|f\|_{\boldsymbol{B}}, \quad \forall h \in \boldsymbol{A}, f \in \boldsymbol{B}. \tag{245}$$

ABanmodul1

These spaces are then again well defined Banach spaces, independent of the use of the particular BUPU in $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ and so on, but there is no analogue of the technical Proposition (WAliatomic 10). Hence we postpone the discussion of these more general spaces. However, there are classical examples of such *Wiener amalgam spaces*, the space denoted by $\boldsymbol{W}(\boldsymbol{L}^p, \boldsymbol{\ell}^q)(\mathbb{R}^d)$, which are using the BUPU of rectangular functions. These amalgam spaces (sometimes also called Wiener-type spaces) are quite popular, especially in sampling theory and time-frequency analysis. (see [51] for an early survey article), and [24] for the - up to now - still most general approach. The convolution theorems have been inspired by [9] or [62] or [63].

In contrast to the regular case sampling theory (the theory which describes the recovery of band-limited functions from irregular samples the use of non-regular BUPUs. If their centers are *well-spread* (no formal definition here, e.g. certain minimal mutual distance of sampling points, but uniform covering of the underlying signal domain), they can also be used to characterize Wiener amalgam spaces. It just requires a bit more geometric arguments and some tricks.

Just in order to visualize the situation let us show a typical picture for the planar situation, i.e. the collection of circles which may serve as supports for a non-regular BUPU in $\mathbb{R}^2$:

UPUsupp2.jpg

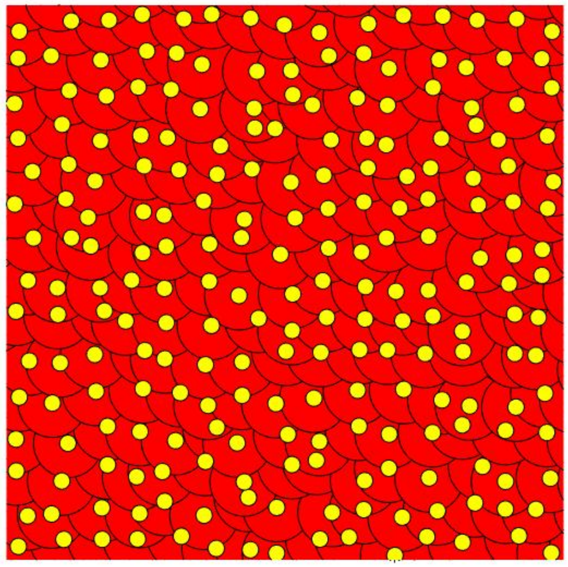

Abbildung 45: BUPUsupp2.jpg: supports of a family $\Psi = (\psi_i)_{i \in I}$ over $\mathbb{R}^2$.

PUcomp02.eps

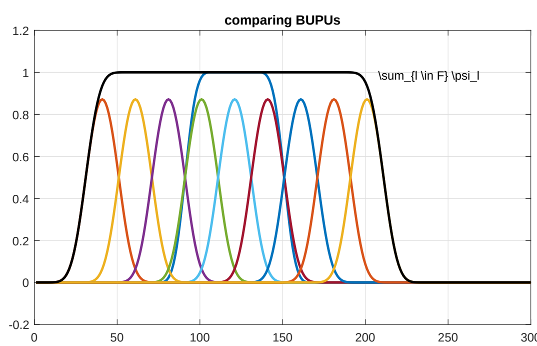

Abbildung 46: BUPUcomp02.eps

InvarPropS0

22.6 Invariance properties of $\boldsymbol{S}_0(\mathbb{R}^d)$

It is important to know that $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ is not only (isometrically) translation and modulation invariant (if we use the “continuous, Gaussian” norm) as well as Fourier invariant, but it is also invariant under dilations.

Recall our conventions concerning dilation: We write $\mathrm{D}_\rho h$ for the *value preserving* dilation given by $\mathrm{D}_\rho h(x) = h(\rho x)$ which is compatible with pointwise multiplication and the sup-norm: $\|\mathrm{D}_\rho h\|_\infty = \|h\|_\infty, h \in \boldsymbol{C}_b(\mathbb{R}^d)$.

Its adjoint is the operators St_ρ, defined for $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$: $[\mathrm{St}_\rho \mu](f) = \mu(\mathrm{D}_\rho f), f \in \boldsymbol{C}_0(\mathbb{R}^d)$, which can be described directly for $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ (or on $\boldsymbol{L}^1(\mathbb{R}^d)$) as $[\mathrm{St}_\rho k](x) = \rho^{-d} k(x/\rho)$. It is compatible with convolution and preserves the $\|\cdot\|_1$-norm.

In fact we have

SOdilatinv **Lemma 44.** *The operators* $(\mathrm{St}_\rho)_{\rho>0}$ *and* $(\mathrm{D}_\rho)_{\rho>0}$ *leave* $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ *invariant, and*

- *The family* $(\mathrm{D}_\rho)_{\rho\geq 1}$ *of value-preserving compression operators acts uniformly bounded on* $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$*;*
- *The family* $(\mathrm{St}_\rho)_{\rho\geq 1}$ *of area-preserving stretching operators acts uniformly bounded on* $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$.

Proof. We start by observing that the isometric St_ρ-invariance of $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ (more or less by the definition of the operators) implies the isometric D_ρ-invariance of $\left(\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_{\mathcal{F}\boldsymbol{L}^1}\right)$. The rest is done, using the atomic description of $\boldsymbol{S}_0(\mathbb{R}^d) = \boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)(\mathbb{R}^d)$. First we observe that one has for any $\rho > 0$

$$\mathrm{supp}(\mathrm{D}_\rho k) = (1/\rho)\,\mathrm{supp}(k), \quad k \in \boldsymbol{C}_c(\mathbb{R}^d).$$

For $\rho < 1$ one has a stretching effect, and if f has an atomic representation using atoms supported on balls $B_R(x_k)$ the same will be true (since the $\mathcal{F}\boldsymbol{L}^1$ norm of atoms is preserved under dilation), *but with larger supports* of the form $B_{R/\rho}(x_k/\rho)$. This implies that one has some constant[22] C_ρ such that

$$\|\mathrm{D}_\rho f\|_{\boldsymbol{S}_0} \leq C_\rho \|f\|_{\boldsymbol{S}_0}, \quad f \in \boldsymbol{S}_0.$$

For $\rho \geq 1$ the supports of atoms are shrinking, hence one has $C_\rho = 1$ in this case.
The effect of $\mathrm{St}_\rho, \rho \geq 1$ is based on the observation that $\mathrm{St}_\rho = \rho^{-d}\,\mathrm{D}_{1/rho}$, we leave the details to the reader (resp. will provide them later on, if time permits). In fact, we need a good estimate for C_ρ for $\rho \to 0$ (it is if the order ρ^{-d}, because a dilated atom hits roughly speaking ρ^{-d} many functions ψ_k). $\square$

By taking adjoint relations (this we do not provide a detailed proof) we have:

SOPdilatinv **Lemma 45.**
- *The family* $(\mathrm{D}_\rho)_{\rho\leq 1}$ *of value-preserving dilation operators acts uniformly bounded on* $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$*;*
- *The family* $(\mathrm{St}_\rho)_{\rho\leq 1}$ *of area-preserving compression operators acts uniformly bounded on* $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$.

It is an interesting consequence of the above dilation results, combined with the (now yet demonstrated result) showing that the *chirp-function* $chirp(t) = e^{i\pi t^2}$ on $\mathbb{R}$ is invariant under the extended Fourier transform, that we have a statement, that appears to be useful in the context of quantum mechanics:

comprchirp **Lemma 46.** *One has* $\boldsymbol{S}_0' - w^* - lim_{\rho\to 0}\,\mathrm{St}_\rho(\mathrm{chirp}) = \delta_0$.

[22] More or less depending on the number of smaller atoms required in order to decompose a large atom, so this is something dependent on the dimension $d \geq 1$ and ρ, some ρ^{-d} probably.

Proof. Since the extended FT preserves w^*-convergence an equivalent statement is

$$\mathcal{F}[\mathrm{St}_\rho(\mathrm{chirp})] = \mathrm{D}_\rho[\mathcal{F}(\mathrm{chirp})] = \mathrm{D}_\rho(\mathrm{chirp}) \to \mathbf{1}, \quad \text{for } \rho \to 0,$$

but this is easy to verify, noting that the involved mild distributions are uniformly bounded in $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$, and hence it is enough to test $\mathrm{D}_\rho(\mathrm{chirp})$ against elements from the dense subspace $\boldsymbol{C}_c(\mathbb{R}^d) \cap \boldsymbol{S}_0(\mathbb{R}^d) = \boldsymbol{C}_c(\mathbb{R}^d) \cap \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) \subset \boldsymbol{C}_c(\mathbb{R}^d)$. But due to the uniform convergence of $e^{\pi i \rho s^2} \to 1$ over the support Q of $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ one clearly has

$$\lim_{\rho\to 0} \int_{\mathbb{R}^d} (\mathrm{D}_\rho \mathrm{chirp})(s) k(s) ds = \lim_{\rho\to 0} \int_Q e^{i\pi(\rho s)^2} k(s) ds \to \int_{\mathbb{R}^d} \mathbf{1}(s) k(s) ds.$$

□

Remark 19. The interesting aspect of the above result (aside from indicating a number of interesting perspectives) is the observation that one can obtain δ_0 also as a limit of unbounded measures.
Clearly a similar argument justifies the statement

$$\boldsymbol{S}_0' - w^* - lim_{\rho\to 0}\, \mathrm{St}_\rho(\mathrm{SINC}) = \delta_0, \tag{246}$$ StrohSINC

because also in this case one also has

$$\mathcal{F}[\mathrm{St}_\rho(\mathrm{SINC})] = \mathrm{D}_\rho(\mathbf{1}_{[-1/2,1/2]}) = \mathbf{1}_{[-1/2\rho,1/2\rho]} \to \mathbf{1} \quad \text{for } \rho \to 0.$$

But $\mathrm{SINC} \in \boldsymbol{L}^2(\mathbb{R}^d)$, although the boundedness of $\mathrm{St}_\rho\, \mathrm{SINC}$, for $\rho \le 1$ is only valid in the $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$-sense, and not with respect to the $\boldsymbol{L}^2$-norm.

23 Various minor issues 25.10.2020

The fact, that $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ contains plateau functions combined with the convolution theorem imply that the Fourier Stieltjes algebra

$$\mathcal{FST}(\mathbb{R}^d) := \{\widehat{\mu}, \mu \in \boldsymbol{M}_b(\mathbb{R}^d)\}$$

has locally the same structure as $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$. Consequently we have:

FSTinCbRd **Lemma 47.**

$$\mathcal{FST}(\mathbb{R}^d) \hookrightarrow \boldsymbol{C}_b(\mathbb{R}^d).$$

Proof. Since we know that the trapezoidal function is the convolution product of the rectangular functions (of different size) we see that their Fourier transforms decay like a pointwise product of SINC-functions, i.e. are dominated by $(1+x^2)^{-1}$ over $\mathbb{R}$. Since these trapezoidal functions can be chosen to be even and in $\boldsymbol{C}_c(\mathbb{R})$ we find that the do have an integrable Fourier transform, hence the belong to the Fourier algebra $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ (because $\mathcal{F}^{-1}(h) = \mathcal{F}(h^{\checkmark}) = \mathcal{F}(h)$ for even functions $h \in \boldsymbol{L}^1(\mathbb{R})$.
For the multi-dimensional case ($d \geq 2$) we can obtain functions in the Fourier algebra which are constant one on cubes of the form $[-R,R]^d$ by taking tensor products of such trapezoidal functions, recalling that one has

$$\mathcal{F}(f \otimes g) = \mathcal{F}(g) \otimes \mathcal{F}(g), \quad \text{and} \quad \|f \otimes g\|_{\boldsymbol{L}^1} = \|(\|_{\boldsymbol{L}^1} f)\|(\|_{\boldsymbol{L}^1} g),$$

and thus

$$\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^m) \otimes \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^k) \subset \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^{m+k}).$$

Let us denote such trapezoidal functions in $(\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1)(\mathbb{R}^d)$ by τ, and their (inverse) Fourier transforms by ϕ.
Given now $\widehat{\mu} \in \mathcal{FST}(\mathbb{R}^d)$ we find that the Fourier transform of $\mu \star \phi \in \boldsymbol{L}^1(\mathbb{R}^d)$ is of the form $\hat{\tau} \cdot \widehat{\mu}$, but it coincides with $\widehat{\mu}$ over any given compact set. Since $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) \subset \boldsymbol{C}_0(\mathbb{R}^d)$ consists only of continuous functions this means that $\widehat{\mu} \in \boldsymbol{C}_b(\mathbb{R}^d)$, since we know already that $|\widehat{\mu}s| \leq \|, \|_{\boldsymbol{M}_b}\ s \in \mathbb{R}^d$. □

Note: The alternative, more technical approach to this fact can be based on the fact that the family $\mathcal{F}(D_\Psi \mu)$ converges uniformly over compacts sets to $\widehat{\mu}$, because this family is not only bounded and w^*-convergent (this by itself would not be sufficient), but it is also uniformly tight. This is best understood by the observation that for any μ with compact support and for all the BUPUs Ψ with $|\Psi| \leq 1$ one has joint support for all the discrete measures $D_\Psi \mu$.
Final remark: Due to the convolution theorem we find that $\mathcal{FST}(\mathbb{R}^d)$ is a pointwise algebra. We can transfer the norm from $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ to $\mathcal{FST}(\mathbb{R}^d)$ by setting $\|\widehat{\mu}\|_{\mathcal{FST}} := \|\mu\|_{\boldsymbol{M}_b}, \mu \in \boldsymbol{M}_b(\mathbb{R}^d)$, because the Fourier-Stieltjes transform $\mu \to \widehat{\mu}$ is injective. In fact, assume that $\widehat{\mu}(s) = 0$ for any $s \in \mathbb{R}^d$, then $\mu \star \phi = 0$ for any $\phi \in \boldsymbol{L}^1(\mathbb{R}^d)$. But ($g \in \boldsymbol{L}^1(\mathbb{R}^d)$ with $\hat{g}(0) = 1$)

$$\mu(f) = \lim_{\rho \to 0} (\mu \star \mathrm{St}_\rho g)(f) = \lim_{\rho \to 0} \mu(\mathrm{St}_\rho g^{\checkmark} \star f) = 0$$

then implies that $\mu = 0$ (as element of $\left(\boldsymbol{C}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{C}_0'}\right)$).

24 BIBO systems represented by L^1-functions

The key observation of **un20 :** [89] concerns the fact, that in various books (he gives prominent examples) on signal processing the false claim is made that every BIBO system is obtained as a convolution operator with some $\boldsymbol{L}^1$-function (usually the argument is presented for $d = 1$, but it is valid also for $d \geq 1$ as well).
The clear example is an ordinary shift operator T_x (or convolution by δ_x), specifically the identity operator, or T_0 (no shift at all, or shift by zero!), which is represented (in our way) as a convolution with the standard Dirac measure δ_0 (often just written as Dirac δ)[23].
It is clear that the identity operator cannot be written as a convolution operator by any $\boldsymbol{L}^1(\mathbb{R}^d)$-function, due to the Riemann Lebesgue Lemma. None of the convolution operators C_f with $f \in \boldsymbol{L}^1(\mathbb{R}^d)$ (equivalently: convolution with the bounded measure induced by the Lebesgue integrable "function") is in fact invertible. We will discuss a partial inverse in the sequel (going by the name of *Wiener's Inversion Theorem.*
But on the positive side let us explain in which sense the *standard claim* can be, at least, partially justified.
Let us put in a practical situation. We cannot (for physical reasons) insert a true Dirac into the system in order to obtain the *impulse response* $T(\delta_0) = \mu \star \delta_0 = \mu$.
What we can do is to *test the system* on a (perhaps large) but finite number of input signals f_n,$n = 1, ..., L$, and try to obtain information from the*input-output relationship*, So assume we know $\mu \star f_n$ for $1 \leq n \leq L$. So we can put the question: Can we decide whether the operator (TILS) T has been realized as a convolution operator by some (finite) discrete measure or as convolution with some $\boldsymbol{L}^1(\mathbb{R}^d)$-function?
We will give several indications that this is a delicate questions, which cannot be answered decisively!
First, we have seen, that for ANY convolution operator $T : f \mapsto \mu \star f$ there is a discrete (without loss of generality even finite discrete) measure producing a close-by result. In fact, given any $g \in \boldsymbol{L}^1(\mathbb{R}^d)$ (better μ_g, or in fact a generic $\mu \in \boldsymbol{M}_d(\mathbb{R}^d)$) we can choose in the given situation for any $\varepsilon > 0$ (think of ε close to the numerical precision of your device) a fine discretization $\nu = \mathrm{D}_\Psi(\sum_{i \in F} \mu\, \psi_i) \in \boldsymbol{M}_d(\mathbb{R}^d)$ such that

$$\|T(f_n) - \nu \star f\|_\infty \leq \varepsilon, \quad \text{for } 1 \leq n \leq L. \tag{247}$$ TILSapprox

This means, that to the observer (even of many) input-output relationships the action of the discrete measure ν and that of the original (perhaps $\boldsymbol{L}^1(\mathbb{R}^d)$ induced) measure μ may not distinguishable (leave alone that an observer may be able to take e.g. only finitely sampling values $T(f_n)(x_k^n)$).
There is another aspect, which indicates that it is in fact *impossible* to recognize from a finite set of observations $T(f_n)$ whether T is generated from a discrete or an absolutely continuous measure (i.e. as $\mu_g, g \in \boldsymbol{L}^1(\mathbb{R}^d)$).

[23] There is also the Kronecker δ, which is describing an $n \times n$-matrix, representing the identity operator on $\mathbb{C}^n$ or $\mathbb{R}^n$, i.e. simply `eye(n)` in a MATLAB setting. Whereas the Kronecker symbol is written as $\delta_{j,k}$, with the value 1 for $j = k$ and zero else, i.e. describing the whole matrix, one can describe it as the collection of (column or) row vectors which have a single 1 at position k of the k-th row. Applied to a column vector $\mathbf{x}$ it thus picks out the k-the coordinate.

The argument is based on the Cohen-Hewitt factorization theorem. Given a finite set f_n and $\varepsilon > 0$ there exists some $g \in \boldsymbol{L}^1(\mathbb{R}^d)$ and $h_n \in \boldsymbol{C}_0(\mathbb{R}^d)$ such that $\|f_n - h_n\|_\infty \leq \varepsilon$ for $1 \leq n \leq L$ such that $f_n = g \star h_n$. Consequently $T(f_n) = \mu \star f_n = (\mu \star g) \star h_n$. Since $\mu \star g \in \boldsymbol{L}^1(\mathbb{R}^d)$ it shows that $T(f_n)$ is in the range (exactly!) of some (slightly different) input signal (compared to the original one).

25 Wiener's Inversion Theorem

Wiener's Inversion Theorem is a way to overcome the lack of invertibility, at least for band-limited functions.
First let us look at the $\boldsymbol{L}^2(\mathbb{R}^d)$-situation. Any TILS on $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$ is given by a bounded transfer function (in fact by some element of $\boldsymbol{L}^\infty(\mathbb{R}^d)$) and is of the form $\mathcal{F}(T(f)) = h \cdot \mathcal{F}(f)$ for some $f \in \boldsymbol{L}^\infty(\mathbb{R}^d)$ with

$$\|h\|_\infty = |||T|||_{\boldsymbol{L}^2(\mathbb{R}^d)} \,. \tag{248}$$

TILSLtRd

In the classical setting this claim is an easy consequence of Plancherel's Theorem plus the work with measurable functions. We still have to obtain a proof in our current setting. The obvious part is that any $h \in \boldsymbol{C}_b(\mathbb{R}^d)$ defines a TILS on $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$ and satisfies (TILSLtRd 248).
If we look, for some compact set Ω, e.g. $\Omega = B_R(0)$ or

$$\Omega = \prod_{k=1}^{d} [-\omega_k, \omega_k]$$

and write

$$\boldsymbol{L}^2_\Omega = \{f \in \boldsymbol{L}^2(\mathbb{R}^d) \mid \operatorname{supp}(\widehat{f}) \subseteq \Omega\}$$

for the space of (Ω-) band-limited elements of $\boldsymbol{L}^2(\mathbb{R}^d)$.
Assume that $h \in \boldsymbol{C}_b(\mathbb{R}^d)$ satisfies $h(z) \neq 0$ for $z \in \Omega$. It is then clear that $1/h(z)$ is well defined and continuous there. Taking any continuation of this local piece, i.e. any bounded function $h_1 \in \boldsymbol{C}_b(\mathbb{R}^d)$ with $h_1(z) = 1/h(z)$ for $z \in \Omega$.
It is then clear that the system $T_1 : f \mapsto \mathcal{F}^{-1}(h_1 \cdot \mathcal{F}(f))$ inverts the operator T on the closed subspace $\boldsymbol{L}^2_\Omega$. In fact, we have (viewing everything on the Fourier transform side) for $f \in \boldsymbol{L}^2_\Omega$:

$$\mathcal{F}[T_1(T(f))](z) = h_1(z)h(z)\widehat{f}(z) = \widehat{f}(z), z \in \mathbb{R}^d,$$

because the result is valid for $z \notin \Omega$ (both expression vanish, because $\widehat{f}(z) = 0$ there), and it is valid for $z \in \Omega$, since $h(z)h_1(z) = 1$ for $z \in \Omega$.
The content of *Wiener's Inversion Theorem* is the interesting fact that a similar result is true for convolution operators from $\boldsymbol{L}^1(\mathbb{R}^d)$. In other words. If we assume that $T(f) = g \star f$ for some $g \in \boldsymbol{L}^1(\mathbb{R}^d)$ there exists some $g_1 \in \boldsymbol{L}^1(\mathbb{R}^d)$ such that

$$\hat{g}_1(z) = 1/\hat{g}(z), z \in \Omega,$$

of course under the natural assumption that $\hat{g}(z) \neq 0$ for $z \in \Omega$.

25.1 The mathematics of Wiener's Inversion Theorem

We will not do the details of Wiener's inversion theorem but rather explain the basic steps towards this remarkable result. Details are found well explained in **re68 :** [78] resp. the new (enlarged) edition **rest00 :** [79], already in the first chapter.
Note, that Reiter uses the symbol M_ρ (indicating that the argument of the function is *multiplied*, in fact with $1/\rho$) for the operator which is written as St_ρ in the current notes, but in the way that preserves the $\boldsymbol{L}^1(\mathbb{R}^d)$-norm.
First let us reformulate the result right in the Fourier setting:

WienerInvThm **Theorem 9.** *Given $h \in \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ and a compact set $\Omega \subset \mathbb{R}^d$ such that $h(z) \neq 0$, then there exists $h_1 \in \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ such that*

$$h_1(z) = 1/h(z), \quad z \in \Omega.$$

Remark 20. The proof does not provide any control about the $\mathcal{F}\boldsymbol{L}^1$-norm of h_1, in fact one cannot obtain such universal estimates without extra conditions.

In the proof one uses the fact that St_ρ defines an isometric automorphism of the Banach convolution algebra $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$:

$$\mathrm{St}_\rho(f \star g) = \mathrm{St}_\rho(f) \star \mathrm{St}_\rho(g), \quad f, g \in \boldsymbol{L}^1(\mathbb{R}^d). \tag{249}$$ StroAuto1

Wiener's Inversion Theorem is a special case of the more general Wiener-Levy Theorem, which shows that one can apply analytic functions to the Fourier transform, as long as one stays away from singularities, and thus one can use power series expansions of analytic functions (Chap.1.3 of [78]).
Roughly speaking the idea is the following: Locally (i.e. over small open balls $B_r(z)$) any analytic function

$$H(u) = \sum_{k=0}^{\infty} c_k u^k$$

can be applied to h, i.e. one can find some $h_2 \in \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ such that

$$h_2(z) = H(h(x)), \quad z \in B_r(x). \tag{250}$$ WLevyLoc

These local pieces can be glued together using fine partitions in $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ to obtain the equality (WLevyLoc 250) over all of Ω.
An important auxiliary result on the way is the following:

IdBAId **Lemma 48.** *The closed ideal $I_0 := \{f \in \boldsymbol{L}^1(\mathbb{R}^d) \,|\, \widehat{f}(0) = 0\}$ has a bounded approximate unit of norm 2, i.e. for any family $f_1, ..., f_L \in I_0$ and $\varepsilon > 0$ there exists $g \in I_0$ with $\|g\|_{\boldsymbol{L}^1} \le 2$ such that*

$$\|g \star f_k - f_k\|_{\boldsymbol{L}^1} \le \varepsilon, \quad 1 \le k \le L.$$

hgfei inserted 2.11.2020

MATERIAL of 02.11.2020

26 Fourier invariant spaces

First let us recall that the *Fourier Inversion Theorem* implies that

LiFLichar1 **Lemma 49.** $(\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1)(\mathbb{R}^d)$ *coincides with the set*

$$\{f \in \boldsymbol{L}^1(\mathbb{R}^d) \mid \widehat{f} \in \boldsymbol{L}^1(\mathbb{R}^d)\}.$$

Proof. Let us start with $f \in \boldsymbol{L}^1(\mathbb{R}^d)$, such that $\widehat{f} \in \boldsymbol{L}^1(\mathbb{R}^d)$. Then $f = \mathcal{F}^{-1}(\mathcal{F}(f)) = \mathcal{F}(\widehat{f}^{\checkmark})$, and clearly $\widehat{f}^{\checkmark} \in \boldsymbol{L}^1(\mathbb{R}^d)$, thus $f \in \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$, as the image of some $\boldsymbol{L}^1$-function under the Fourier transform.
Conversely, assume that $f \in \boldsymbol{L}^1(\mathbb{R}^d)$ is also in $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$, i.e. $f = \widehat{g}$ for some $g \in \boldsymbol{L}^1(\mathbb{R}^d)$. Then of course $\mathcal{F}(f) = \mathcal{F}(\widehat{g}) = g^{\checkmark} \in \boldsymbol{L}^1(\mathbb{R}^d)$.
Also the corresponding natural norms are identical, since

$$\|f\| = \|f\|_{\boldsymbol{L}^1} + \|\widehat{f}\|_{\boldsymbol{L}^1} = \|f\|_{\boldsymbol{L}^1} + \|\,\mathcal{F}^{-1}(f)\|_{\boldsymbol{L}^1}, \tag{251}$$ LiFLinorm

turning that space into a Banach space. □

26.1 Recalling basic properties of $\boldsymbol{S}_0(\mathbb{R}^d)$

We have introduced $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ as $\left(\boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1, \boldsymbol{\ell}^1)(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1,\boldsymbol{\ell}^1)(\mathbb{R}^d)}\right)$, using absolutely convergent decompositions via a (regular) BUPU in $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$, and have established an atomic characterization, which helps to derive more easily a number of basic properties. Let us recall the most important ones:

1. $\boldsymbol{S}_0(\mathbb{R}^d)$ is a Fourier invariant Banach space, continuously embedded into $\left(\boldsymbol{W}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{W}}\right)$ (hence into $\boldsymbol{L}^1(\mathbb{R}^d)$ and $\boldsymbol{L}^2(\mathbb{R}^d)$, and many other spaces);

2. It is a so-called *Segal algebra*, meaning essentially that it is a dense subspace of $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ which is an ideal, thus

$$\boldsymbol{L}^1(\mathbb{R}^d) \star \boldsymbol{S}_0(\mathbb{R}^d) \subseteq \boldsymbol{S}_0(\mathbb{R}^d)$$

combined with the corresponding norm estimates, and $\widehat{g}(0) = 1$ implies that

$$\mathrm{St}_\rho g \star f \to f \text{ in } \left(\boldsymbol{S}_0(\mathbb{R}), \|\cdot\|_{\boldsymbol{S}_0}\right), \text{ for } \rho \to 0.$$

3.
$$\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) \cdot \boldsymbol{S}_0(\mathbb{R}^d) \subseteq \boldsymbol{S}_0(\mathbb{R}^d)$$

combined with the corresponding norm estimates, and $h(0) = 1$ for any $h \in \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ implies that

$$\mathrm{D}_\rho h \cdot f \to f \text{ in } \left(\boldsymbol{S}_0(\mathbb{R}), \|\cdot\|_{\boldsymbol{S}_0}\right), \text{ for } \rho \to 0.$$

bandlLiS0 **Corollary 6.** *Any band-limited function in* $\boldsymbol{L}^1(\mathbb{R}^d)$ *belongs to* $\boldsymbol{S}_0(\mathbb{R}^d)$*. Moreover, for any compact set* $M \subset \mathbb{R}^d$ *there exists a constant* $C = C_M$ *such that*

$$\|f\|_{\boldsymbol{S}_0} \leq C_M \|f\|_{\boldsymbol{L}^1}, \quad \forall f \in \boldsymbol{L}^1(\mathbb{R}^d), \mathrm{supp}(\widehat{f}) \subseteq M. \tag{252}$$ S0Lino1

Proof. Band-limited functions from $\boldsymbol{L}^1(\mathbb{R}^d)$ have a Fourier transform (obviously) in $\boldsymbol{A} = \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$, and have compact support, thus $\widehat{f} \in \boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1, \boldsymbol{\ell}^1)(\mathbb{R}^d)$, and by consequence $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ (by Theorem Fourinv **??**).
Given M there exists some trapezoidal function τ in $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) \cap \boldsymbol{C}_c(\mathbb{R}^d)$ with $\tau(s) = 1$ for $s \in M$, and thus $\widehat{f} \cdot \tau = \widehat{f}$. Taking the IFT one has for $\varphi = \mathcal{F}^{-1}(\tau)$

$$f = \varphi \star f \in \boldsymbol{S}_0(\mathbb{R}^d) \star \boldsymbol{L}^1(\mathbb{R}^d) \subseteq \boldsymbol{S}_0(\mathbb{R}^d), \tag{253}$$ LibldinS0

with the norm estimate (thus $C_M = \|\tau\|_{\boldsymbol{S}_0} = \|\varphi\|_{\boldsymbol{S}_0}$ is a possible choice) estimate:

$$\|f\|_{\boldsymbol{S}_0} \leq \|\varphi\|_{\boldsymbol{S}_0} \|f\|_{\boldsymbol{L}^1}. \tag{254}$$ LibldS0est1

□

Remark 21. Since the norm on $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ is modulation invariant it is possible to replace τ by $T_z\tau$ for any $z \in \mathbb{R}^d$ without spoiling the constant appearing in (S0Lino1 252). This means that the same constant can be used for $z + M$. Thus it is not the maximal frequency arising in M, but only the diameter of M which plays an essential role. Thus in fact one could derive the behaviour of the optimal constant for sets of the form $M = B_R(x)$ as a function of R (order R^d, for $R \to \infty$).

27 The Fourier transform for mild distributions

27.1 Motivation

It is one of the *main goals* of this course to popularize the usefulness of the space $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$, also called the space of *mild distributions*, as described in **fe20-1:** [35][24], mostly with respect to an understanding the role of the Fourier transform, now NOT seen as an integral transform but rather as a more general transform with certain properties. It will also illustrate in which sense the current approach (aside from avoiding Lebesgue integration) overcomes the traditional classification of signals into discrete versus continuous, or periodic versus non-periodic one. This point of view is very much the view-point of abstract harmonic analysis (AHA), which deals with functions (or generalized) functions over LCA (locally compact groups). Depending on the choice of such a group we find that there is always a *version of the Fourier transform*, mapping integrable functions (i.e. elements from $\left(\boldsymbol{L}^1(G), \|\cdot\|_1\right)$) on the group $\mathscr{G}$ into continuous functions $\widehat{f} \in \left(\boldsymbol{C}_0(\widehat{G}), \|\cdot\|_\infty\right)$. There is also a convolution theorem, a Fourier Inversion Theorem, and a Plancherel Theorem and so on, all relying on the existence of a Haar measure on $\mathscr{G}$, the fact that the *dual group* $\widehat{\mathscr{G}}$ is also a LCA group (and Pontrjagin's Theorem stating that the dual group for $\widehat{\mathscr{G}}$ is naturally isomorphic to the original group $\mathscr{G}$, thus giving the "time" and the "frequency" variable, which may live on different "domains" an equivalent role within Fourier Analysis over groups.

For the connection between classical AHA (as presented e.g. in the books of L. Loomis **lo53:** [75] or W. Rudin **ru62:** [80], or in a more modern form in **re68:** [78] and **fo95:** [50]) we may recommend the book entitled "Unified Signal Theory" by G. Ciarolo **ca11:** [11], putting this perspective into the language of engineers.

Traditionally books on Fourier Analysis often have this perspective in the background and take the historical path. Starting from the most classical theory, which is the theory of *Fourier Series for periodic functions*[25] they go on to take limits (the length of the period is going to infinity, so that a-periodic, but decaying signals can undergo a continuous version of the Fourier series expansion, now know as the (integral version) of the Fourier transform, with the well known inversion theorem derived (first) equally vaguely, and so on. Finally, the DFT (Discrete Fourier Transform), of course implemented as FFT (Fast Fourier Transform) is mentioned as a *computational tool*, again analogue to the Fourier series expansion, but now with discrete date, i.e. applied to vectors of finite length, i.e. elements of $\mathbb{C}^N$ (we view them as data in $\boldsymbol{\ell}^2(\mathbb{U}_N)$). This last step can basically be treated in terms of linear algebra and realized (computationally) using e.g. MATLAB.

We will kind of go the converse path. Once we have established and (abstract??, but symmetric) version of the Fourier transform we show that more or less all of the typical realizations of abstract Fourier transforms can be viewed as the consequence of properties

[24] See **feja20:** [41] for a mathematical paper along the lines of the current notes, or **fe20-2:** [36] concerning the possible general plan for a course like this one, or **fe19:** [33] for details concerning the relationship of this approach to the classical one. We will only highlight a few of these aspects here and provide some additional ones, as they are relevant for engineering applications.

[25] This problem was one of the driving forces for the development of many important concepts for mathematical analysis, including the Lebesgue integral, but even set theory, and later on the theory of tempered distributions by Laurent Schwartz (**sc57:** [84]), but in particular it was stimulating the development of Functional Analysis as we are using it a lot in this course.

of specific subspaces of $\boldsymbol{S}_0'(\mathbb{R}^d)$ and their Fourier transforms.

27.2 Special cases of the FT for mild distributions

We have established $\boldsymbol{S}_0(\mathbb{R}^d)$ as a decent Fourier invariant Banach space[26] and define

mildFTdef00 **Definition 12.**

$$\widehat{\sigma}(f) = \sigma(\widehat{f}), \quad \sigma \in \boldsymbol{S}_0', \ f \in \boldsymbol{S}_0. \tag{255}$$ mildFTdef

Given the fact that $\|\widehat{f}\|_{\boldsymbol{S}_0} = \|f\|_{\boldsymbol{S}_0}$ for $f \in \boldsymbol{S}_0$ implies that the Fourier transform defined for $\boldsymbol{S}_0'(\mathbb{R}^d)$ is isometric as well, and an automorphism (i.e. bijective). Any $\tau \in \boldsymbol{S}_0'$ is the Fourier transform of some $\sigma \in \boldsymbol{S}_0'$, and obviously the extended Fourier transform can be inverted by the simple formula

$$[\mathcal{F}^{-1}\,\tau](g) = \tau(\mathcal{F}^{-1}\,g), \quad \tau \in \boldsymbol{S}_0', \ g \in \boldsymbol{S}_0. \tag{256}$$ mildIFTdef

It is also clear that the Fourier-Stieltjes transform can be seen as a special case of the Fourier transform for $\boldsymbol{S}_0'(\mathbb{R}^d)$, and this in turn of course is a generalization of Fourier transform on $\boldsymbol{L}^1(\mathbb{R}^d)$.
This is the so-called *consistency principle*, which can be rephrased in the following short formula, writing for $g \in \boldsymbol{L}^1(\mathbb{R}^d)$

$$\sigma_g(f) = \int_{\mathbb{R}^d} f(x)\,h(x)dx, \quad f \in \boldsymbol{S}_0,$$

with

$$|\sigma_g(f)| \leq \int_{\mathbb{R}^d} |f(x)|\,|g(x)|dx \leq \|f\|_\infty \|g\|_{\boldsymbol{L}^1} \leq (C\|g\|_{\boldsymbol{L}^1})\|f\|_{\boldsymbol{S}_0},$$

satisfies (with $\mathcal{F}$ describing the generalized Fourier transform, while $\widehat{g}$ denotes the integral form of the Fourier transform), in short $\sigma_{\widehat{g}} = \widehat{\sigma_g}$, or:

$$\sigma_{\widehat{g}} = \mathcal{F}(\sigma_g), \quad g \in \boldsymbol{L}^1(\mathbb{R}^d). \tag{257}$$ FTconsist

Remark 22. As we will see later on the definition chosen above (via an adjointness relation) is the unique w^*-w^*-continuous continuation of the ordinary Fourier transform (even just the one defined on $\boldsymbol{W}(\mathbb{R}^d)$, using Riemann integrals).
Later on we will show that the DFT (usually realized via FFT) for discrete periodic signals can also be derived in the framework of the (extended) Fourier transform for $\boldsymbol{S}_0'(\mathbb{R}^d)$, and that it could also be used as a motivation for the definition of the extended FT via w^*-approximation. We will have a separate section on this later on.

[26]... with a lot of other properties, ensuring that it is a rich space, nevertheless contained in most of the function spaces of interest for Fourier analysis, e.g. dense in $\big(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\big)$ or $\big(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\big)$.

27.2.1 Trigonometric polynomials and discrete measures

This already allows us to compute the inverse (or forward) Fourier transform of χ_s:

$$[\mathcal{F}(\chi_s)](f) = \chi_s(\widehat{f}) = \int_{\mathbb{R}^d} \widehat{f}(y)\chi_s(\widehat{f}) = f(s) = \delta_s(f), \quad f \in \boldsymbol{S}_0(\mathbb{R}^d). \tag{258}$$ IFTchis

My standard comment on this is: If a musician listens to a pure frequency χ_s she/he would say: This pure tone I hear has exactly frequency s, and this is nicely expressed by a Dirac delta located at s.

Remark 23. Of course an important special case is obtained by putting $s = 0$, which gives $\mathcal{F}^{-1}(\mathbf{1}) = \delta_0$, which appears in the engineering literature often as the (annoying to mathematicians, divergent) integral

$$\int_{-\infty}^{\infty} 1\, e^{2\pi i s t} = \delta(t). \tag{259}$$ IFTconst1

It is OK to *translate* this formula in the distributional setting as formulated above. Otherwise we have a divergent integral to the left and a generalized function to the right, which cannot be described by pointwise values. The idea that the left hand side is "kind of zero" for $t \neq 0$, and that one obtains $+\infty$ for $t = 0$ (why not $3 \cdot \infty$?) is highly problematic and should not be taken as a solid basis for the further development of a mathematical theory! The question of "existence" of a FT becomes in this way obsolete!

Summarizing our knowledge about finite discrete measures and pure frequencies we have thus altogether:

trigpolFT00 **Lemma 50.** *The (forward or inverse) Fourier transform of a trigonometric polynomial, i.e. $h \in \boldsymbol{C}_b(\mathbb{R}^d)$ which is a finite linear combination of pure frequencies of the form*

$$h(x) = \sum_{i \in F} c_k \chi_{s_k} \tag{260}$$ tripolFT

is a discrete, finite measure of the form (and vice versa!)

$$\sum_{i \in F} c_k \delta_{s_k}, \quad \text{resp.} \quad \mathcal{F}^{-1}(h) = \sum_{i \in F} c_k \delta_{-s_k}.$$

.

27.2.2 Discrete measures on $\mathbb{Z}^d$

We have already seen that the convolution of bounded, discrete measures behaves in the expected way, based on the rule $\delta_x \star \delta_y = \delta_{x+y}, \; x, y \in \mathbb{R}^d$. By consequence we have for any *lattice* $\Lambda \lhd \mathbb{R}^d$ (i.e. a discrete subgroup of the form $\boldsymbol{A} \star \mathbb{Z}^d$, for some non-singular $d \times d$-matrix $\boldsymbol{A}$) a commutative subalgebra of $\boldsymbol{M}_d(\mathbb{R}^d)$, denoted by $\big(\boldsymbol{\ell}^1(\Lambda), \, \|\cdot\|_1\big)$:

liLamdef00 **Lemma 51.**

$$\boldsymbol{\ell}^1(\Lambda) := \{\nu = \sum_{\lambda \in \Lambda} c_\lambda \delta_\lambda, (c_\lambda)_{\lambda \in \Lambda} \in \boldsymbol{\ell}^1(\Lambda)\}, \tag{261}$$ liLamdef

with the natural norm (inherited from $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$)

$$\|\nu\|_{\boldsymbol{\ell}^1(\Lambda)} = \sum_{\lambda \in \Lambda} |c_\lambda| < \infty. \tag{262}$$

liLamprop1 **Proposition 14.** $\left(\boldsymbol{\ell}^1(\Lambda), \|\cdot\|_{\boldsymbol{\ell}^1(\Lambda)}\right)$ *is a commutative Banach algebra with respect to discrete convolution. Given two measures*

$$\nu_1 = \sum_{\lambda\in\Lambda} c^1_\lambda \delta_\lambda, \quad \nu_2 = \sum_{\lambda\in\Lambda} c^2_\lambda \delta_\lambda$$

the convolution product (inherited from $\boldsymbol{M}_b(\mathbb{R}^d)$*) is the discrete measure of the form*

$$\nu = \sum_{\lambda\in\Lambda} c_\lambda \delta_\lambda,$$

with (the Cauchy-product formula):

$$c_\lambda = \sum_{\lambda=\lambda'+\lambda''} c^1_{\lambda'} \cdot c^2_{\lambda''}.$$

For $\Lambda = \mathbb{Z}^d$ the Cauchy product can be better described (and programmed) as

$$c_n = \sum_{k\in\mathbb{Z}^d} c^1_k c^2_{n-k} = \sum_{k\in\mathbb{Z}^d} c^2_k c^1_{n-k}. \tag{263}$$ CauchProd1

Of course this expression can be described as a convolution product of the vectors $(c^1_k)_{k\in\mathbb{Z}^d}$ and $(c^2_j)_{j\in\mathbb{Z}^d}$ using the scalar product

$$\langle \mathbf{x}, \mathbf{y}\rangle = \sum_{k\in\mathbb{Z}^d} x_k y_k,$$

by observing that (CauchProd1 263) can be recast into

$$c_n = \langle \mathbf{c^2}, T_n(\mathbf{c^{1\checkmark}})\rangle,$$

i.e. into the formula already known from the beginning, now for the discrete group $\mathbb{Z}^d$ (and in a similar way for general lattices $\Lambda \lhd \mathbb{R}^d$).

Remark 24. Of course this is reminiscent of the pointwise product of polynomials described at the coefficient level (see (discrconv1 124)).

28 The classical setting

classsetg

The classical setting starts from continuous, periodic functions. Let us consider the most simple case, namely $\Lambda = \alpha\mathbb{Z}^d \lhd \mathbb{R}$, the prototypical periodic function h on $\mathbb{R}$, with period $\alpha > 0$, i.e. $T_{\alpha k} h = h$, for $k \in \mathbb{Z}$. If such a function h is continuous it is in fact *uniformly continuous* (as the uniform compactness of h over the compact fundamental period, e.g. $[0, \alpha]$ or $[-\alpha/2, \alpha/2]$) implies that $f \in \boldsymbol{C}_{ub}(\mathbb{R}^d)$.
Such a function defines (like any $h \in \boldsymbol{C}_b(\mathbb{R}^d)$) an element in $\boldsymbol{S}'_0(\mathbb{R}^d)$ as usual, via

$$\sigma_h(f) = \int_{\mathbb{R}^d} f(x)\, h(x) dx,$$

with

$$|\sigma_h(f)| \le \int_{\mathbb{R}^d} |h(x)|\, |f(x)| dx \le \|h\|_\infty \|f\|_{\boldsymbol{L}^1} \le (C\|h\|_\infty)\|f\|_{\boldsymbol{S}_0},$$

hence

$$\|\sigma_h\|_{\boldsymbol{S}_0'} \leq C\|h\|_\infty, \quad h \in \boldsymbol{C}_b(\mathbb{R}^d),$$

and consequently it has a Fourier transform.
By applying a dilation operator (or a matrix transformation in the multi-dimensional case) we can relate the general question to the standard setting of $\alpha = 2\pi$ (traditional normalization) or better (as we will do) to $\alpha = 1$. We will come back to the question of normalization (change of units on the time axis results in a change of the value alpha in real world applications) later (hopefully!, as time permits).
As we will see the Fourier transform of such a periodic function is a weighted discrete (potentially unbounded!) measure on $\mathbb{Z}$, with amplitudes which can be understood as limits of the situation described in Lemma trigpolFT00 50 (trig. polynomials with the pure frequencies involved being all $\mathbb{Z}$-periodic!).
If we start "from the other side" we consider $\boldsymbol{\ell}^1(\Lambda) \subset \boldsymbol{M}_b(\mathbb{R}^d)$ as a subalgebra of $\boldsymbol{M}_b(\mathbb{R}^d)$ and thus the Fourier transform is well defined via $\widehat{\mu}(s) = \mu(\chi_{-s})$.

ZperiodFT **Lemma 52.** *For $\nu \in \boldsymbol{\ell}^1(\alpha\mathbb{Z}^d)$ the Fourier transform (in the sense of $\boldsymbol{S}_0'(\mathbb{R}^d)$ resp. $\boldsymbol{M}_b(\mathbb{R}^d)$) is a periodic function in $\boldsymbol{C}_{ub}(\mathbb{R}^d)$, with periodicity group $1/\alpha\mathbb{Z}^d$. Since the Fourier transform is injective on $\boldsymbol{M}_b(\mathbb{R}^d)$ we can define*

$$\|\widehat{\mu}\|_{\mathcal{F}\boldsymbol{M}_b} := \|\mu\|_{\boldsymbol{M}_b}.$$

Consequently the range of $\boldsymbol{\ell}^1(\mathbb{Z}^d)$ can be viewed as a Banach algebra with respect to pointwise multiplication, called $\left(\boldsymbol{A}(\mathbb{T}^d), \|\cdot\|_{\boldsymbol{A}(\mathbb{T}^d)}\right)$ in the literature, the algebra of absolutely convergent Fourier series over the d-dimensional torus $\mathbb{T}^d$ (by identifying periodic functions with functions on the torus)[27] *with the norm*

$$\left\| \sum_{k\in\mathbb{Z}^d} c_k \chi_k \right\|_{\boldsymbol{A}} = \sum_{k\in\mathbb{Z}^d} |c_k| < \infty. \tag{264}$$ ATnorm1

Proof. It is clear that any of the characters constituting $\widehat{\nu} = \sum_{k\in\mathbb{Z}^d} c_k \chi_{-\alpha k}$ show the claimed invariance under translation by $\alpha^{-1}j$, for any $j \in \mathbb{Z}^d$. Hence the same is true for $\widehat{\nu}$, the (uniform) limit of these trigonometric polynomials. □

Remark 25. We will discuss the computation of the Fourier coefficients c_k for a given periodic function later on in more detail. For the moment let us just recall that in a "modern terminology" J.P. Fourier's claim from 1822 was that "every periodic function has a Fourier series representation" and that the Fourier transform is computed via certain integrals[28].

Given all the trouble with the problem of pointwise convergence it was a decisive step taken in the early 20th century (mostly by the Riesz brothers, also the time of Hilbert) to recognize that one should look not at summable Fourier coefficients, but at series

[27] This algebra $\left(\boldsymbol{A}(\mathbb{T}), \|\cdot\|_{\boldsymbol{A}}\right)$ over the torus $\mathbb{T}$ (resp. $\mathbb{U}$), i.e. for $d = 1$ is also often called *Wiener's algebra* of absolutely convergent Fourier series, we will try to avoid this terminology in these notes.

[28] Looking back, this was a very courageous claim at that time, and it took decades to find out what the proper concept of a function, what kind of integral one should use (first Riemann, then Lebesgue integral) and finally what kind of convergence is the correct one.

of square summable sequences, because they form a Hilbert space, and the periodic functions, endowed with a very natural scalar product are isometrically isomorphic to the sequence space $\boldsymbol{\ell}^2(\mathbb{Z}^d)$, with the scalar products

$$\langle \mathbf{x}, \mathbf{y} \rangle_{\boldsymbol{\ell}^2} := \sum_{k \in \mathbb{Z}^d} x_k \overline{y_k}, \tag{265}$$

ltspscalp

and (for the case of $\alpha = 1$):

$$\langle h_1, h_2 \rangle_{\boldsymbol{L}^2} := \int_{[0,1]^d} h_1(t) \overline{h_2(t)} dt \tag{266}$$

Ltzoscalp

Our next key result, known in the literature, is the so-called *Parseval's Theorem*. The usual approach is based on the following result:

Parseval

Theorem 10. *The pure frequencies $(\chi_k)_{k \in \mathbb{Z}^d}$ form an complete orthonormal basis for the Hilbert space $\boldsymbol{L}^2(\mathbb{T}^d)$, hence every element $f \in \boldsymbol{L}^2(\mathbb{T}^d)$ can be represented in a unique way as (unconditionally, norm convergent) series of the form*

$$f = \sum_{k \in \mathbb{Z}^d} \langle f, \chi_k \rangle \chi_k, \tag{267}$$

Hilbconf

with the (Fourier coefficients) being given as

$$c_k = \langle f, \chi_k \rangle = \int_{[0,1]^d} f(t) e^{-2\pi i k t} dt = \int_{[-1/2,1/2]^d} f(t) e^{-2\pi i k t} dt, \tag{268}$$

Fourcoffd

satisfying the Parseval Identity

$$\|f\|_{\boldsymbol{L}^2} = \|(c_k)_{k \in \mathbb{Z}^d}\|_{\boldsymbol{\ell}^2}. \tag{269}$$

Parseval00

Remark 26. The pure frequencies $(\chi_k)_{k \in \mathbb{Z}^d}$ are on the one hand the characters of the d-dimensional torus group $\mathbb{T}^d$. Hence the mapping

$$f \mapsto \int_{\mathbb{T}^d} f(u) \chi_k(u) du = \int_{\mathbb{T}^d} f(u) u^k du$$

is just the abstract Fourier transform for $\mathbb{T}^d$.

We will show that the Parseval identity, which results from its polarized version (by choosing $f^1 = f = f^2$), namely

$$\left\langle (c_k^1), (c_k^2) \right\rangle_{\boldsymbol{\ell}^2} = \int_{[-1/2,1/2]^d} f^1(t) \overline{f^2(t)} dt = \langle f^1, f^2 \rangle_{\boldsymbol{L}^2(\mathbb{T}^d)}. \tag{270}$$

ParsevPolar

We will derive this result from the (FRF), by making appropriate choices and taking a limit ($\rho \to 0$). For simplicity we work with $d = 1$.

As in the classical case one verifies the validity of (ParsevPolar 270) for sequences in the dense subspace $\boldsymbol{\ell}^1 \subset \boldsymbol{\ell}^2$ (resp. the corresponding functions in $\left(\boldsymbol{A}(\mathbb{T}), \|\cdot\|_{\boldsymbol{A}}\right)$) and then extends the validity by taking limits on both sides keeping the abstract principle (for bounded bilinear mappings) in mind.

Thus we are looking at the following situation. We start with two sequences $(c_k^1)_{k\in\mathbb{Z}}$ and $(c_k^2)_{k\in\mathbb{Z}}$, in $\boldsymbol{\ell}^1(\mathbb{Z})$, and would like to realize that left hand side as a limit of expression of the form $\int_{\mathbb{R}} f^1(x)\widehat{g_\rho}(x)dx$.
First we turn the sequence (c_k^1) into a function in $f \in \boldsymbol{C}_0(\mathbb{R})$ by putting

$$f = \sum_{k\in\mathbb{Z}} c_k^1 T_k \operatorname{SINC} = \left(\sum_{k\in\mathbb{Z}} c_k^1 \delta_k\right) \star \operatorname{SINC}$$

with

$$\|f\|_\infty \le \left\|\sum_{k\in\mathbb{Z}} c_k^1 \delta_k\right\|_{\boldsymbol{M}_b} \|\operatorname{SINC}\|_\infty \le (C\|\operatorname{SINC}\|_\infty) \left\|(c_k^1)\right\|_{\boldsymbol{\ell}^1(\mathbb{Z})}.$$

Since convolution on the time-side corresponds to a pointwise multiplication on the Fourier side, we have for, observing that $\mathcal{F}(\operatorname{SINC}) = \mathcal{F}^{-1}(\operatorname{SINC}) = \mathbf{1}_{[-1/2,1/2]}$ (by Plancherel's Theorem): $\mathcal{F}(f)(t) = h^1(t) \cdot \mathbf{1}_{[-1/2,1/2]}(t)$, for the periodic function $h^1(t) = \sum_{k\in\mathbb{Z}} c_k^1 \chi_k(t) \in \boldsymbol{A}(\mathbb{T})$. We also set $h^2(t) = \sum_{k\in\mathbb{Z}} c_k^2 \chi_k(t) \in \boldsymbol{A}(\mathbb{T})$.
Due to the fact that $\operatorname{SINC}(0) = 1$ and $\operatorname{SINC}(k) = 0$ for $k \in \mathbb{Z} \setminus \{0\}$ it is clear that the sampling values of f over $\mathbb{Z}$ satisfy: $f(k) = c_k^1$ for $k \in \mathbb{Z}$.
In order to apply the scalar product of the form $\left\langle (c_k^1), (c_k^2) \right\rangle_{\boldsymbol{\ell}^2}$ we would like to apply the bounded (discrete) measure $\nu := \sum_{k\in\mathbb{Z}} \overline{c_k^2} \delta_k$ to this function f, because then one has:

$$\nu(f) = \sum_{k\in\mathbb{Z}} \overline{c_k^2} \delta_k(f) = \sum_{k\in\mathbb{Z}} \overline{c_k^2} f(k) = \sum_{k\in\mathbb{Z}} c_k^1 \overline{c_k^2} = \langle c_k^1, c_k^2 \rangle_{\boldsymbol{\ell}^2(\mathbb{Z})}.$$

We can approximate the weighted sum of Dirac measures ν by choosing Dirac sequences from $\boldsymbol{S}_0(\mathbb{R})$, via the family

$$\widehat{g_\rho} = \sum_{k\in\mathbb{Z}} \overline{c_k^2} (T_k \operatorname{St}_\rho \operatorname{SINC}^2),$$

which is the Fourier transform of

$$g_\rho = \sum_{k\in\mathbb{Z}} \overline{c_k^2} M_k \mathrm{D}_\rho \Delta \in \boldsymbol{S}_0(\mathbb{R}).$$

For fixed $\rho > 0$ this function belongs to $\boldsymbol{S}_0(\mathbb{R}^d)$ and hence the Fourier transform is well defined. The family (g_ρ) is bounded in $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ and since we have $\mathcal{F}(\operatorname{SINC}^2)(0) = \Delta(0) = 1$ we have for each $k \in \mathbb{Z}$: $\lim_{\rho\to 0} T_k \operatorname{St}_\rho \operatorname{SINC}^2 = \delta_k$ in the w^* sense, i.e.

$$\lim_{\rho\to 0} \left[\sum_{k\in\mathbb{Z}} \overline{c_k^2} T_k \operatorname{St}_\rho \operatorname{SINC}^2\right](f) \to \sum_{k\in\mathbb{Z}} \overline{c_k^2} f(k)$$

for $f \in \boldsymbol{C}_0(\mathbb{R})$, in particular for f as chosen above, hence the limit of the left hand side in (FRF) is, as planned:

$$\lim_{\rho\to 0} \int_{\mathbb{R}} f(t)\widehat{g_\rho}(t)dt = \sum_{k\in\mathbb{Z}} \overline{c_k^2} f(k) = \sum_{k\in\mathbb{Z}} c_k^1 \overline{c_k^2} = \langle c_k^1, c_k^2 \rangle_{\boldsymbol{\ell}^2(\mathbb{Z})}. \qquad (271) \quad \boxed{\texttt{FRFL1}}$$

If one looks at the right hand side of (FRF) it is of the form

$$\int_{-\infty}^{\infty} \widehat{f}(s) g_\rho(s) ds = \int_{-\infty}^{\infty} h^1(s) \cdot \mathbf{1}_{[-1/2,1/2]}(s) \cdot g_\rho(s) ds.$$

Since obviously $(M_k \mathrm{D}_\rho \Delta)(s) \to \chi_k(s)$ for $\rho \to 0$ we have the limit (uniformly over compact sets)

$$\lim_{\rho \to 0} g_\rho(s) = h^2(s), \quad \text{uniformly over compact sets!}$$

and thus

$$\lim_{\rho \to 0} \int_{-\infty}^{\infty} \widehat{f}(s) g_\rho(s) = \int_{[-1/2,1/2]} h^1(s)\overline{h^2(s)}ds. \tag{272}$$ FRFR1

Combining the two limits, namely (271) and (272) we see that the (FRF) implies identity formulated in the theorem.
PS: If there are convergence issues for a detailed study of the arguments used above one even can reduce the discussion to the dense subspace of "finite sequences" ((c_k^i) and (c_k^2)).

29 Poisson's Formula

The proof of *Poisson's Formula* is one of the crucial steps for the further development, including the proof of the Shannon Sampling Theorem. We write $\sqcup\!\sqcup = \sum_{k \in \mathbb{Z}} \delta_k$ (also known as Dirac comb). Let us first do a simple version for $\mathscr{G} = \mathbb{R}$:

PoissonR **Theorem 11.**

$$\sum_{k \in \mathbb{Z}} f(k) = \sum_{n \in \mathbb{Z}} \widehat{f}(n), \quad \forall f \in \boldsymbol{S}_0 \tag{273}$$ PoissonForm

with absolute convergence on both sides.

PoissonShahR **Corollary 7.** *For $\sigma = \sqcup\!\sqcup \in \boldsymbol{S}_0'(\mathbb{R})$ one has: $\mathcal{F}(\sqcup\!\sqcup) = \sqcup\!\sqcup$, or*

$$\sum_{k \in \mathbb{Z}} \chi_k = \sum_{n \in \mathbb{Z}} \delta_n,$$

respectively in an engineering setting:

$$\sum_{k \in \mathbb{Z}} e^{2\pi i k t} = \sum_{n \in \mathbb{Z}} \delta_n(t).$$

All the expressions in the left and side a convergent in the w^*-sense for $\boldsymbol{S}_0'(\mathbb{R})$, i.e. after application to any given $f \in \boldsymbol{S}_0(\mathbb{R})$ one observes convergence[29].

Remark 27. For the discrete setting we can observe by a simple MATLAB argument that the FFT of a sequence of unit vectors, given by the command

```
sha = zeros(1,n); sha(1:a:n) = 1; f2sp(sha);
```

shows that the FFT/DFT of such a Dirac comb is up to normalization just another Dirac comb (at distance n/a) if and only if $a|n$, i.e. if a divides the signal length n, and then n/a is again a natural number. The extreme cases are $a = 1$ (i.e. the constant sequence), whose DFT is the unit vector ne_0 (resp. $n\mathbf{e}_1$ in MATLAB notation), or $a = n$ (corresponding to δ_0, with DFT$=$ *const.* $= n$.

[29] In fact, unconditional convergence, i.e. independent of the particular order in the summation process.

Let us give a proof of this theorem.

Proof. We start from the function $f \in \boldsymbol{S}_0(\mathbb{R})$ and produce a $\mathbb{Z}$-periodic version by convolving f by the Dirac comb $⧢ = \sum_{k\in\mathbb{Z}} \delta_k$, i.e. by forming

$$f_{per}(t) = [\left(\sum_{k\in\mathbb{Z}} \delta_k\right) \star f](t) = \sum_{k\in\mathbb{Z}} f(t-k). \tag{274}$$ fperdef

Since $\left(\boldsymbol{S}_0(\mathbb{R}), \|\cdot\|_{\boldsymbol{S}_0}\right) \hookrightarrow \left(\boldsymbol{W}(\mathbb{R}), \|\cdot\|_{\boldsymbol{W}}\right)$ the infinite sum of translates is uniformly convergent (even absolutely in $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$) over any ball of radius $R > 0$! Hence f_{per} is a (uniformly) continuous and bounded function on $\mathbb{R}$ with $\|f_{per}\|_\infty \leq C\|f\|_{\boldsymbol{W}(\mathbb{R})} \leq C_1\|f\|_{\boldsymbol{S}_0(\mathbb{R})}$. And of course we have $T_k f_{per} = f_{per}, \forall k \in \mathbb{Z}$. The left hand side of (273) is just the value $f_{per}(0)$. PoissonForm
On the other hand we have to verify that

$$f_{per}(t) = \sum_{k\in\mathbb{Z}} c_k \chi_k(t) = \sum_{k\in\mathbb{Z}} c_k e^{2\pi i k t},$$

where $c_k = \int_{[-1/2,1/2]} f_{per}(t) e^{-2\pi i k t} dt$ is the Fourier coefficient of f_{per}. But now we can use the $Z-periodicity$ of the pure frequencies $\chi_k, k \in \mathbb{Z}$, in order to make the following important observation: The Fourier coefficients of f_{per} are just the sampling values of $\widehat{f}$ at $k \in \mathbb{Z}$. This is verified as follows: split the integral over $(-\infty, \infty)$ into segments:

$$\widehat{f}(n) = \int_{-\infty}^{\infty} f(t) e^{2\pi i n t} dt = \sum_{k\in\mathbb{Z}} \int_{[k,k+1]} f(t) e^{2\pi i n t} dt = \sum_{k\in\mathbb{Z}} \int_{[0,1]} f(t-k) e^{2\pi i n (t-k)} dt,$$

and observe that $e^{2\pi i n (t-k)} = e^{2\pi i n t}$ for any $k \in \mathbb{Z}$, so that

$$\widehat{f}(n) = \int_{[-1/2,1/2]} f_{per}(t) e^{2\pi i n t} dt = c_n.$$

Since $f \in \boldsymbol{S}_0(\mathbb{R})$ implies $\widehat{f} \in \boldsymbol{S}_0(\mathbb{R}^d) \subset \boldsymbol{W}(\mathbb{R})$ and thus $(c_n)_{n\in\mathbb{Z}} = (\widehat{f}(n))_{n\in\mathbb{Z}} \in \boldsymbol{\ell}^1(\mathbb{Z})$ we find that $f_{per} \in \boldsymbol{A}(\mathbb{T})$ and thus the Fourier series representation is valid in the pointwise sense,

$$f_{per}(y) = \sum_{n\in\mathbb{Z}} c_n \chi_n(y),$$

because the series is uniformly convergent in $\left(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\right)$. Specializing this Fourier series representation of f_{per} to the case $y = 0$ one obtains (since $e^{i0} = 1$!)

$$\sum_{k\in\mathbb{Z}} f(k) = f_{per}(0) = \sum_{n\in\mathbb{Z}} c_n \chi_n(0) = \sum_{n\in\mathbb{Z}} \widehat{f}(n).$$

□

The same argument can be applied for general lattices $\Lambda \lhd \mathbb{R}^d$. In this case the so-called *orthogonal lattice* appears, i.e.

$$\Lambda^\perp = \{\chi_s \in \widehat{\mathbb{R}^d}, \chi_s(\lambda) = 1 \ \forall \lambda \in \Lambda\}.$$

The set $\Lambda^\perp$ is always a group. For a lattice of the form $\Lambda = \boldsymbol{A} \star \mathbb{Z}^d$ the orthogonal lattice appears as $\Lambda^\perp = (\boldsymbol{A}^t)^{-1} \star \mathbb{Z}$. For $\Lambda = \alpha\mathbb{Z}^d$ this gives $\Lambda^\perp = 1/\alpha\mathbb{Z}^d$.
The general version of Poisson's formula then reads.

PoissLam00 **Theorem 12.** *For any lattice $\Lambda \lhd \mathbb{R}^d$ there exists a constant C_Λ such that*

$$\sum_{\lambda\in\Lambda} f(\lambda) = C_\Lambda \sum_{\lambda^\perp\in\Lambda^\perp} \widehat{f}(\lambda^\perp), \quad f \in \boldsymbol{S}_0, \tag{275}$$ PoissLam

with both sides absolutely convergent. Alternatively one has

$$\mathcal{F}(\sqcup\!\sqcup_\Lambda) = C_\Lambda \sqcup\!\sqcup_{\Lambda^\perp}.$$

Remark 28. For $\Lambda = \alpha\mathbb{Z}^d$ we have $\Lambda^\perp = (1/\alpha)\mathbb{Z}$ and $C_\Lambda = \alpha^{-d}$.

Demonstration concerning Shah Distributions (Dirac Combs):
As an illustration of the one-dimensional situation let us look at the following plot:

rband l01.pdf

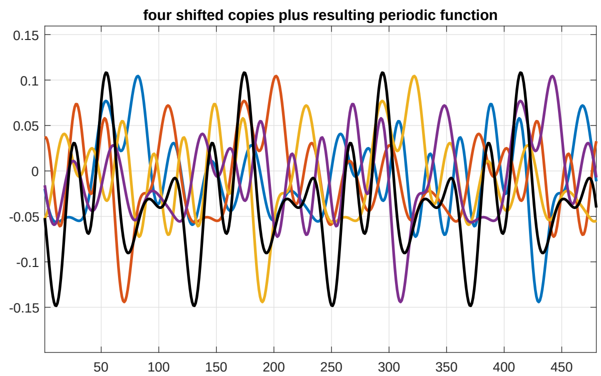

Abbildung 47: fperbandl01.pdf

And the same for a discrete, finite lattice with 9 points:

iodFFT2a.jpg

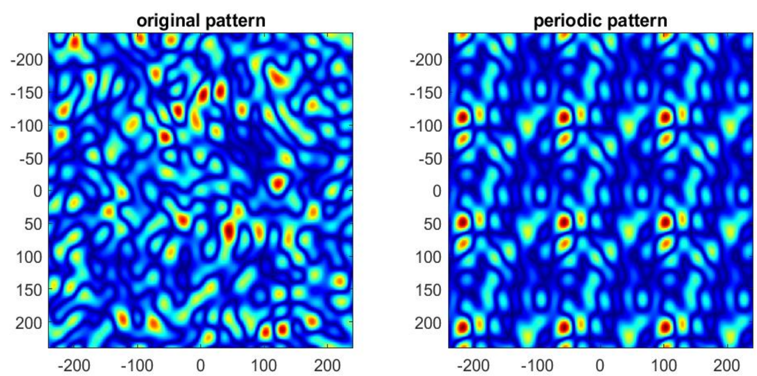

Abbildung 48: periodFFT2a.jpg

As an illustration of the one-dimensional situation let us look at the following plot:
And the same for a discrete, finite lattice with 9 points:

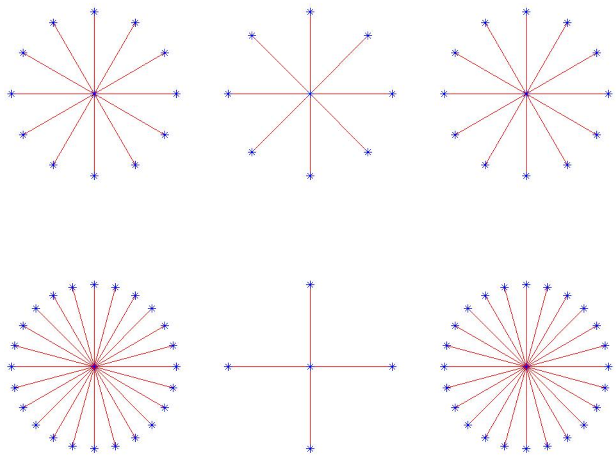

Abbildung 49: shahdem12.pdf:
For a lattice the values which appear, as the sampling values of the trigonometric polynomials $\chi_{-\lambda} = \mathcal{F}(\delta_\lambda), \lambda \in \Lambda$ take the values $\#\Lambda$ (over $\Lambda^\perp$) and zero otherwise: The sum over values which are unit roots of order k, for some $k > 1, k \in \mathbb{N}$.

Remark 29. The absolute convergence of both sides in Poisson's formula is not sufficient for the validity of (PoissonForm 273). But all the sufficient conditions so far imply that $f \in \boldsymbol{S}_0(\mathbb{R})$ (see **gr96:** [58]). All the counter examples rely on the possibility that a periodic function on $\mathbb{R}$ may have a convergent Fourier series at a given point which does not represent the function value at that point.

We will not give a detailed account for this general formula (which is of course specifically useful for any dimension and general lattices Λ), but in principle the arguments for the general case are following the same path.

1. The Dirac-comb $\sqcup\!\sqcup_\Lambda$ belongs to $\boldsymbol{W}(\boldsymbol{M}, \boldsymbol{\ell}^\infty)(\mathbb{R}^d) \subset \boldsymbol{S}_0'(\mathbb{R}^d)$;
2. Convolving $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ by $\sqcup\!\sqcup_\Lambda$ gives a Λ-periodic version $f_{per} = \sqcup\!\sqcup_\Lambda \star f$ of f;
3. The Fourier coefficients of f_{per} (viewed as a function on the fundamental domain of $\mathbb{R}/\Lambda^\perp$) are just the samples $(\widehat{f}(\lambda^\perp))_{\lambda^\perp \in \Lambda^\perp}$;
4. Because $\widehat{f} \in \boldsymbol{S}_0(\mathbb{R}^d)$ the sampling sequence $(\widehat{f}(\lambda^\perp))_{\lambda^\perp \in \Lambda^\perp}$ belongs to $\boldsymbol{\ell}^1(\Lambda^\perp)$,
5. Writing the Fourier series expansion of f_{per} for $t = 0$ implies the validity of the general Poisson formula.

rbandl01.eps

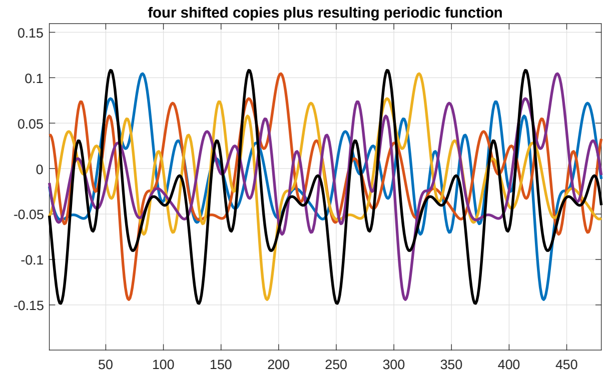

Abbildung 50: fperbandl01.eps

iodFFT2a.jpg

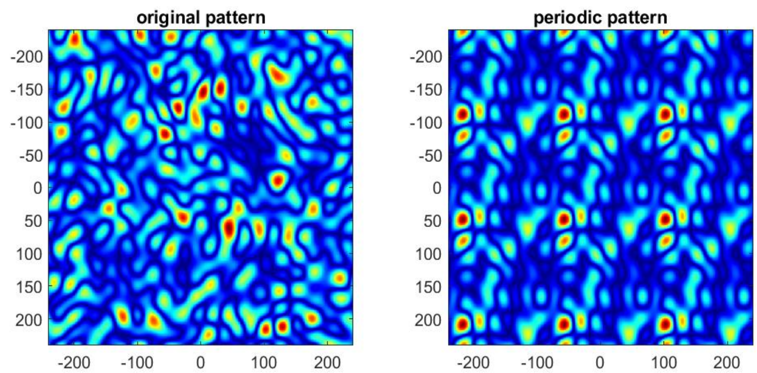

Abbildung 51: periodFFT2a.jpg

30 Wiener algebra and its dual

30.1 Important properties of $\boldsymbol{W}(\mathbb{R}^d)$ and its dual

The Wiener algebra $\boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d)$, which can be obtained using BUPUs as absolutely convergent decompositions, or by the more general *atomic representations* (also absolutely convergent in $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$) has a number of important properties which we will study next. In a second step we will list some important properties of the dual space. Unfortunately either of them is *not invariant under the Fourier transform*, which can be circumvented to some extent by looking at $\boldsymbol{W}(\mathbb{R}^d) \cap \mathcal{F}(\boldsymbol{W}(\mathbb{R}^d))$ (similar to the Fourier invariant space $(\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1)(\mathbb{R}^d)$).
Compared to the space $\boldsymbol{L}^1 \cap \boldsymbol{C}_0(\mathbb{R}^d)$ the Wiener algebra has the important property that it allows sampling, i.e. for any *lattice* $\Lambda = \boldsymbol{A} \star \mathbb{Z}^d$ (e.g. a hexagonal lattice in $\mathbb{R}^2$, or simply $\alpha\mathbb{Z} \lhd \mathbb{R}$ for some $\alpha > 0$) one can find an $\boldsymbol{\ell}^1$-estimate for the "sequence" of sampling values of $f \in \boldsymbol{W}(\mathbb{R}^d)$.

