`Hans G. Feichtinger: Harmonic Analysis based on Functional Analysis`

`http://www.univie.ac.at/nuhag-php/login/skripten/data/AngAnal18Skript.pdf`
and for general functional analysis:
`http://www.univie.ac.at/NuHAG/FEICOURS/ws1314/FAws1314Fei.pdf`
`VERSION of September 2020, ETH Zuerich, hgfei`

Contents

We keep updating these notes, following feedback of readers

`local file name: C:\TOWRITE\AngAnalScript\AngAnalETH20.tex`
`https://www.univie.ac.at/nuhag-php/login/skripten/data/AngAnal19MB.pdf`

A FUNCTIONAL ANALYTIC APPROACH TO APPLIED ANALYSIS: NUHAG: VERSION OF SEPTEMBER 2020

HANS G. FEICHTINGER

2000 *Mathematics Subject Classification.* Primary ; Secondary .
Key words and phrases. Fourier transform, convolution, linear time-invariant systems, bounded measures, Banach modules, approximate units, Banach algebras, essential Banach modules, discrete measures, Plancherel, Fourier Stieltjes transform, compactness, tight, w^*-convergence, strong operator topology, approximate units, support.

ABSTRACT. **Linear algebra** is dealing with *finite dimensional vector spaces and linear mappings.* In order to describe them one uses bases and *matrices* and thus connection to systems of linear equations is given. Solving such a system is more or less the same as inverting the corresponding linear mapping.

Analysis talks a lot about limits, series, continuity, differentiability, and so on. Functional analysis is dealing with non-finite dimensional vector spaces, but in order to express the possibility of infinite series representation one has to explain how limits are to be understood. Instead of using a specific set of linear functionals (the dual basis to the given basis) one often takes all possible linear functionals on a given normed space. And in order to be in a situation more similar to $\mathbb{R}$ (complete!) and not like $\mathbb{Q}$ (rationals) one often assumes that one is working with *Banach spaces.* Occasionally *Hilbert spaces* are used, which allow to define concepts of *orthogonality*, *unitary* mappings etc.. In order to do this properly one has to make use of basic concepts from **functional analysis**. The corresponding concepts are however also of relevance in **numerical analysis**, where one is usually approximating a continuous problem by a corresponding finite dimensional version of the problem (e.g. obtained by *discretization*).

We do not assume that participants of this course have already taken a course in functional or numerical analysis, but we will try to connect the course material with these fields. **The course will emphasize the use of various Banach spaces of functions, their properties and their role for the description of real world problems.**

Note: Usually the role of the Lebesgue spaces $\boldsymbol{L}^p$, $1 \leq p \leq \infty$ is overemphasized. Obviously the cases $p = 1, 2, \infty$ are very important, also for applications, but we will see that many other spaces are at least equally useful.

Typos and inconsistences to be discussed as of Jan. 2020

The use of the symbol $\boldsymbol{C}_c(G)$, $\boldsymbol{C}_c(\mathbb{R}^d)$, to be checked

We start from a variety of "simple spaces" (of continuous, complex-valued functions).

Ccdef1 **Definition 1.**
- $\boldsymbol{C}(G) := \{f : G \to \mathbb{C}, f$ continuous, complex-valued function on G $\}$;
- $\boldsymbol{C}_b(G) := \{f : f \in \boldsymbol{C}(G)$ and bounded $\}$;
- $\boldsymbol{C}_c(G) := \{f \in \boldsymbol{C}(G),$ with $\mathrm{supp}(f)$ *compact* $\}$
  where $\mathrm{supp}(f)$ is defined as $\{x \in G, f(x) \neq 0\}^-$,
  the closure of the set of points in G where f is non-zero;
- $\boldsymbol{C}_{ub}(G) := \{f : f \in \boldsymbol{C}(G)$ and uniformly continuous $\}$.

CubGdef02 **Definition 2.** For $G = \mathbb{R}^d$ (or any metric group) the membership of $f \in \boldsymbol{C}_{ub}(\mathbb{R}^d)$ is given in the usual way: $f \in \boldsymbol{C}_b(\mathbb{R}^d)$ belongs to $\boldsymbol{C}_{ub}(\mathbb{R}^d)$ if and only if $\forall\, \varepsilon > 0\ \exists\, \delta > 0$ such that

$$|x - y| < \delta \quad \Rightarrow \quad |f(x) - f(y)| < \varepsilon.$$

....

1. MATERIAL TO BE USED OR REMOVED LATER

1.1. **Prerequisites.** Normed spaces, Banach spaces, Hilbert spaces dual space, linear functionals, norm of linear functionals, bounded linear operators, operator norm (see [future] appendix).

1.2. **Operators and Conventions.** Many of the operators defined below (such as translation, dilation, Fourier transform or convolution operators) should be considered as quite flexible, in the sense that the general setup of the presentation is made *in such a way that possible ambiguities* will not come to bear when the user makes his interpretation.

Of course this requires some care (e.g. in the definition of the "existence" of convolutions or pointwise products)

1.3. **The old abstract.** We present a functional analytic approach to harmonic analysis, avoiding heavy measure theoretic tools. The results are non-trivial even for the real line $\mathbb{R}$, but are formulated for the $d-$dimensional Euclidean space $\mathbb{R}^d$, viewed as a prototype for a locally compact Abelian (:LCA) group G (with the usual addition of vectors and the topology provided by the Euclidian metric).[1] We start with the description of $\boldsymbol{M}_b(G)$, the space of bounded linear measures, as the dual space of $\boldsymbol{C}_0$, which is naturally endowed with a convolution structure.

Convolution will first be defined for measures (via the identification of impulse responses with the corresponding translation invariant linear systems mapping $\boldsymbol{C}_0(\mathbb{R}^d)$ into $\boldsymbol{C}_0(\mathbb{R}^d)$) and only later specialized to "ordinary functions". The Fourier transform is then introduced as the action of such systems on their joint eigenvectors, namely the pure frequencies.

[1] Since any point has a basis of the neighborhood, consisting of the compact closed balls $B_r(x)$ of radius r around the point x this is a locally compact group.

sec:intro

2. Introduction

We start with a few symbols. First note that we will permanently make use of the fact that $\mathbb{R}^d$ is a **locally compact Abelian group** with respect to addition (dilation will come in only as a convenient but not crucial side aspect Such $\mathbb{R}^d$-specific parts are marked by the symbol $[\mathbf{Rd-specific!}]$!)

This means first of all that we know what convergence is, hence we know what closed sets M are (limits of sequences with elements in a closed set belong to the set), or in term of symbols $M^- = M$ (closure symbol). Closed sets are complements of open sets. A set Q is open if for every element $y \in Q$ there exists some ε-ball $B_\varepsilon(y) \subseteq Q$.

This means that it is clear when a set is open or closed, and that every point has a basis of neighborhoods consisting of compact balls, the closure of the open balls $B_\gamma(x)$ of radius γ around each $x \in \mathbb{R}^d$. Moreover, clearly addition is continuous, i.e. $a_n \to a_0$ and $b_n \to b_0$ for $n \to \infty$ implies $a_n + b_n \to a_0 + b_0$ for $n \to infty$.

First we define the most simple algebra (pointwise, later on with respect to convolution) of continuous *"test functions"*. [2] Because of the *local compactness* of $\mathbb{R}^d$ the following space of *continuous functions with compact support* is a non-trivial (i.e. *not just* the zero-space) linear space of complex-valued functions on G[3]:

CcRddef1 **Definition 3.**

CcRddef2 (1)
$$\boldsymbol{C}_c(\mathbb{R}^d) := \{f : \mathbb{R}^d \to \mathbb{C}, \text{ continuous and with compact support}\} \quad .$$

Here we make use of the standard definition of the support[4] of a function:

defsupp1 **Definition 4.** The *support* of a continuous (!) function, in symbols $\mathrm{supp}(f)$, is defined as the closure of the set of "relevant points": [5]

defsupp2 (2)
$$\mathrm{supp}(f) := \{x \mid f(x) \neq 0\}^-.$$

Lemma 1. *(Ex to Def. 4) A continuous, complex-valued function on $\mathbb{R}^d$ is in $\boldsymbol{C}_c(\mathbb{R}^d)$ if and only if there exists $R = R(f) > 0$ such that $f(x) = 0$ for whenever $|x| \geq R$.*

A list of symbols: Operators, names of operators, and their definitions:
SOME OF THESE DEFINITIONS ARE REPEATED LATER ON!

Definition 5 (Operators defined on functions).

(1) translation by x is the operator T_x given by

translatdef (3)
$$T_x f(z) = [T_x f](z) := f(z - x), \quad x, z \in \mathbb{R}^d;$$

(the graph is preserved but moved by the vector x to another position)[6];

---

[2] They can be defined on any topological group, and the space is interesting and non-trivial for any locally compact group. So much of the material given below extends without difficulty to the setting of locally compact, or at least locally compact Abelian group, except for the statements which involve dilations

[3] The capital "C" stands for *continuous*, and the subscript "c"stands for *compact* support; H. Reiter uses the symbol $\mathcal{K}(G)$ for the space $\boldsymbol{C}_c(G)$.

[4] Why one takes the closure in the above definition will become more clear later on, when the support of a generalized function or distribution will be defined. Otherwise the argument $f \cdot g = 0$ if they do not have any common non-zero argument would be slightly stronger than the claim $\mathrm{supp}(f) \cap \mathrm{supp}(g) = \emptyset \Rightarrow f \cdot g = 0$.

[5] The superscript bar stands for "closure" of a set. Hence $\mathrm{supp}(f)$ is by definition a closed set.

[6] Engineers would rather write $f(z)$ for the function itself, in order to indicate that it a function allowing z from some domain as input. They would then write simple $f(z - x)$ (for fixed x) as the new function, also as a function of z, with values being the same values as that of $f(z)$ at the position $z - x$.

(2) involution $f \mapsto f^{\checkmark}$ with

checkdef1 (4)
$$f^{\checkmark}(z) := f(-z), \quad z \in \mathbb{R}^d;$$

(3) modulation M_s: Multiplication with the character[7] $x \mapsto exp(2\pi is \cdot x)$, i.e.

modopdef1 (5)
$$[M_s f](z) := e^{2\pi is\cdot z} f(z) = \chi_s(t) f(z), \quad x, s \in \mathbb{R}^d;^{8}$$

(4) **[Rd-specific!]**dilation D_ρ (value preserving) and St_ρ (mass preserving); cf. below for details.

(5) Fourier transform $\mathcal{F}, \mathcal{F}^{-1}$ (to be discussed only later): will be given for $f \in \boldsymbol{C}_c(\mathbb{R}^d)$ by the (Riemann or Lebesgue) integral (again with the *normalization* 2π in the exponent!)

Four-intdef (6)
$$\mathcal{F}\colon f \mapsto \hat{f}: \quad \widehat{f}(s) = \int_{\mathbb{R}^d} f(t) e^{-2\pi is\cdot t} dt, \quad f \in \boldsymbol{L}^1(\mathbb{R}^d).$$

Note that $s \cdot t$ stands for the usual scalar product in $\mathbb{R}^d$, so we could also write $e^{-2\pi i[x_1 s_1 + \ldots + x_d s_d]}$ in the integral.

HINTS on Brackets, Bindings, Conventions

Remark 1. During these notes we will make a lot of use of operators (such as the ones just defined), but also of all kinds of multiplications, e.g. pointwise multiplications of functions or between functions and measures, or convolution products (denoted by $*$). When such symbols come together we consider the operators to be completely bound/attached to the function symbol. So clearly $T_x f \cdot g$ is to be interpreted as $(T_x f) \cdot g$,which is different from the expression (which would then be spelled out as such) $T_x(f \cdot g)$. There is no fixed convention when two multiplications come together, hence one has to put some kind of brackets to indicate whether one wants to interpret something (!!! please do not use a symbol like $h * g \cdot f$ which might be interpreted as $(h * g) \cdot f$ or as $h * (g \cdot f)$.

It may also be worthwhile to note that $T_x : f \mapsto T_x f$ is a mapping between functions, and is *not at all* something which acts on a value. So an expression like $T_x[f(t)]$ would be considered as "**non-sense**" in our context[9]. At the beginning this fine distinction my look/sound a bit strange, but there is a lot of practical experience behind these conventions, which have proven to be very convenient and helpful when applied strictly and consistently! Above all one derives certain relations between these operators once and for all times and can then simple use some algebraic identities under various circumstances, without having to repeat the required steps, say change of the argument, applying Fubini's Theorem and the like, over and over again!

Many of the estimates which we will obtain are realized by first establishing a pointwise estimate for two functions, e.g. $|f(x)| \leq |g(x)|$, $\forall x \in G$ (for example). In such a case we will simply write $|f| \leq |g|$ (assuming that the domain is clear from the context), which of course implies e.g. $\|f\|_p \leq \|g\|_p$ for any $p \in [1, \infty]$.

We also want to point out that there is no unique convention within functional analysis, how to denote the dual space. Two conventions are quite standard for the description of the dual space, either the symbol $\boldsymbol{B}'$ (and consequently $\boldsymbol{B}$" for the bi-dual space), or $\boldsymbol{B}^*$ (and then $\boldsymbol{B}^{**}$).

[7]Here we write $x \cdot x$ for the standard scalar product in $\mathbb{R}^d : \sum_{k=1}^d x_k \overline{y}_k$.

[8]The use of the factor 2π in the exponent allows us to interpret s as the frequency number. The real/imginary part of this complex exponential function (cos / sin, recall Euler's formula) have s full periods on any interval of length 1, if $s \in \mathbb{N}$

[9] It is however quite popular among engineers to indicate that f is a function of t, which is then shifted. But probably they would write $t \to f(t - x)$ to describe such a shifted function

Remark 2. More like a FOREWORD/PREFACE to these lecture notes:

The first three operator families defined above are isometric with respect to any of the $\boldsymbol{L}^p$-norms, $1 \leq p \leq \infty$, but we try to introduce these Banach spaces of functions only much later in our notes. In fact, we make an attempt to *avoid the use of the classical Lebesgue spaces* and rather view them as completions of $\boldsymbol{C}_c(G)$ under the usual $\boldsymbol{L}^p$-norms (for $p < \infty$), thus emphasizing the importance of functional-analytic aspects over measure-theoretic considerations. We also hope to better convey the relevance of the concept of *generalized functions*, also called *distributions without basing it on the classical concepts* but rather through a direct (more integration theoretic than measure theoretic) approach.

This feature of our presentation should make it quite unique compared to most of the approaches to Fourier Analysis resp. (Abstract) Harmonic Analysis in the literature, and we hope that both the experts and newcomers to the field may enjoy the path taken!

Despite a lot of comments by many students, coworkers and colleagues the presentation still needs a lot of polishing and hints to short-comings, inconsistencies, repetitions or inaccuracies (solely due to the author) are *very welcome*!

The real stuff begins on the next page, finally! hgfei

3. Operators and conventions

Although we will not use the theory based on Lebesgue integration it is still good to know what the standard rules are on the standard spaces, such as $\boldsymbol{L}^1(\mathbb{R}^d)$. We will come back to this space later on.

Operators		
T_z	$T_z f(x) = f(x-z)$	translation by z
M_s	$M_s f(x) = e^{2\pi i s\cdot x} f(x)$	modulation operator
St_ρ	$\mathrm{St}_\rho f(x) = \rho^{-d} f(x/\rho)$	stretching operator
D_ρ	$\mathrm{D}_\rho f(x) = f(\rho x)$	dilation operator
	$f^{\checkmark}(x) = f(-x)$	flip operator
	$f^*(x) = \overline{f(-x)}$	$\boldsymbol{L}^1$-involution
	$\overline{f}(x) = \overline{f(x)}$	conjugation operator

Translation and modulation are isometric on *all the* $\boldsymbol{L}^p$-spaces, $1 \leq p \leq \infty$. The stretching operator is isometric an $\big(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\big)$, while D_ρ is isometric on $\big(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\big)$ hence $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$ (or $\big(\boldsymbol{L}^\infty(\mathbb{R}^d), \|\cdot\|_\infty\big)$).

Compatibility of operators	
$T_z \circ M_s = e^{-2\pi i s\cdot z} M_s \circ T_z$	translation with modulation
$\mathcal{F} \circ M_s = T_s \circ \mathcal{F}$	translation and Fourier
$M_s(g * f) = M_s f * M_s g$	modulation and convolution
$T_x(h \cdot f) = T_x h \cdot T_x f$	translation and multiplication
$\mathrm{D}_\rho(h \cdot f) = \mathrm{D}_\rho h \cdot \mathrm{D}_\rho f$	dilation and multiplication
$\mathrm{St}_\rho(g * f) = \mathrm{St}_\rho f * \mathrm{St}_\rho g$	stretching and convolution
$(f * g)^* = g^* * f^*$	convolution and involution
$\overline{h \cdot f} = \overline{h} \cdot \overline{f}$	multiplication and conjugation

This formula is describes the deviation (just a scalar of absolute value 1) from the commutative law. It is called the **Heisenberg commutation rule**.

This formula is justifies to call the modulation operators also frequency shift operators. It is equivalent to the statement that $M_s f = \mathcal{F}^{-1}(T_s \widehat{f})$.

We have for $1 \leq p \leq \infty$

Operators	
$\|T_z f\|_p = \|f\|_p$	translation by z
$\|M_s f\|_p = \|f\|_p$	modulation operator
$\|\mathrm{St}_\rho f\|_1 = \|f\|_1$	stretching operator
$\|\mathrm{D}_\rho f\|_\infty = \|f\|_\infty$	dilation operator

We also have a couple of adjointness relations, where the adjoint operators in the sense of the Hilbert space $\big(\boldsymbol{L}^2(\mathbb{R}^d),\, \|\cdot\|_2\big)$ an its standard scalar product, given by

defscalLt
$$\langle f, g\rangle = \int_{\mathbb{R}^d} f(x)\overline{g(x)}dx, \quad f, g \in \boldsymbol{L}^2(\mathbb{R}^d). \tag{7}$$

According to the general definition[10] this is the unique bounded linear operator T' such that

adjointT
$$\langle Tf, g\rangle = \langle f, T'g\rangle \tag{8}$$

Be aware of the difference between adjoint and conjugate operators, and a slight confusion of terminology. In terms of matrices, the difference between $\boldsymbol{A}^t$ and $\boldsymbol{A}'$, the transpose or the transpose conjugate matrix. of course there is no difference for the case of real matrices or real normed spaces.

It is not difficult to derive the existence from the Riesz Representation Theorem (identifying the dual space of $\big(\boldsymbol{L}^2(\mathbb{R}^d),\, \|\cdot\|_2\big)$ with the original space, every linear function on $\big(\boldsymbol{L}^2(\mathbb{R}^d),\, \|\cdot\|_2\big)$ is of the form of a scalar product!). Moreover one has[11]:

adjprop1
$$|||T||| = |||T'||| \quad \text{and} \quad (T')' = T. \tag{9}$$

For example (and as a good exercise to understand the claims in concrete situations) one has: $T_x' = T_{-x}$, $M_s' = M_{-s}$, and ${\mathrm{D}_\rho}' = \mathrm{St}_\rho$ resp. (equivalently) ${\mathrm{St}_\rho}' = \mathrm{D}_\rho$.

Sometimes also a unitary version for the Hilbert space is used, I would choose (an *ad hoc symbol*):

Ltscal1
$$\boxed{S_\rho f(z) = \rho^{-d/2} f(z/\rho), \quad \rho \neq 0.} \tag{10}$$

scalingLt
$$\|S_\rho f\|_2 = \|f\|_2, \qquad S_\rho' = S_{1/\rho} \quad \text{and} \quad \mathrm{supp}(S_\rho f) = \rho\cdot \mathrm{supp}(f). \tag{11}$$

3.1. **Banach spaces of bounded continuous functions.** For the linear space of all bounded (real or complex-valued) functions (by definition of the term *boundedness*, meaning that the range of f, i.e. the set of values $\{f(x), x \in \mathbb{R}^d\}$ is a bounded set, or

boundset
$$\exists\, C > 0 \quad \text{such that } |f(x)| \leq C, \quad \forall x \in \mathbb{R}^d. \tag{12}$$

Since for any such bounded function f the set of possible upper bounds is a non-empty set (and clearly one has that $C' \geq C$ implies that C' is another such bound) there is also a smallest upper bound, which is called

supdef02
$$\|f\|_\infty := \sup_{x\in\mathbb{R}^d}\{|f(x)|\} := \inf\{C \,|\, \text{satisfies estimate (12)}\}. \tag{13}$$

As an illustration of what the "size" (in the sense of the sup-norm) of a given bounded and periodic function f, in fact of a random trigonometric polynomial is, take a look at the figure. The red circle is the smallest circle such that the graph of f (a closed curve in the complex plain) is still inside the circle).

Another way to visualize the distance between two bounded and real-valued functions is the following one: Given $f, g \in \boldsymbol{C}_b(\mathbb{R})$ we have

$$\|f - g\|_\infty \leq \varepsilon$$

[10] In linear algebra this is just the involution of matrices $\boldsymbol{A} \mapsto \boldsymbol{A}'$, the *transpose conjugate* matrix!

[11] Analogous to the introduction of the conjugate transpose matrix for the vector space $\mathbb{C}^n$, where any operator T has a unique representation in the form $T(\mathbf{x}) = \boldsymbol{A} * \mathbf{x}$ for some $n \times n$-matrix $\boldsymbol{A}$. The same is true for the adjoint operator, which is represented by $\boldsymbol{A}'$ (using MATLAB notation for the conjugate transpose matrix).

maxnormdem1

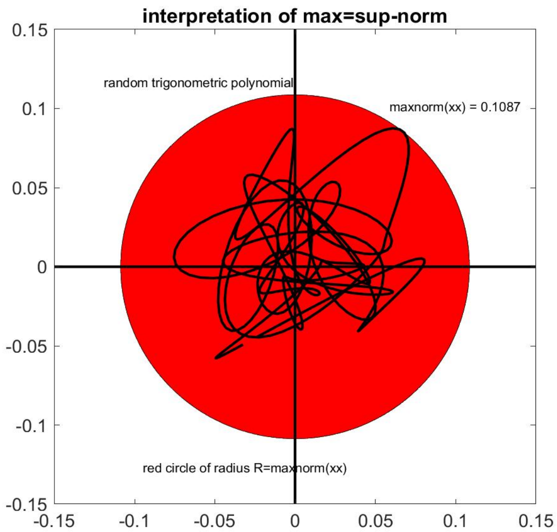


FIGURE 1. Demo of sup-norm for a periodic function [maxnormdem1.jpg]

strip01.eps

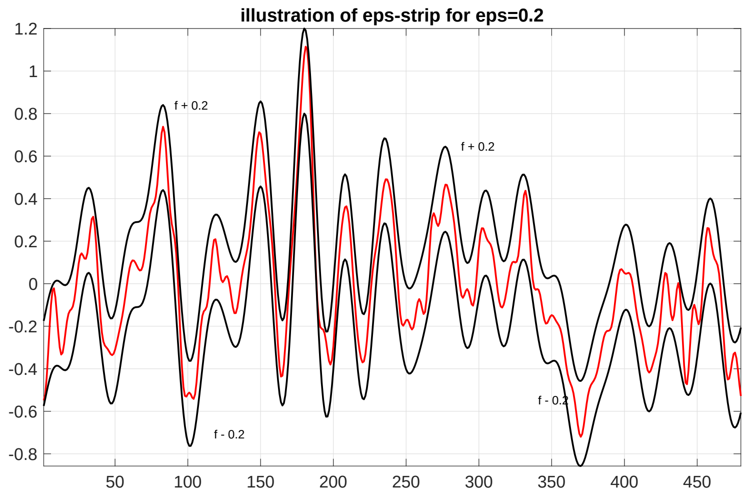


FIGURE 2. epsstrip01.eps

if and only if the graph of g is contained in an ε-strip around the graph of f. In fact, for real-valued functions the condition

$$|f(x) - g(x)| \le \varepsilon \quad \forall x \in \mathbb{R}^d$$

is obviously equivalent to either

$$f(x) - \varepsilon \le g(x) \le f(x) + \varepsilon, \quad \forall x \in \mathbb{R}^d,$$

or also

$$g(x) - \varepsilon \le f(x) \le g(x) + \varepsilon, \quad \forall x \in \mathbb{R}^d.$$

strip02.eps

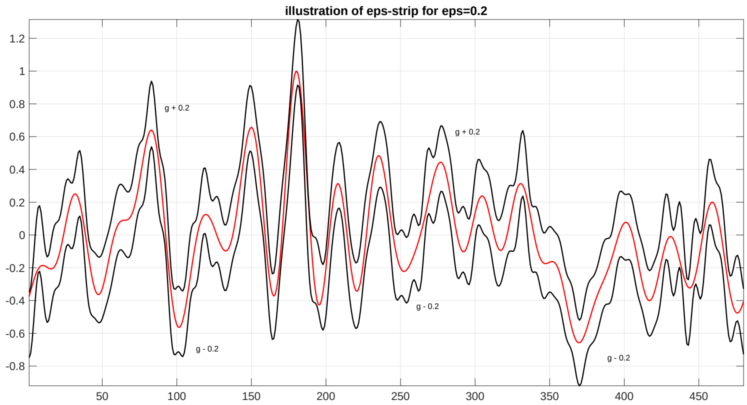


FIGURE 3. Approximation of black curve by a smooth version: epsstrip02.eps

3.2. **Sup-norm and algebras.** Let us recall that a vector space if called an algebra if it is endowed with some (abstract or concrete) multiplication, which is compatible with addition and scalar multiplication, meaning that one can use formulas such as

$$a \bullet (3b + 5c) = 3a \bullet b + 5a \bullet c.$$

We all know such things from many examples, maybe not explicitly by this name. First of all our number systems (the field $\mathbb{Q}$ of rational numbers, or of course also $\mathbb{R}$ and $\mathbb{C}$), but also all bounded functions, or all bounded and continuous functions, of $C^{(k)}(\mathbb{R})$, the set of all function which have continuous derivatives up to order k. So for $k = 2$[12] one assumes that f, f' and f'' are continuous.

The set $\mathscr{P}(\mathbb{R})$ of all polynomial functions over $\mathbb{R}$ is such an algebra. We all know that the product of two polynomials is again a polynomial (and later on we will discuss in great detail the pointwise multiplication of polynomials at the level of coefficients, the so-called *Cauchy product*).

All these examples are *commutative algebras*, because the order of (pointwise) multiplication does not matter.

Of course there are many non-commutative algebras, the most important being the set $\mathcal{M}_{n,n}$ of all $n \times n$-matrices (real or complex valued) with ordinary matrix multiplication.

A normed space $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ which is an algebra with a norm which is compatible with the multiplication $(a, b) \to a \bullet b$, i.e. satisfying the *submultiplicativity* estimate

submultNA (14) $$\|\mathbf{a} \bullet \mathbf{b}\|_{\boldsymbol{B}} \leq \|\mathbf{a}\|_{\boldsymbol{B}} \cdot \|\mathbf{b}\|_{\boldsymbol{B}}, \quad \mathbf{a}, \mathbf{b} \in \boldsymbol{B},$$

Cbdef **Definition 6** (**Banach spaces of continuous functions on $\mathbb{R}^d$**).

$$\boldsymbol{C}_b(\mathbb{R}^d) := \{f : \mathbb{R}^d \mapsto \mathbb{C}, \text{ continuous and bounded, with norm } \|f\|_\infty = sup_{z \in \mathbb{R}^d} |f(z)| \;\}$$

In order to demonstrate the important estimate

supsubmult (15) $$\|h \cdot f\|_\infty \leq \|h\|_\infty \cdot \|f\|_\infty, \quad h, f \in \boldsymbol{C}_b(\mathbb{R}^d),$$

we have to observe that by definition

supest1 (16) $$|f(x)| \leq \|f\|_\infty \quad \text{and} \quad |h(x)| \leq \|h\|_\infty \quad \forall x \in \mathbb{R}^d,$$

[12]These are relevant for curves in computer graphics, i.e. *cubic splines*, to be discussed later on in detail.

ultstr1.eps

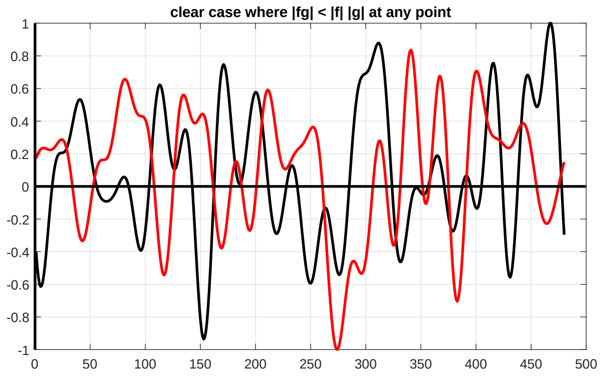


FIGURE 4. submultstr1.eps

and consequently

supest2
$$|[h \cdot f](x)| = |h(x)||f(x)| \leq \|h\|_\infty \|f\|_\infty, \quad \forall x \in \mathbb{R}^d, \tag{17}$$

i.e. $\|h\|_\infty \|f\|_\infty$ is an upper bound for the (absolute value) of the product function $|h \cdot f|$, and consequently the least upper bound (by definition the norm of this product function) is not larger, i.e. we have established the estimate (15).

Note that the triangle inequality (Exercise!) can be done in the same way. Replacing multiplication by addition above one verifies

supsubadd
$$\|f_1 \cdot f_2\|_\infty \leq \|f_1\|_\infty + \|f_2\|_\infty, \quad f_1, f_2 \in \boldsymbol{C}_b(\mathbb{R}^d), \tag{18}$$

supmultrem1 *Remark* 3. The plot below compares the graph of $|h| \cdot |f|$ (in black) with the graph of $|h \cdot f|$ (orange), with $|h|$ and $|f|$ for illustration (blue and red). For this example h and f have been chosen to be smooth (in fact band-limited) and real-valued functions.

In some regions (e.g. around coordinate 440) both functions have the same sign (i.e. either both positive or both negative)and thus the black and the orange curve coincide.

This explains (we leave it to the readers to check out details) that one has equality in (**??**) if and only if the maximum for $|h|$ and $|f|$ are taken at such points.

More generally, one has equality in (15) for complex-valued functions if and only if there exist points $x \in \mathbb{R}^d$ such that the phase of the complex-valued function equals at one of the points where both $|h|$ and $|f|$ attain the maximal value.

The spaces $\boldsymbol{C}_{ub}(\mathbb{R}^d)$ and $\boldsymbol{C}_0(\mathbb{R}^d)$ are defined as the subspaces of $\boldsymbol{C}_b(\mathbb{R}^d)$ consisting of functions which are *uniformly continuous* (and bounded) resp. *decaying at infinity*, i.e.

$$f \in \boldsymbol{C}_0(\mathbb{R}^d) \quad \text{if and only if} \quad \lim_{|x|\to\infty} |f(x)| = 0.$$

Cub-char **Lemma 2.** *i) Characterization of $\boldsymbol{C}_{ub}(\mathbb{R}^d)$ within $\boldsymbol{C}_b(\mathbb{R}^d)$ (or even $\boldsymbol{L}^\infty(\mathbb{R}^d)$).*

Cub-char0
$$\|T_x f - f\|_\infty \to 0 \quad \textit{for} \quad x \to 0 \quad \textit{if and only if } f \in \boldsymbol{C}_{ub}(\mathbb{R}^d). \tag{19}$$

ii) $\boldsymbol{C}_0(\mathbb{R}^d)$ is closed subspace of $\big(\boldsymbol{C}_{ub}(\mathbb{R}^d), \|\cdot\|_\infty\big)$.

Proof. The reformulation of the usual $\varepsilon - \delta$-definition into the format of (19) is left as an important exercise to the interested reader.