RDinSOP1.jpg

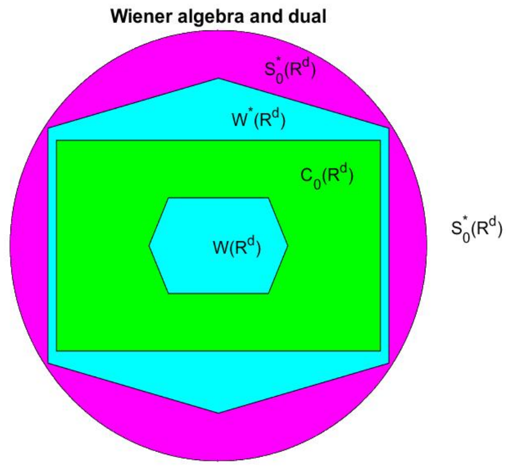

Abbildung 52: WRWRDinSOP1.jpg

samplWRd **Proposition 15.** *Given any lattice $\Lambda \lhd \mathbb{R}^d$ there exists a constant $C_\Lambda > 0$ such that*

$$\|(f(\lambda))_{\lambda\in\Lambda}\|_{\ell^1(\Lambda)} = \sum_{\lambda\in\Lambda} |f(\lambda)| \le C_\Lambda \|f\|_{\boldsymbol{W}(\mathbb{R}^d)}, \quad f \in \boldsymbol{W}(\mathbb{R}^d). \tag{276}$$ liLamsamp

Proof. Again the atomic characterization will be helpful, as it allows to reduce the estimate for general elements in $\boldsymbol{W}(\mathbb{R}^d)$ to the *uniform estimates* with respect to the position of the atoms. So let us assume that we have a function $f \in \boldsymbol{C}_c(\mathbb{R}^d)$ with $\mathrm{supp}(f) \subset B_R(x)$ for some $x \in \mathbb{R}^d$ and $\|f\|_\infty = 1$.
For our intuition let us take again a look at figure coverballs1.jpg (maybe imagine that some non-singular matrix $\boldsymbol{A}$ is applied to this situation, deforming the red balls into red ellipses). [MATLAB experiment should be done to produce such a picture!]
The final estimate (liLamsamp 276) then follows by simple summation over the atoms according to Prop. WAliatomic 10. □

Remark 30. The same argumentation is valid for any discrete set of points (typically a countable set of points) $X = (x_k)_{k=1}^\infty$ in $\mathbb{R}^d$ which is *separated* i.e. with a *positive minimal distance*:

$$|x_k - x_j| \ge \eta > 0, \quad \text{for } k \ne j. \tag{277}$$ mindist1

If course one can take finite unions of such sets, which are called *relatively separated*. They can be characterized by the property that for some $r_1 > 0$ one can control the number of points in any ball $B_{r_1}(x)$ of radius r_1, i.e.

$$\sup_{x\in\mathbb{R}^d} \#\{k \,|\, x \in B_{r_1}(x)\} < \infty. \tag{278}$$ relsepdef

Since balls of any radius $R > 0$ can be covered by a fixed number of balls of radius r_1 this condition is then valid for *any* $R > 0$.[30]

[30] It is not hard to show, but requires some work, to verify that any set satisfying (relsepdef 278) is the finite union of sets which are separated.

DinSOPFT.jpg

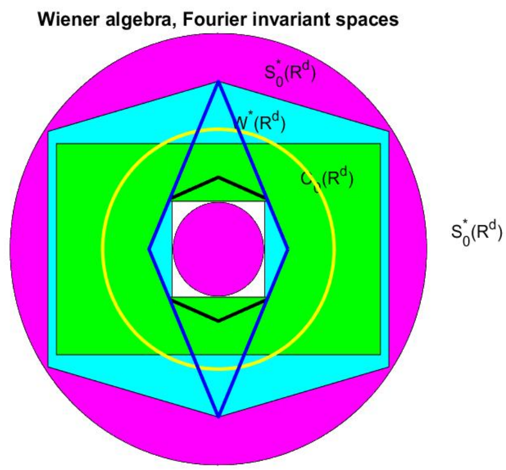

Abbildung 53: WRWRDinSOPFT.jpg

Remark 31. For any relatively separated set X the corresponding Dirac comb $\sqcup\!\sqcup_X = \sum_{k=1}^{\infty} \delta_{x_k}$ is in the Wiener amalgam space $\boldsymbol{W}(\boldsymbol{M}, \boldsymbol{\ell}^\infty)(\mathbb{R}^d)$. In fact, the quantity expressed relsepdef by (278) is equivalent (for any $R > 0$) to the norm of $\sqcup\!\sqcup_X$ as given by any BUPU $\Psi = (\psi_i)_{i \in I}$.

$$\sup_{i \in I} \|\sqcup\!\sqcup_X \cdot \psi_i\|_{\boldsymbol{M}_b}. \tag{279}$$ WMinfBUP1

Recalling that we have for any $h \in \boldsymbol{C}_b(\mathbb{R}^d)$: $\delta_x \cdot h(f) := f(x)\delta_x$, because

$$(\delta_x \cdot h)(f) = \delta_x(hf) = h(x)f(x) = h(x)\delta_x(f), \quad f \in \boldsymbol{C}_0(\mathbb{R}^d),$$

we collect a few simple observations:

relsepprop1 **Lemma 53.** *Given a relatively separated set $X = (x_k)_{k=1}^\infty$ one has:*

1. *For any $h \in \boldsymbol{C}_b(\mathbb{R}^d)$ one has*

$$h \cdot \sqcup\!\sqcup_X = \sum_{k=1}^{\infty} h(x_k)\delta_{x_k} \in \boldsymbol{W}(\boldsymbol{M}, \boldsymbol{\ell}^\infty)(\mathbb{R}^d),$$

with the estimate

$$\|h \cdot \sqcup\!\sqcup_X\|_{\boldsymbol{W}(\boldsymbol{M},\boldsymbol{\ell}^\infty)(\mathbb{R}^d)} \leq \|h\|_\infty \|\sqcup\!\sqcup_X\|_{\boldsymbol{W}(\boldsymbol{M},\boldsymbol{\ell}^\infty)(\mathbb{R}^d)}.$$

2. *For $f \in \boldsymbol{W}(\mathbb{R}^d) = \boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d)$ one has*

$$\|\sqcup\!\sqcup_X \cdot f\|_{\boldsymbol{M}_b} \leq C\|f\|_{\boldsymbol{W}(\boldsymbol{C}_0,\boldsymbol{\ell}^1)(\mathbb{R}^d)}. \tag{280}$$ sampWRdpw

3. *For $h \in \boldsymbol{W}(\mathbb{R}^d) = \boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d)$ one has*

$$\|\sqcup\!\sqcup_X \star f\|_\infty = \left\| \sum_{k=1}^{\infty} T_{x_k} f \right\|_\infty \leq C\|k\|_{\boldsymbol{W}(\boldsymbol{C}_0,\boldsymbol{\ell}^1)(\mathbb{R}^d)}. \tag{281}$$ sampWRdconv

Proof. We leave the proof to the reader resp. will provide more detailed references later on. The arguments are elementary, relying on geometrical arguments and are valid over LCA groups. Let us just give the hint that the important equation (sampWRdconv 281), can be either obtained using the atomic representation of functions in $\boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d)$ or by just observing that one has

$$\mu \star f(y) = \mu(T_u(f^{\checkmark}), \quad f \in \boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d), \mu \in \boldsymbol{W}(\boldsymbol{M}, \boldsymbol{\ell}^\infty)(\mathbb{R}^d),$$

simply by the fact that $\boldsymbol{W}(\boldsymbol{M}, \boldsymbol{\ell}^\infty)(\mathbb{R}^d)$ can be identified with the dual space of Wiener's algebra $\boldsymbol{W}(\mathbb{R}^d) = \Big(\boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d), \, \|\cdot\|_{\boldsymbol{W}}\Big)$. $\square$

Since we have $\Big(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\Big) \hookrightarrow \Big(\boldsymbol{W}(\mathbb{R}^d), \, \|\cdot\|_{\boldsymbol{W}}\Big)$ we note that we have for any $f \in \boldsymbol{S}_0$:

$$\sqcup\!\sqcup_\Lambda \cdot f \in \boldsymbol{M}_d(\mathbb{R}^d) \quad \text{and} \quad \sqcup\!\sqcup_\Lambda \star f \in \boldsymbol{C}_{ub}(\mathbb{R}^d),$$

together with the corresponding norm inequalities.

31 $C_b(\mathbb{R}^d)$ in the dual of L^1

Using measure theory - as we mentioned - the dual space to $\left(L^1(\mathbb{R}^d),\, \|\cdot\|_1\right)$ is identified with $\left(L^\infty(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$, the space of *essentially bounded* measurable functions (in fact: equivalence classes of such functions), with the essential supremum as norm. It is the best upper bound for $|h(x)|, x \in \mathbb{R}^d$, if one allows exceptions on sets of measure zero. The identification is the "usual one", namely via (identification with a so-called Radon measure μ_h):

$$h \leftrightarrow \mu_h : \mu_h(f) = \int_{\mathbb{R}^d} [f(x)h(x)]dx, f \in L^1(\mathbb{R}^d).$$

Of course, based on measure theory, one has to show that the pointwise product fh is well defined (it is indeed a product of measurable functions, and thus itself a measurable function) and it is integrable in the sense of Lebesgue.
There is no problem as long as $h \in C_b(\mathbb{R}^d)$. In fact we have:

CbinLiPRd **Lemma 54.** *The embedding of* $\left(C_b(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$ *into* $\left((L^1)^*(\mathbb{R}^d), \|\cdot\|_{(L^1)^*(\mathbb{R}^d)}\right)$ *via* $h \mapsto \mu_h$ *is isometric, i.e. we have*

$$\|h\|_\infty = \|\mu_h\|_{(L^1)^*(\mathbb{R}^d)}\,, \quad h \in C_b(\mathbb{R}^d). \tag{282}$$ CbLiPisom

In this sense $\left(C_b(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$ *is a closed subspace of* $\left((L^1)^*(\mathbb{R}^d), \|\cdot\|_{(L^1)^*(\mathbb{R}^d)}\right)$, *and* $\left(C_0(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$ *can be viewed as the closure of* $C_c(\mathbb{R}^d)$ *(rather* $\mu_k, k \in C_c(\mathbb{R}^d)$*) inside of* $\left((L^1)^*(\mathbb{R}^d), \|\cdot\|_{(L^1)^*(\mathbb{R}^d)}\right)$. *This space also coincides with the dual space to* $C_c(\mathbb{R}^d)$, *endowed with the* $\|\cdot\|_1$*-norm (using R-integrals).*

CbinSOP **Corollary 8.** *Due to the chain of dense, continuous embeddings:*

$$\left(S_0(\mathbb{R}^d), \|\cdot\|_{S_0}\right) \hookrightarrow \left(W(\mathbb{R}^d),\, \|\cdot\|_W\right) \hookrightarrow \left(L^1(\mathbb{R}^d),\, \|\cdot\|_1\right)$$

the relationship (282) also allows to view $C_b(\mathbb{R}^d)$ *as a subspace of* $W(\mathbb{R}^d)^*$ *or* $S_0'(\mathbb{R}^d)$.

Note: While $\left(C_b(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$ can be viewed as a closed subspace of $\left((L^1)^*(\mathbb{R}^d), \|\cdot\|_{(L^1)^*(\mathbb{R}^d)}\right)$ it is not closed in the larger spaces, but still w^*-dense. This will be verified later on. In fact, even $S_0(\mathbb{R}^d)$ is w^*-dense in $S_0'(\mathbb{R}^d)$.

Proof. It will be enough to understand that any $h \in C_b(\mathbb{R}^d)$ defines a bounded linear function μ_h on $C_c(\mathbb{R}^d)$, endowed with the L^1-norm, but this follows easily from

$$|\mu_h(k)| = \int_{\mathbb{R}^d} h(x)k(x)dx \leq \|h\|_\infty \int_{\mathbb{R}^d} |k(x)|dx = \|h\|_\infty \|k\|_{L^1}, \quad k \in C_c(\mathbb{R}^d), \tag{283}$$ CbLiPestim1

ensuring that $\|\mu_h\|_{(L^1)^*} \leq \|h\|_\infty$.
In order to establish to equality of norms we have to show that the converse is true in the following sense: For any $\varepsilon > 0$ there exists $k \in C_c(\mathbb{R}^d)$ with $\|k\|_\infty = 1$ but $|\mu_h(k)| \geq \|h\|_\infty - \varepsilon$. This can be achieved easily by choosing a bump function $k \in C_c(\mathbb{R}^d)$ which is concentrated near a point where h almost achieves its maximal value. Formally it requires a little bit of analysis and text.

First we can find (for the given $h \in \boldsymbol{C}_b(\mathbb{R}^d)$ and $\varepsilon > 0$) some x_0 such that $h(x_0) > \|h\|_\infty - \varepsilon/4$. In fact, due to the continuity of h we can even ensure that there is a ball $B_s(x_0)$, with some positive radius $s > 0$ such that $|h(x)| \geq \|h\|_\infty - \varepsilon/2$ for all $x \in B_s(x_0)$. If we assume that $h(z) \geq 0$ it is easy to choose any nonzero $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ with $\mathrm{supp}(k) \subseteq B_s(x_0)$, also non-negative, and normalized to have $\|k\|_{\boldsymbol{L}^1} = 1$. The action of μ_h then satisfies

$$\int_{\mathbb{R}^d} k(x)|h(x)|dx = \int_{B_s(x_0)} h(x)k(x)dx \geq (\|h\|_\infty - \varepsilon/2) \int_{B_s(x_0)} |k(x)|dx = \|h\|_\infty - \frac{\varepsilon}{2}. \quad (284)$$ muhlowest1

For the general case of real or complex-valued functions h one has to take phase factors into account. First of all we may assume that $h(x_0) > 0$, otherwise h has to be multiplied by the complex number $|h(x_0)|/h(x_0)$, which is of absolute value 1 and thus does not change the sup-norm of h.
Secondly, we note that the phase $h(x)/|h(x)|$ of a continuous function, which does not vanish on $B_s(x_0)$ and takes the value 1 at x_0 will be very close to 1 in a sufficiently small neighborhood of x_0. Thus, if necessary, by replacing the given ball $B_s(x_0)$ by an even smaller on (we denote the new ball again by $B_s(x_0)$) we can assume that

$$\max\{|h(x) - |h(x)||, |x - x_0| \leq s\} \leq \varepsilon/2$$

so that we have

$$\left|\mu_h(k) - \int_{\mathbb{R}^d} k(x)|h(x)|dx\right| \leq \varepsilon/2 \cdot \|k\|_{\boldsymbol{L}^1} = \varepsilon/2.$$

Combining this last estimate with (284) we achieve the promised lower estimate. muhlowest1
It is easier to understand the multiplication by a phase factor by a simple plot: □

icationk.pdf

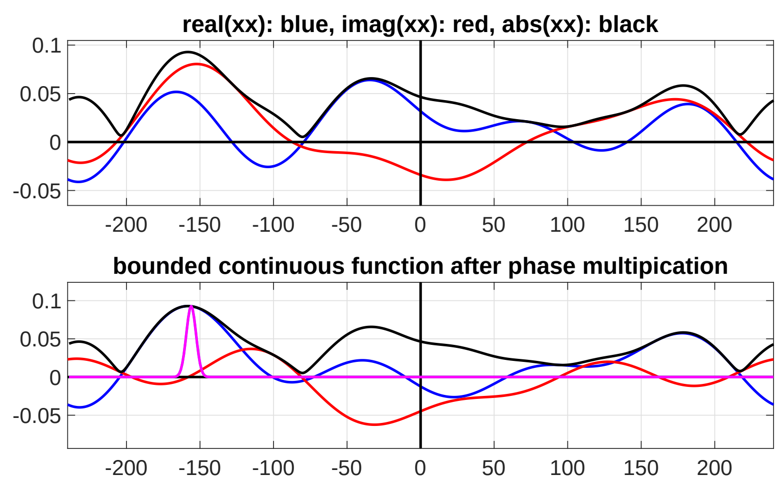

Abbildung 54: phasemultiplicationk.pdf:
Upper plot: real (blue) and imaginary (red) part of f, with $|f|$ in black, taking maximum in the left half. Lower plot: after multiplication with phase factor, $|f|$ is unchanged, real part takes the (real, positive) maximum there, imaginary part equals zero.
Testing bump-function in magenta positioned there, i.e. near -160.

32 Preparing for the Shannon Sampling Theorem

Int01bdl.eps

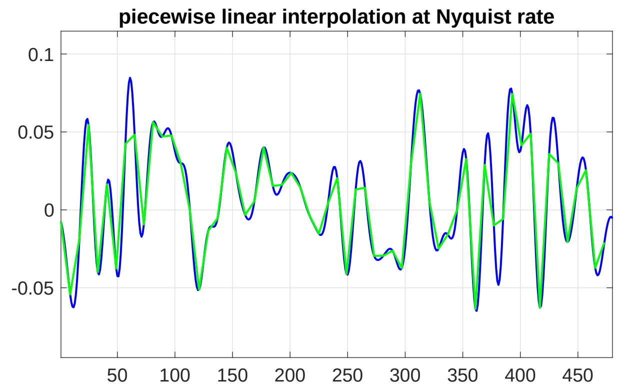

Abbildung 55: pwlinInt01bdl.eps.
blue: real part of band-limited function, green: piecewise linear interpolation, near Nyquist rate, i.e. a=8. Showing quite large deviations!

33 Uniqueness of the extended Fourier transform

Let us recall that $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$ is a closed, proper subspace of $\big(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\big)$. The constant $h(x) \equiv 1$ does not belong to $\boldsymbol{C}_0(\mathbb{R}^d)$, but there are bounded approximate units, e.g. families of functions $\mathrm{D}_\rho(h), \rho \to 0$ if dilated versions of any function $h \in \boldsymbol{C}_0(\mathbb{R}^d)$ with $h(0) = 1$. Since $\|\mathrm{D}_\rho h\|_\infty = \|h\|_\infty$ boundedness is obvious.
In a similar way we can embed $\boldsymbol{S}_0(\mathbb{R}^d)$ into $\boldsymbol{S}_0'(\mathbb{R}^d)$ via $f \mapsto \sigma_f$, but $\boldsymbol{S}_0(\mathbb{R}^d)$ is not dense in $\boldsymbol{S}_0'(\mathbb{R}^d)$ with respect to the norm convergence (similar to the lack of density of $\boldsymbol{C}_c(\mathbb{R}^d)$ in $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b}))$. However, we can work with w^*-convergence to overcome this problem.
Let us recall: A sequence $(\sigma_n)_{n=1}^\infty$ is w^*-convergent to σ_0 if (and only if)

$$\lim_{n\to\infty} \sigma_n(f) = \sigma_0(f), \quad \forall f \in \boldsymbol{S}_0(\mathbb{R}^d). \tag{285}$$ wstconv05

We will write in this situation

$$\sigma_0 = w^*\text{-}\lim \sigma_n.$$

wstconv07 **Lemma 55.** *The (extended) Fourier transform preserves w^*-convergence, i.e.*

$$\sigma_0 = w^*\text{-}\lim \sigma_n \quad \Rightarrow \quad \mathcal{F}(\sigma_0) = w^*\text{-}\lim \mathcal{F}(\sigma_n).$$

Proof. In order to verify the claimed convergence we just have to observe that

$$\mathcal{F}(\sigma_n)(f) = \sigma_n(\mathcal{F}(f))) \to \sigma_0(\mathcal{F}(f)) = \mathcal{F}(\sigma_0)(f)$$

because $\mathcal{F}(f) \in \boldsymbol{S}_0(\mathbb{R}^d)$ is just another element in $\boldsymbol{S}_0(\mathbb{R}^d)$. □

Later on we will discuss that $\boldsymbol{S}_0(\mathbb{R}^d)$ (and hence $\boldsymbol{L}^1(\mathbb{R}^d)$ or $\boldsymbol{M}_b(\mathbb{R}^d)$) are w^*-dense in $\boldsymbol{S}_0'(\mathbb{R}^d)$. Hence we can approximate the (well defined) generalized function (or *mild distribution*) approximately by computing ordinary Fourier transforms of nice functions in $\boldsymbol{S}_0(\mathbb{R}^d)$.

Int02bdl.pdf

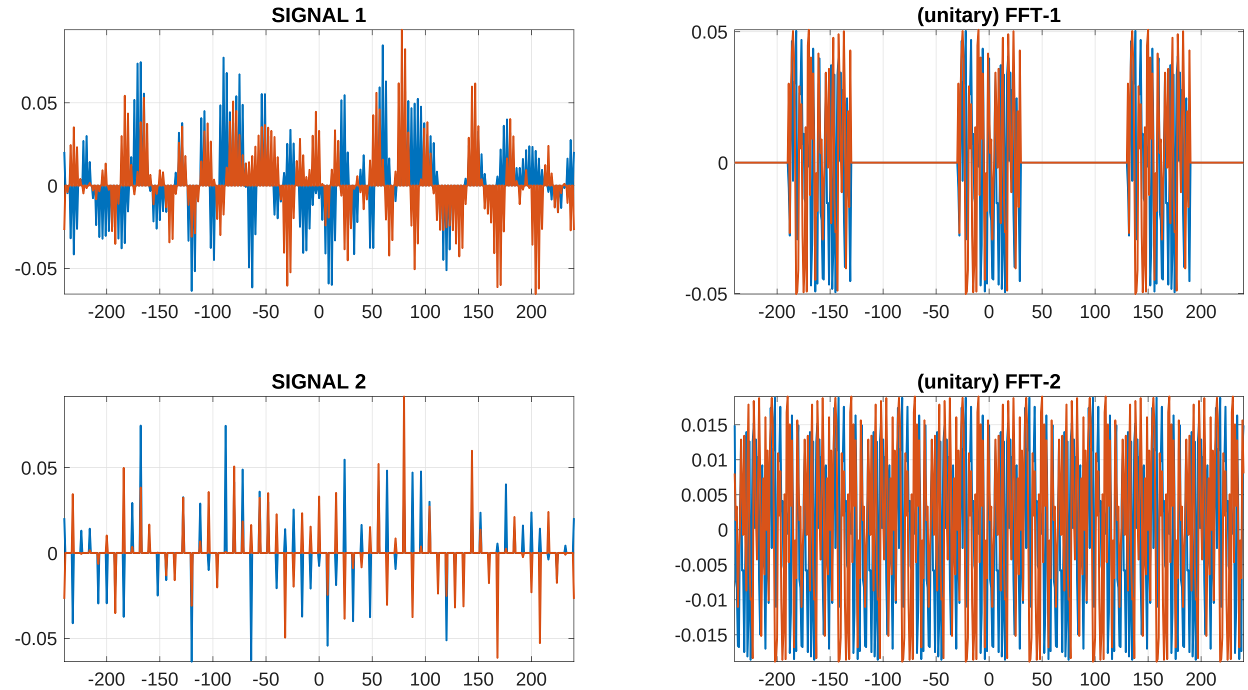

Abbildung 56: pwlinInt02bdl.pdf:
Recovery from dense sampling. Top: one out of three, or three copies of the spectrum: periodicity $N/3$, Below: coarse sampling, repetitions of spectrum are at risk of *aliasing errors*.

34 wst-convergence

We have seen that various sequences (or more generally nets) are w^*-convergent, such as

$$w^*\text{-}\lim D_\Psi \mu \to \mu \quad \text{in } \boldsymbol{M}_b(\mathbb{R}^d),$$

because

$$D_\Psi \mu(f) = \mu(\mathrm{Sp}_\Psi f) \to \mu(f), \quad \forall f \in \boldsymbol{C}_0(\mathbb{R}^d).$$

Clearly this implies w^*-convergence in $\boldsymbol{S}_0'$ (we will describe it $\boldsymbol{S}_0' - w^*$-convergence, or $(\boldsymbol{S}_0', w^*)$ in the future). In fact, if one observes convergence for *any* $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ it will be particularly true for $f \in \boldsymbol{S}_0(\mathbb{R}^d) \subset \boldsymbol{C}_0(\mathbb{R}^d)$.
In certain cases the converse is true as well:

wstMbSOP **Lemma 56.** *Assume that a net (or just a sequence) $(\mu_\alpha)_{\alpha\in I}$ is bounded in $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$, and (only) w^*-convergent in the sense of $\boldsymbol{S}_0'(\mathbb{R}^d)$, then it is also w^*-convergent in the sense of $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$. In fact, convergence for any set of vectors in $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$ whose closed linear span is dense in $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$ implies w^*-convergence in $\boldsymbol{M}_b(\mathbb{R}^d)$.*

Probably the proof has been reproduced in a very similar form earlier.

Proof. Since it is clear that convergence for functions $f = k_1, ..., k_L$ is valid one also has convergence for their finite linear combinations. Hence we can assume that $\mu_\alpha(k) \to$

onrecov1.pdf

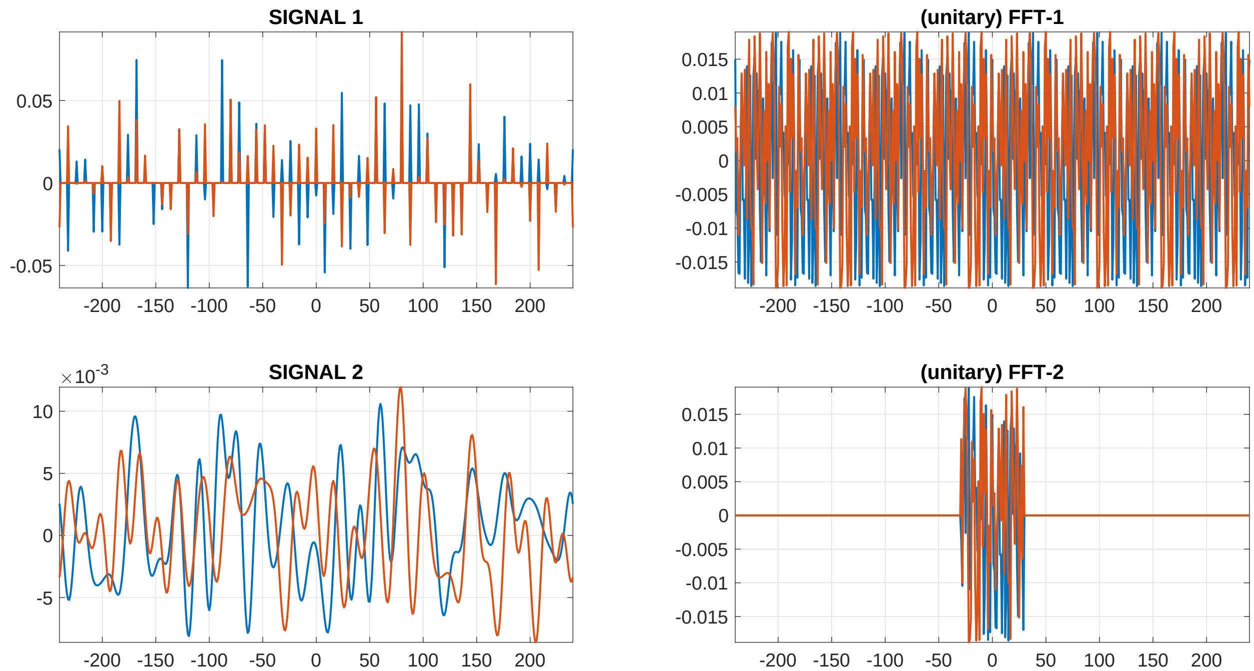

Abbildung 57: Shannonrecov1.pdf:
Given a sampled signal, with sampling rate $a = 8$, hence 8-fold replication of the FT, one can recover, by going (depicted on the right hand side) from the periodi version to the original spectrum by pointwise multiplication with the (ideal?) filter. This example is close to the Nyquist rate, i.e. there is little space between the replications of $\widehat{f}$, so no other choice than using SINC as building blocks for recovery from the samples.

$\mu_0(k)$ for all k from some dense subspace of $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$ (just think of $k \in \boldsymbol{C}_c(\mathbb{R}^d)$, to have something concrete).
Assuming that $\|\mu_\alpha\|_{\boldsymbol{M}_b} \leq C < \infty$ for all α and that for a given function $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ and $\varepsilon > 0$ one has found k from the dense subspace such that $\|f - k\|_\infty < \varepsilon/(3C)$ we apply our usual approximation trick;

$$|\mu_\alpha(f) - \mu_0(f)| \leq |\mu_\alpha(f-k)| + |\mu_\alpha(k) - \mu_0(k)| + |\mu_0(k-f)|.$$

We just have to observe that also, by the density of the subspace

$$\|\mu_0\|_{\boldsymbol{M}_b} = \sup_{k \in \boldsymbol{C}_c(\mathbb{R}^d), \|k\|_\infty \leq 1} |\mu_0(k)|,$$

but

$$|\mu_0(k)| = \lim_\alpha |\mu_\alpha(k)| \leq C\|k\|_\infty = C.$$

Using the fact that (also for $\alpha = 0$)

$$|\mu_\alpha(k-f)| \leq \|\mu_\alpha\|_{\boldsymbol{M}_b} \|k-f\|_\infty < C \cdot \varepsilon/(3C),$$

one only has to put things together in order to achieve: for a suitable choice of α_0 (n_0 for sequences), such that $|\mu_\alpha(k) - \mu_0(k)| < \varepsilon/3$ for $\alpha \succeq \alpha_0$ one gets:

$$|\mu_\alpha(f) - \mu_0(f)| \leq \varepsilon, \quad \text{for } \alpha \succeq \alpha_0.$$

$\square$

singerr1.pdf

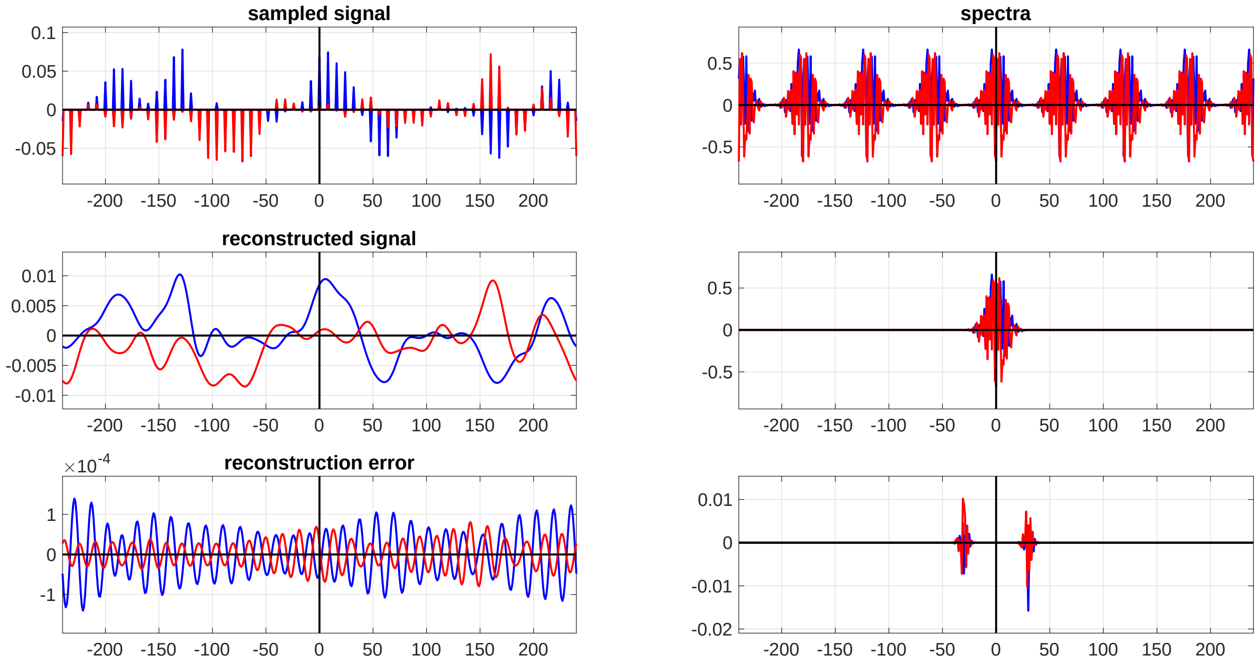

Abbildung 58: aliasingerr1.pdf:
Effect of violation of strict band-limitness condition: spectrum $\widehat{f}$ has Gaussian decay, but it taken as a band-limited function and reconstructed from regular samples. Reconstruction error has a high frequency concentration at the boundaries of the assumed frequency spectrum.

34.1 Uniform convergence over compact subsets

We have observed earlier that for any sequence x_k in $\mathbb{R}$ with $x_k > 0$ and $\lim_{n\to\infty} x_k = 0$ one has

$$\delta_0 = w^* - \lim_{n\to\infty} \delta_{x_k} \quad \text{but} \quad \|\delta_{x_k} - \delta_0\|_{\boldsymbol{M}_b} = 2.$$

Clearly w^*-convergence in $\boldsymbol{M}_b(\mathbb{R})$ implies $\boldsymbol{S}_0' - wst$-convergence, hence also their Fourier transforms, i.e. the sequence χ_{-x_k} has to converge to $\mathcal{F}(\delta_0) = 1$ in the w^*-sense.

ifcompS0Pwst

Lemma 57. *Given a bounded sequence $(h_k)_{k=1}^{\infty}$ in $\left(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\right)$, such for any (compact set, say) ball $B_R(0)$ one has uniform convergence, i.e.*

$$\sup_{x\in B_R(0)} |h_k(x) - h_0(x)| \to 0 \ \textit{for} \ \ k \to \infty,$$

then (viewed as sequence in $\boldsymbol{S}_0'(\mathbb{R}^d)$) one has

$$h_0 = \boldsymbol{S}_0' - w^* - \lim_{n\to\infty} h_n.$$

Proof. Since boundedness of (h_n) in $\left(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\right)$ implies boundedness in $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$ it is enough to check the claimed convergence for compactly supported elements $k \in \boldsymbol{S}_0(\mathbb{R}^d) \cap \boldsymbol{C}_c(\mathbb{R}^d) = \mathcal{F}\boldsymbol{L}^1 \cap \boldsymbol{C}_c(\mathbb{R}^d)$. But for $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ with $\mathrm{supp}(k) \subset B_R(0)$ it is clear that

$$\left|\int_{\mathbb{R}^d} k(x)(h_n(x) - h_0(x)dx\right| \leq \int_{B_R(0)} |k(x)||h_n(x) - h_0(x)|dx \leq \varepsilon \|k\|_{\boldsymbol{L}^1},$$

for $n \geq n_0$ (depending on $\varepsilon > 0$). $\square$

35 Rules for the generalized Fourier Transform

We have DEFINED the Fourier transform for $\sigma \in \boldsymbol{S}_0'$ by the formula $\widehat{\sigma}(f) = \sigma(\widehat{f})$, $f \in \boldsymbol{S}_0$. We have established that this definition is compatible (consistent) with the definition given for $\boldsymbol{C}_c(\mathbb{R}^d)$ (or subsequently $\boldsymbol{L}^1(\mathbb{R}^d)$) via integrals, thus it is justified to call it the *extended Fourier transform*.

There are many different ways to verify that this extended Fourier transform shares all the properties of the ordinary Fourier transform as found in the books[31].

Let us first demonstrate that this defines an invertible linear mapping from $\boldsymbol{S}_0'(\mathbb{R}^d)$ into itself, with the inverse mapping being given by the (expected) formula. In order to avoid confusion of symbols let us write for a moment $\mathcal{F}_g^{-1}$ for the generalized inverse defined in the following way:

$$[\mathcal{F}_g^{-1}(\sigma)](f) = \sigma(\mathcal{F}^{-1}(f)), \quad f \in \boldsymbol{S}_0. \tag{286}$$ IFTgenSOP

The verification is actually a trivial one-line argument:

$$[\mathcal{F}_g^{-1}(\mathcal{F}(\sigma)](f) = [\mathcal{F}(\sigma)](\mathcal{F}^{-1}(f)) = \sigma(\mathcal{F}(\mathcal{F}^{-1}(f)) = \sigma(f), \quad \forall f \in \boldsymbol{S}_0.$$

Also the corresponding concatenation $\mathcal{F} \circ \mathcal{F}_g^{-1} = Id$ is valid, so that the two mappings, both of which are obviously linear and bounded, are inverse to each other. This justifies to use the single symbol $\mathcal{F}^{-1}$ from now on, for both the ordinary (defined on $(\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1)(\mathbb{R}^d)$) and the generalized/extended inverse Fourier transform.

Later on we will use a norm with the extra property that the (ordinary) Fourier is isometric, i.e. satisfying

$$\|\widehat{f}\|_{\boldsymbol{S}_0} = \|f\|_{\boldsymbol{S}_0}, \quad \forall f \in \boldsymbol{S}_0(\mathbb{R}^d). \tag{287}$$ FTS0isom

In this case *also the extended Fourier (and inverse Fourier) transform* are isometric. In fact, we have

$$\|\widehat{\sigma}\|_{\boldsymbol{S}_0'} = \sup_{\|f\|_{\boldsymbol{S}_0} \leq 1} \widehat{\sigma}(f) = \sup_{\|f\|_{\boldsymbol{S}_0} \leq 1} |\sigma(\widehat{f})| = \sup_{\|f\|_{\boldsymbol{S}_0} \leq 1} |\sigma(f)| = \|\sigma\|_{\boldsymbol{S}_0'}, \tag{288}$$ FTS0Pisom

because $f \to \widehat{f}$ is surjective and $\|\widehat{f}\|_{\boldsymbol{S}_0} \leq 1$ if and only if $\|f\|_{\boldsymbol{S}_0} \leq 1$.

We can also extend the involutions, such as $f \to f^{\checkmark}$ to all of $\boldsymbol{S}_0'(\mathbb{R}^d)$ (since $\boldsymbol{S}_0(\mathbb{R}^d)$ is flip-invariant), simply by defining (as we know it already for measures) by

$$\sigma^{\checkmark}(f) = \sigma(f^{\checkmark}), \quad f \in \boldsymbol{S}_0.$$

Of course (exercise) also the extended (inverse) Fourier transform is compatible with the flip-operator, i.e. we have:

$$\mathcal{F}(\sigma^{\checkmark}) = [\mathcal{F}(\sigma)]^{\checkmark}, \quad \sigma \in \boldsymbol{S}_0'. \tag{289}$$ FTgflip1

[31] On the other hand we will describe later, that it can be viewed as the restriction of the theory of Fourier transforms for the space of *tempered distributions* $\boldsymbol{\mathcal{S}}'(\mathbb{R}^d)$, introduced by Laurent Schwartz [84], because his space $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ of *rapidly decreasing functions* on $\mathbb{R}^d$ is a dense subspace of $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$.

The same is true for the pair of dilation operators, namely St_ρ and D_ρ, given by

$$[\mathrm{St}_\rho(\sigma)](f) = \sigma(\mathrm{D}_\rho f) \quad \text{and} \quad [\mathrm{D}_\rho(\sigma)](f) = \sigma(\mathrm{St}_\rho f), \qquad \rho \neq 0. \tag{290}$$

StrhoDrhoSOP

Of course these operators are the unique w^*-w^*--continuous operators which extend the corresponding operators on $\boldsymbol{S}_0(\mathbb{R}^d)$ or $\boldsymbol{M}_b(\mathbb{R}^d)$. Thus (see SCRIPT) on has

$$[\mathrm{St}_\rho(\delta_x)](f) = \delta_x(\mathrm{D}_\rho f) = (\mathrm{D}_\rho f)(x) = f(\rho x) = \delta_{\rho x}(f). \tag{291}$$

Strhodelx

Specifically we can apply these operators to the Dirac comb $\sqcup\!\sqcup$. Since $\mathrm{St}_\rho(\delta_x) = \delta_{\rho x}$ (the mass preserving dilation operator for discrete measures) we get:

$$\mathrm{St}_\rho(\sqcup\!\sqcup) = \sqcup\!\sqcup_{\rho\mathbb{Z}^d} = \sum_{k\in\mathbb{Z}^d} \delta_{\rho k}. \tag{292}$$

Strhoshah

This family of functionals (let us keep $\rho \in [1,\infty)$) is uniformly bounded on $\boldsymbol{S}_0(\mathbb{R}^d)$ (as it is easily seen from the proof that $\sqcup\!\sqcup \in \boldsymbol{W}(\boldsymbol{M},\boldsymbol{\ell}^\infty)(\mathbb{R}^d)$! resp. the fact that the compression operator D_ρ acts non-expansively on $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$. Just think of the atomic characterization (note: for $\rho \geq 1$ the support size shrinks!), and use the fact that

$$\|\mathrm{D}_\rho(\widehat{f})\|_{\mathcal{F}\boldsymbol{L}^1} = \|\mathcal{F}(\mathrm{St}_\rho f)\|_{\mathcal{F}\boldsymbol{L}^1} = \|\mathrm{St}_\rho f\|_{\boldsymbol{L}^1} = \|f\|_{\boldsymbol{L}^1} = \|\widehat{f}\|_{\mathcal{F}\boldsymbol{L}^1}, \quad f \in \boldsymbol{S}_0, \rho > 0. \tag{293}$$

DrhoFLi1

Given the boundedness of the family $\{\mathrm{St}_\rho(\sqcup\!\sqcup), \rho \in [1,\infty)\}$ inside of $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$ it is enough to check the behavior for $\rho \to \infty$ on $k \in \boldsymbol{C}_c(\mathbb{R}^d) \subset \boldsymbol{W}(\mathbb{R}^d)$:
We may assume that $\mathrm{supp}(k) \subset B_R(0)$, hence for sufficiently large (e.g. $\rho > R$) the only point within $\mathrm{supp}(k)$ is 0, since $\rho\mathbb{Z}^d \cap B_R(0) = \{0\}$. This implies that

$$\mathrm{St}_\rho \sqcup\!\sqcup(k) = \sum_{k\in\mathbb{Z}^d} k(\rho k) = k(0) = \delta_0(k),$$

and consequently

$$w^*\text{-}\lim_{\rho\to\infty} \mathrm{St}_\rho = \delta_0.$$

Since the Fourier transform is w^*-w^*-continuous the Fourier transforms have to tend to $\mathbf{1}$, also in the $\boldsymbol{S}_0' - w^*$-sense. We thus have

$$\mathrm{D}_\rho(\sqcup\!\sqcup) = \mathcal{F}(\mathrm{St}_\rho(\sqcup\!\sqcup)) \to \mathcal{F}(\delta_0) = \mathbf{1}, \text{for } \rho \to \infty,$$

in the w^*-sense.
While the St_ρ-operator is *mass preserving*, mapping δ_x to $\delta_{\rho x}$ without changing the amplitudes, the D_ρ-operator is preserving the density, i.e. the integral (at last approximately) of $k \in \boldsymbol{C}_c(\mathbb{R}^d)$, with

$$[\mathrm{D}_\rho(\delta_x)](f) := \delta_x(\mathrm{St}_\rho(x)) = \rho^{-d} f(x/\rho) = [\rho^{-d}\delta_{x/\rho}](f).$$

If we apply these measures we find that the expression $[\mathrm{D}_\rho(\sqcup\!\sqcup)](k)$ corresponds to a Riemannian sum applied to $k \in \boldsymbol{C}_c(\mathbb{R}^d)$, which is thus convergent[32] for $\rho \to \infty$ to

$$\int_{\mathbb{R}^d} k(x)dx = \int_{\mathbb{R}^d} \mathbf{1}\cdot k(x)dx = \sigma_{\mathbf{1}}(k),$$

as predicted by the above argument.

Exercise 8. Check that $\mathrm{St}_\rho(\sqcup\!\sqcup_\alpha) = \sqcup\!\sqcup_{\rho\alpha}$, for $\alpha, \rho > 0$.

[32] We use the symbol $\mathbf{1}$ for the constant function taking the value 1 for any $x \in \mathbb{R}^d$.

36 The embedding of $\boldsymbol{S}_0(\mathbb{R}^d)$ into $\boldsymbol{S}_0'(\mathbb{R}^d)$

Most of the Banach spaces we need for our discussion of the Fourier transform or of questions of time-frequency and in particular Gabor Analysis are sandwiched between $\boldsymbol{S}_0(\mathbb{R}^d)$ and $\boldsymbol{S}_0'(\mathbb{R}^d)$, i.e. satisfy

$$\big(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\big) \hookrightarrow (\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}}) \hookrightarrow (\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'}). \tag{294}$$ SOsandw

In fact, any non-trivial Banach space which is a double module, i.e. which allows

$$\boldsymbol{L}^1(\mathbb{R}^d) \star \boldsymbol{B} \subseteq \boldsymbol{B} \quad \text{and} \quad \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) \cdot \boldsymbol{B} \subseteq \boldsymbol{B}$$

contains $\boldsymbol{S}_0(\mathbb{R}^d)$, because it can be shown to be the *smallest* among all the Banach spaces with this *double module structure*[33].

It is not difficult to show that $\boldsymbol{S}_0(\mathbb{R}^d)$ is *not dense* in $\boldsymbol{S}_0'(\mathbb{R}^d)$. We may consider a Dirac comb as element of $\boldsymbol{S}_0'(\mathbb{R})$ which cannot be approximated (in the norm of $\boldsymbol{S}_0'(\mathbb{R})$) by elements of $\boldsymbol{S}_0(\mathbb{R}^d)$. Essentially one can use (to come) the characterization of the norm closure of $\boldsymbol{S}_0(\mathbb{R}^d)$ within $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$ as the set of all elements from $\boldsymbol{S}_0'(\mathbb{R}^d)$ with a STFT in $\boldsymbol{C}_0(\mathbb{R}^d \times \widehat{\mathbb{R}}^d)$. This will be discussed later on.

However we have a positive statement concerning the w^*-density of $\boldsymbol{S}_0(\mathbb{R}^d)$ in $\boldsymbol{S}_0'(\mathbb{R}^d)$. This can be done using the (very useful) relationships (proof given later):

$$(\boldsymbol{S}_0'(\mathbb{R}^d) \star \boldsymbol{S}_0(\mathbb{R}^d)) \cdot \boldsymbol{S}_0(\mathbb{R}^d) \subseteq \boldsymbol{S}_0(\mathbb{R}^d) \quad \text{and} \quad (\boldsymbol{S}_0'(\mathbb{R}^d) \cdot \boldsymbol{S}_0(\mathbb{R}^d)) \star \boldsymbol{S}_0(\mathbb{R}^d) \subseteq \boldsymbol{S}_0(\mathbb{R}^d),$$

which will be discussed later, but require a characterization of $\boldsymbol{S}_0'(\mathbb{R}^d)$ as Wiener amalgam space $\boldsymbol{W}(\mathcal{F}\boldsymbol{L}^\infty, \boldsymbol{\ell}^\infty)(\mathbb{R}^d)$ and convolution relations between general Wiener amalgams. Instead we do some thing more simple:

regulSOP **Lemma 58.**

$$\boldsymbol{S}_0'(\mathbb{R}^d) \star \boldsymbol{S}_0(\mathbb{R}^d) \subset \boldsymbol{C}_{ub}(\mathbb{R}^d) \quad \textit{and} \quad \|\sigma \star f\|_\infty \leq \|\sigma\|_{\boldsymbol{S}_0'} \|f\|_{\boldsymbol{S}_0}. \tag{295}$$ SOPconvSO

Proof. We just have to recall that (as for the pairing $(\boldsymbol{C}_0(\mathbb{R}^d), \boldsymbol{M}_b(\mathbb{R}^d))$) the convolution can be expressed in a pointwise sense and that translation is continuous in $\big(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\big)$. In fact $\boldsymbol{S}_0(\mathbb{R}^d) \star \boldsymbol{S}_0'(\mathbb{R}^d) \subset \boldsymbol{L}^1(\mathbb{R}^d) \star \boldsymbol{S}_0'(\mathbb{R}^d) \subset \boldsymbol{M}_b(\mathbb{R}^d) \star \boldsymbol{S}_0'(\mathbb{R}^d)$ (together with corresponding norm esimates) follows from the fact that $\boldsymbol{M}_b(\mathbb{R}^d) \star \boldsymbol{S}_0(\mathbb{R}^d) \subseteq \boldsymbol{S}_0(\mathbb{R}^d)$, using the adjointness argument.

That the abstract definition, defined via

$$\mu \star \sigma(f) = \sigma(\mu^{\checkmark} \star f), \quad f \in \boldsymbol{S}_0, \mu \in \boldsymbol{M}_b(\mathbb{R}^d),$$

coincides with the direct, pointwise definition for $\mu = \mu_k, k \in \boldsymbol{S}_0(\mathbb{R}^d)$, namely the equality

$$\mu_k \star \sigma = \sigma_h, \quad \text{with } h = k \star_{pw} \sigma$$

given in a pointwise sense as

$$k \star_{pw} \sigma(y) = \sigma(T_y k^{\checkmark}), k \in \boldsymbol{S}_0(\mathbb{R}^d), \sigma \in \boldsymbol{S}_0', \tag{296}$$ SOSOPpw

[33] To be explained in the SCRIPT, we do not provide a proof here, this is just for the orientation of the reader.

is established in analogy to the proof for $\boldsymbol{C}_0(\mathbb{R}^d) \star \boldsymbol{M}_b(\mathbb{R}^d)$. It is then clear from (296) SOSOPpw that we have

$$|k \star_{pw} \sigma(y)| \leq \|\sigma\|_{\boldsymbol{S}_0'} \|k\|_{\boldsymbol{S}_0}$$

and consequently

$$\|k \star_{pw} \sigma\|_\infty \leq \|\sigma\|_{\boldsymbol{S}_0'} \|k\|_{\boldsymbol{S}_0}, \quad f \in \boldsymbol{S}_0, \sigma \in \boldsymbol{S}_0',$$

and finally, using the fact that convolution commutes with translation:

$$\|T_z(k \star_{pw} \sigma) - k \star_{pw} \sigma\|_\infty \leq \|\sigma\|_{\boldsymbol{S}_0'} \|T_z k - k\|_{\boldsymbol{S}_0} \to 0$$

for $z \to 0$ (see Prop. WAliprops1 11) thus establishing the (uniform) continuity of $k \star_{pw} \sigma$. □

SOdenseSOP **Lemma 59.** *Given $\sigma \in \boldsymbol{S}_0'$, $\varepsilon > 0$ and $f_1, ..., f_L$ in $\boldsymbol{S}_0(\mathbb{R}^d)$ there exists ρ_0 such that for $\rho \in (0, \rho_0]$ one has (writing $g_\rho = \mathrm{St}_\rho g_0$ and $g^\rho = \mathrm{D}_\rho g$):*

$$|\left[([\sigma \star g_\rho] \cdot g^\rho) \star g_\rho\right](f_k) - \sigma(f_k)| < \varepsilon, \quad \text{for } k = 1, ..., L. \tag{297}$$ SOdSOPform

$$(\boldsymbol{C}_{ub}(\mathbb{R}^d) \cdot \boldsymbol{S}_0(\mathbb{R}^d)) \star \boldsymbol{S}_0(\mathbb{R}^d) \subset \boldsymbol{W}(\mathbb{R}^d) \star \boldsymbol{S}_0(\mathbb{R}^d) \subset \boldsymbol{L}^1(\mathbb{R}^d) \star \boldsymbol{S}_0(\mathbb{R}^d) \subset \boldsymbol{S}_0(\mathbb{R}^d) \tag{298}$$ doublereg

implies that $\boldsymbol{S}_0(\mathbb{R}^d)$ is w^*-dense in $\boldsymbol{S}_0'(\mathbb{R}^d)$.

Proof. We do not carry out all the details, but just recall that each of the functions g_ρ and g^ρ is even, n and that the family $g_\rho = \mathrm{St}_\rho g_0$ is a bounded approximate unit in $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$, acting in the expected way on $\boldsymbol{S}_0(\mathbb{R}^d)$, and g^ρ is a bounded approximate unit in $\left(\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_{\mathcal{F}\boldsymbol{L}^1}\right)$, for $\rho \to 0$.
Further more we use the relations which hold for $\tau \in \boldsymbol{S}_0'(\mathbb{R}^d)$, $f \in \boldsymbol{S}_0(\mathbb{R}^d)$:

$$[\tau \star g_\rho](f) = \tau(g_\rho \star f), \quad \text{and} \quad [\tau \cdot g^\rho](f) = \tau(g^\rho \cdot f).$$

So we can apply the Lemma on iterated approximations, i.e. Lemma 29 from the SCRIPT (`doublenet`), using $T_\rho : f \mapsto f \star g_\rho$ and $S_\rho : f \mapsto f \cdot \mathrm{D}_\rho g_0$ (p.47).
Just to indicate the first iteration, the approximation of $\sigma(f)$ by functions of the form $[g^\rho \cdot (g_\rho \star \sigma)](f)$. Without loss of generality (by rescaling) we may assume that $\|\sigma\|_{\boldsymbol{S}_0'} = 1$.
We have, by the usual splitting trick applied to any $f = f_k$:

$$|[g^\rho \cdot (g_\rho \star \sigma)](f) - \sigma(f)| = |\sigma[g_\rho \star (g^\rho \cdot f) - f] \leq \|\sigma\|_{\boldsymbol{S}_0'} \|g_\rho \star (g^\rho \cdot f) - f\|_{\boldsymbol{S}_0}, \tag{299}$$ split006

but the last term can be estimated as

$$\|g_\rho \star (g^\rho \cdot f) - f\|_{\boldsymbol{S}_0} \leq \|g_\rho \star (g^\rho \cdot f - f)\|_{\boldsymbol{S}_0} + \|g_\rho \star f - f\|_{\boldsymbol{S}_0}$$

which in turn can be estimated by the following term, since $\|g_\rho\|_{\boldsymbol{L}^1} = 1$:

$$\|g^\rho \cdot f - f)\|_{\boldsymbol{S}_0} + \|g_\rho \star f - f\|_{\boldsymbol{S}_0} < 2\varepsilon,$$

for $\rho \in (0, \rho_0)$, for $f = f_k$, $k = 1, ..., L$ and ρ_0 properly chosen. □

Looking at the situation in the world of spectrograms (short time Fourier transforms) we see (roughly speaking the situation of the second row of Fig. DiracRegLoD1pdf 36, in Fig. irpsmooth1.jpg 36:

cRegLoc1.pdf

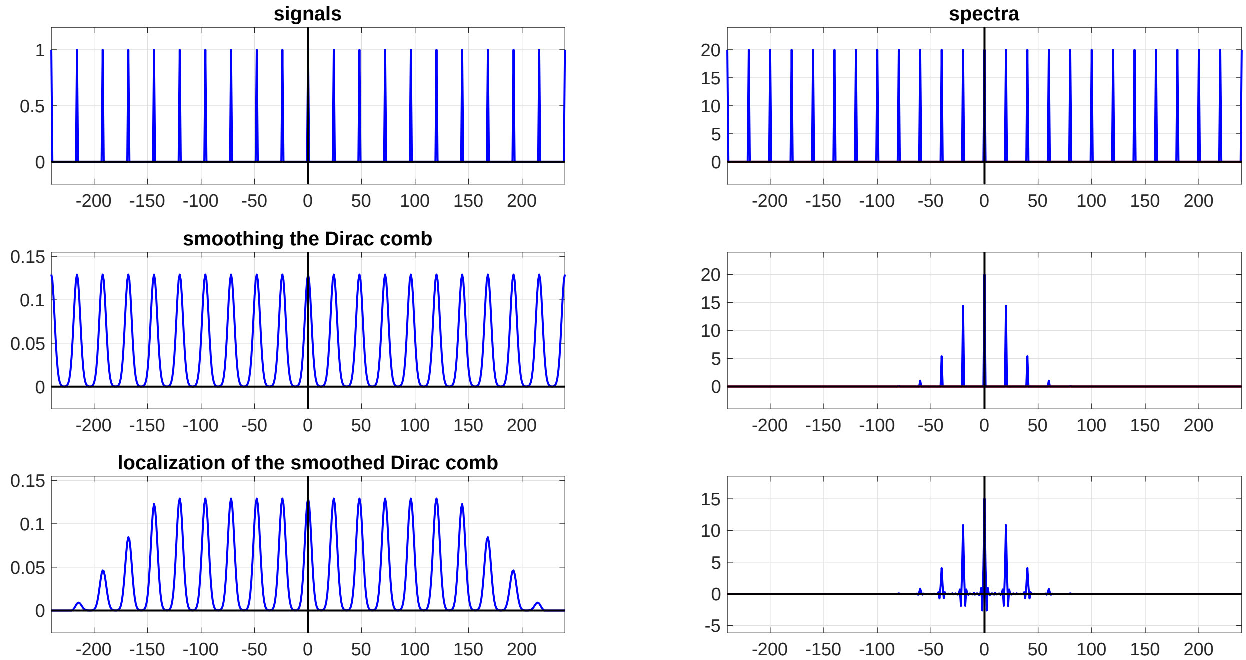

Abbildung 59: DiracRegLoc1.pdf:
Starting from a Dirac comb $⊔⊔ \in \boldsymbol{S}_0'(\mathbb{R})$ e can do smoothing, which corresponds to a damping on the Fourier transform side, and then a periodization, which is a sampling on the FT side, or vice versa. Anyway, the resulting regularization of σ approximates the original mild distribution in the w^*–sense.

Remark 32. The first expression in (S0dSOPform 297) can be interpreted in two ways. Formally it is the application of a mild distributions on each f_k, but the point of the argument is the fact that this mild distribution can be represented by an element from $\boldsymbol{S}_0(\mathbb{R}^d)$ (even a compactly supported one), namely $h = ([\sigma \star g_\rho] \cdot g^\rho) \star g_\rho \in \boldsymbol{S}_0(\mathbb{R}^d)$, in the usual way, as $f \to \mu_h(f) = \int_{\mathbb{R}^d} f(x)h(x)dx$.

MdfinSOP **Lemma 60.** *The finite discrete measures are w^*–dense in $\boldsymbol{S}_0'(\mathbb{R}^d)$.*

Proof. Given $\sigma \in \boldsymbol{S}_0'(\mathbb{R}^d)$, $f_1, ..., f_L \in \boldsymbol{S}_0(\mathbb{R}^d)$ and $\varepsilon > 0$ one has to find some discrete, finite measure $\nu \in \boldsymbol{M}_d(\mathbb{R}^d)$ such that

$$|\sigma(f_j) - \nu(f_j)| < \varepsilon, \quad 1, ..., L. \tag{300}$$ sigmudappr

We know already that $\boldsymbol{S}_0(\mathbb{R}^d) \subset \boldsymbol{L}^1(\mathbb{R}^d) \subset \boldsymbol{M}_b(\mathbb{R}^d)$ is w^*–dense in $\boldsymbol{S}_0'(\mathbb{R}^d)$. Since the compactly supported elements of $\boldsymbol{S}_0(\mathbb{R}^d)$ are dense in $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ it is enough to show that any $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ (or $\boldsymbol{C}_c(\mathbb{R}^d) \cap \boldsymbol{S}_0(\mathbb{R}^d)$) can be approximated in the w^*–sense by a finite discrete measure. But we know that the D_Ψ-operator does exactly this (even with respect to any $f \in \boldsymbol{C}_0(\mathbb{R}^d)$, not only $f \in \boldsymbol{S}_0(\mathbb{R}^d)$).
A combination of these facts completes the argument. We plan to provide more details separately but encourage the reader to fill in the details, starting from the proof of Lemma (S0denseSOP 59).

csmooth1.jpg

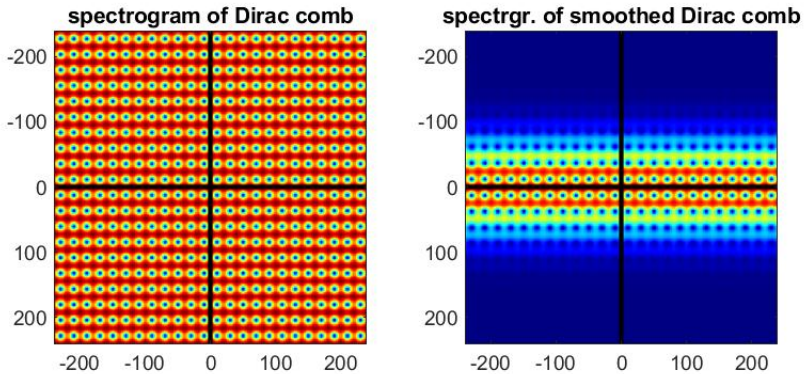

Abbildung 60: Diracsmooth1.jpg: The convolution with a narrow Gaussian corresponds to pointwise multiplication with a wide dilated Gaussian, which narrows down the left over frequency band, so only the lower frequencies constituting the Dirac comb are left intact. Pointwise multiplication has a corresponding effect in the time direction.

split006

In fact, we have by formula (299) observed that for any $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ with $g(0) = 1$ and $\widehat{g}(0) = 1$ one has or $\rho \in (0, \rho_0]$:

$$\|g^{\rho} \cdot (g_{\rho} \star f_k - f_k)\|_{\boldsymbol{S}_0} \leq \varepsilon, \ \ 1, ..., L.$$

If we choose g to have *compact support* the resulting function $k = g^{\rho} \cdot (g_{\rho} \star \sigma)$ belongs to $\boldsymbol{C}_c(\mathbb{R}^d)$. By using the properties of the D_{Ψ}-operator we find that

$$|(\mathrm{D}_{\Psi}\mu_k)(f_k) - \mu_k(f_k)| \to 0, \quad \text{for } \operatorname{diam}(\Psi) \to 0,$$

(which is essentially a statement that $\int_{\mathbb{R}^d} f_k(x)k(x)dx$ can be approximated by finite Riemannian sums).
Combining these facts we obtain

$$|\sigma(f_k) - \mathrm{D}_{\Psi}\mu_k(f_k)| \leq 3\varepsilon, \ \ 1, ..., L,$$

which implies that the given mild distribution $\sigma \in \boldsymbol{S}_0'$ can be approximated by the action of some finite discrete measure, i.e. that finite discrete measures are w^*–dense in $\boldsymbol{S}_0'(\mathbb{R}^d)$. □

Remark 33. There is an abstract, functional analytic argument for this situation. Given the dual space $(\boldsymbol{B}', \|\cdot\|_{\boldsymbol{B}'})$ of some Banach space, we have:
The w^*-closed linear span (i.e. the w^*-closure of the linear span, which is then the

smallest, w^*-closed subspace) of a given set $M \subset \boldsymbol{B}'$ is all of $\boldsymbol{B}'$ (resp. the subspace of all finite linear combinations of the form $\sum_{k=1}^{K} c_k m_k$ is w^*-dense in $\boldsymbol{B}'$) if (and only of) $\sigma(f) = 0$ for any $\sigma \in M$ implies already $f = 0$

For our case the family $M = \{\delta_x, x \in \mathbb{R}^d\}$ can be used, for $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}}) = \Big(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\Big)$. Clearly $\delta_x(f) = 0$ for all $x \in \mathbb{R}^d$ implies of course that $f = 0$. However, this argument is non-constructive, while our proof provides a specific recipe (even a linear mapping) from $\boldsymbol{S}_0'(\mathbb{R}^d)$ to the required subset or subspace. For the case of the target space $\boldsymbol{S}_0(\mathbb{R}^d)$ (our first statement) this is a procedure called *regularization*.

We can even combine the last two observations and show that the *discrete and periodic measures* (which belong to $\boldsymbol{W}^*$ and hence to $\boldsymbol{S}_0'(\mathbb{R}^d)$)are w^*–dense in $\boldsymbol{S}_0'(\mathbb{R}^d)$.

rdiscmeasSOP **Lemma 61.** *Given $\sigma \in \boldsymbol{S}_0'(\mathbb{R}^d)$, a finite set $f_1, ..., f_L$ in $\boldsymbol{S}_0(\mathbb{R}^d)$ and $\varepsilon > 0$, there exists some discrete and periodic measure ν such that*

$$|\nu(f_k) - \sigma(f_k)| \leq \varepsilon, \quad 1, ..., L. \tag{301}$$

perdiscSOPapp

Proof. We just provide the key argument. By the boundedness of the approximation procedure so far we can even verify the boundedness of the family $\sqcup\!\sqcup_\alpha \star \mathrm{D}_\Psi(\mu_k)$, for $\alpha \geq 1$. Given this we observe that it is enough to give a detailed proof under the additional assumption that $f_1, ..., f_L$ are all supported in $B_R(0)$ for some $R > 0$. But in this case (given the joint compactness of the supports of all the discrete measures $\mathrm{D}_\Psi(\mu_k)$ for $k \in \boldsymbol{C}_c(\mathbb{R}^d)$, for $|\Psi| \leq 1$) we can replicate (translates of)$\mathrm{D}_\Psi(\mu_k)$ outside of $B_R(0)$ such that one has in fact:

$$[\sqcup\!\sqcup_\alpha \star \mathrm{D}_\Psi(\mu_k)](f_j) = \mathrm{D}_\Psi(\mu_k)(f_j), \quad 1, ..., L.$$

$\square$

We have used the argument, that for any $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ the family of periodic measures is bounded in $\boldsymbol{W}^*(\mathbb{R}^d)$

$$\{\, \sqcup\!\sqcup_\alpha \star \mu \,|\, \alpha \geq 1\}.$$

For the atoms $\mu\psi_i \in \boldsymbol{M}_b(\mathbb{R}^d)$ this is easy to verify, because for α large enough the αk-translates are living on disjoint cubes and therefore the action on atoms for $\boldsymbol{W}(\mathbb{R}^d)$ is uniformly bounded (Details may be cumbersome).

Remark 34. Note that we did not say that the *space* of periodic and discrete measures is w^*–dense in $\boldsymbol{S}_0'$. In fact, this is just a subset of $\boldsymbol{S}_0'(\mathbb{R}^d)$, but not a subspace, since clearly the sum of two such measures, say $\sqcup\!\sqcup_1$ and $\sqcup\!\sqcup_{\sqrt{2}}$, is a discrete measure on $\mathbb{R}^d$, in fact an element of $\boldsymbol{W}^*(\mathbb{R}^d)$, but not a periodic one.

37 Uniqueness of the extended Fourier transform

The fact, that $\boldsymbol{S}_0(\mathbb{R}^d)$ is w^*–dense in $\boldsymbol{S}_0'(\mathbb{R}^d)$can be used in order to demonstrate that the extended Fourier transform, defined as $\widehat{\sigma}(f) = \sigma(\widehat{f}), f \in \boldsymbol{S}_0$ is the unique extension of the ordinary Fourier transform (defined on $\boldsymbol{S}_0(\mathbb{R}^d)$ via Riemann integrals:

SOPFTuniq **Theorem 13.** *The definition of the extended Fourier transform, via the formula*

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

is equivalent to the following description: Given a sequence $(f_n)_{n=1}^{\infty}$ *in* $\boldsymbol{S}_0(\mathbb{R}^d)$ *with*

$$\sigma = w^*\text{-}\lim{}_{n\to\infty} \sigma_{f_n},$$

the sequence $(\widehat{f}_n)_{n=1}^{\infty}$ *is also* w^*-*convergent and its limit equals* $\widehat{\sigma} \in \boldsymbol{S}_0'(\mathbb{R}^d)$.
Thus the extended FT is the unique w^**-*w^**-continuous extension of the ordinary Fourier transform mapping* $\boldsymbol{S}_0(\mathbb{R}^d)$ *into itself.*

There are now many other characterizations of the extended Fourier transform. We know already, that it maps Dirac measures to pure frequencies and vice versa, since $\mathcal{F}(\delta_x) = \chi_{-x}$ and $\mathcal{F}(\chi_s) = \delta_s$, for $x, s \in \mathbb{R}^d$.
We can use this fact to derive another uniqueness result:

SOPFTDirac **Theorem 14.** *The extended Fourier transform* $\mathcal{F}$ *is the unique,* w^**-*w^**-continuous linear mapping* $\boldsymbol{S}_0'(\mathbb{R}^d) \to \boldsymbol{S}_0'(\mathbb{R}^d)$ *which maps* χ_s *to* δ_s *respectively which maps* δ_x *to* χ_{-x}.