So let us only show how this characterization helps us to easily derive the fact that $\boldsymbol{C}_{ub}(\mathbb{R}^d)$ is a closed subspace of $\big(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\big)$. Assume that we have a sequence $(f_n)_{n=1}^\infty$ is a convergent sequence in $\boldsymbol{C}_{ub}(\mathbb{R}^d)$, with limit $f_0 \in \boldsymbol{C}_b(\mathbb{R}^d)$ (we know already that this is a Banach space). So we have to find for any given $\varepsilon > 0$ some $\delta > 0$ such that we can guarantee that

$$\|T_x f_0 - f_0\|_\infty < \varepsilon \quad \text{if only} \quad |x| \leq \delta.$$

In order to find such a $\delta > 0$ for given $\varepsilon > 0$ we first choose $n_0 \in \mathbb{N}$ such that $\|f_0 - f_{n_0}\|_\infty < \varepsilon/3$. Since $f_{n_0} \in \boldsymbol{C}_{ub}(\mathbb{R}^d)$ we can find for the positive number $\varepsilon/3 > 0$ some $\delta > 0$ such that $|x| < \delta$ implies that we have $\|T_x f_{n_0} - f_{n_0}\|_\infty < \varepsilon/3$. By the triangular equation we obtain the required estimate: Whenever $|x| < \delta$ it is ensured that

Cubclosed2
$$\|T_x f_0 - f_0\|_\infty \leq \|T_x(f_0 - f_{n_0})\|_\infty + \|T_x f_{n_0} - f_{n_0}\|_\infty + \|f_{n_0} - f_0\|_\infty < \varepsilon. \tag{20}$$

`There are still a few claims made above which need verification.` □

pointwise2 **Lemma 3.** (1) $\big(\boldsymbol{C}_{ub}(\mathbb{R}^d), \|\cdot\|_\infty\big)$ *is a Banach algebra under pointwise multiplication;*
(2) $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$ *is a closed ideal in* $\big(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\big)$, *meaning that one has*

$$\boldsymbol{C}_b(\mathbb{R}^d) \cdot \boldsymbol{C}_0(\mathbb{R}^d) \subseteq \boldsymbol{C}_0(\mathbb{R}^d),$$

or

$$h \cdot f \in \boldsymbol{C}_b(\mathbb{R}^d) \quad \forall h \in \boldsymbol{C}_b(\mathbb{R}^d), f \in \boldsymbol{C}_0(\mathbb{R}^d).$$

Proof. The first claim is a consequence of the fact that one has for $f_1, f_2 \in \boldsymbol{C}_{ub}(\mathbb{R}^d)$ [13].

shiftptw1
$$T_x(f_1 \cdot f_2) = T_x f_1 \cdot T_x f_2. \tag{21}$$

In addition we will use the isometric translation invariance of the sup-norm:

inftransiso
$$\|T_x f\|_\infty = \|f\|_\infty, \quad \forall f \in \boldsymbol{C}_b(\mathbb{R}^d), x \in \mathbb{R}^d. \tag{22}$$

We have to show that $f_1 \cdot f_2$ shows good behavior under translation. We will use the triangle inequality (inserting the term $T_x f_1 \cdot f_2$) and write the difference as a sum:

$$T_x(f_1 \cdot f_2) - (f_1 \cdot f_2) = (T_x f_1 - f_1) \cdot T_x f_2 + f_1 \cdot (T_x f_2 - f_2)$$

$$\|T_x(f_1 \cdot f_2) - f_1 \cdot f_2\|_\infty \leq \|(T_x f_1 - f_1) \cdot T_x f_2\|_\infty + \|f_1 \cdot (T_x f_2 - f_2)\|_\infty.$$

Using the *submultiplicativity* of the sup-norm we derive therefrom

upsubmult02
$$\|T_x(f_1 \cdot f_2) - f_1 \cdot f_2\|_\infty \leq \|T_x f_1 - f_1\|_\infty \|T_x f_2\|_\infty + \|f_1\|_\infty \|T_x f_2 - f_2\|_\infty. \tag{23}$$

If some $\varepsilon > 0$ is given, we can choose $\varepsilon' < \varepsilon/(\|f_1\|_\infty + \|f_2\|_\infty)$. Since both f_1 and f_2 belong to $\boldsymbol{C}_{ub}(\mathbb{R}^d)$ we can find some $\delta > 0$ such that

$$\|T_x f_k - f_k\|_\infty \leq \varepsilon' \quad \forall x \text{ with } |x| \leq \delta,\ k = 1, 2.$$

Inserting this estimate above and using (22) implies for $|x| \leq \delta$:

$$\|T_x(f_1 \cdot f_2) - (f_1 \cdot f_2)\|_\infty \leq \varepsilon'(\|f_1\|_\infty + \|f_2\|_\infty) \leq \varepsilon.$$

As for the inclusion $\boldsymbol{C}_b(\mathbb{R}^d) \cdot \boldsymbol{C}_0(\mathbb{R}^d) \subseteq \boldsymbol{C}_0(\mathbb{R}^d)$ let $h \in \boldsymbol{C}_b(\mathbb{R}^d)$ and $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ be given. We have $|h(x)| \leq \|h\|_\infty$ by the definition of the norm. In order to show that $h \cdot f \in \boldsymbol{C}_0(\mathbb{R}^d)$ we have to show for any given (fixed) $\varepsilon > 0$ that there exist some radius $R > 0$ such that for $|x| \geq R$ the values of the product function is smaller than ε.

As $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ we can find for the (known) constant $\varepsilon' = \varepsilon/\|h\|_\infty$ some $R > 0$ such that $|x| \geq R$ implies $|f(x)| \leq \varepsilon'$. This implies for x close to "infinity":

$$|h \cdot f(x)| := |h(x)||f(x)| \leq \|h\|_\infty \varepsilon' \leq \varepsilon,$$

and the expected conclusion has been verified. □

[13] This identity is written in an argument-free way, but when you want to check it one has to verify it pointwise, for all $x \in \mathbb{R}^d$.

pdecay1.eps

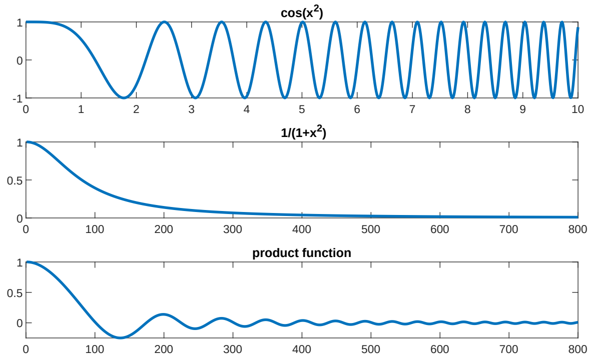


FIGURE 5. chirpdecay1.eps

As a visual illustration let us look at the real part of a chirp signal, i.e. $cos(x^2)$ multiplied with a bump function of the form $1/(1+x^2)$.

We will use the symbol $\mathcal{L}(\boldsymbol{C}_0(\mathbb{R}^d))$ for the Banach space of all bounded and linear operators on the Banach space $\boldsymbol{C}_0(\mathbb{R}^d)$, endowed with the operator norm (using various symbols, listed in the sequel[14]):

Topninf

$$\|T\|_{\mathcal{L}(\boldsymbol{C}_0(\mathbb{R}^d))} = |||T|||_{\infty} = \|T\|_\infty := sup_{\|f\|_\infty \leq 1} \|Tf\|_\infty. \tag{24}$$

opnorm02

Lemma 4. *The space $(\mathcal{L}(\boldsymbol{C}_0(\mathbb{R}^d)), \|\cdot\|_{\mathcal{L}(\boldsymbol{C}_0(\mathbb{R}^d))})$ is a Banach space, in fact a Banach algebra with respect to composition of operators. In particular, one has for $T_1, T_2 \in \mathcal{L}(\boldsymbol{C}_0(\mathbb{R}^d))$:*

opcompest1

$$|||T_2 \circ T_1|||_{\infty} \leq |||T_2|||_{\infty} |||T_1|||_{\infty}. \tag{25}$$

Proof. Let us just summarize the questions which have to be settled.

(1) First one has to show that $|||T|||_{\infty}$ defines a norm on $\mathcal{L}(\boldsymbol{C}_0(\mathbb{R}^d))$;
(2) Second one has to show that one can check completeness of this vector space, hence it is a Banach space;
(3) Finally one has to verify that (25).

WE WILL PUT MORE DETAILS INTO THIS SECTION AS NEEDED LATER ON, 22.09.2020, HGFEI ∎

It is worthwhile to observe that also the collection $\mathcal{M}_{n,n}$ of $n \times n$-matrices can be viewed as a Banach algebra, endowed with the operator norm, which is part of MATLAB. It is computed via the SVD (singular value decomposition) and is characterized by the fact that (for the Euclidean norm) `norm(A * x) <= norm(A) norm(x).`

trnormest01

$$\|\boldsymbol{A} * \mathbf{x}\|_{\mathbb{C}^n} \leq |||\boldsymbol{A}||| \, \|\mathbf{x}\|_{\mathbb{C}^n} \tag{26}$$

with the usual Euclidean norm $\mathbf{x} = \sqrt{\sum_{k=1}^n |x_k|^2}$.

[14]The perhaps most unusual symbol is the triple vertical line for *operator norms*, based on a given norm. So we will measure the size of an operator $T: \boldsymbol{B} \to \boldsymbol{B}$, where $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ is a normed space by the operator norm, and we write $|||T|||_{\boldsymbol{B}}$ in this situation.

4. Banach algebras of bounded and continuous functions

banalg-de **Definition 7** (**Banach algebra**). A Banach space $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ is a *Banach algebra* if it has a *bilinear multiplication* $(a,b) \to a \bullet b$ (or simply $a \cdot b$ or just ab, this means that it is also *associative and distributive*) with the extra property that for some constant $C > 0$[15]

Banalgsubm
$$\|a \bullet b\|_{\boldsymbol{A}} \leq C\|a\|_{\boldsymbol{A}}\|b\|_{\boldsymbol{A}} \quad \forall a, b \in \boldsymbol{A}. \tag{27}$$

Note/comment: Often the situation arises if a given Banach space is somehow embedded into $\mathcal{L}(\boldsymbol{B})$, and if one can show that it is closed under composition. Then the internal multiplication is taken from the composition of operators and associativity is no issue!! This is essentially what we have done when we have introduced *convolution* in $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$. The identification of $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ with the closed subalgebra of all linear operators on $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$ which commute with gives a Banach algebra structure on $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ (which we have called convolution).

Another simple example would be the identification of $\mathbb{R}^n$ with the set of diagonal matrices of size $n \times N$, which act on $\mathbb{R}^n$. Of course the corresponding Banach algebra structure on $\mathbb{R}^n$ is just the algebra structure under coordinate-wise (or pointwise) multiplication.

Again, a good exercise is to give a proof of the following lemma:

BanAlgdens1 **Lemma 5.** *Assume that $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ is a normed algebra, with completion $\widetilde{B} = (\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$. Then $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ is also a Banach algebra (with the same constant $C > 0$), and multiplication extends from a bilinear mapping on $\boldsymbol{B}$ in a unique way to a bilinear mapping on $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$, i.e. whenever $a_0 = \lim_{n\to\infty} a_n$ and $b_0 = \lim_{n\to\infty} b_n$, then*

bilinalg0
$$\lim_{n\to\infty} (a_n \bullet b_n) = a_0 \bullet b_0. \tag{28}$$

In fact, one even has: for $\varepsilon > 0$ there exists n_0 such that $n, m \geq n_0$ implies

bilinalg0
$$\|a_n \bullet b_m - a_0 \bullet b_0\|_{\boldsymbol{A}} < \varepsilon. \tag{29}$$

Hence also iterated limit such as $\lim_{n\to\infty} \lim_{m\to\infty}$ (or $\lim_{m\to\infty} \lim_{n\to\infty}$ exist and are equal to $a_0 \bullet b_0$.

subalg **Definition 8.** A *subalgebra* of a Banach algebra $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ is a closed subspace $\boldsymbol{B}$ of the given algebra, which is a Banach algebra of its own right with respect to the multiplication inherited from $\boldsymbol{A}$, i.e. with the (only) extra property that $\boldsymbol{B} \bullet \boldsymbol{B} \subseteq \boldsymbol{B}$.

Right and *two-sided ideals* are defined correspondingly[16].

BanAlgHom1 **Definition 9.** A bounded linear mapping from one Banach algebra $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ into another Banach algebra is called a *Banach algebra homomorphism* if it also respects multiplication, i.e.if

$$T(\mathbf{a}_1 * \mathbf{a}_2) = T(\mathbf{a}_1) \bullet T(\mathbf{a}_2), \quad \mathbf{a}_1, \mathbf{a}_2 \in \boldsymbol{A}.$$

An *isomorphism* of Banach spaces is a Banach algebra isomorphism if it is also a Banach algebra morphism.

Sometimes Banach algebras allow in addition for some *involution*, i.e. some isometric and linear algebra automorphism $j : \mathbf{b} \to \mathbf{b}^*$ (i.e. with $\|\mathbf{b}^*\|_{\boldsymbol{B}} = \|\mathbf{b}\|_{\boldsymbol{B}}$) and $j(j(\mathbf{b})) = \mathbf{b}$, or $(\mathbf{b}^*)^* = \mathbf{b}$ and $(\mathbf{a} \bullet \mathbf{b})^* = \mathbf{b}^* \bullet \mathbf{a}^*$ for all $\mathbf{a}, \mathbf{b} \in \boldsymbol{B}$. Such Banach algebras are called B^*-*algebras*.

Summarizing what we have verified so far we have the following theorem:

[15]Without loss of generality one can assume $C = 1$, because for the case that $C \geq 1$ one moves on to the equivalent norm $\|a\|'_{\boldsymbol{A}} := C \cdot \|a\|_{\boldsymbol{A}}$.

[16]Obviously any ideal is a subalgebra.

CbC0thm1 **Theorem 1.** *(Banach algebras of continuous functions)*

(1) $\big(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\big)$ *is a Banach algebra with respect to pointwise multiplication, and actually a B^*-algebra. This means that the Banach algebra has an involution, namely $j : f \mapsto \bar{f}$ which is linear, isometric, self-inverse (meaning $\bar{\bar{f}} = f$) and satisfies $\|\bar{f}\|_\infty = \|f\|_\infty$, and compatible with multiplication: $\overline{f \cdot g} = \overline{f} \cdot \overline{g}$. Similar properties hold true for the involutions $f \mapsto f^{\checkmark}$ and $f \mapsto f^* := \overline{f^{\checkmark}} = \overline{f}^{\checkmark}$.*

(2) $\big(\boldsymbol{C}_{ub}(\mathbb{R}^d), \|\cdot\|_\infty\big)$ *is a closed subalgebra of* $\big(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\big)$.

(3) $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$ *is a closed ideal within* $\big(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\big)$.

C0-char1 **Lemma 6.** *Characterization of $\boldsymbol{C}_0(\mathbb{R}^d)$ within $\boldsymbol{C}_b(\mathbb{R}^d)$: $\boldsymbol{C}_0(\mathbb{R}^d)$ coincides with the closure of $\boldsymbol{C}_c(\mathbb{R}^d)$ within $\big(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\big)$.*

Proof. Although this statement is quite intuitive one should try to see what one needs.

The first natural idea (for $d = 1$) is to "cut off" a function and separate the "big value" in some interval $[-N, N]$ from the small values (near "infinity") and then continue the restriction of these values in a continuous way, producing thus a continuous function $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ which is close to a given $f \in \boldsymbol{C}_0(\mathbb{R}^d)$.

However, it is better to think of this process using the idea of an approximation process (not just for an individual function $f \in \boldsymbol{C}_0(\mathbb{R}^d)$, but rather to the whole space), and to do it by pointwise multiplication with a given family of functions which we will call *plateau-like* functions. $\square$

Definition 10. A directed family (a *net* or *sequence*) $(h_\alpha)_{\alpha \in I}$ in a Banach algebra $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ is called a **BAI** (= bounded approximate identity or "approximate unit" for $\boldsymbol{A}$) if

defBAI1 (30)
$$\lim_\alpha \|h_\alpha \cdot h - h\|_{\boldsymbol{A}} = 0 \quad \forall h \in \boldsymbol{A},$$

and if $M := \sup_\alpha \|h_\alpha\|_{\boldsymbol{A}} < \infty$ (the bound of the BAI).

It is a good exercise to verify the following lemma.

testBAI1 **Lemma 7.** *Assume there is a total subset M of a Banach algebra $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$, and a bounded family $(h_\alpha)_{\alpha \in I}$ with satisfies (30) for all $b \in M$, then this net/sequence is a BAI for $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$.*

Proof. First one verifies the validity of (30) for finite linear combinations of elements from M. Then one proceeds by approximation, using the *boundedness* of the family $(h_\alpha)_{\alpha \in I}$. $\square$

It might be good to use a particular symbol for the BAIs, similar to the symbol Psi for BUPUs (22.04.2020, hgfei), then M could be replaced by —FAM— or —FAM—A.

As a TEST let us use the following notation:

$$\clubsuit\mathbf{h} = (h_\alpha)_{\alpha \in I}$$

or, in order to indicate the index set:

$$\clubsuit_I\mathbf{h} = (h_\alpha)_{\alpha \in I}$$

Then one can use another family, using the symbol

$$\clubsuit\mathbf{p} = (p_\beta)_{\beta \in J}$$

or more explicit (and cumbersome)

$$\clubsuit_J\mathbf{p} = (p_\beta)_{\beta \in J}$$

Then we would simply write $M = |\clubsuit\mathbf{h}|$, $|\clubsuit_I\mathbf{h}|$, $|\clubsuit_L\mathbf{p}|$

$$\clubsuit\mathbf{h} = (h_\alpha)_{\alpha\in I}$$

$$\clubsuit_I\mathbf{h} = (h_\alpha)_{\alpha\in I}$$

$$\clubsuit\mathbf{p} = (p_\beta)_{\beta\in J}$$

$$\clubsuit_J\mathbf{p} = (p_\beta)_{\beta\in J}$$

We need the following definition of a value-preserving dilation operator (the function values are just displaced to new positions by some dilation factor):

Drhodef **Definition 11. [Rd-specific!]**The value preserving **dilation operators** are given by:

Drhodef1 (31) $$D_\rho f(z) = f(\rho \cdot z), \ \forall \rho \neq 0, \, z \in \mathbb{R}^d$$

For $\rho = -1$ a special symbol will be used:

checkdef0 (32) $$f^\checkmark(z) = [D_{-1}f](z) = f(-z), \quad \forall f \in \boldsymbol{C}_0(\mathbb{R}^d), z \in \mathbb{R}^d.$$

It is easy to verify that

Drho2 (33) $$\|D_\rho f\|_\infty = \|f\|_\infty \quad \text{and} \quad D_\rho(f\cdot g) = D_\rho(f)\cdot D_\rho(g).$$

For later use let us mention that the dilation and translation operators satisfy the following *commutation relation* **[Rd-specific!]**:

-trans-comm (34) $$D_\rho \circ T_x \ = \ T_{x/\rho} \circ D_\rho\,, \quad \rho > 0, x \in \mathbb{R}^d,$$

and in particular

-trans-comm (35) $$[T_x f]^\checkmark \ = \ T_{-x} f^\checkmark, \quad \text{and} \quad [f^\checkmark]^\checkmark = f, \quad \forall f \in \boldsymbol{C}_0(\mathbb{R}^d).$$

Proof. For any $f \in \boldsymbol{C}_b(\mathbb{R}^d)$ one has for any $z \in \mathbb{R}^d, \rho \neq 0$:

$$[D_\rho(T_x f)](z) = (T_x f)(\rho z) = f(\rho z - x) = f(\rho(z - x/\rho)) = (D_\rho f)(z - x/\rho) = [T_{x/\rho}(D_\rho f)](z).$$

□

Theorem 2. $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$ *is a Banach algebra with bounded approximate units. In fact, any family of functions* $(h_\alpha)_{\alpha\in I}$ *which is uniformly bounded, i.e. with*

supbd1 (36) $$|h_\alpha(x)| \leq C < \infty \qquad \forall x \in \mathbb{R}^d, \forall \alpha{\in}I$$

and satisfies

unifcomp1 (37) $$\lim_\alpha h_\alpha(x) = 1 \quad \textit{uniformly over compact sets}$$

constitutes a BAI (the converse is true as well).

[Rd-specific!]*: In particular, one obtains a BAI by stretching any function* $h_0 \in \boldsymbol{C}_0(\mathbb{R}^d)$ *with* $h_0(0) = 1$*, i.e. by considering the family* $D_\rho h_0(g)(t) := h_0(\rho \cdot t)$*, for* $\rho \to 0$*.*

Proof. First of all the boundedness with respect to the sup-norm is just (36). That the condition (37) is a necessary condition is a necessary condition is obvious, by applying the family $(h_\alpha)_{\alpha\in I}$ to a collection of plateau-like functions.

Given a compact set $K \subset \mathbb{R}^d$ one chooses any function $p \in \boldsymbol{C}_0(\mathbb{R}^d)$ with $\|p\|_\infty \leq 1$ and $p(x) \equiv 1$ for $x \in K$. In order to show that $|1 - h_\alpha(x)| < \varepsilon$ for $\alpha \succeq \alpha_0$ one fixes α_0 such that

$$\|h_\alpha \cdot p - p\|_\infty < \varepsilon, \quad \text{for } \alpha \succeq \alpha_0.$$

This implies for $x \in K$:

$$|1 - h_\alpha(x)| = |(h_\alpha \cdot p - p)(x)| \leq \|h_\alpha \cdot p - p\|_\infty < \varepsilon,$$

thus completing this argument.

Assume conversely that one has (37). Then it is clear that one has norm convergence $\|k - h_\alpha \cdot k\|_\infty\| \to 0$ for any $k \in \boldsymbol{C}_c(\mathbb{R}^d)$, and thus by Lemma 7 for all $f \in \boldsymbol{C}_0(\mathbb{R}^d)$. □

The following elementary lemma is quite standard and could in principle be left to the reader. Probably it should be placed in the appendix. However, it is quite typical for arguments to be used repeatedly throughout these notes and therefore we state it explicitly.

unifcomp01 **Lemma 8.** *Assume that one has a bounded net $(T_\alpha)_{\alpha \in I}$ of operators from a Banach space $(\boldsymbol{B}^1, \|\cdot\|^{(1)})$ to another normed space $(\boldsymbol{B}^2, \|\cdot\|^{(2)})$, such that $T_\alpha \to T_0$ strongly, i.e. for any finite set $F \subset \boldsymbol{B}^{(1)}$ and $\varepsilon > 0$ there exists an index α_0 such that for $\alpha \succeq \alpha_0$ one has $\|T_\alpha f - T_0 f\|_{\boldsymbol{B}^2} < \epsilon$ for all $f \in F$, then one has uniform convergence over compact subsets $M \subset \boldsymbol{B}^1$ (this aspect will be discussed in more detail later on).*

Proof. The argument is based on the usual compactness argument. Assuming that $|||T_\alpha||| \leq C_1$ for all $\alpha \in I$ we can find some finite set $f_1, \dots f_K \in M$ such that balls of radius $\delta = \varepsilon/(3C_1)$ around these points cover M. According to the assumption (and the general properties of nets) one finds α_0 such that

$$\|T_\alpha f_j - T_0 f_j\|_{\boldsymbol{B}^{(2)}} < \varepsilon/3, \quad \text{for } j = 1, 2, \dots K.$$

Consequently one has for any $f \in M$ and suitable chosen index j (with $\|f - f_j\|_{\boldsymbol{B}^{(1)}} < \delta$):

comp-convg
$$\|T_\alpha f - T_0 f\|_{\boldsymbol{B}^{(2)}} < \|T_\alpha(f - f_j)\|_{\boldsymbol{B}^{(2)}} + \|T_\alpha f_j - T_0 f_j\|_{\boldsymbol{B}^{(2)}} + \|T_0(f_j - f)\|_{\boldsymbol{B}^{(2)}}. \tag{38}$$

Since the operator norm of T_0 is also not larger than C_1 (! easy exercise) this implies

comp-convg2
$$\|T_\alpha f - T_0 f\|_{\boldsymbol{B}^{(2)}} \leq 2C_1\|f - f_j\|_{\boldsymbol{B}^{(1)}} + \|T_\alpha f_j - T_0 f_j\|_{\boldsymbol{B}^{(2)}} \leq 2C_1\delta + \varepsilon/3 < \varepsilon. \tag{39}$$

□

total-test1 *Remark* 4. In fact, a similar argument can be used to verify the w^*-convergence of a bounded net of operators, by verifying the convergence only for f from a total subset within its domain. In fact, using linearity implies that convergence is true for linear combinations of the elements for such a set, and by going to a limit one obtains convergence for arbitrary elements in $\boldsymbol{B}^{(1)}$.

The above argument applies also to nets of bounded linear functionals (choosing $\boldsymbol{B}^{(2)} = \mathbb{C}$).

DrhoGrpCo **Lemma 9. [Rd-specific!]***[Group of dilation operators]*
The family of operators $(D_\rho)_{\rho>0}$ is a family of isometric isomorphisms on $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$. Moreover, the mapping $\mathbb{R}^+ \to \mathcal{L}(\boldsymbol{C}_0(\mathbb{R}^d))$ given by $\rho \mapsto D_\rho$, is a group homomorphism from the multiplicative group of positive reals into the isometric linear operators on $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$, i.e. one has

dilcomut1
$$D_{\rho_2} \circ D_{\rho_1} = D_{\rho_1 \cdot \rho_2} = D_{\rho_1} \circ D_{\rho_2} \tag{40}$$

It is however not continuous in the operator sense, since $|||D_{\rho_1} - D_{\rho_2}|||_{\boldsymbol{L}^\infty} \geq 1$ for $\rho_1 \neq \rho_2$, but it is strongly continuous, i.e. for any $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ one has:

Drhostrongc
$$\lim_{\rho \to 1} \|\mathrm{D}_\rho f - f\|_\infty = 0. \tag{41}$$

Proof. The lemma contains a number of useful, but easy to verify claims. The fact that the operator norm of two operators of two different dilation operators does not tend to zero for $\rho_1 \to \rho_2$ follows from the observation that for any such pair of dilation parameters there exist (in fact many) functions $k \in \boldsymbol{C}_c(\mathbb{R}^d), \|k\|_\infty = 1$, which small support such that their dilates have disjoint support. Note $\operatorname{supp}(k) = K$ then $\operatorname{supp}(\mathrm{D}_\rho k) = K/\rho$, so it is enough to choose K such that $\rho_2 K \cap \rho_1 K = \emptyset$.

The positive statement (about strong continuity) follows from the observation that it is a consequence of the *uniform continuity* of $k \in \boldsymbol{C}_c(\mathbb{R}^d)$, plus the usual approximation argument for the general case $f \in \boldsymbol{C}_0(\mathbb{R}^d)$. $\square$

Remark 5. One can make similar claims about general dilations, based on matrices, including *anisotropic dilations*. Given $\boldsymbol{A} \in \mathcal{M}_{d,d}(\mathbb{R})d$ the generalized dilation [17]

dilatmatr1 (42)
$$\mathrm{D}_{\boldsymbol{A}}(f)(\mathbf{z}) := f(\boldsymbol{A} * \mathbf{z}), \quad \mathbf{z} \in \mathbb{R}^d.$$

Then again

DrhostrongM (43)
$$\|\mathrm{D}_{\boldsymbol{A}} f - f\| \to \mathbf{0} \text{ if } \boldsymbol{A} \to \mathbf{Id_d} \text{ in } \mathcal{M}_{d,d}(\mathbb{R})d.$$

-approxunit **Lemma 10.** *Let I be the family of all compact subsets of $\mathbb{R}^d$, and define $K \succeq K'$ if $K \supseteq K'$. If we choose for every such $K \subset\subset \mathbb{R}^d$ [18] a plateau function p_K such that $0 \leq p(x) \leq 1$ on $\mathbb{R}^d$ and $p_K(x) \equiv 1$ on K. Then $(p_K)_{K \in I}$ constitutes a BAI for $\boldsymbol{C}_0(\mathbb{R}^d)$.*

Proof. First we have to show that the "direction" of I is reflexive ($K \supseteq K$ and transitive, i.e. $K_1 \supseteq K_2$ and $K_2 \supseteq K_3$ obviously implies $K_1 \supseteq K_3$). Finally the key property for the index set of a net is easily verified: Given two "indices" K_1, K_2 the set $K_0 := K_1 \cup K_2$ is an element of I, i.e. a compact set, with $K_0 \succeq K_i$ for $i = 1, 2$. Hence $(p_K)_{K \in I}$ is in fact a net.

In order to verify the BAI property let $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ and $\varepsilon > 0$ be given. Then - by definition of $\boldsymbol{C}_0(\mathbb{R}^d)$ there exists some $R > 0$ such that $|x| > R$ implies $|f(x)| \leq \varepsilon$. Then one has $|f(x) \cdot p(x) - f(x)| = |(1 - p(x))||f(x)| \leq |f(x)| \leq \varepsilon$ for $|x| \geq R$. On the other hand $|f(x) \cdot p(x) - f(x)| = |f(x) \cdot 1 - f(x)| = 0$ for $|x| \leq R$, as long as p is a plateau function equal to 1 on the compact set $K_0 = B_R(0)$ [19], i.e. as long as $p = p_K$ has an index with $K \succeq K_0$. Altogether $\|f \cdot p - f\|_\infty \leq \varepsilon$ $\square$

The family of all partial sums of BUPUs (defined below) will satisfy the conditions described in Lemma 10. The following figure illustrates the situation for $G = \mathbb{R}$ (the partial sum adds up to a nice *plateau function* with compact support, covering some region of interest.

char-C0inCb **Lemma 11.** *Another characterization of $\boldsymbol{C}_0(\mathbb{R}^d)$ within $\boldsymbol{C}_b(\mathbb{R}^d)$:*
$h \in \boldsymbol{C}_b(\mathbb{R}^d)$ belongs to $\boldsymbol{C}_0(\mathbb{R}^d)$ if and only if

char-C01 (44)
$$\lim_\alpha \|h_\alpha \cdot h - h\|_\infty = 0$$

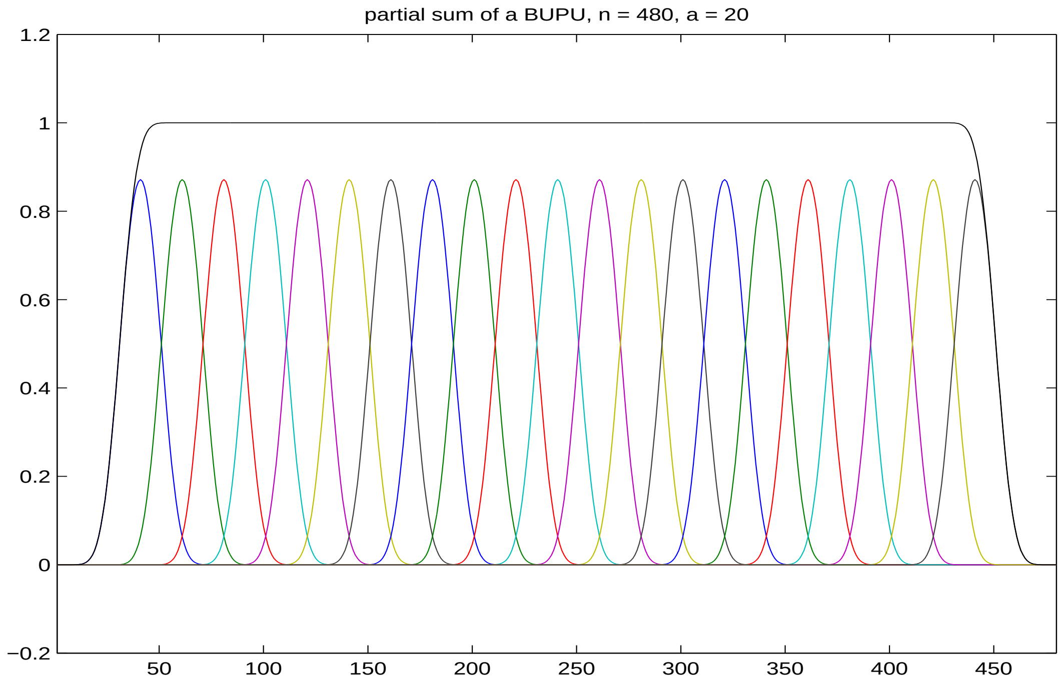
for one (hence all) BAIs for $\boldsymbol{C}_0(\mathbb{R}^d)$ (as described above).