Proof. We know that the extended Fourier transform has these mapping properties. It is also w^*-w^*-continuous .
In order to verify the uniqueness under these circumstances it is enough to verify that the discrete measures are w^*-dense in $\boldsymbol{S}_0'(\mathbb{R}^d)$. By the symmetry of the Fourier (and inverse) FT this implies (and is equivalent) to the fact that the trigonometric polynomials are also w^*-dense in $\boldsymbol{S}_0'(\mathbb{R}^d)$.
This has been observed in Lemma MdfinSOP 60 above. □

Finally let us discuss the action of the extended Fourier transformh $\boldsymbol{S}_0'(\mathbb{R}^d) \to \boldsymbol{S}_0'(\mathbb{R}^d)$ by clarifying the role of discrete and periodic measures. To make sure what we mean: these are measures obtained by periodization of discrete measures, linear functionals on $\nu \in \boldsymbol{S}_0'(\mathbb{R}^d)$ with the property that (for any BUPU Ψ) one has $\nu\psi_i \in \boldsymbol{M}_d(\mathbb{R}^d)$, for any $i \in I$. Equivalently we can say: $\nu\phi \in \boldsymbol{M}_d(\mathbb{R}^d)$, for any $\phi \in \boldsymbol{C}_c(\mathbb{R}^d)$ (so locally it is just a discrete measures).
Recall that $\sigma \in \boldsymbol{S}_0'$ called Λ-periodic, if one has

$$T_\lambda\sigma = \sigma, \quad i.e.\ \sigma(T_{-\lambda}f) = \sigma(f), \ \forall \lambda \in \Lambda, f \in \boldsymbol{S}_0.$$

38 Sampling and Periodization

The observation made by a simple MATLAB experiment that sampling of a signal corresponds to periodization on the Fourier transform side (also easily explained by the fact that a polymial of the form

$$p(x) = a_1 + a_2x^3 + a_3x^6 + a_4x^9$$

can be of course described by a substitution $y = x^3$ as a cubic polynomial

$$q(y) = a_1 + a_2 y + a_3 y^2 + a_4 y^3.$$

For a signal length n which has 3 as a divisor, say $n = 12$, for simple explanation, one has to evaluate that cubic polynomial at the unit roots of order 12, but taking the third power, so we get three copies of $\mathbb{U}_4$, and this explains the replication of the (short, length 4) spectrum (more explanations about the DFT and the evaluation of polynomials at unit roots of order n to be given later on).
Let us state one of the important consequences of Poisson's formula for the case $\Lambda \lhd \mathbb{R}$ (i.e. for $d = 1$):

sampperFT **Proposition 16.** *Given $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ we have $\nu := f \cdot \sqcup\!\sqcup = \sum_{k\in\mathbb{Z}} h(k)\delta_k \in \boldsymbol{M}_d(\mathbb{R})$. It has a Fourier transform is a $\mathbb{Z}$-periodic function, which can be described as $\sqcup\!\sqcup \star \widehat{f}$:*

$$\mathcal{F}(f \cdot \sqcup\!\sqcup) = \widehat{f} \star \sqcup\!\sqcup, \quad f \in \boldsymbol{S}_0. \tag{302}$$ shahconvmult

In a similar way the periodized version of f is a periodic function whose Fourier coefficients are just the samples of $\widehat{f}$:

$$\mathcal{F}(f \star \sqcup\!\sqcup) = \widehat{f} \cdot \sqcup\!\sqcup, \quad f \in \boldsymbol{S}_0. \tag{303}$$ shahmultconv

By combining this result, also with dilations, and choosing the sampling rate $b > a$ of the form $b = a/n$ for some $n \in \mathbb{N}$ we get:

sampper02 **Corollary 9.** *For $c = 1/b, d = 1/a$ we have, up to the constant $C = C_{a,b,c,d}$.*

$$\mathcal{F}[\sqcup\!\sqcup_a \star (\sqcup\!\sqcup_b \cdot f)] = C \cdot [\sqcup\!\sqcup_d \cdot (\sqcup\!\sqcup_c \star f)] = C \cdot [\sqcup\!\sqcup_c \star (\sqcup\!\sqcup_d \cdot \widehat{f})], \quad f \in \boldsymbol{S}_0. \tag{304}$$ sampperF

The resulting discrete measure/distribution on the time-side is a-periodic and supported on $b\mathbb{Z}$, while the Fourier transform in the sense of $\boldsymbol{S}_0'(\mathbb{R}^d)$(!) is d-periodic and supported on $d\mathbb{Z}$, i.e. the support of the FT is the orthogonal lattice of the sampling lattice $b\mathbb{Z}$ and the periodicity on the FT side is the orthogonal lattice to the sampling lattice $b\mathbb{Z}$ on the time side (and vice versa).

Remark 35. There is also a kind of converse of the last statement in the corollary. Any such periodic discrete measure (either on the time side or on the frequency side) is created in this way. We just have to start with a piecewise linear function taking the correct values at the points $0, 1/b, ...(n-1)/b$ (with $n/b = a$).

ampdem1a.eps

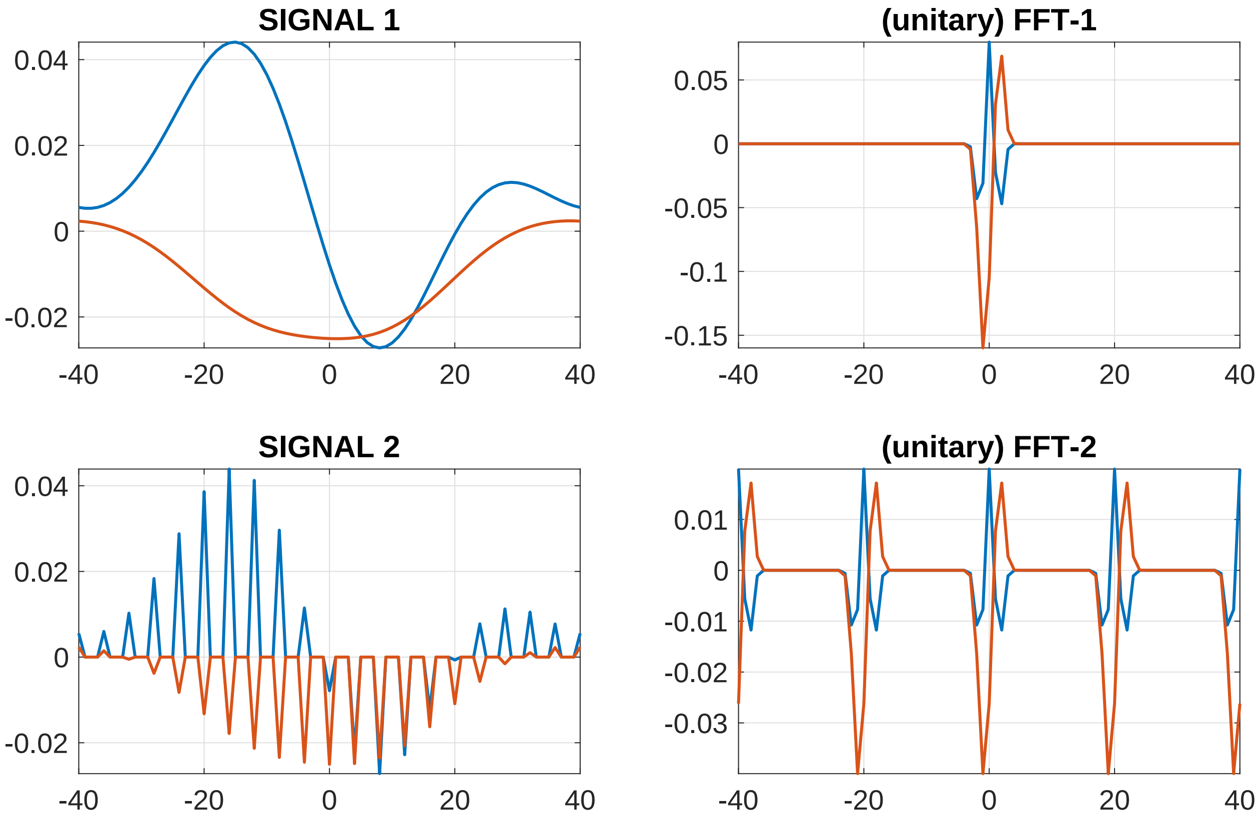

Abbildung 61: persampdem1a.eps:
Given a rather smooth, band-limited signal (low band-width can be recognized on the right hand side, we can expect to recover the function from its sampled version. The information in the sampled version corresponds to a periodized version of $\widehat{f}$.

One can take the samples of the basic period, IN THE RIGHT WAY, i.e. starting with the value at zero, i.e. at the first MATLAB coordinate... then go to the end of the period, ca. at half the signal length and then continue to the end in order to get the coordinates to the LEFT of zero.
One can then observe: The sequence of reduced length **(by the number of periods, multiplied by the sampling gap)** can be put into the DFT/FFT Routine of reduced length, and we will get the same result.
Simple case: $m = 24$, with sampling rate 3 and 2-fold periodization, so the signal plot the signal, cyclically rotated by $m/2 = 12$. Thus we will get two sampled copies, and only the first 12 coordinates are required. But only every 3rd value is non-zero (due to the sampling) So we should start at zero (in the continuous world) which has label/coordinate 1 in the MATLAB world. So we take cordinates $1, 4 = 1 + 3, 7 = 4 + 3$ and $10 = 13 - 3$, the “left” neighbor. We SHOULD take $m + 1 - 3 = 22$, but the value of the periodized version at 22 equals the value at 12.
Details are demonstrated in the MATLAB EXPO talk.

`http://univie.ac.at/nuhag-php/dateien/talks/3609_EXPOFeiE6.pdf` sampperiod2.pdf

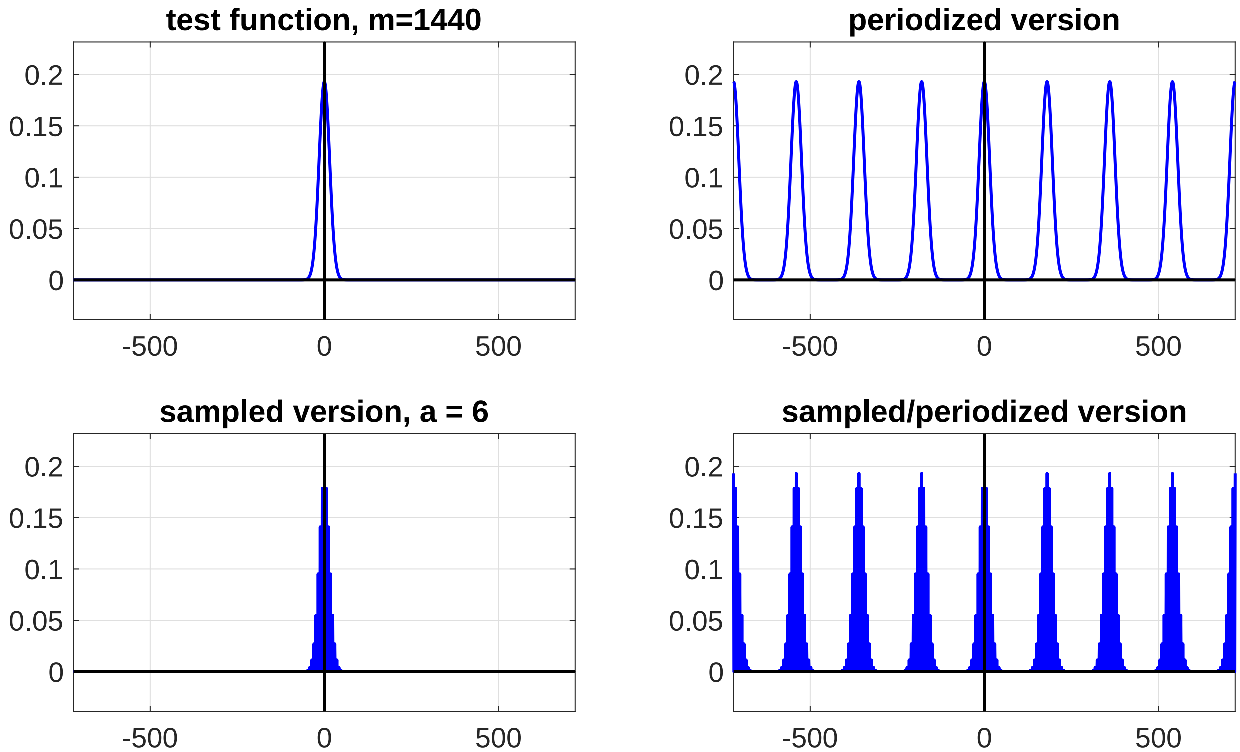

Abbildung 62: sampperiod2.pdf:
Starting from the smooth function to the left, we can either periodize (top right), or sample (left lower corner), or combine it, in fact, if properly done(i.e. if the period is a multiple of the sampling rate!), in any order, BOTH operations.

39 The DFT derived from the FT for $S_0'(\mathbb{R}^d)$

Let us now show that the DFT (usually implemented as FFT) can be described as the realization of the extended Fourier transform, applied to periodic and discrete signals, i.e. series of Dirac measures with amplitudes (real or complex) repeated in a regular manner, i.e. along a lattice. This argument can be applied for any two lattices in $\mathbb{R}^d$, with Λ_0 being a fine lattice (carrying the discrete signal), which is periodic with respect to some sublattice $\Lambda \lhd \Lambda_0$.
Since any lattice Λ_0 is of the form $\Lambda_0 = \boldsymbol{A} \star \mathbb{Z}^d \lhd \mathbb{R}^d$ we can always assume that $\Lambda_0 = \mathbb{Z}^d$ and thus $\Lambda = \boldsymbol{B} \star \mathbb{Z}^d$, for some (other) non-singular matrix (for the general case we assume that $(\boldsymbol{A}^{-1} \star \boldsymbol{B}) \star \mathbb{Z}^d \lhd \mathbb{Z}^d$).
We will explain the general idea by choosing $\Lambda_0 = \mathbb{Z}$ and $\Lambda = N\mathbb{Z}^d$, for $N \in \mathbb{N}$.
An N-periodic signal $(a_k)_{k\in\mathbb{Z}}$ has the property that $a_{k+N} = a_k$ for all $k \in \mathbb{Z}$, and hence it is enough to know the coefficients $\mathbf{a} = [a_0, a_1, ...a_{N-1}] \in \mathbb{C}^N$.
In fact, let us write $⧢_N$ for the Dirac Comb supported by $N\mathbb{Z}$, resp. the sequence which represents the unit vector $\mathbf{e}_0 \in \mathbb{C}^N$. Then we can write the given N-periodic sequence as a finite linear combinations of shifted Dirac-Combs:

$$(a_k)_{k\in\mathbb{Z}} = \sum_{j=0}^{N-1} a_j T_j ⧢_N \qquad (305)$$

perdiscrsign

Lamsdem1.jpg

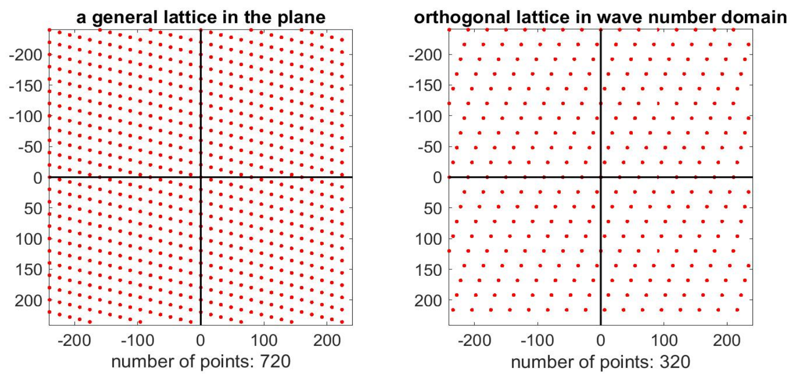

Abbildung 63: dualLamsdem1.jpg:
To the left we see a (periodic) lattice Λ, and to the right the corresponding *orthogonal lattice* $\Lambda^{\perp}$. In this case the lattice Λ has good density, so $\Lambda^{\perp}$ is somewhat sparse. Also the term "orthogonal lattice" may be understood for the case that that the subgroup is $\mathbb{R} \times \{0\} \lhd \mathbb{R}^2$, with orthogonal group $\{0\} \times \mathbb{R}$.

According to the general rules for the extended Fourier transform (as for the usual one) a shift goes over to a modulation by a pure frequency, i.e.

$$\mathcal{F}(T_j \sqcup\!\sqcup_N) = \chi_{-j} \cdot \mathcal{F}(\sqcup\!\sqcup_N) = \chi_{-j} \cdot (1/N) \sqcup\!\sqcup_{1/N} = exp(-2\pi i j/N)(1/N) \sqcup\!\sqcup_{1/N}. \quad (306)$$ shiftcomb

By summation over $j = 0, ..., N-1$ we get, using $\chi_{-j}(k/N) = e^{-2\pi i jk/N}$:

$$\mathcal{F}((c_k)_{k\in\mathbb{Z}}) = \frac{1}{N} \sum_{j=0}^{N-1} a_j e^{-2\pi i j(k/N)} \delta_{k/N} = \sum_{l=0}^{N-1} b_l T_{l/N} \sqcup\!\sqcup. \quad (307)$$ FTpersamp

Clearly any of these expressions is $\mathbb{Z}$-periodic (because the original measure was located over integer points), and therefore a shift by N steps of size $1/N$ brings the Fourier transform back to the original value. We thus only have to determine the first N coefficients which arise in the Fourier representation, **and it turns out that these coefficients, let us call the $\mathbf{b} = [b_0, b_1, ..., b_{N-1}]$ are exactly the DFT values of the original sequence $\mathbf{a} \in \mathbb{C}^N$** . It can also be interpreted as the values of the polynomial with the coefficient sequence **a** (described in the monomial basis in increasing order, i.e. $1, x, x^2, ..x^{N-1}$), evaluated over the unit roots of order N, starting at $1 = \omega^0$, in the clockwise sense:

$$b_k = \sum_{j=0}^{N-1} a_j e^{-2\pi i jk/N}, \quad 1 \leq k \leq N-1. \quad (308)$$ DFTformula

For the more general situation of a discrete signal supported on a lattice of the form $\alpha\mathbb{Z}^d$, but periodic with respect to a coarser lattice $\beta\mathbb{Z}^d = \alpha N \mathbb{Z}^d$, one finds that the Fourier

transform is supported by the lattice $(1/\beta)\mathbb{Z}^d$ and periodic with respect to $(1/\alpha)\mathbb{Z}^d = N(1/\beta)\mathbb{Z}^d$, the orthogonal lattices.
We suggest that the reader tries to do some MATLAB experiment in order to get a more intuitive understanding of the situation.

acCombs2.pdf

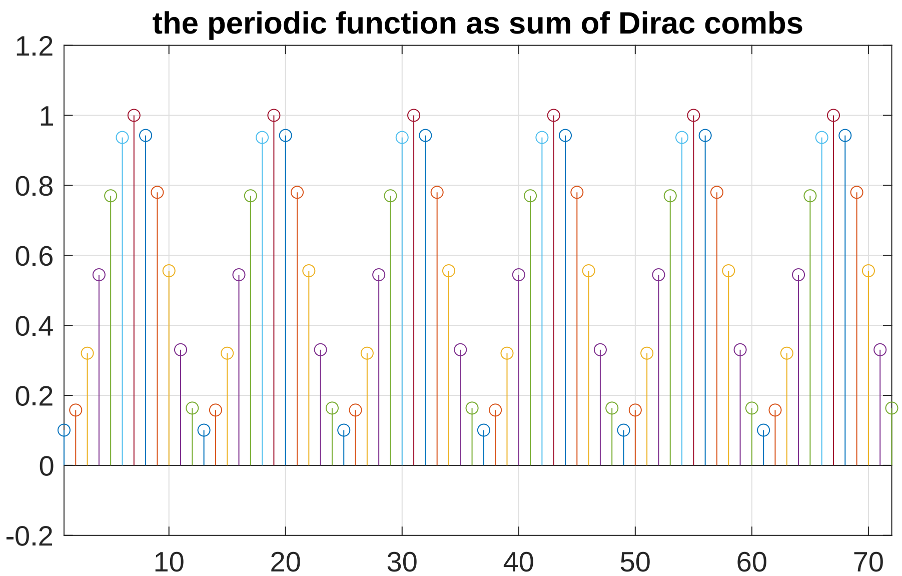

Abbildung 64: DiracCombs2.pdf:
Showing that periodic, discrete measure as a sum of (colored) Dirac combs.

39.1 BUPUs and interpolators

We have used already many times regular BUPUs, i.e. (well localized) partitions of unity consisting of translates of a given function $\psi \in \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) \cap \boldsymbol{C}_c(\mathbb{R}^d)$. Again, let us just discuss the case $d = 1$ and $\Lambda = \mathbb{Z} = \Lambda^\perp \lhd \mathbb{R}$, using $\mathcal{F}(\sqcup\!\sqcup) = \sqcup\!\sqcup$.

BUPUshahZ **Lemma 62.** *Let $\psi \in \boldsymbol{S}_0(\mathbb{R})$ be given, with compact support*[34]*. Then $(\psi_k)_{k\in\mathbb{Z}}$ constitutes a regular BUPUs if and only if $\widehat{\psi}$ has the interpolatory property:*

$$\widehat{\psi}(j) = \delta_{0,j}, \quad i.e.: \ \ \widehat{\psi}(0) = 1 \text{ and } \widehat{\psi}(j) = 0 \text{ for } j \in \mathbb{Z} \setminus \{0\}. \tag{309}$$ interpolZ

Proof. While the BUPU property is equivalent to the condition

$$\sqcup\!\sqcup \star \psi = \sum_{k\in\mathbb{Z}} T_k\psi = \mathbf{1}, \tag{310}$$ BUPUcond1

it is also *equivalent* via the Fourier transform side to

$$\delta_0 = \mathcal{F}(\mathbf{1}) = \mathcal{F}(\sqcup\!\sqcup \star \psi) = \widehat{\sqcup\!\sqcup} \cdot \widehat{\psi} = \sqcup\!\sqcup \cdot \widehat{\psi} = \sum_{k\in\mathbb{Z}} \widehat{\psi}(k)\delta_k \in \boldsymbol{M}_d(\mathbb{R}). \tag{311}$$ BUPUcond2

Comparing coefficients of these two discrete measures (resp. writing $\delta_0 = 1 \cdot \delta_0 + 0$) we observe the equivalence of conditions (310) and (311). ☐

Remark 36. We have seen interesting examples already of this kind, i.e. $\psi = \Delta$, with $\widehat{\psi} = \mathrm{SINC}^2$ (clearly interpolating). This has been used in the proof of Poisson's formula.

FUTURE MATERIAL. HGFEI

The cubic-B-spline generate the space of cubic spline functions, but they are not interpolating (only Δ has this property, but not any higher order B-spline). These functions are cubic polynomials over the intervals $[n, n+1], n \in \mathbb{Z}$, are continuous, and have continuous first and second order derivatives.

The idea of finding the (unique) interpolatory cubic spline is inspired by the idea of turning, on the Fourier transform side, the non-perfect BUPU into perfect BUPU by applying a division process:

We should also discuss the connection between periodic functions, functions on the torus, or functions which have been obtained by periodization. In order to keep the discussion simple let us treat that case of $\mathbb{Z}^d$-periodic functions on $\mathbb{R}^d$ (but everything is valid in the same way for Λ-periodic functions with respect to an arbitrary lattice $\Lambda \lhd \mathbb{R}^d$).

[34] This condition is the same as assuming: $\psi \in \boldsymbol{C}_c(\mathbb{R}^d)$ with $\widehat{\psi} \in \boldsymbol{L}^1(\mathbb{R})$.

riod-ization

Lemma 63. *Given a function $f \in \boldsymbol{S}_0(\mathbb{R}^d)$, the periodized version $f_{per} := \sqcup\!\sqcup_{\mathbb{Z}^d} \star f$ is a $\mathbb{Z}^d$-periodic function on $\mathbb{R}^d$ which has a Fourier transform (in the sense of $\boldsymbol{S}_0'(\mathbb{R}^d)$) of the form $\sqcup\!\sqcup_{\mathbb{Z}^d} \cdot \widehat{f} = \sum_{k \in \mathbb{Z}^d} \widehat{f}(k)\delta_k$. This ensures an absolutely convergent Fourier series expansion*

$$f_{per} = \sum_{n \in \mathbb{Z}^d} \widehat{f}(n)\chi_n.$$

Conversely, every such $\mathbb{Z}^d$-periodic function $h = \sum_{n \in \mathbb{Z}^d} c_n \chi_n$, with $\sum_{k \in \mathbb{Z}^d} |c_n| < \infty$, and

$$\|h\|_{\boldsymbol{A}(\mathbb{T}^d)} = \|(c_n)_{n \in \mathbb{Z}^d}\|_{\boldsymbol{\ell}^1(\mathbb{Z}^d)} < \infty$$

arises in this form.

Proof. The $\mathbb{Z}^d$-periodicity of f_{per} is obvious by direct inspectin. One could also argue that convolution commutes with translation, thus

$$T_n(\sqcup\!\sqcup_{\mathbb{Z}^d} \star f) = (T_n \sqcup\!\sqcup_{\mathbb{Z}^d}) \star f = \sqcup\!\sqcup_{\mathbb{Z}^d} \star f = f_{per}.$$

For the converse take any regular BUPU consisting of $\mathbb{Z}^d$-translates of some $\psi \in \boldsymbol{C}_c(\mathbb{R}^d) \cap \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) \subset \boldsymbol{S}_0(\mathbb{R}^d)$ (e.g. tensor products of triangular functions, or higher order B-spline on $\mathbb{R}^d$). Given the periodic function $h \in \boldsymbol{A}(\mathbb{T}^d)$ we can set $f := h \cdot \psi$. Then we have

$$\|f\|_{\boldsymbol{S}_0} = \|\sum_{n \in \mathbb{Z}^d} c_n M_n \psi\|_{\boldsymbol{S}_0} \leq \sum_{n \in \mathbb{Z}^d} |c_n| \|M_n \psi\|_{\boldsymbol{S}_0} \leq \|(c_n)\|_{\boldsymbol{\ell}^1(\mathbb{Z}^d)} \cdot \|\psi\|_{\boldsymbol{S}_0}. \quad (312)$$ hperloc1

$$f_{per} = \sqcup\!\sqcup_{\mathbb{Z}^d}(h \cdot \psi) = \sum_{k \in \mathbb{Z}^d} (T_k h) \cdot (T_k \psi) = \sum_{k \in \mathbb{Z}^d} h \cdot T_k \psi = h \cdot \sum_{k \in \mathbb{Z}} T_k \psi = h. \quad (313)$$ hlocperiod

$\square$

anning72.jpg

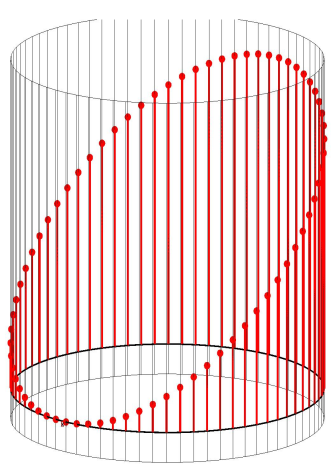

Abbildung 65: stemchanning72.jpg:
JUST A TEST image for now: visualization of a periodic, discrete function as a discrete function whose graph appears on the cylinder.

39.2 PREVIEW: The Banach GELFAND TRIPLE

Trip00XX.jpg

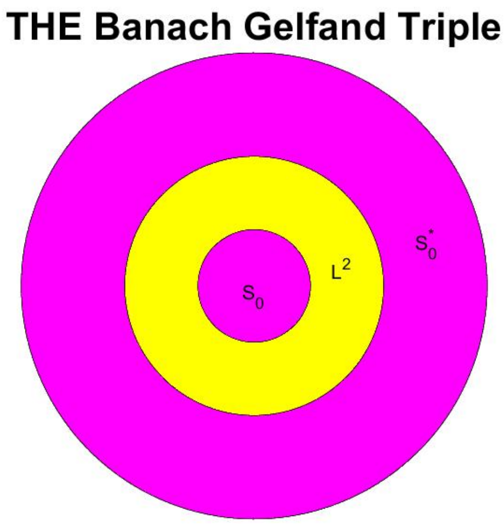

Abbildung 66: BanGelTripXX.jpg

For the purpose of teaching, this triple of spaces can be compared with the number system $\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$, where multiplication and inversion are quite easy in the rational domain, since obviously

$$\frac{3}{4} \cdot \frac{4}{3} = \frac{12}{12} = 1!$$

In a similar way the Fourier inversion theorem or Poisson's formula are quite easy for the $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$-setting (with pointwise convergence, even for Riemann integrals).

The real numbers, resp. the Hilbert space (here $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$) are "more complete than the rationals. We know that $\sqrt{2}$ is well defined in $\mathbb{R}$. Multiplication for *infinite decimals*, meaning for the Cauchy-sequence of longer and longer decimal approximations, is defined multiplying longer and longer approximations and taking the limit of their product. Even inversion can be justified by some limiting procedure (i.e. by taking the limit of the inverses, as they appear in $\mathbb{Q}$). In this way the **symbols** such as $1/\sqrt{2}$, or $1/\pi^2$ are well defined inside of $\mathbb{R}$.

Finally we take the complex numbers as an even larger domain, but with the usual operators (known now for $\mathbb{R}$) being extended in a unique fashion, thus arriving at a complete *field* of complex numbers. It is in the complex domain (now in the environment of "mild distributios") where we obtain the greatest freedom (perhaps with the most "abstract looking objects").

40 FFT for periodic discrete measures

FFTperdiscr

There are many reasons, why w^*-convergence is important and already implicitly used in many contexts and heuristic derivations are turned into solid, mathematical statements. Just to warm up let us consider a few concrete/easy examples, keeping in mind that

$$St_\alpha(\sqcup\!\!\sqcup) = \sqcup\!\!\sqcup_\alpha = \sum_{k\in\mathbb{Z}^d} \delta_{\alpha k}, \tag{314}$$

dilateshah1

is the "mass preserving" method of dilating functions resp. measures (point measures stay point measures and just the underlying points are moved, here from k to αk). The argument is clear, once we know that the (pre-)adjoint operator on $\left(\boldsymbol{W}(\boldsymbol{C}_0,\boldsymbol{\ell}^1)(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{W}}\right)$, namely the operator D_α, is uniformly bounded, hence the operators $\mathrm{St}_\alpha, \alpha \geq 1$ are uniformly bounded in the range of parameters $\alpha \in [1,\infty)$ (compare Lemma S0dilatinv 44 for such arguments).

Thus it is enough to observe that $\sqcup\!\!\sqcup_\alpha(f) = f(0)$ for all sufficiently large values of α, if $f \in \boldsymbol{S}_0(\mathbb{R}^d) \cap \boldsymbol{C}_c(\mathbb{R}^d)$, resp. $f \in \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) \cap \boldsymbol{C}_c(\mathbb{R}^d)$. Hence we have shown:

stretchshah1

Lemma 64.

$$w^*\text{-}\lim_{\alpha\to\infty} \sqcup\!\!\sqcup_\alpha = \delta_0, \tag{315}$$

strechshah

or equivalently

$$\lim_{\alpha\to\infty} \sum_{k\in\mathbb{Z}^d} f(\alpha k) = f(0), \quad \textit{for} \quad f \in \boldsymbol{W}(\boldsymbol{C}_0,\boldsymbol{\ell}^1)(\mathbb{R}^d).$$

This statement is also interesting by comparison with the domain $\boldsymbol{C}_0(\mathbb{R}^d)$. Note that for any $a > 0$ the discrete (*translation-bounded*) measure $\sqcup\!\!\sqcup_a$ is *unbounded* on $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$ (obviously), hence it only makes sense to speak of w^*-convergence within the $(\boldsymbol{W}(\boldsymbol{C}_0,\boldsymbol{\ell}^1), \boldsymbol{W}(\boldsymbol{M},\boldsymbol{\ell}^\infty))-$du
On the other hand the continuous embedding $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right) \hookrightarrow \left(\boldsymbol{W}(\boldsymbol{C}_0,\boldsymbol{\ell}^1)(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{W}}\right)$ implies that in particular the convergence takes place in $\boldsymbol{S}_0'(\mathbb{R}^d)$, endowed with the $wst-$topology (i.e. in the $\sigma(\boldsymbol{S}_0', \boldsymbol{S}_0)$-topology).

Since the Fourier transform of a distribution of the form $St_\alpha\sigma$ is just $D_\alpha\hat{\sigma}$ we obtain as a corollary to Lemma stretchshah1 64:

Rsumsconv1

Lemma 65.

$$b \cdot \sqcup\!\!\sqcup_b \to 1 \quad \textit{in the } w^* - \textit{topology, for} \quad b \to 0. \tag{316}$$

shahRiemconv

In fact, the result can be restated in the following sense. For $f \in \boldsymbol{W}(\boldsymbol{C}_0,\boldsymbol{\ell}^1)(\mathbb{R}^d)$, i.e. a continuous function which is absolutely R-integrable, the Riemannian sums are actually convergent to the R-integral (or the Lebesgue integral) of the given test function f, as the "step-width" of the R-sums tends to 0, again in the $(\boldsymbol{W}(\boldsymbol{C}_0,\boldsymbol{\ell}^1), \boldsymbol{W}(\boldsymbol{M},\boldsymbol{\ell}^\infty))-$duality sense.

41 Approximating different settings

SOPwst2.jpg

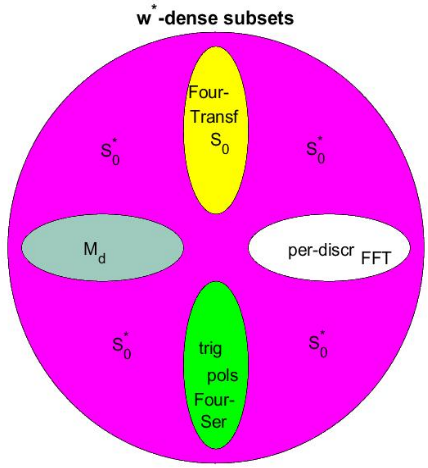

Abbildung 67: SOPwst2.jpg:Different w^*–dense subsets of $\boldsymbol{S}_0'(\mathbb{R}^d)$.

We have now established that four quite different subsets of $\boldsymbol{S}_0'(\mathbb{R}^d)$ are in fact w^*-dense in $\boldsymbol{S}_0'(\mathbb{R}^d)$. Each of these sets has a specific realization of the Fourier transform

1. For the "nice" subalgebra $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ this is course the classical Fourier transform (which can be written as integral expression, using Riemann integral, and which also allows the pointwise recovery of f from $\widehat{f}$ via the Fourier inversion formula).

2. For discrete, non-periodic measure we just have to apply *Fourier synthesis*, which maps discrete measures into trigonometric polynomials, or $\boldsymbol{\ell}^1(\mathbb{Z})$ into the space of *absolutely convergent Fourier series*, which can be identified with Wiener's algebra (another Wiener algebra) $\left(\boldsymbol{A}(\mathbb{T}), \|\cdot\|_{\boldsymbol{A}}\right)$, with the $\boldsymbol{\ell}^1$-norm.

3. The Fourier image of discrete measures, say from $\boldsymbol{\ell}^2(\mathbb{Z})$, we get $\boldsymbol{L}^2(\mathbb{T})$, the classical space of square-integrable $\alpha\mathbb{Z}^d$-periodic functions. This is a Hilbert space with respect to the norm obtained from the scalar product

$$\langle f, g\rangle_{\boldsymbol{L}^2} = \int_{[0,\alpha]^d} f(x)\overline{g(x)}dx.$$

4. For periodic and discrete measures we have the DFT/FFT algorithm in order to determine the Fourier transform, which is then a discrete and periodic (which is also periodic and discrete) measure on the Fourier domain.

First we claim that $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ can be "approximated by periodic versions":

$$f = w^*\text{-}\lim_{a\to\infty} f \star ⧢_a \quad \forall f \in \boldsymbol{S}_0(\mathbb{R}^d). \tag{317}$$ pertononper1

and consequently the same is true the Fourier transform side, i.e. approximation of $f \in \boldsymbol{S}_0(\mathbb{R}^d) \subset \boldsymbol{M}_b(\mathbb{R}^d)$ by discrete measure concentrated on a lattice:

$$\hat{f} = w^*\text{-}\lim_{b\to 0} (b \cdot ⧢_b \cdot \hat{f}), \quad \forall f \in \boldsymbol{S}_0(\mathbb{R}^d). \tag{318}$$ pertononper1f

The interpretation is the following. Choosing e.g. $b = 2^{-n}$ for $n \to \infty$ one can consider the values of the Fourier coefficients of the periodic functions $f \star ⧢_{2^n}$ (up to the normalization factor) as the amplitudes of the sequence of $\delta-$distributions on the integer lattice $2^{-n}\mathbb{Z}$.
In this sense the non-periodic case is the limit of the a-periodic cases, for $a \to \infty$.
On the other hand, we have seen how to approximate the continuous case by spline-type approximations (or by discrete measures). It is important to have the following (non-trivial) result (see feka07 43):

S0quasint **Lemma 66.**

$$\textit{For any } f \in \boldsymbol{S}_0(\mathbb{R}^d) : \|\mathrm{Sp}_\Psi f - f\|_{\boldsymbol{S}_0} \to 0 \quad \textit{for} \quad |\Psi| \to 0. \tag{319}$$ S0quasint1

By taking the adjoint operators on $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$ on derives:

S0Pdiscr **Lemma 67.**

$$\sigma = w^*\text{-}\lim_{|\Psi|\to 0} \mathrm{D}_\Psi \sigma \quad \forall \quad \sigma \in \boldsymbol{S}_0', \tag{320}$$ S0Pdiscr1

and consequently any $\sigma \in \boldsymbol{S}_0'$ can be constructively approximated (in the w^*-sense) by discrete (translation-bounded) measures[35] *of the form* $\mathrm{D}_\Psi(\sigma) = \sum_{k\in\mathbb{Z}^d} \sigma(\psi_k)\delta_{\gamma k}$.

There are also ways to "come back" from the discrete/periodized version of a test function to the test function itself. So first it is enough to check that it is possible to extract $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ from its a-periodized versions (for large $a > 0$). Since periodization of a function f corresponds to sampling of $\hat{f}$ on its (inverse) Fourier transform side, it is enough to show that a good approximate reconstruction can be obtain from a sufficiently dense sampling sequence ([43]):

reconstS0 **Theorem 15.** *For $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ one has for $b = 1/a, a \to 0$:*

$$\|f - (⧢_a \cdot f) \star D_b\Delta\|_{\boldsymbol{S}_0(\mathbb{R}^d)} \to 0.$$

Remark 37. An equivalent view on the above lemma (moving the normalization constant between the factors appropriately, and writing St_a as $a^{-1}D_b$ (?,check) describes the situation by

$$(D_b ⧢ \cdot f) \star St_b\Delta \to (1 \cdot f) \star \delta_0 = f \text{ in} \boldsymbol{S}_0(\mathbb{R}).$$

The PROOF of the above is highly non-trivial (although the fact is very conveniently used), based on properties of Wiener amalgam spaces.
The (approximate) recovery of $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ from any periodized version, in fact from $f \star ⧢_\Lambda$, for $\Lambda = A * \mathbb{Z}^d$, A a dilation matrix, can be obtained by localization. The only property needed required is that one multiplies with a function, which is a Lagrange interpolator in $\boldsymbol{S}_0(\mathbb{R}^d)$, (i.e.with $h(0) = 1, h(\lambda) = 0$ for $\lambda \neq 0$).

[35] This family is bounded in the $\boldsymbol{W}(\boldsymbol{M}, \boldsymbol{\ell}^\infty)$-sense if and only if $\sigma \in \boldsymbol{W}(\boldsymbol{M}, \boldsymbol{\ell}^\infty)(\mathbb{R}^d)$.

othplat1.eps

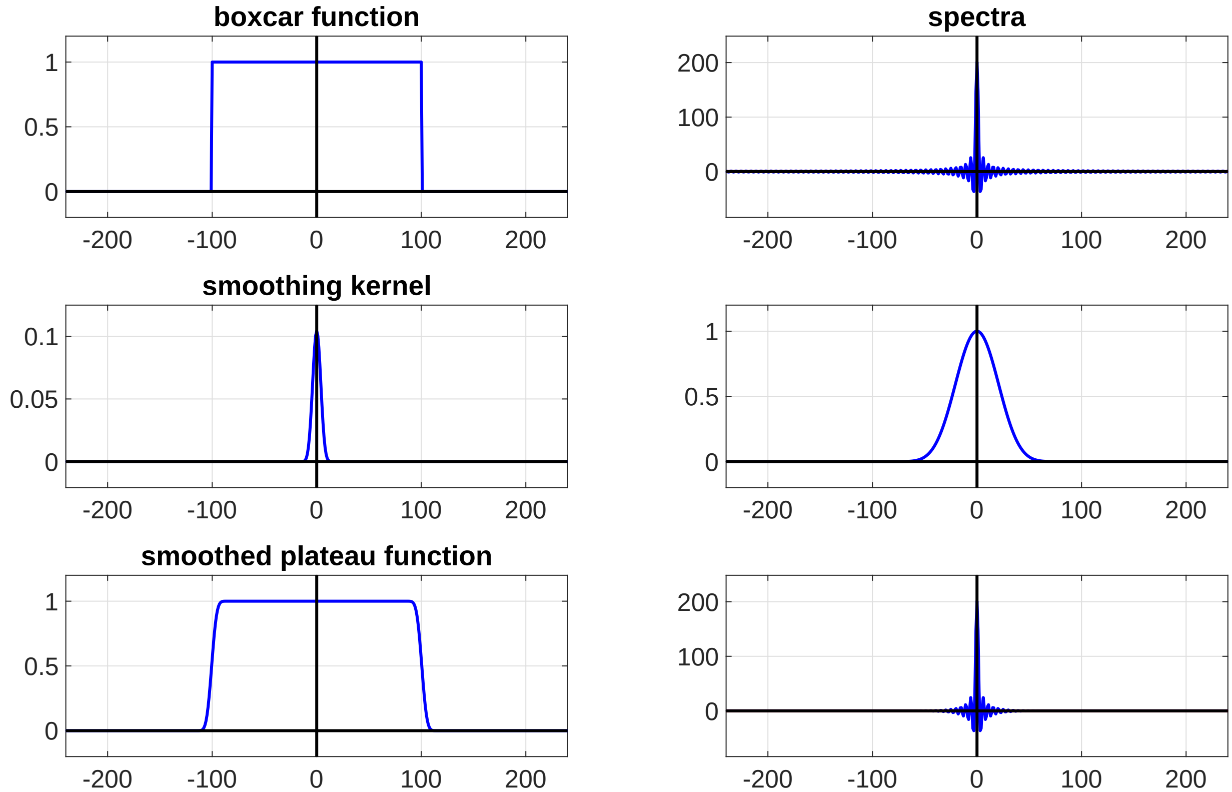

Abbildung 68: smoothplat1.eps

locFTsegal

Lemma 68. *For any compact set $M \subset \mathbb{R}^d$ (e.g. $M = B_R(0)$) the restriction of $\mathcal{F}(\boldsymbol{S}_0(\mathbb{R}^d))$ coincides with $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ resp. $\mathcal{F}(\boldsymbol{M}_b(\mathbb{R}^d))$.*

This results follows from the fact that $\boldsymbol{M}_b(\mathbb{R}^d) \star \boldsymbol{S}_0(\mathbb{R}^d) \subset \boldsymbol{S}_0(\mathbb{R}^d)$ and the existence of smooth plateau functions in $\boldsymbol{S}_0(\mathbb{R}^d) = \mathcal{F}(\boldsymbol{S}_0(\mathbb{R}^d))$.

42 The Hilbert space $\left(\boldsymbol{L}^2(\mathbb{R}^d),\ \|\cdot\|_2\right)$

Let us now come to the introduction ("our own way") of the Hilbert space $\left(\boldsymbol{L}^2(\mathbb{R}^d),\ \|\cdot\|_2\right)$, which is a Banach space with respect to the scalar product

$$\langle f, g\rangle_{\boldsymbol{L}^2(\mathbb{R}^d)} := \int_{\mathbb{R}^d} f(t)\overline{g(x)}dx, \quad f, g \in \boldsymbol{C}_c(\mathbb{R}^d), \tag{321}$$ ScalProdLt

which turns $\boldsymbol{C}_c(\mathbb{R}^d)$ into a normed space with respect to the *associated norm*

$$\|f\|_{\boldsymbol{L}^2} := \sqrt{\langle f, f\rangle_{\boldsymbol{L}^2}} \tag{322}$$ Ltnorm00

The Cauchy-Schwarz inequality

$$|\langle f, g\rangle_{\boldsymbol{L}^2(\mathbb{R}^d)}| \leq \|f\|_{\boldsymbol{L}^2}\|g\|_{\boldsymbol{L}^2} \tag{323}$$ CSinequ

implies that the triangle inequality is valid:

$$\|f+g\|_{\boldsymbol{L}^2} \leq \|f\|_{\boldsymbol{L}^2} + \|g\|_{\boldsymbol{L}^2}, \quad f, g \in \boldsymbol{C}_c(\mathbb{R}^d). \tag{324}$$ Lttriang

By a standard construction in functional analysis one can embed $\boldsymbol{C}_c(\mathbb{R}^d)$ into a Banach space, in fact even a Hilbert space with the extended scalar product (by applying the principle of bounded bilinear mappings) to a space which is denoted by $\left(\boldsymbol{L}^2(\mathbb{R}^d),\ \|\cdot\|_2\right)$ (we write $\boldsymbol{L}^2$-FA version). This procedure is considering the vector space of *equivalence classes of Cauchy sequences* $[(f_n)_{n\geq 1}]$ in $\boldsymbol{C}_c(\mathbb{R}^d)$. Two Cauchy sequences $(f_n)_{n\geq 1}$ and $(g_n)_{n\geq 1}$ are called *equivalent* (w.r.t. the $\boldsymbol{L}^2$-norm) if they have almost in-distinguishable tails (verbal description), or technically speaking, if

$$\forall\, \varepsilon > 0\ \exists n_o \in \mathbb{N} \ \text{ such that } \ n, m \geq n_o \Rightarrow \|f_n - g_m\|_{\boldsymbol{L}^2} \leq \varepsilon. \tag{325}$$ CSequivLt

We use square brackets to denote the *equivalence classes*, i.e. the elements of the *completion* of $(\boldsymbol{C}_c(\mathbb{R}^d), \|\cdot\|_2)$.
Endowed with the norm (shown to work well)

$$\|[(f_n)_{n\geq 1}]\| := \lim_n \|f_n\|_{\boldsymbol{L}^2} \tag{326}$$ Ltlimnorm

it is a complete normed space (hence Banach space), with a norm derived from the (extended) scalar product, given in a similar way (e.g. by polarization):

$$\langle [(f_n)_{n\geq 1}], [(g_n)_{n\geq 1}]\rangle := \lim_{n\to\infty} \langle f_n, g_n\rangle_{\boldsymbol{L}^2} \tag{327}$$ Scalplimnorm

i.e. it is a *Hilbert space* (we do not give the axiomatic description here).
It is an easy exercise to verify that one has, by the density of $\mathcal{F}\boldsymbol{L}^1 \cap \boldsymbol{C}_c(\mathbb{R}^d)$ in $\boldsymbol{C}_c(\mathbb{R}^d)$:

LTSOcompl **Lemma 69.** *$\boldsymbol{L}^2$-FA coincides with the completion of $(\mathcal{F}\boldsymbol{L}^1 \cap \boldsymbol{C}_c(\mathbb{R}^d), \|\cdot\|_2)$.*

Proof. This is a well known fact from functional analysis, but can be explained in simple words: Any equivalence class $[(f_n)]$ has a *representative* (g_n) inside of the dense subspace $\mathcal{F}\boldsymbol{L}^1 \cap \boldsymbol{C}_c(\mathbb{R}^d)$. Just choose any g_n in this subspace with $\|g_n - f_n\|_{\boldsymbol{L}^2} < 1/n$, $n \geq 1$ and check for equivalence. $\square$

There is the alternative to describe $\left(\boldsymbol{L}^2(\mathbb{R}^d),\ \|\cdot\|_2\right)$ as the space of (equivalence classes) of measurable functions which are square integrable over $\mathbb{R}^d$ with respect to the Lebesgue measure. Here *equivalence* of to such measurable function refers to the to condition

$$f(x) = g(x) a.e. \text{ (almost everywhere)}, \tag{328}$$

equivmeas1

i.e.

$$\{x \mid f(x) \neq g(x)\} \ \text{ is a set of Lebesgue measure zero.} \tag{329}$$

equivmeas2

To be on the save side: A set $N \subset \mathbb{R}^d$ is a set of measure zero if for every $\varepsilon > 0$ there exists a sequence of cubes (of different size and position) Q_k of total volume $\sum_{k=1}^{\infty} \mathrm{Vol}(Q_k) \leq \varepsilon$ such that $N \subseteq \bigcup_{k=1}^{\infty} Q_k$.
The Lebesgue approach to $\left(\boldsymbol{L}^2(\mathbb{R}^d),\ \|\cdot\|_2\right)$ is given as follows:

LtLebdef **Definition 13.**

$$\boldsymbol{L}^2(\mathbb{R}^d)-\mathrm{Leb} := \{f \mid \|f\|_{\boldsymbol{L}^2} := \left(\int_{\mathbb{R}^d} |f(x)|^2\right)^{1/2} < \infty\}. \tag{330}$$

LtLebdef1

Using the arguments from *measure theory* one show that the *measurable step functions* and from there in fact the space of continuous functions with compact support form a dense subspace of the Hilbert space $\left(\boldsymbol{L}^2(\mathbb{R}^d),\ \|\cdot\|_2\right)$. In this way the space is isometrically isomorphic to the completion $\boldsymbol{L}^2 - FA$ described above, i.e. all the properties of the corresponding Hilbert space are exactly the same, so that one can identify one with the other, actually in the same way as the Lebesgue space $\boldsymbol{L}^1(\mathbb{R}^d)$ can be viewed as a closed subspace of $\boldsymbol{M}_b(\mathbb{R}^d)$ (by the usual identification $g \to \mu_g$), but now *realized* with respect to the Lebesgue integral.
Just recall that we have (only measurable functions are considered): by the Lebesgue approach to $\left(\boldsymbol{L}^1(\mathbb{R}^d),\ \|\cdot\|_1\right)$ is given as follows:

LiLebdef **Definition 14.**

$$\boldsymbol{L}^1(\mathbb{R}^d)-\mathrm{Leb} := \{f \mid \|f\|_{\boldsymbol{L}^1} := \int_{\mathbb{R}^d} |f(x)| < \infty.\} \tag{331}$$

LiLebdef1

Within a distributional setting we have another (also equivalent) criterion (which is easily verified in the context of a measure theory course, we just provide the statement to proved some form of understanding for the logical connections[36]):

quivbyaction **Lemma 70.** *Assume that we have $g, h \in \boldsymbol{L}^1(\mathbb{R}^d) + \boldsymbol{L}^2(\mathbb{R}^d)$ (i.e. $g = g_1 + g_2$, with $g_1 \in \boldsymbol{L}^1(\mathbb{R}^d)$ and $g_2 \in \boldsymbol{L}^2(\mathbb{R}^d)$, same for h) are given, such that the induced functionals σ_f and σ_h coincide, i.e. with*

$$\int_{\mathbb{R}^d} f(x)g(x)dx = \int_{\mathbb{R}^d} f(x)h(x)dx, \quad f \in \boldsymbol{S}_0 \tag{332}$$

equivsigfh

if and only if $g(x) = h(x)$ a.e.. In fact, it is enough to have (332) for any dense subspace of $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$, e.g. for $f \in \boldsymbol{C}_c(\mathbb{R}^d) \cap \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$.

[36] We use the standard abbrevitoin a.e. for “almost everywhere”, meaning that the exceptional set has Lebesgue measure zero.

43 $\boldsymbol{L}^2(\mathbb{R}^d)$ as part of a Banach Gelfand Triple

In the context of Lebesgue theory the Cauchy-Schwarz inequality reads as follows:

$$\int_{\mathbb{R}^d} |f(x)||g(x)|dx \leq \|f\|_{\boldsymbol{L}^2}\|g\|_{\boldsymbol{L}^2}, \quad f, g \in \boldsymbol{L}^2(\mathbb{R}^d). \tag{333}$$ CSintRd

This implies:

LtinSOP **Proposition 17.** *We have a continuous chain of embeddings*

$$\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right) \hookrightarrow (\boldsymbol{L}^2 - \text{Leb}, \|\cdot\|_2) \hookrightarrow (\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'}), \tag{334}$$ SOLtSOPemb1

with density in the first place and w^*-density for the second part of the continuous embedding.

Proof. We know that $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ is continuously embedded into both $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ and $\left(\boldsymbol{W}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{W}}\right) \hookrightarrow \left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$. Hence it is clear (using Riemann integrals) that we have (in fact even for $f \in \boldsymbol{W}(\mathbb{R}^d)$):

$$(\|f\|_{\boldsymbol{L}^2})^2 = \int_{\mathbb{R}^d} |f(x)|^2 \leq \|f\|_{\boldsymbol{L}^1}\|f\|_\infty \leq C^2\|f\|_{\boldsymbol{S}_0}{}^2, f \in \boldsymbol{S}_0(\mathbb{R}^d), \tag{335}$$ LtWRdest1

for some positive constant $C > 0$, so that we get the required estimate

$$\|f\|_{\boldsymbol{L}^2} \leq C\|f\|_{\boldsymbol{S}_0}, \quad f \in \boldsymbol{S}_0, \tag{336}$$ LtSOest1

providing the first part of the chain of continuous embeddings.
The second one stems from the fact that we can employ the CS-inequality in the form of (CSintRd 333) in order to find that on has for $g \in \boldsymbol{L}^2 - \text{Leb}$

$$|\sigma_g(f)| = \left|\int_{\mathbb{R}^d} f(x)g(x)dx\right| \leq \int_{\mathbb{R}^d} |f(x)| \cdot |g(x)|dx \leq C\|g\|_{\boldsymbol{L}^2}\|f\|_{\boldsymbol{S}_0}, \quad f \in \boldsymbol{S}_0, \tag{337}$$ LtintoSOP

and by consequence that we can estimate the norm of σ_g in $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$ by

$$\|\sigma_g\|_{\boldsymbol{S}_0'} \leq C\|g\|_{\boldsymbol{L}^2}, \quad g \in \boldsymbol{L}^2 - \text{Leb}. \tag{338}$$ LtSOPestim1

□

For us this last chain of embeddings will be a motivation to introduce $\boldsymbol{L}^2(\mathbb{R}^d)$ as a subspace of $\boldsymbol{S}_0'(\mathbb{R}^d)$, in order to establish the **Banach Gelfand Triple** $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$.

43.1 $\boldsymbol{L}^2(\mathbb{R}^d)$ as space of mild distributions

Motivated by the above results found in the literature we provide a *new, but equivalent definition* of $\left(\boldsymbol{L}^2(\mathbb{R}^d),\ \|\cdot\|_2\right)$, as a subspace of $\boldsymbol{S}_0'(\mathbb{R}^d)$:

LtSOPdef **Definition 15.**

$$\boldsymbol{L}^2(\mathbb{R}^d) := \{\sigma \in \boldsymbol{S}_0'(\mathbb{R}^d)\,|\,\sigma = w^*\text{-}\lim{}_n f_n \text{ for some CS}(f_n)_{n=1}^\infty \text{ in } \boldsymbol{S}_0(\mathbb{R}^d) \text{ w.r.t.} \|\cdot\|_2.\} \tag{339}$$ LtSOPdef1

We will use the symbol $\boldsymbol{L}^2 - \mathrm{SOP}(\mathbb{R}^d)$ for a while. As a natural norm we can take

$$\|\sigma\|_{\boldsymbol{L}^2} = \|\sigma\|_{\boldsymbol{L}^2-\mathrm{SOP}} = \lim{}_{n=1}^\infty \|f_n\|_{\boldsymbol{L}^2}. \tag{340}$$ LtSOPnorm

In this way it is clear that we have a well defined subspace of mild distributions $\boldsymbol{L}^2 - \mathrm{SOP}$ which is however “the same space” that we have considered in the previous proposition (where we see how to view $\boldsymbol{L}^2 - \mathrm{Leb}$ as a subspace of $\boldsymbol{S}_0'(\mathbb{R}^d)$).
The advantage of our approach is that we avoid the discussion of equivalence classes, but have right away explicit identification of the given element of $\boldsymbol{L}^2$ with mild distributions, i.e. with elements of $\boldsymbol{S}_0'(\mathbb{R}^d)$.

LTidentif **Theorem 16.** *The space $\boldsymbol{L}^2 - \mathrm{SOP}$ with the norm given by the norm (340) is a Hilbert space, continuously embedded into $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$, which is isometrically isomorphic to $\boldsymbol{L}^2 - \mathrm{Leb}$ (and thus to $\boldsymbol{L}^2 - FA$).*

Proof. There are a couple of things to be done, mostly involving standard considerations in the context of functional analysis. We are only giving some indications of what has to be done.
First of all we observe that the CS-inequality implies that for each Cauchy-sequence $(f_n)_{n=1}^\infty$ in $\boldsymbol{S}_0(\mathbb{R}^d)$ (with respect to $\|\cdot\|_2$!) the corresponding functionals σ_{f_n} produce Cauchy-sequences in $\mathbb{C}$ for every $f \in \boldsymbol{S}_0(\mathbb{R}^d)$, and thus $\sigma(f) := \lim_{n\to\infty}\sigma_{f_n}(f)$ is a well defined linear functional on $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$. Since any Cauchy sequence with respect to the $\|\cdot\|_2$-norm is of course bounded with respect to this norm and hence bounded in the $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$-sense one even finds an estimate, namely

$$|\sigma(f)| \le C\|f\|_{\boldsymbol{S}_0}, \quad f \in \boldsymbol{S}_0(\mathbb{R}^d).$$

Thus Definition 15 is meaningful.
That the norm is well defined, because the Cauchy condition on the sequence (f_n) implies that $(\|f_n\|_{\boldsymbol{L}^2})_{n=1}^\infty$ is a Cauchy sequence in $\mathbb{R}$ (having a well defined limit in $\mathbb{R}$):

$$|\|f_n\|_{\boldsymbol{L}^2} - \|f_m\|_{\boldsymbol{L}^2}| \le \|f_n - f_m\|_{\boldsymbol{L}^2} \to 0 \quad \text{for} \quad n, m \to \infty. \tag{341}$$ Cauchnorms1

□

The completeness is a bit more cumbersome but can be verified by elementary means, namely by showing that absolutely convergent sums define an element in $\boldsymbol{L}^2 - \mathrm{SOP}$ together with the expected norm estimate, valid for and $\varepsilon > 0$:

$$\|\sigma\|_{\boldsymbol{L}^2-\mathrm{SOP}} \le \sum_{k=1}^\infty \|\sigma_k\|_{\boldsymbol{L}^2-\mathrm{SOP}} + \varepsilon.$$

The continuous chain of inclusions is then shown in a similar way as in proposition LtinSOP 17, with slightly modified arguments.
That each $\boldsymbol{L}^2$-Cauchy sequence in $\boldsymbol{S}_0(\mathbb{R}^d)$ has a well defined limit in $\boldsymbol{L}^2 - \text{Leb}$ is clear. Conversely, we can obtain for each $f \in \boldsymbol{L}^2 - \text{Leb}$ some $\boldsymbol{L}^2 - \text{Leb}$-Cauchy-sequence in $\boldsymbol{S}_0(\mathbb{R}^d)$ by approximating f by an expression of the form $(f \cdot \mathrm{D}_\rho h) \star \mathrm{St}_\rho g$, for suitably normalized functions $h, g \in \boldsymbol{C}_c(\mathbb{R}^d) \cap \boldsymbol{S}_0(\mathbb{R}^d)$ with $h(0) = 1$ and $\widehat{g}(0) = 1$, and ρ sufficiently small.
In fact, by the Cauchy Schwartz inequality we have $f \cdot \mathrm{D}_\rho) \subset \boldsymbol{L}^2 - \text{Leb} \cdot \boldsymbol{L}^2 - \text{Leb} \subset \boldsymbol{L}^1 - \text{Leb}$, and thus $(f \cdot \mathrm{D}_\rho h) \star g \in \boldsymbol{L}^1(\mathbb{R}^d) \star \boldsymbol{S}_0(\mathbb{R}^d) \subset \boldsymbol{S}_0(\mathbb{R}^d)$.
The convergence (hence to possibility of producing a Cauchy sequence comes from the fact that we have for $f \in \boldsymbol{L}^2 - \text{Leb}$:

$$\|f - f \cdot \mathrm{D}_\rho h\|_{\boldsymbol{L}^2} \to 0 \quad \text{for} \quad \rho \to 0, \text{ if } h(0) = 1, \tag{342}$$ DrhoLtLeb1

and the corresponding statement for convolutions

$$\|f - f \star \mathrm{St}_\rho g\|_{\boldsymbol{L}^2} \to 0 \quad \text{for} \quad \rho \to 0, \text{ if } \widehat{g}(0) = 1. \tag{343}$$ DrhoLtLeb2

Thus by choosing any such pair $g, h \in \boldsymbol{C}_c(\mathbb{R}^d) \cap \boldsymbol{S}_0(\mathbb{R}^d)$ and any sequence $(\rho_n)_{n=1}^\infty$ we obtain a CS-sequence as required by definition LtSOPdef 15, which then of course (in the setting of $\boldsymbol{L}^2 - \text{Leb}$ is identified with $f \in \boldsymbol{L}^2 - \text{Leb}$, while it is identified with $\sigma_f \in \boldsymbol{S}_0'$ (according to Prop. LtinSOP 17) otherwise.
The fact that this identification is an isometric one is easily verified. Thus we can continue with the new definition given in Def. LtSOPdef 15. It is also true that the chain of embeddings can be verified (more or less by the same arguments, just restricting the estimates to the elements of the approximating Cauchy-sequences and then going to the limit. Thus, also for the new (but equivalent) definition of $\boldsymbol{L}^2(\mathbb{R}^d)$ we have

SOLtnewSOP **Lemma 71.** *We have the following chain of inclusion:*

$$\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right) \hookrightarrow \left(\boldsymbol{L}^2 - \text{SOP}, \|\cdot\|_{\boldsymbol{L}^2-\text{SOP}}\right) \hookrightarrow (\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'}).$$

From now on we will again abolish the attached symbols, like SOP, FA, or Leb, and just talk about $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$
Based on this chain of inclusions we will introduce the following terminology:

Definition 16. The triple $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$ will be called THE Banach Gelfand Triple (based on $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$).

It is THE BGTr which is relevant for Fourier Analysis. Clearly it is a triple of Banach spaces forming a natural chain of inclusions, but there are many more such triples in the literature, which motivates the following more general definition:

44 Banach Gelfand Triples

The goal is to simplify the situation for a discussion of the Fourier transform.

BanGelTrip.jpg

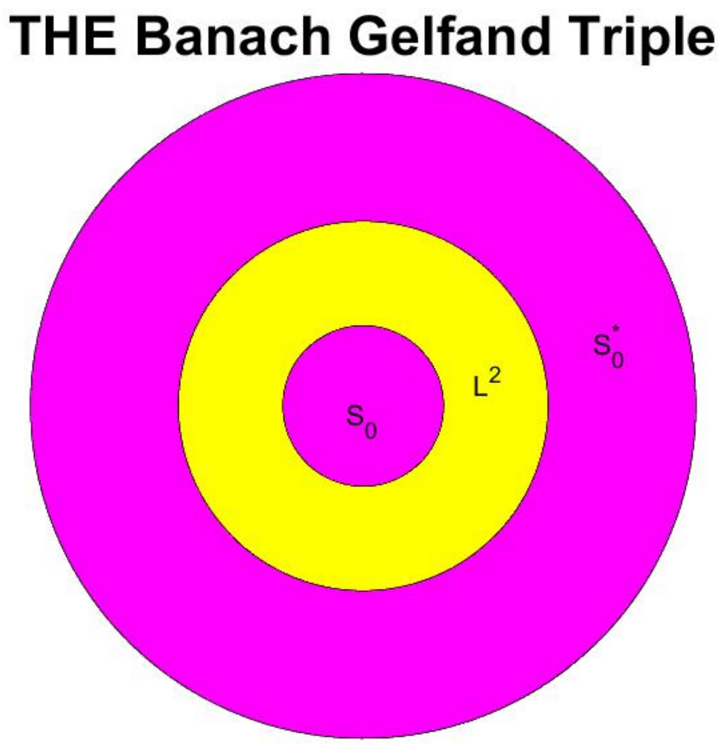

Abbildung 69: BanGelTrip.jpg:
The simple world of the BANACH GELFAND TRIPLE, just consisting of the Segal algebra $\boldsymbol{S}_0(\mathbb{R}^d)$ as a space of "nice test functions" and the dual space, as ambient vector space of "mild distributions", containing everything which is needed for a discussion of digital signal processing, and finally, in the middle, sandwiched by the two, the most beautiful Hilbert space $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$.

DEFINITION: Banach Gelfand Triple
There is quite a large number of talks by the author reporting on the idea of Banach Gelfand Triples, see

`www.nuhag.eu/talks`

The access to the talks there is possible using the username "visitor" and the access code "nuhagtalks". Please contact the author if there is problem to access the talks there. One can also load several "topical queries" e.g. about Banach Gelfand Triples, Conceptual Harmonic Analysis or Banach Gelfand Triples.

45 Discrete/continuous norms on Wiener amalgams

scContS0norm

Let us try to give a short explanation for the various characterizations of $\boldsymbol{S}_0(\mathbb{R}^d)$ (introduced as Wiener amalgam space $(\boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1, \boldsymbol{\ell}^1)(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1,\boldsymbol{\ell}^1)})$, via BUPUs). With the help of the atomic characterization **Ref. WAliatomic** we have seen that different (regular, and in fact even non-regular) BUPUs (in $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$) define the same space with equivalent norms.
The advantage of the *atomic norm* provided by **Ref. atomicnorms** is another equivalent norm with the advantage of providing a norm which is not only *isometrically invariant* under modulations but also under translations,i.e. satisfying

$$\|M_s T_t f\|_{atom} = \|f\|_{atom}, \quad f \in \boldsymbol{S}_0,\, t, s \in \mathbb{R}^d. \tag{344}$$ TFinvarS0

We are going to discuss the more elegant and widely used alternative description of alternative norms using a *continuous control function* using any non-zero function[37] $\varphi \in \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) \cap \boldsymbol{C}_c(\mathbb{R}^d)$:

$$\kappa(f)(z) = \kappa(f, \phi, \mathcal{F}\boldsymbol{L}^1)(z) := \|f \cdot T_z\phi\|_{\mathcal{F}\boldsymbol{L}^1}. \tag{345}$$ CCFFli

We are going to prove the following lemma, already stated earlier without proof, having $\boldsymbol{A} = \mathcal{F}\boldsymbol{L}^1$ in mind:

CSnorm2 **Lemma 72.** *For each non-zero $\phi \in \boldsymbol{A} \cap \boldsymbol{C}_c(\mathbb{R}^d)$ the expression*[38]

$$\|f \,|\, \boldsymbol{W}(\boldsymbol{A}, \boldsymbol{L}^1)\|_{\phi,cont} := \int_{\mathbb{R}^d} \kappa(f, \phi, \boldsymbol{A})(z)dz. \tag{346}$$ CCSnormdef

defines an equivalent norm on $\boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)(\mathbb{R}^d)$, and $f \in \boldsymbol{A}$ belongs to $\boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)(\mathbb{R}^d)$ if and only if this expression is finite.

Proof. First we have to show that we can estimate the "continuous norm", i.e. the expression (CCSnormdef 346), for any $f \in \boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)$, up to some constant.
We have assumed that $K = \mathrm{supp}(\phi)$ is a compact set. Writing $Q = [0,1]^d$ we can split $\mathbb{R}^d$ into a union of cubes of the form $k + Q, k \in \mathbb{Z}^d$. Hence one has for $z \in k + Q$

$$\mathrm{supp}(T_z\phi) = (z + K) \subset k + (K + Q). \tag{347}$$ suppcond

If we take the partial sum of the BUPU $\sum_{k\in F} \psi_k$ such that

$$\Phi(x) = \sum_{k\in\mathbb{Z}^d} T_k\psi(x) = 1 \text{ over} K + Q,$$

and consequently one has

$$T_z\phi = (T_k\Phi) \cdot (T_z\phi), \quad k \in \mathbb{Z}^d, z \in k + Q.$$

[37] Here we go back to the setting already discussed earlier in Lemma **Ref. CCSnorm1**.
[38] The integral appearing here applies to a non-negative, continuous functions, and can be understood in the Riemannian sense!

This implies that we have for any fixed $k \in \mathbb{Z}^d$, $z \in k+Q$, using $\|T_z\phi\|_{\boldsymbol{A}} = \|\phi\|_{\boldsymbol{A}}$:

$$\int_{k+Q} \kappa(f,\phi,\boldsymbol{A})(z)dz = \int_{k+Q} \|f \cdot T_k\Phi \cdot T_z\phi\|_{\boldsymbol{A}}\, dz \leq C_Q \|\phi\|_{\boldsymbol{A}} \|f \cdot T_k\Phi\|_{\boldsymbol{A}} \,.$$

This implies by summation over $\mathbb{Z}^d$:

$$\int_{\mathbb{R}^d} \kappa(f,\phi,\boldsymbol{A})(x)dx = \sum_{k\in\mathbb{Z}^d} \int_{k+Q} \kappa(f,\phi,\boldsymbol{A})(z)dz \leq C_Q \|\phi\|_{\boldsymbol{A}} \sum_{k\in\mathbb{Z}^d} \|f \cdot T_k\Phi\|_{\boldsymbol{A}} \,.$$

But since Φ is a finite sum over elements of the original BUPU Ψ we have of course

$$C_Q \|\phi\|_{\boldsymbol{A}} \|f \cdot T_k\Phi\|_{\boldsymbol{A}} \leq \sum_{j\in F} \|f \cdot T_k(T_j\psi)\|_{\boldsymbol{A}} \,,$$

or together

$$\int_{\mathbb{R}^d} \kappa(f,\phi,\boldsymbol{A})(x)dx \leq C_Q \|\phi\|_{\boldsymbol{A}} \sum_{j\in F} \sum_{k\in\mathbb{Z}^d} \|f \cdot T_k(T_j\psi)\|_{\boldsymbol{A}} \leq (\#(F) C_Q \|\phi\|_{\boldsymbol{A}}) \sum_{l\in\mathbb{Z}^d} \|f \cdot T_l\psi)\|_{\boldsymbol{A}} \,,$$

or in other words, for some constant $C = C(\Psi,\phi)$ we have

$$\int_{\mathbb{R}^d} \kappa(f,\phi,\boldsymbol{A})(x)dx \leq C \|f\|_{\boldsymbol{W}(\boldsymbol{A},\boldsymbol{\ell}^1)(\mathbb{R}^d)} \,, \quad f \in \boldsymbol{S}_0.$$

For the converse we will use an argument which is a bit different from the original argument given in [23], where Wiener's inversion theorem has been applied.
For the converse let us assume that we have some $f \in \boldsymbol{A}$ such that the expression 346 is finite, for some non-zero $\phi \in \boldsymbol{A} \cap \boldsymbol{C}_c(\mathbb{R}^d)$. Let us first show that we may assume without loss of generality that $\phi(x) \geq 0$ for all $x \in \mathbb{R}^d$. In fact, we can replace ϕ by $|\phi|^2$, because we have by the algebra structure of $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$:

$$\|f \cdot (T_z|\phi^2|)\|_{\boldsymbol{A}} \leq \|\overline{T_z\phi}\|_{\boldsymbol{A}} \|f \cdot T_z\phi\|_{\boldsymbol{A}} = \|\overline{\phi}\|_{\boldsymbol{A}} \|f \cdot T_z\phi\|_{\boldsymbol{A}}, \quad z \in \mathbb{R}^d.$$

$$\Rightarrow \quad \int_{\mathbb{R}^d} \kappa(f,|\phi|^2,\boldsymbol{A})(z)dz \leq \|\overline{\phi}\|_{\boldsymbol{A}} \int_{\mathbb{R}^d} \kappa(f,\phi,\boldsymbol{A})(z)dz. \tag{348}$$

This we assume from now on that $\phi(y) \geq 0$ for $y \in \mathbb{R}^d$. Next we use the shift invariance of the $\boldsymbol{A}$-norm in order to find that

$$\kappa(f,T_{x_0}\phi,\boldsymbol{A}) = T_{x_0}\kappa(f,\phi,\boldsymbol{A}), \tag{349}$$

and consequently we have for any fixed $x_0 \in \mathbb{R}^d$:

$$\int_{\mathbb{R}^d} \kappa(f,\phi,\boldsymbol{A})(z)dz = \int_{\mathbb{R}^d} \kappa(f,T_{x_0}\phi,\boldsymbol{A})(z)dz. \tag{350}$$

Given now any non-negative function $p \in \boldsymbol{C}_c(\mathbb{R}^d)$ we know already that $\mathrm{D}_\Psi p$ (more correctly: $\mathrm{D}_\Psi(\mu_p) \in \boldsymbol{M}_b(\mathbb{R}^d)$) is a finite discrete measures of the form $\sum_{i\in F} c_i\delta_{x_i}$ with

$$\sum_{i\in F} |c_i| \leq \|\mu_p\|_{\boldsymbol{M}_b} = \|p\|_{\boldsymbol{L}^1}. \tag{351}$$

This we can estimate the continuous norm using $\varphi = \mathrm{D}_\Psi p$, using (kappaint1 350) as

$$\int_{\mathbb{R}^d} \kappa(f, \varphi, \boldsymbol{A})(z)dz \leq \sum_{i\in F} |c_i| \int_{\mathbb{R}^d} \kappa(f, \phi, \boldsymbol{A})(z)dz \leq \|p\|_{\boldsymbol{L}^1} \int_{\mathbb{R}^d} \kappa(f, \phi, \boldsymbol{A})(z)dz. \tag{352}$$ kappaest2

We know that for $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}}) = \left(\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_{\mathcal{F}\boldsymbol{L}^1}\right)$ one has convergence of $(\mathrm{D}_\Psi p) \star \phi$ to $p \star \phi \in \boldsymbol{L}^1(\mathbb{R}^d) \star \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ (or, due to joint compact supports, in $\boldsymbol{L}^1(\mathbb{R}^d) \star \boldsymbol{S}_0(\mathbb{R}^d)$) and thus we obtain, by taking $|\Psi| \to 0$, for the limit[39]

$$\int_{\mathbb{R}^d} \kappa(f, p \star \phi, \boldsymbol{A})(z)dz \leq \|p\|_{\boldsymbol{L}^1} \int_{\mathbb{R}^d} \kappa(f, \phi, \boldsymbol{A})(z)dz. \tag{353}$$ kappaest3

We can use this in order to replace the original assumption about the finiteness of the integral over the CCF with respect to the given $\phi \in \boldsymbol{A} \cap \boldsymbol{C}_c(\mathbb{R}^d)$ by the finiteness of the corresponding expressions with respect to $\varphi := p \star |\phi|^2$, for any $p \in \boldsymbol{C}_c(\mathbb{R}^d)$. Thus we can choose p in such a way that $p(z) \equiv 1$ on any given compact set Q_1, say on $B_{R_1}(0)$. Let us further assume that $\mathrm{supp}(\phi) = \mathrm{supp}(|\phi|^2) \subseteq B_{R_2}(0)$ and (by rescaling) that $\|\phi\|_{\boldsymbol{L}^2} = 1 = \||\phi|^2\|_{\boldsymbol{L}^1}$.
By the *assumption* $R = R_1 - R_2 > 0$ we obtain for every $z \in B_R(0)$:

$$p \star |\phi^2|(z) = \int_{\mathbb{R}^d} |\phi|^2(y)p(z-y)dy = \int_{B_{R_2}(0)} |\phi|^2(y)p(z-y)dy = \int_{B_{R_2}(0)} |\phi|^2(y)dy = 1, \tag{354}$$

because $|z| \leq R$ and $|y| \leq R_1$ implies $|z - y| \leq R + R_2 \leq R_1$, hence $p(z - y) = 1$ for any such $y \in B_{R_2}(0)$.
With this possibility of using *plateau-like test functions* φ in the continuous norm it is now easy to get the opposite estimate, by trying to estimate the term $\|f\psi_k\|_{\boldsymbol{A}}$ by the corresponding part of the integral norm, i.e. by the integral over $k + Q$. We only need to ensure that $T_z\varphi \cdot \psi_k = \psi_k = T_k\psi$ for $z \in k + Q$ which boils done to the assumption: $T_{z-k}\varphi(x) = 1$ on $\mathrm{supp}(\psi)$, or $varphi(x - y) = 1$ for $x \in \mathrm{supp}(\psi)$ and $y \in Q$, or $\varphi(u) = 1$ on $\mathrm{supp}(\psi) - Q \subset B_R(0)$ for $R > 0$ chosen properly. Using again the algebra properties of $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ we get

$$\|f\psi_k\|_{\boldsymbol{A}} = \|(f\psi_k) \cdot (T_z\varphi)\|_{\boldsymbol{A}} = \|\psi_k \cdot (fT_z\varphi)\|_{\boldsymbol{A}} \leq \|\psi_k\|_{\boldsymbol{A}} \|f \cdot T_z\varphi\|_{\boldsymbol{A}}, \quad z \in k + Q.$$

and thus by integrating over $k + Q$, noting that $\mathrm{Vol}(Q) = 1$ and $\|\psi_k\|_{\boldsymbol{A}} = \|\psi\|_{\boldsymbol{A}}$, by adding over k in $\mathbb{Z}^d$ we get altogether:

$$\sum_{k\in\mathbb{Z}^d} \|f\psi_k\|_{\boldsymbol{A}} \leq \|\psi\|_{\boldsymbol{A}} \sum_{k\in\mathbb{Z}^d} \int_{k+Q} \|f \cdot T_z\varphi\|_{\boldsymbol{A}} dz = \|\psi\|_{\boldsymbol{A}} \int_{\mathbb{R}^d} \|f \cdot T_z\varphi\|_{\boldsymbol{A}} dz, \tag{355}$$ discrcontest0

thus completing the proof of equivalence of norms. □

Remark 38. The final result, but even better the proof of the result allows to understand why different test functions $\phi \in \boldsymbol{A} \cap \boldsymbol{C}_c(\mathbb{R}^d)$ define the same space and equivalent norms.

The proof even implies the following corollary:

[39] We could show, by reduction to elements with compact support, that we have for any sequence $\phi_n \to \phi_0$ in $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ we have convergence of the corresponding integral version of the norm, for $n \to \infty$.

WAlicont2 **Corollary 10.** *Any non-zero function* $g \in \boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)$ *defines an equivalent norm on* $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ *via the continuous control function* $\kappa(f,g)(z) = \|f \cdot T_z g\|_{\mathcal{F}\boldsymbol{L}^1}$. *The most interesting choice is given by* $g = g_0$, *the standard (Fourier invariant) Gaussian.*

This characterization of $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ opens the connection to *time-frequency analysis*, as it allows is to show that one can define $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ also via the **STFT** (the Short-Time Fourier transform) with respect to a *window* g. In fact, for $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}}) = \left(\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_{\mathcal{F}\boldsymbol{L}^1}\right)$ we have

$$\int_{\mathbb{R}^d} \kappa(f,g,\mathcal{F}\boldsymbol{L}^1)(t)dt = \int_{\mathbb{R}^d} \|\mathcal{F}(f \cdot T_t g)\|_{\boldsymbol{L}^1} dt = \int_{\mathbb{R}^d \times \widehat{\mathbb{R}}^d} |V_g f(t,s)| ds dt = \int_{\mathbb{R}^{2d}} |V_g f(z)| d(z), \tag{356}$$ S0defSTFT1

for the STFT of the *signal* f with respect to the *window* g being defined as

$$V_g(f)(t,s) = [\mathcal{F}(f \cdot T_t g)](s) = \int_{\mathbb{R}^d} f(x) g(x-t) e^{-2\pi i t s} dx. \tag{357}$$ Vgfdef00

Of course a short hand description of the norm obtained in such a way is

$$\|f\|_{\boldsymbol{S}_0} := \int_{\mathbb{R}^{2d}} |V_g f(z)| d(z) = \|V_g f\|_{\boldsymbol{L}^1}. \tag{358}$$ S0defSTFT2

etrans01.jpg

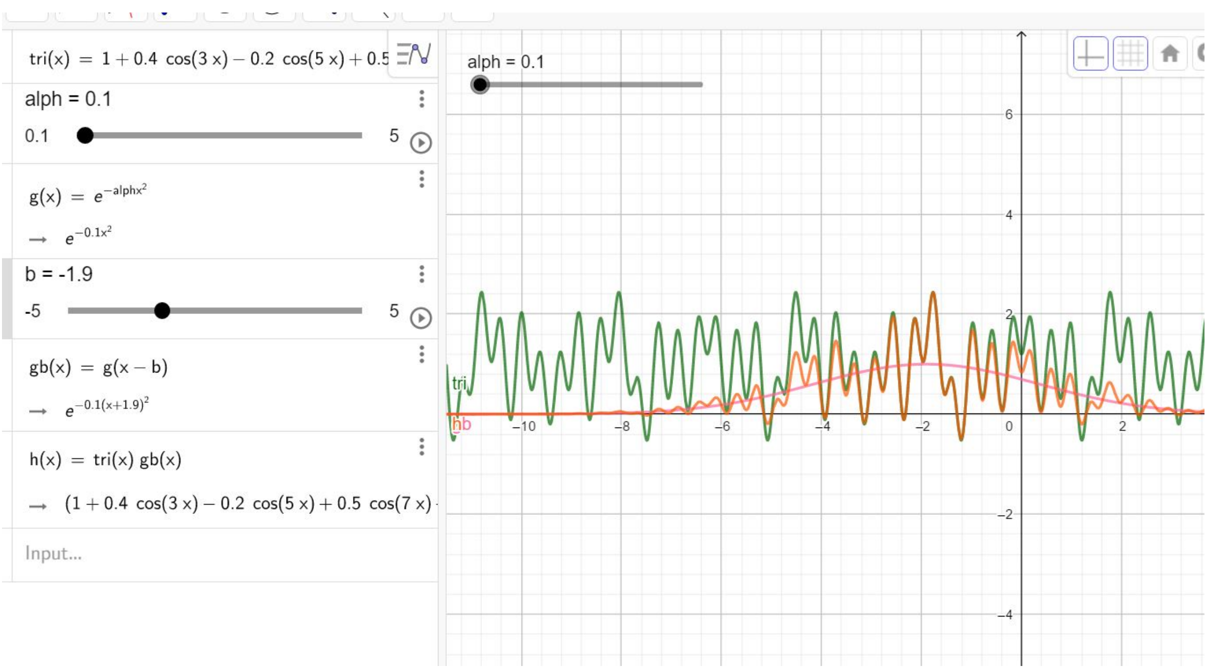

Abbildung 70: Voicetrans01.jpg

Usually the idea is to define the Fourier transform of the *localized version* of f by choosing g centered around zero, maybe even and with non-negative values, but even if the FT should be defined via integrals one can start with $f, g \in \boldsymbol{L}^2(\mathbb{R}^d)$, because then by the Cauchy-Schwarz inequality one has

$$f \cdot T_z g \in \boldsymbol{L}^2(\mathbb{R}^d) \cdot \boldsymbol{L}^2(\mathbb{R}^d) \subset \boldsymbol{L}^1(\mathbb{R}^d).$$

The last integral can also be written as a scalar product (in the sense of $\boldsymbol{L}^2(\mathbb{R}^d)$) between f and a time-frequency shifted version of g:

$$V_g f(t,s) = \langle f, M_s T_t g\rangle_{\boldsymbol{L}^2(\mathbb{R}^d)}, \quad f,g \in \boldsymbol{L}^2(\mathbb{R}^d). \tag{359}$$ VgfTFshift

With this tool it is possible to define $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ as a subspace of $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$ (defined via Lebesgue integration). This is the (elegant) path taken in the book of K. Göchenig (**gr01: [59]**), see Chap.12 (in particular p.246, for detailed comments). In this book $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ is introduced as modulation space $(\boldsymbol{M}^1(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}^1})$.

45.1 $\boldsymbol{S}_0$ introduced as subspace of L^2

For us the definition turns into a characterization (we provide a slightly more general statement than the claims made above): For most mathematicians the assumption $g \in \boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ (the Schwartz space or *rapidly decreasing functions*, which – as we will see later on – is obviously a subspace $\boldsymbol{S}_0(\mathbb{R}^d)$, even continuously embedded) will be most natural. It allows in particular to choose the Gauss function $g_0(t) = e^{\pi|t|^2}$.

SOLtsubsp **Theorem 17.** *For any non-zero $g \in \boldsymbol{\mathcal{S}}(\mathbb{R}^d)$, or $g \in \boldsymbol{L}^1(\mathbb{R}^d)$ and bandlimited resp. $g \in \mathcal{F}\boldsymbol{L}^1 \cap \boldsymbol{C}_c(\mathbb{R}^d)$ one has: $\boldsymbol{S}_0(\mathbb{R}^d)$ coincides with the following subspace of $\boldsymbol{L}^2(\mathbb{R}^d)$ (or in fact $\boldsymbol{L}^2(\mathbb{R}^d)$):*

$$\boldsymbol{S}_{Vgf}(\mathbb{R}^d) := \{f \in \boldsymbol{L}^2(\mathbb{R}^d) \mid V_g f \in \boldsymbol{L}^1(\mathbb{R}^d \times \widehat{\mathbb{R}}^d)\} \tag{360}$$ SOLtsub00

i.e. $\boldsymbol{S}_0(\mathbb{R}^d) = \boldsymbol{S}_{Vgf}$. Moreover, the quantity $\|V_g(f)\|_{\boldsymbol{L}^1(\mathbb{R}^{2d})}$ defines an equivalent norm on $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$. For the choice of a Gaussian window $g = g_0$ this implies immediately the isometric Fourier invariance of $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ (with the continuous, Gaussian norm) due to the fact ([59], p.53):

$$|V_{g_0} f(t,s)| = |V_{g_0} \widehat{f}(s,-t)|, \quad t,s \in \mathbb{R}^d. \tag{361}$$ STFgauFT

Proof. The assumptions on the window g imply that $g \in \boldsymbol{S}_0(\mathbb{R}^d)$, and hence the continuous norm using the window g provides an alternative description of the space $\boldsymbol{S}_0(\mathbb{R}^d)$. The formula (STFgauFT 361) can be derived from Plancherel's Theorem and the properties of the Fourier transform, combined with the Fourier invariance, and the commutation law for TF-shift operators

$$T_t M_s = e^{-2\pi i s t} M_s T_t, \quad (t,s) \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d, \tag{362}$$ TFcommut00

since

$$[T_t(M_s f)](x) = \chi_s(x-t) f(x-t) = \chi_s(-t)[\chi_s(x) f(t-x)] = e^{-2\pi i s t}[M_s(T_t f)](x).$$

The chain of identities leading to (STFgauFT 361) then goes as follows (CORR 17.11.)

$$V_g f(t,s) = \langle f, M_s T_t g\rangle =_{Planch} \langle \widehat{f}, T_s M_{-t} \widehat{g}\rangle = e^{-2\pi i s t}\langle \widehat{f}, M_{-t} T_s \widehat{g}\rangle = e^{-2\pi i s t} V_{\widehat{g}} \widehat{f}(s,-t).$$

Thus, for the Gauss function the Fourier transform looks like a rotation by 90 degrees in the clockwise sense. ☐

The STFT has many good properties. First of all it is *covariant* with respect to time-frequency shifts, which we will denote by $\pi(z) = M_s T_t$, for $z = (t,s) = (z_1, z_2)$, resp. $(t,s) \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d$.

groch313 **Lemma 73.**

$$V_g(T_u M_\eta f)(x,\omega) = e^{-2\pi i u\omega} V_g(x-u, \omega-\eta). \qquad (363)$$ groch3.14

In particular, the spectrogram of $T_u M_\eta f$ is a shifted version of the spectrogram of f:

$$|V_g(T_u M_\eta f)(x,\omega)|^2 = |V_g(x-u,\omega-\eta)|^2 = [T_{(u,\eta)}|V_g|^2](x,\omega). \qquad (364)$$ groch3.14b

Another consequence of Plancherel's Theorem is the fact that $f \to V_g f$ is isometric from $\left(\boldsymbol{L}^2(\mathbb{R}^d),\, \|\cdot\|_2\right)$ to $\left(\boldsymbol{L}^2(\mathbb{R}^d \times \widehat{\mathbb{R}}^d), \|\cdot\|_{\boldsymbol{L}^2(\mathbb{R}^d \times \widehat{\mathbb{R}}^d)}\right)$.

VgisomLt **Theorem 18.** *For two windows g_1, g_2 in $\boldsymbol{L}^2(\mathbb{R}^d)$ one has*

$$\left\langle V_{g_1} f_1, V_{g_2} f_2 \right\rangle_{\boldsymbol{L}^2(\mathbb{R}^{2d})} = \langle f_1, f_2 \rangle_{\boldsymbol{L}^2(\mathbb{R}^d)} \overline{\langle g_1, g_2 \rangle}_{\boldsymbol{L}^2(\mathbb{R}^d)}, \quad f_1, f_2 \in \boldsymbol{L}^2(\mathbb{R}^d). \qquad (365)$$ orthogSTFT321

An important result will be Lemma STFTSOSO 76 below, showing that one has for $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ with $\|g\|_{\boldsymbol{L}^2} = 1$ the mapping $f \mapsto V_g f$ is not only isometric from $\left(\boldsymbol{L}^2(\mathbb{R}^d),\, \|\cdot\|_2\right)$ to $\left(\boldsymbol{L}^2(\mathbb{R}^{2d}),\, \|\cdot\|_2\right)$, but it is also a bounded linear mapping (by restriction to $\boldsymbol{S}_0(\mathbb{R}^d)$) from $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ to $\left(\boldsymbol{S}_0(\mathbb{R}^{2d}), \|\cdot\|_{\boldsymbol{S}_0}\right)$.
Since we know that $\boldsymbol{S}_0 \hookrightarrow \boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)$ this also implies
that for any lattice $\Lambda \lhd \mathbb{R}^d \times \widehat{\mathbb{R}}^d$ the mapping

$$f \mapsto (V_g f(\lambda))_{\lambda \in \Lambda} = V_g f|_\Lambda$$

is bounded from $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ into $\left(\boldsymbol{\ell}^1(\Lambda),\, \|\cdot\|_1\right)$, i.e. there exists $C_1 > 0$ such that

$$\sum_{\lambda \in \Lambda} |V_g f(\lambda)| \leq C_1 \|f\|_{\boldsymbol{S}_0}, \quad f \in \boldsymbol{S}_0. \qquad (366)$$ S0toliLam

This follows from a combination of Lemma XXX and Lemma DcombSOP 42.

46 The reconstruction formula

A variant of the Shannon Sampling Theorem combined with the BUPU decomposition of elements of $\boldsymbol{S}_0(\mathbb{R}^d)$, in fact also of $\boldsymbol{L}^2(\mathbb{R}^d)$ and $\boldsymbol{S}_0'(\mathbb{R}^d)$ allow to locally decompose a signal in any of these space into a Fourier series (which gives a periodic function) which then just has to be localized.
So we note: Given one of the building blocks $f\psi_k$ for $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ we can replace it by a (coarse) periodic version, which then has an *absolutely convergent Fourier series expansion*, and then the "local piece" can be recovered from the periodic signal by multiplying back with some trapezoidal function (it just has to belong to $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) \cap \boldsymbol{C}_c(\mathbb{R}^d)$).
The abstract principle is the following

invisomemb **Lemma 74.** *Let T be an isometric embedding from a Hilbert space $(\mathcal{H}_1, \|\cdot\|_{\mathcal{H}_1})$ into another Hilbert space $(\mathcal{H}_2, \|\cdot\|_{\mathcal{H}_2})$. Then the inverse of T on the range of T, i.e. on the closed subspace $T(\mathcal{H}_1) \subset \mathcal{H}_2$ is the conjugate operator T^*.*

Proof. For an operator T between two Hilbert space the conjugate operator (corresponding to the transpose conjugate of a matrix $\boldsymbol{A}$ in the MATLAB world!) is characterized by the formula

$$\langle \mathbf{y}, T(\mathbf{x}) \rangle_{\mathcal{H}_2} = \langle T^*(\mathbf{y}), \mathbf{x} \rangle_{\mathcal{H}_1}, \quad \mathbf{x}, \in \mathcal{H}_1, \mathbf{y} \in \mathcal{H}_2. \tag{367}$$ conjopdef1

If T is isometric we have $\|T(\mathbf{x})\|_{\mathcal{H}_2} = \|\mathbf{x}\|_{\mathcal{H}_1}, \mathbf{x} \in \mathcal{H}_1$, or

$$\langle \mathbf{x}, \mathbf{x} \rangle_{\mathcal{H}_1} = \langle T(\mathbf{x}), T(\mathbf{x}) \rangle_{\mathcal{H}_2} = \langle T^*T(\mathbf{x}), \mathbf{x} \rangle_{\mathcal{H}_1}, \quad \mathbf{x} \in \mathcal{H}_1.$$

By the polarization principle[40] this implies

$$\langle \mathbf{x}, \mathbf{z} \rangle_{\mathcal{H}_1} = \langle T^*T(\mathbf{x}), \mathbf{z} \rangle_{\mathcal{H}_1}, \quad \mathbf{x}, \mathbf{z} \in \mathcal{H}_1,$$

which in turn implies

$$T^*T(\mathbf{x}) = \mathbf{x}, \quad \mathbf{x} \in \mathcal{H}_1, \quad \text{or } T^*T = Id_{\mathcal{H}_1}.$$

□

Remark 39. A good exercise to understand this principle is to think of $\mathcal{H}_1 = \mathbb{R}^2$ and $\mathcal{H}_2 = \mathbb{R}^3$, with the trivial embedding $T([x,y]) = [x,y,0]$ ($x-y$-plane viewed as a subset of $\mathbb{R}^3$). Form the matrix representing T (it has to be a simple 3×2 matrix) and then multiply it with the transpose = transpose conjugate from the left and check that you get the unit matrix of size 2×2! What if you change the order?

[40] This shows, based on the so-called parallelogram law, expressing scalar products in any Hilbert space by norms!

Let us now verify the isometric embedding $f \mapsto V_k f$ for the case that $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ and $\|k\|_{\boldsymbol{L}^2} = 1$ (for more general $f \in \boldsymbol{L}^2(\mathbb{R}^d)$ one can obtain the formula by approximation.

Proof. The first step is to check that one has the following identity, under the assumption (for convenience) that $k(x) \geq 0$, $k = k^\checkmark$ (k is even), and the (important) normalization $\|k\|_{\boldsymbol{L}^2} = 1 = \int_{\mathbb{R}^d} k^2(x)dx$:

$$LS := \int_{\mathbb{R}^d} \|T_x k \cdot f\|_2^2 \, dx = \|f\|_2^2, \quad f, k \in \boldsymbol{C}_c(\mathbb{R}^d). \tag{368}$$ contLtnorm1

In fact, the integrand of the left hand side (LS) equals

$$\int_{\mathbb{R}^d} k^2(y-x)|f|^2(y)dy = [k^2 \star |f|^2](x),$$

and therefore we have, using $\mathbf{1} \star |k|^2 = \mathbf{1}$ (by the condition $\int_{\mathbb{R}^d} k^2(z)dz = 1$):

$$LS = \int_{\mathbb{R}^d} \mathbf{1}(x)(k^2 \star |f|^2) \, dx = \sigma_{\mathbf{1}}(k^2 \star |f|^2) = (\sigma_{\mathbf{1}} \star k^2)(|f|^2) = \sigma_{\mathbf{1}}(|f|^2) = \|f\|_2^2.$$

In the next step we apply Plancherel's Theorem, which gives for $\varphi := \mathcal{F}(k)$:

$$\|T_x k \cdot f\|_{\boldsymbol{L}^2} = \|M_{-x}\varphi \star \widehat{f}\|_{\boldsymbol{L}^2},$$

and consequently we have (for details see **gr01:** [59]):

$$\|f\|_{\boldsymbol{L}^2}^2 = \int_{\mathbb{R}^d} \|M_{-x}\varphi \star \widehat{f}\|_{\boldsymbol{L}^2}{}^2.$$

$\square$

Remark 40. The mapping $T : f \mapsto V_g f$ is, under the conditions formulated above, an isometric embedding. Therefore it makes sense to identify the *conjugate mapping* $T^* : \boldsymbol{L}^2(\mathbb{R}^d \times \widehat{\mathbb{R}}^d) \to \boldsymbol{L}^2(\mathbb{R}^d)$ (hence we write V_g^*), characterized by the

$$\langle T^* x, y \rangle = \langle x, Ty \rangle. \tag{369}$$ conjopchar

The claim in **gr01:** [59], p.41, is to define it via vector-valued integration:

$$V_g^*(F) = \int_{\mathbb{R}^d \times \widehat{\mathbb{R}}^d} F(z)\pi(z)g dz, \quad \forall F \in \boldsymbol{L}^2(\mathbb{R}^d \times \widehat{\mathbb{R}}^d).$$

Of course we take it as a formal integral for $F \in \boldsymbol{C}_c(\mathbb{R}^d)$ and once it is identified with the adjoint operator to T above we extend it by taking limits, but it is also (relatively) easy to show that, upon replacement of $F \in \boldsymbol{C}_c(\mathbb{R}^d) \subset \boldsymbol{M}_b(\mathbb{R}^d \times \widehat{\mathbb{R}}^d)$ which acts as (again we take the integral formal)

$$\int_{\mathbb{R}^d \times \widehat{\mathbb{R}}^d} \pi(z) g \mathrm{D}_\Psi(F) dz = \sum_{i \in I} F(\psi_i)\pi(z_i)g$$

which is a Cauchy net in $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ and hence in $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$. See formula reconGab01 439 in Theorem goodGabor 24 below!

46.1 Factorization of the STFT

In **gr01:** [59],p.40 the STFT is described/factorized in the following way:

lemma312 **Lemma 75.** *If* $f, g \in \boldsymbol{L}^2(\mathbb{R}^d)$ *then*

$$V_g f = \mathcal{F}_2 \mathcal{T}_a (f \otimes g). \tag{370}$$ STFTfact1

Here we use the *asymmetric coordinate transformation* $\mathcal{T}_a$ on $\mathbb{R}^{2d}$, given by

$$[\mathcal{T}_a F](u, v) = F(v, v-u).$$

and $\mathcal{F}_2$ is the partial Fourier transform with respect to the second variable.
NOTE: It is among the special properties of $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ that we have $\boldsymbol{S}_0(\mathbb{R}^{2d}) = \boldsymbol{S}_0(\mathbb{R}^d) \widehat{\otimes} \boldsymbol{S}_0(\mathbb{R}^d)$, with equivalence of the corresponding norms.
This means, that for any $F \in \boldsymbol{S}_0(\mathbb{R}^{2d})$ one finds an absolutely convergent representation as a series of elementary tensors $[f \otimes g](t, s) = f(t) g(s)$:

$$F = \sum_{k=1}^{\infty} f_k \otimes g_k, \quad \sum_{k=1}^{\infty} \|f_k\|_{\boldsymbol{S}_0} \|g_k\|_{\boldsymbol{S}_0} < \infty. \tag{371}$$ tensprojrep

Of course there are then many such representations, and hence one takes the following expression as a norm (and shows that it is equivalent to any of the usual norms for $\left(\boldsymbol{S}_0(\mathbb{R}^{2d}), \|\cdot\|_{\boldsymbol{S}_0}\right)$):

$$\|F\|_{\boldsymbol{S}_0 \widehat{\otimes} \boldsymbol{S}_0} = \inf \left\{ \sum_{k=1}^{\infty} \|f_k\|_{\boldsymbol{S}_0} \|g_k\|_{\boldsymbol{S}_0} \right\} \tag{372}$$ tensprojnorm

where the infimum is taken over all the (so-called *admissible*) representations (371) of F.
Given this fact we obtain:

STFTSOSO **Lemma 76.** *The mapping* $(f, g) \mapsto V_g f$ *maps* $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right) \times \left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ *into* $\left(\boldsymbol{S}_0(\mathbb{R}^{2d}), \|\cdot\|_{\boldsymbol{S}_0}\right)$ *in a bounded way: There exists* $C > 0$ *such that*

$$\|V_g f\|_{\boldsymbol{S}_0(\mathbb{R}^{2d})} \leq C \|f\|_{\boldsymbol{S}_0} \|g\|_{\boldsymbol{S}_0}, \quad f, g \in \boldsymbol{S}_0(\mathbb{R}^d). \tag{373}$$ STFSOSOest

Proof. Starting from the factorization result (370) and the boundedness of $\mathcal{T}_a$ on $\boldsymbol{S}_0(\mathbb{R}^{2d})$, which arises from the invariance of $\left(\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_{\mathcal{F}\boldsymbol{L}^1}\right)$ and the atomic characterization of $\left(\boldsymbol{S}_0(\mathbb{R}^{2d}), \|\cdot\|_{\boldsymbol{S}_0}\right)$, combined with the boundedness of the partial Fourier transform

$$\mathcal{F}_2(\boldsymbol{S}_0 \otimes \boldsymbol{S}_0) = \boldsymbol{S}_0 \otimes (\mathcal{F}\boldsymbol{S}_0) = \boldsymbol{S}_0 \otimes \boldsymbol{S}_0,$$

it is clear that we can get the estimate (373). $\square$

Vgstar1 **Lemma 77.** *The adjoint mapping* $V_g^* : (\boldsymbol{S}_0'(\mathbb{R}^{2d}), \|\cdot\|_{\boldsymbol{S}_0'(\mathbb{R}^{2d})}) \to (\boldsymbol{S}_0'(\mathbb{R}), \|\cdot\|_{\boldsymbol{S}_0'})$ *is characterized by the property*

$$V_g^*(\delta_z) = \pi(z)g = M_s T_t g, \quad z = (t,s) \in \mathbb{R}^{2d}. \tag{374}$$ Vgstzzz

Proof. Since $V_g : f \to V_g f$ maps $\boldsymbol{S}_0(\mathbb{R}^d)$ boundedly into $\left(\boldsymbol{S}_0(\mathbb{R}^{2d}), \|\cdot\|_{\boldsymbol{S}_0}\right)$ and by the adjoint mapping $V_g^* : \left(\boldsymbol{S}_0'(\mathbb{R}^{2d}), \|\cdot\|_{\boldsymbol{S}_0'(\mathbb{R}^{2d})}\right) \to (\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$ can be applied to $\delta_z \in \boldsymbol{M}_b(\mathbb{R}^{2d}) \subset \boldsymbol{S}_0'(\mathbb{R}^{2d})$, and we have for $f \in \boldsymbol{S}_0(\mathbb{R}^d)$:

$$V_g^*(\delta_z)(f) = \delta_z(V_g f) = V_g f(z) = \langle f, \pi(z)g\rangle_{\boldsymbol{L}^2(\mathbb{R}^d)}. \tag{375}$$ Vgstdelz

This implies that $V_g^*(\delta_z) \in \boldsymbol{S}_0'(\mathbb{R}^d)$ coincides with $\mu_{\pi(z)g}$, or with the usual identification[41]

$$V_g^*(\delta_z) = \pi(z)g, \quad z \in \mathbb{R}^{2d}. \tag{376}$$ Vgstdelzsum

□

Since we start from $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ we get $V_g^*(\boldsymbol{M}_b(\mathbb{R}^{2d})) \subset \boldsymbol{S}_0(\mathbb{R}^d)$, and in fact, one can verify, using the continuity of $z \to \pi(z)(g) : \quad \mathbb{R}^{2d} \to \boldsymbol{S}_0(\mathbb{R}^d)$ one even can verify that for any measure $\mu \in \boldsymbol{M}_b(\mathbb{R}^{2d})$ one has $V_g^*(\mu) = \lim_{|\Psi|\to 0} V_g^*(D_\Psi \mu)$ in $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$, hence the mapping

$$\mathrm{D}_\Psi(F) = \sum_{i\in I} F(\psi_i)\pi(z_i)g$$

is a Cauchy-net in the Banach space $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$. This statement is stronger than the usual, which talks about weak convergence in $\boldsymbol{L}^2(\mathbb{R}^d)$ only. The norm convergence in $\boldsymbol{L}^2(\mathbb{R}^d)$ (due to the window $g \in \boldsymbol{S}_0(\mathbb{R}^d)$) is also valid and has been shown by classical tools by F. Weisz.
The isometric property of $V_g : \left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right) \to \left(\boldsymbol{L}^2(\mathbb{R}^{2d}), \|\cdot\|_2\right)$ implies that V_g^* maps $\left(\boldsymbol{L}^2(\mathbb{R}^{2d}), \|\cdot\|_2\right)$ boundedly back to $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$. Note that the kernel of V_g^* (meaning the nullspace: $\mathcal{N}_{V_g^*} = \{F \in \boldsymbol{L}^2(\mathbb{R}^{2d}) \,|\, V_g^*(F) = 0\}$) is just the orthogonal complement of the range of V_g!
Finally we can extend the STFT to all of $\boldsymbol{S}_0'(\mathbb{R}^d)$, since it is easy to see that we have for any $\sigma_0 \in \boldsymbol{S}_0'(\mathbb{R}^d)$ with $\sigma_0 = \boldsymbol{S}_0' - w^* - \lim_{n\to\infty}$ for some sequence in $\boldsymbol{S}_0(\mathbb{R}^d)$, that one can take the limit (pointwise) and obtain (still with the assumption $g \in \boldsymbol{S}_0(\mathbb{R}^d)$)[42]

$$V_g\sigma(z) = \sigma(\overline{\pi(z)g}), \quad z \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d. \tag{377}$$ STFTforSOP

which is obviously a bounded and continuous function:

$$|V_g\sigma(z)| \leq \|\sigma\|_{\boldsymbol{S}_0'} \|\pi(z)(g)\|_{\boldsymbol{S}_0} = \|\sigma\|_{\boldsymbol{S}_0'} \|g\|_{\boldsymbol{S}_0}, z \in \mathbb{R}^{2d}. \tag{378}$$ VgsigCb

The continuity follows once more from the continuity of $z \mapsto \pi(z)g$ from $\mathbb{R}^{2d}$ into $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$. NOTE ALSO THAT THE MAPPING V_g^* IS w^*-w^*-CONTINUOUS !

[41] The consistency of arguments, using both bilinear duality mappings and sesquilinear mappings, has be the checked separately, not a final answer yet.

[42] We have corrected from $\pi(z)g$ to $\overline{\pi(z)}$ in order to re-establish consistence with the $\boldsymbol{L}^2(\mathbb{R}^d)$-case, but $f \mapsto \bar{f}$ is isometric on $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$, thus there is no real change of any further argument. 04.01.2021

47 From Fourier Series to the Fourier Transform

47.1 The Classical Approach

The classical approach to Fourier Analysis often takes that historical path, i.e. it start with the theory of Fourier series expansions, explaining that the "correct setting" is that of the *Hilbert space* $\left(\boldsymbol{L}^2(\mathbb{T}),\, \|\cdot\|_2\right)$ (identified with $\boldsymbol{L}^2([0,1])$ in our case (by the choice of characters $\chi_k(x) = e^{2\pi i k x}, k \in \mathbb{Z}$), which form a *complete*[43] *orthogonal system*, which implies that every "function" in that space has a representation as infinite sum

$$f = \sum_{k\in\mathbb{Z}} \widehat{f}(k)\chi_k,$$

with (unconditional) convergence taking place in the sense of the quadratic mean. This means: Given $f \in \boldsymbol{L}^2(\mathbb{T})$ and $\varepsilon > 0$ there exists a finite set $F_0 \subset \mathbb{Z}$ such that for any other finite set $F \supset F_0$ one has

$$\|f - \sum_{k\in F} \widehat{f}(k)\chi_k\|_{\boldsymbol{L}^2} \leq \varepsilon,$$

meaning that a certain quality of approximation can be guaranteed once all the (finitely) many relevant coefficients are used (and potentially a few more). Also $\|f\|_{\boldsymbol{L}^2} = \|(\widehat{f}(k))_{k\in\mathbb{Z}}\|_2$, this is Parseval's theorem, so one could take for $\gamma > 0$ (and small)

$$F_0 := \{k \in \mathbb{Z} \,|\, |\widehat{f}(k)| \geq \gamma\}.$$

It was a sensation that Lennart Carleson was able to prove a long-standing conjecture of Lusin (from 1922) that the Fourier series is almost everywhere convergent (see [12], Acta Mathematica), i.e. that one has indeed

$$f(x) = \lim_{K\to\infty} \sum_{k=-K}^{K} \widehat{f}(k) e^{2\pi i k x} \text{ almost everywhere.} \tag{379}$$

This is a rather deep mathematical result, but I have not seen any strong impact on applications, where it is fine to have the pointwise convergence under mild extra conditions. So the strength of the theorem, that it is valid for *any* $f \in \boldsymbol{L}^2(\mathbb{T})$, does not help much, because the exceptional set can even any countable set. Well, the situation is much better than for the case of $f \in \boldsymbol{L}^1(\mathbb{T})$, where one may have pointwise divergence everywhere! (not just on a sets of positive measure). But *this essentially means* that - as far as applications is concerned - one should not bother too much about questions of pointwise convergence of Fourier series.
Of course there is a huge body of papers concerning summability methods and either norm convergence of almost everywhere convergence for particular summability methods (comparable to the ideas needed in the proof of the Fourier inversion theorem), but this leads as far away from the focus of this course.

[43] Warning: Here the word complete is used in a different way compared to the completeness for Banach spaces!

47.2 Letting the period go to infinity

Very often the Fourier transform form $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ is introduced as the limit of period versions of k, with the period tending to infinity.
We will describe the situation in the following way:

shahp **Proposition 18.** *The family* $\mathrm{St}_\rho ⧢ = ⧢_\alpha$*, for* $\alpha = 1/rho \geq 1$*, is uniformly bounded in* $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$*, and satisfies*

$$\boldsymbol{S}_0' - w^* - lim_{\alpha\to\infty} ⧢_\alpha = \delta_0. \tag{380}$$ wstshapinf

Consequently $⧢_\alpha \star k = \sum_{k\in\mathbb{Z}^d} T_k k \in \boldsymbol{C}_b(\mathbb{R}^d)$ *is a bounded family in* $\Big(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\Big)$*, hence also bounded in* $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$*. In fact, one obtains* $\widehat{f}$ *as a* w^*-*limit:*

$$\boldsymbol{S}_0' - w^* - lim_{\alpha\to\infty} \mathcal{F}(⧢_\alpha \star f) = \widehat{f}, \quad f \in \boldsymbol{S}_0. \tag{381}$$ FSperlimFT

Proof. Starting from the fact that $⧢ \in \boldsymbol{S}_0'(\mathbb{R}^d)$ and the observation that mass-preserving dilation operators are uniformly bounded on $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$ (see Lemma SOPdilatinv 45) we obtain the boundedness of the α-periodized versions of k by the simple observation that we have for the case that $\mathrm{supp}(k) \subseteq B_R(0)$ that for $\alpha > 2R$ the translates of k along $\alpha\mathbb{Z}^d$ are disjoint, thus $\|⧢_\alpha \star k\|_\infty = \|k\|_\infty$.
Since convolution by k preserves w^*-convergence, due to the adjointness relation

$$\sigma \star k(f) = \sigma(k^\checkmark \star f), f \in \boldsymbol{S}_0(\mathbb{R}^d), k \in LiRd,$$

and the (extended) Fourier transform is w^*-w^*-continuous , we conclude that

$$\boldsymbol{S}_0' - w^* lim_{\alpha\to\infty} ⧢_\alpha \star k = \boldsymbol{S}_0' - w^* lim_{\alpha\to\infty} \mathrm{D}_\alpha ⧢ \cdot \widehat{k} = \widehat{k}. \tag{382}$$ FourSerTransf

□

Sometimes a similar argument is used in order to derive the Fourier inversion formula, but a mathematical justification of the manipulations required to actually show that the intuition arising from such transitions ($\alpha \to \infty$) can be justified is certainly more cumbersome than directly verifying the validity that the forward and the inverse Fourier transform (originally just NAMES like this) are in fact inverse mapping to each other on proper domains, such as $(\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1)(\mathbb{R}^d)$ (and thus $\boldsymbol{S}_0(\mathbb{R}^d)$).

Remark 41. Compared to a simple computational verification of the statements above the formulas above show much more: Given alpha, the Fourier transform of a α periodic version of $f \in \boldsymbol{S}_0(\mathbb{R}^d) \cap \boldsymbol{C}_c(\mathbb{R}^d)$ is obviously (write again $\rho = 1/alpha \leq 1$):

$$(⧢_\alpha \star f) = \mathrm{D}_\alpha(⧢) \cdot \widehat{f} = \rho^d \sum_{k\in\mathbb{Z}^d} \widehat{f}(n)\delta_{\rho k} \in \boldsymbol{M}_d(\mathbb{R}^d),$$

is in fact a bounded and discrete measure (since $\sum_{k\in\mathbb{Z}^d} |\widehat{f}(k)| < \infty$ for $f \in \boldsymbol{S}_0(\mathbb{R}^d)$), supported on $\rho\mathbb{Z}^d$. In fact, these discrete measures, applied to $\varphi \in \boldsymbol{S}_0(\mathbb{R}^d)$ represent Riemannian sums to the integral

$$\int_{\mathbb{R}^d} \varphi(s)\widehat{f}(s)ds = \mu_{\widehat{f}}(\varphi).$$

48 Approximation by Sampling

The considerations above indicate that the periodized version of a function in $\boldsymbol{S}_0(\mathbb{R}^d)$ approximate the original (non-periodic) version of $f \in \boldsymbol{S}_0(\mathbb{R}^d)$. In fact, we have convergence in the w^*–sense on both the time and the frequency side. But now these two views can be seen as equivalent. In other words, we could start to store more and more information about $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ by sampling it on a grid which is more and more refined. By the above considerations we should not multiply f by $\sqcup\!\sqcup_\beta$, letting β go to zero, but rather multiply of by $\mathrm{D}_\rho \sqcup\!\sqcup = \alpha^d \sqcup\!\sqcup_\alpha$ (for $\rho = 1/\alpha$). Then we can expect (and actually have the valid claim) that

$$\boldsymbol{S}_0' - w^* - \lim_{\alpha\to\infty} \alpha^d \sqcup\!\sqcup_\alpha \cdot h = \boldsymbol{S}_0' - w^* - \lim_{\alpha\to\infty} \alpha^d \sum_{k\in\mathbb{Z}^d} h(\alpha k)\delta_{\alpha k} = h, \tag{383}$$

sampconvSOP

certainly for $h \in \boldsymbol{C}_b(\mathbb{R}^d)$.
In fact, we have for any $f \in \boldsymbol{S}_0(\mathbb{R}^d)$:

$$|\alpha^d \sqcup\!\sqcup_\alpha \cdot h(f)| \leq \|h\|_\infty \|\mathrm{D}_\rho \sqcup\!\sqcup\|_{\boldsymbol{S}_0'} \|f\|_{\boldsymbol{S}_0} \leq [\|h\|_\infty \cdot C \|\sqcup\!\sqcup\|_{\boldsymbol{S}_0'}] \|f\|_{\boldsymbol{S}_0},$$

and therefore we can verify w^*–convergence by watching the action of these discrete measures on $k \in \boldsymbol{C}_c(\mathbb{R}^d) \cap \boldsymbol{S}_0(\mathbb{R}^d)$, but these are just Riemannian sums for the integral $\int_{\mathbb{R}^d} h(x)k(x)dx$ (see above).

48.1 Recovery from regular samples

The information contained in the samples of $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ is getting more and more. We know, that for any $f \in \left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$ one has norm convergence of $\mathrm{Sp}_\Psi f$ to f with respect to the sup-norm, as $|\Psi| \to 0$ (think of piecewise linear interpolation over $\mathbb{R}$).
Since the sequence of compressed triangular functions $\mathrm{St}_\rho \Delta, \rho \to 0$ forms a Dirac family one may expect that

$$\mathrm{St}_\rho \Delta \star [\alpha^d \sum_{k\in\mathbb{Z}^d} f(\alpha k)\delta_{\alpha k}] \to \delta_0 \star (\mathbf{1} \cdot f) \to f, \quad f \in \boldsymbol{S}_0. \tag{384}$$

pwlininterpol

We have to make two observations: First of all this is in fact an alternative description of piecewise linear interpolation, since $\mathrm{D}_{1\alpha} \Delta$ is just a triangular function with basis $[-\alpha, \alpha]$.

$$\mathrm{St}_\rho \Delta \star [\alpha^d \sum_{k\in\mathbb{Z}^d} f(\alpha k)\delta_{\alpha k}] = \sum_{k\in\mathbb{Z}^d} f(\alpha k) T_{\alpha k} \mathrm{D}_\rho(\Delta). \tag{385}$$

pwlinconf02

A highly relevant result (which however requires some work)concerning the quality of the reconstruction. Instead of requiring a lot of smoothness and decay in order to achieve e.g. uniform of $\boldsymbol{L}^2$-convergence we can apply the result to any $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ and obtain as a result convergence in the sense of $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$. Among others this implies that one has convergence (at least for all $f \in \boldsymbol{S}_0(\mathbb{R}^d)$) in the uniform sense (easy), but also in the sense of $\boldsymbol{L}^1$ (in fact, in the sense of $\left(\boldsymbol{W}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{W}}\right)$ and hence also in the sense of $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$. The details are found in **feka07:** [43].
reconstS0
The following theorem is a generalization of Theorem 15:

ntbdl5bz.jpg

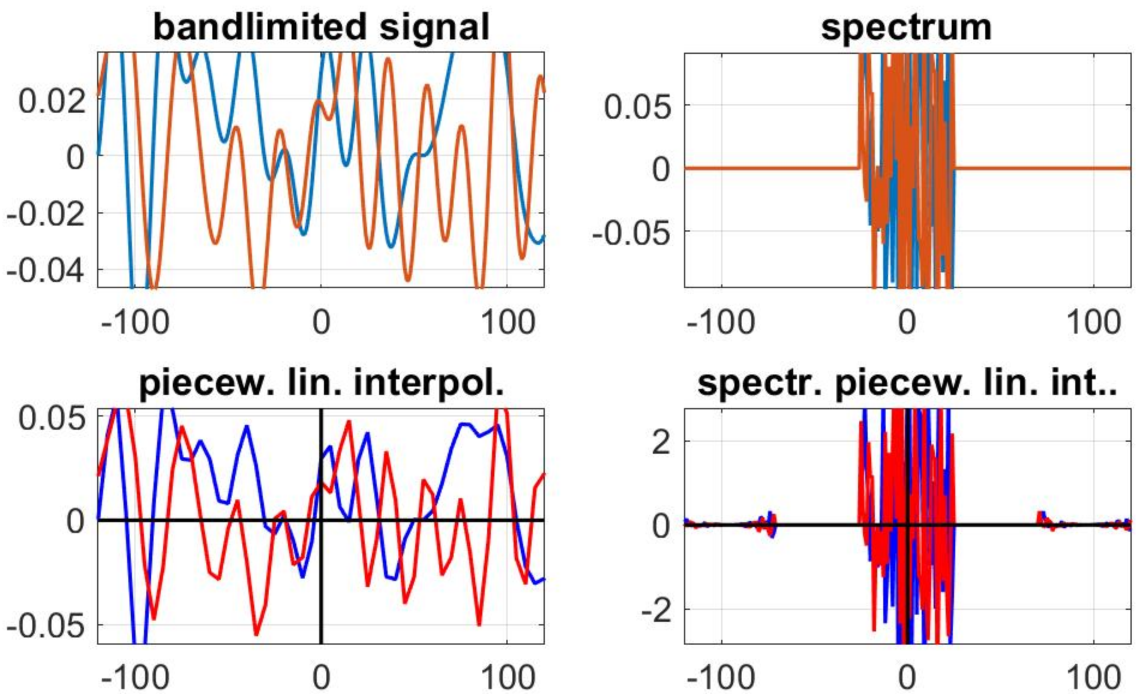

Abbildung 71: plnintbdl5bz.jpg:
piecewise linear interpolation of a smooth, i.e. here band-limited signal (partial zoom in, to show the lack of smoothness). The edges require some high frequency contributions.

quasintSORd **Theorem 19.** *Assume that* $\Psi = (T_k\psi)_{k\in\mathbb{Z}^d}$ *defines a BUPU in* $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$[44] *and write* $\mathrm{D}_\rho\Psi$ *for the family* $\mathrm{D}_\rho(T_k\psi) = (T_{\alpha k}\mathrm{D}_\rho\Delta)_{k\in\mathbb{Z}^d}$, *with* $\alpha = 1/\rho \to 0$. *Then* $|\mathrm{D}_\rho\Psi| \leq r\alpha \to 0$ *for* $\alpha \to 0$, *and*

$$\|f - \alpha^d \sum_{k\in\mathbb{Z}^d} f(\alpha k) T_{\alpha k} \operatorname{St}_\alpha \psi\|_{\boldsymbol{S}_0} \to 0, \quad \text{for } \alpha \to 0, \forall f \in \boldsymbol{S}_0. \tag{386}$$ SODeltaconv

This result was the cornerstone for the subsequent result published by Norbert Kaiblinger, concerning the approximate computation of the Fourier transform of a "nice function" $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ via an FFT routine, applied to a sequence of samples of the original function (more correctly: applied to a periodic and discrete sequence $\mathbf{x}$ of length N, obtained by appropriate sampling (at rate β) and periodizing (at rate $\alpha = N\beta$) the original function, and then applying the *quasi-interpolation operator* Sp_Ψ as in Thm. SODeltaconv 386 to `y = fft(x)` (associated to the corresponding lattice points in the frequency domain). So let us describe mostly the procedure, justified by the main result of ka05 67 (Thm.2, in particular).

[44]As it is required for the definition of $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$, for some $\psi \in \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ with $\operatorname{supp}(\psi) \subset B_r(0)$.

48.2 Fourier transforms via FFT

As we have pointed out earlier the FFT algorithm (resp. the DFT) allows to compute the extended Fourier transform to a discrete and periodic signal, by viewing it as an element of $\boldsymbol{S}_0'(\mathbb{R}^d)$, as a finite linear combination of Dirac combs (this is possible even in the multi-dimensional setting), because the coarse periodicity lattice Λ_p is a sublattice (hence of finite index) of the sampling lattice Λ_s.
For our discussion it is OK to think of the one-dimensional case, with $\Lambda_s = \beta\mathbb{Z}$ and $\Lambda_p = \alpha\mathbb{Z}$, for $\alpha = N\beta$. For $d \geq 2$ one can reduce the discussion to the case $\Lambda_s = \mathbb{Z}^d$ and $\Lambda_p = \boldsymbol{A} \star \mathbb{Z}^d$ for some integer-valued $d \times d$-matrix.
The procedure suggested by Theorem 2 of ka05 67 suggest the following:

1. First one should produce a coarsely periodized version of the given function. If the function has compact support one can choose Λ_p wide enough to guarantee that the shifted copies of $f \in \boldsymbol{C}_c(\mathbb{R}^d)$ do not overlap, meaning that there is no information lost by the periodization. (see Fig. [**?**]). For the general case one keeps control over the error obtained by overlapping tails of the well decaying functions. At least on can show that the overall error is getting smaller and smaller (in the $\big(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\big)$-sense) as the periodization lattice Λ_p is getting more and more coarse.

2. Then we go on the sample the periodized function sufficiently fine. Of course, we would have perfect information if the function was band-limited, but there is a small error (depending on the degree of smoothness of the sampled function, understood for now in a naive way).

3. Since we assume that $\Lambda_p \lhd \Lambda_s$ we can also revert the order of sampling and periodization, i.e. we will get the same result by first sampling and then periodization. In terms of formulas we have

$$\sqcup\!\sqcup_s \cdot (\sqcup\!\sqcup_p \star f) = \sqcup\!\sqcup_p \star (\sqcup\!\sqcup_s \cdot f), \quad f \in \boldsymbol{S}_0. \tag{387}$$

sampperiod00

4. An application of the quasi-interpolation result with the possibility of approximate reconstruction of $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ from the coarsely periodized version allows to recover f, up to some error, by a suitable (linear reconstruction principle, using BUPUs etc.), we just write $\mathcal{R}_{s,p}$, such that we have for a given $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ and $\varepsilon > 0$: if Λ_s is fine enough and Λ_p is coarse enough then we can guarantee that

$$\|\mathcal{R}_{s,p}[\sqcup\!\sqcup_s \cdot (\sqcup\!\sqcup_p \star f)] - f\|_{\boldsymbol{S}_0} \leq \varepsilon. \tag{388}$$

recover

5. Finally we obtain the sampled an periodized version of $\widehat{f}$ by the DFT (the finite version of the Fourier transform). Since we know that $\widehat{f} \in \boldsymbol{S}_0(\mathbb{R}^d)$ as well we can approximately recover $\widehat{f}$ from its sampled and periodized version (the periodicity lattice if $\Lambda_s^\perp$ and the sampling lattice is $\Lambda_p^\perp$, and of course these are coarse and fine respectively of their partner lattices are fine and coarse).

49 Discussion of various approaches

Let us give a short discussions of the many possible ways to introduce $\boldsymbol{S}_0(\mathbb{R}^d)$ and its dual $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$.
The ETH course (autumn 2020) starts "from nothing", i.e. assumes only the use of the Riemann integral, concepts of linear algebra, and the willingness to of the reader to learn fundamental principles of functional analysis, which are required in order to prove identities using norm-convergence in complete normed vector spaces, i.e. Banach spaces, and sometimes the alternative concept of strong operator convergence (for operators) respectively of (the so-called) w^*-convergence in dual Banach spaces.
The benefit is of course that it is not required to take first a course in measure theory in order to understand the Lebesgue spaces $\left(\boldsymbol{L}^p(\mathbb{R}^d),\ \|\cdot\|_p\right)$ (specifically for $p = 1, 2, \infty$) as vector spaces (of equivalence classes) of measurable functions which are Lebesgue integrable $(\boldsymbol{L}^1(\mathbb{R}^d))$, square integrable (the Hilbert space $\boldsymbol{L}^2(\mathbb{R}^d)$) or essentially bounded respectively $\boldsymbol{L}^\infty(\mathbb{R}^d)$). We also do not work with topological vector spaces which are required in order to derive mathematically correct the basic facts about the Schwartz space $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ of rapidly decreasing functions and its dual, the space of *tempered distributions*.
The goal of the course is to establish a way to understand the *extended Fourier transform* which is well defined on $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$, also called the space of *mild distributions*, which is the dual of the Segal algebra $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ (also called *Feichtinger's algebra*. It is *not based on Lebesgue integration theory*, but of course it still introduces $\left(\boldsymbol{L}^1(\mathbb{R}^d),\ \|\cdot\|_1\right)$ (as a subspace of the bounded measures, using the Riemann integral resp. the Haar measure for $\mathbb{R}^d$), and - in order to present Plancherel's Theorem also $\boldsymbol{L}^2(\mathbb{R}^d)$, but as a subspace of $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$ (which is defined first).
Technically speaking the setting is chosen in such a way that most of the issues can be verified in the context of general LCA groups $\mathscr{G}$, once the existence of the Haar measure is established (here with the Riemann integral). The convolution of measures is introduced by the identification of $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b}) = \left(\boldsymbol{C}_0'(\mathbb{R}^d),\ \|\cdot\|_{\boldsymbol{C}_0'}\right)$ with the space $\mathcal{H}_{\mathbb{R}^d}(\boldsymbol{C}_0(\mathbb{R}^d))$ of all translation-invariant operators on $\left(\boldsymbol{C}_0(\mathbb{R}^d),\ \|\cdot\|_\infty\right)$ (so called BIBO-systems), which form a commutative Banach algebra with respect to composition.
Of course many things would like unnatural if one would try to completely avoid specific tools available in the context of the Euclidean space $\mathscr{G} = \mathbb{R}^d$, namely the use of dilation (allowing the give a concrete description of Dirac sequences in $\left(\boldsymbol{L}^1(\mathbb{R}^d),\ \|\cdot\|_1\right)$), or dilation operators providing - in the spirit of summability kernels - approximate units for the Fourier algebra $\left(\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d),\ \|\cdot\|_{\mathcal{F}\boldsymbol{L}^1}\right)$ (with pointwise multiplication).
In addition, the Gauss-function plays an important role, mostly because of its Fourier invariance and the fact that it is optimally concentrated in a time-frequency sense, which makes it (according to D. Gabor, **ga46:** [53]) an ideal "window" for the STFT (short-time Fourier transform).

50 The Shannon Sampling Theorem

One can say that in some sense the Shannon Sampling theorem is a consequence of Plancherel's Theorem (telling us that the Fourier transform is a *unitary linear operator* on $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$) and the fact that the *pure frequencies* define an orthonormal system for the $\boldsymbol{L}^2$-space over the fundamental domain, e.g. the unit cube $[-1/2, 1/2]^d$ within $\mathbb{R}^d$, viewed as the subspace of $\boldsymbol{L}^2(\mathbb{R}^d)$ which is obtained by multiplying (on the FT side) all the $\boldsymbol{L}^2$-functions with the indicator function (box function) of the fundamental domain. In order to formulate it in a way that is compatible with the usual conventions let us call Λ to be a lattice, i.e. a discrete subgroup of $\mathbb{R}^d$, or more concretely $\Lambda = A \star (\mathbb{Z}^d)$ for some non-singular $d \times d$-matrix. Furthermore $\Lambda^\perp$ is by definition the so-called *orthogonal lattice*, which is the set of all pure frequencies χ_s such that $\chi_s(\lambda) \equiv 1$ for all $\lambda \in \Lambda$.

ShannononRd **Theorem 20.** *Assume that $\Omega \subset \mathbb{R}^d$ is a bounded subset with the property Ω is disjoint from all its (non-trivial) $\Lambda^\perp$-translates, i.e. $\lambda^\perp + \Omega \cap \Omega = \emptyset$ for all $\lambda^\perp \neq 0, \lambda^\perp \in \Lambda^\perp$. Then every Ω-band-limited function $f \in \boldsymbol{L}^2(\mathbb{R}^d)$[45], i.e. every f which can be written as an inverse [Plancherel] Fourier transform of the form*

$$f(t) = \int_\Omega h(s) e^{2\pi i s \cdot t} ds \tag{389}$$ bandlimrepr1

can be completely be recovered from the samples $(f(\lambda))_{\lambda \in \Lambda}$ by the following series expansion (called the Shannon Sampling expansion): For a suitable chosen constant $C_\Lambda > 0$, one has:

$$f(t) = C_\Lambda \sum_{\lambda \in \Lambda} f(\lambda) T_\lambda SINC_\Omega(t) = C_\Lambda \sum_{\lambda \in \Lambda} f(\lambda) SINC_\Omega(t - \lambda), \tag{390}$$ Shannenserrep

where $SINC_\Omega = IFFT(\mathbf{1}_\Omega)$. The convergence takes place in the $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$-sense, but also uniformly.

Remark 42. For the standard choice, i.e. $\Omega = [-1/2, 1/2]^d$ we simply have

$$SINC_\Omega(t_1, \cdots, t_d) = SINC(t_1) \cdots SINC(t_d),$$

using the classical SINC-function, $SINC = IFFT(\mathbf{1}_{[-1/2,1/2]})$.
Moreover we have $\Lambda^\perp = \mathbb{Z}^d$ for $\Lambda = \mathbb{Z}^d$, and $\Lambda^\perp = \frac{1}{a}\mathbb{Z}^d$ for $\Lambda = a \cdot \mathbb{Z}^d$.

[45] Or in a more practical description: if $h = \hat{f}$ is supported inside Ω, resp. is zero outside of Ω, except possibly for a set of measure zero.

Proof. We will give the proof (for convenience) for $\Omega = [-1/2, 1/2] \subset \mathbb{R}$ (occasionally we will write Q, thinking of $[-1/2, 1/2]^d \subset \mathbb{R}^d$). d
First let us observe that (via Plancherel's) Theorem we have

$$\mathrm{SINC} \star \mathrm{SINC} = \mathrm{SINC}, \tag{391}$$ SINCconv2

as a simple consequence of the pointwise relation

$$\mathbf{1}_{[-1/2,1/2]} \cdot \mathbf{1}_{[-1/2,1/2]} = \mathbf{1}_{[-1/2,1/2]}. \tag{392}$$ boxcarsqu

This implies among others, that we have $\varphi := \widehat{f} \in \boldsymbol{L}^2(Q)$ the pointwise relationship $\varphi = \varphi \cdot \mathbf{1}_{[-1/2,1/2]}$, or on the other hand the convolution relation

$$f = (f \star \mathrm{SINC}), \tag{393}$$ reprSINC1

or

$$f(x) = \int_{\mathbb{R}^d} f(y)\, \mathrm{SINC}(x-y)dy. \tag{394}$$ reprSCIN2

which by the Cauchy-Schwarz inequality implies

$$|f(x)| \leq \|SINC\|_{\boldsymbol{L}^2} \|f\|_{\boldsymbol{L}^2}, \tag{395}$$ bandlsupest

and furthermore, due to the fact that $\boldsymbol{L}^2(\mathbb{R}^d) \star \boldsymbol{L}^2(\mathbb{R}^d) \subset \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) \subset \boldsymbol{C}_0(\mathbb{R}^d)$ (again by Plancherel's Theorem) and the continuous shift property in $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$ (i.e. $\lim_{x\to 0} \|T_x f - f\|_{\boldsymbol{L}^2} = 0$) we find that band-limited $\boldsymbol{L}^2$-functions also belong to $\boldsymbol{C}_0(\mathbb{R}^d)$ and that $\|\cdot\|_2$-convergence of a series in the space of Ω-bandlimited functions also implies uniform convergence!
The main part of the proof is then a combination of the Fourier series expansion of $\varphi := \widehat{f} \in \boldsymbol{L}^2(Q)$, combined with the observation (Plancherel's Theorem) that

$$\mathcal{F}^{-1}(\mathbf{1}_{[-1/2,1/2]} \cdot \chi_k) = \mathcal{F}^{-1}(\mathbf{1}_{[-1/2,1/2]}) \star \mathcal{F}^{-1}(\chi_k) = \mathrm{SINC} \star \delta_{-k} = T_{-k}\, \mathrm{SINC}, \text{ for } k \in \mathbb{Z}^d.$$

Furthermore we have (again by Plancherel, for $\boldsymbol{L}^2(\mathbb{R}^d)$):

$$\widehat{\varphi}(k) = \langle \varphi, \chi_k \rangle = \langle \mathcal{F}^{-1} \varphi, \mathcal{F}^{-1}(\mathbf{1}_{[-1/2,1/2]} \cdot \chi_k) \rangle = f(-k),$$

(in fact, this is valid after one shows that f is in fact continuous, which is OK). Thus we obtain the following representation of $f = \mathcal{F}^{-1} \varphi$ (with convergence in $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$ as well as in $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$), Thus

$$\varphi = \sum_{k \in \mathbb{Z}^d} \widehat{\varphi}(k)(\chi_k \cdot \mathbf{1}_{[-1/2,1/2]})$$

implies by applying $\mathcal{F}^{-1}$ (which preserves convergence in $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$):

$$f = \mathcal{F}^{-1}(\varphi) = \left(\sum_{k \in \mathbb{Z}^d} f(-k)\, \mathcal{F}^{-1}(\mathbf{1}_{[-1/2,1/2]} \cdot \chi_k)\right) = \left(\sum_{k \in \mathbb{Z}^d} f(-k) T_{-k}\, \mathrm{SINC}\right),$$

which, upon substitution $k \to -k \in \mathbb{Z}^d$ gives the required result. □

51 The S_O-version of Shannon's Sampling Theorem

While the typical Shannon Sampling Theorem for band-limited functions in $\boldsymbol{L}^2(\mathbb{R}^d)$ makes use of *orthogonal expansions* in Hilbert spaces (using the pure frequencies over the domain Ω). But practically speaking the ideal filter, which corresponds to pointwise multiplication by $\mathbf{1}_\Omega$ on the Fourier transform side, so for $d=1$ and $\Omega=[-1/2,1/2]$ we have to do convolution with the SINC-function, which is smooth, but not well decaying. Therefore the following version, with a slightly restricted domain, is more useful.

ShannonSO1 **Theorem 21.** *Given a compact set $\Omega \subset \widehat{\mathbb{R}^d}$, and some lattice $\Lambda \lhd \mathbb{R}^d$ with the property that the $\Lambda^\perp$-translates of Ω are pairwise disjoint, then one recover any $f \in \boldsymbol{L}^1(\mathbb{R}^d)$ with* $\mathrm{supp}(\widehat{f}) \subset \Omega$ *can be recovered from the Λ-samples of f by the series expansion*

$$f(t) = \sum_{\lambda \in \Lambda} f(\lambda) g(t-\lambda), \tag{396}$$ ShannonrecSO

with convergence in $\big(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\big)$, in particular absolutely for every $t \in \mathbb{R}^d$ and uniformly (in $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$), for any $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ with $\widehat{g}(\omega) \equiv 1$ for $\omega \in \Omega$ and $\mathrm{supp}(\widehat{g}) \cap \lambda^\perp + \Omega$ *for any $\lambda^\perp \in \Lambda^\perp$, $\lambda^\perp \neq 0$.*

Proof. Any band-limited function in $\boldsymbol{L}^1(\mathbb{R}^d)$ belongs to $\boldsymbol{S}_0(\mathbb{R}^d)$ and we have seen that one has for fixed, compact set

$$\|f\|_{\boldsymbol{S}_0} \leq C_\Omega \|f\|_{\boldsymbol{L}^1}, \quad f \in \boldsymbol{L}^1(\mathbb{R}^d), \, \mathrm{supp}(\widehat{f}) \subset \Omega. \tag{397}$$ SOLiest02

This implies also that $(f(\lambda))_{\lambda \in \Lambda} \in \boldsymbol{\ell}^1(\Lambda)$ and altogether

$$\sum_{\lambda \in \Lambda} |f(\lambda)| \leq C'_\Omega \|f\|_{\boldsymbol{L}^1}. \tag{398}$$ sampestLam01

The task is to reconstruct f from

$$\sqcup\!\sqcup_\Lambda \cdot f = \sum_{\lambda \in \Lambda} f(\lambda) \delta_\lambda,$$

or equivalently the same information on the Fourier transform side:

$$\mathcal{F}(\sqcup\!\sqcup_\Lambda \cdot f) = C_\Lambda \sqcup\!\sqcup_{\Lambda^\perp} \star \widehat{f} \in \boldsymbol{S}'_0(\mathbb{R}^d).$$

Then the properties of g resp. $\widehat{g}$ imply

$$\widehat{g} \cdot [C_\Lambda \sqcup\!\sqcup_{\Lambda^\perp} \star \widehat{f}] = \widehat{f}.$$

Taking the inverse Fourier transform we obtain (recalling $\delta_\lambda \star g = T_\lambda g$)

$$f = \mathcal{F}^{-1}[\mathcal{F}(f)] = \mathcal{F}^{-1}(\widehat{g}) \star (\sqcup\!\sqcup_\Lambda \cdot f) = [\sum_{\lambda \in \Lambda} f(\lambda)\delta_\lambda] \star g = \sum_{\lambda \in \Lambda} f(\lambda) T_\lambda g,$$

with absolute convergence in $\big(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\big)$. $\square$

51.1 Pictorial Illustration

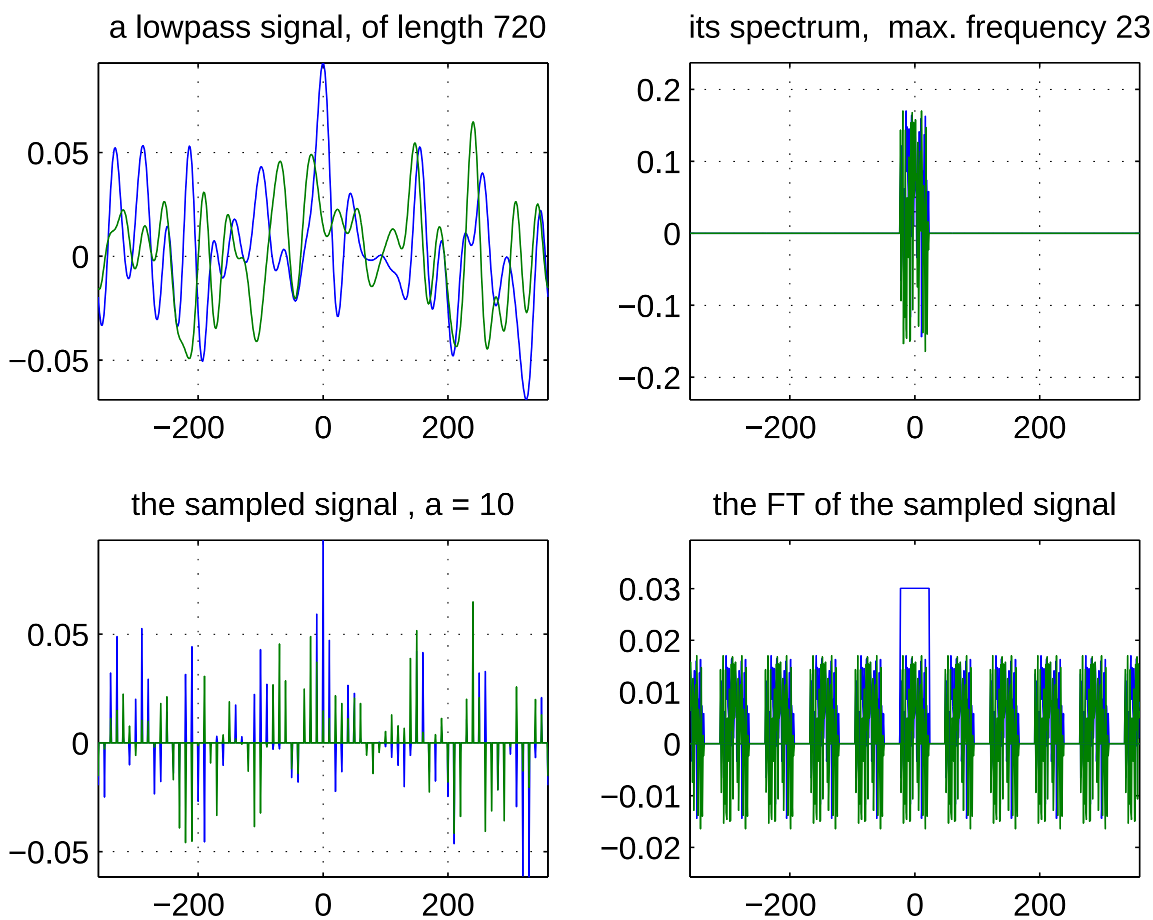

Abbildung 72: shannondem1b.pdf: Signal and sampled signal (to the left), and the corresponding spectrum/FT on the right hand side.

```
see  shannondem2.m MATLAB code by NuHAG (hgfei)
```

hanndem4.pdf

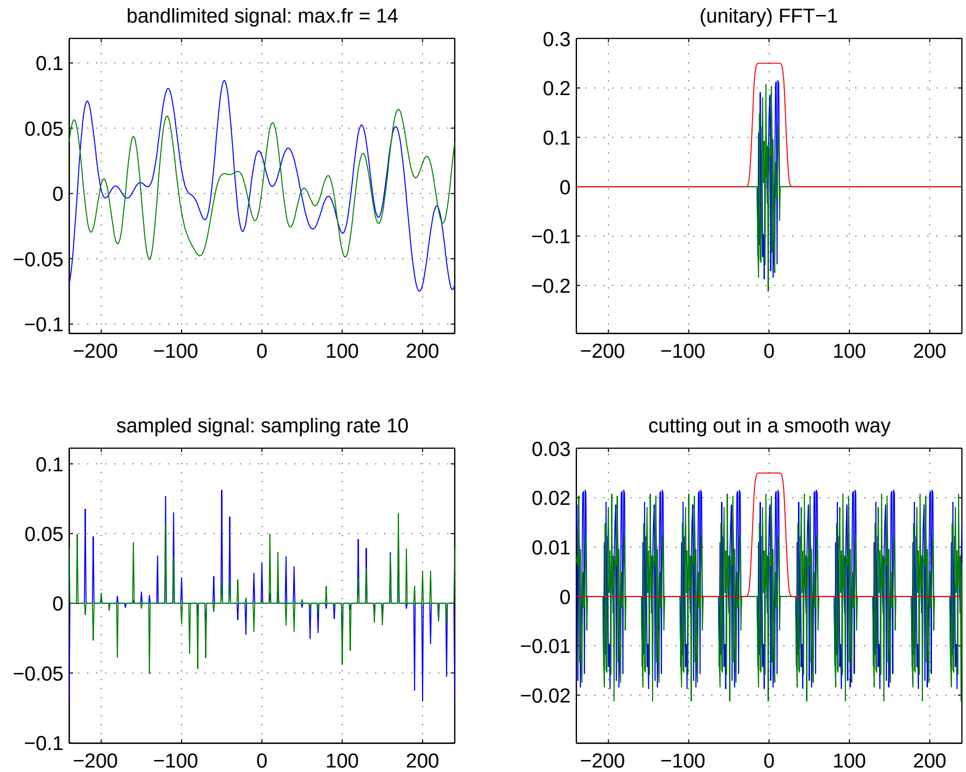

Abbildung 73: TEST4 shanndem4.pdf:
If there is a bit of oversampling, i.e. the s Dirac comb used for sampling is a bit more fine than the minimal requirement (Nyquist criterion) then one has more freedom in the choice of $\widehat{g}$, filtering out the basic period of $\widehat{f}$ from its periodic repetition. The red picture shows that one can take a *smooth* trapezoidal-like function, which shows better localization on the time side, see Fig. shannondem2B.pdf 51.1 below. The contributions of the individual shifted copis of g is then shown in Fig. shannondem2C.pdf 51.1.