Remark 6. It is a good exercise that such a BAI acts *uniformly* on compact subsets, i.e. for any relatively compact set M in $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$ one finds: Given $\varepsilon > 0$ one can find some α_0 such that for $\alpha \succeq \alpha_0$ implies

char-C02 (45)
$$\|h_\alpha \cdot h - h\|_{\boldsymbol{B}} \leq \varepsilon \quad \forall h \in M.$$

[17] We write $\boldsymbol{A} * \mathbf{z}$ for matrix-vector multiplication, following MATLAB^{TM}.
[18] We write $K \subset\subset \mathbb{R}^d$ to indicate that K is a *compact* subset of $\mathbb{R}^d$.
[19] $B_R(0)$ denotes the ball of radius R around 0.

labelbupu480dem1.pdf



The following simple lemma will be useful later on (characterization of LTIS, i.e. *linear time-invariant systems*, also called TILS, i.e. *translation-invariant linear systems* for $d = 2$ we can talk about linear space-invariant operators, as they are used for image processing applications):

homog-BF1 **Lemma 12.** *A function* $f \in \boldsymbol{C}_b(\mathbb{R}^d)$ *belongs to* $\boldsymbol{C}_{ub}(\mathbb{R}^d)$ *if and only if the (non-linear) mapping* $z \mapsto T_z f$ *is continuous from* $\mathbb{R}^d$ *into* $\big(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\big)$*. In fact, such a mapping is continuous at zero if and only if it is uniformly continuous.*

Proof. It is clear that continuity is a necessary condition (and we have already seen that it is equivalent to uniform continuity for $f \in \boldsymbol{C}_b(\mathbb{R}^d)$). Conversely assume continuity at zero, i.e. $\|T_z f - f\|_\infty < \varepsilon$ for sufficiently small z ($|z| < \delta$). Then one derives continuity at x as follows:

$$\|T_{x+z} f - T_x f\|_\infty = \|T_x(T_z f - f)\|_\infty = \|T_z f - f\|_\infty < \varepsilon$$

implying uniform continuity of the discussed mapping. [20] □

Remark 7. One can say, that the mapping $x \mapsto T_x$ from $\mathbb{R}^d$ into $\mathcal{L}(\boldsymbol{C}_0(\mathbb{R}^d))$ is a representation of the Abelian group $\mathbb{R}^d$ on the Banach space $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$, which by definition means that the mapping is a homomorphism between the additive group $\mathbb{R}^d$ and the group of invertible (even isometric) operators on $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$. The additional property that $z \mapsto T_z f$ is continuous for $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ is referred to as the *strong continuity* of this representation. The same mapping into

[20] The proof gives an argument, that a homomorphism from an Abelian topological group G ($\mathbb{R}^d$ in our case) into a group of operators on a Banach space $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ (here $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$) is strongly continuous, i.e. satisfies the discussed continuity property analogue to $f \mapsto T_z f$, is continuous from G into $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ if and only it is continuous at zero (the identity element of G).

the larger space $\mathcal{L}(\boldsymbol{C}_b(\mathbb{R}^d))$ would not be strongly continuous, because for $f \in \boldsymbol{C}_b(\mathbb{R}^d) \setminus \boldsymbol{C}_{ub}(\mathbb{R}^d)$ this mapping fails to be continuous (cf. below)

Remark 8. One can derive from the above definition that the mapping $(x, f) \to T_x f$, which maps $\mathbb{R}^d \times \boldsymbol{C}_0(\mathbb{R}^d)$ into $\boldsymbol{C}_0(\mathbb{R}^d)$ is continuous with respect to the product topology (Exercise).

Definition 12. We denote the dual space of $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$ with $(\boldsymbol{M}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}})$. Sometimes the symbol $\boldsymbol{M}_b(\mathbb{R}^d)$ is used in order to emphasize that one has "bounded" (regular Borel) measures.

The *Riesz Representation Theorem* (see e.g. [6], p.112, as well as many other sources, probably also [7], or [2]) provides the justification for this definition and establishes the link to the concept explained in *measure theory.* In that case the action of μ on the test function is of course written in the form

meas-act (46)
$$\mu(f) = \int_{\mathbb{R}^d} f(t) d\mu(t).$$

In the classical case of functionals on $\boldsymbol{C}(I)$, where $I = [a, b]$ is some interval, one can describe these integrals using Riemann-Stieltjes integrals. They make use of functions F of bounded variation. The distribution function [21] F is connected with the *measure* μ, defined on the σ-algebra of Borel sets in I via

dist-meas (47)
$$F(x) = \int_a^x d\mu(x); \qquad \int_I f(x) d\mu(x) = lim_{\delta \to 0} \sum_i f(\xi_i)[F(x_i) - F(x_{i-1})].$$

The original approach to $\big(\boldsymbol{L}^1(\mathbb{R}), \|\cdot\|_1\big)$ was within this context: A closed subspace of the linear space of all functions of bounded variation $BV(\mathbb{R})$ is the space of *absolutely continuous functions* (cf. historical notes, in particular see the article of 1929 by Plessner [76], where he shows that these are exactly the elements within BV which have continuous translation, hence are approximated using e.g. convolution by nice summability kernels.[22])

In this context also the "main theorem of analysis" (popularly known by the claim: taking the anti-derivative F is the an operation which is - up to additive constants - an operation which is inverse to the differentiation operator): A function is the distribution function of some function in $\boldsymbol{L}^1$ if and only if it is absolutely continuous. In this case one can be assured that $F'(x)$ (taken in the *classical sense*) exists almost everywhere, that the resulting function is measurable and in fact in $\boldsymbol{L}^1$, and the absolutely continuous function F can be recovered (up to an additive constant) from its derivative by taking the indefinite integral (area function).

Of course the norm (measurement of the *size of total variation*), the so called BV-norm of F and the norm of the linear functional turn out to be the same. Moreover, there is natural way to split a function F of bounded variation into a difference of two non-decreasing functions, intuitively the "increasing" and the "decreasing part"[23]. The sum of these two parts (as opposed to the difference) is a monotonous function which induces therefore a non-negative measure of the same total variation, which is called $|\mu|$ in the general measure theory (Bochner's decomposition), the absolute value of the measure μ (just a symbolic operation). Normally the

[21] This use of the word distribution is quite different from the use of distributions in the sense of generalized functions, as it is used in the rest of these notes.

[22] It is also interesting to see how most of the still relevant facts concerning Fourier series, Lebesgue integral and so on appear already in a rather clear form in those early papers and books on the subject. However, this was before the age of Banach space theory and in fact not at all referring to abstract Hilbert space theory, as it is seen nowadays as a corner stone. For further comments see the [forthcoming] section on historical notes.

[23] Think of a hiker how walks up and down in the mountains and reports on the total height of ascent and descent part.

(non-linear) mapping $\mu \mapsto |\mu|$ requires quite some measure theory (see e.g. [2]). There is a way to provide a proof of this decomposition using the approximation by discrete measures, in combination with the "obvious" decomposition of a real-valued discrete measure into a positive and a negative a part (splitting the non-zero coefficients into two groups: the positive and the negative ones). The proof (the author just has some handwritten notes) is however technically somewhat involved and will be given in a separate publication. It goes beyond the scope of these lecture notes, and will not be needed elsewhere.

Remark 9. As a compromise between the simple Riemannian integral and the high-level Lebesgue integral (which is *complete* with respect to the $\boldsymbol{L}^1$-norm!) people sometimes restrict their attention to the closure of the step functions in the sup-norm (this is a space containing both the step functions and the continuous functions), if we work over an interval or bounded subset of $\mathbb{R}^d$. One could use the closure of the simple (i.e. defined by indicator functions of $d-$dimensional boxes or cubes) in the sense of the Wiener norm of $\boldsymbol{W}(\boldsymbol{L}^1, \boldsymbol{\ell}^\infty)(\mathbb{R}^d)$ (see Sections 11 and 33 below).

EXAMPLES: point measures $\delta_0 : f \mapsto f(0)$, $\delta_u : f \mapsto f(u)$ for some fixed $u \in \mathbb{R}^d$, or integrals over bounded sets: $f \mapsto \int_a^b f(x)dx$, or more generally $\mu_k : f \mapsto \int_{\mathbb{R}^d} f(x)k(x)dx$, where integration is taken in the sense of Riemannian (or Lebesgue) integrals, for $k \in \boldsymbol{C}_c(\mathbb{R}^d)$. For the setting of locally compact groups G one will use the *Haar measure* for the definition of such integrals.

Remark 10. Again we have to comment on the use of symbols in the engineering literature. What we denote by δ_0 (the *Dirac measure* or *point measure* sitting at 0 is often just called (by a kind of misuse of terminology) the *Dirac delta-"function"* and denoted by $\delta(t)$ (the argument again indicating that it a "function" of e.g. the time-variable t). Consequently the shifted version of this function is described by the symbol $\delta(t-u)$ (corresponding exactly to our δ_u).

Note: the norm of $\mu \in \boldsymbol{M}(\mathbb{R}^d)$ is of course just the functional norm, i.e. (by the density of $\boldsymbol{C}_c(\mathbb{R}^d)$ in $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$:

munorm1

$$\|\mu\|_{\boldsymbol{M}} := \sup_{\|f\|_\infty \le 1} |\mu(f)| = \sup_{\|f\|_\infty = 1} |\mu(f)| = \sup_{k \in \boldsymbol{C}_c(\mathbb{R}^d) \,|\, \|k\|_\infty = 1} |\mu(f)|. \tag{48}$$

EXERCISE: $\|\delta_t\|_{\boldsymbol{M}} = 1$, and more generally:

Theorem 3. *For any finite linear combinations of Dirac-measures, i.e. for $\mu = \sum_{k\in F} c_k \delta_{t_k}$ (where F is some finite index set and we assume that one has the natural representation, with $t_k \neq t_{k'}$) one has $\|\mu\|_{\boldsymbol{M}_b} = \sum_{k\in F} |c_k|$.*

Proof. The upper estimate is simple. Applying $\sum_{k\in F} c_k \delta_{t_k}$ to f one gets the estimate

$$\left|\sum_{k\in F} c_k f(t_k)\right| \le \sum_{k\in F} |c_k|\,|f(t_k)| \le \|f\|_\infty \sum_{k\in F} |c_k|.$$

Conversely, one cannot get a better estimate, because it is possible for any sequence (t_k) of distinct points and corresponding complex coefficients to find a continuous functions (e.g. a linear combination of sufficiently narrow triangular functions) such that $\|f\|_\infty = 1$ and $f(t_k) = c_k/|c_k|$ (whenever $c_k \neq 0$). But then

$$\sum_{k\in F} |c_k f(t_k)| = \sum_{k\in F} |c_k|,$$

and the proof is complete. $\square$

iracnonconv **Corollary 1.**

$$\|\delta_x - \delta_y\|_{\boldsymbol{M}_b(\mathbb{R}^d)} = 2 \quad \textit{whenever} \quad x \neq y.$$

In particular, $x_n \to x_0$ does NOT imply $\delta_{x_n} \to \delta_{x_0}$ in the norm of $\boldsymbol{M}_b(\mathbb{R}^d)$.

In contrast, one has of course $\delta_{x_n}(f) \to \delta_{x_0}(f)$, since $f(x_n) \to f(x_0)$ for $f \in \boldsymbol{C}_b(\mathbb{R}^d)$.

A simple lemma helps us to identify the closed linear subspace of $(\boldsymbol{M}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}})$ generated from the (linear) subspace of finite-discrete measures. We will call it the space of *discrete measures.*

Lemma 13. *Let V be a linear subspace of a normed space $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$. Then the closure of V coincides with the space obtained by taking the absolutely convergent sequences in $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ with elements from V:*

bs-convser1 (49)
$$Abs(\boldsymbol{B}) := \{x = \sum_n v_n, \textit{ with } \quad \sum_n \|v_n\|_{\boldsymbol{B}} < \infty\}$$

Proof. It is obvious that $Abs(\boldsymbol{B}) \subseteq V^-$, the closure of V in $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$. Conversely, any element $x = lim_{k\to\infty} v_k$ can also be written as a limit of a (sub)sequence with $\|x - v_k\|_{\boldsymbol{B}} < 2^{-k}$. In fact, one can choose by induction a strictly increasing subsequence (v_{n_k}) of the original sequence such that this is valid (and relabel the elements afterwards), and may therefore be rewritten as a telescopic sum, with $y_1 = v_1$, $y_{n+1} = v_{n+1} - v_n$. .... □

Working out the details of the above argument constitutes a good exercise for the interested reader.

sconv-dense *Remark* 11. A simple but powerful variant of the above lemma is obtained if one allows the elements v_n being only taken from a dense subset of V (density of course in the sense of the $\boldsymbol{B}-$norm. Since such a set has the same closure this is an immediate corollary from the above lemma).

As a consequence (of the last two results) one finds that the closed linear subspace generated by the finite discrete measures coincides with absolutely convergent series of Dirac measures:

defMdRdef **Definition 13.** [Bounded discrete measures] The space of *bounded discrete measures* is defined by:

defMdRd1 (50)
$$\boldsymbol{M}_d(\mathbb{R}^d) := \{\mu \in \boldsymbol{M}(\mathbb{R}^d) : \mu = \sum_{k=1}^{\infty} c_k \delta_{t_k} \mid \sum_{k=1}^{\infty} |c_k| < \infty\}$$

The elements of $\boldsymbol{M}_d(\mathbb{R}^d)$ are called the *bounded discrete measures*, and thus we claim that they form a (proper) closed subspace of $(\boldsymbol{M}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}})$, and coincides with the norm-closure of the finite discrete measures within $(\boldsymbol{M}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}})$.

The following statement will be understandable once the concept of tightness is formally introduced (see Definition 40):

Definition 14. A sequence of measures $(\mu_n)_{n\geq 1}$ is *Bernoulli convergent* to some $\mu_0 \in \boldsymbol{M}(\mathbb{R}^d)$ according to Bochner (bo55-1 [10, p.15), if it is bounded in $(\boldsymbol{M}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}})$ and if

$$\mu_n(f) \to \mu_0(f) \quad \text{for all} \quad f \in \boldsymbol{C}_b(\mathbb{R}^d).$$

CbExtMb02 **Lemma 14.** *The action of any $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ extends in a natural way to all of $\boldsymbol{C}_b(\mathbb{R}^d)$, and the extended functional (which we simply denote by the same symbol) has the same norm as element of $\big(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\big)'$. The convolution operator C_μ extends to a bounded linear operator on $\big(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\big)$ of norm $\|\mu\|_{\boldsymbol{M}_b}$, commuting with translation.*

Proof. We make use of Thm. (4). For any given BUPU Ψ we can define the action of μ as the sum of the action of all the local pieces $\mu\psi_i, i \in I$ (and of course this definition is independent of Ψ, we leave this verification to the reader), we simply set

measCbact2
$$\mu(h) := \sum_{i \in I} \mu(\psi_i \cdot h), \quad h \in \boldsymbol{C}_b(\mathbb{R}^d). \tag{51}$$

In order to check that the action of μ on $\boldsymbol{C}_b(\mathbb{R}^d)$ is well defined we have to introduce the function

$$\psi_i^* := \sum_{j: \psi_j \cdot \psi_i \neq 0} \psi_j, \quad i \in I.$$

As a finite sum of terms (the number if neighbors is uniformly bounded) also the supports of the family $\Psi^* := (\psi_i^*)_{i \in I}$ is of uniform size.It is also clear that one has for all $i \in I$:

$$\psi_i = \psi_i \cdot 1 = \psi_i \cdot \sum_{j \in I} \psi_j = \psi_i \cdot \psi_i^*.$$

Obviously the sum well defined, since $\mu(\psi_i \cdot h) = (\mu \cdot \psi_i)(\psi_i^* \cdot h)$, and all these functions $\psi_i^* \cdot h$ belong to $\boldsymbol{C}_c(\mathbb{R}^d) \subset \boldsymbol{C}_0(\mathbb{R}^d)$. Finally thanks to formula (57) we have

MbActCb02
$$\sum_{i \in I} |(\mu\psi_i)(\psi_i^* \cdot h)| \leq \sum_{i \in I} \|\mu\psi_i\|_{\boldsymbol{M}_b} \|\psi_i^* \cdot h\|_\infty \leq \|\mu\|_{\boldsymbol{M}_b} \|h\|_\infty, \tag{52}$$

so the new functional is also bounded by $\|\mu\|_{\boldsymbol{M}_b(\mathbb{R}^d)}$. It is clear that the new functional is linear and extends the given on $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$, and hence the two norms are equal.

In the next step we have to verify that the convolution operator by C_μ also extends to $\big(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\big)$. The extension to the space $\big(\boldsymbol{C}_{ub}(\mathbb{R}^d), \|\cdot\|_\infty\big)$. is in fact easy to verify, using the argument just given and the preservation of uniform continuity of translation invariant linear operators already used in the proof of the representation theorem for TLIS.

Now we want to argue that the pointwise defined result of the convolution operator, namely $C_\mu(h)(x) = \mu(T_x h^\checkmark)$ gives a *continuous function*, even if h just belongs to $\boldsymbol{C}_b(\mathbb{R}^d)$.

Since continuity is a local property, and since we have reduced the problem to measures of the form $\mu\,\psi_i$. It is enough to show that $\psi_j \cdot C_{\mu\,\psi_i}(h)$ is a continuous function, because then

$$\left(\sum_{j \in I} \psi_j\right) \cdot C_{\mu\,\psi_i}(h) = C_{\mu\,\psi_i}(h)$$

is a continuous function. Recall that $\operatorname{supp}(\psi_j) \subset B_\delta(x_j)$, and let us assume for simplicity that we have $\delta \leq 1$. Since it is easy to verify that $\operatorname{supp}(C_{\mu\,\psi_i})(\psi_l f) \subset B_2(x_i + x_l)$ it is clear that we can ignore most of the terms $\psi_j f$, because they will provide only contributions outside of the support of ψ_j. In fact, for any fixed $i \in I$ the family

$$F_{i,j} := \{l \,|\, B_2(x_i + x_l) \cap B_1(x_j) \neq \emptyset\} = \{l \,|\, x_l \in B_3(x_j - x_i)\}, \quad i, j \in I.$$

is finite (even uniformly bounded in number), hence we can say that

$$(C_{\mu\,\psi_i}(h)) \cdot \psi_l = C_{\mu\,\psi_i} \cdot \left(\sum_{l \in F_{i,j}} h\psi_l\right) \cdot \psi_i.$$

But this auxiliary expression $h_{i,j} := \sum_{l \in F_{i,j}} h\psi_l$ is a function in $\boldsymbol{C}_c(\mathbb{R}^d)$ and hence $C_{\mu\,\psi_i}(h_{i,j})$ is in $\boldsymbol{C}_c(\mathbb{R}^d)$ as well. This completes our proof. $\square$

BernConv1 **Lemma 15.** *[Bernoulli convergence for tight sequences] Any bounded, tight sequence $(\mu_n)_{n\geq 1}$ in $(\boldsymbol{M}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}})$ with $\mu_0 = w^* - lim_n \mu_n$ is also Bernoulli convergent to μ_0.*

The proof if this claim is a good exercise and thus is left to the interested reader. It is using the fact that the action of bounded measures (originally only on $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ extends in a natural way to all of $\boldsymbol{C}_b(\mathbb{R}^d)$).

Another variation of the theme is provided by the next proposition.

wststrop1 **Proposition 1.** *Assume that $(\mu_n)_{n\geq 1}$ is a bounded and tight sequence in $(\boldsymbol{M}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}})$. Then $\mu_0 = w^* - lim_n\, \mu_n$ if and only if*

$$\mu_n * f \to \mu_0 * f, \quad \forall f \in \boldsymbol{C}_0(\mathbb{R}^d),$$

uniformly over compact sets.

Proof. Of course we make use of the details provided by Thm. (7) which will be given below.

That the convergence of operators (in the strong operator topology) implies the converge of the associated measure in the w^*- sense follows easily from equation (75), but OVERALL WE HAVE TO MOVE THIS STATEMENT TO A LATER POSITION. □

Despite the fact that the discrete measures form a closed (and obviously property) subspace of $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ we claim that they form a w^*-dense subspace of $\boldsymbol{M}(\mathbb{R}^d)$. Before verifying this we will define the adjoint of the pointwise multiplication within $\boldsymbol{C}_0(\mathbb{R}^d)$ on its dual space, i.e. how to define $h \cdot \mu$ (for $h \in \boldsymbol{C}_0(\mathbb{R}^d)$ or even $h \in \boldsymbol{C}_b(\mathbb{R}^d)$).

nctimesmeas **Definition 15.** $\mu \cdot h(f) := \mu(h \cdot f), \quad h \in \boldsymbol{C}_b(\mathbb{R}^d), f \in \boldsymbol{C}_0(\mathbb{R}^d)$.

easmodlemma **Lemma 16.** *For $h \in \boldsymbol{C}_b(\mathbb{R}^d)$ and $\mu \in \boldsymbol{M}(\mathbb{R}^d)$ one has*[24]

ncmeasnorms
$$\|\mu \cdot h\|_{\boldsymbol{M}} \leq \|h\|_\infty \|\mu\|_{\boldsymbol{M}}. \tag{53}$$

In fact, the mapping from the pointwise algebra $\big(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\big)$ into the algebra of pointwise multiplication operators $M_h : \mu \mapsto \mu \cdot h$ is in fact an isometric embedding.

Proof. The statements are more an exercise in terminology than a deep mathematical statement and are therefore left to the interested reader. □

Remark 12. Being an adjoint mapping it is also w^*-w^*-continuous, i.e. for any bounded and w^*-convergent net $(\mu_\alpha)_{\alpha\in I}$ with limit μ_0 and any fixed $h \in \boldsymbol{C}_b(\mathbb{R}^d)$ on has

$$h \cdot \mu_\alpha \to h \cdot \mu_0.$$

Since "ordinary functions" (e.g. the space $\boldsymbol{C}_c(\mathbb{R}^d)$) form a w^*-dense subspace of $\boldsymbol{M}_b(\mathbb{R}^d)$ one can show that the definition of the pointwise product is the *uniquely defined* w^*-w^*--continuous extension of the ordinary concept of *pointwise products* for continuous functions.

For this reason we will introduce a simple tool, the so-called BUPUs, the "bounded uniform partitions of unity". For simplicity we only consider the regular case, i.e. BUPUs which are obtained as translates of a single function:

[24] We have changed from $h \cdot \mu$ in the original version to $\mu \cdot h$ in Jan. 2018. This change can be ignored, since pointwise multiplication is a commutative procedure, but it can also better be interpreted as "right Banach module action". Of course we will write soon just μh.

regBUPUdef **Definition 16.** A (countable) *family of translates* $\Psi = (T_\lambda\psi)_{\lambda\in\Lambda}$, where ψ is a compactly supported function (i.e. $\psi \in \boldsymbol{C}_c(\mathbb{R}^d)$), and $\Lambda = \boldsymbol{A}(\mathbb{Z}^d)$ a lattice in $\mathbb{R}^d$ (for some non-singular $d \times d$-matrix $\boldsymbol{A}$) is called a **regular BUPU** (or more precisely a Λ-regular BUPU, or a Λ-invariant BUPU) if

LamregBUPU1 (54)
$$\sum_\lambda \psi(x-\lambda) \equiv 1.$$

We say that $\operatorname{diam}(\Psi) \leq \gamma$ if $\operatorname{supp}(\psi) \subset B_\gamma(0)$.

We will be interested in the case of "finer and finer" BUPUs, i.e.the case $\operatorname{diam}(\Psi) \to 0$. Sometimes we will simply write $|\Psi|$ instead of $diam(\Psi)$.

Regular BUPUs are sufficient for our purposes. They are a special case for a more general concept of (unrestricted) BUPUs:

BUPU-gen0 **Definition 17.** A BUPU, a so-called **b**ounded **u**niform **p**artition of **u**nity[25] in some Banach algebra $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ of continuous functions on G is a family $\Psi = (\psi_i)_{i\in I}$ of non-negative functions on G, if the following set of conditions is satisfied:

(1) There exists some *relatively compact* neighborhood U of the identity element of the group G that for each $i \in I$ there exists $x_i \in G$ such that $\operatorname{supp}(\psi_i) \subseteq x_i + U$ for all $i \in I$;
(2) The family Ψ is bounded in $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$, i.e. there exists $C_{\boldsymbol{A}} > 0$ such that $\|\psi_i\|_{\boldsymbol{A}} \leq C_{\boldsymbol{A}}$ for all $i \in I$;
(3) The family of supports $(x_i + U)_{i\in I}$ is *relatively separated*, i.e. the number of intersecting neighbors is uniformly bounded, i.e.there exists $C_0 < \infty$ such that[26]

BUPUoverlap (55)
$$\#\{j \,|\, (x_i + U) \cap (x_j + U) \neq \varnothing\} \leq C_0;$$

(4) $\sum_{i\in I} \psi_i(x) \equiv 1$.

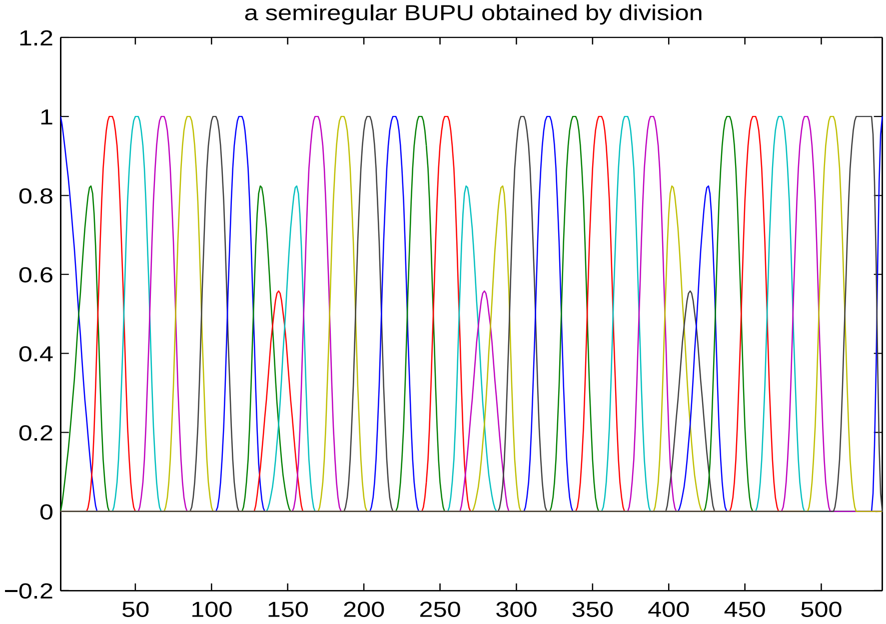
CORRECTION: WE NEED, ACCORDING TO [31] A SLIGHTLY DIFFERENT CONDITION: For any compact set (in fact, just for some compact set with non-void interior) Q one has:

overlapCORR (56)
$$\#\{j \,|\, (x_i + U) \cap x + Q \neq \varnothing\} \leq C_Q;$$

Occasionally we will refer to U as the *size of the* BUPU. Any constant C_A above is just an upper bound for the bounded set $\{\psi_i, i \in I\}$ (the best constant could be called the $\boldsymbol{A}$-norm of Ψ) and C_0 is a kind of *overlapping constant* of the family. Note that it follows from (3) that for each point the sum is finite, so even if one does not make any positivity assumption on the functions ψ_i one has no convergence problems because sums are finite at every point.

[25]Here the uniformity of the partition is referring to the *uniform size* of the partition, while the boundedness, resp. the more precisely the $\boldsymbol{A}$-boundedness refers to the boundedness of the family in the Banach algebra $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$. Of course $C_{\boldsymbol{A}}$ could be chosen to be just $\sup_{i\in I} \|\psi_i\|_{\boldsymbol{A}}$. Of course for e.g. for C^∞-BUPUs there could be various different constants, depending on the choice of $\boldsymbol{A}$, e.g. $\boldsymbol{A} = \boldsymbol{C}^{(k)}$

[26]Check that the subsequent condition is automatically satisfied for regular BUPUs.



Let $\Psi = (\psi_i)_{i\in I}$ be a BUPU, i.e. a bounded uniform partition of unity.

asabsdecomp **Theorem 4.** *Let $\Psi = (\psi_i)_{i\in I}$ be a non-negative BUPU. Then*[27]

absconvmbo (57)
$$\|\mu\|_{\boldsymbol{M}_b} = \sum_{i\in I} \|\mu\psi_i\|_{\boldsymbol{M}_b},$$

hence in particular $\mu = \sum_{i\in I} \mu\psi_i$ is absolutely convergent for $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$.

Proof. The estimate $\|\mu\|_{\boldsymbol{M}} \le \sum_{i\in I} \|\mu\psi_i\|_{\boldsymbol{M}}$ is obvious, by the triangular inequality of the norm and the completeness of $(\boldsymbol{M}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}})$.

In order to prove the opposite inequality (and in fact the finiteness of the series on the right hand side) it will be sufficient to go for the estimate $\|\mu\|_{\boldsymbol{M}} + \varepsilon \ge \sum_{i\in F} \|\mu\psi_i\|_{\boldsymbol{M}}$ for an arbitrary given $\varepsilon > 0$ and an appropriately chosen sufficiently large finite set $F \subset I$.

Recalling that $\|\mu\|_{\boldsymbol{M}_b} = \sup_{\|f\|_\infty \le 1} |\mu(f)|$ it is enough to show that for every $\varepsilon > 0$ there exists some $f \in \boldsymbol{C}_c(\mathbb{R}^d)$ with $\|f\|_\infty \le 1$ such that $|\mu(f)| \ge \sum_{i\in I} \|\mu\psi_i\|_{\boldsymbol{M}} - 3\varepsilon$.

For the given $\varepsilon > 0$ we choose first a sequence $\varepsilon_i > 0$ such that $\sum_{i\in I} \varepsilon_i < \varepsilon$. By the definition of $\|\mu\psi_i\|_{\boldsymbol{M}}$ we can find $f_i \in \boldsymbol{C}_0(\mathbb{R}^d)$ with $\|f_i\|_\infty = 1$, such that $|\mu\psi_i(f_i)| = |\mu(\psi_i f_i)| > \|\mu\psi_i\|_{\boldsymbol{M}} - \varepsilon_i$. Without loss of generality (by changing the phase of f_i if necessary) we can assume that $\mu(\psi_i f_i)$ is real-valued and in fact non-negative, i.e. absolute values can be omitted.