52 Preparing for Gabor representations

The first step towards a general theory of Gabor expansions of signals can be based directly on the description of $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ as Wiener amalgam space $\boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1, \boldsymbol{\ell}^1)(\mathbb{R}^d)$. We call a norm convergent representation (as an unconditional sum) of some function or (mild) distribution f with respect to a pair (g, Λ), where $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ and $\Lambda \lhd \mathbb{R}^d \times \widehat{\mathbb{R}}^d$ is some lattice, any representation of f as (unconditionally convergent)

$$f = \sum_{\lambda \in \Lambda} c_\lambda \pi(\lambda) g, \quad \pi(t,s) = M_s T_t, \; t, s \in \mathbb{R}^d. \tag{399}$$ Gaboexpans00

The terminology is based on the early paper by D. Gabor from 1946 (**ga46:** [53]) who suggested to use the Gauss function $g_0(t) = e^{-\pi t^2}$ over $\mathbb{R}$ and $\Lambda = \mathbb{Z} \times \mathbb{Z} \lhd \mathbb{R} \times \widehat{\mathbb{R}}$.
Let us first treat the case which is inspired by the atomic characterization of Besov-Triebel-Lizorkin spaces realized by Frazier-Jawerth (see **frja90:** [52]), which also was the motivation for the corresponding atomic characterization of *modulation spaces* in **fe89-1:** [26]. Here we only (first) claim that such a representation is possible using

ondem2A.pdf

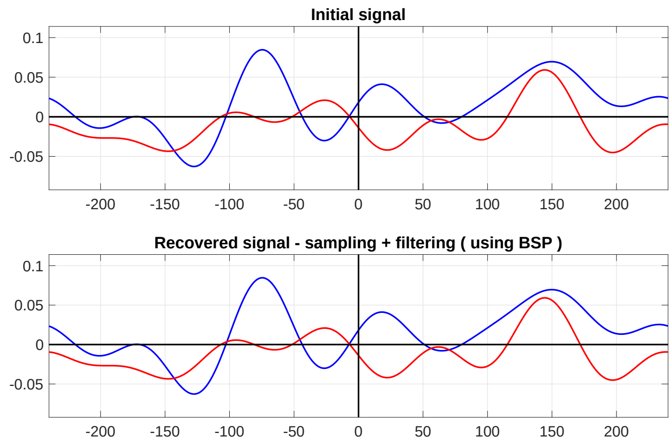

Abbildung 74: shannondem2A.pdf:
Showing a low-pass signal (random trigonometric polynomial with some maximal frequency): blue: real part, red: imaginary part. The lower plot gives the reconstructed by either method (i.e. not distinguishable from the original.)

simple arguments, by choosing $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ properly. Later on, we will collect information about the more general case, which is typically presented in the following form:
Question: Under which conditions on the pair (g, Λ) is it possible to find a bounded linear mapping $f \mapsto (c_\lambda)_{\lambda \in \Lambda}$ from $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right) \in (\boldsymbol{\ell}^2(\Lambda), \|\cdot\|_{\ell^2(\Lambda)})$ such that any $f \in \boldsymbol{L}^2(\mathbb{R}^d)$ one has a representation of the form (399). Let us do the case $d = 1$ first:

ondem2B.pdf

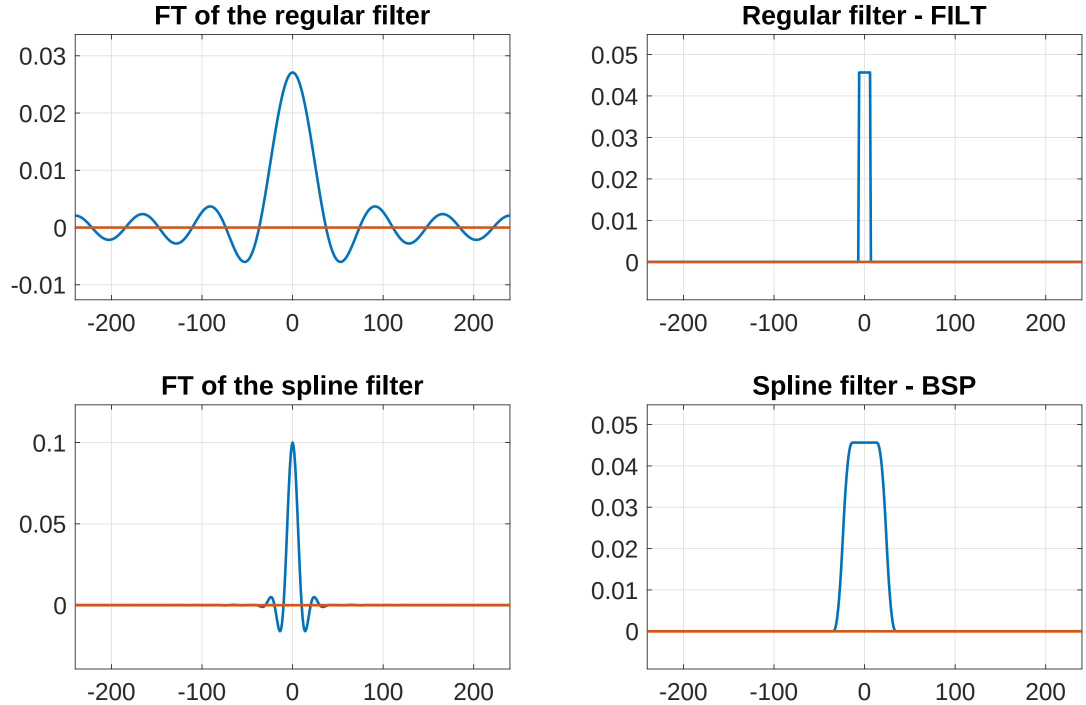

Abbildung 75: shannondem2B.pdf: different filters, watch for time-concentration!

GaborexpFrJ **Theorem 22.** *Let $\tau \in \boldsymbol{C}_c(\mathbb{R}^d)$ with $\widehat{\tau} \in \boldsymbol{L}^1(\mathbb{R})$, be some function which has a plateau near zero. Then there exists $a_0 > 0, b_0 > 0$ such that for any $a \in (0, a_0]$ and $b \in (0, b_0]$ one has: The Gabor family $(M_{bn}T_{ak}\tau)_{(n,k)\in\mathbb{Z}\times\mathbb{Z}}$ allows to expand any $f \in \boldsymbol{S}_0(\mathbb{R})$ in the following way: By choosing an appropriate coefficient mapping $C : \left(\boldsymbol{S}_0(\mathbb{R}), \|\cdot\|_{\boldsymbol{S}_0}\right) \to (\boldsymbol{\ell}^1(\mathbb{Z}\times\mathbb{Z}), \|\cdot\|_{\boldsymbol{\ell}^1(\mathbb{Z}\times\mathbb{Z})}),\ f \to (c_{n,k})_{(n,k)\in\mathbb{Z}\times\mathbb{Z}}$ such that these coefficients allow the following Gabor expansion of f:*

$$f = \sum_{(n,k)\in\mathbb{Z}\times\mathbb{Z}} c_{n,k} M_{bn} T_{ak} \tau. \tag{400}$$ GaborexpZZ

Remark 43. Since we claim that the coefficients are in $\boldsymbol{\ell}^1(\mathbb{Z}\times\mathbb{Z})$ it is clear that the sum is absolutely convergent in $\left(\boldsymbol{S}_0(\mathbb{R}), \|\cdot\|_{\boldsymbol{S}_0}\right)$, hence in $\left(\boldsymbol{C}_0(\mathbb{R}), \|\cdot\|_\infty\right)$ and pointwise. Thus the summation can taken either way, because it is clearly unconditional. This means, we can sum in *any order* and choose, for example, to create the double sum as an iterated sum of the form $\sum_{n\in\mathbb{Z}}\sum_{k\in\mathbb{Z}}$ or also $\sum_{k\in\mathbb{Z}}\sum_{n\in\mathbb{Z}}$, or as a limit of finite partial sums (e.g. by taking the big coefficients first).

Proof. Based on the assumption that $\tau(z) = 1$ over some interval $[-\gamma, \gamma]$ we start by using a any γ-BUPU, i.e. any $\psi \in \mathcal{F}\boldsymbol{L}^1(\mathbb{R})$ with $\mathrm{supp}(\psi) \subset [-\gamma, \gamma]$ and generate from it decomposition of the form $f = \sum_{k\in\mathbb{Z}} f \cdot \psi_k$, with $\mathrm{supp}(\psi_k) \subseteq [ak-\gamma, ak+\gamma]$, and thus $T_{ak}\tau \cdot \psi_k = \psi_k$, $k \in \mathbb{Z}$.
The idea is now do develop for each $k \in \mathbb{Z}$ a coarse, periodic version of $f \cdot \psi_k$ into a Fourier series. We use essentially a Fourier version the Shannon Sampling Theorem: For any b which is small enough ($b \in (0, b_0]$) we can assure (see Fig. 52) that locperpfat2b.pdf

$$T_{ak}\tau \cdot (f \cdot \psi_k) = T_{ak}\tau \cdot [(f \cdot \psi_k) \star Ш_{1/b}]. \tag{401}$$ Fougrrepr01

ondem2C.pdf

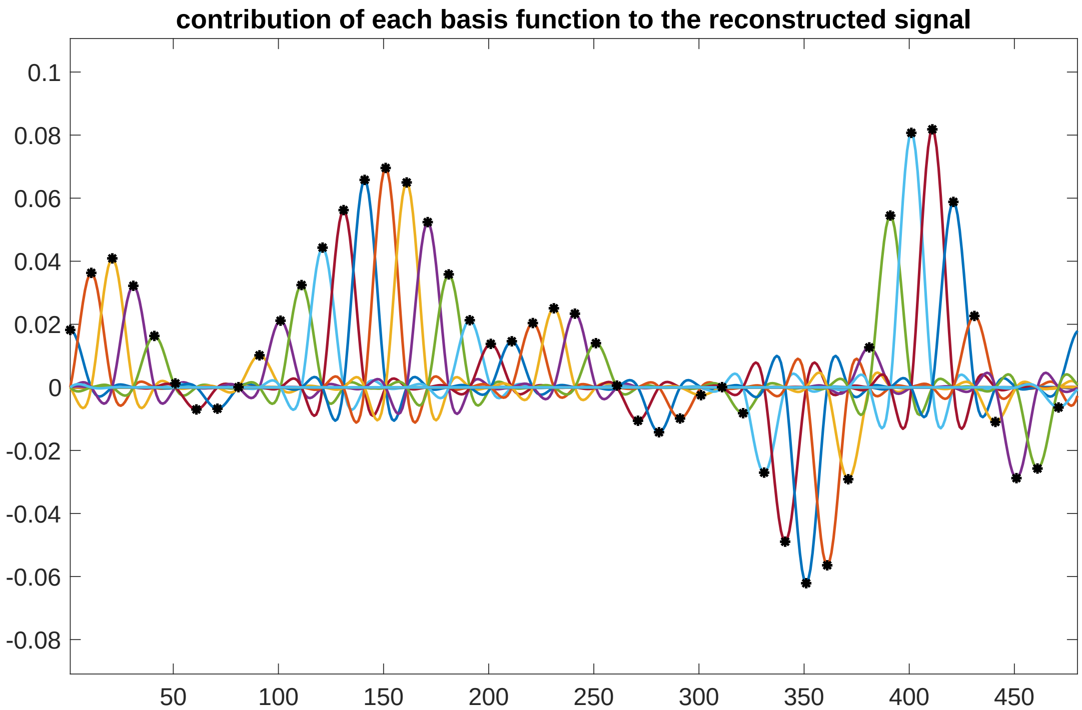

Abbildung 76: shannondem2C.pdf:
Showing the individual contributions of the well-localized version for the reconstruction of the **real part** of the signal.

It is enough to ensure that $\mathrm{supp}(\tau) \cap (j/b + \cdot B_r(ak)) = \emptyset$ for $j \neq 0$ (perhaps this should be made more clear, i.e. an estimate for b_0 should be given!).
The Fourier coefficients of the $[(f \cdot \psi_k) \star \sqcup\!\sqcup_{1/b}]$, the $1/b$-periodic version of $f\psi_k$ has Fourier coefficients which are the samples of $\mathcal{F}(f \cdot \psi_k)$ over $b\mathbb{Z}$, but since $f \cdot \psi_k \in \mathcal{F}\boldsymbol{L}^1(\mathbb{R})$ (by assumption) we can estimate these Fourier coefficients in the following form: There exists $C_1 > 0$ such that

$$\sum_{l \in \mathbb{Z}} |c_{k,l}| \leq C_1 \|f \cdot \psi_k\|_{\mathcal{F}\boldsymbol{L}^1}. \tag{402}$$ FourCofest02

The Fourier representation of $[(f \cdot \psi_k) \star \sqcup\!\sqcup_{1/b}]$ gives then for each $k \in \mathbb{Z}$:

$$[(f \cdot \psi_k) \star \sqcup\!\sqcup_{1/b}](x) = \sum_{n \in \mathbb{Z}} c_{n,k} \chi_{bn}(x). \tag{403}$$ FOurdecomp02

Subsequent multiplication by $T_{ak}\tau$ cuts out $f \cdot \psi_k$ and summation over $k \in \mathbb{Z}$ (absolutely convergent in $\left(\boldsymbol{S}_0(\mathbb{R}), \|\cdot\|_{\boldsymbol{S}_0}\right)$ is granted):

$$f = \sum_{k \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} c_{n,k} M_{bn} T_{ak} \tau. \tag{404}$$ Gaborexp007

$\square$

Remark 44. Note that in the above proof the coefficients are obtained making use of the BUPU Ψ (with small support), while the synthesis mapping makes use of the plateau-type function τ. In this sense the representation is a bit asymmetric. See Thm. 23 below for a more symmetric situation. tightFrJ ondem2D.pdf

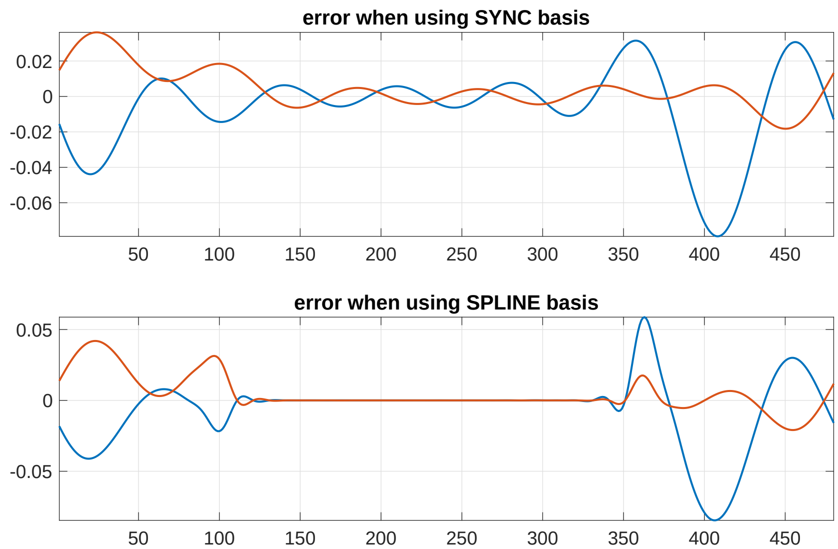

Abbildung 77: shannondem2D.pdf:
Showing the **significant effect** on local reconstruction, if not all the samples are available, but just those in the middle segment, in some interval around zero. We see that the alternative choice of g (with $\widehat{g}$ being the smooth cut-out function in Figs. 51.1) does a much better job compared to the "ideal" SINC-like building block. shannondem2A.pdf

The same expansion method can be used to extend the coefficient mapping to a bounded coefficient mapping from $\left(\boldsymbol{L}^2(\mathbb{R}), \|\cdot\|_2\right)$ into $(\boldsymbol{\ell}^2(\mathbb{Z}\times\mathbb{Z}), \|\cdot\|_{\ell^2(\mathbb{Z}\times\mathbb{Z})})$.

Remark 45. In the finite dimensional case, i.e. signals on $\mathbb{U}_n$ respectively elements of $\mathbb{C}^n$ the STFT maps into $(\boldsymbol{\ell}^2(\mathbb{Z}_N\times\mathbb{Z}_N), \|\cdot\|_{\ell^2(\mathbb{Z}_N\times\mathbb{Z}_N)})$, which can of course be identified with $\mathbb{C}^{N^2}$, with the usual Euclidean structure, very much as in our Gabor setting, where we identify

```
STxxg = stft(xx,g,a,b);    % typical n = 480, a = 20, b = 16
```

which comes in matrix format of size $n/b \times n/a$ in our experiments (so 30×24) , with

```
cofs = xx*G';
```

which is the collection of all (720) Gabor coefficients (taken in natural, i.e. in MATLAB order, i.e. with the order of the points as in the command

```
LAM = lattp(n,a,b);  ndLAM = find(LAM);     % gives 1,17,32,..
```

which is arranged *in such a way that we have*

```
STxx(:) == cofs(:);  % test: compnorm(STxx,cofs);
```

Remark 46. Although we have for every pair of functions $f,g\in\boldsymbol{L}^2(\mathbb{R}^d)$: $V_gf\in\boldsymbol{L}^2(\mathbb{R}^{2d})\cap \boldsymbol{C}_0(\mathbb{R}^{2d})$, it is not true anymore that the mapping from $f\in\boldsymbol{L}^2(\mathbb{R}^d)$ to the sampling values $(V_gf(an,bk))_{(n,k)\in\mathbb{Z}^d\times\mathbb{Z}^d}$ over $\Lambda = a\mathbb{Z}\times b\mathbb{Z}\lhd\mathbb{R}^d\times\widehat{\mathbb{R}}^d$ (with $a,b>0$) maps $\boldsymbol{L}^2(\mathbb{R}^d)$ boundedly into $\boldsymbol{\ell}^2(\mathbb{Z}^{2d})$. Expressed in the language of *frame theory* (see **ch16:** [13]) this means: **Not every** $g\in\boldsymbol{L}^2(\mathbb{R}^d)$ **has the property** that for $\Lambda = a\mathbb{Z}\times b\mathbb{Z}$ the family $(\pi(\lambda)g)_{\lambda\in\Lambda}$ is a *Bessel family*.

orsplit1.pdf

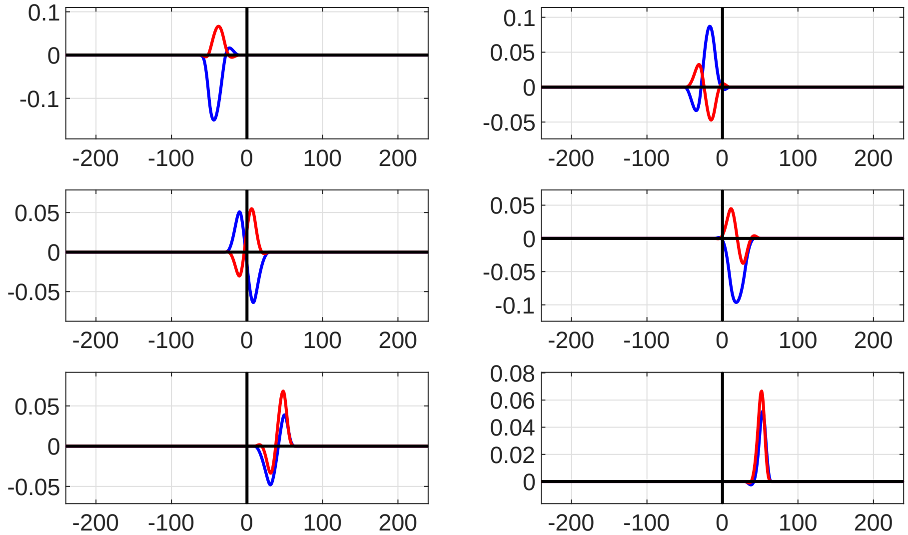

Abbildung 78: Gaborsplit1.pdf: Splitting a function into Gabor building blocks.

But let us recall this definition:

Besselfamdef

Definition 17. An indexed family $(f_i)_{i\in I}$ in a Hilbert space $\mathcal{H}$ is called a *Bessel family* if the mapping $f \mapsto \langle f, f_i\rangle_{\mathcal{H}}$ defines a bounded linear mapping from $\mathcal{H}$ into $(\boldsymbol{\ell}^2(I), \|\cdot\|_2)$.

Avoiding infinite summation and other technical issues, a good way to describe the Bessel property of a family is given by the following simple lemma:

charBessel1

Lemma 78. *An indexed family $(f_i)_{i\in I}$ in a Hilbert space $\mathcal{H}$ is a Bessel family if and only if there exists some constant (the optimal such constant may be called the Bessel bound of the family), such that for every finite subset $F \subset I$ one has:*

$$\sum_{i\in F} |\langle f, f_i\rangle_{\mathcal{H}} \leq C\|f\|_{\mathcal{H}}, \quad \forall f \in \mathcal{H}. \tag{405}$$ Besselest00

A detailed discussion of such questions is given in **feja00:** [42].
Let us state (for now without proof), that $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ is also a sufficient condition in this context:

S0GabBess

Lemma 79. *For $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ the family $(\pi(\lambda)g)_{\lambda\in\Lambda}$ is a Bessel family for any $\Lambda \lhd \mathbb{R}^d \times \widehat{\mathbb{R}}^d$.*

There is a certain asymmetry in the construction of the Gaborian expansion as described in Theorem GaborexpFrJ 22. Furthermore the *ingredients* have to be chosen properly, i.e.for the analysis step we chose a sufficiently fine regular BUPU, while for the synthesis the plateau-property of the synthesis function τ was playing an important role.
The natural question arising in this context is: Could one perhaps choose $\psi = \tau$? This is indeed possible, if we choose start from a non-negative BUPU $(\varphi_\lambda)_{\lambda\in\Lambda}$ in $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ and then choose $\tau = \psi = \sqrt{\varphi}$.
This is indeed possible!! We just observe, that we might start from a given BUPU $(\phi_\lambda)_{\lambda\in\Lambda}$, which is a periodic function which *does not vanish*! (Exercise) $\Phi(t) = \sum_{k\in\mathbb{Z}^d} T_k\phi^2$ (see Fig. sumBUPsqu2.eps) Hence we can apply the operator $z \to 1/\sqrt{z}$ to this function in

erplat2b.pdf

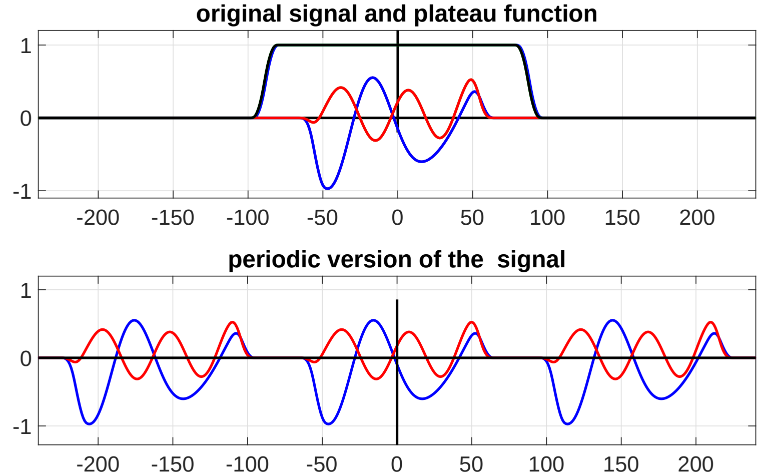

Abbildung 79: locperplat2b.pdf:
periodic repetition, cut out via plateau-type function (black)

order to find that it is (by the Wiener-Levy Theorem) an absolutely convergent sum of pure frequencies, and thus a multiplier for $\boldsymbol{S}_0(\mathbb{R}^d)$, hence we and set

$$\varphi := \phi/\sqrt{\Phi}$$

and be ensured that we have both:

$$\sum_{k\in\mathbb{Z}^d} T_k\varphi \equiv 1, \quad \text{and} \quad \varphi \in \boldsymbol{S}_0(\mathbb{R}^d).$$

The resulting statement can be formulated in a theorem:

tightFrJ

Theorem 23. *There exist functions $\varphi \in \boldsymbol{S}_0(\mathbb{R}^d)$ such that for (a,b) small enough (both) one has: The family $(M_{bn}T_{ak}\varphi)_{(n,k)\in\mathbb{Z}^d\times\mathbb{Z}^d}$ is a tight Gabor frame, i.e. we have for a suitable constant $C = C_{a,b} > 0$:*

$$f = C_{a,b} \sum_{(n,k)\in\mathbb{Z}\times\mathbb{Z}} \langle f, M_{bn}T_{ak}\varphi\rangle M_{bn}T_{ak}\varphi. \tag{406}$$ GaborexpZZ

It is one of the striking results of the (even more general) *coorbit theory* of Feichtinger/Gröchenig (see **fegr89:** [38] and related papers) that one can even start with any $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ (as potential generator) for a Gabor frame, and that a formula as the above one will be "approximately true" for sufficiently small lattice constants (a,b).

mBUPsqu2.eps

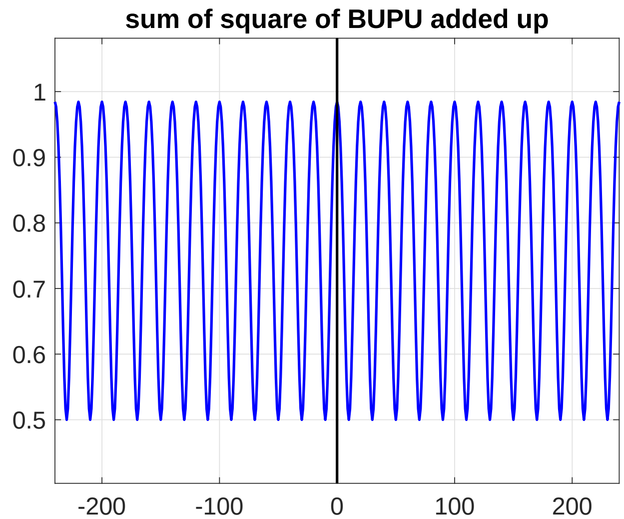

Abbildung 80: sumBUPsqu2.eps:
Showing the behaviour of the period version of g^2 in Prop. 83 SQBUPconstr

53 Pointwise multiplier for $\boldsymbol{S}_0(\mathbb{R}^d)$

ptwmultSO

The first result is a simple observation (cf. Lemma 63 above) period-ization

periodAT

Proposition 19. *Given $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ and $\Lambda \lhd \mathbb{R}^d$, $f_{per} = ⊔⊔_\Lambda \star f$ is a Λ-periodic function which has an absolutely convergent Fourier series expansions of the form*

$$f_{per} = \sum_{\lambda^\perp \in \Lambda^\perp} c_{\lambda^\perp} \chi_{\lambda^\perp},$$

hence f_{per} is a (uniformly!) continuous function.
If f_{per} is free of zeros, then the function $h^{(1)}_{per}(t) = 1/f_{per}(t)$ as well as the functions $h^{(2)}_{per}(t) = 1/\sqrt{f_{per}(t)}$ have an absolutely convergent Fourier series representation.
In particular each of these functions defines a pointwise multiplier to $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$.

Proof. We have seen already earlier (during the discussion of Poisson's formula) that the Fourier coefficients $c_{\lambda^\perp}$ of f_{per} are just the sampling values of $\widehat{f}$, i.e.

$$(c_{\lambda^\perp})_{\lambda^\perp \in \Lambda^\perp} = \widehat{f}|_{\Lambda^\perp} \in \boldsymbol{\ell}^1(\Lambda^\perp)$$

together with a corresponding norm estimate.
Thus the (classical) version of the Wiener Inversion Theorem resp. of the Wiener-Levy Theorem (functional calculus of analytic functions operating on the Fourier transform resp. Fourier coefficients) applies. We just have to check (the obvious fact) that the function $1/z$ or $1/\sqrt{z}$ are analytic on the range of f_{per}, i.e. on the set $\{f_{per}(t), tRd\}$, but this is obvious due to the fact that this is a *compact* subset of $\mathbb{C}$ (because the range is the same as the range of just a compact fundamental domain for $\Lambda^\perp$) and by assumption that it does not contain zero, thus there is a whole ball of the form $B_r(0)$ disjoint to the range of f_{per}.

Assuming now that $h^{(1)}$ or $h^{(2)}$ can be written as absolutely convergent sums of pure frequencies we conclude from the fact that modulation operators are uniformly bounded (if not isometric) on $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ that also these infinite (but absolutely convergent) series are also pointwise multipliers[46]. $\square$

The fact the $(\boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1, \boldsymbol{\ell}^\infty)(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1,\boldsymbol{\ell}^\infty)(\mathbb{R}^d)})$ can be identified with the space of pointwise multipliers of $\boldsymbol{S}_0(\mathbb{R}^d) = \boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1, \boldsymbol{\ell}^1)(\mathbb{R}^d)$ is an easy exercise. Let us state it explicitly as a lemma (for better citation).

SOptwmult **Lemma 80.** *A function $h \in \boldsymbol{C}_b(\mathbb{R}^d)$ is a pointwise multiplier on $\boldsymbol{S}_0(\mathbb{R}^d)$ if and only if it belongs locally to the Fourier algebra, and for some (and then any)$k \in \boldsymbol{C}_c(\mathbb{R}^d) \cap \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ one has:*

$$\sup_{x\in\mathbb{R}^d} \|h \cdot T_x k\|_{\mathcal{F}\boldsymbol{L}^1} < \infty. \tag{407}$$ WFLilinfccs

In fact, for each such k the expression (407) defines a norm on this subspace of $\left(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\right)$ which is equivalent to the norm of the multiplication operator: $M_h : f \mapsto h \cdot f$ as an operator on $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$, i.e. to $|||M_h|||_{\boldsymbol{S}_0(\mathbb{R}^d)}$.

Proof. Assume that h satisfies (407), then it is clear that $\sup_{x\in\mathbb{R}^d}\|h \cdot T_x k\|_\infty < \infty$ and hence this space (we could say a description of the Wiener amalgam space $\boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1, \boldsymbol{\ell}^\infty)(\mathbb{R}^d)$ via a continuous norm) ensures a continuous embedding into $\left(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\right)$.
Let us first assume that $M_h : f \mapsto h \cdot f$ defines a linear operator from $\boldsymbol{S}_0(\mathbb{R}^d)$ into $\boldsymbol{S}_0(\mathbb{R}^d)$. Since $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ is a Banach space we want to establish boundedness of this operator first. In fact, this is a typical situation where the Closed Graph Theorem (CGT) from Functional Analysis can be applied: We only have to show, that the multiplication mapping M_h has a closed graph. Boundedness then follows by general principles.
We just have to show (this is the elegant version of the CGT) that $f_n \to 0$ in $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ and $h \cdot f_n \to g$ in $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ for some $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ one has to show that $g = 0$. But this is easy to check, because convergence in $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ implies pointwise convergence. Thus $\lim_{n\to\infty} h(x) f_n(x) = h(x) \lim_{n\to\infty} f_n(x) = 0$ for any $x \in \mathbb{R}^d$. But the assumption $M_h f_n \to g$ in $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ also implies pointwise convergence, thus we have $g(x) = 0$, as was to be shown.
Since M_h is now a bounded operator on $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ we conclude that

$$\sup_{x\in\mathbb{R}^d} \|h \cdot T_x k\|_{\boldsymbol{S}_0} \leq |||M_h|||_{\boldsymbol{S}_0(\mathbb{R}^d)} < \infty$$

because $\{T_x k, x \in \mathbb{R}^d\}$ is a bounded subset of $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$. Consequently also the supremum over the $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$-norms is finite (due to the fact that $\mathrm{supp}(h \cdot T_x k) \subset \mathrm{supp}(T_x k) \subset B_R(x)$ for some $R > 0$), in fact this is an equivalent statement.
Let us now assume conversely assume that (407) is satisfies. We start from the atomic characterization of $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ given in Proposition 10. Since the finiteness of (407) for some $k \in \boldsymbol{C}_c(\mathbb{R}^d) \cap \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ also implies that it is valid for some other $p \in \boldsymbol{C}_c(\mathbb{R}^d) \cap$

[46] In general there is no generally valid estimate which can be derived from this abstract principle.

$\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ which has the plateau property, as in the proof of Lemma CSnorm2 72, with $p(x) = 1$ on a given compact set Q_1, i.e.

$$\sup_{x\in\mathbb{R}^d} \|h \cdot T_z p\|_{\mathcal{F}\boldsymbol{L}^1} = C_5 < \infty.$$

Let us now start with an atomic characterization with atoms of size $r > 0$, i.e. $B_r(0) \subseteq Q_1$. Then we have for an atom h_k with $\mathrm{supp}(h_k) \subseteq B_r(0)$:

$$\|h \cdot h_k\|_{\mathcal{F}\boldsymbol{L}^1} = \|h \cdot h_k \cdot T_x p\|_{\mathcal{F}\boldsymbol{L}^1} \leq \|h \cdot T_x p\|_{\mathcal{F}\boldsymbol{L}^1}\|h_k\|_{\mathcal{F}\boldsymbol{L}^1} \leq [\sup_{x\in\mathbb{R}^d} \|h \cdot T_x p\|_{\mathcal{F}\boldsymbol{L}^1}]\|h_k\|_{\mathcal{F}\boldsymbol{L}^1}. \tag{408}$$ Mhestim01

Since obviously $\mathrm{supp}(h \cdot h_k) \subseteq \mathrm{supp}(h_k) \subset B_r(x)$ this implies that $M_h(h_k) = h \cdot h_k$ is again an atom with the expected control of the norm in $\left(\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_{\mathcal{F}\boldsymbol{L}^1}\right)$. □

53.1 An alternative description of $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$

Since we do not have enough (technical) information about the dual space of $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ (usually described as $\left(\boldsymbol{L}^\infty(\mathbb{R}^d), \|\cdot\|_\infty\right)$, the space of *essentially bounded, measurable functions*) the following characterization will be useful:

0PcharlinfCb **Lemma 81.** *For any $\sigma \in \boldsymbol{S}_0'$ and any BUPU $\Psi = (\psi_k)$ in $\left(\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_{\mathcal{F}\boldsymbol{L}^1}\right)$ we have:*

$$\sigma = \sum_{k\in\mathbb{Z}^d} \sigma\psi_k, \quad \text{with } \mathcal{F}(\sigma\psi_k) \in \boldsymbol{C}_b(\mathbb{R}^d), \text{ and } \sup_{k\in\mathbb{Z}^d} \|\widehat{\sigma\psi_k}\|_\infty < \infty.$$

Conversely, for any $R > 0$ one has: Assume that $\sigma_k \in \mathcal{F}\boldsymbol{C}_b$, with

$$\mathrm{supp}(\sigma_k) \subset B_R(k) \text{ and } \sup_{k\in\mathbb{Z}^d} \|\widehat{\sigma_k}\|_\infty = C_2 < \infty,$$

then the w^*-convergent series $\sum_{k\in\mathbb{Z}^d} \sigma_k \in \boldsymbol{S}_0'(\mathbb{R}^d)$ defines some $\sigma \in \boldsymbol{S}_0'$ and up to some constant C_R one has: $\|\sigma\|_{\boldsymbol{S}_0'} \leq C_R C_2$.

In the language of Wiener amalgams we show that

$$\boldsymbol{S}_0'(\mathbb{R}^d) = \boldsymbol{W}(\boldsymbol{C}_b, \boldsymbol{\ell}^\infty)(\mathbb{R}^d) = \boldsymbol{W}(\boldsymbol{S}_0', \boldsymbol{\ell}^\infty)(\mathbb{R}^d),$$

in contrast to the identification of ${\boldsymbol{S}_0}^* = (\boldsymbol{M}^1(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}^1})^*$ as $\boldsymbol{W}(\mathcal{F}\boldsymbol{L}^\infty, \ell^\infty)(\mathbb{R}^d)$ respectively $(\boldsymbol{M}^\infty(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}^\infty})$ (in **gr01:** [59]).

Proof. As in the proof of Thm. S0Fourinv 8, implying the Fourier invariance of $\boldsymbol{S}_0(\mathbb{R}^d)$ we make use of plateau functions $p \in \boldsymbol{S}_0(\mathbb{R}^d)$, satisfying (platrepro2 409)

$$\sigma\psi_k = \sigma\psi_k \cdot T_k p, \quad k \in \mathbb{Z}^d. \tag{409}$$ platrepro2

By consequence, using $\boldsymbol{S}_0'(\mathbb{R}^d) \star \boldsymbol{S}_0(\mathbb{R}^d) \subset \boldsymbol{C}_{ub}(\mathbb{R}^d)$, we have for any $k \geq 1$:

$$\mathcal{F}(\sigma\psi_k) = \mathcal{F}(\sigma\psi_k) \star \mathcal{F}(p), \quad k \in \mathbb{Z}^d \tag{410}$$ platFTrepr2

$$\|\mathcal{F}(\sigma\psi_k)\|_{\boldsymbol{C}_b(\mathbb{R}^d)} \leq \|\mathcal{F}(\sigma\psi_k)\|_{\boldsymbol{S}_0'}\|\mathcal{F}(p)\|_{\boldsymbol{S}_0} = \|\mathcal{F}(p)\|_{\boldsymbol{S}_0} \cdot \|\sigma\psi_k\|_{\boldsymbol{S}_0'}.$$

The converse direction is a bit technical (and will not provided here), but is elementary in nature, based on the *atomic characterization* of $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$. □

This characterization has a very useful consequence:

Corollary 11. *Given $\sigma \in \boldsymbol{S}_0'(\mathbb{R}^d)$ and $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ we have: $\sigma \star f \in \boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1, \boldsymbol{\ell}^\infty)(\mathbb{R}^d)$, together with corresponding norm estimates, i.e. there exists $C_3 > 0$ such that:*

$$\|\sigma \star f\|_{\boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1,\boldsymbol{\ell}^\infty)} \leq C_4 \|\sigma\|_{\boldsymbol{S}_0'} \cdot \|f\|_{\boldsymbol{S}_0}. \tag{411}$$ S0PconvS0est

Proof. We only have to control the action on atoms of the form

$$(\sigma\psi_j) \star (f\psi_l) \in \mathcal{F}\boldsymbol{C}_b \star \mathcal{F}\boldsymbol{L}^1 \subset \mathcal{F}\boldsymbol{L}^1$$

because the convolution product $\mathcal{F}\boldsymbol{C}_b(\mathbb{R}^d) \star \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ corresponds to pointwise multiplication on the Fourier transform side, and obviously we have $\boldsymbol{C}_b(\mathbb{R}^d) \cdot \boldsymbol{L}^1(\mathbb{R}^d) \subset \boldsymbol{L}^1(\mathbb{R}^d)$, together with the obvious norm estimate:

$$\|h \cdot k\|_{\boldsymbol{L}^1} \leq \|h\|_\infty \cdot \|k\|_{\boldsymbol{L}^1}, \quad h \in \boldsymbol{C}_b(\mathbb{R}^d), k \in \boldsymbol{C}_c(\mathbb{R}^d),$$

extended by approximation[47] to $g \in \boldsymbol{L}^1(\mathbb{R}^d)$. □

Note that $\boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1, \boldsymbol{\ell}^\infty)(\mathbb{R}^d)$ is the space of pointwise multipliers of $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$, and hence the corollary implies

$$(\boldsymbol{S}_0'(\mathbb{R}^d) \star \boldsymbol{S}_0(\mathbb{R}^d)) \cdot \boldsymbol{S}_0(\mathbb{R}^d) \subset \boldsymbol{S}_0(\mathbb{R}^d), \tag{412}$$ S0PconvS0mS0

and the same relationship, on the Fourier transform side reads:

$$(\boldsymbol{S}_0'(\mathbb{R}^d) \cdot \boldsymbol{S}_0(\mathbb{R}^d)) \star \boldsymbol{S}_0(\mathbb{R}^d) \subset \boldsymbol{S}_0(\mathbb{R}^d). \tag{413}$$ S0PmS0convS0

Remark 47. Lemma 81 (S0PcharlinfCb) actually is equivalent to the following statement (in the terminology of Wiener amalgams):

$$\boldsymbol{S}_0'(\mathbb{R}^d) = \boldsymbol{W}(\boldsymbol{S}_0', \boldsymbol{\ell}^\infty)(\mathbb{R}^d) = \boldsymbol{W}(\mathcal{F}\boldsymbol{C}_b, \boldsymbol{\ell}^\infty)(\mathbb{R}^d), \tag{414}$$ S0PWlinfchar

with equivalence of the corresponding norms. One could also argue (by taking the Fourier perspective), that Ω band-limited elements of $\boldsymbol{S}_0'(\mathbb{R}^d)$ are represented by bounded and continuous functions, and the $\boldsymbol{S}_0'$-norm is equivalent to the sup-norm.

Remark 48. Let $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ be any Banach space embedded into $\boldsymbol{S}_0'(\mathbb{R}^d)$ such that $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ is a pointwise Banach module over $\left(\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_{\mathcal{F}\boldsymbol{L}^1}\right)$, i.e. with $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) \cdot \boldsymbol{B} \subseteq \boldsymbol{B}$ with corresponding norm estimate. This is true if $\boldsymbol{S}_0(\mathbb{R}^d)$ is dense in $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ and modulation operators are isometric.

[47] We could not rely on the Lebesgue argument $\boldsymbol{L}^\infty(\mathbb{R}^d) \cdot \boldsymbol{L}^1(\mathbb{R}^d) \subset \boldsymbol{L}^1(\mathbb{R}^d)$!

SEPARATE LEMMA! 17/11/2020

Then we have an equivalence of the BUPU description (using $\Psi = (\psi_k)$)

$$\boldsymbol{W}(\boldsymbol{B}, \boldsymbol{\ell}^\infty)(\mathbb{R}^d) := \{f \in \boldsymbol{S}_0'(\mathbb{R}^d) \,|\, \forall k \in \mathbb{Z}^d, f\psi_k \in \boldsymbol{B}, \sup_{k\in\mathbb{Z}^d} \|f\psi_k\|_{\boldsymbol{B}} < \infty\} \tag{415}$$ WBlinfBUP

coincides with the atomic characterization similar to Lemma 81 above. SOPcharlinfCb

Limodul1 **Lemma 82.** *Assume that $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ is a Banach space continuously embedded into $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$, containing $\boldsymbol{S}_0(\mathbb{R}^d)$ as a dense subspace. If the norm is isometrically translation invariant, then $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ is an essential Banach module over $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ with respect to convolution.*

Proof. We know that for any $f \in \boldsymbol{B}$ and $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ the convolution $\mu \star f \in \boldsymbol{M}_b(\mathbb{R}^d) \star \boldsymbol{S}_0'(\mathbb{R}^d) \subset \boldsymbol{S}_0'(\mathbb{R}^d)$ is *well defined*.
But - and this is a lengthly, but elementary argument - it is the (w^*) limit of expressions of the form $D_\Psi \mu \star f = \sum_{i\in I} \mu(\psi_i) T_{x_i} f$.
Since translation is isometric in $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ and continuous in $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ with respect to the $\boldsymbol{S}_0$-norm, we can argue by density that it is also continuous in $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$:

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

This, in turn, implies that the family $(D_\Psi \mu \star f)_{|\Psi|\to 0}$ is a Cauchy sequence in $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$, which - due to the completeness of $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ - has a limit. This limit is also a limit with respect to the norm of $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$ (due to the embedding $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}}) \hookrightarrow (\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$) and thus coincides with $\mu \star f \in \boldsymbol{S}_0'(\mathbb{R}^d)$ as defined already (as a w^*-limit in $\boldsymbol{S}_0'(\mathbb{R}^d)$). $\square$

Note that the general assumptions are Fourier invariant. If we have a Banach space sandwiched between $\boldsymbol{S}_0(\mathbb{R}^d)$ and $\boldsymbol{S}_0'(\mathbb{R}^d)$ which is modulation invariant, then $\mathcal{F}\boldsymbol{B}$ (with the norm $\|\widehat{f}\|_{\mathcal{F}\boldsymbol{B}} = \|f\|_{\boldsymbol{B}}$) is isometrically translation invariant. Thus the corresponding Fourier version of the above lemma reads as follows:

FLimodul1 **Corollary 12.** *Assume that $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ is a Banach space continuously embedded into $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$, containing $\boldsymbol{S}_0(\mathbb{R}^d)$ as a dense subspace. If the norm is isometrically modulation invariant, then $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ is an essential Banach module over $\left(\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_{\mathcal{F}\boldsymbol{L}^1}\right)$ with respect to pointwise multiplication.*

53.2 Towards tight Gabor frames

Tight Gabor frames are particularly useful for the construction of Gabor multipliers, because in this situation real-valued (or non-negative) multiplier sequences define symmetric (resp. non-negative) operators.

SQBUPU **Definition 18.** A function ϕ defines a *regular SQ-BUPU* for the lattice $\Lambda \lhd \mathbb{R}^d$ if one has

$$\sum_{\lambda\in\Lambda} T_\lambda \phi^2 \equiv 1.$$

We will see that it is a necessary condition for a tight Gabor frame for $\left(\boldsymbol{L}^2(\mathbb{R}), \|\cdot\|_2\right)$ generated by (g,a,b) that g defines a SQ-BUPU (because this condition implies that the main diagonal of the Gabor frame matrix coincides behaves correctly).

SQBUPconstr **Lemma 83.** *Given a function $\psi \in \boldsymbol{S}_0(\mathbb{R}^d)$ and $\Lambda \lhd \mathbb{R}^d$ such that $\Psi = \sum_{\lambda\in\Lambda} T_\lambda|\psi|^2$ is free of zeros. Then $\phi := \psi/\sqrt{\Psi}$ defines a SQ-BUPU.*

Proof. Since ψ belongs to $\boldsymbol{S}_0(\mathbb{R}^d)$ the function Ψ is a periodic function which is free of zeros (by assumption), but the Fourier coefficients are just the sampling values (perhaps up to some constant) of $(\widehat{\psi}(\lambda^\perp))_{\lambda^\perp\in\Lambda^\perp} \in \boldsymbol{\ell}^1(\Lambda^\perp)$. By the Wiener-Levy Theorem the pointwise inverse square root, which is obviously also Λ-periodic, also has a Fourier series expansion with $\boldsymbol{\ell}^1$-coefficients. But this implies that it defines a pointwise multiplier of $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ (since multiplication with pure frequencies is isometric on $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ and we are talking about an absolutely convergent sum of such operators). This implies first of all the $\psi \in \boldsymbol{S}_0(\mathbb{R}^d)$.
Finally we have to check that it generates a SQ-BUPU. This we check first that

$$T_\lambda\psi^2 = T_\lambda\phi^2 \cdot \frac{1}{T_\lambda\Phi} = T_\lambda\phi^2/\Phi, \quad \lambda\in\Lambda,$$

and thus by summing over $\lambda\in\Lambda$ we obtain

$$\sum_{\lambda\in\Lambda} T_\lambda\psi^2(t) = [\sum_{\lambda\in\Lambda} T_\lambda\phi^2(t)]/\Phi(t) = 1, \quad \forall x \in \mathbb{R}^d.$$

$\square$

This concept (and also the fact that it is "constructively realizable") is relevant when one searches for *tight Gabor frame*, it is related to the so-called *painless case* of Gabor analysis, which applies to compactly supported windows (proofs given later):

painless **Proposition 20** (Painless Case:)**.** *Assume that the with $\tau \in \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) \cap \boldsymbol{C}_c(\mathbb{R}^d)$ (i.e. $g \in \boldsymbol{C}_c(\mathbb{R}^d)$ is given, with $\widehat{g} \in \boldsymbol{L}^1(\mathbb{R}^d)$), and assume furthermore that $\Lambda \lhd \mathbb{R}^d$ is such that the function $\Phi_{\Lambda-per} = ⧢_\Lambda \star g$ is free of zeros. Then there exists b_0 such that $b \in (0, b_0]$ one has: The frame operator $S = S_{g,\Lambda}$ is just a simple multiplication operator, and thus the canonical dual window is given by $\widetilde{g} = g/(⧢_\Lambda \star |g|^2)$.*
In particular, $\widetilde{g}$ also has compact support and belongs to $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ as well, and $\widetilde{g} = g$ if g generates an SQ-BUPU, i.e. it (g,a,b) defines a tight Gabor frame.

53.3 Viewing $\boldsymbol{S}_0'(\mathbb{R}^d)$ as a completion of $\boldsymbol{S}_0(\mathbb{R}^d)$

For the case of tempered distributions the books of Lighthill (see ([74]) or Jones ([66] offer an approach which avoids topological vector spaces but uses instead *equivalence classes of good functions*. This is fine, and one can do a lot with this approach, but it turns out that one needs (the) heavy machinery of topological vector spaces (including the so-called *barreled* spaces) in order to establish the mathematical equivalence of the original concept of L. Schwartz with the sequential approach (which goes back to work of G. Temple (**te53:** [87] and **te55:** [88], to my knowledge). In contrast, I have shown in the recent paper **fe20-1:** [35], entitled "A sequential approach to mild distributions" that indeed it is relatively easy to establish the equivalence of these concepts.
One direction relies on the regularization operators in combination with relation (412) SOPconvSOmSO (using bounded approximate units in $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ resp. $\left(\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_{\mathcal{F}\boldsymbol{L}^1}\right)$, with elements from the dense subspace $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$, in order to obtain for a given $\sigma \in \boldsymbol{S}_0'(\mathbb{R}^d)$ a mild Cauchy sequence.
Starting from $\left(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\right)$, a space of functions which have obviously a bounded spectrogram (i.e.$\|(\|_\infty V_g h) < \infty$) one tried to enlarge the reservoir of "generalized functions" which still have a bounded spectrogram.
From our approach we expect that this will contain in fact $\boldsymbol{S}_0'(\mathbb{R}^d)$, since

$$|V_g\sigma(s,t)| \leq \|\sigma\|_{\boldsymbol{S}_0'}\|g\|_{\boldsymbol{S}_0}, \quad \sigma \in \boldsymbol{S}_0', f \in \boldsymbol{S}_0.$$

The basic idea is to define "*mild Cauchy sequences*" $(h_n)_{n=1}^\infty$ in $\boldsymbol{C}_b(\mathbb{R}^d)$, if one as pointwise convergence of the spectrograms $V_g(h_n)$ for every point in phase space (i.e. $(t,s) \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d$. (p.8) of the cited paper).
The main result of **fe20-1:** [35] established an equivalence between the elements $\sigma \in \boldsymbol{S}_0(\mathbb{R}^d)$ with *equivalence classes of mild Cauchy sequences*, so called **ECmiCS** (Thm.2., p.10).
The proof there relies on Prop.1 ([35]) which in turn makes use of general convolution and pointwise relations for Wiener amalgam spaces. In THIS note we have tried to avoid the direct use if these statements, in particular, the use of $\mathcal{F}\boldsymbol{L}^\infty(\mathbb{R}^d)$ (including he convolution relation $\mathcal{F}\boldsymbol{L}^\infty(\mathbb{R}^d) \star \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) \subset \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$, which relies on the simple argument - in the Lebesgue context - that one has a pointwise relation together with the natural norm estimate:

$$\boldsymbol{L}^\infty(\mathbb{R}^d) \cdot \boldsymbol{L}^1(\mathbb{R}^d) \subset \boldsymbol{L}^1(\mathbb{R}^d), \quad \|h\dot{f}\|_{\boldsymbol{L}^1} \leq \|h\|_\infty \|f\|_{\boldsymbol{L}^1}. \tag{416}$$

LinfptwLi

53.4 The Riemann-integral paper

In the paper **feja20:** [41] an alternative approach to the spaces $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ and $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$, only relying on the Riemann integral, is given.
It provides all the details, including the Fourier invariance of the space, starting from Riemann integration, and thus defining $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ as a subspace of $\left(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\right)$, but with the integrability condition on $V_g f$. We will discuss this later on, as time permits.

54 BIBOS II

54.1 Periodic signals

Given any $\Lambda \lhd \mathbb{R}^d$ it is clear that the subspace of all Λ-periodic elements is a closed subspace of $\big(\boldsymbol{C}_b(\mathbb{R}^d), \, \|\cdot\|_\infty\big)$.We write

$$\boldsymbol{C}_\Lambda(\mathbb{R}^d) := \{f \in \boldsymbol{C}_b(\mathbb{R}^d) \,|\, T_\lambda f = f \ \ \forall \lambda \in \Lambda\}.$$

Note that the suppression of the latter b (for bounded and continuous in the symbol $\boldsymbol{C}_b(\mathbb{R}^d)$) is justified, because any periodic function which is continuous is automatically *bounded and uniformly continuous*. This is due to the fact that the fundamental domain for the lattice Λ is relatively compact, or in other words, there is a compact set Q such that $\mathbb{R}^d = \bigcup_{\lambda \in \Lambda} \lambda + Q$, and hence the properties of $f \in \boldsymbol{C}_\Lambda$ are completely determined by the restriction of f to such a compact set: By the compactness $|f|$ attains a maximum over Q (which determines $\|f\|_\infty$) and it is equicontinuous over Q (or $Q + B_1(0)$ to be on the safe side).

LamperBUP **Lemma 84.** *For any regular BUPU in $\boldsymbol{C}_c(\mathbb{R}^d)$ of the form $(\psi_\lambda)_{\lambda \in \Lambda}$, with $\psi_\lambda = T_\lambda \psi$ (for some $\psi \in \boldsymbol{C}_c(\mathbb{R}^d)$) one has:*

$$f \in \boldsymbol{C}_\Lambda \quad \Leftrightarrow \quad f = \sqcup\!\sqcup_\Lambda \star (f\psi). \tag{417}$$ periodBUP

On the other hand, the convolution operator $f \mapsto \sqcup\!\sqcup_\Lambda \star f$ maps $\big(\boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d), \, \|\cdot\|_{\boldsymbol{W}}\big)$ into $(\boldsymbol{C}_\Lambda, \|\cdot\|_\infty)$. Thus it is in fact enough to assume that ψ above is in $\boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d)$.

Let us now check what a BIBO-system does to the space $\boldsymbol{C}_\Lambda \subset \boldsymbol{C}_{ub}(\mathbb{R}^d) \subset \boldsymbol{C}_b(\mathbb{R}^d)$:

BIBOCLam **Lemma 85.** *Given any BIBOs operator on $\big(\boldsymbol{C}_0(\mathbb{R}^d), \, \|\cdot\|_\infty\big)$, then its extension to $\big(\boldsymbol{C}_b(\mathbb{R}^d), \, \|\cdot\|_\infty\big)$ also maps $\boldsymbol{C}_\Lambda$ into $\boldsymbol{C}_\Lambda$.*
However, uniqueness is (drastically) lost: If μ is a representative with the property $T(f) = \mu \star f$ for all $f \in \boldsymbol{C}_\Lambda$, then any linear combination (or series) of the form

$$\nu = \sum_{k=1}^\infty c_k T_{\lambda_k} \mu, \quad \text{with } \sum_{k=1}^\infty |c_k| < \infty \text{ and } \sum_{k=1}^\infty c_k = 1$$

will give the same action on $\boldsymbol{C}_\Lambda$, i.e.

$$\mu \star f = \nu \star f, \quad \forall f \in \boldsymbol{C}_\Lambda. \tag{418}$$ BIBOCLam2

Proof. The proof is quite obvious: we have

$$T_{\lambda_k} f = f \quad \forall f \in \boldsymbol{C}_\Lambda$$

and hence (since the series is obviously absolutely convergent in $\big(\boldsymbol{C}_b(\mathbb{R}^d), \, \|\cdot\|_\infty\big)$), using the translation invariance of $T: f \mapsto \mu \star f$:

$$\nu \star f = \sum_{k=1}^\infty c_k (T_{\lambda_k} \mu \star f) = \sum_{k=1}^\infty c_k \mu \star (T_{\lambda_k} f) = \sum_{k=1}^\infty (c_k \mu \star f) = \mu \star f.$$

$\square$

In order to reduce (for fixed Λ) this ambiguity we recall that it is possible to split any measure, with respect to any BUPU, into an absolutely convergent sum. Thus we can also use the Λ-invariant BUPU used above and have:

$$\mu = \sum_{\lambda\in\Lambda} \mu\psi_\lambda \quad \text{with} \sum_{\lambda\in\Lambda} \|\mu\psi_\lambda\|_{\boldsymbol{M}_b} = \|\mu\|_{\boldsymbol{M}_b}, \quad \forall \mu \in (\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b}). \tag{419}$$

muLamBUP

This implies the following fact:

Lemma 86. *For any $\Lambda \lhd \mathbb{R}^d$ there exists a compact set $Q \subset \mathbb{R}^d$ (or some ball $B_R(0)$) such that the action of a BIBOS on $\boldsymbol{C}_\Lambda$ can be described as the convolution operator by some ν with* $\operatorname{supp}(\nu) \subset Q$.

BIBOmuloc

Proof. We can use (the proof of) Lemma 85 and choose

$$\nu = \sum_{\lambda\in\Lambda} T_{-\lambda}(\mu\psi_\lambda). \tag{420}$$

BIBOCLam3

Then due to (419) we have absolute convergence of the series, and the action on $f \in \boldsymbol{C}_\Lambda$ equals the action of μ on $\boldsymbol{C}_\Lambda$, since

$$\mu \star f = \sum_{\lambda\in\Lambda} (\mu\psi_\lambda) \star f = \sum_{\lambda\in\Lambda} T_{-\lambda}(\mu\psi_\lambda) \star T_\lambda f = \sum_{\lambda\in\Lambda} T_{-\lambda}(\mu\psi_\lambda) \star f = \nu \star f, \tag{421}$$

BIBOClam5

But obviously $T_{-\lambda}(\mu \cdot T_\lambda\psi) = T_{-\lambda}\mu \cdot \psi$ implies that we have for some $R > 0$:

$$\operatorname{supp}(T_{-\lambda}(\mu\psi_\lambda)) \subseteq \operatorname{supp}(\psi) =: Q \subseteq B_R(0).$$

$\square$

This observation greatly reduces the variety of bounded measures representing the *action* of a BIBO system T on $\boldsymbol{C}_\Lambda$, but it still does not make it unique.
Let us therefor restrict our attention to the case of an absolutely continuous measure, i.e. some $f \in \boldsymbol{L}^1(\mathbb{R}^d)$, and let us discuss the general case above for the lattice $\Lambda = \mathbb{Z}^d \lhd \mathbb{R}^d$[48]. In this case it is natural to choose as a fundamental domain $Q = [-1/2, 1/2]$ (this has the advantage of being symmetric with respect to the origin $0 \in \mathbb{R}^d$), or, following the tradition: $Q = [0,1]^d$, the d-dimensional unit cube. This makes the representation unique:

BIBOZd

Lemma 87. *Given a BIBOs $T : f \mapsto g \star f$, with some $g \in \boldsymbol{L}^1(\mathbb{R}^d)$, acting on $\boldsymbol{C}_\Lambda$ (with $\Lambda = \mathbb{Z}$), there exists a uniquely determined $g_Q = (\sqcup\!\sqcup \star g)\mathbf{1}_Q \in \boldsymbol{L}^1(\mathbb{R}^d)$ with* $\operatorname{supp}(g_Q) \subseteq Q$ *and*

$$\|g_Q\|_{\boldsymbol{L}^1} \leq \|g\|_{\boldsymbol{L}^1} \quad \text{and} \quad g_Q \star f = g \star f = T(f), \forall f \in \boldsymbol{C}_\Lambda.$$

Proof. Following Lemma 54.1 we can represent the action of a convolution operator by another, obtained with the help of an arbitrary BUPU. In the present case we can work with pointwise multiplication by $T_k(\mathbf{1}_Q)$, which is - so to say - a spline BUPU of order 1

[48] The general case can be easily transformed into this case, recalling that for general lattices there are (also non-unique) matrices $\boldsymbol{A}$ such that $\Lambda = \boldsymbol{A} \star \mathbb{Z}^d$.

(or degree zero). Since the boundary of $Q = [-1,1]^d$ is a set of measure zero the partition of unity property is valid in the a.e.-sense, i.e.

$$\sum_{k\in\mathbb{Z}^d} T_k(\mathbf{1}_Q)(z) = \sum_{k\in\mathbb{Z}^d} \mathbf{1}_{k+Q} = 1 \ a.e. \text{on } \mathbb{R}^d.$$

Thus g_Q as described is just the realization of ν as described in Lemma BIBOClam **??** above, under the concrete circumstances ($\Lambda = \mathbb{Z}^d, \mu = \mu_g$ using the trivial BUPU with $\psi = \mathbf{1}_{[0,1)}$). The norm estimate is clear because we have

$$\|g_Q\|_{\boldsymbol{L}^1} \leq \sum_{k\in\mathbb{Z}^d} \|g\cdot \mathbf{1}_{k+Q}\|_{\boldsymbol{L}^1} = \|g\|_{\boldsymbol{L}^1(\mathbb{R}^d)}\,.$$

Finally we have to show that the reshaping of g to g_Q induces uniqueness. So let us assume that we have two functions g_Q^1 and g_Q^2, both vanishing outside of Q and such that $g_Q^1 \star f = g_Q^2 \star f$ for all $f \in \boldsymbol{C}_\Lambda$ or $\boldsymbol{C}_{\mathbb{Z}^d}$.
Of course we can write $f = \sum_{k\in\mathbb{Z}^d} f\cdot \mathbf{1}_{k+Q}$, or $f = \sqcup\!\sqcup_{\mathbb{Z}^d} \star f_Q$ with $f_Q = f\cdot \mathbf{1}_Q$. Then

$$g_Q^1 \star (\sqcup\!\sqcup_{\mathbb{Z}^d} \star f_Q) = g_Q^2 \star (\sqcup\!\sqcup_{\mathbb{Z}^d} \star f_Q),$$

for any continuous function (in the sense of $\mathbb{R}^d$) f_Q, or equivalently

$$(g_Q^1 \star \sqcup\!\sqcup_{\mathbb{Z}^d}) \star f_Q = (g_Q^2 \star \sqcup\!\sqcup_{\mathbb{Z}^d}) \star f_Q.$$

We can now choose $f = \chi_k$, $k \in \mathbb{Z}^d$, the pure frequencies, and find that g_Q^1 and g_Q^2 have the same Fourier coefficients! This implies that $g_Q^1(x) = g_Q^2(x)$ are equal almost everywhere, or equal as elements of $\boldsymbol{L}^1(Q)$. Thus we have established the *uniqueness*, under the assumption that Q is some (!) fundamental domain for the periodicity lattice (here $\Lambda = \mathbb{Z}^d$). □

This viewpoint helps us to understand why "cyclic convolution" is arising in a natural way. We describe it at the level of functions first:
Given two functions $g^1, g^2 \in \boldsymbol{L}^1(Q)$ each of them describes a convolution operator on $\boldsymbol{C}_\Lambda$. It is clear that the concatenation of two such operators gives another BIBO system, and we know already that it is described by the ordinary convolution product $g^1 \star g^2$ (inside $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$, or via $\mu_{g^1} \star \mu_{g^2}$, just realized in a pointwise sense).
But we have

$$\mathrm{supp}(g^1 \star g^2) \subseteq \mathrm{supp}(g^1) + \mathrm{supp}(g^2) \subseteq Q + Q.$$

In order to come back to the unique representative inside of $\boldsymbol{L}^1(Q)$ we have to form

$$g^3 := [\sqcup\!\sqcup_\Lambda \star (g^1 \star g^2)] \cdot \mathbf{1}_Q$$

which turns out to be another interpretation of the usual *cyclic convolution.* **The remarkable property is the fact that it can be represented by a pointwise multiplication of Fourier coefficients!!**

BIBOLamper **Lemma 88.** *Two measures μ_1 and μ_2 have the same action on $\boldsymbol{C}_\Lambda$ if and only if*

$$\widehat{\mu_1}(\lambda^\perp) = \widehat{\mu_2}(\lambda^\perp) \quad \forall \lambda^\perp \in \Lambda^\perp.$$

Proof. We have

$$\mu_1 \star f_{per} = \mu_2 \star f_{per}, \quad \forall f_{per} \in \boldsymbol{C}_\Lambda,$$

but since this family coincides with the collection

$$f_{per} = f \star ⧢_\Lambda, \quad f \in \boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d)$$

we have

$$(\mu_1 \star ⧢_\Lambda) \star f = (\mu_2 \star ⧢_\Lambda) \star f, \quad \forall f \in \boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d).$$

But equality of systems on the dense subspace $\boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d) \subset \boldsymbol{C}_0(\mathbb{R}^d)$ implies equality of the functionals (now viewed as elements of $\boldsymbol{S}_0'(\mathbb{R}^d)$!)

$$\mu_1 \star ⧢_\Lambda = \mu_2 \star ⧢_\Lambda$$

or equivalently, by taking $\boldsymbol{S}_0'$-FTs:

$$\mathcal{F}(\mu_1 \star ⧢_\Lambda) = \mathcal{F}(\mu_2 \star ⧢_\Lambda), \tag{422}$$ mu1mu2hatA

but by the general Poisson formula we get

$$\widehat{\mu_1} \cdot ⧢_{\Lambda^\perp} = \widehat{\mu_2} \cdot ⧢_{\Lambda^\perp},$$

which in turn is equivalent to the statement

$$\widehat{\mu_1}(\lambda^\perp) = \widehat{\mu_2}(\lambda^\perp), \quad \forall \lambda^\perp \in \Lambda^\perp. \tag{423}$$ mu1mu2hat

□

Remark 49. To complete the picture one should also check that the cyclic convolution corresponds exactly to the *coordinate-wise* multiplication of the Fourier coefficients of the localized impulse response functions $g_Q \in \boldsymbol{L}^1([0,1])$. This can be verified by showing that it is valid for Dirac measures. Their Fourier coefficients are just the samples of their Fourier transforms, i.e. pure frequencies over $\mathbb{Z}$. But once we add $x+y$ in $\mathbb{R}$ and pass the value 1 in the exponent we can reduce everything modula Z (we should give a concrete example here, e.g. $z = 2/3$, $\delta_z \star_{cycl} \delta_z = \delta_{-z} = \delta_{1/3}$).

Remark 50. From a group theoretical point of view we may consider a Λ-periodic function as a function on the *quotient group* $\mathbb{R}^d/\Lambda$, of all equivalence classes of vectors in $\mathbb{R}^d$ modulo Λ, i.e. by the equivalence relation

$$\mathbf{x} \sim_\Lambda \mathbf{y} \quad \Leftrightarrow \quad \mathbf{x} - \mathbf{y} \in \Lambda, \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^d. \tag{424}$$ LamEquiv

It is clear that the equivalence classes are just the set $\mathbf{x} + \Lambda$ (since obviously $\mathbf{x} + \Lambda = \{\mathbf{x} + \lambda, \lambda \in \Lambda\}$ coincides with $\mathbf{y} + \Lambda$ if and only if $\mathbf{x} - yy \in \Lambda$ (hence $\mathbf{y} - \mathbf{x} \in \Lambda$. Let us write $[\mathbf{x}]_\Lambda$ or $[\mathbf{x}]_\Lambda$ respectively $[\mathbf{y}]_\Lambda$ for the corresponding equivalence classes.
On can go on and show that they form themselves a group by the group law:

$$[\mathbf{x}]_\Lambda [+] [\mathbf{y}]_\Lambda := [\mathbf{x} + \mathbf{y}]_\Lambda, \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^d. \tag{425}$$ Lamclmult

Of course one has to show that this is well defined and that for example the additive inverse to $[\mathbf{x}]_\Lambda$, usually written $-[\mathbf{x}]_\Lambda$ is given as $[-\mathbf{x}]_\Lambda$ and so on.
One then goes on to show that the dual group, i.e. all the characters of this new group, which is written as $H := \mathbb{R}^d/\Lambda$ can be realized as the subgroup of characters which are Λ-periodic, which is equivalent to the assumption that $s \in \Lambda^\perp$. (Lemma?)