Hence for any finite set $F \subset I$ a function $f \in \boldsymbol{C}_c(\mathbb{R}^d)$ can be defined by $f := \sum_{i\in F} f_i\psi_i$ we have $|f(t)| \le \sum_{i\in F} \|f_i\|_\infty |\psi_i(t)| \le \sum_{i\in F} \psi_i(t) = 1$. Since μ is a linear functional we have therefore positivity of $\mu(f)$ due to

$$\mu(f) = \sum_{i\in F} \mu(\psi_i f_i) = \sum_{i\in F} \mu(\psi_i f_i) > \sum_{i\in F} (\|\mu\psi_i\|_{\boldsymbol{M}} - \varepsilon_i) \ge \sum_{i\in F} \|\mu\psi_i\|_{\boldsymbol{M}} - \varepsilon.$$

The proof is finished by observing that the estimate ensures (absolute) convergence of the sum $\sum_{i\in I} \|\mu\psi_i\|_{\boldsymbol{M}}$ and therefore one can choose F_0 such that $\sum_{i\in I\setminus F_0} \|\mu\psi_i\|_{\boldsymbol{M}} < \varepsilon$ and hence for a

[27] The reader should not worry about the order in which we write a pointwise product. Strictly speaking we might prefer to use the symbol $\psi_i \cdot \mu$ for what is written below as $\mu\psi_i$.

suitably defined $f \in \boldsymbol{C}_c(\mathbb{R}^d)$ with $\|f\|_\infty = 1$ one has

$$|\mu(f)| \geq \sum_{i \in I} \|\mu\psi_i\|_{\boldsymbol{M}} - 2\varepsilon,$$

and the proof is complete. □

lemeastight **Corollary 2.** *Every measure μ is a limit of its finite partial sums $\sum_{i\in F}\psi_i\mu$, where F runs through the family of finite subsets of I, in the following sense:*

Given $\varepsilon > 0$ there exists a finite set $F_0 \subset I$ such that for any finite set F with $F \supset F_0$ one has

approxmu (58)
$$\|\mu - \sum_{i\in F}\mu\,\psi_i\|_{\boldsymbol{M}} < \varepsilon.$$

Hence the compactly supported measures[28] *are dense in $(\boldsymbol{M}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}})$. In particular, $(\boldsymbol{M}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}})$ is an essential Banach module over $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$ with respect to pointwise multiplications. One possible approximate unit consists of the families ψ_J, where the index J is running through the finite subsets of I and is given as $\psi_J = \sum_{i\in J}\psi_i$.*

Later on we will discuss (at an abstract level) that a set shows "uniform approximability" by some approximation process (approximate unit in a Banach algebra) if and only if one specific approximating process is working well. See for details in the section on Banach modules.

One can easily verify by the same arguments as above that one obtains an equivalent norm on $(\boldsymbol{M}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}})$ if one considers the expression $\sum_{i\in I}\|h_i\cdot\mu\|_{\boldsymbol{M}}$, for some family of non-negative functions with $0 < \delta_0 \leq \sum_{i\in I} h_i(x) \leq C_1 < \infty, \forall x \in \mathbb{R}^d$.

Definition 18. For any BUPU Ψ the *Spline-type* or *Quasi-interpolation operator Sp_Φ* can be defined on the space of continuous functions on $\mathbb{R}^d$ via follows:

SpPsidef (59)
$$f \mapsto \mathrm{Sp}_\Psi(f) := \sum_{i\in I} f(x_i)\psi_i.$$

Note that the assumptions imply that pointwise, even uniformly over sets with a given diameter and independent of the position (!) the number of non-zero terms which can be added are always finite.

Lemma 17. *The operators Sp_Ψ map continuous functions into continuous functions, and are in fact a uniformly bounded family of operators on $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$ and $\big(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\big)$ respectively. Moreover $\|\mathrm{Sp}_\Psi(f) - f\|_\infty \to 0$ for $|U| \to 0$ whenever f is uniformly continuous, i.e. $f \in \boldsymbol{C}_{ub}(\mathbb{R}^d)$.*

For the proof of these statements one can/should use the following pointwise estimates:

sc-estimspl (60)
$$|(\mathrm{Sp}_\Psi g - g)(x)| \;\leq\; \mathrm{osc}_\delta(g)(x),\; \forall x \in \mathbb{R}^d.$$

where the local (δ-) *oscillation* of a function g is given by

oscdef (61)
$$\mathrm{osc}_\delta(g)(x) := sup_{|h|\leq\delta}\{|g(x) - g(x-h)|\}.$$

Of course the sup can be replaced by a max if g is a continuous function! We choose the version which measures the oscillation with respect to the center (the variable x), but of course for theoretical estimates one could look for the maximal distance of two vales of g taken within the ball of radius δ around x. This would lead to equivalent estimates.

[28] A formal discussion of the support of a measure and then of a distribution is postponed to a later section. Let us take it as a (possible) terminology for now: We will say that "a measure has compact support" (you may think of the property of "support-boundedness" being introduced before the more involved notion of a support is provided) if there exists some $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ such that μ with $\mu = \mu \cdot k$.

Proof. Note that we only have to do a pointwise estimate between $\mathrm{Sp}_\Psi g(x)$ and $g(x) = \sum_{i\in I} \psi_i(x)g(x)$, where $\mathrm{supp}(\psi_i) \subseteq B_\delta(x_i)$. So let us fix $x \in \mathbb{R}^d$ for a moment. Then

$$|\,\mathrm{Sp}_\Psi g(x) - g(x)| \leq \sum_{i\in I} |g(x_i) - g(x)| \cdot |\psi_i(x)|$$

has to be estimated, and of course the sum can be reduced to the subset $I(x) \subset I$ given by $I(x) := \{i \in I \,|\, \psi_i(x) \neq 0\}$. For $i \in I(x)$ however we have $x \in \mathrm{supp}(\psi_i)$ hence $|x - x_i| \leq \delta$, because we assume that Ψ is a δ-BUPU. Hence for each such $i \in I(x)$ we have $|g(x_i) - g(x)| \leq \mathrm{osc}_\delta(g)(x)$. This implies the desired estimate:

$$|\,\mathrm{Sp}_\Psi g(x) - g(x)| \leq \sum_{i\in I(x)} \mathrm{osc}_\delta(g)(x)\psi_i(x) \leq \mathrm{osc}_\delta(g)(x),$$

summarized in argument-free form as

estimSpminf
$$|\,\mathrm{Sp}_\Psi g - g| \leq \mathrm{osc}_\delta(g) \tag{62}$$

In order to finish the proof we only have to observe that the concept of *uniform continuity* is equivalent to the vanishing of $\mathrm{osc}_\delta(f)$ for $\delta \to 0$, in other words, for $f \in \boldsymbol{C}_b(\mathbb{R}^d)$ one has:

univcontosc
$$f \in \boldsymbol{C}_{ub}(\mathbb{R}^d) \quad \Leftrightarrow \quad \lim_{\delta\to 0} \|\mathrm{osc}_\delta\, f\|_\infty = 0. \tag{63}$$

and the obvious observation that this, combined with (62) implies the required statement. □

The family of Spline-type operators forms a *net* of operators, indexed by the neighborhood system of the identity element. In other words, we consider some $U_1 - BUPU\Psi^{(1)}$ to be superior to some $U_2 - BUPU\Psi^{(2)}$ if $U_1 \subset U_2$ (i.e. if it is associated to the smaller neighborhood of the identity element). It is also clear that in this way the index set (namely the neighborhood system of the neutral element of the group) is oriented in terms of "concentration", and consequently the observation can just be reformulated by saying:

The bounded net of operators $(\mathrm{Sp}_\Psi)_{|\Psi|\to 0}$ is *strongly convergent* to the identity operator on $\big(\boldsymbol{C}_0(\mathbb{R}^d),\, \|\cdot\|_\infty\big)$. Let us formulate this as a lemma:

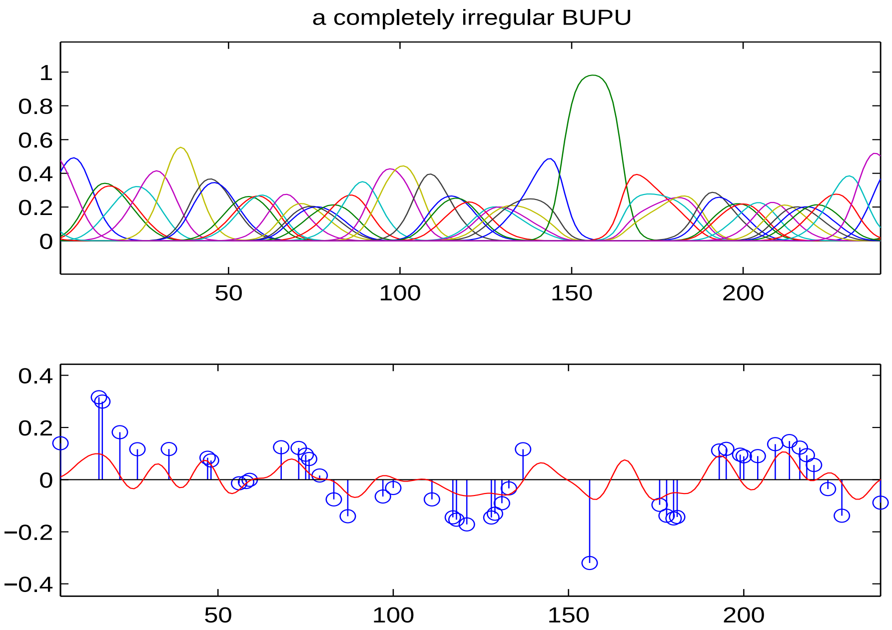
Lemma 18. *For $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ the sum defining $Sp_\Phi f$ is (unconditionally) norm convergent in $\big(\boldsymbol{C}_0(\mathbb{R}^d),\, \|\cdot\|_\infty\big)$ (even finite at each point), and $\|Sp_\Phi(f)\|_\infty \leq \|f\|_\infty$, i.e. Sp_Φ is a linear and non-expansive mapping on $\big(\boldsymbol{C}_0(\mathbb{R}^d),\, \|\cdot\|_\infty\big)$. In particular, the family of operators $(Sp_\Phi)_\Phi$, where Φ is running through the family of all (regular) BUPUs of uniform size (that means that the support size of ϕ with $\|\phi\|_\infty$ is limited) is uniformly bounded on $\big(\boldsymbol{C}_0(\mathbb{R}^d),\, \|\cdot\|_\infty\big)$.* [29] *Moreover, these spline-type quasi-interpolants are norm convergent to f as $|\Phi|$ (the maximal diameter of members of Φ) tends to zero. In other words, we claim: For every $\varepsilon > 0$ and any finite subset $F \subset \boldsymbol{C}_0(\mathbb{R}^d)$ there exists $\delta > 0$ such that $\|Sp_\Phi f - f\|_\infty < \varepsilon$ if only $|\Phi| \leq \delta_0$.*

Remark 13. In some cases it might be of interest to look at BUPUs which preserve uniform continuity, i.e. which have the property that $\mathrm{Sp}_\Psi(f) \in \boldsymbol{C}_{ub}(G)$ for any $f \in \boldsymbol{C}_b(G)$. This is certainly the case if one has a regular BUPU but not for general BUPUs.

Next we study the *adjoint operator* associated to Sp_Ψ, which we will call *discretization* operator and denote by D_Ψ, mapping bounded measures into discrete measures. Thus is justified by the following consideration: The concrete form of the dual resp. adjoint operator is given by:

SpPsi-dual0
$$[\mathrm{Sp}_\Psi'(\mu)](f) = \mu(\mathrm{Sp}_\Psi(f)) = \mu\left(\sum_{i\in I} f(x_i)\psi_i\right) = \sum_{i\in I} \mu(\psi_i) f(x_i) = \left[\sum_{i\in I} \mu(\psi_i)\delta_{x_i}\right](f). \tag{64}$$

[29] these statements have to be simplified



This motivates the following definition

DPsi-def1 **Definition 19.** Given a BUPU Ψ we define the corresponding *discretization* operator by

DPsidef (65)
$$D_\Psi \mu \quad (\text{or better}) \quad D_\Psi \mu := \sum_{i \in I} \mu(\psi_i) \delta_{x_i}.$$

For every BUPU Ψ we denote the adjoint mapping to Sp_Ψ by D_Ψ, the *discretization operator generated by the partition of unity* Ψ. It maps $\boldsymbol{M}(\mathbb{R}^d)$ into itself (in a linear way), and since Sp_Ψ is non-expansive the same is true for D_Ψ We will allow to apply operators of this kind also to families of operators, i.e. we will shortly write $D_\rho \Phi$ for the family $(D_\rho(T_\lambda \varphi))_{\lambda \in \Lambda}$. Since D_ρ preserves values of functions (it moves the values via stretching or compression to "other places"), hence $D_\rho \mathbf{1}_{\mathbb{R}^d} = \mathbf{1}_{\mathbb{R}^d}$[30]. Consequently Φ is a (regular) BUPU if and only if $D_\rho \Phi$ is a BUPU with respect to the lattice $1/\rho \cdot \Lambda$ for any $\rho > 0$.

In order to understand the engineering terminology of an "impulse response" uniquely describing the behavior of a linear time-*invariant* system let us take a quick look at the situation over the group $G = \mathbb{Z}_n$, the cyclic group (of complex unit roots of order n). Obviously $c_0(\mathbb{Z}_n)$ is just $\mathbb{C}^n$, and the (group-) translation is just cyclic index shift (mod n).

Cn-TILS **Lemma 19.** *A matrix A represents a "translation-invariant" linear mapping on $\mathbb{C}^n$ if and only if it is circulant, i.e. if it is constant along side-diagonals (we will also call such matrices "convolution matrices");*

Proof. We can start with the "first unit vector", which is mapped onto some column vector in $\mathbb{C}^n$ by the linear mapping $x \mapsto A * x$. Since we can interpret all further vectors as the image of the other unit vectors, but these are obtained from the first unit vector by cyclic shift, we see immediately that the columns of the matrix are obtained by cyclic shift of the first column, hence the matrix A has to be circulant (in the cyclic sense).

[30] We use the symbol $\mathbf{1}_M$ to indicate the *indicator function* of a set, which equals 1 on M and zeros elsewhere.

Usually engineers call the first column of this matrix, which is the output corresponding to an "impulse like" input (the first unit vector) the *impulse response of the translation invariant linear system* A.

Since the converse is easily verified we leave it to the interested reader. In fact, it may be interesting to verify the translation invariance by deriving first for a general matrix action A the action of $T_{-1} \circ A \circ T_1$ and to observe subsequently that this operation does not change circulant matrices. □

For the "continuous domain" (i.e.for linear systems over $\mathbb{R}$ resp. over $\mathbb{R}^d$) one has to invoke functionals in order to be able to "represent" the translation invariant systems. We give the details below.

Nevertheless we would like to recall the usual argument (and later how it should be modified). Standard textbooks explain the behavior of a TILS (a translation invariant linear system[31]) by the following illustration. On decomposes, resp. approximates the given input signal (here a smooth, well localized complex-valued function, with real part plotted in blue, imaginary part in green) by a step function, or in other words, a superposition of shifted rectangular functions. Since the linear system $T : f \mapsto T(f)$ (input-output relation) is translation invariant, the output of such a box function is some output signal, which one has to store, say $g = T(box)$.

Then the output is a linear combination of shifted copies of g (same shifts with the same coefficients/amplitudes), see Fig.1 below.

The argument then continues with the "observation" that one obtains, by making the rectangles shorter and shorter (suitably normalized, such that they all have area one) an **impulse** and the corresponding limit of the output results $T(g)$ "tend to some **impulse response** " function (or object, in whatever sense), so that one can explain the behavior of a linear system by translating the impulse response (with amplitudes coming from the input signal f). In quasi-mathematical writing one could say: We start from the so-called *sifting property* (see the literature of signal processing)

sifting (66)
$$f = \int_{\mathbb{R}^d} f(x)\delta_x dx = \int_{\mathbb{R}^d} f(x)T_x(\delta_0)dx$$

to the realization of a TILS T by

Tilsift1 (67)
$$Tf = \int_{\mathbb{R}^d} f(x)T_x[T(\delta_0)]dx.$$

Writing $h := T(\delta_0)$ for the impulse response function (assuming it is a function) we get

TILSsift2 (68)
$$Tf(z) = \int_{\mathbb{R}^d} f(x)\, T_x h(z)dx = \int_{\mathbb{R}^d} h(z-x)f(x)dx,$$

which is exactly the well-known formula for convolution of functions, which at first sight appears to justify the reasoning.

We go on (following the engineers arguments) to say, that it is plausible that this calculus is valid, because the step-functions plotted in the above figure "tend to the limiting" function, and therefore (by the continuity of the system) one can expect that also the output is convergent to $T(f)$, the output to the true (smooth, well decaying) function f used as input.

[31]Systems are mathematically speaking just linear operators, or input-output relations which satisfy the so-called *superposition principle* $T(f+g) = T(f) + T(g)$, which practically equivalent with linearity of T. Often the domain of T remains vague, i.e. it may change, or the information about output depends on the information of the input. Engineers often use box-diagrams, viewing T as a "black box" with an input arrow and a corresponding output arrow

stepapprox1

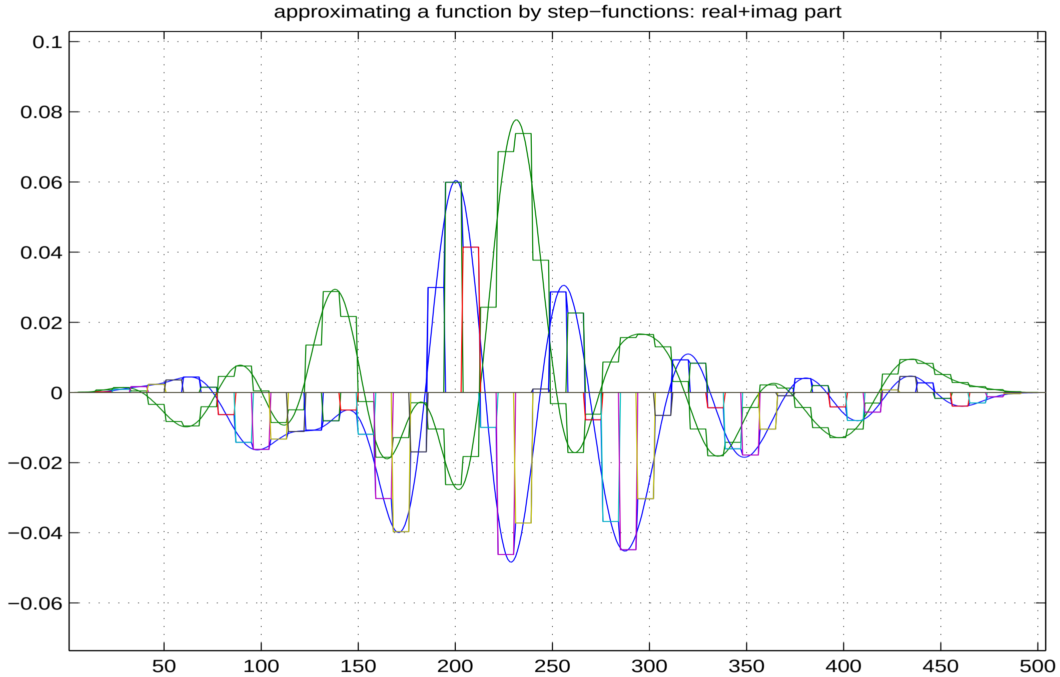


FIGURE 6. A typical illustration of an approximation to the input signal of a translation invariant linear system, preparing for the use of the Dirac impulse as limit of the rectangular functions. But in which sense are these step function convergent to the input signal f, and how are the steps determined?

This is a nice *plausibility explanation*, but it **leaves more things open than providing explanations**. In which sense do we have convergence (uniform convergence, e.g. if the input function is uniformly continuous). But what are the assumption on the TILS. Is it mapping bounded continuous functions into bounded continuous functions? Is it what engineers would call a BIBOS system (bounded input to bounded output system)? And how can one understand that the limit of the output signals to smaller and smaller box-functions (centered at zero, $\boldsymbol{L}^1$-norm $= 1$) exists, and in which sense? But if the domain are the bounded and (uniformly?) continuous functions, how can we be sure that the step-functions (approximating them) are also in the domain of the operator (after all they are *discontinuous*!).

Just to illustrate the situation by a plot, which is similar to those found in most books on engineering applications let us look at a step-function approximation of some decent (bounded and decaying) continuous function f. Those books would argue: Feeding the step-function approximation into the system (instead of the continuous function itself) will (but based on which reasoning?) give a very similar output, so that it is enough to know the output for the compressed box function in order to predict the output $T(f)$. The typical picture in standard text books looks like Fig. 4 below (cf. [70], Sec.1.4 introducing the "Faltungsintegral"). I.Sandberg has argued (receiving a lot of sceptical voices within the engineering community) that there is a scandal in the literature ([89, 90], as well as many more articles going in that direction), but how can we get out of this trap? A partial answer is given in [38].

We will use a slightly different model, by assuming that TILS are defined on the Banach space $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$ and bounded there. Since $\boldsymbol{C}_c(\mathbb{R}^d)$ is a dense subspace this really means that we only assume that TILS are defined for compactly supported and continuous functions,

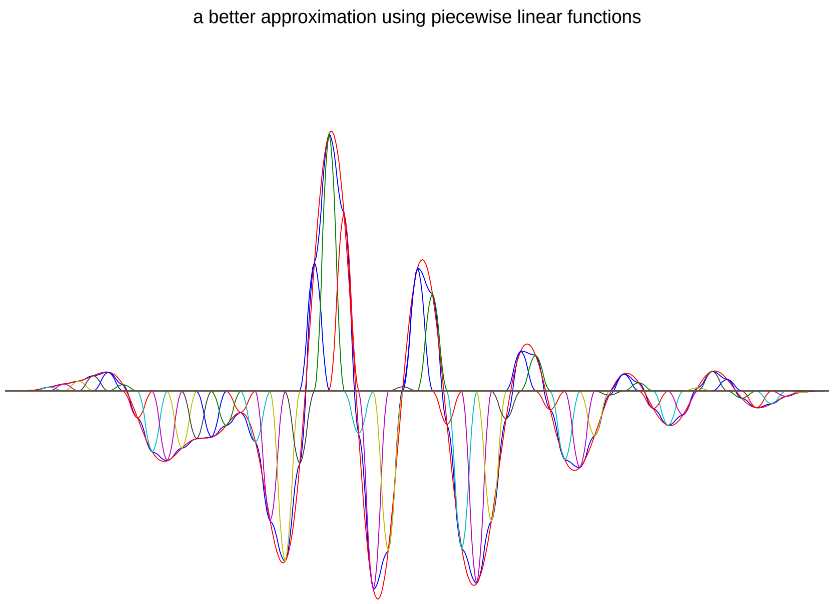


FIGURE 7. The piecewise linear functions are special cases of our approximations of the form $\mathrm{Sp}_\Psi(f)$ and thus converge to f in the uniform norm. Hence by the BIBOs assumption we can also ensure that the output $T(\mathrm{Sp}_\Psi(f))$ will also converge to $T(f)$ uniformly. Note that $\mathrm{Sp}_\Psi(f) = \sum_{i\in I} f(\lambda) T_\lambda \psi_0$ and thus $T(\mathrm{Sp}_\Psi(f)) = \sum_{i\in I} f(\lambda) T_\lambda(T\psi)$. The argument is then, that a (rescaled!!) version of ψ_0 (for $|\Psi| \to 0$) tends to δ_0.

but with the extra property of boundedness, i.e.that there exists $C = C(T) > 0$ such that

BIBOS-est1

$$\|T(f)\|_\infty \le C\|f\|_\infty, \quad \forall f \in \boldsymbol{C}_0(\mathbb{R}^d). \tag{69}$$

We will see that in this way the above reasoning can be made sharp. Moreover, in order to approximate within the space $\boldsymbol{C}_0(\mathbb{R}^d)$ it is better to use for the approximation piecewise linear functions, which in fact are superpositions of shifted triangular functions (which in turn form the basis for spline-functions of order 2 resp. degree 1). We will see that under these circumstances the convolution is with respect to some measure, which in turn is the limit of elementary (discrete) measures.

The common part for both viewpoints is the observation, that for a TILS T the image of a function of the form $\sum_{i\in I} c_i T_{x_i} h$ under T obviously will be of the form

$$T\left(\sum_{i\in I} c_i T_{x_i} h\right) = \sum_{i\in I} c_i T_{x_i}[T(h)],$$

where $T(h)$ is the approximate impulse response of T.

We may explain, in which way the action of $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ extends in a very natural way from the action only on $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$ (original domain) to all of $\big(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\big)$ (instead of considering all possible bounded linear functionals on $\big(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\big)$, including those which vanish on $\boldsymbol{C}_0(\mathbb{R}^d)$ without being the zero functional on all of $\big(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\big)$!).

Given $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ and some BUPU Ψ, let us recall that we have (according to Thm. 4) following equation (**??**) BELOW! we have:

absconvmbo1 (70)
$$\|\mu\|_{\boldsymbol{M}} = \sum_{i\in I} \|\mu\psi_i\|_{\boldsymbol{M}},$$

and hence we can put for $h \in \boldsymbol{C}_b(\mathbb{R}^d)$:

$$\mu(h) := \sum_{i\in I} (\psi_i\mu)(h) = \sum_{i\in I} \mu(\psi_i \cdot h),$$

which is well defined and can be estimated via

$$|\mu(h)| \leq \sum_{i\in I} |\mu(\psi_i \cdot h) \leq \sum_{i\in I} \|\mu\|_{\boldsymbol{M}_b(\mathbb{R}^d)} \sup_{i\in I} \|\psi_i\|_\infty \|h\|_\infty \leq \|\mu\| \|h\|_\infty.$$

Of course it is not difficult to verify that this definition does *not depend* on the particular choice of the BUPU and that, for a uniformly bounded net (f_α) of functions in $\boldsymbol{C}_b(\mathbb{R}^d)$ it suffices to have uniform convergence over compact sets in order to guarantee that $\mu(f_\alpha) \to \mu(f_0)$. This argument will help us to show that $\hat{\mu}$ is a continuous function on the dual group!

ompact-CORd **Theorem 5.** *A bounded and closed subset $M \subset \boldsymbol{C}_0(\mathbb{R}^d)$ is compact if and only if it is uniformly tight and uniformly equicontinuous* [32]*, i.e. if the following conditions are satisfied:*

- *for $\varepsilon > 0$ there exists $\delta > 0$ such that $|y| \leq \delta \Rightarrow \|T_y f - f\|_\infty \leq \varepsilon \quad \forall f \in M$;*
- *for $\varepsilon > 0$ there exists some $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ such that $\|f - kf\|_\infty \leq \varepsilon, \quad \forall f \in M$;*

Proof. It is obvious that finite subsets $M \subset \boldsymbol{C}_0(\mathbb{R}^d)$ have these two properties, and it is easy to derive these two properties for compact sets by the usual approximation argument.

So we have to show the converse. We observe first that $\|pf - f\|_\infty \to 0$, if p is a sufficiently large plateau-function. The set $\{pf \mid f \in M\}$ is still equicontinuous (cf. the proof, that $\boldsymbol{C}_{ub}(\mathbb{R}^d)$ is a Banach algebra with respect to pointwise multiplication). We may assume that p has compact support. Then we apply a (sufficiently) fine BUPU to ensure that $\|pf - Sp_\Psi(pf)\|_\infty \leq \varepsilon$. Since p has compact support only finitely many terms make up $Sp_\Psi(pf)$, i.e. one can approximate by finite linear combinations of the elements of Ψ. In other words, up to some $\varepsilon > 0$ one can approximate the elements of M by the elements of a bounded set in a finite-dimensional vector space, which is of course (by the usual Heine-Borel argument) a relatively compact set.

Since a set which can be approximated arbitrarily well by a relatively compact set is relatively compact itself (think of the corresponding ε-coverings with the finite covering property!) the claim follows[33]. □

In the literature (see e.g. [18], p.175) the so-called Arzela-Ascoli Theorem is the prototype characterizing relative compactness in $(\boldsymbol{C}(X), \|\cdot\|_\infty)$ over compact domains.

ArzAscoThm **Theorem 6.** *[Arzela-Ascoli Theorem] Let X be a compact (and hence completely regular) Hausdorff topological space, and M a bounded and closed subset of $(\boldsymbol{C}(X), \|\cdot\|_\infty)$. Then M is compact if and only if it is equicontinuous, which means: For every $\varepsilon > 0$ and $x_0 \in X$ there exists some neighborhood U_0 of x_0 such that*

$$|f(z) - f(x_0)| \leq \varepsilon, \quad \forall f \in M.$$

---

[32] The concept of pointwise equicontinuity of functions could be formulated, but is rarely relevant, hence in many cases on just speaks of *equicontinuity* of M.

[33] In this part of the argument we make use of the fact that in a complete metric space relative compactness (resp. the assumption that a set has compact closure) and precompactness (for every $\varepsilon > 0$ there exists a finite covering with balls of radius ε) are equivalent properties.

Just in order to demonstrate that the technical details are in fact not very deep here let us formulate the statement:

precomp **Lemma 20.** *Given $\varepsilon > 0$ and a set H in some metric space, we denote by H_ε the ε-hull of H, defined as $H_\varepsilon := \bigcup_{h\in H} B_\varepsilon(h)$.*

If there is a sequence of finite sets H^n and a sequence ε_n such that $H \subset H^n_{\varepsilon_n}$ for all $n \geq 1$, with $\varepsilon_n \to 0$ for $n \to \infty$, then H is a precompact set.

Proof. Given $\varepsilon > 0$ we have to show that H has a finite covering of balls of radius ε. For this purpose let us choose n_0 such that $\varepsilon_{n_0} < \varepsilon$. Using now that $H \subset H^{n_0}_{\varepsilon_{n_0}}$ we find that H covered by a finite collection of balls of radius ε[34]. □

ef-BIBOTLIS **Definition 20.** The Banach space of all "translation invariant linear systems" (TLIS) on $\boldsymbol{C}_0(\mathbb{R}^d)$ is denoted by[35]

HGCORdDef
$$\mathcal{H}_{\mathbb{R}^d}(\boldsymbol{C}_0(\mathbb{R}^d)) = \{T : \boldsymbol{C}_0(\mathbb{R}^d) \to \boldsymbol{C}_0(\mathbb{R}^d), \text{ bounded, linear} : T \circ T_x = T_x \circ T, \forall x \in \mathbb{R}^d\} \tag{71}$$

Remark 14. [36] It is easy to show that $\mathcal{H}_{\mathbb{R}^d}(\boldsymbol{C}_0(\mathbb{R}^d))$ is a closed subalgebra of the Banach algebra of $\mathcal{L}(\boldsymbol{C}_0(\mathbb{R}^d))$ (in fact it is closed with respect to the strong operator topology), hence it is a Banach algebra of its own right (with respect to composition as multiplication). We will see later that it is in fact a commutative Banach algebra.

Let us formulate this a separate lemma.

GCORdclosed **Lemma 21.**
(1) *The space $\mathcal{H}_{\mathbb{R}^d}(\boldsymbol{C}_0(\mathbb{R}^d))$ is in fact a closed subalgebra of $\mathcal{L}(\boldsymbol{C}_0(\mathbb{R}^d))$ (with operatornorm), hence endowed with the operator norm it is*
(2) *$\mathcal{H}_{\mathbb{R}^d}(\boldsymbol{C}_0(\mathbb{R}^d))$ is closed with respect to the strong operator topology, i.e.if you have a sequence of operator $(T_n)_{n\geq 1}$ in $\mathcal{L}(\boldsymbol{C}_0(\mathbb{R}^d))$ with the property that*
$$\|T_n f - T_0 f\|_\infty = 0, \quad \forall f \in \boldsymbol{C}_0(\mathbb{R}^d),$$
then the limiting operator also belongs to $\mathcal{H}_{\mathbb{R}^d}(\boldsymbol{C}_0(\mathbb{R}^d))$.
(3) *Clearly $\mathcal{H}_{\mathbb{R}^d}(\boldsymbol{C}_0(\mathbb{R}^d))$ contains all the translation operators $T_x, x \in \mathbb{R}^d$, and their closed linear span forms a commutative subalgebra of $(\mathcal{H}_{\mathbb{R}^d}(\boldsymbol{C}_0(\mathbb{R}^d)), |||\cdot|||)$.*

It will be one of the goals of these notes to show that the strong closure of the translation operators is all of $\mathcal{H}_{\mathbb{R}^d}(\boldsymbol{C}_0(\mathbb{R}^d))$ and that consequently the algebra is commutative.