Remark 51. Let us return to the more concrete situation of $\Lambda = \mathbb{Z} \lhd \mathbb{R}$ [49], where most readers will be familiar with addition modulo $\mathbb{Z}$ on $\mathbb{R}$. In fact, this is a built in MATLAB routine, which allows to define (a kind of cyclic) addition on $[0, 1)$, the representative system for $\mathbb{R}/Zst$.
It is however more convenient to find a group homomorphism from $\mathbb{R}$ to $\mathbb{U}$ which has exactly $\Lambda = \mathbb{Z}$ as the kernel of the homomorphism, i.e. with $\chi(t) = 1$ if and only if $t \in \mathbb{Z}$. The natural (and more or less only choice) is $\chi_1(t) = e^{2\pi i t}$ (we can call it the *wrap-around* function.
This function maps $\mathbb{R}$ onto the unit circle in the complex plain (with ordinary multiplication of complex numbers, they form obviously a group, with $\exp(-2\pi i t) = \overline{\exp(2\pi i t)}$ being the inverse.
It is easy to verify that it establishes a natural isomorphism between the "abstract quotient group" $\mathbb{R}/\mathbb{Z}$ and $\mathbb{U}$.
The dual group of $\mathbb{U}$ (compact) is just $\mathbb{Z}$ again, now seen as the set of all labels for our Fourier coefficients (written as complex exponentials $\chi_k(t) = \exp(2\pi i k t)$. That we have $\mathbb{Z}^\perp = \mathbb{Z}$ (hence also ${\mathbb{Z}^d}^\perp = \mathbb{Z}^d$ for $\mathbb{Z}^d \lhd \mathbb{R}^d$, $d \geq 1$) has already shown up on the proof of Poisson's formula, if we reconsider what has been done there.
For details concerning the abstract point of view in the context of LCA (locally compact Abelian) groups see also **re68:** [78] or **rest00:** [79]. The content of Lemma BIBOZd 87 is more or less described equivalently (but really in a very abstract way) in the form of so-called *Weil's formula* (for normal subgroups). It reads there as (written additively):

$$\int_{\mathscr{G}/\mathscr{H}} \left\{ \int_{\mathscr{H}} f(x+\xi) d\xi \right\} d\dot{x} = \int_{\mathscr{G}} f(x) dx, \quad f \in \boldsymbol{C}_c(\mathscr{G}). \tag{426}$$ WeilsChap333

which describes the connection between Haar measures (invariant measures, or linear functionals on $\boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathscr{G})$) on the group $\mathscr{G}$, the subgroup $\mathscr{H}$ and quotient group $\mathscr{G}/\mathscr{H}$ (with Haar measure $d\dot{x}$).
Here we have the case $\mathscr{G} = \mathbb{R}^d$, $\mathscr{H} = \Lambda$ and topologically as well as from the measure-theoretic viewpoint.
This formula is used in order to derive the surjectivity of the (non-expansive) mapping $T_{\mathscr{H}} : \left(\boldsymbol{L}^1(\mathscr{G}), \|\cdot\|_1\right) \to (\boldsymbol{L}^1(\mathscr{H}), \|\cdot\|_1)$, which is given by the formula (on our case):

$$T_{\mathscr{H}} f(\dot{x}) = \int_{\mathscr{H}} f(x+\xi) d\xi, \quad \text{with } \dot{x} = [x]_{\mathscr{H}} = x + H \in \mathscr{G}/\mathscr{H}.$$

The argument relies on the fact that one has

$$\|T_{\mathscr{H}} k\|_{\boldsymbol{L}^1} \leq \|k\|_{\boldsymbol{L}^1}, \quad k \in \boldsymbol{C}_c(\mathscr{G}),$$

and the surjectivity of this mapping from $\boldsymbol{C}_c(\mathscr{G})$ to $\boldsymbol{C}_c(\mathscr{H})$ (see Reiter's book [78], Chap.3.4.2, p.68). Reiter calls this definition an *abuse of notation* (because it disguises bit the identification of periodic functions with functions on the quotient group!). The surjectivity is then described in Chap.3.4.4, establishing the fact that $\boldsymbol{L}^1(\mathscr{G}/\mathscr{H})$ is a quotient space of $\boldsymbol{L}^1(\mathscr{G})$ (divided by the null-space of the *canonical mapping* $T_{\mathscr{H}}$).

[49] For people who prefer the description of characters via the simple exponential functions e^{ist}, $s \in \mathbb{R}^d$, the choice $\Lambda = 2\pi\mathbb{Z}$ would be natural.

55 The support and spectrum of a mild distribution

For ordinary functions (either $f \in \boldsymbol{C}_c(\mathbb{R}^d)$ or $f \in \boldsymbol{C}_b(\mathbb{R}^d)$) the support is defined as the closure of the set of "relevant positions", i.e.of the set $\{x \mid f(x) \neq 0\}$.
One may ask way the closure is taken. Also, one may ask about the support of a function in $\boldsymbol{L}^1(\mathbb{R}^d)$ or in $\boldsymbol{L}^2(\mathbb{R}^d)$, but one quickly recognizes that on cannot use the simple-minded concept for these *equivalence classes*, because changing the values in a countable set (certainly a set of measure zero!) would have a strong impact on the closure (just add a sequence of points in the unit cube which is dense there, and you would add the whole unit cube to the support defined otherwise). In short, there is no chance to use pointwise values for the definition of a support.
We also remember the usual description of the Dirac delta (measure), which is said to be equal to zero on $\mathbb{R} \setminus \{0\}$, but still having the property that $\int_{-\infty}^{\infty} \delta(t)dt = 1$. We will give this (contradictory) formulation a concrete meaning.
Finally, less us recall that there is a notion of the spectrum of a function in $h \in \boldsymbol{L}^\infty(\mathbb{R})$, defined in Chap. 7 of Reiter's book, entitled "The spectrum and its applications" (**re68:** [78], resp. **rest00:** [79]), but this has to be done indirectly via the cospectrum of the closed ideal in $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ which is annihilated by the convolution operator $f \mapsto h \star f$. But this is a rather indirect approach (which of course avoids talking about a generalized Fourier transform). Again, this approach gives for $h = \chi_s$ of course the spectrum (which is always a closed subset of the frequency domain $\widehat{\mathbb{R}^d}$) is just $\{s\}$ as expected, since the only "relevant frequency" for χ_s is the frequency $s \in \mathbb{R}^d$.
As we will see, the idea of defining the support of $f \in \boldsymbol{C}_b(\mathbb{R}^d)$ should extend the definition of the support generalized functions, in our case to $\boldsymbol{S}_0(\mathbb{R}^d)$ in a meaningful (and of course consistent) form. Moreover, we anticipate that $\operatorname{spec}(\sigma) = \operatorname{supp}(\widehat{\sigma})$. To get started we introduce the idea, that there are (open) regions, where the given distribution acts in a trivial way, i.e.puts every test function "living in that area" to zero. This suggests the following definition:
See the original LN, section 22, p.130, Def.57 and the subsequent session, including the *definition* $\operatorname{spec}(\sigma) := \operatorname{supp}(\widehat{\sigma})$, for $\sigma \in \boldsymbol{S}_0'$, and the verification that for the case of $h \in \boldsymbol{L}^\infty(\mathbb{R}^d)$ is equivalent to the indirect definition provided in Reiter's book. In fact, (p.139 of [78]) one defines for $\phi \in \boldsymbol{L}^\infty(\mathbb{R}^d)$:

$$I_\phi = \{f \mid f \in \boldsymbol{L}^1(\mathbb{R}^d), f^* \star \phi = 0\}$$

and defines

$$\operatorname{spec}_{Reiter}(\phi) := \mathbb{R}^d \setminus \operatorname{cosp}(I_\phi), \text{ with } \operatorname{cosp}(\phi) := \{s \mid \widehat{f}(s) = 0, \forall f \in I_\phi\}.$$

From a practical point of view one can *characterize* $\operatorname{spec}(\sigma)$ as the "relevant set for the Fourier transform".

suppSOP **Definition 19.** The support $\operatorname{supp}(\sigma)$ of $\sigma \in \boldsymbol{S}_0'$ is defined as the complement of the (open) area of *trivial action*. A point $x \in \mathbb{R}^d$ is said to belong the to area of trivial action, if there is a some open ball $B_\eta(x)$ (for some positive $\eta > 0$) such that

$$f \in \boldsymbol{S}_0(\mathbb{R}^d) \text{ with } \operatorname{supp}(f) \subset B_\eta(x) \quad \Rightarrow \quad \sigma(f) = 0.$$

Remark 52. It is to check that this area of trivial action is an open subset of $\mathbb{R}$ (simply because any $y \in B_{\eta/2}(x)$ the ball $B_{\eta/2}(y) \subset B_\eta(x)$). Hence the support (as defined above) is a *closed subset* of $\mathbb{R}^d$. This is illustrated by the following figure (55). The yellow domain is the circle around the center x, the point y is inside the ball with half radius, and the red ball (with half radius) is still completely inside of the yellow one!.

BxRbyRh.jpg

BxRbyRh.jpg

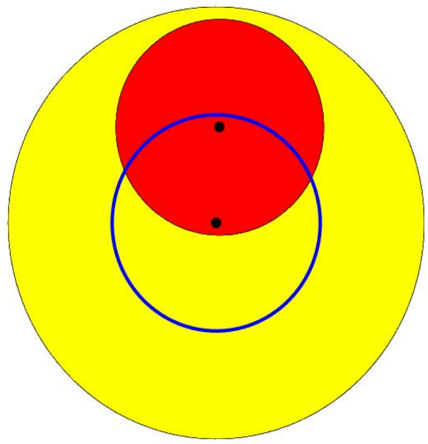

Abbildung 81: BxRbyRh.jpg

Exercise 9. Show that $\mathrm{supp}(\mu_k)$ as defined above coincides with $\mathrm{supp}(k)$ defined as usual (i.e.$\{x \,|\, k(x) \neq 0\}^-$), i.e. that the new definition is a *consistent* extension of the known concept. Similar arguments apply to σ_h for $h \in \boldsymbol{C}_b(\mathbb{R}^d)$.
In contrast, the naive definition of a support (based on function values) is not meaningful anymore in the context of $\boldsymbol{L}^1(\mathbb{R}^d)$ or $\boldsymbol{L}^2(\mathbb{R}^d)$!

We can give the following characterization of the spectrum of $\sigma \in \boldsymbol{S}_0(\mathbb{R}^d)$ (later we will show that it coincides with the concept of a spectrum as introduced in **re68:** [78] for $h \in \boldsymbol{L}^\infty(\mathbb{R}^d)$ $(\subset \boldsymbol{S}_0'(\mathbb{R}^d))$.

specfilt

Lemma 89. *A point $s \in \widehat{\mathbb{R}^d}$ belongs to* $\mathrm{spec}(\sigma)$ *if and only if there is a sequence of test functions $(h_n)_{n=1}^\infty$ in $\boldsymbol{S}_0(\mathbb{R}^d)$ such that*

$$\chi_s = w^*\text{-}\lim{}_{n\to\infty}\, h_n \star \sigma.$$

55.1 Mild distributions with discrete support

We want to discuss the fact that a mild distribution $\sigma \in \boldsymbol{S}_0'$ with $\mathrm{supp}(\sigma) \subset \Lambda \lhd \mathbb{R}^d$ is a weighted Dirac comb, more precisely, $\sigma = \sum_{\lambda\in\Lambda} c_\lambda \delta_\lambda$ with $(c_\lambda)_{\lambda\in\Lambda} \in \boldsymbol{\ell}^\infty(\Lambda)$.
Let us recall the definition of a *closed ideal* in a Banach algebra. We are mostly interested in commutative Banach algebra, where no distinction between left and right and two-sided ideals has to be made.

closedideal **Definition 20.** A closed subspace $\boldsymbol{I}$ of a Banach algebra $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ (with multiplication written as $\bullet$) is called a closed left (resp. twosided) *ideal if one has*

$$\boldsymbol{A} \bullet \boldsymbol{I} \subseteq \boldsymbol{I} \quad (\textit{and} \quad \boldsymbol{I} \bullet \boldsymbol{A} \subseteq \boldsymbol{I}).$$

Note that $\boldsymbol{I}$ inherits the norm from $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$, thus it is obvious that any such ideal is a closed subalgebra and hence a Banach algebra of its own right.
Let us first discuss the case of a single point. In the proof we will make use of the so-called *Gelfand-Mazur Theorem*, which can be found with a more or less complete proof in the German WIKIPEDIA page

`https://de.wikipedia.org/wiki/Satz_von_Gelfand-Mazur`

GelfandMazur **Lemma 90.** *Assume that a Banach algebra $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ over the field of complex numbers is a skew-field (also called a division-ring*[50]*, which means that every non-zero element has an inverse element!), then $\boldsymbol{A} = \mathbb{C}$ (as algebra, and hence as Banach algebra).*

maxideal **Corollary 13.** *For any $\mathbb{C}$-Banach algebra $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$, with multiplication written as $(f,g) \mapsto f \bullet g$, the closed maximal ideals $\boldsymbol{I}$ correspond to the null-spaces of the multiplicative linear functionals $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}}) \to \mathbb{C}$, i.e. those bounded and linear functionals σ on $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ which satisfy in addition to linearity also*

$$\sigma(f \bullet g) = \sigma(f) \cdot \sigma(g), \quad f, g \in \boldsymbol{A}. \tag{427}$$
multlinfunc

More precisely, for any such linear functional we have

$$\boldsymbol{I}_\sigma = \{f \in \boldsymbol{A} \,|\, \sigma(f) = 0\}, \tag{428}$$
maxidsigma

and conversely, given a maximal ideal I we can define $\sigma_{\boldsymbol{I}}$ by the identification described in Gelfand-Mazur as $j : \boldsymbol{A}/\boldsymbol{I} \to \mathbb{C}$:

$$\sigma_{\boldsymbol{I}}(f) = j(f + \boldsymbol{I}), quadf \in \boldsymbol{A}. \tag{429}$$
sigmaI

Of course the mappings $\boldsymbol{I} \to \sigma_{\boldsymbol{I}}$ and $\sigma \to \boldsymbol{I}_\sigma$ are inverse to each other.

[50] `https://en.wikipedia.org/wiki/Division_ring`, well described in the English version

I is an important fact concerning $\big(\boldsymbol{L}^1(\mathbb{R}^d),\, \|\cdot\|_1\big)$ that the *Gelfand space* of this Banach algebra (with respect to convolution), can be well identified with the *dual group*, i.e. the only multiplicative linear functionals are of the form $\sigma = \chi_s$. More precisely, we have:

multLFLi **Proposition 21.** *Any multiplicative linear functional σ on $\big(\boldsymbol{L}^1(\mathbb{R}^d),\, \|\cdot\|_1\big)$ is of the form:*

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

for a uniquely determined $s \in \widehat{\mathbb{R}^d}$.

Proof. The direct part is easy. By the convolution theorem (108) we have

$$\mathcal{F}(f \star g)(s) = \widehat{f}(s) \cdot \widehat{g}(s), \quad f, g \in \boldsymbol{L}^1(\mathbb{R}^d), s \in \widehat{\mathbb{R}^d}, \tag{430}$$ convLiRd

and since we have

$$|\widehat{f}(s)| \leq \|\widehat{f}\|_\infty \leq \|f\|_{\boldsymbol{L}^1}, \quad f \in \boldsymbol{L}^1(\mathbb{R}^d)$$

they are all linear functionals (of norm 1) on $\boldsymbol{L}^1(\mathbb{R}^d)$.

Now let us assume that we have such a multiplicative linear functional, call it $\sigma_0 \in (\boldsymbol{L}^1)^*(\mathbb{R}^d) \subset \boldsymbol{S}_0'(\mathbb{R}^d))$. As it is easy to show (see elsewhere) that

$$(\boldsymbol{L}^1)^*(\mathbb{R}^d) \star \boldsymbol{L}^1(\mathbb{R}^d) \subset \boldsymbol{C}_{ub}(\mathbb{R}^d), \tag{431}$$ LiPRdconv

together with the corresponding norm estimates

$$\|\sigma_0 \star f\|_\infty \leq \|\sigma\|_{(\boldsymbol{L}^1)^*(\mathbb{R}^d)} \cdot \|f\|_{\boldsymbol{L}^1(\mathbb{R}^d)}\,,$$

and that it can be represented in a pointwise sense using the standard formula (obvious for $f \in \boldsymbol{S}_0(\mathbb{R}^d)$):

$$[\sigma_0 \star f](y) = \sigma_0(T_y f^\checkmark), \quad y \in \mathbb{R}^d, f \in \boldsymbol{L}^1(\mathbb{R}^d).$$

Let us define $h := \sigma_0 \star k^\checkmark \in \boldsymbol{C}_{ub}(\mathbb{R}^d)$. In fact, this function has even derivatives of any order which are all bounded.

Upon replacing k (if necessary) by λk for a suitable number $\lambda \in \mathbb{C}$ we can assume that $\sigma(k) = 1$. But then we have

$$(\sigma_0 \star k^\checkmark)(f) = \sigma_0(k \star f) = \sigma_0(f)\sigma_0(k) = \sigma_0(f), \quad f \in \boldsymbol{L}^1(\mathbb{R}^d),$$

$$\text{or in short:} \quad \sigma_0 = \sigma_0 \star k = \sigma_h,$$

for some smooth and bounded function h.

The multiplicativity (427) can now be used (standard limiting argument now) to derive, for any pair $x, y \in \mathbb{R}^d$, by choosing Dirac sequences $f = T_x \mathrm{St}_\rho g_0$ and $g = T_y \mathrm{St}_\rho g_0$ and taking suitable limits that one derives from (427) the multiplicative law (another form of the exponential law),

$$h(x+y) = h(x) \cdot h(y), \quad x, y \in \mathbb{R}^d. \tag{432}$$ charlaw

We have $|h(x)| = 1$ because otherwise the sequence $|h(nx)| = |h(x)|^n, n \in \mathbb{Z}$ would be unbounded! Thus h is a *character* (or a multiplicative group-homomorphism) from $(\mathbb{R}^d, +)$ into $(\mathbb{U}, \cdot)$, and they are all of the form $h = \chi_s, s \in \widehat{\mathbb{R}^d}$, thus finishing the argument. The uniqueness of this choice is obvious (but should be checked by the reader). $\square$

We do not go into the details of the following observation: By characterizing the set of all multiplicative linear functionals for $\boldsymbol{L}^1(\mathbb{R}^d)$ with the pure frequencies, which are at the same time the joint eigen-vectors of BIBOs systems arising from convolution operators with $\boldsymbol{L}^1(\mathbb{R}^d)$ (this a very general form of diagonalization via the Fourier transform) on obtains another, natural topology on the set of multiplicative linear functionals.
In fact, if we take the set

$$MulSOP := \{\sigma \mid \sigma \in \boldsymbol{S}_0', \sigma(f \star g) = \sigma(f)\sigma(g), f, g \in \boldsymbol{S}_0(\mathbb{R}^d)\}$$

then we can show that $MulSOP$ is a (norm) bounded and w^*-closed subset of $\boldsymbol{S}_0'(\mathbb{R}^d)$, and that it coincides with

$$\{0\} \cup \{\chi_s, s \in \widehat{\mathbb{R}^d}\}.$$

Due to the *Banach-Alaoglu Theorem* (certainly the most important reason for working with the w^*-topology) this set is compact with respect to the w^*-convergence. In our case this means: for every sequence in such a set there is a sub-sequence which is w^*-convergent to an element of the same set (in fact, either we have $\chi_{s_n} \to_{w^*} \chi_{s_0}$ for $s_n \to s_0$, or $\chi_{s_n} \to_{w^*} 0$ for $s_n \to \pm\infty$).
Of we remove from any compact set a single point (here we remove $\{0\}$, the rest remains a locally compact space. Obviously the characters form a group, with the trivial character $h_0(t) = 1, \forall t \in \mathbb{R}^d$ and inverse element $h^{-1}(t) = 1/h(t) = \overline{h(t)}, t \in \mathbb{R}^d$. This is the construction of the dual group (another LCA group) for the general context.
Of course this abstract viewpoint is compatible with the natural one used so far, where we have been talking about ordinary convergence of sequences (s_n) in $\mathbb{R}^d$, and where we have used already to fact that it implies (and is in fact equivalent) with uniform convergence over compact sets. But let us still formulate it in an intuitive way (without proof):

GelftopRd **Lemma 91.** *A sequence of linear functionals χ_{s_n} is w^* convergent to χ_{s_0}, or explicitly*

$$\widehat{f}(s_n) \to \widehat{f}(s_0), \quad \forall f \in \boldsymbol{L}^1(\mathbb{R}^d), \tag{433}$$ FTpointwconv

if and only if one has for any ball $B_R(0)$ $(R > 0)$ uniform convergence, i.e. given $\varepsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that $n \geq n_0$ implies

$$|\chi_{s_n}(t) - \chi_{s_0}(t)| \leq \varepsilon, \quad \forall |x| \leq R.$$

On the other hand, we have for unbounded sequences, that there exists a subsequence such that $|s_n| \to \infty$ and in this case we have $\chi_{s_n} \to_{w^*} 0$ (and also conversely).

The key step towards the proof that any $\sigma \in \boldsymbol{S}_0'(\mathbb{R}^d)$ with $\mathrm{supp}(\sigma) = \{0\}$ is of the form $\sigma = \lambda\delta_0$ for some $\lambda \in \mathbb{C}$, is the following observation:

bdapprIzero **Lemma 92.** *The closed ideal in $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$*

$$I_0 = \{f \in \boldsymbol{L}^1(\mathbb{R}^d) \mid \widehat{f}(0) = 0\},$$

is a Banach algebra (for convolution) with bounded approximate units (of norm 2).

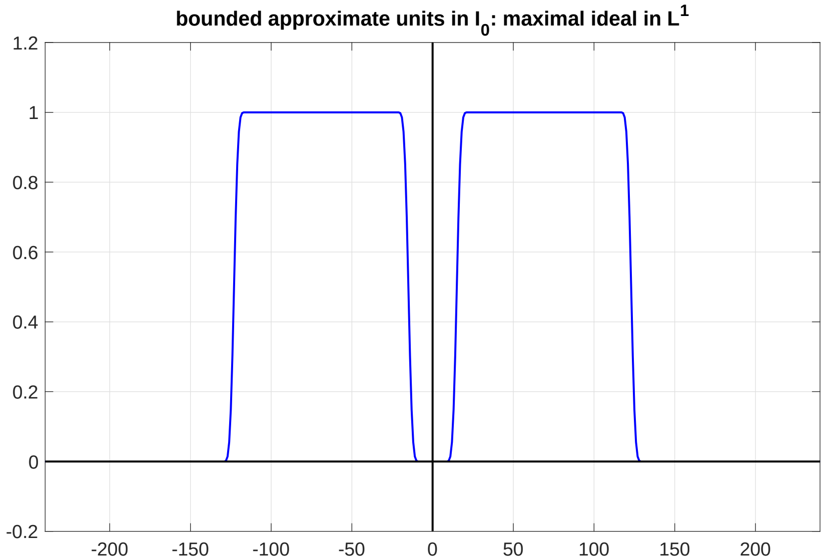

Abbildung 82: approxunitIz1.eps

Proof. We know that $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$, as Banach algebra with respect to convolution, has bounded approximate units. In fact, given f and $\varepsilon > 0$, we can take obtain a Dirac sequence from ANY $g \in \boldsymbol{L}^1(\mathbb{R}^d)$ with $\widehat{g}(0) = 1$: $(\mathrm{St}_\rho g)_{\rho\to 0}$, with $\|\mathrm{St}_\rho g\|_{\boldsymbol{L}^1} = \|g\|_{\boldsymbol{L}^1} = 1$ for all ρ, such that for $\rho \in (0, \rho_0]$ one has

$$\|\mathrm{St}_\rho g \star f - f\|_{\boldsymbol{L}^1} \le \varepsilon/2. \tag{434}$$

But of course these functions satisfy $\mathcal{F}(\mathrm{St}_\rho g)(0) = \mathrm{D}_\rho \widehat{g}(0) = 1$ for any ρ and thus they do not belong to I_0.
But f satisfies $\widehat{f}(0) = 0$ and thus we can find some $\rho_1 \le \rho_0$ such that we also have:

$$\|\,\mathrm{St}_{1/\rho}\, g \star f\|_{\boldsymbol{L}^1} = \|\,\mathrm{St}_{1/\rho}(g \star \mathrm{St}_\rho f)\|_{\boldsymbol{L}^1} = \|g \star \mathrm{St}_\rho f\|_{\boldsymbol{L}^1} = \|g \star \mathrm{St}_\rho f - \widehat{f}(0) g\|_{\boldsymbol{L}^1} \le \varepsilon/2. \tag{435}$$

Obviously we have now for $f_\rho := \mathrm{St}_\rho g_0 - \mathrm{St}_{1\rho}\, g_0$ the norm estimate $\|f_\rho\|_{\boldsymbol{L}^1} \le 2$, but also $f_\rho \in I_0$ (please check yourself!). Finally the combination of (434) with (435) ensures the validity of the required estimate

$$\|f - f_\rho \star f\|_{\boldsymbol{L}^1} \le \varepsilon.$$

□

The freedom of choosing $g \in \boldsymbol{L}^1(\mathbb{R}^d)$ is a valuable feature of proposition 92. So we can choose for g the so-called De LaVallee-Pousson kernels, which is nothing else but trapezoidal functions on the Fourier transform side. The approximate units arising in this way are then the difference of a widely stretched trapezoidal functions (arising from the Dirac sequence in the time domain) minus a compressed (in the sense of D_ρ) version, i.e. by subtracting a narrow trapezoidal function. Thus overall $\widehat{f_\rho}$ becomes in this case as depicted in Fig. 55.1.
We thus conclude:

apprbdlIzore **Lemma 93.** *For any $f \in I_0$ (as in Lemma* bdapprIzero *92) and $\varepsilon > 0$ there exist and approximation (of the form $f \star f_\rho$) such that* $\mathrm{spec}(f \star f_\rho)$ *is a compact set disjoint from $\{0\}$, i.e. $f \star f_\rho$ is a bandlimited element in $\boldsymbol{S}_0(\mathbb{R}^d)$ with a Fourier transform vanishing near 0.*

Proof. For f_ρ described as above we have $\widehat{f_\rho}(t) = 0$ near zeros (the difference of two trapezoidal functions of equal height near 0), and also the compactness of the support of the trapezoidal functions guarantees the band-limitedness. □

LET US NOT FORGET TO EXPLAIN THE IDEAL THEOREM FOR SEGAL ALGEBRAS!

56 Good properties of "nice functions" from $\boldsymbol{S}_0(\mathbb{R}^d)$

1. We have seen, that the Fourier Inversion Theorem applies, even with the help of Riemann integrals only, since $\mathcal{F}(\boldsymbol{S}_0(\mathbb{R}^d)) = \boldsymbol{S}_0(\mathbb{R}^d) \hookrightarrow \boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d)$;

2. We have the validity of Poisson's formula, which relies on the fact that $\mathcal{F}(\boldsymbol{S}_0(\mathbb{R}^d)) = \boldsymbol{S}_0(\mathbb{R}^d)$ guarantees that the restriction to any lattice $\Lambda \lhd \mathbb{R}^d$ gives a sequence in $\boldsymbol{\ell}^1(\Lambda)$. This can be used to show that the Λ-periodized version of $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ has an absolutely convergent Fourier series representation, which in turn implies that the relationship is valid pointwise. So the particular choice $t = 0$ is valid, but $f_{per}(0) = \sum_{k\in\mathbb{Z}} f(k)$ (for the case $\Lambda = \mathbb{Z} \lhd \mathbb{R}$), just as an illustration. But the Fourier series at zero is just the sum of the Fourier coefficients (because $e^0 = 1$), and this explains the right hand side of the Poisson formula.

   The Poisson formula is also valid for $f \in \boldsymbol{W}(\mathbb{R}^d)$ with $\widehat{f} \in \boldsymbol{W}(\mathbb{R}^d)$ but the only natural argument for checking that $\widehat{f} \in \boldsymbol{W}(\mathbb{R}^d)$ is to verify that $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ implies $\widehat{f} \in \boldsymbol{S}_0(\mathbb{R}^d) \subset \boldsymbol{W}(\mathbb{R}^d)$. But strictly logical this other space is (slightly) larger than $\boldsymbol{S}_0(\mathbb{R}^d)$.

3. As it has been shown in the work of F. Weisz, all the classical summability kernels belong to $\boldsymbol{S}_0(\mathbb{R}^d)$. Let us discuss the

   Let us recall the FRFT (just switching side)

$$\int_{\mathbb{R}^d} f(x)\hat{g}(x)dx = \int_{\mathbb{R}^d} \hat{f}(y)g(y)dy \quad f, g \in \boldsymbol{L}^1(\mathbb{R}^d). \tag{436}$$ fundamFour02b

   The inversion theorem has been established by choosing g as $M_t \mathrm{D}_\rho h$, because this means that it is the Fourier transform of some Dirac family at centered at t, so that we can pick out the value $f(t)$ in the limit, as $\delta_x(f)$.

   In fact, what we need is just the condition $\hat{h}(0) = \int_{\mathbb{R}^d} h(y)dy = 1$ (so for non-negative functions $h \in \boldsymbol{L}^1(\mathbb{R}^d)$ this means $\|h\|_{\boldsymbol{L}^1} = 1$). It will be convenient to consider $h = h^{\checkmark}$, so that both h and $\hat{h}$ are symmetrically concentrated around 0. and to use for

$$\mathcal{F}^{-1}(M_t \mathrm{D}_\rho h) = T_t \mathrm{St}_\rho \hat{h}.$$

   If we choose $\hat{h} = g = g^{\checkmark} \in \boldsymbol{S}_0(\mathbb{R}^d)$ we also have $h = \mathcal{F}^{-1}(g) = (\mathcal{F}(h))^{\checkmark} = g^{\checkmark} = g$, and we come up with the inversion formula for $\widehat{f} \cdot \hat{h}$, because the expression on the right hand side is just

$$\int_{\mathbb{R}^d} \widehat{f}(s)\hat{h}(s)e^{2\pi ist}ds.$$

This is now a well defined element of $\boldsymbol{S}_0(\mathbb{R}^d)$ (hence Riemann integrable) because $\boldsymbol{L}^1(\mathbb{R}^d) \star \boldsymbol{S}_0(\mathbb{R}^d) \subseteq \boldsymbol{S}_0(\mathbb{R}^d)$, or equivalently $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) \cdot \boldsymbol{S}_0(\mathbb{R}^d) \subseteq \boldsymbol{S}_0(\mathbb{R}^d)$!

So the function $\hat{h}$ has the role of a (mild) damping factor, which ensures that the right hand side of the inversion formula is valid (right now we keep ρ fixed and do not care about the limit for $\rho \to 0$!).

Since we look into $\widehat{f} \cdot \hat{h}$ on the Fourier transform side, we look correspondingly into $f \star h$ (resp. $f \star \mathrm{St}_\rho h$, for fixed ρ) on the time side. Again, this is a decent function in $\boldsymbol{L}^1(\mathbb{R}^d) \star \boldsymbol{S}_0(\mathbb{R}^d) \subseteq \boldsymbol{S}_0(\mathbb{R}^d)$ and thus a continuous and integrable function.

Summability kernels are often chosen to be positive (there is a big collection of them). The Fejer kernel is the most famous one, corresponding to the triangular function Δ on the Fourier transform side , and this is the one used in our considerations for the proof of the inversion theorem. It corresponds to SINC^2 on the time-side, which is non-negative. This has some additional advantages. Let us do the details outside of this listing.

Lemma 94. *Assume that you have some function $f \in \boldsymbol{L}^1(\mathbb{R}^d)$ (or $\boldsymbol{L}^1(\mathbb{R}^d) \cap \boldsymbol{C}_0(\mathbb{R}^d)$) which takes only real values between A and $B > A$ (e.g. only values in $[0, 1]$), then one has for any non-negative function $h \in \boldsymbol{L}^1(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d} h(y)dy = \hat{h}(0) = \|h\|_{\boldsymbol{L}^1} = 1$:*

$$A \leq f \star h(z) \leq B, \quad \forall z \in \mathbb{R}^d.$$

Proof. First note that non-negativity implies that

$$\|h\|_{\boldsymbol{L}^1} = \int_{\mathbb{R}^d} |h(z)|dz = \int_{\mathbb{R}^d} h(z)dz = \hat{h}(0),$$

but we have

$$|h(z)| \leq \|\hat{h}\|_\infty \leq \|h\|_{\boldsymbol{L}^1}, \quad z \in \widehat{\mathbb{R}^d},$$

showing that $\hat{h}$ takes its maximum at $z = 0$.
Since we have under the given conditions $h \star \mathbf{1} = \mathbf{1}$ restrict our attention to the case $0 \leq f(z) \leq B - A$ (just replace f by $f - A\mathbf{1}$, the convolutions still exist!).
Now clearly the convolution still is non-negative now (since only positive values appear in the integral) and we have the estimate

$$f \star h(z) = \int_{\mathbb{R}^d} f(x-y)h(y)dy \leq \|f\|_\infty \int_{\mathbb{R}^d} h(y)dy = B - A,$$

and the proof is complete. □

This is in contrast to the "ideal filter", which corresponds to pointwise multiplication by a (dilated) version of the boxcar function $\mathbf{1}_{[-1/2,1/2]}$, or convolution by the SINC-function on the time side. As it is known this function shows poor decay and has of course negative values.
As a consequence the ideal filter, applied to $f = \mathbf{1}_{[-1/2,1/2]}$ gives some overshot and undershoot (more or less independently in terms of hight, which is ca. 9 percent of the height of the jump. Thus the range of an "ideally filtered" boxcar function is around -0.09 to 1.09.
We can easily illustrate this using MATLAB

56.1 Gabor atoms in $S_0(\mathbb{R}^d)$ and fine lattices

Here we will about Gabor representations. While both the lattice and the building block had to satisfy certain conditions it is now possible to freely choose $g \in S_0(\mathbb{R}^d)$ (of course non-zero).

goodGabor **Theorem 24.** *Given $g \in S_0(\mathbb{R}^d)$ there exists some $\delta_0 > 0$ such that for any δ_0-dense lattice $\Lambda \lhd \mathbb{R}^d \times \widehat{\mathbb{R}}^d$*[51] *any $f \in S_0(\mathbb{R}^d)$ can be represented as an absolutely convergent Gabor sum of the form*

$$f = \sum_{\lambda\in\Lambda} c_\lambda \pi(\lambda) g, \quad \textit{with} \quad \|(c_\lambda)\|_{\ell^1(\Lambda)} \leq C\|f\|_{S_0}. \tag{437}$$ GabS0atom

In fact, one can choose the mapping $f \to (c_\lambda)_{\lambda\in\Lambda}$ to be linear, and being given as the samples of the STFT of f at the lattice Λ, with another window $\widetilde{g} \in S_0(\mathbb{R}^d)$.

Proof. We will not give a detailed proof here, but rather an indication of the idea of the proof. The starting point is the reconstruction formula which allows to reconstruct $f \in S_0(\mathbb{R}^d) \subset L^2(\mathbb{R}^d)$ from $V_g f$ via the adjoint mapping V_g^*. This was based on some adjointness relation and the fact that $\|g\|_{L^2} = 1$ (which we will assume from now on) implies that we have the important relation

$$f = V_g^*(V_g f) = \int_{\mathbb{R}^d} V_g f(z)\pi(z) g dz = \int_{\mathbb{R}^d} \pi(z) g d\mu_F(z), \tag{438}$$ STFrecon02

with the usual convention of writing for $F := V_g f \in S_0(\mathbb{R}^d \times \widehat{\mathbb{R}}^d) \subset L^1(\mathbb{R}^d \times \widehat{\mathbb{R}}^d)$ and μ_F for the corresponding *bounded measure* on $\mathbb{R}^d \times \widehat{\mathbb{R}}^d$.

It is thus plausible that we could replace this bounded measure by corresponding discrete measures, which could be concentrated on sufficiently dense lattice. *This is in fact a good heuristic which can be turned into a clean mathematical argument.*

It may not be surprising to observe that we have (even in the norm of $\big(S_0(\mathbb{R}^d), \|\cdot\|_{S_0}\big)$, so that we obtain, choosing regular BUPUs $(\psi_\lambda) = (T_\lambda \psi)$ on $\mathbb{R}^d \times \widehat{\mathbb{R}}^d$:

$$f = \lim_{\mathrm{diam}(\Psi)\to 0} \int_{\mathbb{R}^d} \pi(z) g \; d\mathrm{D}_\Psi \mu_F(z) = \sum_{\lambda\in\Lambda} \mu_F(\psi_\lambda)\pi(\lambda) g. \tag{439}$$ reconGab01

The mapping $f \mapsto V_g f \mapsto \int_{\mathbb{R}^d} V_g f(z)\psi(z-\lambda)$ is of course linear, and since we know that $f \mapsto V_g f|_\Lambda$ is bounded into $(\ell^1(\Lambda), \|\cdot\|_{\ell^1(\Lambda)})$ it is altogether bounded linear mapping.

The key issue if the formal part of the proof is however that the speed of convergence does NOT depend on the individual f (each time we look at $V_g f$ for a different f we have a new function, and the speed of convergence ($\mathrm{diam}(\Psi) \to 0$) might (!) depend on f, but it is the "good quality of g" (the assumption $g \in S_0(\mathbb{R}^d)$ excludes unpleasant or pathological case!) that allows to prove that one has approximation in the operator norm sense, i.e. one can find $\delta_0 > 0$ such that one has for some $\gamma < 1$:

$$\|f - \sum_{\lambda\in\Lambda} \mu_F(\psi_\lambda)\pi(\lambda) g\|_{S_0} \leq \gamma \|f\|_{S_0}, \quad f \in S_0. \tag{440}$$ approxfinGab

[51] This means of course that $\bigcup_{\lambda\in\Lambda} B_{\delta_0}(\lambda) = \mathbb{R}^d \times \widehat{\mathbb{R}}^d$.

In other words we have two operators, the identity operator and the approximation operator $\boldsymbol{A} : f \mapsto \sum_{\lambda\in\Lambda} \mu_F(\psi_\lambda)\pi(\lambda)g$ which are close enough to guarantee the invertibility of the operator $\boldsymbol{A}$, because the operator norm of the difference is small enough.
In fact we can apply just $\boldsymbol{A}$ in order to get a first approximation f_1 to our f. Then we can compute the sampling values of $V_g(f_1)$ at Λ in order to compare them with the given sampling values of $V_g f$. The will not be exact, but the correction term will get smaller and smaller after each **iteration**. Details to be given later. But it turns out that this value $\widehat{\psi}(0)$ depends not on the particular BUPU, but only on the lattice, in fact only on the size of the fundamental domain of Λ, resp. the redundancy $n/(a*b)$ for the simple, regular case ($1/ab$ for the continuous case).

$$f \approx C_\Lambda \sum_{\lambda\in\Lambda} V_g f(\lambda)\pi(\lambda)g. \tag{441}$$ reconGab02

This last expression is called the *frame operator* for the Gabor family $(g_\lambda)_{\lambda\in\Lambda}$. More precisely, we should call it the *normalized frame operator* because, following the current standard literature the Gabor frame operator for the Gabor family $(g_\lambda)_{\lambda\in\Lambda}$ is usually denoted by $S_{(g,\Lambda)}$ OR $S_{g,\Lambda}$ is defined as

$$S_{g,\Lambda}(f) := \sum_{\lambda\in\Lambda} V_g f(\lambda)\pi(\lambda)g = \sum_{\lambda\in\Lambda} P_\lambda(f). \tag{442}$$ SgLamdef

In some sense it is better to introduce the local projection operators

$$P_\lambda : f \mapsto \langle f, g_\lambda\rangle g_\lambda, \quad \lambda\in\Lambda,$$

and thus rewrite the reconstruction formula as

$$f = \int_{\mathbb{R}^d\times\widehat{\mathbb{R}}^d} V_g f(\lambda)\pi(\lambda)g d\lambda = \int_{\mathbb{R}^d\times\widehat{\mathbb{R}}^d} P_\lambda(f) d\lambda, \tag{443}$$ reconstrPlam

or even more compact (and perhaps more abstract, at the level of operators)

$$\mathrm{Id}_{\boldsymbol{L}^2} = \int_{\mathbb{R}^d\times\widehat{\mathbb{R}}^d} P_\lambda \, d\lambda. \tag{444}$$ reconstrPlam

Such a formula is often described as a *continuous resolution of the identity*. The family $(g_\lambda)_{\lambda\in\mathbb{R}^d\times\widehat{\mathbb{R}}^d}$ is then called the *family of coherent states* (see the work of J. Klauder, J.P. Gazeau, J.P. Antoine and others).
In fact, there is not such a big trouble with respect to the highly oscillating phase factors, and thus the frame operator can be (correctly) be seen as a Riemannian sum for this last integral representation, in other words, it makes use of the relation

$$\mathrm{Id}_{\boldsymbol{L}^2} = \int_{\mathbb{R}^d\times\widehat{\mathbb{R}}^d} P_\lambda \, d\lambda \, \approx \, C_\Lambda \sum_{\lambda\in\Lambda} P_\lambda \tag{445}$$ reconstrPLAM

where the natural choice for C_Λ is the size (the volume) of the (any) fundamental domain for the lattice Λ, or $C_\Lambda = |\det(\boldsymbol{A})|$ for $\Lambda = \boldsymbol{A}\star\mathbb{Z}^d$, with $\det(\boldsymbol{A}) \neq 0$. Again this confirms that this factor tends to zero for $\Lambda \to \mathbb{R}^d \times \widehat{\mathbb{R}}^d$ (meaning that the density of the lattice is increasing, or that corresponding $BUPUs$ $(\psi_\lambda)_{\lambda\in\Lambda}$ can be chosen in a way such that one has $\mathrm{diam}(\Psi) \to 0$). $\square$

Just recall: a matrix is the sum of its entries, but the entry $a_{3,5}$ corresponds to the contribution of the rank one operator $(\mathbf{e}_3' \star \mathbf{e}5)$ to the overall matrix. In other words, an $n \times n$ matrix $\boldsymbol{A}$ is the sum of n^2 rank one operators. Only a diagonal
matrix can be written as a sum of projection operators onto the unit vectors. In a mixed MATLAB notation we have diag$([1,2,3])$ is just a linear combination of $\sum_{k=1}^{3} kP_k$, where P_k is the projection on the k-th unit vector in $\mathbb{R}^3$.

Remark 53. The fact that both g and $\widetilde{g}$ belong to $\boldsymbol{S}_0(\mathbb{R}^d)$ guarantees that one can extend the mapping $f \to (c_\lambda)$ from $\boldsymbol{S}_0(\mathbb{R}^d)$ to $\boldsymbol{L}^2(\mathbb{R}^d)$ and thus create a bounded analysis mapping into $(\boldsymbol{\ell}^2(\Lambda), \|\cdot\|_{\boldsymbol{\ell}^2(\Lambda)})$. Also the synthesis mapping $(c_\lambda) \to \sum_{\lambda \in \Lambda} c_\lambda \pi(\lambda)(\widetilde{g})$ is bounded back from $\boldsymbol{\ell}^2(\Lambda)$ into $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$. *Such a statement would not be possible for general elements g or $\widetilde{g}$ in $\boldsymbol{L}^2(\mathbb{R}^d)$ only!*
It is even possible to extend these mappings (in a w^*-w^*-continuous manner) to mappings from $\boldsymbol{S}_0'(\mathbb{R}^d)$ into $\boldsymbol{\ell}^\infty(\Lambda)$ and back! We will discuss this in more detail using the terminology of Banach Gelfand Triples.

Looking again into the approximate reconstruction formula

$$f \approx \sum_{\lambda \in \Lambda} \mu_F(\psi_\lambda)\pi(\lambda)g \tag{446}$$

reconGab01b

we observe that the actual computation of

$$\mu_F(\psi_\lambda) = \int_{\mathbb{R}^d} \psi(y-\lambda)V_g(f)(y)dy$$

is computational expensive and also it requires the full information about the STFT $V_g f$. However it is much easier to compute an even more simple expression, derived from the assumption that one has

$$\int_{\mathbb{R}^d} \psi(y-\lambda)V_g(f)(y)dy \approx V_g f(\lambda) \int_{\mathbb{R}^d} \psi(y-\lambda)dy = V_g f(\lambda) \int_{\mathbb{R}^d} \psi(y)dy = V_g f(\lambda)\widehat{\psi}(0). \tag{447}$$

simplBUPint

Summing now again over Λ we can hope that

Remark 54. A practical experiment even reveals that the high oscillations of the STFT (recall that $V_g f$ is a complex-valued function, which appears to have large oscillations, and thus a mean value taken naively leads to a significant amount of cancellations, as shown for the case of the lattice with $a = 4 = b$ for $n = 480$ in the experiment below.

ighfreq1.jpg

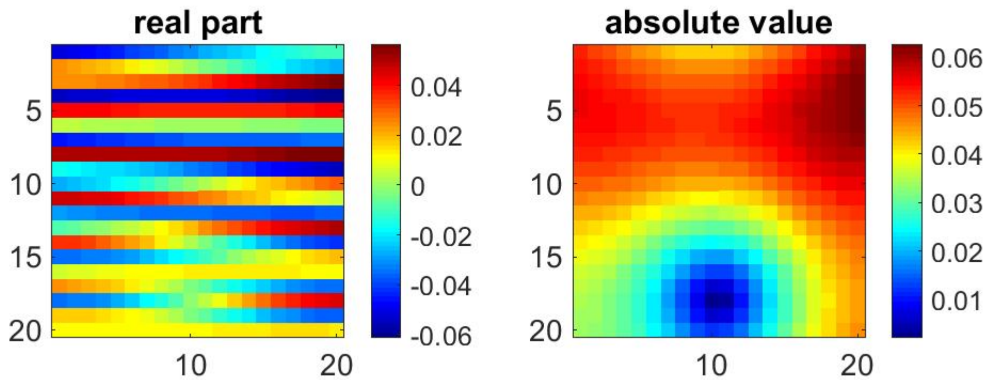

Abbildung 83: STFThighfreq1.jpg: This image shows the high oscillations which occur in the real and the imaginary part of the STFT.

SUPPORT001.tex !30.

57 Support and Spectrum of Distributions

So far we have tried to avoid a detailed discussion of the concept of a support of a generalized function, and even the support of a measure has not been defined.
First of all we have to extend the notion of a support which is clear for continuous functions (e.g. $k \in \boldsymbol{C}_c(\mathbb{R}^d)$) to generalized functions. Clearly this has to be done through the action of a distribution on the corresponding test functions. It is also clear that both for the case of $\boldsymbol{C}_0(\mathbb{R}^d)$ and $\boldsymbol{S}_0(\mathbb{R}^d)$ one has *regular* function algebras, i.e.there are functions with arbitrary small support at any position. The approach to the notion of support (as a set of relevant points) is through its complement, namely the (open!) set of irrelevant points:

loctrivdist **Definition 21.** A mild distribution $\sigma \in \boldsymbol{S}_0'(\mathbb{R}^d)$ is *acting trivially at a given point* $t \in \mathbb{R}^d$ if one has for some $k \in \boldsymbol{S}_0(\mathbb{R}^d)$, with $k(t) = 1$: $\sigma \cdot k = 0$, or $\sigma(k \cdot f) = 0, \forall f \in \boldsymbol{S}_0(\mathbb{R}^d)$.

Remark 55. It is important to note that the definition does not depend on the exact space. Because any $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ can be approximated (in the norm of $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$) by elements from $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) \cap \boldsymbol{C}_c(\mathbb{R}^d)$ (e.g. by convolving with some Dirac-kernel with compact support in $\boldsymbol{S}_0(\mathbb{R}^d) \subset \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$) one has for a bounded measure $\mu \in \boldsymbol{M}_b(\mathbb{R}^d) \subset \boldsymbol{S}_0'(\mathbb{R}^d)$: μ is acting trivially near $t \in \mathbb{R}^d$ if and only if there exists $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ with $k(x) = 1$ and $\sigma(k \cdot f) = 0, \forall f \in \boldsymbol{C}_0(\mathbb{R}^d)$.

eanval01.jpg

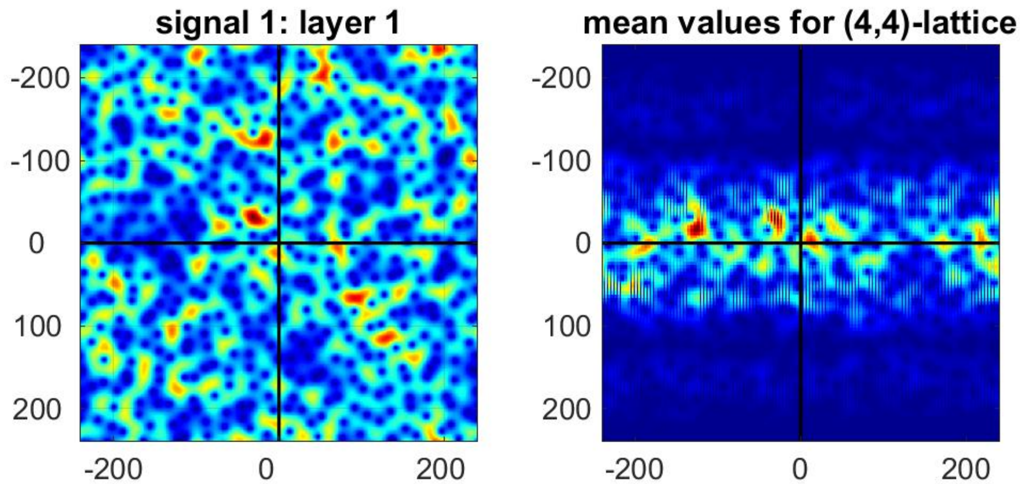

Abbildung 84: STxxmeanval01.jpg: showing strong oscillations!

The following, slightly more involved formulation may be more intuitive, but perhaps less practical for use (because it talks about a lot of functions instead of a single function with specific properties).

cospalt1 **Lemma 95.** *A point t does NOT belong to the support of σ (resp. σ shows "trivial action near t) if there exists some neighborhood U of the point t, e.g. some open ball $U = B_\varepsilon(t)$ such that $\sigma(f) = 0$ for all $f \in \boldsymbol{C}_c(\mathbb{R}^d) \cap \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ with $\operatorname{supp}(f) \subseteq U$.*

Proof. Assume that $\sigma \cdot k = 0$, then it is clear that $0 = (\sigma \cdot k)(f) = \sigma(k \cdot f)$ for any $f \in \boldsymbol{C}_0(\mathbb{R}^d)$. But by continuity of k (and in the case of the space $\boldsymbol{S}_0(\mathbb{R}^d)$ the much deeper result, the so-called Wiener's inversion theorem provides a similar fact, see [79]) one can find some function $\phi \in \boldsymbol{C}_c(\mathbb{R}^d)$ respectively in $\boldsymbol{C}_c(\mathbb{R}^d) \cap \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$) with $k(t) \cdot \phi(t) \equiv 1$ on some small (closed) ball $U = B_\varepsilon(t)$. For $f \in \boldsymbol{C}_c(\mathbb{R}^d)$ with small support inside of U we have of course $f = k \cdot \phi \cdot f$, or $f = k \cdot (\phi \cdot f)$, with $\phi \cdot f \in \boldsymbol{C}_c(\mathbb{R}^d)$, and hence $\sigma(f) = 0$ by the observation at the beginning of the proof.
As for the converse.. One may decompose a measure into a sum of "small pieces". This is done early on for bounded measures, but it is also valid for $\sigma \in \boldsymbol{S}_0'(\mathbb{R}^d)$, using the BUPU characterization of $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$. Let us make sure that we have a fairly fine BUPU, consisting of plateau-functions, so each ψ_i which is used (with $\sum_{i\in I}\psi(x) \equiv 1$) has a plateau near its center which we denote by x_i. Without loss of generality we can assume that $x_1 = t$ and that $\operatorname{supp}(\psi_1) = 1$ (it would be enough to have it different from zero!). But then

$$[\sigma \cdot \psi_1](f) = \sigma(\psi_1 \cdot f) = 0,$$

by the assumption, since

$$\operatorname{supp}(f \cdot \psi_1) \subseteq \operatorname{supp}(\psi_1) \subseteq U.$$

This argument completes the proof. TO BE ADAPTED OR CHECKED! □

It is a good exercise to verify that the set of "points of trivial action" (e.g. for a given bounded measure) does *not depend* on the choice of the underlying algebra of test functions. Whether one assumes k with $k(t) = 1$ to be from $\boldsymbol{C}_c(\mathbb{R}^d)$ or from $\boldsymbol{S}_0(\mathbb{R}^d)$ does not play a role, since on can approximate one by the other (in the sup-norm). See above remark.

opentrivact1 **Lemma 96.** *The set of points for which a distribution $\sigma \in \boldsymbol{S}_0'(\mathbb{R}^d)$ (resp. : a bounded measure $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$) shows trivial action is an open set.*

Proof. For the case of bounded measures we can argue as follows: *Assume that t_0 allows for some $k \in \boldsymbol{S}_0(\mathbb{R}^d)$ (we may also assume by density) with compact support, such that $k_0(t_0) = 1$ and $\sigma \cdot k_0 = 0$. Clearly $|k_0(y)| > 0$ for t near t_0. Hence we find that $k := k_0/k_0(y)$ is now a function with $k(y) = 1$, but still $\sigma \cdot k = 0$.*
For the general case let us just make use of the alternative characterization of "trivial action" and refer the to Figure 57.1 for an illustration of this situation. □

delxBUPU1.eps

Now it is time to define the support of a distribution.

suppSOPRd **Definition 22.** The **support** $\mathrm{supp}(\sigma)$ of $\sigma \in \boldsymbol{S}_0'(\mathbb{R}^d)$ is defined as the complement of the set of points of trivial action.

As the complement of an open set it is clear that $\mathrm{supp}(\sigma)$ is a closed set. Since we have now for $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ several possible concepts of supports we have to show that there is consistency:

suppfctdist1 **Lemma 97.** *Given $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ we find that the classical support*[52]

$$\mathrm{supp}_{Classic}(k) := \{t \,|\, k(t) \neq 0\}^-$$

and the distributional support $\mathrm{supp}(\sigma_k)$ coincide. Furthermore, the notion of support does not depend on the algebra of test functions used.

Proof.

- Assume that $t \notin \mathrm{supp}(k)$ (classical). Since $\mathrm{supp}(k)$ is closed, there exists some neighborhood U of t such that $k(z) \equiv 0$ on U. Choosing now any $k \in \boldsymbol{C}_c(\mathbb{R}^d) \cap \boldsymbol{S}_0(\mathbb{R}^d)$ with $k_0(t) = 1$ and $\mathrm{supp}(k_0) \subseteq U$, hence $k \cdot k_0 = 0$, as a function and consequently as a generalized function.
- Conversely, assume that $t \in \mathrm{supp}(k)$, i.e. for *any* neighborhood U of t there exists some point z such that $k(z) \neq 0$. Choosing $h \in \boldsymbol{C}_c(\mathbb{R}^d)$ with small support around centered at z and very close (in the w^*–sense to δ_z we obtain a good approximation to δ_z, hence δ_t. (more details should be given eventually).

□

Note that it is *not meaningful* to define the support for $f \in \boldsymbol{L}^p(\mathbb{R}^d)$ in the "usual way", because their elements are just *equivalence classes*. Having e.g. the only representative for $k \in \boldsymbol{C}_c(G)$ with $Q := \mathrm{supp}(k)$ we could always change the values (but not the class in $\boldsymbol{L}^1(G)$, say) of k by putting them equal to 1 at all the rational points of some compact set K disjoint from Q. In

[52] The bar at the end denotes the *closure* of the underlying set.

this way the closure of $\{x \,|\, k(x) \neq 0\}$ would become $Q \cup K$, different from the "true support", namely Q.
A typical argument is: If we apply a measure μ to a function $f \in \boldsymbol{C}_b(\mathbb{R}^d)$ which is zero on $\operatorname{supp}(\mu)$ then the result should be (and is in fact) zero.
The analysis behind this claim (of course it is enough to verify the result for compactly supported functions $f \in \boldsymbol{C}_c(\mathbb{R}^d)$) is essentially based on the fact that it is possible to approximate any measure (in the w^* sense with respect to $\left(\boldsymbol{C}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{C}_0'}\right)$) by discrete measures *whose support* is contained in $\operatorname{supp}(\mu)$, i.e. no other Dirac measures than those from the set $\operatorname{supp}(\mu)$ are needed. This is plausible, but not at all trivial, because the standard approximation (e.g. using regular BUPUs) makes of course only use of "near-by Diracs" (this is an easy claim), but not strictly from within the support. So *some modification* is required to ensure this claim.

Since generalized functions $\sigma \in \boldsymbol{S}_0'(\mathbb{R}^d)$ have a Fourier transform the support of that Fourier transform (so-to-say the "relevant frequencies" contained in σ) is a meaningful concept, which deserves a separate name:

spectrumdef **Definition 23.** For any $\sigma \in \boldsymbol{S}_0'(\mathbb{R}^d)$ the *spectrum* of σ is defined as $\operatorname{supp}(\widehat{\sigma})$.

Of course it is of interest to characterize the spectrum of σ (in fact it is not easier or more complicated to characterize the spectrum of $h \in \boldsymbol{L}^\infty(\mathbb{R}^d)$!) in several other, alternative ways:

harspectrum1 **Lemma 98.** *A pure frequency χ_s is in the spectrum (more precisely $s \in \operatorname{spec}(\sigma)$) if and only if χ_s is in the w^*-closure of the* **linear span of the translates of** *σ within $\boldsymbol{S}_0'(\mathbb{R}^d)$. In words we could say, knowing that w^*-closure coincides with the vector space $\boldsymbol{S}_0(\mathbb{R}^d) \star \sigma \subset \boldsymbol{C}_{ub}(\mathbb{R}^d)$, that those frequency can be "approximately" (in the w^*-sense) be "filtered out of the given distribution σ by suitable input from $\boldsymbol{S}_0$ (or equivalently $\boldsymbol{L}^1(\mathbb{R}^d)$, due to density).*
Thus we can formulate the characterization as follows: $s \in \operatorname{spec}(\sigma)$ is equivalent to the existence of some net (g_α) in $\boldsymbol{L}^1(\mathbb{R}^d)$ (possibly unbounded, and it may even be required to come from the dense subspace $\boldsymbol{S}_0(\mathbb{R}^d)$) such that $g_\alpha \star \sigma \to \chi_s$ in the w^*-topology.

Even more: the w^*-closure of the set of all translates (which of course the same as the w^*-closure of all functions elements of the form $\nu \star \sigma$, with $\nu \in \boldsymbol{M}_d(\mathbb{R}^d)$ (discrete measures) or on the w^*-closure of the linear span of elements of the form $\boldsymbol{L}^1(\mathbb{R}^d) \star \sigma$, or even just the w^*-closure of the more decent (because continuous and bounded) functions of the form $\boldsymbol{S}_0(\mathbb{R}^d) \star \sigma$ is the same, due to mutual w^*-approximation.
Since $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ is separable one can even find a *sequence* with these properties, for every single $s \in \operatorname{supp} \widehat{\sigma}$.
First of all the idea of local (pointwise) inversion comes into play. This is trivial in the case of $k \in \boldsymbol{C}_c(\mathbb{R}^d)$, since the function $1/k(t)$ is obviously continuous in some neighborhood of t and can be extended, with small support. Similar arguments apply to functions having m-continuous derivatives on $\mathbb{R}$, using the quotient rule of differentiation.
For the case of $\boldsymbol{S}_0(\mathbb{R}^d)$ this is not trivial anymore and one has to resort to Wiener's inversion theorem, which is of a quite different structure and in a way much deeper. It tells us that for any given $h \in \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ with $h(t_0) \neq 0$ there exists some other function

$g \in \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ with $g(s) = 1/h(s)$ near t_0, and since $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ is a pointwise algebra the product $h \cdot g \equiv 1$ near t_0.
Alternative approach: Assume that $\sigma \cdot k = 0$ for all functions with sufficiently small support! Then one can smooth them out (resp. integrate or convolve) and get that there are also some functions having small plateaus near t_0. Technically this variant of producing local plateau-functions is easier, thus we leave it to the reader to work out the details.

57.1 Characterization of the support

We want to show that the support of a distribution can be characterized as the set

$$LOC_\sigma = \{t \,|\, \delta_t = w^*\text{-}\lim{}_\alpha \, \sigma k_\alpha, \;\, k_\alpha \in \boldsymbol{S}_0(\mathbb{R}^d)\} \tag{448}$$ suppchar001

The concrete formulation of the statement is contained in the following proposition, which also shows that the w^*-approximating net (σk_α) can in fact be chosen to be uniformly concentrated (joint support on the time side) and bounded in the $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$-sense (in fact, uniformly bounded in the $\left(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\right)$-sense on the FT side).
We present the following important, but unfortunately technically involved result, which however demonstrates the advantages of having a well defined Fourier transform and avoids Lebesgue integration arguments:

charsupp01 **Proposition 22.** *Given $\sigma \in \boldsymbol{S}_0'$ we have the following characterization: $t \in \mathrm{supp}(\sigma)$ if and only if for any finite set $F = \{f_1, ..., f_L\} \subset \boldsymbol{S}_0(\mathbb{R}^d)$ and $\varepsilon > 0$ there exists some $k \in \boldsymbol{C}_c(\mathbb{R}^d) \cap \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) \subset \boldsymbol{S}_0(\mathbb{R}^d)$ such that one has:*

$$|\sigma \cdot k(f) - \delta_t(f)| = |[\sigma \cdot k](f) - f(t)| \leq \varepsilon, \quad f \in F. \tag{449}$$ charsupp

Proof. The idea of the proof is to choose the sequence of functions $k \in \boldsymbol{C}_c(\mathbb{R}^d) \cap \boldsymbol{S}_0(\mathbb{R}^d)$ in such way that $\|\widehat{\sigma k}\|_\infty \leq 2$, with compact, uniform convergence to $\widehat{\delta_t} = \chi_{-t}$.
First of all we observe that it is no restriction of the generality (and makes the description of the arguments more convenient) to assume that $\|\sigma\|_{\boldsymbol{S}_0'} = 1$ and that $t = 0$. The general case can be reduced to this case by a simple translation argument:

$$T_{-t}(\sigma \cdot k) = T_{-t}\sigma \cdot T_{-t}k, \quad t \in \mathbb{R}^d.$$

In order to achieve (449) (now for $t = 0$) we will assume that $\varepsilon > 0$ and F are given. It will be convenient two use write $g = \widehat{f}^{\checkmark} = \mathcal{F}^{-1}(f)$, for $f \in F$.
Without loss of generality we assume that $\|g\|_{\boldsymbol{L}^1} = \|\widehat{f}\|_{\boldsymbol{L}^1} = 1$ for $f \in F$. In order to show that we can approximate these functions by band-limited functions we make use of the Fourier invariance of $\boldsymbol{S}_0(\mathbb{R}^d)$, and find some $p = p^{\checkmark} \in \boldsymbol{C}_c(\mathbb{R}^d) \cap \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ (typically some finite partial sum of the form $\sum_{i \in E} \psi_k$ for some BUPU Ψ characterizing $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$), such that $\|p\|_\infty = 1$ and

$$\|g - g \cdot p^{\checkmark}\|_{\boldsymbol{L}^1} = \|\widehat{f} - \widehat{f} \cdot p\|_{\boldsymbol{L}^1} \leq \|\widehat{f} - \widehat{f} \cdot p\|_{\boldsymbol{S}_0} \leq \varepsilon/6, \;\; g = \mathcal{F}^{-1}(f), f \in F. \tag{450}$$ apprbandl

Let us fix R such that $\mathrm{supp}(p) \subset B_R(0)$, hence $\mathrm{supp}(p^{\checkmark}) \subseteq B_R(0)$.

Next we choose some $\tau \in \boldsymbol{S}_0(\mathbb{R}^d)$ such that $\tau(r) = 1$ on $B_1(0)$ (e.g. some trapezoidal function, or again, a finite partial sum from the BUPU). For the classical case one could choose $\mathcal{F}(\tau)$ to be a de La-Vallee Poussin kernel (belonging to $\boldsymbol{S}_0(\mathbb{R})$).
This implies that $\widehat{\tau} \in \boldsymbol{S}_0(\mathbb{R}^d) \subset \boldsymbol{L}^1(\mathbb{R}^d)$ and thus we can find $\delta > 0$ such that

$$\|T_z\widehat{\tau} - \widehat{\tau}\|_{\boldsymbol{L}^1} \le \varepsilon/4, \quad \text{for} \quad |z| \le \delta. \tag{451}$$

hattauLino

Applying the dilation operator D_ρ (for large ρ) to τ we have

$$\mathrm{D}_\rho\tau(u) = 1 \quad \text{for} \quad |u| \le 1/\rho$$

and for its Fourier transform $\mathrm{St}_\rho\widehat{\tau}$ we have, by choosing $\rho = R/\delta$ for $v \in Q$ obviously $|v/\rho| \le R/\rho = \delta$, and thus a significant stability with respect to shifts:

$$\|T_v\mathrm{St}_\rho\widehat{\tau} - \mathrm{St}_\rho\widehat{\tau}\|_{\boldsymbol{L}^1} = \|\mathrm{St}_\rho[T_{v/\rho}\widehat{\tau} - \widehat{\tau}]\|_{\boldsymbol{L}^1} = \|T_{v/\rho}\widehat{\tau} - \widehat{\tau}\|_{\boldsymbol{L}^1} \le \varepsilon/4, \quad v \in B_R(0). \tag{452}$$

DrhotauLino

By the definition of the support the assumption $0 \in \mathrm{supp}(\sigma)$ implies that for the open ball $B_{1/\rho}(0)$ can choose a sequence of functions k_n, with shrinking support (e.g. $\mathrm{supp}(k_n) \subset B_{1/n}(0)$ (and $n \ge n_0$ with $n_0 \ge \rho$), such that $\sigma \cdot k_n \ne 0$, thus $\|\widehat{\sigma k_n}\|_\infty > 0$.
We choose n in such a way that for the given choice of τ and ρ we have $\mathrm{supp}(k_n) \subseteq B_{1/\rho}(0)$. Once n is fixed we can write $k := k_n$, and see how we can make use of the conditions accumulated.
Since $\mathrm{supp}(\sigma \cdot k) \subset B_{1/\rho}(0)$ it is clear that $k \cdot \mathrm{D}_\rho\tau = k$, hence we have $(\sigma \cdot k) \cdot \mathrm{D}_\rho\tau = \sigma \cdot k$, and (by taking Fourier transforms) we observe that

$$\widehat{\sigma k} = \widehat{\sigma k} \star \mathrm{St}_\rho\widehat{\tau} \subset \boldsymbol{S}_0'(\mathbb{R}^d) \star \boldsymbol{S}_0(\mathbb{R}^d) \subset \boldsymbol{C}_{ub}(\mathbb{R}^d). \tag{453}$$

reproduce01

By suitably renormalizing k (replacing k by λk for a suitable $\lambda \in \mathbb{C}$), we can assume that we have $\widehat{\sigma k}(s_0) = 1$ for some $s_0 \in \mathbb{R}^d$ and[53] $\|\widehat{\sigma k}\|_\infty \le 2$. We also may assume that $s_0 = 0$ upon replacing k by $M_{-s_0}k$, which preserves the support of k and membership of k in $\boldsymbol{S}_0(\mathbb{R}^d)$.
Let now use the symbol $h := \widehat{\sigma k}$ (for convenience). We have ensured by our preparations that we have $h(0) = \widehat{\sigma k}(0) = 1$, $\|h\|_\infty \le 2$ and $h = h \star \mathrm{St}_\rho\widehat{\tau}$. But the variance condition (451) we obtain for $z \in B_R(0)$

$$|1 - h(z)| = |h(0) - h(z)| = |h \star \mathrm{St}_\rho\widehat{\tau}(0) - h \star T_z\mathrm{St}_\rho\widehat{\tau}(0)| \le \|h\|_\infty \|\mathrm{St}_\rho\widehat{\tau} - T_z\mathrm{St}_\rho\widehat{\tau}\|_{\boldsymbol{L}^1} \le 2\varepsilon/4.$$

This implies (still working on the FT side)

$$|\sigma_{\mathbf{1}}(gp) - \sigma_h(gp)| \le \int_{B_R(0)} |g(y)|\,|(1 - \widehat{\sigma k}(y))||p(y)|dy \le \frac{\varepsilon}{2}\|g\|_{\boldsymbol{L}^1} = \varepsilon/2, \quad f \in F,$$

and consequently we have with the help of (450) and the estimate $\|h\|_\infty \le 2$:

$$|\sigma_{\mathbf{1}}(g) - \sigma_h(g)| \le \varepsilon/2 + |\sigma_{\mathbf{1}}(g - pg)| + |\sigma_h(gp - g)| \le \varepsilon/2 + \varepsilon/6 + 2\varepsilon/6 = \varepsilon.$$

[53] By choosing some position s_0 with $|\widehat{\sigma k}(s_0)| \ge \|\widehat{\sigma k}\|_\infty/2 > 0$ and subsequently multiplying k by the inverse of $\widehat{\sigma k}(s_0)$.

By applying the inverse FT, given by $[\mathcal{F}^{-1}\sigma](f) = \sigma(\mathcal{F}^{-1}f)$, thus $(\mathcal{F}^{-1}\tau)(\mathcal{F}(f)) = \tau(f)$, for $f \in \boldsymbol{S}_0$, $\sigma, \tau \in \boldsymbol{S}_0'$, and specifically $\mathcal{F}^{-1}\mathbf{1} = \delta_0$ we get, using these identities

$$\sigma_{\mathbf{1}}(g) = \sigma_{\mathbf{1}}(\widehat{f}^{\checkmark}) = \mathcal{F}^{-1}(\mathbf{1})(f) = \delta_0(f), \quad \text{and} \quad \sigma_h(g) = \widehat{\sigma k}(\mathcal{F}^{-1}f) = (\sigma k)(f), \qquad (454)$$ backtotime01

we arrive at the required estimate (recall $g = \mathcal{F}^{-1}(f)$):

$$|\delta_0(f) - (\sigma k)(f)| = |f(0) - \sigma k(f)| = |\sigma_{\mathbf{1}}(g) - \sigma_h(g)| \leq \varepsilon, \quad f \in F. \qquad (455)$$ pullback112

The converse condition is quite easy. Assume that $t \notin \operatorname{supp}(\sigma)$, i.e. there exists $\gamma > 0$ such that $B_\gamma(t)$ is an area of trivial action, and thus we have $\operatorname{supp}(\sigma k) \cap B_\gamma(t) = \emptyset$. But this makes it impossible to approximate δ_t by expressions of the form σk!

$\square$

specsig **Corollary 14.** *Given $\sigma \in \boldsymbol{S}_0'(\mathbb{R}^d)$ we have: $s \in \operatorname{spec}\sigma$ if and only if there exists a sequence g_n in $\boldsymbol{S}_0(\mathbb{R}^d)$ such that*

$$\chi_s = w^*\text{-}\lim_{n\to\infty} g_n \star \sigma.$$

Viewing this result one can say: χ_s belongs to the spectrum of σ if and only if the pure frequency χ_s can be "filtered out of the signal σ".
This naturally gives rise to the following questions: If $\operatorname{spec}\sigma$ described the relevant sequences 'found inside of σ (from the analysis point of view), can one reconstitute the signal σ from those frequencies? To make this more precise, we can formulate it in the sense: Is there a sequence of finite linear combinations of the from $\sum_{i \in F} c_k \chi_{s_k}$ of trigonometric polynomials with $s_k \in \operatorname{spec}(\sigma)$, which is w^*-convergent to σ?
This questions, known as the question of *spectral synthesis*, is at the heart of abstract Fourier Analysis. It was a sensational insight that there are cases where the expected result cannot be achieved.
We will not discuss the delicate issue here in detail, but let us state a few results in this direction.

1. If one finds that $\operatorname{supp}(\sigma)$ is a nice set, e.g. a discrete set, or a subgroup (e.g. a lattice $\Lambda \lhd \mathbb{R}^d$, or the xy-plane in $\mathbb{R}^3$), then spectral synthesis is valid.

2. if the functional is a bounded measure, then synthesis is possible

3. there exists (for $\mathbb{R}^3$ first found by L. Schwartz sets which are not of spectral synthesis, i.e. where one can define functionals in $\boldsymbol{S}_0'(\mathbb{R}^d)$ which cannot be approximated in the w^*-sense by the trigonometric polynomials using *only frequencies* from $\operatorname{spec}(\sigma)$. The description does not use the pair $(SORd, \boldsymbol{S}_0'(\mathbb{R}^d))$, but rather $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ and its dual, but thanks to the so-called ideal theorem for Segal algebras one can show (this is not so difficult) that these questions are equivalent, because the local structure of $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ and of $\mathcal{F}(\boldsymbol{S}_0(\mathbb{R}^d)) = \boldsymbol{S}_0(\mathbb{R}^d)$ are the same!

Reiterspec **Lemma 99.** *The spectrum of $h \in \boldsymbol{L}^\infty(\mathbb{R}^d)$ as defined in* **re68:** [78] *coincides with* $\operatorname{supp}(\mathcal{F}(h))$, *i.e. the definition of the spectrum given here is consistent with the definition given in [78]:* $\operatorname{spec}_{Reiter}(h) = \operatorname{spec}(\sigma_h)$.

Another, perhaps not very significant comments tells us that the support of a measure can be computed locally, i.e. we have (the proof is left to the interested reader)

suppsigpsi **Lemma 100.** *For any BUPU* $\Psi = (\psi_i)_{i\in I}$ *and* $\sigma \in \boldsymbol{S}_0'$ *one has:*

$$\mathrm{supp}(\sigma) = \bigcup_{i\in I} \mathrm{supp}(\sigma\psi_i).$$

It is easy to show that one has

disjsupp **Lemma 101.** *Given* $\sigma \in \boldsymbol{S}_0'$, $f \in \boldsymbol{S}_0$, *with* $\mathrm{supp}(\sigma) \cap \mathrm{supp}(f) = \emptyset$. *Then* $\sigma(f) = 0$.

Proof. Since $f = \lim_{F\nearrow I} \sum_{i\in F} f\psi_i$ we can assume that f has *compact support* (the general case follows from this by taking limits). But any compact set (now $\mathrm{supp}(f) = \sum_{i\in F} f\psi_i$ for some fixed finite set) has a positive distance from a closed set, here $M := \mathrm{supp}(\sigma)$ (because the distance $d(x, M) = \min(|x - m|, m \in M)$ is a continuous function of x, which takes a (naturally strictly positive (!)) minimum over $\mathrm{supp}(f)$).
Thus by using a sufficiently fine BUPU $\Psi = (\psi_i)_{i\in I}$ we can *split the index set* I into two disjoint subsets I_1 such that $\sigma = \sum_{i\in I_1} \sigma\psi_i$ and $f = \sum_{i\in I_2} \psi_i f$. Consequently

$$\sigma(f) = \sigma\left(\sum_{i\in I} f\psi_i\right) = \sum_{i\in I_1} (\sigma\psi_i)(f) + \sum_{i\in I_2} \sigma(\psi_i f) = 0 + 0 = 0.$$

□

nearzero01.eps

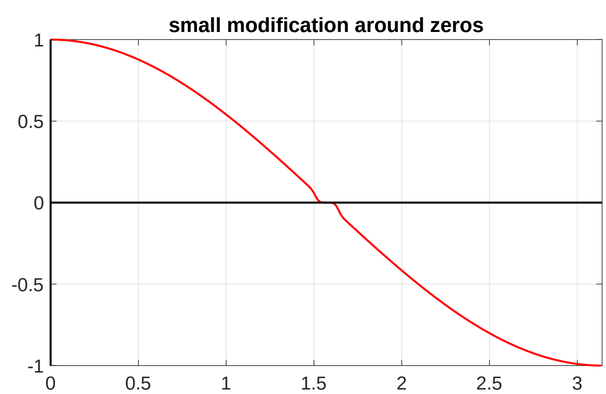

Abbildung 85: nearzero01.eps: continuous function flattened around zero

If we consider σ_h, a regular (mild) distributions induced by the continuous function $h \in \boldsymbol{C}_b(\mathbb{R}^d)$ (in the usual way) it is clear that $\mathrm{supp}(h) \cap \mathrm{supp}(f) = \emptyset$ implies $\sigma_h(f) = \int_{\mathbb{R}^d} h(x)f(x)dx$, because obviously the assumptions imply that $h(x)f(x) = 0$ for any $x \in \mathbb{R}^d$.
For bounded measures (and without providing details, for any $\sigma \in \boldsymbol{W}^*(\mathbb{R}^d)$ (the space of translation-bounded *measures* the situation is even better:

suppmeas **Lemma 102.** *For* $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ *ones has*

$$\mathrm{supp}(\mu) \cap \{x | f(z) = 0\} = \emptyset \quad\Rightarrow\quad \mu(f) = 0.$$

arzero02.eps

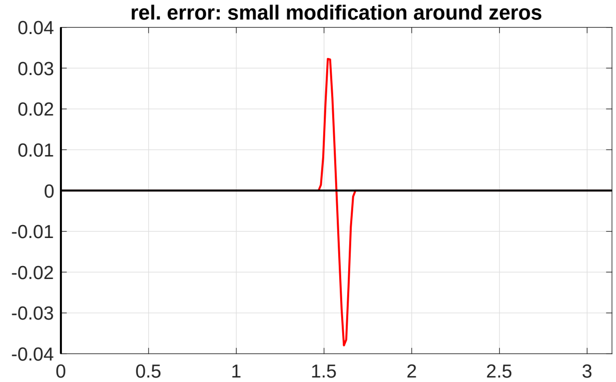

Abbildung 86: nearzero02.eps: observe the scaling

Proof. For any $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ (and even for $f \in \boldsymbol{C}_0(\mathbb{R}^d)$) one can modify the function (by replacing f by zero in some small neighborhood of its set of zeros (sometimes denoted by $f^{-1}(0) = \{z, f(z) = 0\}$, using the *inverse image notation for general mappings*) by zero.
Since this modification is (obviously, please take a look at the corresponding plots) can obviously be arranged to be small in the sense of the $\|\cdot\|_\infty$ we are done.
For $d \geq 2$ this arguments requires a little bit of topology argument, because in this case the zero-set $f^{-1}(0)$ can be quite strange, but the Lemma of Urysohn helps to overcome this problem.
A detailed technical realization can be obtained also by using the existence of arbitrary fine BUPUs. Given $f \in \boldsymbol{C}_{ub}(\mathbb{R}^d)$ with zero-set $Z(f) = \{z, f(z) = 0\}$ and $\varepsilon > 0$ we can find some $\delta > 0$ such that $|f(x-y) - f(x)| \leq \varepsilon$ whenever $|y| \leq 3\delta$ (this will be more convenient that the usual δ-condition).
Next we fix any δ-fine BUPU $\Psi = (\psi_i)_{i \in I}$, i.e. with $\mathrm{supp}(\psi_i) \subset B_\delta(x_i)$ for $i \in I$ and suitable points $x_i \in \mathbb{R}^d$. Without loss of generality we may assume non-negativity of ψ_i, or $|\psi_i(x)| = \psi_i(x)$ for $i \in I$.
Let us now define the index set corresponding to those members of the BUPU which are close to the zero set of f:

$$I_1 := \{i \in I \,|\, \mathrm{dist}(\mathrm{supp}(\psi_i), Z) \leq \delta\},$$

and $I_2 = I \setminus I_1$.
Now we have for $x \in \mathrm{supp}\, \psi_i \subseteq B_\delta(x_i)$

$$dist(x, Z) \leq dist(x_i, Z) + |x - x_i| \leq 2\delta.$$

But $f(z) = 0$ for any $z \in Z$ and thus by the assumed equicontinuity of f we have $|f(x)| \leq \varepsilon$ for all $x \in \mathrm{supp}\, \psi_i$. Summing over all the contributions from I_1 we have

$$\left| \sum_{i \in I_1} f(x)\psi_i(x) \right| \leq \sum_{i \in I_1} |f(x)||\psi_i(x)| \leq \varepsilon \sum_{i \in I} \psi_i = \varepsilon, \quad x \in \mathbb{R}^d.$$

Thus changing from $f = \sum_{i \in I} f\psi_i$ to $f_1 = \sum_{i \in I_1} f\psi_i$ is a minor change, but obviously the remainder, i.e.

$$f_2 = f - f_1 = \sum_{i \in I_2} f\psi_i$$

has support with a δ-separation from Z (because it is constituted only by those contributions $\psi_i, i \in I_2$ which are at some minimal distance δ from Z), so that $\mu(f_2) = 0$ (by Lemma disjsupp 101). This implies that

$$|\mu(f)| \leq |\mu(f_1)| + |\mu(f_2)| \leq \|\mu\|_{\boldsymbol{M}_b} \|f_1\|_\infty + 0 \leq \varepsilon,$$

but since this estimate is valid for every positive $\varepsilon > 0$ we have $|\mu(f)| = 0$ or just $\mu(f) = 0$ as claimed. $\square$

delxBUPU1.eps

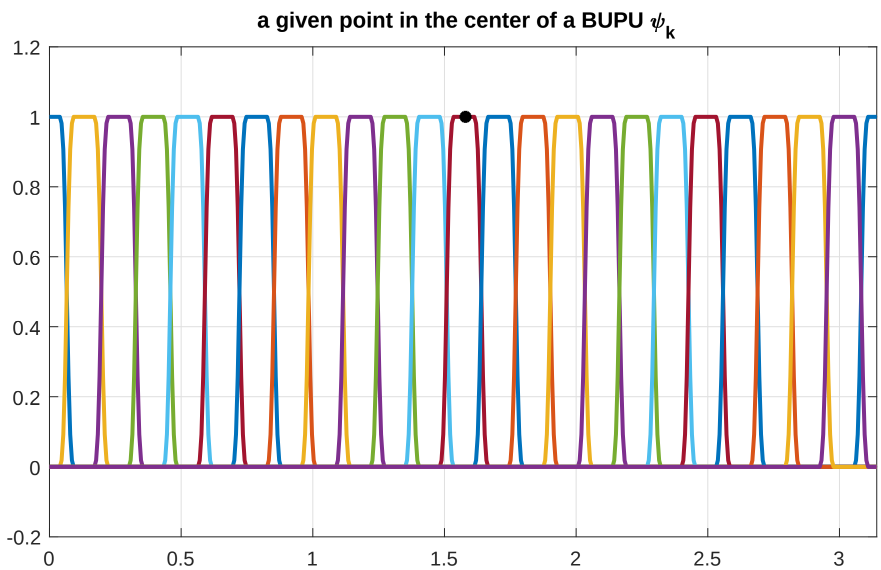

Abbildung 87: delxBUPU1.eps

oneptsupp **Lemma 103.** *Assume that* $\mathrm{supp}(\sigma) = \{t_0\}$*, for some* $t_0 \in \mathbb{R}^d$*. Then we have, for arbitrary fine, well chosen BUPUs* $\Psi = (\psi_k)_{k \in \mathbb{Z}^d}$*:* $\sigma = \sigma\psi_{k_0}$ *for some* $k_0 \in \mathbb{Z}^d$*.*

Proof. First we define a BUPU which is built from a generator ψ_0 such that $\psi_0(x) = 1$ for $|x| \leq \delta$. By suitable dilation (if necessary) we can produce a dilated version, for the lattice $\alpha\mathbb{Z}^d$ for some (small) value α, such that $x = \alpha k_0$ for some integer. Then $\sigma = \sigma\psi_{k_0}$, and $\mathrm{supp}(\psi_{k_0}) \subset B_\delta(x)$ for any given value $\delta > 0$ (just by choosing $\alpha > 0$ appropriately). $\square$

onepointsupp **Lemma 104.** *Given* $\sigma \in \boldsymbol{S}_0'$ *with* $\mathrm{supp}(\sigma) = \{t_0\}$*. Then one has* $\sigma(h) = 0$ *for any* $h \in \boldsymbol{S}_0(\mathbb{R}^d)$ *with* $h(z) = 0$ *in* $B_\gamma(t_0)$*, for some* $\gamma > 0$*.*

Proof. By adjusting a BUPU $\Psi = (\psi_k)_{k \in \mathbb{Z}^d}$ in such a way that only $\sigma = \sigma\psi_{k_0}$, but also $\mathrm{supp}(\psi_{k_0}) \cap \mathrm{supp}(h) = \emptyset$. Hence one has $\sigma\psi_{k_0}(h) = \sigma(\psi_{k_0}h) = 0$ while $\sigma\psi_k = 0$ for $k \neq k_0$ and consequently $\sigma(h) = \sum_{k \in \mathbb{Z}^d} \sigma(\psi_k h) = 0$. $\square$

oneptsynth **Lemma 105.** *Assume* $0 \neq \sigma \in \boldsymbol{S}_0'(\mathbb{R}^d)$*, such that* $\mathrm{supp}(\sigma) = \{t_0\}$ *for some* $t_0 \in \mathbb{R}^d$*. then* $\sigma = \lambda\delta_{t_0}$*, for some* $\lambda \in \mathbb{C}$*.*

Proof. Without loss of generality we may assume that $t_0 = 0$. We consider the closed subspace. $\mathcal{N}_\sigma = \{f \in \boldsymbol{S}_0(\mathbb{R}^d), \sigma(f) = 0\}$.
By Lemma onepointsupp 104 the set of

$$I_0 = \{h \in \boldsymbol{S}_0(\mathbb{R}^d) \,|\, h(0) = 0\}$$

is a closed ideal (for pointwise multiplication) in $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$, which has bounded approximate units with compact support, and consequently the elements $h \in I_0$ with $\mathrm{supp}(h)$ compact and disjoint from $B_\gamma(0)$ form a dense subspace of I_0.
Hence for given $\varepsilon > 0$ any $h \in I_0$ can be split as $h = h_1 + (h - h_1)$ with $\|h - h_1\|_{\boldsymbol{S}_0} \leq \varepsilon$. [needs some refinement!]
Thus

$$\sigma(h) = \sigma(h_1) + \sigma(h - h_1) = 0 + \sigma(h - h_1),$$

but this implies that one has for any ε:

$$|\sigma(h)| \leq \|\sigma\|_{\boldsymbol{S}_0'} \|h - h_1\|_{\boldsymbol{S}_0} \leq \varepsilon \|\sigma\|_{\boldsymbol{S}_0'},$$

which in turn implies (since $\varepsilon > 0$ was arbitrary) $|sigma(h)| = 0$, or $\sigma(h) = 0$ for $h \in I_0$. In other words $I_0 \subset \mathcal{N}_\sigma$. Since clearly $\mathcal{N}_\sigma \neq \boldsymbol{S}_0(\mathbb{R}^d)$ (because $\sigma \neq 0$) and because I_0 is maximal we have equality: $I_0 = \mathcal{N}_\sigma$.
This means that $\sigma(f) = \sigma(f + h)$ for any $h \in I_0$, and consequently it can be viewed as a mapping from the quotient space $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)/I_0$ into $\mathbb{C}$. But (due to the maximality of I_0) this quotient space is one-dimensional and one has (by the principle of Gelfand-Mazur, see the theory of Banach algebras) we have $\sigma(f_1) = \sigma(f_2)$ if and only if $f_1 + I_0 = f_2 + I_0$ or otherwise, if and only if $f_1(0) = f_2(0)$. Take any function f_0 with $f_0(1) = 1$. Let us put $\lambda = \sigma(f_0)$.
As already observed any other function $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ with $f(0) = 1$ we get as well $\sigma(f) = \lambda$ (due to the equivalence relation).
Obviously, for an arbitrary function $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ we have (unless $g(0) = 0$, but then we know already that $\sigma(g) = 0$) $f_1 = f/g(0)$ is a function with $f_1(0) = 1$, and thus

$$\sigma(g) = \sigma(g(0)f_1) = g(0)\lambda = \lambda\delta_0(g)$$

i.e. we obtain the claimed realization of σ as

$$\sigma = \lambda\delta_0.$$

□

We can come up with a characterization of mild distributions supported on a lattice $\Lambda \lhd \mathbb{R}^d$. In fact, with minor modifications one can even take perturbed lattice or more generally (among others) sets which are η-separated.

etaseparat **Definition 24.** A discrete set $\Lambda \subset \mathbb{R}^d$ is called η - separated for some $\eta > 0$, if one has for any two points from Λ:

$$|\lambda - \lambda'| \geq \eta, \mathrm{for} \lambda \neq \lambda'.$$

We can summarize the observations resulting from the characterization of one-point support in the following way:

elspreaddist **Proposition 23.** *For any $\eta > 0$ there exists a constant $C_\eta > 0$ such that one has: any $\sigma \in \boldsymbol{S}_0'$ with $\operatorname{supp}(\sigma) = X$ for some η-separated set $X = (x_j)_{j \in J}$ one has:*

$$\sigma = \sum_{j \in J} c_j \delta_{x_j}, \quad \textit{with} \quad \sup_{j \in J} |c_j| \leq \|\sigma\|_{\boldsymbol{S}_0'}.$$

Conversely, there exists some constant $C_\eta > 0$ such that any (discrete and possibly unbounded) measure $\mu = \sum_{j \in J} c_j \delta_{x_j}$ defines a bounded linear functional on $\left(\boldsymbol{W}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{W}}\right) = \left(\boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{W}}\right)$ with a uniform norm estimate of the form

$$\|\sigma\|_{\boldsymbol{S}_0'} \leq \|\mu\|_{\boldsymbol{W}^*(\mathbb{R}^d)} \leq C_\eta \sup_{j \in J} |c_j| \tag{456}$$ estsepsupport

for any η-separated set X.

Proof. First we observe that one has: Given $\eta > 0$ one can choose some $\delta > 0$ (essentially $\delta < \eta/3$) such that $\psi_i(x_j) \neq 0$ for some $i \in I$ and $j \in J$ one has $x_j \in B_\delta(x_i)$, or $|x_i - x_j| \leq \delta$.
Consequently for any FIXED $j \in J$ we have for any *index* $i \in I$ such that $\psi_i(x_j) \neq 0$: $\operatorname{supp}(\psi_i) \subset B_{2\delta}(x_j)$. Let us write $I_j := \{i \in I, \psi_i(x_j) \neq 0\}$.
Thus we have $I_j \cap I_{j'} = \emptyset$ for $j \neq j'$ in J, because clearly $x_j \notin B_{2\delta}(x_i)$ because

$$|x_{j'} - x_i| \geq |x_{j'} - x_j| - |x_j - x_i| \geq \eta - 2\delta \geq \delta > 0.$$

In other words, we have (in the w^*–sense)

$$\sigma = \sum_{i \in I} \sigma \psi_i = \sum_{j \in J} \sum_{i \in I_j} \sigma \psi_i,$$

with the inner sum having the property

$$\operatorname{supp}\left(\sum_{i \in I_x} \sigma \psi_i\right) = \{x_j\}, \quad j \in J.$$

As a consequence of Lemma oneptsynth 105 this inner sum is of the form $c_j \delta_{x_j}$, for some $c_j \in \mathbb{C}$. Altogether we have now

$$\sigma = \sum_{i \in I} \sigma \psi_i = \sum_{j \in J} \sum_{i \in I_j} \sigma \psi_i = \sum_{j \in J} c_j \delta_{x_j}.$$

It is *important to note that for all the* η-separated sets $X = (x_j)_{j \in J}$ the same BUPU can be used. One characterization of $\boldsymbol{S}_0'(\mathbb{R}^d)$ would be to compare its norm with the expression $\sup_{i \in I} \|\mathcal{F}(\sigma\psi_i)\|_\infty$ for a fixed BUPU $\Psi = (\psi_i)_{i \in I}$. We will not use this argument here.
But this estimate is easy to achieve by choosing any test function $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ with small support (for $d = 1$ it could be just a compressed triangular function of the form $\mathrm{D}_\rho \Delta$) with $f(0) = 1$. Then, by applying σ to $T_{x_j} f$ with $\|T_{x_j} f\|_{\boldsymbol{S}_0} = \|f\|_{\boldsymbol{S}_0}$ and $T_{x_j} f(x_j) = f(x_j - x_j) = f(0) = 1$ and $T_{x_j}(x_{j'} = 0$ for any $j' \in J, j' \neq j$, hence

$$\sigma(f) = \sum_{j \in J} c_j \delta_{x_j}(f) = c_j.$$

But this implies $|c_j| = |\sigma(f)| \leq \|\sigma\|_{\boldsymbol{S}_0'} \|f\|_{\boldsymbol{S}_0}, \quad j \in J.$
The converse is verified with the help of the atomic characterization of $\boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d)$. In fact, for atoms of a given size, say supported on balls of the form $B_1(x)$ for some $x \in \mathbb{R}^d$ each such ball hits at most a limited number of points $x_j, j \in J$, say at most K of them (this number depends ONLY on $\eta > 0$!). Thus we can estimate the action on $k \in \boldsymbol{C}_0(\mathbb{R}^d)$ with $\mathrm{supp}(k) \subseteq B_1(x)$ and $\|k\|_\infty = 1$:

$$|\sigma(k)| \leq K \left(\sup_{j \in J} |c_j| k(x_j) \right) \leq K \sup_{j \in J} |c_j|,$$

thus completing the proof. □

relsep01.jpg

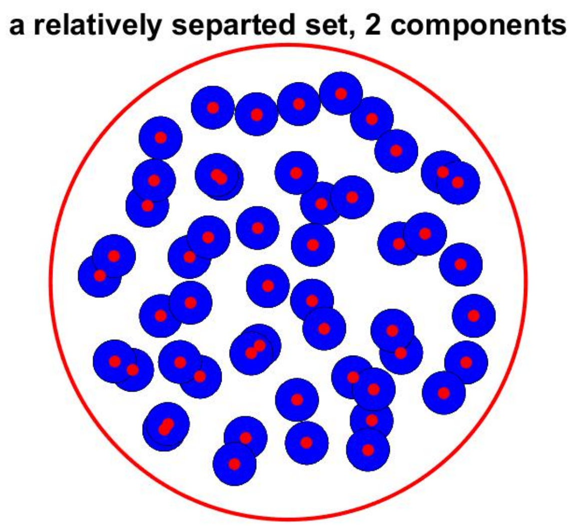

Abbildung 88: relsep01.jpg

relsep02.jpg

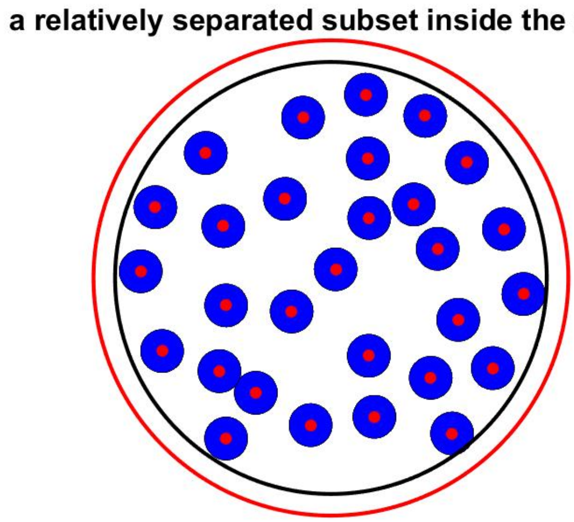

Abbildung 89: relsep02.jpg

58 Back to the theory of Multipliers

The starting point of our journey was the characterization of BIBOS systems as convolution operators with the help of bounded measures. We also have seen that these convolution operators (each such operator is the adjoint of another such operator, with the measure $\nu := \mu^{\checkmark}$) extend from the domain $\boldsymbol{C}_c(\mathbb{R}^d)$ to all of $\boldsymbol{M}_b(\mathbb{R}^d)$, in a unique way to a w^*-w^*-continuous manner.
So we should return to this topic for more general settings (cf. the recent article **un20:** [89]).

59 Outlook on further areas of applications

Further topics which have not been discussed during the course:

59.1 Laurent Schwartz: Tempered Distributions

The standard approach to an *extended Fourier* for *tempered distributions* starts with the Schwartz space of *rapidly distributions*, defined as follows:

$$\boldsymbol{\mathcal{S}}(\mathbb{R}^d) = \{f \in \boldsymbol{C}_b(\mathbb{R}^d), \forall k \in \mathbb{N} : Df(x)(1+|x|^2)^k \in \boldsymbol{C}_b(\mathbb{R}^d)\} \qquad (457)$$ ScRddef

for any partial differential operator D. The definition immediately implies that $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ is also stable under differentiation (unlike $\boldsymbol{S}_0(\mathbb{R}^d)$) and in fact (much less obvious) under the Fourier transform. The defining expression $\|Df(x)(1+|x|^2)^k\|_\infty$, with D and k running through their natural parameter spaces, define a (countable) family of seminorms, which allows to define a *metric* (but not a norm) which allows to describe convergence in a

relsep04.jpg

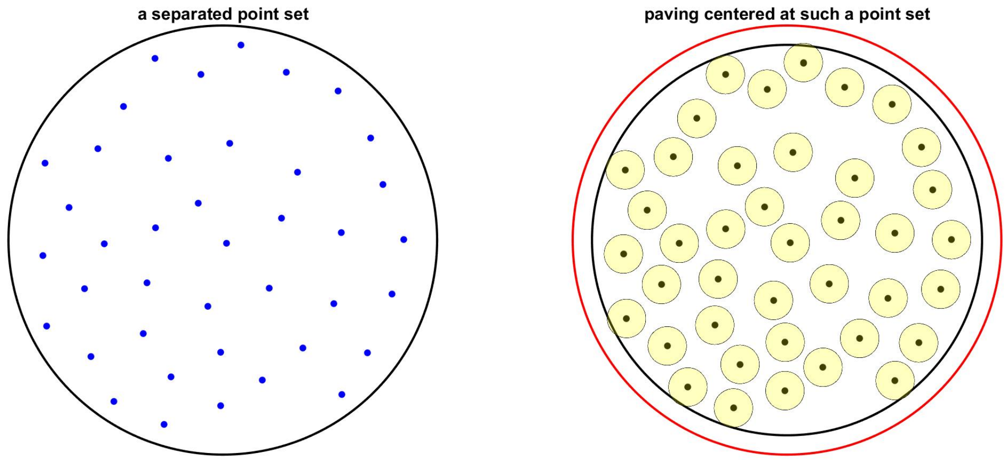

Abbildung 90: relsep04.jpg

somewhat involved way, which turns $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ into a *topological vector space*, in fact, into a so-called *nuclear Frechet space*.
The dual space, the set of all linear functionals, is called the space of $\boldsymbol{\mathcal{S}}'(\mathbb{R}^d)$ of *tempered distributions*. The continuity can be (a bit more practicable) described as follows: Given $\sigma \in \boldsymbol{\mathcal{S}}'(\mathbb{R}^d)$ one can find a finite family of seminorms (the selection depends on σ) such that one can assure that for *any sequence* in $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ $f_n \to f_0$ with respect to *these seminorms* (in $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$) guarantees that $\sigma(f_n) \to \sigma(f_0)$.
One goes on the check that $\mathcal{F}\colon \boldsymbol{\mathcal{S}}(\mathbb{R}^d) \to \boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ is a *continuous mapping* (i.e. respecting this kind of convergence) and thus one can extend it to $\boldsymbol{\mathcal{S}}'(\mathbb{R}^d)$ by the standard trick $\widehat{\sigma}(f) = \sigma(\widehat{f})$.
It is easy to check (due to the density of $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ in $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$) that this is in fact a further extension of the extended Fourier transform on $\boldsymbol{S}_0'(\mathbb{R}^d)$, or expressed differently. $\boldsymbol{S}_0'(\mathbb{R}^d)$ is subspace of $\boldsymbol{\mathcal{S}}'(\mathbb{R}^d)$ which is invariant under the extended (Schwartz) FT for tempered distribution. These properties make $\boldsymbol{\mathcal{S}}'(\mathbb{R}^d)$ so useful for PDEs (partial differential equations), because it turns a partial differential equations (you can think of it as convolution with a derivative of the Dirac Delta) into a pointwise multiplication operator.
It is also easy to characterize $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ and $\boldsymbol{\mathcal{S}}'(\mathbb{R}^d)$ by the short-time Fourier transform:

ScRdinSORd **Lemma 106.** *A function $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ belongs to the subspace if*

$$|V_g f(t,s)| \cdot (1 + |t|^2 + |s|^2)^k \in \boldsymbol{C}_b(\mathbb{R}^d \times \widehat{\mathbb{R}}^d), \quad \forall k \in \mathbb{N}.$$

On the other hand, one can define $V_g\sigma$ for any $\sigma \in \boldsymbol{\mathcal{S}}'(\mathbb{R}^d)$ is continuous and there exists $K \in \mathbb{N}$ such that

$$\sup_{(t,s)\in\mathbb{R}^d\times\widehat{\mathbb{R}}^d} |V_g\sigma(t,s)| \leq (1 + |t|^2 + |s|^2)^K < \infty.$$

relsep05.jpg

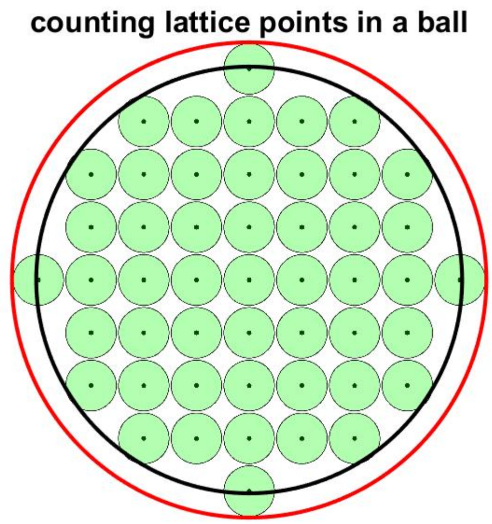

Abbildung 91: relsep05.jpg

One also has a similar regularization setting as for $\boldsymbol{S}_0'(\mathbb{R}^d)$, namely

$$\Big(\boldsymbol{\mathcal{S}}'(\mathbb{R}^d) \star \boldsymbol{\mathcal{S}}(\mathbb{R}^d)\Big) \cdot \boldsymbol{\mathcal{S}}(\mathbb{R}^d) \subseteq \boldsymbol{\mathcal{S}}(\mathbb{R}^d), \quad \Big(\boldsymbol{\mathcal{S}}'(\mathbb{R}^d) \cdot \boldsymbol{\mathcal{S}}(\mathbb{R}^d)\Big) \star \boldsymbol{\mathcal{S}}(\mathbb{R}^d) \subseteq \boldsymbol{\mathcal{S}}(\mathbb{R}^d).$$

There is also a sequential approach, prominently presented in the book of Lighthill (see **li62:** [73]). A similar approach can be given for the space $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$ as explained in **fe20-1:** [35].

59.2 Quasi-Crystals

In the theory of *quasi-crystals* (which also can be formulated over general LCA groups), see the work of C. Baake and coauthors ([3,4]) makes use of a Fourier transform for unbounded measures. This is a concept developed by J. Gil de Lamadrid and L. Argabright ([2,56]), which can be described by means of translation bounded measures. Essentially one has

ildeLamChar1 **Lemma 107.** *A translation bounded measures is transformable in the sense of $A-GdL$ with a transformable FT $\widehat{\mu}$ (in the sense of $\boldsymbol{S}_0'(\mathbb{R}^d)$ or $\boldsymbol{S}_0'(G)$) if and only if both μ and $\widehat{\mu}$ belong to the dual of Wiener's algebra, i.e. to $\boldsymbol{W}(\boldsymbol{M}, \boldsymbol{\ell}^\infty)(\mathbb{R}^d)$.*

60 Compactness in Function Spaces

There is a simple general criterion for compactness in function spaces having a double module structure (as described in [8]), namely: a bounded, closed subset M of $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ is compact if and only if it is *tight* and *equicontinuous*, i.e. the continuity of shifts is uniform with respect to the set.

relsep05b.jpg

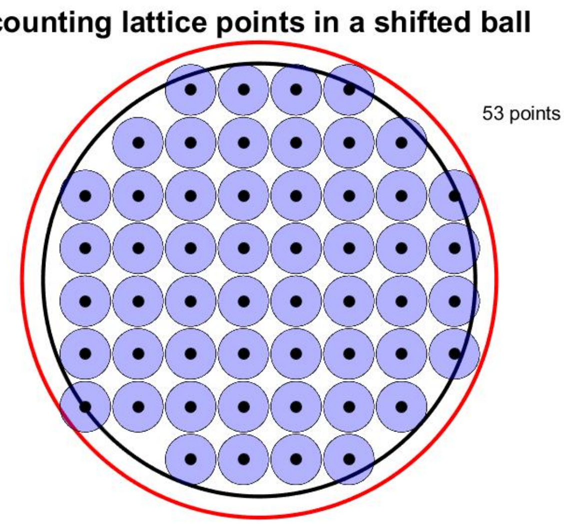

Abbildung 92: relsep05b.jpg

With the help of approximate units in (for example) $\left(\boldsymbol{L}^1(\mathbb{R}^d),\, \|\cdot\|_1\right)$ and $\left(\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d),\, \|\cdot\|_{\mathcal{F}\boldsymbol{L}^1}\right)$ this can be expressed in a way which makes clear (in a more explicit) way that these two properties are symmetric under the Fourier transform (which of course maps compact subsets of $(\boldsymbol{B},\, \|\cdot\|_{\boldsymbol{B}})$ into compact subsets of $\mathcal{F}\boldsymbol{B}$, endowed with its natural norm).
Specifically for $\mathscr{G} = \mathbb{R}^d$ it can be characterized as follows, making use of the classical Gauss-function with $\widehat{g}(0) = 1$.

tightness **Lemma 108.** *A bounded set M in $(\boldsymbol{B},\, \|\cdot\|_{\boldsymbol{B}})$ is equicontinuous if and only if for $\varepsilon > 0$ there exists ρ_0 such that for $\rho \in (0, \rho_0]$ one has:*

$$\|\mathrm{St}_\rho g_0 \star f - f\|_{\boldsymbol{B}} \le \varepsilon, \quad \forall f \in M \tag{458}$$ apprconv01

tightness **Lemma 109.** *A bounded set M in $(\boldsymbol{B},\, \|\cdot\|_{\boldsymbol{B}})$ is tight if and only if for $\varepsilon > 0$ there exists ρ_0 such that for $\rho \in (0, \rho_0]$ one has:*

$$\|\mathrm{D}_\rho g_0 \star f - f\|_{\boldsymbol{B}} \le \varepsilon, \quad \forall f \in M \tag{459}$$ tight03

Thus clearly tightness on the time-side corresponds to equicontinuity on the Fourier transform side and vice versa. The most important special cases are those for Fourier invariant spaces such as $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ or $\left(\boldsymbol{L}^2(\mathbb{R}^d),\, \|\cdot\|_2\right)$ (see the paper of L. Pego **pe85-1:** [77]).

PegoS0 **Corollary 15.** *A closed, bounded subset M of $\left(\boldsymbol{L}^2(\mathbb{R}^d),\, \|\cdot\|_2\right)$ (or also $\boldsymbol{S}_0'(\mathbb{R}^d)$) is compact in $\left(\boldsymbol{L}^2(\mathbb{R}^d),\, \|\cdot\|_2\right)$ (resp. $\boldsymbol{S}_0'(\mathbb{R}^d)$) if both M and $\mathcal{F}(M)$ are tight in $\left(\boldsymbol{L}^2(\mathbb{R}^d),\, \|\cdot\|_2\right)$ (resp. $\boldsymbol{S}_0'(\mathbb{R}^d)$).*

This statement is a variant of the following principle: Good decay of the Fourier transform implies that (by the Fourier inversion theorem) the function has to have few contributions from high frequencies, and this of course means that the function is smooth. On the other hand, if you start with a smooth function then the Fourier transform will have good decay properties. In principle, integration a function with is very flat, or almost constant over a fixed interval, will contribute very little to $\widehat{f}(s)$ if s is large enough, since the pure frequencies $\cos(2\pi sx)$ (and $\sin(2\pi sx)$) both have integral zero (also integrated against any constant function!).

Recall, that a subset M of a normed space $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ is called compact if and only if every sequence $(h_n)_{n=1}^{\infty}$ in $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ has a convergent subsequence $(h_{n_k})_{k=1}^{\infty}$, i.e. there exists some $h_0 \in M$ such that

$$\lim_{k\to\infty} h_{n_k} = h_0 \quad \text{in } (\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}}).$$

It is not hard to show that any compact set in $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ has to be bounded AND closed. For the finite dimensional case also the opposite is true (by the Heine-Borel Theorem for $\mathbb{R}^{d^d}$).

A characterization if compact sets is the following:

relcompchar **Definition 25.** A subset M of a normed space $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ is called *relatively compact* if the following is true: For any $\varepsilon > 0$ there is a finite subset $F = \{m_1, ..., m_L\} \subset M$ such that

$$M \subseteq \bigcup_{k=1}^{L} B_{\varepsilon}(m_k).$$

Without proof we provide the important equivalence statement (found in functional analysis books):

charcomp **Lemma 110.** *A closed subset M of a normed space $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ is compact if and only if it is relatively compact.*

Since bounded (closed) sets in a non-finite-dimensional Banach space $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ may not be compact the following theorem from functional analysis is important for applications, and was in fact used in several places during the course:

Alaoglu **Theorem 25.** *Any bounded subset M of a dual Banach space $(\boldsymbol{B}', \|\cdot\|_{\boldsymbol{B}'})$ is w^*-compact, i.e. any sequence has a w^*-convergent subsequence (if $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ is separable), or at least a net which is convergent in $\boldsymbol{B}'$ in the sense that*

$$w^*\text{-}\lim{}_{\alpha}\sigma_{\alpha} = \sigma_0 \in \boldsymbol{B}'.$$

This important theorem is based on the *Theorem of Tychonoff* (telling that the product of any number of compact topological spaces, with a suitable mode of convergence) is compact again. The proof makes use of the *Axiom of Choice* respectively the (equivalent) Hausdorff Maximality Principle. In other words, there is no constructive proof for finding the convergent subsequence *in general*.

We have used (at least implicitly) the Alaoglu Theorem in many places, but most of the time just as a justification for working with the w^*-topology and convergence. *However*, we have always tried to provide *constructive approximation procedures*, e.g. the use of the discretization operator D_{Ψ} for the w^*-approximation of bounded measures by discrete measures.

61 Survey-type papers by the author

feja20: [41] is a kind of preparatory step for the current notes, indicating how one can develop in a mathematical correct way the foundations of Fourier and Time-Frequency Analysis. There are still a lot of detail that have to be treated and to be worked out, in some case we have correct arguments, but it would be great to refine the approach and obtain more elegant or more intuitive lines of arguments, or to reorganize the overall approach to make it is compact and striking as possible, without dropping the intuitive connection to the finite case.

fe20-1: [35] describes the simplicity of the *sequential approach to mild distributions*, which follows the ideas exploited for the context of tempered distributions by Lighthill (or Jones and others). However, the verification of the fact (which of course might not bother the user) that the realization of the equality of the objects obtained by the duality approach on the one hand and the *completion* realized by the equivalence classes of nice functions is much less involved and uses again just basic principles of Banach space theory (namely the principle of uniform boundedness), combined with the specific properties of $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ and $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$.

fe20-2: [36] puts the theory into a context outlining more or less the program executed in this course in general terms. It also emphasizes the role of the Banach Gelfand Triple $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$, which has been introduced earlier, and is described (starting from the standard tools, and allowing Lebesgue integration to define $\boldsymbol{L}^1(\mathbb{R}^d), \boldsymbol{L}^2(\mathbb{R}^d)$ and $\boldsymbol{L}^\infty(\mathbb{R}^d)$) is given in **cofelu08:** [15]. Another summary of the basic facts of the Banach Gelfand triple is provided by **fe18-3:** [32].

The first formulation of concepts related to the current course are found in **fe14:** [29]. The subsequent paper **fe15:** [28] describes a number of principles which are used in the literature (and by the author) in order to create function spaces, or to combine different principles (like decomposing a function into pieces, using a BUPU, or doing the same on the Fourier transform side, in order to obtain *modulation spaces*). But it also puts an emphasis on the question, which function space are most suitable for the treatment of a given problem. In fact, the idea of providing something like *consumer's reports* is ventilated in this paper. Of course, the pair $(\boldsymbol{S}_0, \boldsymbol{S}_0')$ plays an important role from this view-point due to its universal usefulness, not only inside of Gabor and time-frequency analysis.

The connections between abstract and numerical harmonic analysis are addressed by the idea of *Conceptual Harmonic Analysis* first promoted in **fe16-1:** [30] which is an important direction to be developed further in the coming years. The role of numerical work for the development of Gabor analysis (in the past and I think also for the future) is partially addressed in **fe19-2:** [34].

The connections between classical Fourier Analysis (worked out in much more detail now in the current lecture notes) is summarized for the first time in a systematic way in **fe19:** [33].

The kernel theorem is one of the most important additional features (in addition to the Fourier invariance) of the Banach Gelfand Triple $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$. It implies that any bounded linear operator $T : \left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right) \to (\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$ can be characterized by a $2d$-dimensional kernel $\sigma \in \boldsymbol{S}_0'(\mathbb{R}^d \times \widehat{\mathbb{R}}^d)$.

61.1 Generalized Stochastic Processes

This theory is the common generalization of the term *stochastic process* (uncertain outcome at a given time) and *generalized functions*, which give an average over time: a precise number, given the time (something like the idea of thinking of the bins in the histogram as expressions of the form $\mu(\psi_i)$).
Combining the two ideas is very natural and requires of course corresponding mathematical tools, namely bounded linear operators defined on the space of test functions (here $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$) into some (abstract) Hilbert space which is describing the probabilistic background, so usually it is a space of square integrable random variables on some abstract probability space, modified in such a way that one subtracts the meanwhile (the expected value) of the random variable, with the good consequence that *uncorrelated random variables* are then exactly those which are orthogonal in such an abstract Hilbert space. Obviously one can define for each such *generalized stochastic process*, which is a bounded linear operator $\rho : \left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right) \to (\mathcal{H}, \|\cdot\|_{\mathcal{H}})$ the *spectral process* (we would say the Fourier transform of the process), given by $\widehat{T}(f) = T(\widehat{f})$, $f \in \boldsymbol{S}_0$.
The theory of *generalized stochastic processes* also naturally employs the space $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ and $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$, see **feho14:** [40], based on the master thesis and PhD of Wolfgang Hörmann (**ho89:** [64]), or B. Keville (**ke03:** [70]).

62 PINV, Frames and Riesz Bases

From linear algebra we know the PINV, the (Moore-Penrose) pseudo-inverse of a matrix, or in MATLAB pinv($\boldsymbol{A}$). It has the format of the transpose of a matrix. In case the mapping $\mathbf{x} \to \boldsymbol{A} \star \mathbf{x}$ is surjective onto $\mathbb{C}^m$, i.e. if the column-space of $\boldsymbol{A}$ is all of $\mathbb{C}^m$ every $\mathbf{y} \in \mathbb{C}^m$ has (usually many) representations of the form $\mathbf{y} = \boldsymbol{A} \star \mathbf{x}$, for suitable elements $\mathbf{x} \in \mathbb{C}^n$. The variety of all those vectors defines and *affine subspace* of $\mathbb{C}^n$ and thus has a (unique) element of *minimal norm*. The pseudo-inverse delivers exactly this element. Thus we have in this case

$$\boldsymbol{A} \star \mathrm{pinv}(\boldsymbol{A}) = \mathrm{Id}_{\mathbb{C}^m}$$

and if $\mathbf{z} \in \mathbb{C}^n$ satisfies $\boldsymbol{A} \star \mathbf{z} = \mathbf{y}$ then we know:

$$\|\mathrm{pinv}(\boldsymbol{A}) \star \mathbf{x}\| \leq \|\mathbf{z}\|.$$

Frames, i.e. countable families of vectors $(f_i)_{i\in I}$ with similar properties as called *frames* in $(\mathcal{H}, \|\cdot\|_{\mathcal{H}})$. The coefficient functionals for the minimal norm representation of any $f \in \mathcal{H}$ depend in a linear way on element $f \in \mathcal{H}$ which should be represented, and thus we can describe it by a set of linear functionals $f \mapsto \langle f, \widetilde{f}_i\rangle_{\mathcal{H}}$. The corresponding family $(\widetilde{f}_i)_{i\in I}$ is called the *dual frame*. It is actually also a generating system with the same properties and the dual frame for the dual frame is (as the terminology suggests) the original frame.
Usually frame in Hilbert spaces are described as sequences of vectors $(f_n)_{n=1}^{\infty}$ which satisfy for some positive numbers $A, B > 0$ the following inequality

$$A\|f\|^2 \leq \sum_{n=1}^{\infty} |\langle f, f_n\rangle_{\mathcal{H}}|^2 \leq B\|f\|^2, \quad f \in \mathcal{H}, \tag{460}$$ framinequ01

which in turn implies that the so-called frame operator

$$Sf = \sum_{n=1}^{\infty} \langle f, f_n \rangle_{\mathcal{H}} f_n, \quad f \in \mathcal{H} \tag{461}$$ framopdef01

is well defined and (positive definite hence) *invertible*. The dual frame then can be computed (at least abstractly) as the family$(\tilde{f}_n)_{n=1}^{\infty} = (S^{-1}(f_n))_{n=1}^{\infty}$.
We also could have discussed *Riesz bases* and *Riesz basic sequences*.

63 Banach modules and approximate units

This is another (abstract) chapter, which would provide further unified notation describing the situation encountered during this course (going back to the work of Mark Rieffel).
Some more information is in the original lecture notes for the course, as held at TUM in 2017 or 2018.

63.1 Spline Type spaces

Spline type spaces generalize, in some sense, the idea of band-limited functions. Instead of taking the orthonormal basis obtained from the shifted SINC-functions (with all of their properties) one takes other functions, such as the so-called *B-splines*. Here the letter B stands for basic spline functions, because for any degree on can use the collection of shifted copies $(T_k\psi)_{k\in\mathbb{Z}^d}$ as a *Riesz basis* for their linear span, the space of spline functions with nodes at the integers (for $d = 1$, and tensor products thereof for $d \geq 2$). The most important case is the family of *cubic splines* which are C^2-functions on $\mathbb{R}$ (meaning two times continuously differentiable functions) which are represented by cubic polynomials on each of the intervals $[n, n+1]$.
The orthogonal projection of any given function in $\boldsymbol{L}^2(\mathbb{R}^d)$ can be computed efficiently with the help of the *biorthogonal family* which turns out to be of the same form, i.e. $(T_k\tilde{\psi})_{k\in\mathbb{Z}^d}$.

64 Balian-Low Theorem

Lapped transforms, **bofrgrhl96:** [?], **bofrgrhl97:** [?].
The Balian-Low Theorem states, that there is no *Riesz basis* which is a Gabor family with $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ (or even $g \in \boldsymbol{W}(\mathbb{R}^d)$). The only simple system, which even allows to generate a an ONB (orthonormal basis) for $\big(\boldsymbol{L}^2(\mathbb{R}), \|\cdot\|_2\big)$, starting from $g = \mathbf{1}_{[-1/2,1/2]}$, with the TF-lattice $\mathbb{Z} \times \mathbb{Z} \lhd \mathbb{R} \times \widehat{\mathbb{R}}$ has the disadvantage of *poor decay* of SINC $= \mathcal{F}(\mathbf{1}_{[-1/2,1/2]})$ in the frequency domain, hence bad TF-locality of the corresponding Gabor system. In other words, the decay (or concentration of the corresponding Gabor coefficients) of the signal will be not good even if the analyzed function is a smooth function.
In the finite dimensional setting the situation becomes clear by an experiment of the following type: Take a critical Gabor lattice, i.e. a TF-lattice with redundancy 1, i.e. with $b = n/a$, so that the lattice has n (signal length) lattice points, and the corresponding

xcarbp1.jpg

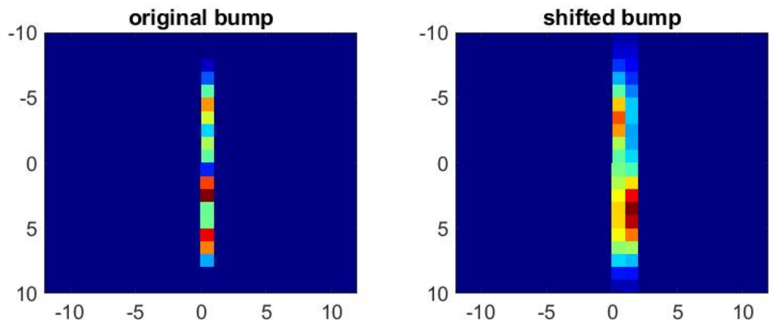

Abbildung 93: stftboxcarbp1.jpg

Gabor matrix `gabbasp(g,a,b)` is a square matrix (e.g. $n = 480, a = 20, b = 24$), and $g = gaussnk(n)$; Then it turns out that not only that the matrix (and hence the frame operator) is not invertible, but simply that there is on dimension missing in the linear span, namely the alternating Dirac comb located at the shifted lattice $10 : a : n$. Moreover, the generator of the dual family (obtained by the command `pinv(gabbasp(g,a,n/a))` is badly localized.
For the continuous case this corresponds to the case $g = g_0, a = b = 1$ (as it was proposed by D. Gabor in 1946), with some very natural (but not strictly valid) argumentation, why the case $ab > 1$ should be avoided (correct: the closed linear span is all of $\left(\boldsymbol{L}^2(\mathbb{R}), \|\cdot\|_2\right)$) and $ab < 1$ as well, to avoid redundancy (no problem, one has to take the minimal $\boldsymbol{\ell}^2$-norm choice for the coefficients to make them unique). So D. Gabor was not just unable to *prove* his conjecture, but it turned out that he was too optimistic, namely that his "critical Gabor system" could be a (well localized) Riesz basis. Even if it was such a system he overlooked that (as it turn out) the question of finding coefficients might be more and more non-local the closer one gets to critical scheme.

64.1 Wilson Basis

The alternative to critical Gabor families is to use so-called *Wilson bases* (related to Malvar's *lapped transform*, or local cosine bases.
The idea in the background is to start from a tight Gabor frame of redundancy 2 and then create linear combinations (very much like Euler's formula used complex exponentials in order to build cos and sin-functions, which kind of have the same frequency. The *trick* is

secrit02.jpg

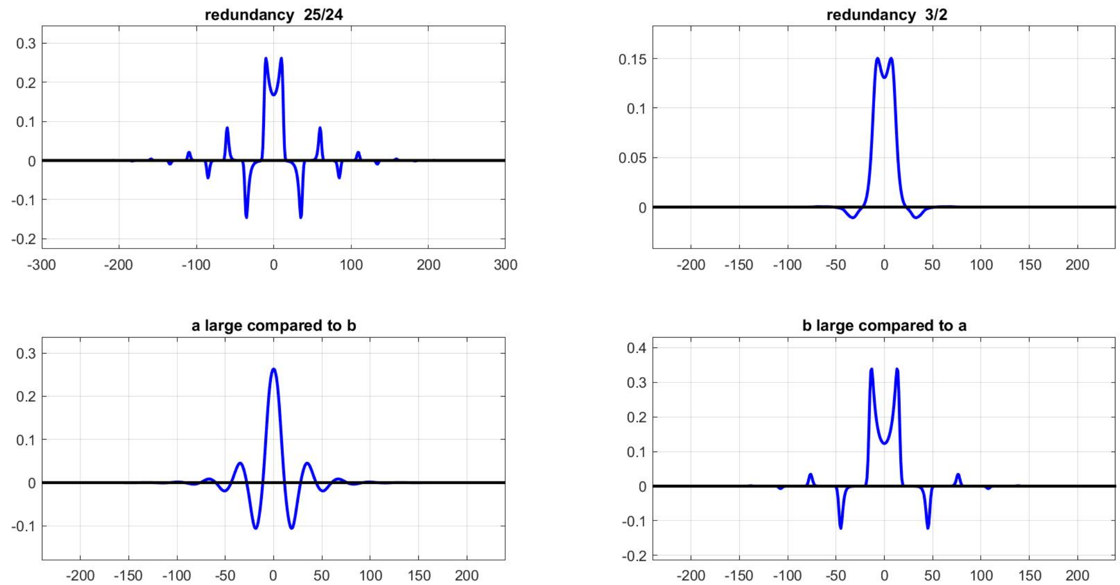

Abbildung 94: closecrit02.jpg

to combine them pairwise **alternating** in adjacent columns. Despite the fact that each building block has still a clear frequency the building blocks have (in the TF-sense) to constituents which are further and further apart, as the frequencies increase.
Nevertheless such bases have been used in the detection of *gravitational waves* (which resulted in the Physics Nobel Prize in the year 2017).
We have done implementations of discrete Wilson bases already in 1996/97 (**bofegrhl96:** [5],**bofegrhl97:** [6]).
There is a number of papers working on Wilson bases:
The first and most influential paper on WILSON BASIS was written by Daubechies, Jaffard and Journe ([16]).
It was taken up to be useful as unconditional bases for modulation spaces in [39].
There is a very interesting and general paper by G. Kutyniok and T. Strohmer ([71]). They show how one can generated Wilson orthonormal bases from tight Gabor frames generated in a symplectic way. For non-symplectic lattices not much seems to be known.

65 Weight functions and weighted spaces

A short summary of definitions if weight functions as used since my thesis:

`Download http://univie.ac.at/nuhag-php/bibtex/open_files/1407_9550001.pdf`

65.1 Submultiplicative, moderate, weakly subadditive weight functions.

There are different types of weight functions. For our purpose it is not problem to allow only *continuous* and *strictly positive* weight functions (in fact: for all the cases of interest we can replace a (moderate) weight function by and *equivalent* having these mild extra properties, while still defining the same space.
A weight function v on $\mathbb{R}^{2d}$ is called **submultiplicative** or a *Beurling weight* (according to [78, 79]) , if

$$v(z_1 + z_2) \leq v(z_1)v(z_2), \tag{462}$$

submultdef

for all $z_1 = (x_1, \xi_1), z_2 = (x_2, \xi_2) \in \mathbb{R}^{2d}$.
A weight function w on $\mathbb{R}^{2d}$ is called *v-moderate* (or moderate with respect to the submultiplicative weight function v), if

$$w(z_1 + z_2) \leq cv(z_1)w(z_2)(*) \tag{463}$$

for all $z_1 = (x_1, \xi_1), z_2 = (x_2, \xi_2) \in \mathbb{R}^{2d}$.
Moderate weights have been introduced more or less by R. Edwards and G. Gaudry for the study of translation invariant weighted $\boldsymbol{L}^p$-spaces ([19], [54]).
Two weights w_1 and w_2 are called *equivalent*, written $w_1 \asymp w_2$, if

$$C^{-1}w_1(z) \leq w_2(z) \leq Cw_1(z) \tag{464}$$

for all $z = (x, \xi) \in \mathbb{R}^{2d}$ and some positive constant C.
It is well known that any moderate weight function is equivalent to a weight function which is continuous and strictly positive, this is why we can restrict our attention (which is mostly motivated by the definition of suitable weighted spaces) to such weight functions.
DEFINITIONS HAVE BEEN TYPED FOR TFD!

WSA-def **Definition 26.** A weight function w is called *weakly subadditive* (WSA) if there exists some $C > 0$ such that

$$w(x + y) \leq C \cdot [w(x) + w(y)] \quad \forall x, y \in G. \tag{465}$$

WSA-def1

Remark 56. A radial symmetric function w, i.e. a function of the form $w(x) = h(|x|)$ for some strictly positive function h on $[0, \infty)$ is

weakly submultiplicative (at least up to equivalence), if it satisfies

$$h(2r) \leq h(r)^2, \quad \forall r > 0;$$

- weakly subadditive if (and only if) it satisfies

$$h(2r) \leq C\, h(r) \quad \forall r > 0.$$

Note: there are some inconsistencies between the usage symbols for weight functions between the work of Feichtinger resp. Gröchenig. While hgfei tried to restrict the use of the symbol w to submultiplicative (Beurling) weights, and rather use the letter m, m_1, m_2 for moderate functions in Gröchenigs work the role of the (general, moderate) weight m ("m" for **m**oderate) has become w ("w" for **w**eight) and most often the corresponding submultiplicative weight is denoted by v. Of course the Japanese brackets or polynomial weights $\langle y\rangle^s = (1+|y|^2)^{s/2}$ is such as submultiplicative weight of *polynomial growth* (for any $s \geq 0$, which is of course a WSA weight).

Personally I consider the paper of Brandenburg as one of the most interesting (and relatively simple) ones, providing some statements about spectral radius of elements in weighted spaces with WSA weights! [7]
Each of those weights satisfies the so-called (BD), *Beurling-Domar condition* ([78], Chap.3.1., p.132), namely

$$\sum_{n\geq 1} \frac{w(nx)}{n^2} < \infty \quad \forall x \in \mathbb{R}^d. \tag{466}$$ BDcond00

Hence it is clear that compactly supported elements are dense in the corresponding *Fourier-Beurling algebra*, defined as $\mathcal{F}(\boldsymbol{L}^1_w(G))$.
[17] , [78] (Beurling algebras), Beurling Domar condition (BD) [79]: updated edition of H. Reiter's book (by his former student I. Stegeman).
Two important consequences of the conditions on the weights are:

submconv **Lemma 111.** *Fix any SM weight (or Beurling weight).*
Assume that two measurable functions f, g are given such that the pointwise convolution product of fw with gw exists. Then $f \star g$ is well defined and we have (pointwise a.e.) the following estimate:

$$|(f\star g)(x)w(x)| \leq [|fw| \star |gw|](x). \tag{467}$$ subconv

In particular, for any solid Banach convolution algebra $\boldsymbol{A}$ (meaning that $g \in \boldsymbol{A}$ and $|f(x)| \leq |g(x)|$ a.e. implies that $f \in \boldsymbol{A}$ and $\|f\|_{\boldsymbol{A}} \leq \|g\|_{\boldsymbol{A}}$) then also

$$\boldsymbol{A}_w := \{f \mid f\cdot w \in \boldsymbol{A}\}$$

is a (solid) Banach convolution algebra with respect to the norm $\|f\|_{\boldsymbol{A}_w} := \|fw\|_{\boldsymbol{A}}$.

By specializing to $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}}) := \left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ we get:

BeurlAlgebra **Corollary 16.** *For any submultiplicative weight w the weighted $\boldsymbol{L}^1$-space $\left(\boldsymbol{L}^1_w(\mathbb{R}^d), \|\cdot\|_{1,w}\right)$ is a (solid) Banach algebra with respect to convolution, called Beurling algebra.*

For WSA weights w we have the following result:

WSAconv **Proposition 24.** *For any WSA weight w we have: $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$, with $\boldsymbol{B} := \boldsymbol{L}^1(\mathbb{R}^d) \cap \boldsymbol{L}^p_w(\mathbb{R}^d)$ and natural norm $\|f\|_{\boldsymbol{B}} = \|f\|_{\boldsymbol{L}^1} + \|fw\|_{\boldsymbol{L}^p}$ is a Banach algebra with respect to convolution.*

Proof. The crucial *pointwise estimate* is the following one:

$$|f \star g) \cdot w|(x) \leq C(|f| \star |gw| + |fw| \star |g|)(x), \quad x \in \mathbb{R}^d. \tag{468}$$ WSAestim02

In fact, we just have to split the weight factor and use that

$$w(x) \leq C\left(w(x-y) + w(y)\right), \quad x, y \in \mathbb{R}^d.$$

Since the estimate

$$\|[f \star g]w\|_{\boldsymbol{L}^1_w(\mathbb{R}^d)} \leq \|[fw\|_{\boldsymbol{L}^1_w(\mathbb{R}^d)} \|gw\|_{\boldsymbol{L}^1_w(\mathbb{R}^d)}$$

is valid due to the Beurling weight property we get the estimate from the observation that

$$\|[f \star g]w\|_{\boldsymbol{L}^p} \leq C\left(\|f\|_{\boldsymbol{L}^1}\|fw\|_{\boldsymbol{L}^p} + \|g\|_{\boldsymbol{L}^1}\|gw\|_{\boldsymbol{L}^p}\right).$$

□

66 Sobolev algebras

Sobolev spaces have been introduced in analysis in order to describe the smoothness of functions. One can start with concepts such as *continuity* or *differentiability*.
The most natural way to describe increasing smoothness is the number of continuous derivatives a functions has. For example, we call a function $f(x)$ on $\mathbb{R}$ a $C^{(k)}$-function for some $k \in \mathbb{N}$ if f, f', f'' etc. up to the derivative of order k, written as $f^{(k)}$ are all continuous functions.
If we look into B-splines (the convolution powers of the boxcar functions $\mathbf{1}_{[-1/2,1/2]}$) we find that the B-spline of order 2 (or degree 1), i.e. the triangular function Δ, is continuous but not everywhere differentiable, hence it is in $\boldsymbol{C}_c(\mathbb{R}^d)$ resp. $C^{(0)}(\mathbb{R})$. The next one is continuously differentiable (but only first order), and locally glued together by quadratic polynomials. Finally cubic splines are constituted by locally cubic polynomials (degree 3, order 4), which have continuous derivatives up to order 2 (continuous *curvature*), so they constitute $C^{(2)}$-functions and are thus very suitable to plot curves.
For $\boldsymbol{L}^2(\mathbb{R}^d)$ one can has a similar concept, namely, to assume that the derivative exists almost everywhere and that the class constituted in this way also belongs to $\boldsymbol{L}^2(\mathbb{R})$, and so on (by induction).
Looking at the behaviour of the derivative operation (which is of course a [unbounded] translation invariant procedure) we expect that it a multiplication operator on the Fourier transform side, and looking into the derivative of $x \mapsto \chi_s(x) = exp(2\pi i s x)$ is of course (by the exponential law) $\chi_s'(x) = 2\pi i s \chi_s(x)$, we may not be surprised that the standard formula (for our normalization of the FT) is:

ferentiation **Lemma 112.** *Assume that $f \in \boldsymbol{L}^1(\mathbb{R}^d) \cap \boldsymbol{C}_0(\mathbb{R}^d)$ has a continuous, integrable derivative (i.e. that f is an absolutely continuous functions (in the classical sense), then we have*

$$\mathcal{F}(f')(s) = 2\pi\, i\, s\, \widehat{f}(s), \quad s \in \mathbb{R}. \tag{469}$$ derivFT01

In the context of $\left(\boldsymbol{L}^2(\mathbb{R}), \|\cdot\|_2\right)$ one can describe therefore the number of derivatives by the square integrability of the corresponding derivatives in $\boldsymbol{L}^2(\mathbb{R})$, and thus by the membership of $\widehat{f}$ in some corresponding weighted $\boldsymbol{L}^2$-space on $\mathbb{R}$. This is how *Sobolev spaces* can be introduced via the Fourier transform as weighted $\boldsymbol{L}^2$-spaces. This also allows to talk about *fractional differentiability*!
weakly subadditivity and submultiplicativity >> Beurling algebras, see TEXBLOCK on weight functions ([21] or English summary in [20])
We are using the weights $v_s(y) := \langle y \rangle^s = (1+|y|^2)^{s/2}$ for $s \geq 0$ and define *Sobolev spaces* (of positive order $s \geq 0$) as subspaces of $\boldsymbol{L}^2(\mathbb{R}^d)$ in the following way

HsRddef **Definition 27.**

$$\mathcal{H}_s(\mathbb{R}^d) := \{f \mid \widehat{f} \cdot v_s \in \boldsymbol{L}^2(\mathbb{R}^d)\} \tag{470}$$ HSRddef00

with the norm $\|f\|_{\mathcal{H}_s(\mathbb{R}^d)} := \|\widehat{f}v_s\|_{\boldsymbol{L}^2(\mathbb{R}^d)}$.

Proof. We write $\boldsymbol{L}^2_w(\mathbb{R}^d)$ for the weighted space

$$\{g \mid g \cdot v_s \in \boldsymbol{L}^2(\mathbb{R}^d)\}.$$

Observing that $1/v_s \in \boldsymbol{L}^2(\mathbb{R}^d)$ if (and only if) $s > d/2$ on finds that $\boldsymbol{L}^2_w(\mathbb{R}^d) \hookrightarrow \boldsymbol{L}^1(\mathbb{R}^d)$ via Cauchy-Schwartz, writing $h \in \boldsymbol{L}^2_w(\mathbb{R}^d)$ as a product of hw with $1/w$:

$$\|h\|_1 = \|hw \cdot 1/w\|_1 \leq \|hw\|_2 \cdot \|1/w\|_2. \tag{471}$$ LtwinLi

Applying the Fourier transform on both sides we obtain that $\left(\mathcal{H}_s(\mathbb{R}^d), \|\cdot\|_{\mathcal{H}_s}\right) \hookrightarrow \left(\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_{\mathcal{F}\boldsymbol{L}^1}\right) \hookrightarrow \left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$, with the natural norm estimates. □

Note that the above argument shows that $\mathcal{H}_s(\mathbb{R}^d)$ is not only continuously embedded into $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$ (according to Plancherel, since $\boldsymbol{L}^2_w(\mathbb{R}^d) \hookrightarrow \boldsymbol{L}^2(\mathbb{R}^d)$), but also in $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$, hence into $\boldsymbol{L}^2 \cap \boldsymbol{C}_0(\mathbb{R}^d)$ with the natural norm (sum of the two norms).
Using the so-called theory of Wiener amalgam spaces on can show that $\mathcal{H}_s(\mathbb{R}^d) \hookrightarrow \boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^2)(\mathbb{R}^d)$, by the argument that $\boldsymbol{L}^2_w(\mathbb{R}^d) = \boldsymbol{W}(\boldsymbol{L}^2, \boldsymbol{\ell}^2_w) = \boldsymbol{W}(\mathcal{F}\boldsymbol{L}^2, \boldsymbol{\ell}^2_w) \hookrightarrow \boldsymbol{W}(\mathcal{F}\boldsymbol{L}^2, \boldsymbol{\ell}^1)$ (using again Cauchy Schwartz in the last inclusion), which by a variant of the Hausdorff-Young inequality for Wiener amalgam spaces gives

$$\mathcal{H}_s(\mathbb{R}^d) = \mathcal{F}^{-1}\,\boldsymbol{L}^2_w \hookrightarrow \mathcal{F}^{-1}\,\boldsymbol{W}(\mathcal{F}\boldsymbol{L}^2, \boldsymbol{\ell}^1) \hookrightarrow \boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1, \boldsymbol{\ell}^2) \hookrightarrow \boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^2)(\mathbb{R}^d),$$

which is a space strictly contained in $\boldsymbol{L}^2 \cap \boldsymbol{C}_0(\mathbb{R}^d)$.
Note that $\mathcal{H}_s(\mathbb{R}^d)$ is not only a Banach space with respect to the natural norm $\|f\|_{\mathcal{H}_s} := \|\hat{f}w\|_2$, but even a Hilbert space, because the norm obviously comes from the scalar product

$$\langle f, g\rangle_{\mathcal{H}_s} := \langle \hat{f}w, \hat{g}w\rangle_{\boldsymbol{L}^2} = \int_{\mathbb{R}^d} \hat{f}(s)\,\overline{\hat{g}(s)}\; w^2(s)ds. \tag{472}$$ HS-scalprod

Hs-RKHS **Corollary 17.** *$\mathcal{H}_s(\mathbb{R}^d)$ is a reproducing kernel Hilbert space, i.e. for each $t \in \mathbb{R}^d$ the Dirac measure $\delta_x : f \mapsto f(x)$ is a continuous linear functional on the Hilbert space $\mathcal{H}_s(\mathbb{R}^d)$. The kernel $K(x,y)$, consisting of functions $K(\cdot, y) = k_x(y)$ in $\mathcal{H}_s(\mathbb{R}^d)$ with*

$$f(x) = \langle f, k_x\rangle_{\mathcal{H}_s} \quad \textit{for all} \quad x \in \mathbb{R}^d, f \in \mathcal{H}_s \tag{473}$$ Hs-RKHSphi

is obtain as collections of shifts $T_x\varphi$, with $\varphi = \mathcal{F}^{-1}(1/w^2)$.

Proof. Since $\mathcal{H}_s(\mathbb{R}^d)$ is continuously embedded into $\left(\boldsymbol{C}_0(\mathbb{R}^d),\ \|\cdot\|_\infty\right)$ it is clear that the family of point measures acts even *uniformly bounded* on $\mathcal{H}_s(\mathbb{R}^d)$. Hence we only have to prove the explicit representation of δ_0 on $\mathcal{H}_s(\mathbb{R}^d)$ and then (using the definitions) the covariance of the situation: shifting the point evaluation (δ_0 to δ_x) corresponds to shift the generator representing φ. We get the representation using the Fourier inversion formula (it is easy to recognize that we do not need that f itself is in $\boldsymbol{L}^1(\mathbb{R}^d)$, the inverse Fourier transform a priori defined as an $\boldsymbol{L}^2$-FT has the usual form as integral if $\hat{f} \in \boldsymbol{L}^1(\mathbb{R}^d)$, which is true for any $f \in \mathcal{H}_s(\mathbb{R}^d)$, if $s > d/2$):

$$f(0) = \int_{\mathbb{R}^d} \hat{f}(s)ds = \int_{\mathbb{R}^d} \hat{f}(s)w(s) \cdot \frac{1}{w^2}(s)w(s)ds =: \int_{\mathbb{R}^d} \hat{f}(s)\,\overline{\hat{\varphi}(s)}\ w^2(s)ds, \qquad (474)$$ Hs-delta0rep

if we put $\varphi = \mathcal{F}^{-1}(1/w^2)$.
As a matter of fact we have $\langle y \rangle^r \in \boldsymbol{S}_0(\mathbb{R}^d)$ if (and only if) $r < -d/2$.
Now we can apply the shift invariance of the scalar product (exercise):

$$\langle T_x f, T_x h\rangle_{\mathcal{H}_s} = \langle f, h\rangle_{\mathcal{H}_s} \quad \text{for all} \quad f, h \in \mathcal{H}_s(\mathbb{R}^d), \qquad (475)$$ HS-scalshift

because we know that translation goes to modulation on the Fourier transform side, but having the same modulations within a scalar products means (due to the fact that one is taking a conjugation in the scalar product) that it is unchanged. Technically speaking one could argue that translation goes into modulation, but for *all weights* modulations are unitary on the corresponding weighted $\boldsymbol{L}^2$-spaces. $\square$

The pointwise algebra properties for Sobolev spaces is a consequence of the WSA property of $\langle y \rangle^s$ (for non-negative s) and the fact that one has for $s > d/2$ that $v_{-s} \in \boldsymbol{L}^2(\mathbb{R}^d)$, which implies, with the help of the Cauchy Schwartz inequality that $\boldsymbol{L}^2_{v_s}(\mathbb{R}^d) \hookrightarrow \boldsymbol{L}^1(\mathbb{R}^d)$, or equivalently $\boldsymbol{L}^1(\mathbb{R}^d) \cap \boldsymbol{L}^2_{v_s}(\mathbb{R}^d) = \boldsymbol{L}^2_{v_s}(\mathbb{R}^d)$ (with equivalence of the natural norms). But by
Probably some more details can be found in the paper on robustness ([46]), where it is shown, that somehow the reproducing kernels depend in a continuous way on the weight, e.g. on the parameter s for the case of the weights $w = w_s$ on $\mathbb{R}^d$.
Different choices of weights have been treated in the (recent) PhD thesis of Rosa Aceska, treating the case of *variable smoothness* ([1]).

For references see [20, 21, 46, 60].