It will be done using the "representation" of such a *translation invariant system* using its *impulse response*. We think that it is more natural to view these operators as *moving averages*, but we will explain this (hopefully) later in more detail. In fact, the moving average makes use of the flipped version of the impulse response, and this is why in the usual (well justified) conventions concerning convolution we have so many flip operators!

What will be important for later is the following fact: Assume we know that there is some abstract (hopefully isometric) isomorphism between the (then) *commutative Banach algebra* $(\mathcal{H}_{\mathbb{R}^d}(\boldsymbol{C}_0(\mathbb{R}^d)), |||\cdot|||)$ and WHATEVER Banach space, then of course this Banach space can be turned into a Banach algebra by **transfer of structure**. We will do this with the concrete Banach space of bounded measures, $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$, which for the sake of convenience is described as simply the dual of $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$.

---

[34]It would in fact be sufficient that H_n is a precompact set for every $n \geq 1$.

[35]The letter $\mathcal{H}$ in the definition refers to *homomorphism* [between normed spaces], while the subscript G in the symbol refers to "commuting with the action of the underlying group $G = \mathbb{R}^d$ realized by the so-called regular representation, i.e. via ordinary translations.

[36]Sometimes we will write $[T, T_z] \equiv 0$ in order to express the commutation formula using the commutator symbol, and the "$\equiv$"–symbol to express that this relation holds true $\forall z \in \mathbb{R}^d$.

This is in fact in complete analogy to the situation for linear mappings from $\mathbb{R}^n$ to $\mathbb{R}^n$. They also clearly form a Banach algebra with respect to composition. Using for simplicity the standard basis $(\mathbf{e}_k)_{k=1}^n$ for $\mathbb{R}^n$ we know that we can identify this algebra with $\mathbb{C}^{n^2}$, usually written in the form of a matrix $\boldsymbol{A}$, with $\vec{\mathbf{a}_k} = T(\mathbf{e}_k)$, the corresponding "multiplication" on these $n \times n$ matrices is of course just matrix multiplication as we have learned it in our linear algebra courses!

Recall the notion of a FLIP operator:

$$\check{f}(z) = f^{\checkmark}(z) = f(-z).$$

mustarf1 **Definition 21.** Given $\mu \in \boldsymbol{M}(\mathbb{R}^d)$ we define the *convolution operator* C_μ by:

Cmudef1
$$C_\mu(f)(z) := \mu(T_z f^{\checkmark}). \tag{72}$$

The *reverse mapping* R recovers the corresponding measure $\mu = \mu_T$ from the system T via

muTdef1
$$\mu(f) := T(f^{\checkmark})(0). \tag{73}$$

zation-TLIS **Theorem 7.** *[Characterization of LTISs on $\boldsymbol{C}_0(\mathbb{R}^d)$]*
There is a natural isometric isomorphism between the Banach space $\mathcal{H}_{\mathbb{R}^d}(\boldsymbol{C}_0(\mathbb{R}^d))$, endowed with the operator norm, and $(\boldsymbol{M}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}})$, the dual of $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$, by means of the following pair of mappings:

(1) *Given a bounded measure $\mu \in \boldsymbol{M}(\mathbb{R}^d)$ we define the operator C_μ (to be called convolution operator with convolution kernel μ later on) via:*

onv-measdef
$$C_\mu f(x) = \mu(T_x f^{\checkmark}). \tag{74}$$

(2) *Conversely we define $T \in \mathcal{H}_{\mathbb{R}^d}(\boldsymbol{C}_0(\mathbb{R}^d))$ the linear functional $\mu = \mu_T$ by*

impuls-def
$$\mu_T(f) = [Tf^{\checkmark}](0). \tag{75}$$

The claim is that both of these mappings: $C: \mu \mapsto C_\mu$ and the mapping $T \mapsto \mu_T$ are linear, non-expansive, andcd inverse to each other. Consequently they establish an isometric isomorphism between the two Banach spaces with

LTIS-isomet
$$\|\mu_T\|_{\boldsymbol{M}} = \|T\|_{\mathcal{L}(\boldsymbol{C}_0(\mathbb{R}^d))} \quad \textit{and} \quad \|C_\mu\|_{\mathcal{L}(\boldsymbol{C}_0(\mathbb{R}^d))} = \|\mu\|_{\boldsymbol{M}}. \tag{76}$$

THIS IS A RATHER IMPORTANT THEOREM, ALLOWING TO DEFINE CONVOLUTION, AND ALSO TO DERIVE THE CONVOLUTION THEOREM, NOT JUST IN AN $\boldsymbol{L}^1$-CONTEXT, BUT FOR GENERAL MEASURES!

Proof. The proof has a number of partial steps carried out in the course. Some of them are quite elementary (e.g. that the resulting objects are really linear systems resp. linear functionals, or that the mappings C and its inverse are linear mappings) and are left to the interested reader. We concentrate on the most interesting parts.

First of all it is easy to check that the definition of C_μ really defines an operator on $\boldsymbol{C}_0(\mathbb{R}^d)$, with the output being a bounded function.[37] Since

conv-bd1
$$|C_\mu(f)(x)| \leq \|\mu\|_{\boldsymbol{M}} \|T_x f^{\checkmark}\|_\infty = \|\mu\|_{\boldsymbol{M}} \|f\|_\infty, \tag{77}$$

and hence the operator norm of C_μ[38] satisfies $\|C_\mu\|_{\mathcal{L}(\boldsymbol{C}_0(\mathbb{R}^d))} \leq \|\mu\|_{\boldsymbol{M}(\mathbb{R}^d)}$.

[37] Engineers talk of BIBOS, i.e. Bounded Input [gives] Bounded Output Systems.

[38] Of course we think of the "convolution by the measure μ, in conventional terms, and this is also the reason for using the symbol C.

Furthermore it commutes with translations, since $D_{-1}(T_z f) = T_{-z} D_{-1} f = T_{-z}\check{f}$

onv-commut1
$$C_\mu(T_z f)(x) = \mu(T_x T_{-z}\check{f}) = \mu(T_{x-z}\check{f}) = C_\mu(f)(x-z) = T_z C_\mu f(x). \tag{78}$$

It is also easy to check - using this fact - that (uniform) continuity is preserved, since

conv-cont
$$\|T_z(C_\mu(f)) - C_\mu(f)\|_\infty = \|C\mu(T_z f - f)\| \leq \|\mu\|_{\boldsymbol{M}} \|T_z f - f\|_\infty \to 0 \quad \text{for} \quad |z| \to 0. \tag{79}$$

Finally we have to verify that the output of $f \to C_\mu(f)$ is not just continuous but also decaying at infinity. Due to the closedness of $\boldsymbol{C}_0(\mathbb{R}^d)$ within $\big(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\big)$ it is sufficient to approximate $C_\mu(f)$ for $f \in \boldsymbol{C}_c(\mathbb{R}^d)$ by compactly supported (continuous) functions. This is achieved using Theorem 4 resp. Corollary 3. Given $\varepsilon > 0$ one can choose $h \in \boldsymbol{C}_c(\mathbb{R}^d)$ with $\|\mu - h\mu\|_{\boldsymbol{M}} < \varepsilon$. Then the difference

compapprox
$$\|C_{h\mu} f - C_\mu f\|_\infty \leq \|h \cdot \mu - \mu\|_{\boldsymbol{M}} \|f\|_\infty \tag{80}$$

can be made arbitrarily small, while on the other hand

ompsuppconv
$$C_{h\mu} f(x) = h\mu(T_x f^{\checkmark}) = \mu(h \cdot T_x f^{\checkmark}) = 0 \tag{81}$$

if $\mathrm{supp}(h) \cap \mathrm{supp}(T_x f^{\checkmark}) = \mathrm{supp}(h) + (x - \mathrm{supp}(f)) \neq \emptyset$, i.e. if $x \notin \mathrm{supp}(h) + \mathrm{supp}(f)$, since we have assumed that f has compact support. Thus altogether we have found that $f \mapsto C_\mu(f)$ is in $\mathcal{H}_{\mathbb{R}^d}(\boldsymbol{C}_0(\mathbb{R}^d))$, with a control of the operator norm on $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$ by $\|\mu\|_{\boldsymbol{M}}$.

Let us verify next that the two mappings are inverse to each other. Let us verify first that the system associated with μ_T is the original system, in other words we have to show that $C_{\mu_T} = T$ for all $T \in \mathcal{H}_{\mathbb{R}^d}(\boldsymbol{C}_0(\mathbb{R}^d))$. Using the relevant definitions this follows from the following chain of equalities

sysmeassys1
$$[C_{\mu_T}(f)](x) = \mu_T(T_x f^{\checkmark}) = T((T_x f^{\checkmark})^{\checkmark})(0) = T(T_{-x} f)(0) \tag{82}$$

Because T commutes with all the translations we continue by

sysmeassys
$$[C_{\mu_T}(f)](x) = T(T_{-x} f)(0) = T_{-x} T f(0) = T f(0 - (-x)) = T f(x). \tag{83}$$

Since this is valid for all $x \in \mathbb{R}^d$ and $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ one direction is shown. It also shows that *every* $T \in \mathcal{H}_{\mathbb{R}^d}(\boldsymbol{C}_0(\mathbb{R}^d))$ is a convolution operator by a uniquely determined measure $\mu_T \in \boldsymbol{M}_b(\mathbb{R}^d)$, and that $\|\mu_T\|_{\boldsymbol{M}} = \|T\|_{\mathcal{L}(\boldsymbol{C}_0(\mathbb{R}^d))}$. In fact, since we know already that $\|\mu_T\|_{\boldsymbol{M}} \leq \|T\|_{\mathcal{L}(\boldsymbol{C}_0(\mathbb{R}^d))}$ we only have to verify that a strict inequality would imply a contradiction. Thus using (83) we come to the following exact isometry:

$$\|T\|_{\mathcal{L}(\boldsymbol{C}_0(\mathbb{R}^d))} = \|C_{\mu_T}\|_{\mathcal{L}(\boldsymbol{C}_0(\mathbb{R}^d))} \leq \|\mu_T\| \leq \|T\|_{\mathcal{L}(\boldsymbol{C}_0(\mathbb{R}^d))}.$$

Let us now look for the converse identity: the measure associated to the convolution operator C_μ associated with some given $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ is just the original measure, since

measysmeas1
$$\mu_{C_\mu}(f) = C_\mu(f^{\checkmark})(0) == \mu(T_0(f^{\checkmark})^{\checkmark}) = \mu(f). \tag{84}$$

which in turn implies that the inverse mapping is surjective, i.e. that the (already known to be isometric) bijection between bounded measures and systems is in fact surjective onto the bounded measures. Thus our proof is complete. □

Note also that there is a matter of consistency to be verified. For "ordinary functions" $g \in \boldsymbol{C}_c(\mathbb{R}^d)$ [39] one can define a unique bounded measure

[39] Resp. for $g \in \boldsymbol{L}^1(\mathbb{R}^d)$ if one is familiar with Lebesgue integration, but we prefer to avoid this part of analysis, mostly in order to demonstrate that it is not of the importance usual given to it in most books.

regbdmeas1 **Definition 22.**

regbdmeas (85)
$$\mu_g(f) = \int_{\mathbb{R}^d} f(x)g(x)dx,$$

Within measure theory one would say that $d\mu(x) = g(x)dx$ describes a measure which is *absolutely continuous* with respect to the Lebesgue measure.[40]

As was explained already this embedding is isometric into $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$, if one endows the function space with the usual $\|\cdot\|_1$-norm (taken in the Riemannian or Lebesgue sense).

4.1. **Many ways to introduce convolutions!** The previous characterization allows to introduce in a natural way a Banach algebra structure on $\boldsymbol{M}(\mathbb{R}^d)$. In fact, given μ_1 and μ_2 the translation invariant system $C_{\mu_1} \circ C_{\mu_2}$ is represented by a bounded measure μ. In other words, we can define a new (so-called) *convolution product* $\mu = \mu_1 * \mu_2$ of the two bounded measures such that the relation (completely characterizing the measure $\mu_1 * \mu_2$)

CONVO1 (86)
$$C_{\mu_1 * \mu_2} = C_{\mu_1} \circ C_{\mu_2}$$

TURN THE NEXT STATEMENT INTO A FORMAL DEFINITION

It is immediately clear from this definition that $(\boldsymbol{M}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}})$ is a **Banach algebra with respect to convolution**! Associativity is given for free, but *commutativity of the new convolution is not so obvious* (and will follow only later, clearly as a consequence of the commutativity of the underlying group).

The translation operators themselves, i.e. T_z are elements of $\mathcal{H}_{\mathbb{R}^d}(\boldsymbol{C}_0(\mathbb{R}^d))$, which correspond exactly to the Dirac measures δ_z, $z \in \mathbb{R}^d$, due to the following simple consideration[41].

conv-deltax (87)
$$C_{\delta_x} f(z) = \delta_x(T_z f^{\checkmark}) = [T_z f^{\checkmark}](x) = f^{\checkmark}(x - z) = f(z - x) = [T_x f](z),$$

and thus $C_{\delta_x} = T_x$, in particular $C_{\delta_0} = T_0 = Id$, resp.

convDirac0 (88)
$$f = \delta_0 * f \quad \text{and} \quad T_x f = \delta_x * f, \quad \forall f \in \boldsymbol{C}_0(\mathbb{R}^d).$$

HERE OR AT ANOTHER PLACE ON SHOULD EXPLAIN THE CONNECTION BETWEEN THIS RELATION AND THE SO-CALLED SIFTING PROPERTY OF THE DIRAC 'DELTA FUNCTION' $\delta(t)$, AS IT IS USED IN ENGINEERING BOOKS QUITE A LOT!

Remark 15. In the literature there are other, quite non-obvious definitions of the convolution of two bounded measures. As a consequences of this "usual" approach it is required to provide a (non-trivial) proof for the associativity, which is *not necessary* in the case of our definition. In fact, if convolution is defined in an individual manner in the context of generalized functions this associativity rule may even *fail to be valid* (cf. [4], p. 111/112). A similar point arises in H. Shapiro's description of (abstract) *homogeneous Banach spaces* where he also has to assume that the action of (generalized convolution, i.e. the integrated group action) is associative, and he gives a proof for the case of locally integrable functions ([91], Chap.9).

The classical approach describes the convolution of two measures (for now the order matters) as follows: Given $\mu_1, \mu_2 \in \boldsymbol{M}_b(\mathbb{R}^d)$ the action of $\mu_1 * \mu_2$ on $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ can be characterized by

conv-meas1 (89)
$$\mu_1 * \mu_2(f) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(y + x) d\mu_2(y) d\mu_1(x).$$

[40] Cf. the Theorem of Radon-Nikodym!

[41] We start writing $\mu * f$ for $C_\mu(f)$, the convolution of μ with f. This is justified by the fact that in the case of double interpretation, e.g. for $\mu = \mu_g$ for some $g \in \boldsymbol{C}_c(\mathbb{R}^d)$ either as a pointwise defined integral or the action of a measure on test functions one can *easily* check that there is consistency among the different view-points. From now on we will not discuss this issue in all details and leave the verification of details to the reader.

Proof. First recall the definition, i.e. that $\mu_1 * \mu_2$ is the linear functional corresponding to $T := C_{\mu_1} \circ C_{\mu_2} \in \mathcal{H}_{\mathbb{R}^d}(\boldsymbol{C}_0(\mathbb{R}^d))$ satisfies

$$\mu_1 * \mu_2(f) = T(f^{\checkmark})(0) = [(C_{\mu_1} \circ C_{\mu_2})(f^{\checkmark})](0) = [C_{\mu_1}(C_{\mu_2}(f^{\checkmark}))](0) = \mu_1(g),$$

with $g(x) = [C_{\mu_2}(f^{\checkmark})]^{\checkmark}(x) = C_{\mu_2}(f^{\checkmark})(-x) = \mu_2(T_x f)$. Putting everything into the standard notation we have $g(x) = \int_{\mathbb{R}^d} f(y+x) d\mu_2(y)$ which implies the result stated above. □

Remark 16. The different realizations of convolution can conveniently be disguised by using (mostly from now on throughout the rest of the manuscript) the symbol $*$ for convolution, i.e. we write $\mu * f$ instead of $C_\mu(f)$ from now on, or $g*f$ for the convolution of $\boldsymbol{L}^1(\mathbb{R}^d)$-functions using the Lebesgue integral. We have discussed that the reinterpretation of those things in varying contexts is - from a pedantic point of view - a delicate matter, but having discussed this problem in detail for a few cases we leave the rest to the reader if there is interest in the subject. We just want to convince all our readers that e.g. associativity or commutativity of convolution is always well justified in the context that we are providing, and that in this sense the user of the calculus does not have to worry about possible inconsistencies.

Although the above result, combined with Fubini's theorem indicates that convolution is *commutative* we do not make use of this fact, because we do not want to invoke results from measures theory at this point. In fact, using an approximation argument (using discrete measures) we can get the result. Looking back we can say that this is also more or less the way how we prove Fubini, using suitable resummation of suitable Riemannian sums.

4.2. Convolution and the commutativity question.

Psiconvappr **Proposition 2.** *For every $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ and $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ one has*

convdiscr1 (90)
$$lim_{|\Psi|\to 0} \|D_\Psi \mu * f - \mu * f\|_\infty = 0.$$

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

CompOpenCB1 *Remark* 17. We will demonstrate separately (see [38] for details) that we can say something similar for these convolution operators acting on $\boldsymbol{C}_b(\mathbb{R}^d)$. However, in this case we cannot guarantee uniform convergence anymore, but only *uniform convergence over compact sets.*

convDpismu1 *Remark* 18. Although it looks as if the convolution with $D_\Psi\mu$ on f was just the same (by the adjointness of D_Ψ and Sp_Ψ) as convolving μ with $\mathrm{Sp}_\Psi(f)$ (which in turn is close to f), but this is not so. Since the spline quasi-interpolation operators Sp_Ψ does NOT commute with all translations (or the inflection $f \mapsto f^{\checkmark}$) we have in general

$$\mu(T_x(\mathrm{Sp}_\Psi)^{\checkmark}) \neq \mu(\mathrm{Sp}_\Psi(T_x f^{\checkmark})).$$

siOscEstim1 *Remark* 19. There are different ways to verify this statement. The proof in the course was based on the fact that one can derive (from the definition) first pointwise convergence, hence (by the use of ε-coverings) uniform convergence over compact sets, and finally by the *uniform concentration* (approximation of both μ and f by compactly supported elements in the respective spaces) uniform convergence overall. This is OK but it seems that we can give a better argument, using a pointwise oscillation estimate:

scmeasconv2 (91)
$$[(D_\Psi\mu - \mu) * f](x) \le [|\mu| * \mathrm{osc}_\delta(f)](x), \, \forall x \in \mathbb{R}^d.$$

Unfortunately this argument requires (at the moment) the use of a measure $|\mu|$ (with the same norm as μ, i.e. with $\| \, |\mu| \, \|_{\boldsymbol{M}_b} = \|\mu\|_{\boldsymbol{M}_b}$).

Proof. There are different ways to verify this statement. The proof in the course was based on the fact that one can derive (from the definition) first pointwise convergence, hence (by the use of ε-coverings) uniform convergence over compact sets, and finally by the *uniform concentration* (approximation of both μ and f by compactly supported elements in the respective spaces) uniform convergence overall.

Let us try now to outline the details of the proof, we do it in different steps.

(1) First we prove that we have pointwise convergence. This is easy because convolution can be rewritten (pointwise) as the action of the convolving measure on $T_x f^{\checkmark}$, thus

$$(D_\Psi \mu - \mu)(T_x f^{\checkmark}) \to 0 \ \text{ for } \ |\Psi| \to 0.$$

(2) Using the properties of a net (or sequence) one can easily show that this convergence is uniform over arbitrary finite sets;

(3) As a next step we verify that it is uniform over compact subsets. This is now a general functional analytic argument. Given a compact set $K \subset \mathbb{R}^d$ (or a LCA group) and $\varepsilon > 0$ we find that there is some open (relatively compact) neighborhood U of the identity such that

$$\|T_x f^{\checkmark} - T_{x+u} f^{\checkmark}\|_\infty = \|f - T_{x+u} f\|_\infty < \varepsilon/(3\|\mu\|_{\boldsymbol{M}})$$

for all $u \in U$. Thus the family $(x+U)_{x \in K}$ is a covering of K by open subsets, and since K is compact we can find a finite sequence $(x_k)_{k=1}^K$ such that $(x_k + U)_{k=1}^K$ covers all of K. Choosing now $\delta > 0$ in such a way that we can be sure the

$$|(D_\Psi \mu - \mu)(T_{x_k} f^{\checkmark})| < \varepsilon/3 \ \text{ if } \ |\Psi| \le \delta$$

we can put things together. For every given $x \in K$ we can find at least one $u \in U$ and x_k in the finite list such that $x = x_k + u$, and thus we can estimate as follows

$$\|T_{x_k} f^{\checkmark} - T_{x_k+u} f^{\checkmark}\|_\infty \le \varepsilon/(3\|\mu\|_{\boldsymbol{M}}),$$

and as a consequence

$$|C_\mu(f)(x) - C_\mu(f)(x_k)| \le \|\mu\|_{\boldsymbol{M}} \cdot \varepsilon/(3\|\mu\|_{\boldsymbol{M}}) = \varepsilon/3 \text{ for } x = x_k + u.$$

But since $\|D_\Psi \mu\|_{\boldsymbol{M}} \le \|\mu\|_{\boldsymbol{M}}$ we can argue that in a similar way we also have

$$|C_{D_\Psi \mu}(f)(x) - C_{D_\Psi \mu}(f)(x_k)| \le \varepsilon/3 \text{ for } x = x_k + u.$$

(4) Putting these inequalities together, knowing that one has first to choose the finitely many points (for the estimate which depends only on the required $\varepsilon > 0$ and the known norm of the measure) and *then*, after choosing an appropriate sequence (x_k), and then a common parameter $\gamma \le \delta$ such $|\Psi| \le \gamma$ implies

convestK (92)
$$|C_\mu(f)(x) - C_{D_\Psi \mu}(f)(x)| \le \varepsilon, \ \text{ for all } x \in K. \tag{92}$$

(5) What we have implies so far is: whenever μ and f have compact support, i.e. are of the form $\mu = \mu \cdot p$ for some $p \in \boldsymbol{C}_c(\mathbb{R}^d)$ and $f \in \boldsymbol{C}_c(\mathbb{R}^d)$ then $\mu * f(x) = 0$ outside the compact set $K = \text{supp}(p) + \text{supp}(f)$. We thus have under this extra assumption, in a more compact form

$$\|C_\mu(f) - C_{D_\Psi \mu}(f)\|_\infty \le \varepsilon, \ \text{ whenever } \ |\Psi| \le \gamma, \ \text{ for } x = x_k + u.$$

(6) Next we show that not only for compactly supported measures or functions f (for such functions it is enough to observe the convergence over some compact set, because the convolution product is zero outside of some compact set!) the convergence is possible, but also for general measure $\mu \in \boldsymbol{M}(\mathbb{R}^d)$ and functions $f \in \boldsymbol{C}_0(\mathbb{R}^d)$, using an approximation argument.

(7) The trick will be to approximate μ by measures of the form $\mu \cdot p$, with $p \in \boldsymbol{C}_c(G)$ (plateau functions with compact support, e.g. a finite partial sum of a BUPU), and in a similar way $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ by some $k \in \boldsymbol{C}_c(G)$ (possibly also of the form $k = p_1 \cdot f$. We need then an estimate of the form

convlocap1 (93)
$$\|\mu * f - (\mu \cdot p) * (f \cdot p_1)\|_\infty$$

and a similar estimate with $D_\Psi \mu$, using essentially the boundedness of the bilinear mapping $(\mu, f) \to \mu * f$, in the following manner:

Using the triangle inequality we get the following estimate:

$$\|\mu * f - (\mu \cdot p) * (f \cdot p_1)\|_\infty \leq \|(\mu - \mu \cdot p) * (f \cdot p_1)\|_\infty + \|(\mu \cdot p)(f \cdot p_1 - f)\|_\infty$$
$$\leq \|\mu - \mu \cdot p\|_{\boldsymbol{M}} \|f \cdot p_1\|_\infty + \|\mu \cdot p\|_{\boldsymbol{M}} \|f \cdot p_1 - f\|_\infty.$$

Given $\varepsilon > 0$, and knowing the constants $\|f \cdot p_1\|_\infty \leq \|f\|_\infty$ and $\|\mu \cdot p\|_{\boldsymbol{M}} \leq \|\mu\|_{\boldsymbol{M}}$ we can of course choose $p, p_1 \in \boldsymbol{C}_c(\mathbb{R}^d)$ in such a way that the overall expression is smaller than $\varepsilon/5$, say.

□

THIS UPCOMING STATEMENT SHOULD BE TURNED INTO A PROPOSITION OR EVEN BETTER THEOREM (:FA HINT!) The internal convolution (this is how we have defined it as a multiplication on the bounded measures, realizing the composition of TILS) is exactly the *adjoint action* (up to inversion = flip) of measures via convolution operators on the predual of $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$, which is $\boldsymbol{C}_0(\mathbb{R}^d)$.

So we have the following situation: On the one hand there exists an external action of a bounded measure μ on the (homogeneous) Banach space $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$, via the *convolution operator* C_μ, which was used to define the *internal* multiplication which we called convolution. Recall that this has been defined in such a way that one enforces the associativity law $C_{\mu_1} \circ C_{\mu_2} = C_{\mu_1 * \mu_2}$. But since $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ is the dual space of $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$ we also have the adjoint action, i.e. the operator C'_μ defined as usual by the identity

adjdef02 (94)
$$C'_\mu(\nu)(f) = \nu(C_\mu(f)), \quad f \in \boldsymbol{C}_0(\mathbb{R}^d).$$

It is a good exercise to verify that one has:

rnaladjoint **Theorem 8.** *The operator C'_μ can be described by*

adjconv01 (95)
$$\nu \mapsto \mu^{\checkmark} * \nu, \quad \nu \in \boldsymbol{M}_b(\mathbb{R}^d).$$

Remark 20. The fact that the Banach space $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$ on which the action if the (now) Banach algebra $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ (with convolution) acts, via $\mu \bullet f = C_\mu(f)$ and the fact that this Banach algebra happens to be the dual space to $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$ implies that there is also the adjoint action on $\big(\boldsymbol{C}'_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{C}'_0}\big)$.

The above formula tells us that this adjoint action does not define another multiplication on $\boldsymbol{M}_b(\mathbb{R}^d)$, it is just the already defined convolution combined with the involution $\mu \to \mu^{\checkmark}$.

flip-convC0 (96)
$$(\mu, f) \mapsto \mu^{\checkmark} * f = (\mu * f^{\checkmark})^{\checkmark} \quad \forall \mu \in \boldsymbol{M}_b(\mathbb{R}^d), f \in \boldsymbol{C}_0(\mathbb{R}^d).$$

Preliminary comment: Here it might be the right place to comment on the compatibility of involutions, such as $f \mapsto f^{\checkmark}$ with convolution! (even over groups!)

lipconvmeas (97)
$$(\mu_1 * \mu_2)^{\checkmark} = {\mu_2}^{\checkmark} * {\mu_1}^{\checkmark}, \quad \mu_1, \mu_2 \in \boldsymbol{M}_b(G).$$

Remark 21. First we observe that this involution is 'self-dual' in the following sense: assume that you have a measure, then one can try to define the dual operation on measures, which is then an involution on $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$. On the other hand one can try to apply this involution directly to ordinary functions and then extend it (expecting that it should be w^*-w^*--continuous) to an involution on all of $\boldsymbol{M}_b(\mathbb{R}^d)$. In addition one could go and verify that ${\delta_x}^{\checkmark} = \delta_{-x}$, hence in particular δ_0 is an "even" measure with ${\delta_0}^{\checkmark} = \delta_0$ and infer from this that the transition from μ to $\mu^{\checkmark}$ can also be justified via w^*-approximation using discrete bounded measures.

We are supposed to check the mapping $T'(\mu)$ for $T : f \mapsto f^{\checkmark}$, or $T'(\mu)(f) = \mu(f^{\checkmark})$. For discrete measures it is clear that one has $T'(\delta_x)(f) = \delta_x(f^{\checkmark}(x) = f(-x) = \delta_{-x}(f)$, or $T'(\delta_x) = \delta_{-x}$. In fact, we can write $f^{\checkmark} = D_{-1}(f)$, hence $T' = \mathrm{St}_{-1}$, consistent with our operations.

One can also make a similar considerations for other involutions, e.g. $f \mapsto f^*$, given by $f^*(x) = \overline{f(-x)}$, which is obviously a combination (the order does *not* matter in this case) of the involution $f \mapsto f^{\checkmark}$ and the natural involution on function spaces coming from complex conjugation: $f \mapsto \overline{f}$.

In order to verify this other consistency we can argue as follows: By definition we have

$$C_{\mu_1 * \mu_2} h(z) = C_{\mu_1}(C_{\mu_2} h)(z)$$

for all $h \in \boldsymbol{C}_0(\mathbb{R}^d)$ and $z \in \mathbb{R}^d$, hence for $z = 0$, hence by our conventions

$$(\mu_1 * \mu_2)(h^{\checkmark}) = \mu_1((\mu_2 * h)^{\checkmark}) = \mu_1({\mu_2}^{\checkmark} * h^{\checkmark})$$

or by writing $f = h^{\checkmark} \in \boldsymbol{C}_0(\mathbb{R}^d)$:

dist-conv
$$(\mu_1 * \mu_2)(f) = \mu_1({\mu_2}^{\checkmark} * f) = \mu_1((\mu_2 * f^{\checkmark})^{\checkmark}). \tag{98}$$

This form of defining the convolution of two "abstract objects" is very much like the typical definition of the convolution of distributions. It makes sense as long as the convolution of the functional μ_2 with test functions results in another test function [42], here a member of $\boldsymbol{C}_0(\mathbb{R}^d)$.

Since the algebra $\mathcal{L}(\boldsymbol{C}_0(\mathbb{R}^d))$ is *not* commutative it is not at all clear from this definition why $(\boldsymbol{M}(\mathbb{R}^d), \|\cdot\|_M, *)$ should be a commutative Banach algebra, which is in fact true.

In order to prepare for this statement we have to provide a few more statements.

The compatibility of the (isometric) dilation operators D_ρ with translations, i.e. the rule **[Rd-specific!]**

dil-trans
$$D_\rho T_z = T_{z/\rho} D_\rho \tag{99}$$

makes it possible to define another norm preserving *automorphism* for the Banach algebra $(\boldsymbol{M}(\mathbb{R}^d), \|\cdot\|_M, *)$, given as follows: [43].

Strhodef **Definition 23.** The *adjoint* action of the group $\mathbb{R}^+$ on $\boldsymbol{M}(\mathbb{R}^d)$ is defined as the family of adjoint operators, i.e. $\mathrm{St}_\rho := {\mathrm{D}_\rho}'$ is defined on $\boldsymbol{M}(\mathbb{R}^d)$ via:

def-Stroh
$$\mathrm{St}_\rho \mu(f) := \mu(D_\rho f), \quad \forall f \in \boldsymbol{C}_0(\mathbb{R}^d),\ \rho \neq 0. \tag{100}$$

We consider this the *mass-preserving* dilation operator (the letters "St" can be read as *stretching*[44] operator). It is easy to check that for $k \in \boldsymbol{C}_c(\mathbb{R}^d) \subset \boldsymbol{L}^1(\mathbb{R}^d) \cap \boldsymbol{C}_0(\mathbb{R}^d)$ one has:

rodrohcomp1
$$\mathrm{St}_\rho[\mu_k] = \rho^{-d} \cdot \mu_{[D_{1/\rho} k]}, \quad \text{and} \quad \mathrm{D}_\rho[\mu_k] = \rho^d \mu_{[\mathrm{St}_\rho k]}, \tag{101}$$

[42]Then the right hand side makes sense and can be used to give the expression $\mu_1 * \mu_2$ a meaning. Note that this is *dangerous terrain*, because it may happen that such an individually defined convolution between distributions - similar to pointwise products - may turn out to be *non-associative*!