67 Convolution operators on $\boldsymbol{S}_0(\mathbb{R}^d)$

THIS MATERIAL is taken from the last section of **fe19: [33]**.
So at the end of this section let us take a look at multipliers (operators commuting with translations) defined on $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$. Since we have

$$\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right) \hookrightarrow \left(\boldsymbol{L}^p(\mathbb{R}^d),\ \|\cdot\|_p\right) \hookrightarrow (\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$$

for $1 \le p \le \infty$ let us consider operators from $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$ into $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$.
Using mild distributions we can provide the following characterization, so the say the most general multipliers on $\boldsymbol{S}_0(\mathbb{R}^d)$.

armultSOSOP **Theorem 26.** *A bounded linear operator from* $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ *to* $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$ *commutes with translation if and only if it is a convolution operator by some (uniquely determined) mild distribution* $\sigma \in \boldsymbol{S}_0'$, *i.e. for* $x \in \mathbb{R}^d$*:*

$$Tf(x) = \sigma \star f(x) = \sigma(T_x f^{\checkmark}), \quad \forall f \in \boldsymbol{S}_0(\mathbb{R}^d). \tag{476}$$ convSOSOP1

In fact, one has

$$\boldsymbol{H}_{\mathbb{R}^d}(\boldsymbol{S}_0, \boldsymbol{S}_0') = \boldsymbol{H}_{\mathbb{R}^d}(\boldsymbol{S}_0, \boldsymbol{C}_b)$$

which allows to obtain σ *from* T *by the rule* $\sigma(f) = T(f^{\checkmark})(0)$. *Moreover, one has equivalence of norms:*

$$\|T\|_{\boldsymbol{S}_0 \to \boldsymbol{S}_0'} \approx \|\sigma\|_{\boldsymbol{S}_0'}.$$

Proof. Due to the Fourier invariance of $\boldsymbol{S}_0(\mathbb{R}^d)$ we have: $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ if and only if $\widehat{f} \in \boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1, \boldsymbol{\ell}^1)(\mathbb{R}^d)$, i.e. if and only if the sum

$$\widehat{f} = \sum_{k \in \mathbb{Z}^d} (T_k \psi) \cdot \widehat{f}$$

is absolutely convergent in $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$, but since $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ belongs to the pointwise multipliers of $\boldsymbol{S}_0(\mathbb{R}^d) = \boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1, \boldsymbol{\ell}^1)(\mathbb{R}^d)$ it is easy to show that the sum is actually absolutely convergent in $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$.
Choosing any $\psi^* \in \boldsymbol{S}_0(\mathbb{R}^d)$ (e.g. $\psi^* \in \boldsymbol{\mathcal{S}}(\mathbb{R}^d)$) such that $\psi(s) \cdot \psi^*(s) = \psi(s)$ for all $s \in \mathbb{R}^d$ one finds that

$$\|T_k\psi \cdot \widehat{f}\|_{\boldsymbol{S}_0(\mathbb{R}^d)} = \|T_k\psi^* \cdot T_k\psi \cdot \widehat{f}\|_{\boldsymbol{S}_0(\mathbb{R}^d)} \leq \|T_k\psi^*\|_{\boldsymbol{S}_0} \|T_k\psi \cdot \widehat{f}\|_{\mathcal{F}\boldsymbol{L}^1}$$

and consequently one has for any $f \in \boldsymbol{S}_0(\mathbb{R}^d)$:

$$\sum_{k \in \mathbb{Z}^d} \|(T_k\psi) \cdot \widehat{f}\|_{\boldsymbol{S}_0} \leq \|\psi^*\|_{\boldsymbol{S}_0} \sum_{k \in \mathbb{Z}^d} \|(T_k\psi) \cdot \widehat{f}\|_{\mathcal{F}\boldsymbol{L}^1} \leq C\|\psi^*\|_{\boldsymbol{S}_0} \|f\|_{\boldsymbol{S}_0}.$$

Note that $\varphi^* := \mathcal{F}^{-1}\psi^*$ also belongs to $\boldsymbol{S}_0(\mathbb{R}^d)$ and that we have now

$$f = \sum_{k \in \mathbb{Z}^d} M_k\varphi^* \star M_k\varphi \star f.$$

Given $T \in \mathcal{L}(\boldsymbol{S}_0, \boldsymbol{S}_0')$, a bounded linear operator which commutes with translation and hence with convolution by $M_k\varphi^*$ for each $k \in \mathbb{Z}^d$, we obtain:

$$T(f) = \sum_{k \in \mathbb{Z}^d} M_k\varphi^* \star T[M_k\varphi \star f], \tag{477}$$ 114

and further

$$\|T(f)\|_\infty \leq \|\varphi^*\|_{\boldsymbol{S}_0} |||T|||_{\boldsymbol{S}_0 \to \boldsymbol{S}_0'} \sum_{k \in \mathbb{Z}^d} \sum_{n \in \mathbb{Z}^d} \|M_k\varphi \star f\|_{\boldsymbol{S}_0}, \tag{478}$$ 115

or altogether: For some constant $C > 0$ one has

$$\|T(f)\|_\infty \leq C|||T|||_{\boldsymbol{S}_0 \to \boldsymbol{S}_0'} \|f\|_{\boldsymbol{S}_0}, \quad \forall f \in \boldsymbol{S}_0. \tag{479}$$ 116

Since we have $\|T_x f - f\|_{\boldsymbol{S}_0(\mathbb{R}^d)} \to 0$ for $x \to 0$ for any $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ this also implies for each $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ the *uniform continuity* of Tf by the following argument:

$$\|T_x(Tf) - Tf\|_\infty = \|T(T_x f - f)\|_\infty \to 0, \quad \text{as } x \to 0.$$

The rest of the argument is now the usual procedure: We just have to check that $\sigma(f) := T(f^\checkmark)(0)$ is a well defined linear function on $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ and in fact can be used to *represent* the linear operator. We leave details to the interested reader. $\square$

As an immediate corollary to this result we have the following consequences for multipliers between (potentially different) $\boldsymbol{L}^p$-spaces over $\mathbb{R}^d$:

`LpLqconv1` **Theorem 27.** *Given $p \in [1, \infty)$ and $q \in [1, \infty]$. Then any bounded linear operator T from $\left(\boldsymbol{L}^p(\mathbb{R}^d), \|\cdot\|_p\right)$ to $\left(\boldsymbol{L}^q(\mathbb{R}^d), \|\cdot\|_q\right)$ which commutes with all translations $T_y, y \in \mathbb{R}^d$ is a convolution operator by a uniquely determined $\sigma \in \boldsymbol{S}_0'$, meaning that*

$$Tf(x) = \sigma \star f(x) = \sigma(T_x f^\checkmark), \;\; f \in \boldsymbol{S}_0(\mathbb{R}^d), x \in \mathbb{R}^d.$$

Equivalently, T can be represented as Fourier multiplier:

$$\widehat{Tf} = \tau \cdot \widehat{f}, \quad f \in \boldsymbol{S}_0(\mathbb{R}^d)$$

for $\tau = \widehat{\sigma} \in \boldsymbol{S}_0'(\mathbb{R}^d)$ and vice versa (with $\sigma = \mathcal{F}^{-1}(\tau)$).

Note that we do not claim that the representations are valid for all $f \in \boldsymbol{L}^p(\mathbb{R}^d)$. Of course it is enough to explicitly describe the action of a bounded linear operator on a dense subspace, while the extension to the full Banach space is carried out by approximation ($\boldsymbol{S}_0(\mathbb{R}^d)$ is dense in $\left(\boldsymbol{L}^p(\mathbb{R}^d), \|\cdot\|_p\right)$ for $p < \infty$).
An important example for a non-trivial Fourier multiplier is the chirp signal $ch(t) = exp(i\pi x^2) \in \boldsymbol{C}_b(\mathbb{R})$. It has the interesting property of being mapped onto itself by the (generalized) Fourier transform. Although $SINC := \mathcal{F}^{-1}(box)$ evidently belongs to $\boldsymbol{L}^2(\mathbb{R})$ but not to $\boldsymbol{L}^1(\mathbb{R})$ it is *not possible* to describe the action of $\sigma = ch$ via convolution on the SINC function by a pointwise integral, because $\int_{\mathbb{R}} ch(t) T_x(SINC)(t)dt$ does not exist (in the Lebesgue sense) for any $x \in \mathbb{R}$.
There is another result, which is more similar to the situation for $\boldsymbol{L}^2(\mathbb{R}^d)$:

`convS0Rd` **Proposition 25.** *Any bounded linear operator on $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ which commutes with translations is a Fourier multiplier by some $h \in \boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1, \boldsymbol{\ell}^\infty)(\mathbb{R}^d)$, i.e. is of the form $\widehat{f} \mapsto \widehat{Tf} = h \cdot \widehat{f}$.*
Moreover, the operator norm of T on $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ and the Wiener amalgam norm of h, namely the expression

$$\sup_{k \in \mathbb{Z}^d} \|h \cdot T_k \psi\|_{\mathcal{F}\boldsymbol{L}^1} = \sup_{k \in \mathbb{Z}^d} \|\sigma \star M_k \varphi\|_\infty$$

define equivalent norms.

This result follows from a general principle concerning pointwise multipliers of decomposition spaces (specialized to the setting of Wiener amalgams), see [37], or [45].

67.1 Summary concerning TILS

Let us recall: We have started with the consideration of linear systems, i.e. with linear operators commuting with translation. Our first model was the BIBO-assumption: in the mathematical formulation we made the assumption that T maps also $\left(\boldsymbol{C}_0(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$ into itself. The first main result was a characterization of such BIBO systems as moving average with the help of some continuous/bounded linear functional on the space $\left(\boldsymbol{C}_0(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$, we called these bounded linear functionals *bounded measures*.
Then we went on and found out that the adjoint action of a convolution is just the internal convolution (up to the flip operator). Since any adjoint operator is not only norm-to-norm continuous on the dual spaces (here now from $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ into itself) but also w^*-w^*-continuous we found out that the usual identification of the *impulse response*, i.e. the limit (in whatever sense) of the output to any Dirac sequence is justified: It really does not depend on the "shape" of the input (be it a sequence of rectangular, triangular or SINC^2 impulse, as long as they are (Riemann) integrable with $\widehat{g}(0) = 1$ they would have δ_0 as the limit of the family $\mathrm{St}_\rho g, \rho \to 0$, taken in the w^*–sense in $\boldsymbol{M}_b(\mathbb{R}^d)$.
In the mathematical literature there is a well-known result that shows that also the linear operators on $\left(\boldsymbol{L}^1(\mathbb{R}^d),\, \|\cdot\|_1\right)$ which commute with translations, can be characterized as well as convolution operators with bounded measures. This is known as Wendel's Theorem (see the introduction to **la71:** [72]), but this is very close to BIBOS system, just starting from the $\left(\boldsymbol{L}^1(\mathbb{R}^d),\, \|\cdot\|_1\right)$ side. In fact, the adjoint operation defines a w^*-w^*--continuous operator from $\boldsymbol{L}^\infty(\mathbb{R}^d)$ into itself which is therefore well defined on $\left(\boldsymbol{C}_0(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$. By showing that $\left(\boldsymbol{C}_0(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$ is w^*–dense in $\boldsymbol{L}^\infty(\mathbb{R}^d)$ (or rather $(\boldsymbol{L}^1)^*(\mathbb{R}^d)$) one can reduce the domain of the adjoint operator to $\left(\boldsymbol{C}_0(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$, ensuring that it maps at least $\left(\boldsymbol{C}_0(\mathbb{R}^d),\, \|\cdot\|_\infty\right)$ into $\boldsymbol{C}_{ub}(\mathbb{R}^d)$. This is enough to go back to the situation of BIBO systems, but we will not go deeper into this direction here.
But now we also can cover many other cases. Most importantly we can now fully discuss the case of operators from $\left(\boldsymbol{L}^p(\mathbb{R}^d),\, \|\cdot\|_p\right)$ into $\left(\boldsymbol{L}^q(\mathbb{R}^d),\, \|\cdot\|_q\right)$ (for any pair of parameters $1 \leq p < \infty, 1 \leq q \leq \infty$.) Let us just look into this case for $p = 2 = q$. In this case the classical results is:

LtTILS **Proposition 26.** *Any $T \in \boldsymbol{H}_{\mathbb{R}^d}(\boldsymbol{L}^2(\mathbb{R}^d), \boldsymbol{L}^2(\mathbb{R}^d))$ can be described as a pointwise multiplier with some $h \in \boldsymbol{L}^\infty(\mathbb{R}^d)$:*

$$\widehat{Tf} = h \cdot \widehat{f}, \quad f \in \boldsymbol{L}^2(\mathbb{R}^d)$$

(where the FT is taken in the sense of the Plancherel Theorem).

An important example is the chirp signal $h = \mathrm{ch}(s) = exp^{i\pi s^2}$. This is known to be *invariant under the extended FT* (defined on $\boldsymbol{S}_0'(\mathbb{R}^d)$)!.
In this case w^*–convergence has to be taken in the sense of the duality of the pair $(\boldsymbol{S}_0(\mathbb{R}^d), \boldsymbol{S}_0'(\mathbb{R}^d))$.

68 Sampling and Periodization

We have seen that sampling in the time-domain corresponds to periodization in the frequency domain and conversely: periodization in the time (or frequency) domain corresponds to sampling in the frequency (or time) domain.
If the sampling rate and the periodization are compatible (rationally related, i.e. their ratio is some rational number, ideally if the period is an integer multiple of the sampling rate then one obtains a discrete and periodic signal from a given function. Clearly we obtain from $f \in \boldsymbol{S}_0(\mathbb{R}^d) \subset \boldsymbol{W}(\mathbb{R}^d)$ a bounded function of the form $⊔⊔_\Lambda \star f \in \boldsymbol{C}_{ub}(\mathbb{R}^d)$ and from that by pointwise multiplication with another Dirac comb an weighted Dirac comb with bounded (because periodic)

period23.jpg

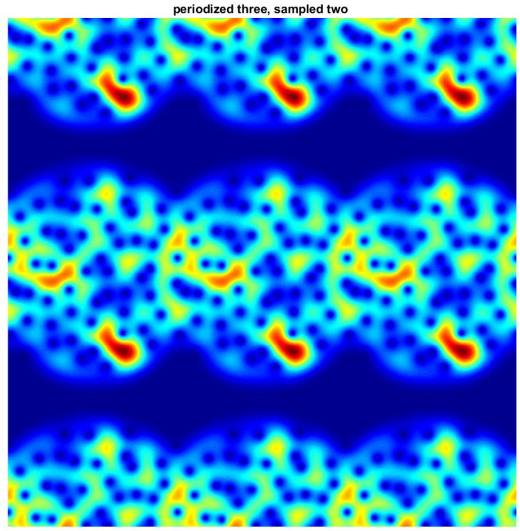

Abbildung 95: sampperiod23.jpg

We have seen in the proof of the Shannon sampling Theorem (see Fig. shanndem4.pdf 51.1) that it is possible to recover a band-limited function from its regular samples, if only the *Nyquist criterion* is satisfied.
If we take a *naive approach* we can hope to (partially recover the function by a simple piecewise interpolation (this is what MATLAB is doing automatically for the standard plot of discrete samples of a given function) or more sophisticated methods (like BUPU-quasi-interpolation, using cubic splines etc.).
The situation is more clear on the "other side", namely how to recover a well concentrated function from its periodic version. Here the ideal case is that the function itself is compact support and a coarse periodization which produces *no overlaps*, as depicted in Fig. locperplat2b.pdf 52.
In the STFT-domain one can see the effect of both of these operations: They are simply periodic repetitions in two directions, and since at the end every function in $\boldsymbol{S}_0(\mathbb{R}^d)$ is a limit of functions which have their STFT (with respect to a Gaussian

eriod23B.jpg

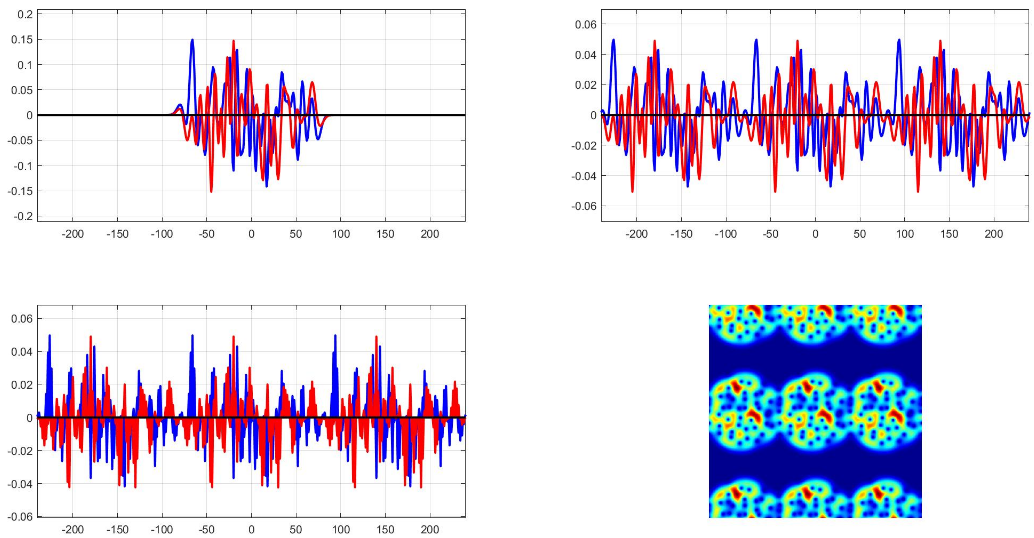

Abbildung 96: sampperiod23B.jpg

window, as shown in the illustrating figure) the observer, who is watching just a big rectangular domain in phase space will see only the STFT of the original function. This is what w^*–convergence is all about: better and better (uniform) convergence of the STFT over larger and larger domains. The same effect can be observed by filtering with larger and larger filters (the effect of summability methods, which essentially reduce the spectrogram to a horizontal stripe). Of course one can also multiply on the time side with a sequence of plateau functions, e.g. $D_\rho \tau$, for $\rho \to 0$. In this case one creates broader and broader horizontal stripes (with full frequency resolution).

68.1 Partially reverting the procedures

We have thus seen, that we can recover perfectly a band-limited function from (sufficiently fine) regular samples, and we can recover a compactly supported function from its (sufficiently coarse) periodization. But there is no non-trivial band-limited function with compact support, so there is always a loss of information.
Still, by taking more and more samples and periodizing more and more coarsely one can try to preserve as much information from an original function. This "intuitive" expectation can be realized for functions in $\boldsymbol{S}_0(\mathbb{R}^d)$ (!), based on the result in **feka07:** [43], see Theorem reconstS0 15 above.

68.2 Periodization on the FT side

The characterization of periodic functions as periodization variants of localized functions corresponds on the "other side" to the fact that any distribution which is supported on a lattice (and has bounded coefficients) is arising by pointwise multiplication of a continuous function with a corresponding Dirac comb.

One direction is clear: If we multiply a function $h \in \boldsymbol{C}_b(\mathbb{R}^d)$ with a Dirac comb $\sqcup\!\sqcup_\Lambda$ we obtain the *weighted Dirac comb* $h \cdot \sqcup\!\sqcup_\Lambda = \sum_{\lambda\in\Lambda} h(\lambda)\delta_\lambda$, and clearly the resulting weighted Dirac comb is supported inside of the lattice Λ. The above characterization allows to go the other way round. Given $\sigma \in \boldsymbol{S}_0'(\mathbb{R}^d)$ with $\mathrm{supp}(\sigma) \subset \Lambda$ one can assure that it arises in exactly this way.

perFTsamp **Proposition 27.** *Assume that $\sigma \in \boldsymbol{S}_0'$ is Λ-periodic for some $\Lambda \lhd \mathbb{R}^d$, i.e.*

$$T_\lambda(\sigma) = \sigma, \quad \lambda{\in}\Lambda,$$

then $\mathrm{spec}(\sigma) \subset \Lambda^\perp$. **Conversely***: If* $\mathrm{spec}(\sigma) \subset \Lambda^\perp$ *then sigma is Λ-periodic.*

Proof. If σ is Λ-periodic, then it can be obtained by periodization from $\sigma\psi$ (where ψ generates a regular Λ-BUPU!). But this means that

$$h_\sigma := \mathcal{F}(\sigma\psi) \in \boldsymbol{S}_0'(\mathbb{R}^d) \star \boldsymbol{S}_0(\mathbb{R}^d) \subset \boldsymbol{C}_{ub}(\mathbb{R}^d)$$

and thus

$$\mathcal{F}(\sigma) = \mathcal{F}(\sqcup\!\sqcup_\Lambda) \star \mathcal{F}(\sigma\psi) = \sqcup\!\sqcup_{\Lambda^\perp} \cdot h_\sigma,$$

which obviously has support in $\Lambda^\perp$.

Conversely, if we have $\mathrm{supp}(\widehat{\sigma}) \subset \Lambda^\perp$ then we can write it as a sum of Dirac measures supported on $\Lambda^\perp$ with bounded coefficients. So we can put shifted triangular functions to those (regular) positions and find some $h \in \boldsymbol{C}_{ub}(\mathbb{R}^d)$ such that $h(\lambda^\perp)$ has the appropriate values. But the most important fact is that we can be sure that multiplication with pure frequencies χ_λ (we are now looking into functions on the frequency domain) leave the Fourier transform invariant, because $\chi_\lambda(\lambda^\perp) = 1$ by definition of the pairing $(\Lambda, \Lambda^\perp)$. But $M_\lambda\widehat{\sigma} = \widehat{\sigma}$ for all $\lambda{\in}\Lambda$ is equivalent to the condition $T_\lambda\sigma = \sigma$ for all $\lambda{\in}\Lambda$. □

69 Kernel Theorem

kernthm001

This is a kind of analogue of the matrix representation of operators for continuous variables, somehow in the spirit of considering the families $(\delta_x)_{x\in\mathbb{R}^d}$ as a *continuous basis*, and equivalently $(\chi_s)_{s\in\mathbb{R}^d}$ as another such basis, with the Fourier transform representing a *change of basis*.

The alternative representation using the *Kohn-Nirenberg symbol* of a linear operator and the (for us even more important) *spreading representation* (first presented in detail in **feko98:** [44]).

The KERNEL THEOREM for $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$

The *kernel theorem* for the Schwartz space can be read as follows:

Theorem 28. *For every continuous linear mapping T from $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ into $\boldsymbol{\mathcal{S}}'(\mathbb{R}^d)$ there exists a unique tempered distribution $\sigma \in \boldsymbol{\mathcal{S}}'(\mathbb{R}^{2d})$ such that*

$$T(f)(g) = \sigma(f \otimes g), \quad f, g \in \boldsymbol{\mathcal{S}}(\mathbb{R}^d). \tag{480}$$ kernelthm00

Conversely, any such $\sigma \in \boldsymbol{\mathcal{S}}'(\mathbb{R}^{2d})$ induces a (unique) operator T such that (480) holds.

The proof of this theorem is based on the fact that $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ is a *nuclear Frechet space*, i.e. has the topology generated by a sequence of semi-norms, can be described by a metric which

turns $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ into a complete metric space.

70 The Kernel Theorem

The KERNEL THEOREM for S_0 I

Tensor products are also most suitable in order to describe the set of all operators with certain mapping properties. The backbone of the corresponding theorems are the *kernel-theorem* which reads as follows (!! despite the fact that $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ is NOT a *nuclear Frechet space*)
One of the corner stones for the kernel theorem is: One of the most important properties of $\boldsymbol{S}_0(\mathbb{R}^d)$ (leading to a characterization given by V. Losert, [76]) is the tensor-product factorization:

0tensor-prod **Lemma 113.**

$$\boldsymbol{S}_0(\mathbb{R}^k)\hat{\otimes}\boldsymbol{S}_0(\mathbb{R}^n) \cong \boldsymbol{S}_0(\mathbb{R}^{k+n}), \tag{481}$$ S0tens-repr

with equivalence of the corresponding norms.

A few more details

Given two functions f^1 and f^2 on $\mathbb{R}^d$ respectively, we set $f^1 \otimes f^2$

$$f^1 \otimes f^2(x_1, x_2) = f^1(x_1)f^2(x_2), \; x_i \in \mathbb{R}^d, i = 1, 2.$$

Given two Banach spaces $\boldsymbol{B}^1$ and $\boldsymbol{B}^2$ embedded into $\boldsymbol{\mathcal{S}}'(\mathbb{R}^d)$, $\boldsymbol{B}^1\hat{\otimes}\boldsymbol{B}^2$ denotes their *projective tensor product*, i.e.

$$\left\{ f \mid f = \sum f_n^1 \otimes f_n^2, \; \sum \|f_n^1\|_{\boldsymbol{B}^1}\|f_n^2\|_{\boldsymbol{B}^2} < \infty \right\}; \tag{482}$$ bibt-tensor

Any such representation of f will be called an *admissible representation*, and of course this is not unique, because one can add certain terms and subtract part of it later on. There is also often no optimal or canonical representation (as we have it for finite, discrete measures). It is easy to show that (482) defines a Banach space of tempered distributions on $\mathbb{R}^{2d}$, in our case a Banach space of mild distributions or even with respect to the folowing (natural quotient) norm:

$$\|f\|_{\hat{\otimes}} := inf \left\{ \sum \|f_n^1\|_{B^1}\|f_n^2\|_{B^2}, \right\}, \tag{483}$$ quot-norm

where the infimum is taken over all *admissible representations*.
One of the most important properties of $\boldsymbol{S}_0(\mathbb{R}^d)$ (leading to a characterization given by V. Losert, [76]) is the tensor-product factorization:

0tensor-prod **Lemma 114.**

$$\boldsymbol{S}_0(\mathbb{R}^k)\hat{\otimes}\boldsymbol{S}_0(\mathbb{R}^n) \cong \boldsymbol{S}_0(\mathbb{R}^{k+n}). \tag{484}$$ S0tens-repr

The easiest way to realize this relationship is to make use of the atomic decomposition of $\boldsymbol{S}_0(\mathbb{R}^m)$, observing that both time- and frequency shifts, but also the multi-dimensional Gaussian function factorize into lower dimensional partial ingredients.

This tensor product property of $\boldsymbol{S}_0$ is on the other hand the basis for the realization of the so called kernel theorem (see [44], Chap.7.4).

The KERNEL THEOREM for S_0 II

The **Kernel Theorem** for general operators in $\mathcal{L}(\boldsymbol{S}_0, \boldsymbol{S}_0')$:

Theorem 29. *If K is a bounded operator from $\boldsymbol{S}_0(\mathbb{R}^d)$ to $\boldsymbol{S}_0'(\mathbb{R}^d)$, then there exists a unique kernel $k \in \boldsymbol{S}_0'(\mathbb{R}^{2d})$ such that $\langle Kf, g\rangle = \langle k, g \otimes f\rangle$ for $f, g \in \boldsymbol{S}_0(\mathbb{R}^d)$, where $g \otimes f(x,y) = g(x)f(y)$.*

Formally sometimes one writes by "abuse of language"

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

with the understanding that one can define the action of the functional $Kf \in \boldsymbol{S}_0'(\mathbb{R}^d)$ as

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

The KERNEL THEOREM for S_0 III

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

The KERNEL THEOREM for S_0 IV

Theorem 30. *The classical kernel theorem for Hilbert Schmidt operators is unitary at the Hilbert spaces level, with $\langle T, S\rangle_{\mathcal{HS}} = trace(T * S')$ as scalar product on $\mathcal{HS}$ and the usual Hilbert space structure on $\boldsymbol{L}^2(\mathbb{R}^{2d})$ on the kernels.*
Moreover, such an operator has a kernel in $\boldsymbol{S}_0(\mathbb{R}^{2d})$ if and only if the corresponding operator K maps $\boldsymbol{S}_0'(\mathbb{R}^d)$ into $\boldsymbol{S}_0(\mathbb{R}^d)$, but not only in a bounded way, but also continuously from w^*-topology into the norm topology of $\boldsymbol{S}_0(\mathbb{R}^d)$.

In analogy to the matrix case, where the entries of the matrix

$$a_{k,,j} = T(\mathbf{e}_j)_k = \langle T(\mathbf{e}_j), \mathbf{e}_k\rangle$$

we have for $K \in \boldsymbol{S}_0$ for the corresponding operator $T = T_K$ the continuous version of this principle:

$$K(x,y) = \delta_x(T(\delta_y), \quad x, y \in \mathbb{R}^d.$$

The Kernel Theorem as a BGT isomorphism

The different version of the kernel theorem for operators between $\boldsymbol{S}_0$ and $\boldsymbol{S}_0'$ can be summarized using the terminology of Banach Gelfand Triples (BGTR) as follows.

Theorem 31. *There is a unique Banach Gelfand Triple isomorphism between the Banach Gelfand triple of kernels* $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^{2d})$ *and the operator Gelfand triple around the Hilbert space* $\mathcal{HS}$ *of Hilbert Schmidt operators, namely* $(\mathcal{L}(\boldsymbol{S}_0', \boldsymbol{S}_0), \mathcal{HS}, \mathcal{L}(\boldsymbol{S}_0, \boldsymbol{S}_0'))$*, where the first set is understood as the* w^* *to norm continuous operators from* $\boldsymbol{S}_0'(\mathbb{R}^d)$ *to* $\boldsymbol{S}_0(\mathbb{R}^d)$*, the so-called regularizing operators.*

Spreading function and Kohn-Nirenberg symbol

1. For $\sigma \in \boldsymbol{S}_0'(\mathbb{R}^d)$ the *pseudodifferential operator* with *Kohn-Nirenberg symbol* σ is given by:
$$T_\sigma f(x) = \int_{\mathbb{R}^d} \sigma(x,\omega)\hat{f}(\omega)e^{2\pi i x\cdot\omega}d\omega$$
The formula for the integral kernel $K(x,y)$ is obtained
$$\begin{aligned} T_\sigma f(x) &= \int_{\mathbb{R}^d}\Big(\int_{\mathbb{R}^d}\sigma(x,\omega)e^{-2\pi i(y-x)\cdot\omega}d\omega\Big)f(y)dy \\ &= \int_{\mathbb{R}^d} k(x,y)f(y)dy. \end{aligned}$$

2. The *spreading representation* of T_σ arises from
$$T_\sigma f(x) = \iint_{\mathbb{R}^{2d}} \widehat{\sigma}(\eta,u)M_\eta T_{-u}f(x)dud\eta.$$
$\widehat{\sigma}$ is called the spreading function of T_σ.

Further details concerning Kohn-Nirenberg symbol
(courtesy of Goetz Pfander (Eichstätt):)

- *Symmetric coordinate transform:* $\mathcal{T}_s F(x,y) = F(x+\frac{y}{2}, x-\frac{y}{2})$
- *Anti-symmetric coordinate transform:* $\mathcal{T}_a F(x,y) = F(x, y-x)$
- *Reflection:* $\mathcal{I}_2 F(x,y) = F(x,-y)$
- *partial Fourier transform in the first variable:* $\mathcal{F}_1$
- *partial Fourier transform in the second variable:* $\mathcal{F}_2$

The kernel $k(x,y)$ can be described as follows (imitating our MATLAB viewpoint: first reshape the matrix to side-diagonal form, then take a one-dimensional Fourier transform along the diagonals).

$$k(x,y) = \mathcal{F}_2\sigma(\eta, y-x) = \mathcal{F}_1^{-1}\widehat{\sigma}(x, y-x) = \int_{\mathbb{R}^d}\widehat{\sigma}(\eta, y-x)\cdot e^{2\pi i\eta\cdot x}d\eta.$$

Kohn-Nirenberg symbol and spreading function II

$$
\begin{array}{cc}
\text{operator } H & Hf(x) \\
\updownarrow & = \\
\text{kernel } \kappa_H & \int \kappa_H(x,s)f(s)\,ds \\
\updownarrow & = \\
\text{Kohn–Nirenberg symbol } \sigma_H & \int \sigma_H(x,\omega)\widehat{f}(\omega)e^{2\pi i x\cdot\omega}\,d\omega \\
\updownarrow & = \\
\text{time–varying impulse response } h_H & \int h_H(t,x)f(x-t)\,dt \\
\updownarrow & = \\
\text{spreading function } \eta_H & \int\int \eta_H(t,\nu)f(x-t)e^{2\pi i x\cdot\nu}\,dt\,d\nu \\
 & = \\
 & \int\int \eta_H(t,\nu)M_\nu T_t f(x),dt\,d\nu\,,
\end{array}
$$

Spreading representation and commutation relations

The description of operators through the spreading function and allows to understand a number of commutation relations.

If an operator is a limit (in the strong operator topology) of translation operators it is just a convolution operator with some $\tau \in \boldsymbol{S}_0'(\mathbb{R}^d)$, resp. its spreading representation is just an element concentrated on the *time axis* (more or less representing $\hat{\tau}$, the "individual frequency contributions".

Similarly, multiplication operators require just the use of modulation operators, so their spreading function is concentrated in the frequency axis of the TF-plane.

Finally typical *Gabor frame operators* arising from a family of Gabor atoms (g_λ), where $\lambda \in \Lambda$, some lattice within $\mathbb{R}^d \times \widehat{\mathbb{R}}^d$ typically commute with TF-shift operators, one can say that

they are obtained by periodizing the projection operator

$f \mapsto \langle f, g\rangle g$ along the lattice.

The symplectic Fourier transform

The *symplectic Fourier transform* connects the Kohn-Nirenberg symbol with the spreading function, i.e.

$$\mathcal{F}_s(\sigma(T)) = \eta(T) \quad \text{resp.} \quad \mathcal{F}_s(\eta(T)) = \sigma(T). \tag{485}$$ sigmatoeta

$$(\mathcal{F}_{symp} f)(k,l) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} f(x,y)e^{-2\pi i(k\cdot y - l\cdot x)}; \quad f \in \boldsymbol{S}_0(\mathbb{R}^d \times \widehat{\mathbb{R}}^d). \tag{486}$$ sympldef1

It is completely characterized by its action on elementary tensors:

$$\mathcal{F}_{symp}(f \otimes \hat{g}) = g \otimes \hat{f}, \quad f,g \in \boldsymbol{S}_0(\mathbb{R}^d), \tag{487}$$ FTsymptens1

and extends from there in a unique way to a $w^* - w^*$ continuous mapping from $\boldsymbol{S}_0'(\mathbb{R}^{2d})$ to $\boldsymbol{S}_0'(\mathbb{R}^{2d})$, also ${\mathcal{F}_s}^2 = Id$.

Best approximation by Gabor Multipliers

The Kohn-Nirenberg setting is also the best place, not only to compute a Gabor multiplier (given the ingredients, so ideally a pair (gt, Λ), with $gt \in \boldsymbol{S}_0(\mathbb{R}^d)$, such that the Gabor family $(\pi(\lambda)(gt))_{\lambda\in\Lambda}$ defines a tight Gabor frame) but also to find for a given Hilbert-Schmidt operator T (in the MATLAB situation: a given matrix) the best approximation by a Gabor multiplier based on the pair (gt, Λ), i.e. a collection of coefficients $(m_\lambda)_{\lambda\in\Lambda}$ such that

$$\left\| \sum_{\lambda\in\Lambda} m_\lambda P_\lambda - T \right\|_{\mathcal{HS}} := \min!! \tag{488}$$ HSapprGabmul

In the KNS setting this is like the problem to approximate a given function $f \in \boldsymbol{L}^2(\mathbb{R})$ by linear combinations of shifted B-splines (a task which has a nice, FFT-based solution), since the unitary equivalence between Hilbert Schmidt operators (resp. their integral kernels in $\boldsymbol{L}^2(\mathbb{R}^{2d})$) and the corresponding KNS symbol gives the following equivalent task:

$$\left\| \sum_{\lambda\in\Lambda} m_\lambda T_\lambda(\sigma(P_g)) - \sigma(T) \right\|_{\mathcal{HS}} := \min!! \tag{489}$$ HSapprGMKNS

Understanding the Janssen representation

The spreading representation of operators has properties very similar to the ordinary Fourier expansion for functions!
Periodization at one side corresponds to sampling on the transform side, if we understand "translation" either at the level of ordinary translation of the Kohn-Nirenberg symbol (which is the *symplectic Fourier transform* of the spreading function), OR by conjugation of an operator by the corresponding TF-shifts.
In other words: for any given operator T and $\lambda \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d$ we can **define** [recall $\pi(x,\omega) = M_\omega T_x$ for $\lambda = (x,\omega)$]

$$\pi \otimes \pi^*(\lambda)(T) = \pi(\lambda) \circ T \circ \pi(\lambda)^*, \tag{490}$$ LAMrepr

providing the important *covariance property* for KNS:

$$\sigma[\pi \otimes \pi^*(\lambda)(T)] = T_\lambda[\sigma(T)], \quad \lambda \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d. \tag{491}$$ KNScovar03

Periodization goes over to sampling
If we have a "nice operator" T_0 we can form its periodic version $\sum_{\lambda\in\Lambda} \pi \otimes \pi^*(\lambda)(T_0)$ and it is still a well defined operator from $\boldsymbol{S}_0(\mathbb{R}^d)$ to $\boldsymbol{S}_0'(\mathbb{R}^d)$. Its KNS is just the Λ-periodization of T_0.
Consequently its spreading function is obtained by sampling of $\eta(T) \in \boldsymbol{S}_0(\mathbb{R}^d \times \widehat{\mathbb{R}}^d)$, over the *adjoint lattice* Λ° and obtain in this case an $\boldsymbol{\ell}^1$-sequence.
The adjoint lattice Λ° can be characterized by the fact that

$$\mathcal{F}_s(⊔⊔_\Lambda) = C_\Lambda ⊔⊔_{\Lambda^\circ}. \tag{492}$$ symplatt

For the projection on the Gabor atom $P_g : f \mapsto \langle f, g\rangle g$ the spreading functions is essentially

$$[\eta(P_g)](\lambda) = V_g g(\lambda) = \langle g, \pi(\lambda) g\rangle, \quad \lambda \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d.$$

For the concrete case of the Gaussian atom g_0 one can verify (experimentally or analytically) that the KNS symbol of the projection operator P_{g_0} is, in absolute terms exactly the same as the two-dimensional Gauss function, i.e.

$$|\sigma(P_{g_0})| = |P_{g_0}|| = P_{g_0},$$

(meaning the absolute value of the matrix or kernel of the rank one operator $f \mapsto \langle f, g_0 \rangle g_0$, on the right hand side).

Janssen representation II

An important insight concerning the connection between the Gabor atom g, the TF-lattice $\Lambda \lhd \mathbb{R}^d \times \widehat{\mathbb{R}}^d$ and the quality of the resulting Gabor frame resp. Gabor Riesz basis (e.g. condition number) clearly comes from the *Janssen representation* of the *Gabor frame operator* for any $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ with $\|g\|_2 = 1$:

$$S_{g,\Lambda}(f) := \sum_{\lambda \in \Lambda} \langle f, g_\lambda \rangle g_\lambda = \sum_{\lambda \in \Lambda} P_{g_\lambda}(f) = \sum_{\lambda \in \Lambda} \pi \otimes \pi^*(\lambda)[P_g]. \tag{493}$$

Here $P_g(f) = \langle f, g_\lambda \rangle$ is the orthogonal projection onto the one-dimensional subspace generated by the scalar multiples of the atom g, and we could write $P_{g_\lambda}(f) = \langle f, g_\lambda \rangle g_\lambda$, for any $\lambda \in \Lambda$, observing that one has

$$P_\lambda(f) := P_{g_\lambda}(f) = \langle f, g_\lambda \rangle g_\lambda = \langle f, \pi(\lambda) g \rangle \pi(\lambda) g =_{unitary} \pi(\lambda)[\langle \pi(\lambda)^*(f), g \rangle g].$$

which, in a more compact form is just the claim that

$$P_\lambda = [\pi \otimes \pi^*(\lambda)] P_g := \pi(\lambda) \circ P_g \circ \pi(\lambda)^{-1}, \quad \lambda \in \Lambda.$$

This variant of the "representation" of the groupg $\Lambda \lhd \mathbb{R}^d \times \widehat{\mathbb{R}}^d$ (on the Hilbert space of Hilbert-Schmidt operators) is in fact a unitary representation, i.e. it satisfies the nice composition rule (without phase factors), for $\lambda_1, \lambda_2 \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d$:

$$[\pi \otimes \pi^*(\lambda_1)] \circ [\pi \otimes \pi^*(\lambda_2)] = \pi \otimes \pi^*(\lambda_1 + \lambda_2) = [\pi \otimes \pi^*(\lambda_2)] \circ [\pi \otimes \pi^*(\lambda_1)]. \tag{494}$$ pipisrules

This is consistent with the observation that the action of $\pi \otimes \pi^*(\lambda)$ on any operator (here the projection operator P_g) translates into ordinary translation in the Kohn-Nirenberg setting.
This is actually a way to compute the matrix describing a Gabor multiplier! The periodization principle gives the Janssen representation

$$S_{g,\Lambda} = \eta^{-1}[\eta(S_{g,\Lambda})] = C_\Lambda \sum_{\lambda^\circ \in \Lambda^\circ} V_g g(\lambda^\circ) \pi(\lambda^\circ), \tag{495}$$ Jansrep07

as an absolutely convergent sum of TF-shifts from Λ°, for some $C_\Lambda > 0$.
Following the ideas of Lemma 24 (Mbinvert) one can derive from the Janssen representation a very practical criterion for the inversion of the frame operator $S_{g,\Lambda}$, by checking that (for a normalized window $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ with $\|g\|_{\boldsymbol{L}^2} = 1$, hence $V_g g(0,0) = \|g\|_{\boldsymbol{L}^2}{}^2 = 1$)

$$\sum_{\lambda^\circ \in \Lambda^\circ} |V_g g(\lambda^\circ)| < 2$$

because this implies that

$$\sum_{(0,0)\neq\lambda^\circ\in\Lambda^\circ} |V_g g(\lambda^\circ)| < V_g g(0,0) = 1$$

and consequently one can estimate the operator norm on $\left(\boldsymbol{L}^2(\mathbb{R}^d),\ \|\cdot\|_2\right)$, but also on $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ (and $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'}))$ as

$$|||C_\Lambda \operatorname{Id}_{\boldsymbol{S}_0(\mathbb{R}^d)} - S_{g,\Lambda}|||_{\boldsymbol{S}_0(\mathbb{R}^d)} \le \sum_{(0,0)\neq\lambda^\circ\in\Lambda^\circ} |V_g g(\lambda^\circ)| < C_\Lambda. \tag{496}$$ Janscrit01

which implies that $S_{g,\Lambda}$ is invertible as an operator on $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$. This also guarantees that $S_{g,\Lambda}^{-1}$ is also bounded on $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$, and this in turn implies that $\widetilde{g} := S_{g,\Lambda}^{-1}(g) \in \left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$.

71 Engineering Questions and Mathematical Answers

At the end of this course let us take a critical look back on the content of the course, to check what kind of "take home messages" we might be able to formulate.
Given many years of interaction with engineers and physicists I hope to be able to contribute a little bit to the mutual understanding of the mathematical methods in their relationship to applications. Ideally we have something like Gabor Analysis, where the development of the last 40 years has allowed to develop quite useful theory which is highly relevant for applications (e.g. in mobile communication), where "slowly varying systems" (so not perfectly translation invariant, but approximately) arise in many concrete situations, due to relative motion of vehicles or trains. We should look for mathematical theory which supports the development of good and efficient (FFT-based?) algorithms, and we should be able to derive proofs for their performance (for example, such a proof could ensure a "guaranteed convergence rate"), or find at least a variant which is more safe towards failure compared to naive approaches, making use of experiment and error only.

71.1 The Dirac function $\delta(t)$

The Dirac functional $f \mapsto f(0)$ is a continuous linear functional on any Banach space of continuous functions, as long as convergence in $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ implies pointwise convergence. If we choose $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}}) = \left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$, then we may consider δ (no variable in the argument) as a bounded measure, with $\|\delta\|_{\boldsymbol{M}_b} = 1$.
It is *does not take the value zero for* $t \neq 0$, but we have $\mathrm{supp}(\delta) = \{0\}$, or equivalently: every $t \neq 0$ is part of the set of trivial action of δ. Similarly we have $\mathrm{supp}(\delta_x) = \mathrm{supp}(T_x\delta) = \{x\}$. The fact that $\int_{-\infty}^{\infty} \delta(t)dt = 1$ is just reflecting the fact that the functional (now extended to all of $\left(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\right)$) δ, applied to $\mathbf{1}$, the constant functions gives of course the value 1 (which arises also at $t = 0$), so we should translate the integral statement as $\delta(\mathbf{1}) = 1$.

71.2 From Fourier Series to Fourier Transforms

The usual motivation to introduce the Fourier transform as an integral transform comes from the observation that "formally" the Fourier series expansions of a function $f \in \boldsymbol{C}_c(\mathbb{R}^d)$, with the period tending to infinity, will create a continuous forward and inverse transform, which actually happens to look like a pair of mutually inverse transforms (because the "approximating" formulas of Fourier analysis: compute the Fourier coefficients, and Fourier synthesis: write the given signal as a series of the pure frequencies, with the Fourier coefficients as "coefficients"), but are they really inverse to each other.
Given the view-point that one should make use Lebesgue integrals to write the two transforms we have seen that they are in fact inverse to each other (with some caution), e.g. in $(\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1)(\mathbb{R}^d)$ or more general $\boldsymbol{S}_0(\mathbb{R}^d)$. Once this fact has been

established it is *fair to use* the so-called inverse Fourier transform as the *inverse transform* for the forward FT.

71.3 The extended Fourier transform

The extended Fourier transform can be introduced for mild distributions, using the fact that $\mathcal{F}$ and $\mathcal{F}^{-1}$ leave $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ (isometrically) invariant, by the well known formula $\widehat{\sigma}(f) = \sigma(\widehat{f}), f \in \boldsymbol{S}_0(\mathbb{R}^d)$.
Using this simple formula we obtain the unique w^*-w^*-continuous extension of the ordinary Fourier transform (base on the Riemann integral).
The forward Fourier transform of δ_0 is (of course) $\mathbf{1}$, and thus one expects (rightfully) that

$$\mathcal{F}^{-1}\,\mathbf{1} = \delta_0. \tag{497}$$ IFTbfi

But for ordinary functions the inverse Fourier transform is written as

$$\int_{-\infty}^{\infty} \textit{invisible } \mathbf{1}\; e^{2\pi i s t} ds = \delta_0,$$

which from a purely mathematical interpretation is not at all a meaningful integral.
The argument, that "in the mean" the integral is "kind of zero" while for $s = 0$ we have

$$\int_{-\infty}^{\infty} e^{2\pi i s 0} ds = \int_{-\infty}^{\infty} 1 dx = +\infty,$$

sounds like a reasonable argument, because usually one *says that $\delta(0) = \infty$, in fact "so strong that the integral is 1". But how can one distinguish δ_0 from $3\delta_0$?? And what about Dirac sequences for which the limit of the point-values at zero is zero, or tends to minus infinity, or oscillates (all this is possible)?*

71.4 Convolution and the sifting property

For the characterization of TILS one can use the output to the unit vector (at zero) if one is in a discrete setting (periodic or non-periodic). However, for the continuous variable case one often starts from the so-called sifting property of Dirac delta.

$$f(t) = \int_{-\infty}^{\infty} f(x)\delta(t - x)dx.$$

For us this is related to the convolution relation

$$\delta_0 \star f = T_0(f) = f, \quad f \in \boldsymbol{C}_0(\mathbb{R}^d),$$

expressing the identity operator as shift by zero. Thus it is not surprising to claim: If we apply the identity operator to some function $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ and ask for the value of $f = \mathbf{Id}(f)(y)$ then the answer is of course: "f(y)". In fact, this is not very surprising (and just a game with symbols[54]).

[54] I would almost say, not even mathematical symbols

71.5 Impulse response and w^*–convergence

The usual proof, observing that the input arising from applying the TILS to a Dirac sequence (e.g. a BIBO system) is justified in the section about the adjoint action of the ordinary convolution operator (which acts on $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$), and this adjoint operator maps the dual space, hence $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ into itself. But it turns out that the adjoint action can be described as *internal convolution*, but not by $\nu \in \boldsymbol{M}_b(\mathbb{R}^d)$, but by $\nu^{\checkmark}$. But since for any $\nu \in \boldsymbol{M}_b(\mathbb{R}^d)$ also $\nu^{\checkmark}$, defined by $\nu^{\checkmark}(f) = \nu(f^{\checkmark})$, belongs to $\boldsymbol{M}_b(\mathbb{R}^d)$ and we even have $\|\nu\|_{\boldsymbol{M}_b} = \|\nu^{\checkmark}\|_{\boldsymbol{M}_b}$.
Since adjoint operators area always w^*-w^*-continuous we can now safely claim (at least for BIBO system) that the output for the Dirac sequence is in fact w^*–convergent to the impulse response (as defined at the beginning of the course).

72 What we have NOT done: Missing topics

We did not find enough time to deal with concrete questions for the multi-dimensional case (at least not as far as it concerns MATLAB coding or concrete applications to image processing).

73 Gabor Analysis in detail

73.1 Gabor expansions and linear operators

One of the early studies in Gabor analysis was the fact that they allow to provide something like an *almost diagonalization* of *slowly varying systems* [55].
This was the starting point of joint projects with the team around F. Hlawatsch and G. Matz, which finally lead to two international patents in mobile communication. One to detect the distortion of a slowly varying channel (beyond the idea of just having eigenvectors) and the the second one to (partially) correct the distortion before doing the encoding.
On the way of these studies one has to show that different Gabor families with similar qualities (mostly decay and smoothness of g) generate matrix representation of a similar type.
One draw-back of the situation (based on the Balian-Low principle) however is the fact that due to the unavoidable redundancy of such a "good Gabor system" one cannot expect that the matrix, describing a system is an invertible matrix, even if the operator is a invertible. It turns out that in this case one has to resort to the *pseudo inverse* of the corresponding Gabor matrix and that one can actually derive similar claims (work of M. Fornasier and others).
This means, that e.g. Gaussian beams (i.e. modulated Gaussians) are "good approximated eigenvectors" of such a channel. Thus it is fair to approximate $T(g_\lambda)$ as $\rho_\lambda g_\lambda$, so that the decoding on the receiver side can be simple done by estimating ρ (e.g. by round-off quantization, if the data are quantized, say just the values from a finite set of coefficients, such as $\{0,1\}$ in the extreme case).
The mathematical development lead to a discussion of Banach algebras of matrices which are describing such operators via the *Gabor frame matrix representation*. It is (almost) like the ordinary matrix representation of a linear operator: one stores all the information about the operator in an (double infinite) matrix, which stores the Gabor coefficients $T(g_\lambda)$ as columns in such a matrix. Thus the entries of such a matrix are of the form

$$(\langle T(g_\lambda), \widetilde{g_\lambda} \rangle)_{\lambda,\lambda' \in \Lambda}$$

and that these entries decay if the distance $|\lambda - \lambda'|$ increases. We talk about well *localized Gabor frames* (for smooth and well decaying atoms, so at least $g \in \boldsymbol{S}_0(\mathbb{R}^d)$). It

[55]This is the time when Werner Kozek, at his time at NuHAG coming from the Comm. Theory Dept. TU Vienna, from the group of Franz Hlawatsch, under Prof. Wolfgang Mecklenbräucker, who had introduced the Wigner transform to signal processing. Helmut Bölcskei, at that time a PhD student from that group, also regularly attended the NuHAG seminar at this time, and we did a couple of joint papers on Gabor expansions and filter banks.

was an interesting mathematical development to find out that such Banach algebras of matrices with good off-diagonal decay (there are many variants of this principle) have the property that an invertible element in such an algebra (in our case: invertibility as operator on the Hilbert space $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$ is the starting point) *automatically have their inverse in the smaller Banach algbra* of matrices. This entails, that for many good properties of g one can guarantee qualitatively the same properties of $\widetilde{g}$ (or the tight Gabor atom $gt = S_{g,\Lambda}^{1/2}(g)$).

73.2 What is a good Gabor frame

There are many, many ways to choose a Gabor system. It should be a *well localized system*, meaning both in time and frequency direction, or better in the sense of *phase space* $\mathbb{R}^d \times \widehat{\mathbb{R}}^d$. Otherwise we fall into the trap of the original suggestion which kind of "just misses the stable representation" by choosing minimal redundancy, while at the same time loosing locality of the (then unique) coefficient mapping.
But unlike the expectation of D. Gabor was formulated (and put into a modern mathematical terminology): The critical Gabor system with $a = 1 = b$, starting from the Gauss-function g_0, is neither a frame nor a Riesz basic sequence, leave alone a *Riesz basis*.
It is true that this Gabor family is *total* in the Hilbert space $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$, which means that we have the following situation

$$\forall f \in \boldsymbol{L}^2(\mathbb{R}^d)\ \forall \varepsilon > 0\ \exists \text{ such that } \left\| f - \sum_{\lambda \in F} c_\lambda g_\lambda \right\|_{\boldsymbol{L}^2(\mathbb{R}^d)} < \varepsilon, \qquad (498)$$ Gabortotal

for some finite set $F \subset \Lambda$ and suitable coefficients $(c_\lambda)_{\lambda \in F}$. HOWEVER, it turns out, that the cost of the approximation, measured by the $\boldsymbol{\ell}^2$-norm of the set of coefficients is not limited. In other words, one cannot assume that $\sum_{\lambda\in\Lambda} |c_\lambda|^2 \leq C\|f\|_{\boldsymbol{L}^2}{}^2$ for all $f \in \boldsymbol{L}^2(\mathbb{R}^d)$. Consequently one cannot represent an arbitrary element in this way by taking limits of coefficients (for $\varepsilon \to 0$). For details see the paper by Genossar and Porat (**gepo92:** [55]).

73.3 The duality principle of Gabor analysis

All this relies on the fact that one has for any lattice Λ such that (g, Λ) (with $g \in \boldsymbol{S}_0(\mathbb{R}^d)$) generates a Gabor frame, in fact not only for $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$, but then (due to the good quality of $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ and the nice properties of $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$) also $\widetilde{g} \in \boldsymbol{S}_0(\mathbb{R}^d)$ and the family $(g_\lambda)_{\lambda\in\Lambda}$ also is a Gabor frame for the Banach Gelfand triple $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$.
In other words, the coefficient mapping is a bounded BGTr morphism from $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$ into $(\boldsymbol{\ell}^1, \boldsymbol{\ell}^2, \boldsymbol{\ell}^\infty)$, and the synthesis mapping (using $\widetilde{g}$) is a left inverse to the coefficient building (meaning it allows to recover the signal in a norm convergent fashion from the Gabor coefficients), using the dual Gabor atom $\widetilde{g} \in \boldsymbol{S}_0(\mathbb{R}^d)$. Technical details in this direction are already presented in **fezi98:** [47].
While the density of a lattice $\Lambda \lhd \mathbb{R}^d \times \widehat{\mathbb{R}}^d$ gives good chances that the corresponding Gabor family with atom $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ forms such a Gabor frame, the dual lattice will be

ZebraG.jpg

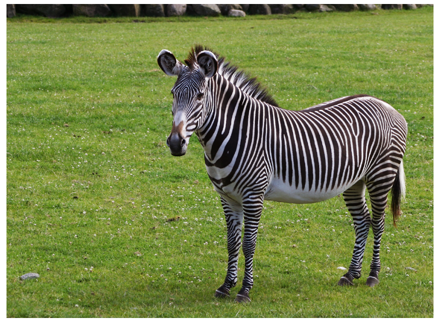

Abbildung 97: ZebraG.jpg The Zebra as a typical test image for 2D Gabor analysis.

less and less redundant, meaning it will become more and more sparse (so the chances of applying the Janssen criterion for the invertibility of the Gabor frame operator $S_{g,\Lambda}$ are getting better and better).
It is one of the key facts of Gabor analysis that one has the following principle.

GaborDuality

Theorem 32. *Given $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ and $\Lambda \lhd \mathbb{R}^d \times \widehat{\mathbb{R}}^d$. The family gLam is a Gabor frame if and only if the (adjoint) family $(g_{\lambda^\circ})_{\lambda^\circ \in \Lambda^\circ}$ is a Riesz basic sequence, i.e. a Riesz basis for its closed linear span.*
Up to normalization the dual Gabor atom $\tilde{g}$ also generates the biorthogonal Riesz basic sequence for the elements in the closed linear span of the adjoint family. In other words, one can describe the orthogonal projection onto that closed subspace of $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$ (hence the identity operator for elements from the subspace) as

$$P(f) = C_{\Lambda^\circ} \sum_{\lambda^\circ \in \Lambda^\circ} \langle f, g_{\lambda^\circ} \rangle \tilde{g}_{\lambda^\circ}. \qquad (499)$$

projadjGab

73.4 Two-dimensional Gabor Analysis

The building blocks for the Gabor expansions of an audio signal are the localized pure frequencies. The mathematical description given already covers the multi-dimensional case. But the practical realization is quite different. Concrete implementations require not only the more cumbersome computation of dual atoms (not a problem anymore, nowadays) but rather also questions of memory management and so on. Just take into account that we have already a redundancy of n for harmless one-dimensional signals of size n. Hence the selection of suitable lattices of redundancy in the range of say typically $red \in [1.1, 2]$ is very reasonable. This problem is even more important in the 2 and 3-dimensional case. Also the variety of possible lattices already increases dramatically, even for the finite case of square images of format $n \times n$ (pixel images, so to say, viewed as periodic and discrete signals).
Still it is interesting to recall that in the 2-dimensional case the pure frequencies are "plane waves" and thus localized pure frequencies are (smooth and well localized) blobs with a clear local frequency of a certain wavenumber (meaning direction and density of the ripples). Thus a good test image is a Zebra.

Literatur

[1] R. Aceska. *Functions of Variable Bandwidth: a Time-Frequency Approach.* PhD thesis, University of Vienna, 2009.

[2] L. N. Argabright and J. Gil de Lamadrid. *Fourier Analysis of Unbounded Measures on Locally Compact Abelian Groups.* Deutscher Taschenbuch Verlag, München, 1974.

[3] M. Baake and U. Grimm. *Aperiodic Order. Vol. 1*, volume 149 of *Encyclopedia of Mathematics and its Applications.* Cambridge University Press, Cambridge, 2013.

[4] M. Baake and U. Grimm. *Aperiodic Order: Volume 2, Crystallography and Almost Periodicity*, volume 166. Cambridge University Press, 2017.

[5] H. Bölcskei, H. G. Feichtinger, K. Gröchenig, and F. Hlawatsch. Discrete-time Wilson expansions. In *Proc. IEEE-SP Int.Symp. on Time-Frequency and Time-Scale Analysis*, pages 525–528, Paris, France, 1996.

[6] H. Bölcskei, K. Gröchenig, F. Hlawatsch, and H. G. Feichtinger. Oversampled Wilson expansions. *IEEE Signal Proc. Letters*, 4(4):106–108, 1997.

[7] L. H. Brandenburg. On identifying the maximal ideals in Banach algebras. *J. Math. Anal. Appl.*, 50:489–510, 1975.

[8] W. Braun and H. G. Feichtinger. Banach spaces of distributions having two module structures. *J. Funct. Anal.*, 51:174–212, 1983.

[9] R. C. Busby and H. A. Smith. Product-convolution operators and mixed-norm spaces. *Trans. Amer. Math. Soc.*, 263:309–341, 1981.

[10] P. L. Butzer and R. J. Nessel. *Fourier Analysis and Approximation. Vol. 1: One-dimensional Theory.* Birkhäuser, Stuttgart, 1971.

[11] G. Cariolaro. *Unified Signal Theory.* Springer, London, 2011.

[12] L. Carleson. On convergence and growth of partial sums of Fourier series. *Acta Math.*, 116:135–157, 1966.

[13] O. Christensen. *An Introduction to Frames and Riesz Bases.* Applied and Numerical Harmonic Analysis. Birkhäuser Basel, Second edition, 2016.

[14] D. Cohn. *Measure Theory.* Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks]. Birkhäuser/Springer, New York, Second edition, 2013.

[15] E. Cordero, H. G. Feichtinger, and F. Luef. Banach Gelfand triples for Gabor analysis. In *Pseudo-differential Operators*, volume 1949 of *Lecture Notes in Mathematics*, pages 1–33. Springer, Berlin, 2008.

[16] I. Daubechies, S. Jaffard, and J. L. Journé. A simple Wilson orthonormal basis with exponential decay. *SIAM J. Math. Anal.*, 22:554–573, 1991.

[17] Y. Domar. Harmonic analysis based on certain commutative Banach algebras. *Acta Math.*, 96:1–66, 1956.

[18] J. Duoandikoetxea. *Fourier Analysis*, volume 29 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, 2001.

[19] R. E. Edwards. The stability of weighted Lebesgue spaces. *Trans. Amer. Math. Soc.*, 93:369–394, 1959.

[20] H. G. Feichtinger. English translation of: Gewichtsfunktionen auf lokalkompakten Gruppen (download at www.nuhag.eu/bibtex). *Sitzungsber.d.österr. Akad.Wiss.*, 188, 1979.

[21] H. G. Feichtinger. Gewichtsfunktionen auf lokalkompakten Gruppen. *Sitzber. d. österr. Akad. Wiss.*, 188:451–471, 1979.

[22] H. G. Feichtinger. A characterization of minimal homogeneous Banach spaces. *Proc. Amer. Math. Soc.*, 81(1):55–61, 1981.

[23] H. G. Feichtinger. On a new Segal algebra. *Monatsh. Math.*, 92:269–289, 1981.

[24] H. G. Feichtinger. Banach convolution algebras of Wiener type. In *Proc. Conf. on Functions, Series, Operators, Budapest 1980*, volume 35 of *Colloq. Math. Soc. Janos Bolyai*, pages 509–524. North-Holland, Amsterdam, Eds. B. Sz.-Nagy and J. Szabados. edition, 1983.

[25] H. G. Feichtinger. Modulation spaces on locally compact Abelian groups. Technical report, University of Vienna, January 1983.

[26] H. G. Feichtinger. Atomic characterizations of modulation spaces through Gabor-type representations. In *Proc. Conf. Constructive Function Theory*, volume 19 of *Rocky Mountain J. Math.*, pages 113–126, 1989.

[27] H. G. Feichtinger. Modulation spaces on locally compact Abelian groups. In R. Radha, M. Krishna, and S. Thangavelu, editors, *Proc. Internat. Conf. on Wavelets and Applications*, pages 1–56, Chennai, January 2002, 2003. New Delhi Allied Publishers.

[28] H. G. Feichtinger. *Choosing Function Spaces in Harmonic Analysis*, volume 4 of *The February Fourier Talks at the Norbert Wiener Center, Appl. Numer. Harmon. Anal.*, pages 65–101. Birkhäuser/Springer, Cham, 2015.

[29] 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.

[30] H. G. Feichtinger. Thoughts on Numerical and Conceptual Harmonic Analysis. In A. Aldroubi, C. Cabrelli, S. Jaffard, and U. Molter, editors, *New Trends in Applied Harmonic Analysis. Sparse Representations, Compressed Sensing, and Multifractal Analysis*, Applied and Numerical Harmonic Analysis, pages 301–329. Birkhäuser, Cham, 2016.

[31] H. G. Feichtinger. A novel mathematical approach to the theory of translation invariant linear systems. In Peter J. Bentley and I. Pesenson, editors, *Novel Methods in Harmonic Analysis with Applications to Numerical Analysis and Data Processing*, pages 483–516. Birkhäuser, Cham, 2017.

[32] H. G. Feichtinger. Banach Gelfand Triples and some Applications in Harmonic Analysis. In J. Feuto and M. Essoh, editors, *Proc. Conf. Harmonic Analysis (Abidjan, May 2018)*, pages 1–21, 2018.

[33] H. G. Feichtinger. Classical Fourier Analysis via mild distributions. *MESA, Non-linear Studies*, 26(4):783–804, 2019.

[34] H. G. Feichtinger. Gabor expansions of signals: computational aspects and open questions. In *Landscapes of Time-Frequency Analysis*, volume ATFA17, pages 173–206. Birkhäuser/Springer, 2019.

[35] H. G. Feichtinger. A sequential approach to mild distributions. *Axioms*, 9(1):1–25, 2020.

[36] H. G. Feichtinger. Ingredients for Applied Fourier Analysis. In *Sharda Conference Feb. 2018*, pages 1–22. Taylor and Francis, 2020.

[37] H. G. Feichtinger and P. Gröbner. Banach spaces of distributions defined by decomposition methods. I. *Math. Nachr.*, 123:97–120, 1985.

[38] H. G. Feichtinger and K. Gröchenig. Banach spaces related to integrable group representations and their atomic decompositions, I. *J. Funct. Anal.*, 86(2):307–340, 1989.

[39] H. G. Feichtinger, K. Gröchenig, and D. F. Walnut. Wilson bases and modulation spaces. *Math. Nachr.*, 155:7–17, 1992.

[40] H. G. Feichtinger and W. Hörmann. A distributional approach to generalized stochastic processes on locally compact abelian groups. In G. Schmeisser and R. Stens, editors, *New Perspectives on Approximation and Sampling Theory. Festschrift in honor of Paul Butzer's 85th birthday*, pages 423–446. Cham: Birkhäuser/Springer, 2014.

[41] H. G. Feichtinger and M. S. Jakobsen. Distribution theory by Riemann integrals. *Mathematical Modelling, Optimization, Analytic and Numerical Solutions*, pages 33–76, 2020.

[42] H. G. Feichtinger and A. J. E. M. Janssen. Validity of WH-frame bound conditions depends on lattice parameters. *Appl. Comput. Harmon. Anal.*, 8(1):104–112, 2000.

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

[44] H. G. Feichtinger and W. Kozek. Quantization of TF lattice-invariant operators on elementary LCA groups. In H. G. Feichtinger and T. Strohmer, editors, *Gabor analysis and algorithms*, Appl. Numer. Harmon. Anal., pages 233–266. Birkhäuser, Boston, MA, 1998.

[45] H. G. Feichtinger and G. Narimani. Fourier multipliers of classical modulation spaces. *Appl. Comput. Harmon. Anal.*, 21(3):349–359, 2006.

[46] H. G. Feichtinger and T. Werther. Robustness of regular sampling in Sobolev algebras. In J. Benedetto, editor, *Sampling, Wavelets and Tomography*, pages 83–113. Birkhäuser, 2004.

[47] 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*, Applied and Numerical Harmonic Analysis, pages 123–170. Birkhäuser Boston, 1998.

[48] G. B. Folland. *Fourier Analysis and its Applications.* Wadsworth and Brooks, CA, 1992.

[49] G. B. Folland. *A Course in Abstract Harmonic Analysis.* CRC Press, 1994.

[50] G. B. Folland. *A Course in Abstract Harmonic Analysis.* Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1995.

[51] J. J. F. Fournier and J. Stewart. Amalgams of L^p and ℓ^q. *Bull. Amer. Math. Soc. (N.S.)*, 13:1–21, 1985.

[52] M. Frazier and B. Jawerth. A discrete transform and decompositions of distribution spaces. *J. Funct. Anal.*, 93(1):34–170, 1990.

[53] D. Gabor. Theory of communication. *J. IEE*, 93(26):429–457, 1946.

[54] G. I. Gaudry. Multipliers of weighted Lebesgue and measure spaces. *Proc. Lond. Math. Soc., III. Ser.*, 19:327–340, 1969.

[55] T. Genossar and M. Porat. Can one evaluate the Gabor expansion using Gabor's iterative algorithm? *IEEE Trans. Signal Process.*, 40(8):1852–1861, 1992.

[56] J. Gil de Lamadrid and L. N. Argabright. *Almost Periodic Measures.* Deutscher Taschenbuch Verlag, München, 1990.

[57] L. Grafakos. *Classical and Modern Fourier Analysis.* Prentice Hall, 2004.

[58] K. Gröchenig. An uncertainty principle related to the Poisson summation formula. *Studia Math.*, 121(1):87–104, 1996.

[59] K. Gröchenig. *Foundations of Time-Frequency Analysis.* Appl. Numer. Harmon. Anal. Birkhäuser, Boston, MA, 2001.

[60] K. Gröchenig. Weight functions in time-frequency analysis. In L. Rodino and et al., editors, *Pseudodifferential Operators: Partial Differential Equations and Time-Frequency Analysis*, volume 52 of *Fields Inst. Commun.*, pages 343–366. Amer. Math. Soc., Providence, RI, 2007.

[61] E. Hewitt and K. A. Ross. *Abstract Harmonic Analysis. Vol. II: Structure and Analysis for Compact Groups. Analysis on Locally Compact Abelian Groups.* Springer, Berlin, Heidelberg, New York, 1970.

[62] F. Holland. Harmonic analysis on amalgams of L^p and ℓ^q. *J. London Math. Soc.*, 10:295–305, 1975.

[63] F. Holland. On the representation of functions as Fourier transforms of unbounded measures. *Proc. Lond. Math. Soc., III. Ser.*, 30:347–365, 1975.

[64] W. Hörmann. *Generalized Stochastic Processes and Wigner Distribution.* PhD thesis, University of Vienna, (AUSTRIA), 1989.

[65] 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.

[66] D. S. Jones. *Generalised Functions.* McGraw-Hill Book Co., New York-Toronto, 1966.

[67] N. Kaiblinger. Approximation of the Fourier transform and the dual Gabor window. *J. Fourier Anal. Appl.*, 11(1):25–42, 2005.

[68] R. P. Kanwal. *Generalized Functions. Theory and Applications. 3rd Revised ed.* Birkhäuser, MA, 2004.

[69] Y. Katznelson. *An Introduction to Harmonic Analysis. 2nd corr. ed.* Dover Publications Inc., New York, 1976.

[70] B. Keville. *Multidimensional Second Order Generalised Stochastic Processes on Locally Compact Abelian Groups.* PhD thesis, Trinity College Dublin, 2003.

[71] G. Kutyniok and T. Strohmer. Wilson bases for general time-frequency lattices. *SIAM J. Math. Anal.*, 37(3):685–711, 2005.

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

[73] M. J. Lighthill. *Introduction to Fourier Analysis and Generalised Functions. (Students' Edition).* Cambridge University Press, Cambridge, 1962.

[74] M. J. Lighthill. *Einführung in die Theorie der Fourieranalysis und der Verallgemeinerten Funktionen.* Bibliographisches Institut, Mannheim, 1966.

[75] L. H. Loomis. *An Introduction to Abstract Harmonic Analysis.* Van Nostrand and Co., 1953.

[76] V. Losert. A characterization of the minimal strongly character invariant Segal algebra. *Ann. Inst. Fourier (Grenoble)*, 30:129–139, 1980.

[77] R. L. Pego. Compactness in L^2 and the Fourier transform. *Proc. Amer. Math. Soc.*, 95:252–254, 1985.

[78] H. Reiter. *Classical Harmonic Analysis and Locally Compact Groups.* Clarendon Press, Oxford, 1968.

[79] H. Reiter and J. D. Stegeman. *Classical Harmonic Analysis and Locally Compact Groups. 2nd ed.* Clarendon Press, Oxford, 2000.

[80] W. Rudin. *Fourier Analysis on Groups.* Interscience Publishers, New York, London, 1962.

[81] I. W. Sandberg. The reconstruction of band-limited signals from nonuniformly spaced samples. *IEEE Trans. Circuits Syst., I, Fundam. Theory Appl.*, 41(1):64–66, 1994.

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

[83] I. W. Sandberg. Continuous-time linear systems: folklore and fact. *Circuits Syst. Signal Process.*, 21(3):337–343, 2002.

[84] L. Schwartz. *Théorie des Distributions. (Distribution Theory). Nouveau Tirage. Vols. 1.* Paris: Hermann. xii, 420p., 1957.

[85] R. Shakarchi and E. M. Stein. *Fourier Analysis: An Introduction.* Princeton Lectures in Analysis. Princeton University Press, 2003.

[86] A. Sokal. A really simple elementary proof of the uniform boundedness theorem. *Amer. Math. Monthly*, 118(5):450–452, 2011.

[87] G. Temple. Theories and applications of generalized functions. *J. London Math. Soc.*, 28:134–148, 1953.

[88] G. Temple. The theory of generalized functions. *Proc. Roy. Soc. London. Ser. A.*, 228:175–190, 1955.

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

[90] A. Vretblad. *Fourier Analysis and its Applications.* Springer-Verlag, NY, 2003.