[43]We will see that these two operations on $\boldsymbol{C}_0(\mathbb{R}^d)$ resp. $\boldsymbol{M}(\mathbb{R}^d)$ are adjoint to each other.

[44]In German language we use the words "STrecken" for "stretching" ($\rho > 1$) and "STauchen" for "compression, which occurs for $\rho < 1$, so it German the symbol fits even better for this Stauch-Streck operators.

which justifies the equation

StroDro2
$$\mathrm{St}_\rho = \rho^{-d}\,\mathrm{D}_{1/\rho} \quad \text{resp.} \quad \mathrm{D}_{1/\rho} = \rho^d\,\mathrm{St}_\rho,\ \rho \neq 0. \tag{102}$$

Later on (for the sampling theorem) we will need two observations, involving the Shah-distribution $\sqcup\!\sqcup := \sum_{n\in\mathbb{Z}^d}\delta_n$, often also called a *Dirac comb* [45].

Note that $\rho = \alpha$ we have $\mathrm{St}_\rho(\sqcup\!\sqcup) = \sqcup\!\sqcup_\alpha := \sum_{n\in\mathbb{Z}^d}\delta_{\alpha n}$ and, according to the above formulas $\mathrm{D}_\rho\sqcup\!\sqcup = (1/\rho)\cdot\sqcup\!\sqcup_{1/\rho}$.

The corresponding variant of convergence is then the so-called *vague convergence*:

vagconvdef2 **Definition 24.** A sequence (resp. net) of (bounded or unbounded) measures μ_α is convergent in the *vague sense* to the limit measure μ_0 if one has

$$\lim_\alpha \mu_\alpha(k) = \mu_0(k), \quad \forall k \in \boldsymbol{C}_c(\mathbb{R}^d).$$

Remark 22. It is an easy exercise to check that a bounded net in $(\boldsymbol{M}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}})$ is vaguely convergent if and only if it is w^*-convergent with respect to the $\boldsymbol{C}_0(\mathbb{R}^d)/MbRd$ duality. In other words, one can extend the range of validity of the limit relation above from $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ to all $k \in \boldsymbol{C}_0(\mathbb{R}^d)$.

Using this concept we can show

shah-conv **Lemma 22.** *One has the following vague convergence of (unbounded) measures:*

- $\lim_{\rho\to\infty}\mathrm{St}_\rho\sqcup\!\sqcup = \lim_{\rho\to\infty}\sqcup\!\sqcup_\rho \ = \ \delta_0;$
- $\lim_{\rho\to\infty}\mathrm{D}_\rho\sqcup\!\sqcup \ = \lim_{\rho\to\infty}\rho^{-d}\sqcup\!\sqcup_{1/\rho} \ = \ 1.$

Proof. i) Given $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ one finds $R > 0$ such that $k(x) = 0$ if only $|x| \geq R$. It is then clear that $\mathrm{D}_\rho k(z) = 0$ for $|z| > 1$, as long as $\rho > R$, and thus in fact

dilStrsha2
$$\mathrm{St}_\rho\sqcup\!\sqcup(k) = \sum_{n\in\mathbb{Z}^d}\delta_n(\mathrm{D}_\rho(k)) = \delta_0(\mathrm{D}_\rho(k)) = k(\rho 0) = k(0). \tag{103}$$

ii) As for the second argument it is sufficient to verify that the functional $\mathrm{D}_\rho\sqcup\!\sqcup = \rho^{-d}\sqcup\!\sqcup_{1/\rho}$ represents exactly a typical Riemannian sum applied to $f \in \boldsymbol{C}_c(\mathbb{R}^d)$, namely $\rho^{-d}\sum_{n\in\mathbb{Z}^d} f(n/\rho)$ which is known to tend (for $\rho \to \infty$) to the integral, which itself can be written in the form of an integrator against the constant function $\mathbf{1}$, i.e. the limit is $\int_{\mathbb{R}^d} f(x)\mathbf{1}(x)dx$. This arguments completes the proof. $\square$

Being defined as adjoint operators each of the operators St_ρ is not only isometric on $\boldsymbol{M}(\mathbb{R}^d)$, but also w^*-w^*--continuous on $\boldsymbol{M}(\mathbb{R}^d)$.

StroAltdef **Lemma 23.** *Given $\mu \in \boldsymbol{M}(\mathbb{R}^d)$ the uniquely determined measure corresponding to the operator $D_\rho^{-1} \circ C_\mu \circ D_\rho$ coincides with $\mathrm{St}_\rho\mu$. Also from this characterization the fact that St_ρ defines an automorphism on $(\boldsymbol{M}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}})$ is evident.*

Proof. One has to apply both operators to an arbitrary $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ and verify that the result is the same. $\square$

Remark 23. Reformulated in a more usual way (using convolutions), one has the identity:

stromsconv1
$$\mathrm{St}_\rho\mu * f = {\mathrm{D}_\rho}^{-1}(\mu * \mathrm{D}_\rho(f)),\ f \in \boldsymbol{C}_0(\mathbb{R}^d). \tag{104}$$

[45] It is of course an unbounded measure, but not too bad. We will see that it can be viewed as a bounded linear functional on the so-called *Wiener algebra* $\boldsymbol{W}(\mathbb{R}^d)$, see Def.43.

We collect the basic facts for this new mapping:

conv-strho
$$\mathrm{St}_\rho \mu_1 * \mathrm{St}_\rho \mu_2 = \mathrm{St}_\rho(\mu_1 * \mu_2) \tag{105}$$

Proof. To be provided later. It shows that the two alternative definitions above are indeed equivalent!! □

We have **[Rd-specific!]**

Strho-isom
$$\|\mathrm{St}_\rho \mu\|_{\boldsymbol{M}} = \|\mu\|_{\boldsymbol{M}}. \tag{106}$$

While the (mass-preserving) operator St_ρ is well compatible with convolution, the (value-preserving) operators D_ρ is ideally compatible with pointwise multiplications, since obviously

Drho-ptwmul
$$\mathrm{D}_\rho(f \cdot g) = \mathrm{D}_\rho f \cdot \mathrm{D}_\rho g. \tag{107}$$

While D_ρ is well adapted to pointwise multiplication and St_ρ to convolution, it is still possible to use D_ρ in a convolution context, observing the following relations:

Drho-conv1
$$\mathrm{St}_\rho \mu * D_{1/\rho} f = D_{1/\rho}(\mu * f) \quad \text{resp.} \quad \mathrm{St}_{1/\rho}\, \mu * \mathrm{D}_\rho f = \mathrm{D}_\rho(\mu * f). \tag{108}$$

One could state similar formulas for St_ρ with respect to pointwise multiplication.

xuxuxu **Lemma 24.** *The group $\mathbb{R}^+$ is acting strongly continuous on $\boldsymbol{C}_0(\mathbb{R}^d)$ via D_ρ, i.e. the mapping $x \to \mathrm{D}_\rho f$ is a continuous mapping from $\mathbb{R}^+$ to $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big), \forall f \in \boldsymbol{C}_0(\mathbb{R}^d)$. In fact, $(\rho, f) \mapsto \mathrm{D}_\rho f$ is jointly continuous.*

Proof. The proof can be done in an elementary way, checking continuity at $\rho = 1$ for $f \in \boldsymbol{C}_c(\mathbb{R}^d)$ (using the uniform continuity of f and the joint compact support of all the functions $(\mathrm{D}_\rho f)_{\rho \in [1/2,2]}$). The general case of $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ follows then by approximation (Try and check the details if you are not sure about the argument!)[46] □

Stroh-map **Lemma 25.** *The mapping $\rho \mapsto \mathrm{St}_\rho$ is continuous from $\mathbb{R}^+$ into $\mathcal{L}(\boldsymbol{M}(\mathbb{R}^d))$, endowed with the strong operator topology. It is also true that the mapping $(f, \rho) \mapsto C_{\mathrm{St}_\rho \mu}(f)$ is continuous from $\boldsymbol{C}_0(\mathbb{R}^d) \times \mathbb{R}^+$ into $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$. Obviously the operators $(\mathrm{St}_\rho)_{\rho>0}$ form a commutative group of isometric operators on $\boldsymbol{M}(\mathbb{R}^d)$ (but also - since they leave $\boldsymbol{L}^1(\mathbb{R}^d)$ invariant - on $\big(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\big)$).*

stro-group
$$\mathrm{St}_{\rho_1} \circ \mathrm{St}_{\rho_2} = \mathrm{St}_{\rho_1 \cdot \rho_2} = \mathrm{St}_{\rho_2} \circ \mathrm{St}_{\rho_1} \qquad \rho_1, \rho_2 > 0. \tag{109}$$

Proof. Proof to be typed later on. The isometric action on any of the $\boldsymbol{L}^p$-spaces results from the transformation rule for integrals, while the composition law follows by easy direction computation.

Among others it is obvious that the property $\|\mathrm{D}_\rho f\|_\infty = \|f\|_\infty$ for all $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ implies $\|\mathrm{St}_\rho \mu\|_{\boldsymbol{M}_b} = \|\mu\|_{\boldsymbol{M}_b}$ for all $\mu \in (\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$.[47] □

trho-deltas
$$\mathrm{St}_\rho \delta_x = \delta_{\rho x}, \quad \rho \in \mathbb{R}, \tag{110}$$

and based on

Drho-deltas
$$\mathrm{D}_\rho \delta_x = |\rho^{-d}|\, \delta_{x/\rho}, \quad \rho \neq 0. \tag{111}$$

[46] This can be seen as a general fact: The mapping $\rho \to \mathrm{D}_\rho$ is a group representation of $(\mathbb{R}_+, \cdot)$ as a group of isometric transformation on the Banach space $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$, i.e. one can use the fact that $\mathrm{D}_{\rho_2} \circ \mathrm{D}_{\rho_1} = \mathrm{D}_{\rho_2 \cdot \rho_1}$.

[47] Generally speaking: the adjoint operator to a given isometric operator is again isometric on the dual Banach space.

In engineering books this last relation is derived using formal integral, but this is just the same idea in a different notation. So in order to derive the identity

dildelt3 (112)
$$\delta(at) = 1/|a|\,\delta(t/a),$$

one uses the transformation rule in order to derive it:

dildelt4 (113)
$$\int_{-\infty}^{\infty} \delta(at)h(t)dt = \int_{-\infty}^{\infty} \delta(s)h(t/a)ds/|a| = \frac{1}{|a|}h(0) = \int_{-\infty}^{\infty} [\frac{1}{|a|}\delta(s)]h(s)ds,$$

which is identical (almost step by step) with our argumentation:

dildelt5 (114)
$$[\mathrm{D}_\rho \delta_x](h) = \delta_x(\mathrm{St}_\rho h) = [\mathrm{St}_\rho h](x) = |\rho^{-d}|h(x/\rho) = [|\rho^{-d}|\delta_{x/\rho}](h).$$

One can describe the equation (112) the use of a transformation rule normally valid for ordinary functions, e.g. for $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ instead of $\delta(t)$. In some sense it is justified, because St_ρ resp. D_ρ are the w^*-w^*--continuous extensions of the ordinary operators and one might derive the validity by taking limits in the w^*-sense, or in other words replace the functional δ_x by an approximating sequence of the form $T_x \mathrm{St}_\rho g$, for $\rho \to 0$.

Remark 24. Due to the w^*-density of finite discrete measures in $\boldsymbol{M}(\mathbb{R}^d)$ one can characterize St_ρ as the uniquely determined norm-to-norm continuous *and* w^*-w^*--continuous mapping which is isometric on $(\boldsymbol{M}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}})$ and satisfies formula (110).

Remark 25. There is an interesting connection to probability theory and random variables. For each random variable X with values in $\mathbb{R}^d$ there exists exactly on probability measure μ on $\mathbb{R}^d$ such that for every Borel set $M \subseteq \mathbb{R}^d$ one has

$$E\,(X(t) \in M) = \int_M d\mu(x).$$

In fact, any probability statement about the random variable $X(t)$ can be expressed knowing only μ.

Remark 26. In this context convolution comes into the game as follows: Assume that X_1, X_2 are two *independent random variables* (this has to be defined and explained separately, just think for now of things which have “nothing to do with each other in terms of probability”) with corresponding measures μ_1, μ_2. Then the random variable $X_1 + X_2$ has of course another measure μ associated to it. The claim is then: $\mu = \mu_1 * \mu_2(!)$.

This is easy to understand in the case of constant random variables. If it is 100% certain that $X(t) = x, \forall t \in \Omega$ ad $Y(t) = y, \forall t \in \Omega$, or in other words $\mu_X = \delta_x$ and $\mu_Y = \delta_y$, then it is obvious that $(X+Y)(t) = x+y, \forall t \in \Omega$, hence $\mu_{X+Y} = \delta_{x+y}$ which matches perfectly with the observation that $\delta_{x+y} = \delta_x * \delta_y$.

For discrete random variables, i.e. for random variables taking only a finite number of discrete values with positive probability (such as a dice with $E(D = k) = 1/6$ for $k = 1, \ldots, 6$ similar rules apply, because the corresponding discrete measures (in our case $\mu_D = \sum_{k=1}^6 \delta_k/6$) have to be convolved.

A good exercise is to check what can be said about the probabilities of a sum of two independent dices (wether one has two independent dices thrown in parallel or two independent repetitions of the same dice does not matter!). Obviously there are altogether 36 possibilities, one of them being the pair $(1,1)$ (resp. $(6,6)$). Another three possibilities (giving exactly the sum of 4) are the pairs $(1,3), (2,2), (3,1)$, etc.. Checking this at a formal level reveals the analogy to convolution in a discrete setting, resp. the Cauchy product formula for the multiplication of polynomials (as we learn it at school).

HINT: DISTRIBUTION FUNCTION: $F(z) = \int_{-\infty}^{z} d\mu(t)$.

The usual way to provide visual information about a probability measure is the look at its distribution function, or – in the case of experimental data – some *histogram.* Depending on the available amount of data and the variance of the measure one can have a good representation with a certain number of bins, not too small but also not too big. The law of the large number tells us essentially that we can expect that for a large number of realizations the probability distribution can be "read off from the histogram", or from a different view-point: The histogram will give a realistic picture about the probability distribution, in some discretized way.

Assume that we have a real-valued *random variable* $X : \Omega \to \mathbb{R}$, with a *probability density function* $\mu = p_X$ and *cumulative distribution function* F_X, i.e. with $F_X(x) = P(X \leq x), x \in \mathbb{R}$, hence

$$\int_y^z p_X(x)dx = F_X(z) - F_X(y) = \int \mathbf{1}_{[y,z]}(x)\, d\mu(x).$$

Then of course we expect a histogram to contain good approximative information about integrals over the collection of such integrals, for a collection of intervals $[k/n, (k+1)/n]$, $k \in \mathbb{Z}$, or equivalently, the numerical sequence $\mu(\mathbf{1}_{[k/n,(k+1)/n]})_{k\in\mathbb{Z}}$, i.e. one has information about a collection $\mu(\psi_k)$, where $(\psi_k)_{k\in\mathbb{Z}}$ forms a partition of unity for $\mathbb{R}$ of diameter $1/n$, for some given $n \in \mathbb{N}$. If the range of X is in $[-1, 1]$ (for example) it is of course enough to make use of bins supported in that region of $\mathbb{R}$.

We will make use of a so-called BUPUs which will serve a similar purpose: They will allow us to localize information provided by the measure μ, and recollect the global information by summing up all these local pieces. Again one can define BUPUs in the setting of general locally compact (LC) groups G, but for a more convenient presentation of the underlying ideas we will mostly make use of so-called regular BUPUs, because they are easy to generate on $\mathbb{R}^d$ and make proofs much shorter. Nevertheless we take care that the properties that we use are the same ones which by definition are satisfied by the more general BUPUs.

Remark 27. Note that the central limit theorem can also be reformulated as a statement[48] about $\mathrm{St}_{1/\sqrt{N}}(\mu^{N*})$, which is convergent to the normal distribution if $\mu(\mathbf{1}) = 1$.

There is also a nice compatibility of the Fourier transform with dilations, which in fact provides another argument for the use of two versions of the dilation operators.

Fourdil1 **Lemma 26.** $\mathcal{F}(\mathrm{St}_\rho \mu) = \mathrm{D}_\rho(\mathcal{F}\mu), \ \ \rho \neq 0, \quad \mu \in \boldsymbol{M}_b(\mathbb{R}^d)$.

In other words, the Fourier transform (and also its inverse) is an *intertwining operator* for these two types of dilation operators:

ntertwindil (115)
$$\mathcal{F} \circ \mathrm{St}_\rho = \mathrm{D}_\rho \circ \mathcal{F} \quad \text{for } \rho \neq 0.$$

For the Shannon sampling theorem one uses dilated (compressed) interpolators, and convolves with the sampled functions

$$[f \cdot ⊔⊔_\rho] * D_{1/\rho}\,\varphi = [f \cdot \mathrm{D}_\rho ⊔⊔] * \mathrm{St}_\rho \varphi$$

Observing that $\mathrm{D}_\rho ⊔⊔ \to \mathbf{1}$ for $\rho \to \infty$ (in the vague sense), hence one may conclude

$$f \cdot \mathrm{D}_\rho ⊔⊔ \to f \cdot \mathbf{1} = f$$

in the w^*-sense,

$$f = w^* - \lim f \cdot \mathrm{D}_\rho ⊔⊔ \to f \cdot \mathbf{1} = f$$

[48]It would be interesting to know under which conditions one has relative compactness hence norm convergence of this sequence in some nice, hopefully small function space, e.g. in $(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0})$. E.g. under the assumption that $\mu = g \in \boldsymbol{Q}_s(\mathbb{R}^d)$ for sufficiently large $s > d$, for example.

and moreover $\mathrm{St}_\rho\varphi \to \delta_0$ if only $\int_{\mathbb{R}^d}\varphi(x)dx = 1 = \hat{\varphi}(1)$, also for $\rho \to 0$. Altogether we may conclude that (at least for $f \in \boldsymbol{C}_c(\mathbb{R}^d)$):

dilconv04 (116)
$$[f \cdot \mathrm{D}_\rho \sqcup\!\sqcup] * \mathrm{St}_\rho\varphi \to_{w^*} [f \cdot 1] * \delta_0 = f, \quad \text{for } \rho \to 0.$$

5. Basic functional analytic considerations

A series of lemmata (lemmas) making use of density.

Lemma 27. *Assume a bounded linear mapping between two Banach spaces is given on a dense subset only, then it can be extended in a unique way to a bounded linear operator of equal norm on the full space.*

Proof. Obviously it is enough to know a bounded, linear mapping T on a dense subset D of a Banach (or in fact just normed) space, in order to observe that for any element v in the domain of T one has a convergent sequence $(d_n)_{n\in\mathbb{N}}$ in D with limit x. Hence $(d_n)_{n\in\mathbb{N}}$ is a Cauchy sequence, so that the boundedness of T implies that $(Td_n)_{n\in\mathbb{N}}$ is a again a Cauchy sequence in the range, hence (by completeness) has a limit. The uniqueness of this extension and the fact that this extension has the same norm, i.e. that

$$|||T||| := \|T\|_{Op} = sup_{\|d\|\leq 1}\|Td\|$$

is also an immediate consequence. □

AIsubspace **Lemma 28. Test of BAI on a dense subspace:**
A bounded family $(e_\alpha)_{\alpha\in I}$ in a Banach algebra $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ is a BAI for $\boldsymbol{A}$ if (only) for some total subset $D \subseteq \boldsymbol{A}$ one has:

totalst1 (117)
$$\|e_\alpha \cdot d - d\|_{\boldsymbol{A}} \to 0 \quad \forall d \in D.$$

Proof. Obviously the validity of (117) is necessary. Conversely, it is easy to derive - using the properties of a net - that one has (uniform) convergence for finite linear combinations of elements from D, and then by a density argument one verifies convergence for all elements. [49] In this last step the (a priori) uniform boundedness of the approximate unit plays a crucial role. □

5.1. Iterated limits of nets and interchange of order.

The following result should be compared with a theorem on *iterated limits* provided in Kelley's book ([65], p.69).

doublenet **Lemma 29.** *Assume that $(T_\alpha)_{\alpha\in I}$ and $(S_\beta)_{\beta\in J}$ are two bounded nets of operators in $\mathcal{L}(\boldsymbol{V})$, which are strongly convergent to limits T_0, and S_0 resp. i.e.*

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

Then the net $(T_\alpha \circ S_\beta)_{(\alpha,\beta)}$ (with index set $I \times J$ and natural order[50]) is also strongly convergent, with limit $T_0 \circ S_0$, i.e. for each $\mathbf{v} \in \mathbf{V}$ one has:

doubnetconv (118)
$$T_o[S_0(\mathbf{v})] = [T_0 \circ S_0](\mathbf{v}) = \lim_{\alpha,\beta}[T_\alpha \circ S_\beta](\mathbf{v}).$$

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

doubnetest (119)
$$\|T_0(S_0(\mathbf{v})) - T_\alpha(S_\beta(\mathbf{v}))\| \leq \varepsilon.$$

[49] Of course this is more or less a statement about strong operator convergence for a net of bounded operators.
[50] Of course I and J may have completely different order!

In particular we have in the sense of strong limits:

doublenet2
$$T_0 \circ S_0 = \lim_\alpha \lim_\beta T_\alpha \circ S_\beta = \lim_\beta \lim_\alpha T_\alpha \circ S_\beta \tag{120}$$

Proof. Note that $|||T_\alpha||| \leq C < \infty$. The statement depends on the following estimate

oublenetest
$$\|T_\alpha[S_\beta(\mathbf{v})] - T_0[S_0(\mathbf{v})]\| \leq \|T_\alpha[S_\beta(\mathbf{v})] - T_\alpha[S_0(\mathbf{v})]\| + \|T_\alpha[S_0(\mathbf{v})] - T_0[S_0(\mathbf{v})]\| \tag{121}$$

The first expression can be estimated as follows:

doubest1
$$\|T_\alpha[S_\beta(\mathbf{v})] - T_\alpha[S_0(\mathbf{v})]\| \leq |||T_\alpha|||\, \|S_\beta(\mathbf{v}) - S_0(\mathbf{v})\| \leq C\|S_\beta(\mathbf{v}) - S_0(\mathbf{v})\|, \tag{122}$$

which gets $< \varepsilon/2$ for $\beta \succeq \beta_0$ (chosen for ε/C), while the second term can be estimated by

doubest2
$$\|T_\alpha[S_0(\mathbf{v})] - T_0[S_0(\mathbf{v})]\| \leq \varepsilon/2, \tag{123}$$

for any α with $\alpha \succeq \alpha_0$ (choosing $\mathbf{w} = S_0(\mathbf{v})$). Combining these estimates we obtain

doubest5
$$\|T_\alpha[S_\beta(\mathbf{v})] - T_0[S_0(\mathbf{v})]\| \leq \varepsilon \quad \forall \beta \succeq \beta_0, \alpha \succeq \alpha_0. \tag{124}$$

Finally we have to check that the validity of (124) implies that also the iterated limits exist (of course with the same limit). We elaborate on the first iterated limit (namely $\lim_\alpha \lim_\beta$). The other one works in the same way[51]. So for the situation in (119) we can first fix any α and look at the net $(T_\alpha(S_\beta(\mathbf{v})))_{\beta \in J}$. By assumption the limit $S_0(\mathbf{v}) = lim_\beta S_\beta(\mathbf{v})$ exists. Since T_α (for fixed α first) is a bounded linear operator also the following limits exist and can be estimated by

$$\|T_0(S_0(\mathbf{v})) - T_\alpha(\lim_\beta S_\beta(\mathbf{v}))\| \leq \varepsilon, \quad \text{if only} \quad \alpha \succeq \alpha_0.$$

Setting $\mathbf{w}_0 := S_0(\mathbf{v})$ we see that by assumption also $(T_\alpha(\mathbf{w}_0))$ is convergent, and of course the last estimate remains true in the limit (due to the continuity of the norm), hence

$$\|T_0[S_0(\mathbf{v})] - \lim_\alpha \lim_\beta T_\alpha[S_\beta(\mathbf{v})]\| \leq \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, equation (120) is valid. □

Remark 28. Note that only the boundedness of the net (T_α) is only required, not that of (S_β). Moreover, in the case of sequences (instead of nets) the *uniform boundedness principle* can be applied, which tells us that the strong convergence of operators implies the norm boundedness of the corresponding sequence. This situation could be formulated as a corollary then, i.e. strong convergence of (T_n) and (S_k) implies

doubesequ
$$\lim_n \lim_k T_n[S_k(\mathbf{v})] = T_0[S_0(\mathbf{v})] = \lim_k \lim_n T_n[S_k(\mathbf{v})] \tag{125}$$

AIoncomp **Lemma 30. One can choose the BAI elements from a dense SUBSPACE .** *Let $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ be a Banach algebra with a bounded approximate unit, and D some dense subset of $\boldsymbol{A}$. Then there exists also approximate units $(d_\alpha)_{\alpha \in I}$ in D*[52].

Proof. One just has to choose d_α close enough to e_α, especially for α large enough (to be expressed properly). By the density of D one can do this. If D is a subspace (not just a subset) one can renormalizing the new elements so that they have the same norm in $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ as the original elements $(e_\alpha)_{\alpha \in I}$. Details are left to the reader. □

[51]Observe that the *order* in which the operators are applied does NOT change!!!

[52]If D is a dense *subspace* the new family $(d_\alpha)_{\alpha \in I}$ often can be chosen to be of equal norm, as one can verify easily by renormalizing the approximating sequence.

This fact is useful, if one wants to conclude from the existence of bounded approximate units in the Banach convolution algebra $\big(\boldsymbol{L}^1(\mathbb{R}^d),\, \|\cdot\|_1\big)$ that one can have compactly supported ones (the usual approach, with shrinking support), but also band-limited ones (because the band-limited functions form a dense ideal in $\big(\boldsymbol{L}^1(\mathbb{R}^d),\, \|\cdot\|_1\big)$).

Lemma 31. Uniform action of BAIs on compact subsets
By the definition a family $(e_\alpha)_{\alpha\in I}$ acts pointwise like an identity in the limit case, for each "point". This action is even uniformly over finite sets, and hence over compact sets, by approximation (we shall use the acronym unifcomp-convergence for this situation in the future).

Proof. That one has uniform convergence over finite subsets is easily verified, using the property of a net. The inductive step is based on the following argument: Assume that one has found α_0 such that

$$\|e_\alpha \cdot a_i - a_i\| \leq \varepsilon \quad \forall \alpha \succ \alpha_0,\, 1 \leq i \leq m.$$

Since $\|e_\alpha \cdot a_{m+1} - a_{m+1}\| \leq \varepsilon$ for all $\alpha \succ \alpha_1$ we just have to choose some index α_2 with $\alpha_2 \succ \alpha_0$ *and* $\alpha_2 \succ \alpha_1$ (which is possible due to the definition of directed sets). Obviously

$$\|e_\alpha \cdot a_i - a_i\| \leq \varepsilon \quad \forall \alpha \succ \alpha_2,\, 1 \leq i \leq m+1.$$

In order to come up with uniform convergence over compact sets, we use again a typical approximation argument. Given any compact set $M \subseteq \boldsymbol{C}_0(\mathbb{R}^d)$ and $\varepsilon > 0$ we have to find some index α_3 such that

$$\|e_\alpha \cdot a - a\| \leq \varepsilon \quad \forall a \in M.$$

Recalling that $(e_\alpha)_{\alpha\in I}$ is bounded, i.e. $\|e_\alpha\| \leq C$ for some $C \geq 1$ for all $\alpha \in I$, we may choose some finite subset $F \subseteq M$ such that for any $a \in M$ there exists $a_i \in F$ with $\|a_i - a\| \leq \varepsilon/(3C) > 0$ (which is just another positive constant, known once we know C and ε). Hence for any given $a \in M$ one can use one of such elements $a_i \in F$ in order to argue that the triangular inequality implies (adding and subtracting the term $e_\alpha \cdot a_i$).

$$\|e_\alpha \cdot a - a\| \leq \|e_\alpha \cdot (a - a_i)\| + \|e_\alpha \cdot a_i - a_i\| + \|a_i - a\|.$$

If we choose now α_3 such that $\|e_\alpha \cdot a_i - a_i\| \leq \varepsilon/3 \quad \forall \alpha \succ \alpha_3$ and $a_i \in F$ we obtain altogether (more details are left to the reader, ...),

$$\|e_\alpha \cdot a - a\| \leq \varepsilon.$$

□

boundedAI1 **Lemma 32.** *The following properties are equivalent for a Banach algebra $(\boldsymbol{A},\, \|\cdot\|_{\boldsymbol{A}})$:*

- *there is a bounded approximate identity in $(\boldsymbol{A},\, \|\cdot\|_{\boldsymbol{A}})$;*
- *there exists $C > 0$ such that for every finite subset $F \in \boldsymbol{A}$ and $\varepsilon > 0$ there exists element $h \in \boldsymbol{A}$ with $\|h\| \leq C$, such that*

$$\|h \cdot a - a\| \leq \varepsilon \quad \forall a \in F.$$

The argument to turn this family into a bounded net is the obvious one. One just has to set $\alpha := (F, \varepsilon)$, and defining such a pair "stronger than another pair $\alpha_1 := (F_1, \varepsilon_1)$ if $F_1 \supseteq F$ and $\varepsilon_1 \leq \varepsilon$. It is easy to verify that this defines a directed set, and that the choice $e_\alpha = h$ (corresponding to the pair (K, ε) as describe above) turn $(e_\alpha)_{\alpha\in I}$ into a bounded and convergent net, hence constitutes a BAI for $(\boldsymbol{A},\, \|\cdot\|_{\boldsymbol{A}})$.

Note: In words: The existence of a BAI is equivalent to the existence of a bounded net in $(\boldsymbol{A},\, \|\cdot\|_{\boldsymbol{A}})$ such that the action of "pointwise multiplication" (each element e_α is identified with the left algebra multiplication operator $a \mapsto e_\alpha \cdot a$) is convergent to the identity operator in the strong operator norm topology (which is just the pointwise convergence of operators).

Note: Sometimes one observes that one has unbounded approximate identities, where the "cost" (i.e. the norm of e_α grows as the required quality of approximation, expressed by the smallness of ε), tends to the ideal limit zero. It makes sense to think of limited costs (given by $C > 0$) for arbitrary good quality of approximation which makes the BAI so useful.

Theorem 9. *(The Cohen-Hewitt factorization theorem, without proof, see [58])*
Let $(\boldsymbol{A}, \cdot, \|\cdot\|_{\boldsymbol{A}})$ be a Banach algebra with some BAI, then the algebra factorizes, which means that for every $a \in \boldsymbol{A}$ there exists a pair $a', h' \in \boldsymbol{A}$ such that $a = h' \cdot a'$, in short: $\boldsymbol{A} = \boldsymbol{A} \cdot \boldsymbol{A}$. In fact, one can even choose $\|a - a'\| \leq \varepsilon$ and $\|h'\| \leq C$.

There is a more general result, involving the terminology of **Banach modules**.

BanMod **Definition 25.** A Banach space $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ is a *Banach module* over a Banach algebra $(\boldsymbol{A}, \cdot, \|\cdot\|_{\boldsymbol{A}})$ if one has a bilinear mapping $(a, b) \mapsto a \bullet b$, from $\boldsymbol{A} \times \boldsymbol{B}$ into $\boldsymbol{B}$ with

BanModEst
$$\|a \bullet b\|_{\boldsymbol{B}} \leq \|a\|_{\boldsymbol{A}} \|b\|_{\boldsymbol{B}} \quad \forall\, a \in \boldsymbol{A}, b \in \boldsymbol{B}, \tag{126}$$

which behaves like an ordinary multiplication, i.e. is associative, distributive, etc.:

BanModAss
$$a_1 \bullet (a_2 \bullet b) = (a_1 \cdot a_2) \bullet b \quad \forall a_1, a_2 \in \boldsymbol{A}, b \in \boldsymbol{B}. \tag{127}$$

A special case of a Banach module is a Banach module inside the Banach algebra:

BanIdDef **Definition 26.** A Banach space $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ is a *Banach ideal* in (or within, or of) a Banach algebra $(\boldsymbol{A}, \cdot, \|\cdot\|_{\boldsymbol{A}})$ if $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ is continuously embedded into $(\boldsymbol{A}, \cdot, \|\cdot\|_{\boldsymbol{A}})$, and if in addition (317) is valid with respect to the internal multiplication inherited from $\boldsymbol{A}$.

A variant of this definition is found in the literature on so-called *abstract Segal algebras.* Although there is not a completely exact match in the terminology one can roughly say, that Hans Reiter's *Segal algebras* are just the abstract Segal algebras of the Banach convolution algebra $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ resp. $\left(\boldsymbol{L}^1(G), \|\cdot\|_1\right)$ over a LCA group.

bsSegalAlg1 **Definition 27.** A Banach ideal $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ within a Banach algebra $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ is called an *abstract Segal algebra* if it is also dense in $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ and an essential Banach module with respect to the internal multiplication.

Practically, for the case of a Banach algebra $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ with BAI (bd. approx. identity) $(\mathbf{e}_\alpha)$ it means that for all $\mathbf{b} \in \boldsymbol{B}$ one has:

$$\|\mathbf{e}_\alpha \bullet \mathbf{b} - \mathbf{b}\|_{\boldsymbol{B}} \to 0 \quad \text{for} \quad \alpha \to \infty.$$

Ex1Banmod **Lemma 33.** *(Ex. to Def. 25) A Banach space $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ is an (abstract) Banach module over a Banach algebra $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ [53] if and only if there is a non-expansive (hence continuous) linear algebra homomorphism J from $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ into $\mathcal{L}(\boldsymbol{B})$.[54]*

Proof. For a Banach module the mapping: $a \mapsto J(a) : [b \mapsto a \bullet b]$ defines a linear mapping from $\boldsymbol{A}$ into $\mathcal{L}(\boldsymbol{B})$.

Conversely, one can define a $\boldsymbol{A}$-Banach module structure on $\boldsymbol{B}$ by the definition: $a \bullet b := J(a)(b)$.

[53] $\boldsymbol{A}$ may be commutative or non-commutative, with our without unit.

[54] This viewpoint immediately suggests to discern between general embedding and those having special properties, e.g. the case where J is an injective mapping, i.e. for every $a \in \boldsymbol{A}$ there exists some $b \in \boldsymbol{B}$ such that $a \bullet b \neq 0$. Such modules are called *faithful* (or sometimes *true*) modules.

Without going into all necessary details let us recall that the associativity law

assocBM (128)
$$a_1 \bullet (a_2 \bullet b) = (a_1 \cdot a_2) \bullet b \quad \forall a_1, a_2 \in \boldsymbol{A}, b \in \boldsymbol{B}$$

is a immediate consequence of the homomorphism property of J: The left hand side $a_1 \bullet (a_2 \bullet b)$ translates into $J(a_1)[J(a_2)b]$, while the right hand side equals $J(a_1 \cdot a_2)(b)$. $\square$

ess-Banmod **Definition 28.** A Banach module $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ over some Banach algebra $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ is called *essential* if it coincides with the closed linear span of $\boldsymbol{A} \bullet \boldsymbol{B} = \{a \bullet b \,|\, a \in \boldsymbol{A}, b \in \boldsymbol{B}\}$. The same terminology applies to Banach ideals.

For a general Banach module $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ the closed linear span of $\boldsymbol{A} \bullet \boldsymbol{B}$ is denoted by $\boldsymbol{B}_e$ or $\boldsymbol{B}_{\boldsymbol{A}}$. It is called the *essential part* of the Banach module $\boldsymbol{B}$ (with respect to $\boldsymbol{A}$).

Despite the fact that the terminology obviously suggest this result it is something that should be proved (nevertheless we leave the easy proof to the reader):

essmodess1 **Lemma 34.** *For every Banach algebra with $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ with bounded approximate units the essential part of $\boldsymbol{B}$ is an essential submodule of $\boldsymbol{B}$, i.e. a Banach module over $\boldsymbol{A}$. In fact, it contains all other essential submodules over $\boldsymbol{A}$ within $\boldsymbol{B}$.*

5.1.1. *Essential Banach modules and BAIs.* The usual way to define the *essential part* $\boldsymbol{B}_e$ of a Banach module $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ with respect to some Banach algebra action $(\mathbf{a}, \mathbf{b}) \mapsto \mathbf{a} \bullet \mathbf{b}$ is the define it as the closed linear span of $\boldsymbol{A} \bullet \boldsymbol{B}$ within $\big(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}}\big)$. It is interesting that this subspace can be shown to have another nice characterization using BAIs (bounded approximate units) in $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$, if they exist:

sspartchar1 **Lemma 35.** *Assume that $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ is a Banach algebra with BAI $(\mathbf{e}_\alpha)_{\alpha \in I}$. Then*

haresspart1 (129)
$$\boldsymbol{B}_e = \{\mathbf{b} \in \boldsymbol{B} \,|\, \lim_\alpha \mathbf{e}_\alpha \bullet \mathbf{b} = \mathbf{b}\}.$$

In particular one has: Let $(\mathbf{e}_\alpha)_{\alpha \in I}$ and $(\mathbf{u}_b)_{\beta \in J}$ be two bounded approximate units (i.e. bounded nets within $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ acting in the limit almost like an identity in the Banach algebra $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$). Then one has ATTENTION COMES TWICE, SEE BELOW

BAIequiv2c (130)
$$\lim_\alpha \mathbf{e}_\alpha \bullet \mathbf{b} = \mathbf{b} \quad \Leftrightarrow \quad \lim_\beta \mathbf{u}_\beta \bullet \mathbf{b} = \mathbf{b}.$$

Although this relation follows immediately from the above characterization it is instructive to look at a direct proof of this equivalence.

Proof. Assume that $(\mathbf{e}_\alpha)$ acts *'properly'* on a given element $\mathbf{b} \in \boldsymbol{B}$, i.e. we assume that the left hand side of (130) is valid. We want to estimate the action of $\mathbf{u}_\beta$ (for "good indices" $\beta \succeq \beta_0$) on the same given element $\mathbf{b} \in \boldsymbol{B}$.

So for any given (but then fixed) $\varepsilon > 0$ we can fix beforehand and for convenience some constants: $\gamma := \varepsilon/(3\|\mathbf{b}\|_{\boldsymbol{B}})$, $C_0 = \sup_{\beta \in J} \|\mathbf{u}_\beta\|_{\boldsymbol{A}}$ and $C_1 := 1 + C_0$. We can then *fix*[55] some α_0 such that $\|\mathbf{b} - \mathbf{e}_{\alpha_0} \bullet \mathbf{b}\|_{\boldsymbol{B}} < \varepsilon/2C_1$.

Since $(\mathbf{u}_\beta)$ is a BAI for $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ we can find β_0 such that for $\beta \succeq \beta_0$ one has $\|\mathbf{u}_\beta \bullet \mathbf{e}_{\alpha_0} - \mathbf{e}_{\alpha_0}\| < \gamma$ and hence

$$\|\mathbf{u}_\beta \bullet \mathbf{e}_{\alpha_0} \bullet \mathbf{b} - \mathbf{e}_{\alpha_0} \bullet \mathbf{b}\|_{\boldsymbol{B}} \le \|\mathbf{u}_\beta \bullet \mathbf{e}_{\alpha_0} - \mathbf{e}_{\alpha_0}\|_{\boldsymbol{A}} \cdot \|\mathbf{b}\|_{\boldsymbol{B}} < \gamma \cdot \|\mathbf{b}\|_{\boldsymbol{B}} = \varepsilon/3.$$

But also $\mathbf{u}_\beta \bullet \mathbf{b}$ is not far away since

$$\|\mathbf{u}_\beta \bullet \mathbf{e}_{\alpha_0} \bullet \mathbf{b} - \mathbf{u}_\beta \bullet \mathbf{b}\|_{\boldsymbol{B}} \le \|\mathbf{u}_\beta\|_{\boldsymbol{A}} \cdot \|\mathbf{e}_{\alpha_0} \bullet \mathbf{b} - \mathbf{b}\|_{\boldsymbol{B}} \le C_0 \cdot \|\mathbf{e}_{\alpha_0} \bullet \mathbf{b} - \mathbf{b}\|_{\boldsymbol{B}}.$$

[55] For the *subsequent* estimates it will be important to fix α, despite the fact that we could take any α in the first upcoming step!

Putting all this together one obtains as an estimate for $\|\mathbf{b} - \mathbf{u}_\beta \bullet \mathbf{b}\|_{\boldsymbol{B}} \leq$

$$\leq \|\mathbf{b} - \mathbf{e}_{\alpha_0} \bullet \mathbf{b}\|_{\boldsymbol{B}} + \|\mathbf{e}_{\alpha_0} \bullet \mathbf{b} - \mathbf{u}_\beta \bullet \mathbf{e}_{\alpha_0} \bullet \mathbf{b}\|_{\boldsymbol{B}} + \|\mathbf{u}_\beta \bullet \mathbf{e}_{\alpha_0} \bullet \mathbf{b} - \mathbf{u}_\beta \bullet \mathbf{b}\|_{\boldsymbol{B}} \leq$$

$$\leq \varepsilon/3 + (1 + C_0)\|\mathbf{u}_\beta \bullet \mathbf{e}_{\alpha_0} \bullet \mathbf{b}\|_{\boldsymbol{B}} \leq \varepsilon/3 + \varepsilon/2 < \varepsilon, \quad \forall \beta \succeq \beta_0$$

as was requested, hence $\lim_\beta \|\mathbf{u}_\beta \bullet \mathbf{b} - \mathbf{b}\|_{\boldsymbol{B}} = 0$. □

The notion of "essential Banach modules" is of course trivial in case the Banach algebra $\boldsymbol{A}$ has a unit element which is mapped into the identity operator, i.e. if there exists $u \in \boldsymbol{A}$ such that $u \cdot a = a \, \forall a \in \boldsymbol{A}$ and also $u \bullet b = b$, $\forall b \in \boldsymbol{B}$. This may not follow formally from the definition but is in fact easily verified in all cases of interest. One might want to add it to the axioms in an abstract definition.

Lemma 36. *Let $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ be a Banach algebra with BAIs $(e_\alpha)_{\alpha \in I}$. Then a Banach module $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ is essential if and only if*

essmod-BAI (131)
$$\|e_\alpha \bullet b - b\|_{\boldsymbol{B}} \to 0 \quad \forall b \in \boldsymbol{B}$$

Hence relation (131) holds true for one such BAI if and only if it is true for every BAI in $\boldsymbol{A}$.

Proof. The argument for the equivalence of the usual definition and the "expected action" of BAIs is relatively easy[56] First of all it is clear that condition (131) above implies that even $\boldsymbol{A} \bullet \boldsymbol{B}$ is dense in $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$, even without taking linear combinations.

Assume on the other hand that linear combinations from $\boldsymbol{A} \bullet \boldsymbol{B}$ are dense in $\boldsymbol{B}$. Then we can choose α_0 as follows, for given $\varepsilon > 0$ and $f \in \boldsymbol{B}$.

First we find a finite linear combination $g = \sum_{k=1}^n c_k \mathbf{a}_k \bullet \mathbf{b}_k$ with $\|f - g\|_{\boldsymbol{B}} \leq \varepsilon/2$. Via simple renormalization of the elements $\mathbf{b}_k$ (moving the scaling factors over to $\mathbf{a}_k$) we can assume that $\|\mathbf{b}_k\|_{\boldsymbol{B}} = 1$ for $1 \leq k \leq n$. Then choose α_0 such that for $\alpha \succeq \alpha_0$ one has

$$\|e_\alpha \cdot \mathbf{a}_k - \mathbf{a}_k\|_{\boldsymbol{A}} < \frac{\varepsilon}{2n}$$

because this clearly implies

$$\|e_\alpha \cdot g - g\|_{\boldsymbol{A}} \leq \sum_{k=1}^n \|e_\alpha \cdot \mathbf{a}_k - \mathbf{a}_k\|_{\boldsymbol{A}} \|\mathbf{b}_k\|_{\boldsymbol{B}} < \varepsilon/2$$

which of course gives together

$$\|f - \mathbf{e}_\alpha \bullet f\|_{\boldsymbol{B}} \leq \|f - g\|_{\boldsymbol{B}} + \|g - \mathbf{e}_\alpha \cdot g\|_{\boldsymbol{A}} \cdot 1 < \varepsilon > 0,$$

thus completing the proof. □

As already mentioned earlier it is interesting to consider bounded subsets $M \subset \boldsymbol{B}$ with the property that the convergence in (131) is taking place in a uniform manner. Let us formalize this in a definition (which is not widely spread in the literature, maybe it is found in some paper that we are not aware of).

unifapprdef **Definition 29.** A bounded subset $M \subset \boldsymbol{B}$ within an (essential) Banach module $\boldsymbol{B}$ over a Banach algebra $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ with bounded approximate units is called *uniformly approximable* of the following situation occurs: For every bounded approximate unit (e_α) in $\boldsymbol{A}$ one has: Given $\varepsilon > 0$ one can find α_0 such that for $\alpha \succeq \alpha_0$ one has

nifapprdef2 (132)
$$\|e_\alpha \bullet b - b\|_{\boldsymbol{B}} \leq \varepsilon, \quad \forall f \in M.$$

[56]It does not work so for unbounded approximate identities which may occur in same cases.

We have chosen the "practical version" of the definition, instead of the axiomatic/minimalistic, which would be to assume the property (132) only for one (out of many possible) approximate units. As we show next that would not make a difference (and since the above choice is more useful in practice we have given that, and thus the subsequent lemma just provides a sufficient condition).

onetoallBAI **Lemma 37.** *Assume that in the situation described in Def. 29 the condition (132) is satisfied for only one specific BAI (e_α). Then it is also valid for any other approximate unit $(u_\beta)_{\beta \in J}$.*

Proof. The proof given here is very similar Lemma 89.

Assume that the condition of uniform approximability is satisfied and $(u_\beta)_{\beta \in J}$ is some other BAI for $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$. Since M is bounded we can fix $C := \sup_{m \in M} \|b\|_{\boldsymbol{B}}$.

Given $\varepsilon > 0$ we can thus find α_0 such that

$$\|e_\alpha \bullet b - b\|_{\boldsymbol{B}} \leq \varepsilon/2 \quad \forall \alpha \succeq \alpha_0.$$

Choosing β_0 suitably one has

$$\|u_\beta \bullet e_{\alpha_0} - e_{\alpha_0}\|_{\boldsymbol{A}} \leq \varepsilon/(2C)$$

for all $\beta \succeq \beta_0$, and consequently

$$\|u_\beta \bullet e_{\alpha_0} \bullet b - e_{\alpha_0} \bullet b\|_{\boldsymbol{B}} < \varepsilon/2 \quad \forall b \in M.$$

□

dual-alg **Definition 30.** For any Banach module $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ over the Banach algebra $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ the dual space$(\boldsymbol{B}', \|\cdot\|_{\boldsymbol{B}'})$ can be turned naturally into a $\boldsymbol{A}$-Banach module via the action

dual-alg1 (133) $$[\mathbf{a} \bullet' \sigma](\mathbf{b}) := \sigma(\mathbf{a} \bullet \mathbf{b}), \quad \forall \mathbf{a} \in \boldsymbol{A}, \mathbf{b} \in \boldsymbol{B}, \sigma \in \boldsymbol{B}'.$$

Note that the dual space of a *left* Banach module is a *right* Banach module (over the same Banach algebra $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$), which means that one has the "right" associativity law:

rightass1 (134) $$a_2 \bullet' (a_1 \bullet' \sigma) = (a_1 \cdot a_2) \bullet' \sigma.$$

Essentially the idea behind this definition is the observation that for any Banach algebra $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ of operators on $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ the collection

$$\{T' \,|\, T(\mathbf{b}) = \mathbf{a} \bullet \mathbf{b}, \mathbf{a} \in \boldsymbol{A}\}$$

of transposed operators is now a Banach algebra of operators on $(\boldsymbol{B}', \|\cdot\|_{\boldsymbol{B}'})$. Clearly the "right associativity" results from the usual observation that $(T_1 \circ T_2)' = T_2' \circ T_1'$.

For this reason one may prefer an alternative way of writing the "transpose action", by using the symbol $\sigma \bullet \mathbf{a}$. This has the advantage that the convention (133) can be rewritten as

dual-algR (135) $$(\sigma \bullet' \mathbf{a})(\mathbf{b}) = \sigma(\mathbf{a} \bullet \mathbf{b}), \quad \mathbf{a} \in \boldsymbol{A}, \mathbf{b} \in \boldsymbol{B}.$$

This also results in a more natural form of the "right associativity", namely

rightass (136) $$(\sigma \bullet' \mathbf{a}1) \bullet' \mathbf{a}_2 = \sigma \bullet' (\mathbf{a}_1 \cdot \mathbf{a}_2), \quad \mathbf{a}_1, \mathbf{a}_2 \in \boldsymbol{A}.$$

In this way it is always possible to turn the dual space $(\boldsymbol{A}', \|\cdot\|_{\boldsymbol{A}'})$ of a given Banach algebra $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ into a (right or left) Banach module over $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$.

Of course the change of order is irrelevant in the case of commutative Banach algebras, but it has to be taken into account for the general case. Maybe it appears as more natural to write the action on the dual space from the right. Then we have the following alternative definition:

Most likely this first version will be eliminated later, because:

dual-algalt **Definition 31.** For any Banach module $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ over the Banach algebra $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ the dual space$(\boldsymbol{B}', \|\cdot\|_{\boldsymbol{B}'})$ can be turned naturally into a $\boldsymbol{A}$-Banach module via the action

ual-algalt2
$$[\sigma \bullet' \mathbf{a}](\mathbf{b}) := \sigma(\mathbf{a} \bullet \mathbf{b}), \quad \forall \mathbf{a} \in \boldsymbol{A}, \mathbf{b} \in \boldsymbol{B}, \sigma \in \boldsymbol{B}'. \tag{137}$$

There is also an obvious compatibility of the original action and the adjoint action, described by the following identity

leftright1
$$(\sigma \bullet' \mathbf{c})(\mathbf{a} \bullet \mathbf{b}) = \sigma((\mathbf{c} \cdot \mathbf{a}) \bullet \mathbf{b}) = [\sigma \bullet' (\mathbf{c} \cdot \mathbf{a})](\mathbf{b}) \tag{138}$$

Remark 29. This way of writing is most natural if we consider the situation familiar from linear algebra! Choose $\boldsymbol{A} = \mathcal{M}_{n,n}(\mathbb{R})$, the Banach algebra of $n \times n$-matrices over $\mathbb{R}$, acting on $\boldsymbol{B} = \mathbb{R}^n$. We know that in this case the dual space is, by definition the space $\mathcal{L}\mathbb{R}^n, \mathbb{R}$ which - by general linear algebra reasoning - can be identified with the set of $1 \times n$-matrices, which we would normally view as $\mathbb{R}^n$, but considered as the space of real row-vectors. The $\bullet$-operation of $\mathbf{a} \in \mathcal{M}_{n,n}(\mathbb{R})$ is then in both cases through matrix multiplication.

While the formula (31) is a definition it is easy to check that (due to the associativity law for matrix multiplication the dual action of $\mathbf{a}$ on a row vector $\sigma \in \boldsymbol{B}'$ is just through (right) matrix multiplication, since one has

$$(\sigma * \mathbf{a}) * \mathbf{b} = \sigma * (\mathbf{a} * \mathbf{b}), \quad \mathbf{a} \in \mathcal{M}_{n,n}(\mathbb{R}), \mathbf{b} \in \mathbb{R}^n.$$

The “natural” interpretation of the action as described in the first variant (133) would be to identify $\mathcal{L}\mathbb{R}^n, \mathbb{R}$ again with $\mathbb{R}^n$, viewed now as a system of column vectors (via the transposition mapping, or alternatively by choosing the unit vectors as a basis) and then have $\mathcal{M}_{n,n}(\mathbb{R})$ act on those column vectors by matrix multiplication with the *transpose matrix* from the right. Of course, these two viewpoints describe (isometrically) isomorphic (right) Banach module action, thanks to the well known linear algebra formula

$$(\mathbf{h} * \boldsymbol{C})^t = \boldsymbol{C}^t * \mathbf{h}^t, \quad \mathbf{h} \in \mathbb{R}^n, \boldsymbol{C} \in \mathcal{M}_{n,n}(\mathbb{R}).$$

DiracMult **Lemma 38.** *(Ex. to Def. (30)) Multiplication of a Dirac measure is realized as scalar multiplication of this Dirac measure by the point value of the continuous function h at that point:*

Diracmult1
$$h \cdot \delta_x = \delta_x \cdot h = h(x)\delta_x, \quad h \in \boldsymbol{C}_b. \tag{139}$$

For *convenience* we will write $h \cdot \mu$ instead of the abstract symbol $\bullet$ in order to indicate pointwise multiplication between functionals μ and functions $h \in \boldsymbol{C}_0(\mathbb{R}^d)$.

bRdessCOMod **Theorem 10.** *The Banach space $(\boldsymbol{M}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}})$ is an essential Banach module over $\boldsymbol{C}_0(\mathbb{R}^d)$ with respect to the natural (dual) action of pointwise multiplication.*

Proof. In fact we will show much more: for any BUPU $\Phi = (\phi_i)$ one has

$$\mu = \sum_{i \in I} \mu \cdot \phi_i$$

as absolutely convergent sum in $(\boldsymbol{M}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}})$, hence

artsumconv1
$$\|\mu - \mu \cdot \sum_{i \in F} \phi_i\|_{\boldsymbol{M}} \to 0 \quad \text{for} \quad F \text{ finite}, F \uparrow I. \tag{140}$$

□

Psipartsum1 **Definition 32.** It will be convenient to use for any set $J \subseteq I$ the symbol ψ_J for the function $\psi_J := \sum_{i \in J} \psi_i$. For the choice $F = F(i) = \{j | \psi_j \cdot \psi_i \neq 0\}$ we write ψ_i^* for $\psi_{F(i)}$.

A very useful relation which will be used later on is:

platrepr (141)
$$\psi_i \cdot \psi_i^* = \psi_i \quad \text{for all} \quad i \in I.$$

Relation (140) above is thus equivalent to

$$\lim_{F \to I} \|\mu \cdot \psi_F - \mu\|_{\boldsymbol{M}} = 0$$

which of course translates into the more explicit variant:

Given a BUPU Ψ, we can find for any $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ and $\varepsilon > 0$ a finite set $F_0 \subset I$ such that for $F \supset F_0$ on has:

$$\|\mu - \mu \cdot \psi_F\|_{\boldsymbol{M}} < \varepsilon.$$

Another useful observation is this one:

measactloc **Lemma 39.** *Let $\mu \in \boldsymbol{M}(\mathbb{R}^d)$ and $\varepsilon > 0$ be given. Then there exists $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ with $\|k\|_\infty \le 1$ such that $\mu(k) \ge 0$ and $\mu(k) \ge \|\mu\|_{\boldsymbol{M}}(1 - \varepsilon)$.*

Proof. Using the definition of the functional norm and the density of $\boldsymbol{C}_c(\mathbb{R}^d)$ in $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$, one only has to take into account phase factors and multiply, if necessary, k by $|\mu(k)|/\mu(k)$, which does not change $\|k\|_\infty$[57]. Then it is guaranteed that $\mu(k)$ is a non-negative real number. □

eascompsupp **Corollary 3.** *Every $\mu \in \boldsymbol{M}(\mathbb{R}^d)$ is the limit of "compactly supported" measures of the form $(\psi_F \cdot \mu)$, where F is running through the finite subsets of I.*

We will derive therefrom that one can also "integrate" arbitrary elements $h \in \boldsymbol{C}_b(\mathbb{R}^d)$.

integrCbRd1 **Lemma 40. Integration of bounded functions against bounded measures**
For any $h \in \boldsymbol{C}_b(\mathbb{R}^d)$ and $\mu \in \boldsymbol{M}(\mathbb{R}^d)$ the net $(p_K \cdot \mu)(h) = \mu(p_K \cdot h)$ is a Cauchy-net in $\mathbb{C}$. Therefore it makes sense to define $\mu(h) = \lim_K \mu(p_K \cdot h)$. It is clear that in this way μ extends in a unique way to a bounded linear functional on $\big(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\big)$, and that the norm of this extension equals $\|\mu\|_{\boldsymbol{M}}$.

The above statement does of course not imply that the dual space of $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$ and that of $\big(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\big)$ are equal, because functionals on $\boldsymbol{C}_b(\mathbb{R}^d)$ which are equal on $\boldsymbol{C}_0(\mathbb{R}^d)$ need not be equal on $\boldsymbol{C}_b(\mathbb{R}^d)$.

`Alternative formulation`

integrCbRd2 **Lemma 41.** *Let $(p_\beta)_{\beta \in J}$ be any bounded approximate identity for the (pointwise) Banach algebra $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$. Then for any $h \in \boldsymbol{C}_b(\mathbb{R}^d)$ and $\mu \in \boldsymbol{M}(\mathbb{R}^d)$ the net $\mu(p_\beta \cdot h)$ is a Cauchy-net in $\mathbb{C}$. Therefore it makes sense to define $\mu(h) = \lim_K \mu(p_\beta \cdot h)$. It is clear that in this way μ extends in a unique way to a bounded linear functional on $\big(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\big)$, i.e. the extension does not depend on the specific BAI (p_β), and the norm of this extension (still) equals $\|\mu\|_{\boldsymbol{M}}$. Clearly it cannot be smaller, but also the norm is preserved by all the approximating functions $p_\beta \in \boldsymbol{C}_0(\mathbb{R}^d)$.*

Remark 30. For any bounded set of convolution operators (resp. translation invariant systems) the image of a (bounded and) equicontinuous set is again equicontinuous. This is clear since $T_y(Tf) - Tf = T(T_y f - f)$, which tends to zero uniformly if f is chosen from an equicontinuous set.

$$\|T_y(Tf) - Tf\| \le \|T\| \|(T_y f - f)\|_\infty \le C \|T_y f - f\|_\infty \to 0.$$

[57]See the embedding of $(\boldsymbol{C}_c(\mathbb{R}^d), \|\cdot\|_1)$ into $\boldsymbol{M}_b$ given below for more details.

Remark 31. If the net of bounded measures is (bounded and) tight, then it is an exercise to show, that it is *vaguely convergent* [58] if and only if it is w^*-convergent, resp. if and only if $\mu_\alpha(h) \to \mu_0(h)\ \forall h \in \boldsymbol{C}_b(\mathbb{R}^d)$. It is a bit more work (but still an exercise) to check out that this "pointwise convergence" actually takes place *uniformly over compact subsets* of $\boldsymbol{C}_b(\mathbb{R}^d)$. Since obviously the collection $\{\chi_s, s \in K\}$ is a compact subset of $\boldsymbol{C}_b(\mathbb{R}^d)$ for any compact subset $K \subset \mathbb{R}^d)$ (as the continuous image of a compact set) this implies that in this case one obtains uniform convergence of the Fourier (Stieltjes) transform of measures over compact sets.

It may be a good exercise to consider the question, whether (resp. why probably not) the same statement fails to be true for non-tight sets. What about δ_{x_n} with $|x_n| \to \infty$.

Remark 32. Recall that the "mass-preserving" stretching/compression operator St_ρ can be extended to $\boldsymbol{M}_b(\mathbb{R}^d)$ by the definition **[Rd-specific!]**

Strhodef4
$$\mathrm{St}_\rho\, \mu(f) := \mu(\mathrm{D}_\rho f). \tag{142}$$

Check that one has $\mathrm{St}_\rho\, \delta_x = \delta_{\rho x}$, and that $lim_{\rho \to 0} \mathrm{St}_\rho\, \mu = \mu(1)\delta_0$. In fact, for each $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ one has: $\mathrm{D}_\rho f$ is uniformly bounded with respect the sup-norm for $\rho \to 0$ one has: $f(x) \to f(0)$, uniformly over compact sets. Since we can approximate the measure μ (by localizing it) to a measure with compact support its action is defined on sufficiently large compact sets, where $\mathrm{D}_\rho f$ is like the constant function $f(0) = \delta_0(f)$, while the action on $Const \equiv 1$ is denoted by $\mu(1)$.

One can use this fact to find out that it is even possible to extend the convolution operators $f \mapsto C_\mu f$ to all of $\boldsymbol{C}_b(\mathbb{R}^d)$ (still with the equality of operator norm on $\big(\boldsymbol{C}_b(\mathbb{R}^d),\, \|\cdot\|_\infty\big)$ with the functional norm of μ), and with the property that the operators arising in such a way commute with translations. Applying the Hahn-Banach theorem one can construct translation invariant *means* on $\big(\boldsymbol{C}_b(\mathbb{R}^d),\, \|\cdot\|_\infty\big)$ which in turn allow to construct bounded linear operators on $\boldsymbol{C}_b(\mathbb{R}^d)$ which commute with all the translation operators without being of the form of a convolution by some bounded measure. In fact, those operators are non-zero operators on $\boldsymbol{C}_b(\mathbb{R}^d)$, but they map all of $\boldsymbol{C}_0(\mathbb{R}^d)$ onto the zero function. It is also not much more than a simple exercise to find out (using the characterization of $\boldsymbol{C}_{ub}(\mathbb{R}^d)$ within $\boldsymbol{C}_b(\mathbb{R}^d)$ given early on) to check that any operator on $\boldsymbol{C}_b(\mathbb{R}^d)$ commuting with translations will map $\boldsymbol{C}_{ub}(\mathbb{R}^d)$ into itself (in fact, this argument was used at the beginning of the identification theorem.)

Dual groups and characters

We are now in the position to define the **Fourier transform** of a bounded measure. In the classical literature this is often referred to as the Fourier-Stieltjes transform of a measure, because it can be carried out technically over $\mathbb{R}$ using Riemann-Stieltjes integrals. Such a R-St-integral is the difference of two R-St-integrals with respect to bounded, non-decreasing "distribution" functions F. So in such a definition the ordinary Riemannian sum is replaced by a sum of the same form, but instead of the "natural length" of the interval $[a, b]$, which is $|b - a|$, one uses the length in the sense of F which is $F(b) - F(a)$.

aracterdef1 **Definition 33.** A *character* on a group G is a continuous homomorphism from a topological group G into the torus group $\mathbb{T} = \{z \in \mathbb{C}\,|\,|z| = 1\}$. In other words, χ is a character if

chardef0
$$\chi(x + y) = \chi(x) \cdot \chi(y) \quad \forall x, y \in G. \tag{143}$$

[58] This means that it is convergent in the $\sigma(\boldsymbol{M}_b(\mathbb{R}^d), \boldsymbol{C}_c(\mathbb{R}^d))$-topology, resp. $\mu_\alpha(k) \to \mu_0(k) \quad \forall k \in \boldsymbol{C}_c(\mathbb{R}^d)$.

dualgroup **Definition 34.** The set of all character is called the *dual group*, because those *characters* form an Abelian group under pointwise multiplication (!Exercise!).

The *neutral element* in this group is of course the *trivial* character $\chi_0(t) \equiv 1$. Consequently, since $|\chi(x)| = 1$ one has $\overline{\chi(x)} = 1/\chi(x)$ for all $x \in G$, as inverse element.

We write $\widehat{G}$ for the dual group [59] corresponding to G.

Remark 33. In the engineering context these characters are named "pure frequencies" and for the image processing context ($\mathbb{R}^2$) one talks about "plane waves".

charRd **Theorem 11.** *In the case of $(\mathbb{R}^d, +)$ the dual group consists of the characters χ of the form $\chi_s : t \mapsto exp(2\pi i s \cdot t)$, with $s \in \mathbb{R}^d$. Due to the exponential law the pointwise multiplication of characters turns into addition of the parameters s (describing the "frequency content" of χ_s).*

Moreover, one can show that the convergence of a sequence s_n in $\mathbb{R}^d$ is equivalent to the uniform convergence over compact sets, i.e. the metric topology on $\mathbb{R}^d$ and the compact-open topology on $\widehat{\mathbb{R}^d}$ (of characters) turn this identification into an isomorphism of (locally compact) topological groups.

Proof. We omit the proof here, indicating that it is found in various books on the field (e.g. [23]). It is not spelled out in Reiter's book [79] or Rudin's classical book [87], if I remember correctly [to be verified!]. In any case it can be reduced (by induction) to the case $d = 1$.

Typically one verifies first that the continuity (in fact even measurability is sufficient!) of a given character χ implies in fact continuous differentiability and from that the validity of a simple ordinary differential equation. From this it can be concluded that $\chi(t) = exp(\alpha t)$, for some $\alpha \in \mathbb{C}$. But since χ also has to be bounded the exponent has to be purely imaginary! $\square$

Remark 34. We have two remarkable properties, all resulting from the well-known (and all-important) exponential law:

expo-law
$$e^{z_1} \cdot e^{z_2} = e^{z_1 + z_2}, \quad z_1, z_2 \in \mathbb{C}. \tag{144}$$

which on the one hand (obviously) gives

expo-law1
$$\chi_s(t_1 + t_2) = \chi_s(t_1) \cdot \chi_s(t_2), \quad t_1, t_2 \in \mathbb{R}, \tag{145}$$

but also on the other hand

expo-law2
$$\chi_{s_1}(t) \cdot \chi_{s_2}(t) = \chi_{s_1+s_2}(t), \quad \text{resp.} \quad \chi_{s_1} \cdot \chi_{s_2} = \chi_{s_1+s_2}. \tag{146}$$

which corresponds to the fact that each function χ_s defines a character, while on the other hand (146) implies that the abstract multiplication in $\hat{G}$ is isomorphic to the usual group operation (namely addition) in $(\mathbb{R}, +)$. Corresponding laws are valid for $\mathbb{R}^d$ by induction, but also in the context of general LCA groups!

urStieltDef **Definition 35.** The Fourier transform[60] of $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ is defined by

meas-FT-def
$$\hat{\mu}(s) = \mu(\chi_{-s}) = \mu(\overline{\chi_s}). \tag{147}$$

Remark 35. Later on, with the help of distributions (specifically the Banach Gelfand triple $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$) we will show that this is a special case of the more general approach. This in turn can be further extended to the setting developed by L. Schwartz and his theory of *tempered distributions*

[59] Often the group law written as addition in order to emphasize that it is a (obviously) a *commutative* group!

[60] Using the classical terminology, e.g. of Bochner, one often talks about the Fourier-Stieltjes transform of the measure. This has to do with the classical view-point, which does not talk about abstract measures but uses instead the corresponding distribution function $F_\mu(z) = \int_{-\infty}^{z} d\mu(z)$ and realized integrals using Riemann-Stieltjes sums and their limits, cf. 47 above.

[Rd-specific!]

There is also a nice compatibility between pure frequencies and dilations. It will be more convenient to use the dilation using the D_ρ-operators, since they move pure frequencies into new frequencies because this dilation operator does NOT change the values of a function (all in the unit circle).

dilpurfr1 **Lemma 42.**

$$\mathrm{D}_\rho \chi_s = \chi_{s/\rho}, \quad \rho \neq 0, s \in \mathbb{R}^d.$$

The proof is left to the readers as an easy exercise!

Note that $\chi_s \in \boldsymbol{C}_{ub}(\mathbb{R}^d)$ and that $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ can be applied according to Lemma 40.

The argument is based on the following general result:

onvmeasfcts **Proposition 3.** *Assume that a (bounded) and tight net (μ_α) is w^*-convergent to μ_0 and that (h_β) is a bounded net in $\big(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\big)$ which is uniformly convergent to h_0 over compact subsets. Then for every $\varepsilon > 0$ there exist indices α_0 and β_0 such that for every $\alpha \geq \alpha_0$ and $\beta \geq \beta_0$*

convmeas2
$$|\mu_\alpha(h_\beta) - \mu_0(h_0)| \leq \varepsilon. \tag{148}$$

which is expressed equally by the formulation

$$\lim_{\alpha,\beta} \mu_\alpha(h_\beta) = \mu_0(h_0).$$

Proof. Given $\varepsilon > 0$ we recall that $\|h_\beta\|_\infty \leq C_0$ and $\|\mu_\alpha\|_{\boldsymbol{M}} \leq C_1$ for all α. By the tightness of the net (μ_α) we can find some plateau-like function $p \in \boldsymbol{C}_c(\mathbb{R}^d)$ with $\|p\|_\infty \leq 1$ such that $\|\mu_\alpha - p \cdot p \cdot \mu_\alpha\|_{\boldsymbol{M}} < \varepsilon C_0{}^{-1}$ for all α. Writing $K := \mathrm{supp}(p)$ we can find some index β_0 such that $|h_\beta(x) - h_0(x)| \leq \varepsilon / C_1$ for all $x \in K$ and all $\beta \geq \beta_0$. Altogether we then have

$$|\mu_\alpha(h_\beta) - \mu_0(h_0)| \leq$$
$$|\mu_\alpha(h_\beta) - p \cdot \mu_\alpha(h_\beta)| + |\mu_\alpha(p \cdot h_\beta) - \mu_\alpha(p \cdot h_0)| + |\mu_\alpha(p \cdot h_0) - \mu_0(p \cdot h_0)| + |p \cdot \mu_0(h_0) - \mu_0(h_0)| \leq$$
$$|[\mu_\alpha - p \cdot \mu_\alpha](h_\beta)| + |\mu_\alpha(p \cdot (h_\beta - h_0))| + |[\mu_\alpha - \mu_0](p \cdot h_0)| + |[p \cdot \mu_0 - \mu_0](h_0)| \leq$$
$$\|\mu_\alpha - p \cdot \mu_\alpha\|_{\boldsymbol{M}} \|h_\beta\|_\infty + \|\mu_\alpha\|_{\boldsymbol{M}} \|p \cdot (h_\beta - h_0)\|_\infty + |[\mu_\alpha - \mu_0](p \cdot h_0)| + \|p \cdot \mu_0 - \mu_0\|_{\boldsymbol{M}} \|h_0\|_\infty \leq$$
$$\varepsilon \cdot C_0 + C_1 \cdot \|p \cdot (h_\beta - h_0)\|_\infty + |[\mu_\alpha - \mu_0](p \cdot h_0)| + \varepsilon C_0 \leq C_2 \cdot \varepsilon.$$

for some constant $C_2 > 0$, and for all $\alpha \succ \alpha_0, \beta \succ \beta_0$. WELL: WE HAVE TO CHECK THE DETAILS Note that the w^*-convergence of the net (μ_α) implies that $\|\mu_0\|_{\boldsymbol{M}} \leq \liminf_\alpha \|\mu_\alpha\|_{\boldsymbol{M}}$. The first and last part is due to the tightness and boundedness, independent from the , while the is due to uniform convergence of (h_β) over $K = \mathrm{supp}(p)$ and the third term is getting small due to the w^*-convergence of μ_α to μ_0. □

FTDpsequi **Corollary 4.** *Given $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$. Then the family $\{\widehat{D_\Psi \mu}, |\Psi| \leq 1\}$ is equicontinuous, i.e. for all $\varepsilon > 0$ there exists $\delta > 0$ such that*

ererer
$$|\widehat{D_\Psi \mu}(s) - \widehat{D_\Psi \mu}(s-h)| \leq \varepsilon \quad \textit{for all} \quad |h| \leq \delta, \quad |\Psi| \leq 1. \tag{149}$$

: THIS IS A GENERAL FACT: TIGHTNESS IN THE TIME-DOMAIN IMPLIES EQUICONTINUITY IN THE FOURIER DOMAIN (AND IN FACT VICE VERSA)!

iemStiLemma **Proposition 4** (Riemann-Stieltjes transform properties)**.** *The Fourier transform is a linear and non-expansive mapping from $\boldsymbol{M}_b(\mathbb{R}^d)$ into $\big(\boldsymbol{C}_{ub}(\mathbb{R}^d), \|\cdot\|_\infty\big)$*[61]*.*

[61] In the same way as the Fourier transform from $\boldsymbol{L}^1(\mathbb{R}^d)$ into $\boldsymbol{C}_0(\mathbb{R}^d)$ is injective, but not surjective we also have here (for "local reasons") that the mapping is injective but not surjective. Only for non-negative measures, i.e. if $f \geq 0$ implies $\mu(f) \geq 0$ then we have $\hat{\mu}(0) = \mu(1) = \|\mu\|_{\boldsymbol{M}}$ and thus $\|\hat{\mu}\|_\infty = \|\mu\|_{\boldsymbol{M}}$ in this case. This applies in particular to probability measures, with $\|\hat{\mu}\|_\infty = 1 = \|\mu\|_{\boldsymbol{M}}$

Proof. The uniform continuity results follows from the essential concentration of bounded measures over compact sets, the "usual rule" $T_s\mathcal{F}(\mu) = \mathcal{F}(M_s\mu)$, and the properties of characters.

In fact, given $\varepsilon > 0$, and $\chi_0 \in \hat{G}$, we can find a compact subset $Q \subseteq G$ such that $\|\mu - \psi_Q\mu\|_{\boldsymbol{M}} < \varepsilon/3$. Hence there is a neighborhood W of the identity in $\widehat{G}$ such that for all $\chi \in \chi_0 + W$ one has $|\chi(y) - \chi_0(y)| = |\chi \cdot \overline{\chi_0}(y) - 1| < \varepsilon/6$ for all $y \in Q$, by the definition of neighborhoods in $\hat{G}$ (compact open topology), and the fact that the quotient $\chi/\chi_0 (= \chi \cdot \overline{\chi_0})$ belongs to W. Expressed differently we have $\|\psi_Q(\chi_0 - \chi)\|_\infty < \varepsilon/6$, and thus

$$|[\mathcal{F}(\mu\psi_Q)](\chi_0 - \chi)| \leq 2|\mu(\psi_Q\overline{(\chi_0 - \chi)})|.$$

Altogether we have

$$|\hat{\mu}(\chi_0) - \hat{\mu}(\chi)| \leq |\mathcal{F}(\mu - \mu\psi_Q)(\chi_0)| + |\mathcal{F}(\mu\psi_Q)(\chi_0 - \chi)| + |\mathcal{F}(\mu\psi_Q - \mu)(\chi)|$$

and consequently

$$|\hat{\mu}(\chi_0) - \hat{\mu}(\chi)| \leq 2\|\mu - \mu\psi_Q\|_{\boldsymbol{M}} + \varepsilon/3 \leq \varepsilon.$$

□

In general very little can be said about the behaviour of the Fourier transforms of w^*-convergent nets or sequences. Obviously $w^*\text{-}\lim_{n\to\infty}\delta_{-n} = 0$, but $\mathcal{F}(\delta_{-n}) = \chi_n$, which is not even pointwise convergent.

The situation is quite different for w^*-convergence (in $\boldsymbol{M}(\mathbb{R}^d)$) combined with tightness:

wsttoFOUR1 **Lemma 43.** *Let (μ_α) be a w^*-convergent and tight net in $\boldsymbol{M}_b(\mathbb{R}^d)$, with $\mu_0 = w^* - lim_\alpha \mu_\alpha$. Then we have $\widehat{\mu_\alpha}(s) \to \widehat{\mu_0}(s)$, uniformly over compact subsets of $\widehat{G}$.*

Proof. Pointwise convergence of the Fourier transform is a consequence of the fact that due to the tightness of the whole family the action can be reduced to the case of measures with compact joint support. SHOULD BE MORE EXPLICIT!. □

DpsiFourCO1 **Corollary 5.** *Consider the family $\mathrm{D}_\Psi\mu$ with $|\Psi| \leq 1$. Then their Fourier (Stieltjes) transforms are uniformly convergent to $\hat{\mu}$ over compact sets of $\mathbb{R}^d$.*

The starting point of our further development is a simple re-interpretation of the value of the exponential function in two natural ways:

deltFdelt
$$\delta_y(\widehat{\delta_x}) = e^{-2\pi i x\cdot y} = \delta_x(\widehat{\delta_y}). \tag{150}$$

The formula (150) immediately implies that for any two BUPUs Φ and Ψ one has, due to the absolute convergence of the involved sequences on both sides:

measFmeas3
$$D_\Psi\mu(\widehat{D_\Phi\nu}) = D_\Phi\nu(\widehat{D_\Psi\mu}) \quad \text{for all} \quad \mu,\nu \in \boldsymbol{M}_b(\mathbb{R}^d) \text{ and any BUPUs } \Phi,\Psi. \tag{151}$$

By linearity and then passing to the limit (using Prop. 3 below) we can derive therefrom the following important **Fundamental Formula for the Fourier Stieltjes Transform**:

measFmeas
$$\nu(\hat{\mu}) = \mu(\hat{\nu}) \quad \text{for any pair} \quad \mu,\nu \in \boldsymbol{M}_b(\mathbb{R}^d). \tag{152}$$

For integrable functions this relation turns into what H. Reiter has called in his courses in Vienna in the early 1970th the
Fundamental Relation for the Fourier Transform ((FRF) for $\boldsymbol{L}^1(\mathbb{R}^d)$):

L1FourL1
$$\int_{\mathbb{R}^d} f(t)\hat{g}(t)dt = \int_{\mathbb{R}^d} g(s)\hat{f}(s)ds, \quad \text{for all} \quad f,g \in \boldsymbol{L}^1(\mathbb{R}^d). \tag{153}$$

Of course it can be verified directly, using Fubini's theorem, since $(t,s) \to f(t)g(s) \in \boldsymbol{L}^1(\mathbb{R}^{2d})$.

As an immediate consequence of identity (152) we can derive the uniqueness theorem for Fourier Stieltjes transforms. If $\hat{\mu}$ is the zero-function, we can apply $\nu = \mu_g$, for any possible

$g \in \boldsymbol{C}_c(\mathbb{R}^d)$ (or $\boldsymbol{L}^1(\mathbb{R}^d)$). Hence $\mu(\hat{g}) = 0$ for every $\hat{g} \in \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$. Since $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ is dense in $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$ with respect to the $\|\cdot\|_1$-norm this implies that μ is the zero functional.

Note that an easy way to show that $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ is dense in $\boldsymbol{C}_0(\mathbb{R}^d)$ is based on the following sequence of observations:

(1) $\mathcal{F}\boldsymbol{L}^1(\mathbb{R})$ contains the triangular function Δ, which can be viewed as B-spline of order 1 (piecewise linear function) and whose Fourier transform is the (non-negative and even) function $sinc^2$;

(2) the shifted triangular functions $T_n\Delta$ forms a BUPU (bounded in $\big(\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_{\mathcal{F}\boldsymbol{L}^1}\big)$, since translation corresponds to modulation in $\big(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\big)$ and this is isometric)

(3) for $d > 1$ one obtains BUPUs by looking at tensor products of $1D$-Δ-functions. Moreover one has $\|f \otimes g\|_1 = \|f\|_1\|g\|_1$ and

tensorFour
$$\mathcal{F}(f \otimes g) = \mathcal{F}(f) \otimes \mathcal{F}(g), \tag{154}$$

based on the following convention (using the *tensor product symbol*):

tensordef0
$$f \otimes g(x, y) := f(x) \cdot g(y), \quad x, y \in \mathbb{R}^d, \tag{155}$$

(4) Since $\big(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\big)$ is isometrically dilation invariant under $\mathrm{St}_\rho, \rho > 0$ we find that $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ is isometrically dilation invariant under the family $\mathrm{D}_\rho, \rho > 0$, hence one obtains arbitrary fine BUPUs of the form $\mathrm{D}_\rho T_n\Delta, n \in \mathbb{Z}^d, \rho > 0$;

(5) for $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ one has convergence of the corresponding spline-type approximations with respect to the $\|\cdot\|_\infty$-norm, i.e. $\|Sp_{\rho\Delta}k - k\|_\infty \to 0$.

Using the last lemma we can derive the formula (152) from (150) via the following relation: *Fundamental Relation for the Fourier-Stieltjes transform*:

measFmeas2
$$\mu(\hat{\nu}) = \nu(\hat{\mu}) \quad \text{for} \quad \mu, \nu \in \boldsymbol{M}_b(\mathbb{R}^d). \tag{156}$$

then follows using exactly Proposition 3 and Lemma 43.

convFTmeas
Theorem 12. *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 crisp formulas which are included in this verbal statement read as follows:

convFT00
$$\mathcal{F}(\mu_1 * \mu_2) = \mathcal{F}(\mu_1) \cdot \mathcal{F}(\mu_2), \qquad \mu_1, \mu_2 \in \boldsymbol{M}_b(\mathbb{R}^d), \tag{157}$$

or in the more usual form:

convFThat
$$\widehat{\mu_1 * \mu_2} = \widehat{\mu_1} \cdot \widehat{\mu_2}, \tag{158}$$

The compatibility with convolution is an easy exercise for discrete measures, and can be transferred to the general case using a weak-star argument. Recall again that w^*-convergence of bounded nets of measures implies pointwise convergence of their Fourier transforms.

Recall that a net (or just sequence) of measures $(\mu_\alpha)_{\alpha \in I}$ is w^*-convergent to μ_0 if and only if one has: for any $f \in \boldsymbol{C}_0$ and $\varepsilon > 0$ one can find some index α_0 such that $\alpha \succeq \alpha_0$ implies:

$$|\mu_\alpha(f) - \mu_0(f)| \leq \varepsilon.$$

There is an alternative way of proving commutativity of convolution. It is easy to see that the convolution of (finite) discrete measures is commutative, and the general case follows from this (by approximation in the strong operator topology).

We still have to give the *ordinary Fourier inversion formula using summability kernels.* This Fourier inversion is a direct consequence of the
Fundamental Relationship for the Fourier transform of $\boldsymbol{L}^1(\mathbb{R}^d)$-functions (we give a proof in the Euclidean context, making use of the fine properties of the dilation operators):

undamFour02 (159)
$$\int_{\mathbb{R}^d} \hat{f}(y)g(y)dy \quad = \quad \int_{\mathbb{R}^d} f(x)\hat{g}(x)dx.$$

A paper giving a short proof based on Poisson's formula is [84]: Alain Robert: A short proof of the Fourier inversion formula, which can be said to be based on the Fourier invariance of the Shah-distributions (Dirac comb) (probably in [85] or [86]).

iFourInvers **Theorem 13.** *For $f \in \boldsymbol{L}^1(\mathbb{R}^d)$ with $\hat{f} \in \boldsymbol{L}^1(\mathbb{R}^d)$* [62] *one has the pointwise (a.e.) relation:*

ourinvForm1 (160)
$$f(x) = \int_{\mathbb{R}^d} \hat{f}(s)e^{2\pi i x\cdot s}ds.$$

Proof. There are many proofs to this results, most of them involving some kind of summability kernels, helping to derive (160) from (153). These arguments typically differ in the way how auxiliary functions g (and the corresponding $\hat{g}$) are chosen in order to achieve the desired goal.

One option is to make use of the **Gauss function**, given by

Gaussdef00 (161)
$$g_0(x) = e^{-\pi x^2} \quad \text{with} \quad \hat{g}_0(0) = \|g_0\|_{\boldsymbol{L}^1(\mathbb{R}^d)} = 1 = \|g_0\|_{\boldsymbol{L}^2(\mathbb{R}^d)}\,.$$

In addition we will make use of dilation properties of the Fourier transform, i.e. use the dilation operators St_ρ and D_ρ, $\rho > 0$ respectively.

We are going to prove this relationship for an individual point $x_0 \in \mathbb{R}^d$. In order to obtain the point value on the left hand side one would like to put $\hat{g} = \delta_{x_0}$, and correspondingly one may expect that $g = \chi_s$ could be the pure frequency (since we know that $\widehat{\chi_s} = \delta_s$), but this is not directly available in the $\boldsymbol{L}^1$-context. □

There is also another possibility that would allow us to verify that there is a dense subspace in $\big(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\big)$ which is Fourier invariant and for which the Fourier inversion formula is valid, namely the so-called exotic Banach space described in a joint paper with G. Zimmermann ([48]).

SC: NOT SURE WHETHER THIS FORMULA HASN'T BEEN GIVEN ELSEWHERE

It is an easy exercise to verify the following result: **[Rd-specific!]**

FTdilcomp **Lemma 44.** *For any $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ and $\rho \neq 0$ one has*

FTdilat1 (162)
$$\mathcal{F}(\mathrm{St}_\rho\mu) = \mathrm{D}_\rho(\mathcal{F}\mu) \quad \textit{and} \quad \mathcal{F}(\mathrm{D}_\rho\mu) = \mathrm{St}_\rho(\mathcal{F}\mu), \forall \rho \neq 0.$$

Proof. The proof of this so-called *intertwining* property for the Fourier transform is easily derived using the standard substitution. It is a special case of more general *automorphism* applied to "frequency domain". Of course the formulas are also valid for e.g. anisotropic dilations. □

Remark 36. Here one could discuss the action of an automorphism on the dual group. Given $\alpha \in \mathrm{Aut}(G)$ the adjoint automorphism α^* on $\widehat{G}$ is defined by

$$\alpha^*(\chi)(x) = \chi(\alpha(x)), \quad \chi \in \widehat{G}, x \in G.$$

And there are corresponding rules for the Fourier transform.

[62] There are of course many $\boldsymbol{L}^1$-functions, such as the BOX-function, whose FT, namely the SINC function, are *not* in $\boldsymbol{L}^1(\mathbb{R}^d)$. A lot of details about the behaviour of the SINC function can be found at the WIKIPEDIA page `https://en.wikipedia.org/wiki/Sinc_function`.

Remark 37. For real life applications this simply means that running a tape at double speed will result in double frequencies (we call it the "Mickey-Mouse-Effect"). Correspondingly running at lower speed produces deeper voices. The same is true for images. Resizing an image by making it small all the plan waves are resized, i.e. get finer (so-called higher frequency in the wave-number domain). For images one can use rotations as another important family of automorphisms of $\mathbb{R}^2$. Then one goes on an shows that rotating an image will have a corresponding effect on its spectrum (FFT2).

5.2. **A short proof of the Fourier Inversion Theorem.** Gauß function:

$$g_0(x) = e^{-\pi x^2} \quad \text{with} \quad \|g_0\|_{\boldsymbol{L}^2(\mathbb{R}^d)} = 1\,.$$

Show $\widehat{g_0} = g_0$ (proof: GZ)

Note that the $\boldsymbol{L}^1$-normalized dilation operator D^1_ρ is denoted by St_ρ in the notes of hgfei, (where "St" stands for stretching, or Stauch-Streck operator in German).

INVERSION THEOREM: for $f \in \boldsymbol{L}^1 \cap \boldsymbol{C}_0(\mathbb{R}^d)$ with $\widehat{f} \in \boldsymbol{L}^1$.

Remark 38. One can then argue separately that $\hat{f} \in \boldsymbol{L}^1(\mathbb{R}^d)$ already implies that there is exactly one *continuous* representative in the class representing f as member of $\boldsymbol{L}^1(\mathbb{R}^d)$ (defined as space of equivalence classes of measurable functions), to which the inversion formula converges, but this is outside of the scope of these notes.

$$\begin{aligned} f(x) &= \lim_{\rho\to\infty} \int f(t)\,[\mathrm{T}_x\,\mathrm{D}^1_\rho\, g_0](t)\,dt \\ &= \lim_{\rho\to\infty} \int \widehat{f}(t)\,[\mathrm{M}_x\,\mathrm{D}^0_{1/\rho}\, g_o](t)\,dt = \int \widehat{f}(t)\,e^{2\pi i t x}\,dt \end{aligned}$$

i.e. Inversion Theorem. Thus $\boldsymbol{A}\cap\boldsymbol{L}^1 = \{f \in \boldsymbol{L}^1 : \widehat{f} \in \boldsymbol{L}^1\}$

ALTERNATIVE writeup:
Combing the formula (which is a simple consequence of Fubini's theorem)

$$\int_{\mathbb{R}^d} f(y)\hat{g}(y)dy = \int_{\mathbb{R}^d} \hat{f}(z)g(z)dz$$

for any pair of integrable functions $f, g \in \boldsymbol{L}^1(\mathbb{R}^d)$ with the fact that, due to the properties of the Fourier transform we have

$$\mathcal{F}(M_x \mathrm{D}_\rho h) = T_x \mathrm{St}_\rho \mathcal{F}(h), \quad x \in \mathbb{R}^d, \rho > 0,$$

specialized to $h = g_0$ and $\hat{g} = T_x \mathrm{St}_\rho h$ we get, at least for $f \in \boldsymbol{L}^1 \cap \boldsymbol{C}_0(\mathbb{R}^d)$:

ourInvFast2 (163) $$f(x) = \lim_{\rho\to\infty} \int f(t)\,[\mathrm{T}_x\,\mathrm{St}_\rho g_0](t)\,dt = \lim_{\rho\to\infty} \int \widehat{f}(t)\,[\mathrm{M}_x\,\mathrm{D}_\rho g_o](t)\,dt = \int \widehat{f}(t)\,e^{2\pi i t x}\,dt$$

A detailed proof of the Fourier invariance of the Gauss function can be found in Example 1.3.3 of [6]. For a technical report on this see [108] (Technical Report by G. Zimmermann).

Remark 39. One can then argue separately that $\hat{f} \in \boldsymbol{L}^1(\mathbb{R}^d)$ already implies that there is exactly one *continuous* representative in the class representing f as member of $\boldsymbol{L}^1(\mathbb{R}^d)$ (defined as space of equivalence classes of measurable functions), to which the inversion formula converges, but this is outside of the scope of these notes.

$$\begin{aligned} f(x) &= \lim_{\rho\to\infty} \int f(t)\,[\mathrm{T}_x\,\mathrm{D}^1_\rho\, g_0](t)\,dt \\ &= \lim_{\rho\to\infty} \int \widehat{f}(t)\,[\mathrm{M}_x\,\mathrm{D}^0_{1/\rho}\, g_o](t)\,dt = \int \widehat{f}(t)\,e^{2\pi i t x}\,dt \end{aligned}$$

i.e. Inversion Theorem, which has as a natural domain the Banach space $\boldsymbol{A}\cap\boldsymbol{L}^1 = \{f \in \boldsymbol{L}^1 : \widehat{f} \in \boldsymbol{L}^1\}$, with its natural norm $\|f\| := \|f\|_1 + \|\widehat{f}|_1$.

ALTERNATIVE writeup:
Combing the formula (which is a simple consequence of Fubini's theorem)

$$\int_{\mathbb{R}^d} f(y)\hat{g}(y)dy = \int_{\mathbb{R}^d} \hat{f}(z)g(z)dz$$

for any pair of integrable functions $f, g \in \boldsymbol{L}^1(\mathbb{R}^d)$ with the fact that, due to the properties of the Fourier transform we have

$$\mathcal{F}(M_x \mathrm{D}_\rho h = T_x \mathrm{St}_\rho \hat{h}, \quad x \in \mathbb{R}^d, \rho > 0,$$

specialized to $h = g_0$ and $\hat{g} = T_x \mathrm{St}_\rho h$ we get, at least (exactly) for $f \in \boldsymbol{L}^1 \cap \boldsymbol{C}_0(\mathbb{R}^d)$:

ourInvFast4 (164)

$$f(x) = \lim_{\rho\to 0} \int f(y)\,[\mathrm{T}_x\,\mathrm{St}_\rho g_0](y)\,dy = \lim_{\rho\to 0} \int \widehat{f}(y)\,[\mathrm{M}_x\,\mathrm{D}_\rho g_o](y)\,dy = \int \widehat{f}(y)\,e^{2\pi i y x}\,dy = \int \widehat{f}(y)\,e^{2\pi i x y}\,dy.$$

This chain of equations can be rewritten as

chiFTdx (165) $$\delta_x = w^*\text{-}\lim{}_{\rho\to 0} T_x \mathrm{St}_\rho g_0 = w^*\text{-}\lim{}_{\rho\to 0} \mathcal{F}(M_x \mathrm{D}_\rho g_0) = \widehat{\chi_x}$$

In the last step we could argue that we have essentially

$$w^*\text{-}\lim{}_{\rho\to 0} \mathcal{F}(M_x \mathrm{D}_\rho g_0) = \mathcal{F}(M_x w^*\text{-}\lim{}_{\rho\to 0} \mathrm{D}_\rho g_0) == \mathcal{F}(\chi_x \cdot \mathbf{1}) = \widehat{\chi_x}.$$

A detailed proof of the Fourier invariance of the Gauss function can be found in Example 1.3.3 of [6]. For a technical report on this see [108] (Technical Report by G. Zimmermann), but there should be many others, e.g. Stein's book [93]. (CHECK)

The first limit is justified because $\mathrm{St}_\rho g_0 \to \delta$ in the w^*-sense, or $\int_{\mathbb{R}^d} h(x)\mathrm{St}_\rho g_0 = h(0)$ for any $h \in \boldsymbol{C}_0(\mathbb{R}^d)$, applied to $h = T_{-x}f \in \boldsymbol{W}(\mathbb{R}^d) \subset \boldsymbol{C}_0(\mathbb{R}^d)$, resulting in

$$f(x) = f(0+x) = T_{-x}f(0) = \lim_{\rho\to 0}\int_{\mathbb{R}^d} f(t+x)\mathrm{St}_\rho g_0(t)dt,$$

which is equal to the expression in the first limit.

For the convergence of the second argument we use the fact that $\widehat{f} \in \boldsymbol{W}(\mathbb{R}^d)$ by the density of $\boldsymbol{C}_c(\mathbb{R}^d)$ in $\big(\boldsymbol{W}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{W}}\big)$ one can restrict the attention to convergence of $\mathrm{D}_\rho g_o(t) \to 1$ for $\rho \to 0$, uniformly over compact sets. Details are left to the reader.

Reading the left hand side as a function of x it is easily reinterpreted as $\mathrm{St}_\rho g_0 * f(x)$, which tends to $f(x)$ uniformly for any $f \in \boldsymbol{C}_0(\mathbb{R}^d)$, but also in the Wiener norm for $f \in \big(\boldsymbol{W}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{W}}\big)$.

Material on Banach Modules

The Banach module is called *true* if the mapping J described above is injective, i.e.if $\mathbf{a} \bullet \mathbf{b} = 0 \forall \mathbf{b} \in \boldsymbol{B}$ implies that $\mathbf{a} = 0$. For the case of a "non-true" Banach module one should factor out some closed ideal in the Banach algebra in order to turn the Banach module into a true Banach module with respect to another Banach algebra (which then "truly acts" on $\boldsymbol{B}$);

If one only has a continuous (but not necessarily non-expansive algebra homomorphism J) one can replace the norm on $\boldsymbol{A}$ by another equivalent norm (just some constant multiple of the original one) in order to ensure this (harmless) extra property.

Recall the notions of *weak topology* on any Banach space (such as $\big(\boldsymbol{C}_0(\mathbb{R}^d),\, \|\cdot\|_\infty\big)$, and the *$w^*$-topology*) on any dual space, such as $(\boldsymbol{M}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}})$.

Theorem 14. *A sequence (or indeed a bounded net) of functions in $\big(\boldsymbol{C}_0(\mathbb{R}^d),\, \|\cdot\|_\infty\big)$ is weakly convergent if and only if it is pointwise convergent (while in contrast norm-convergence means uniform convergence over $\mathbb{R}^d$).*

Proof. Since the Dirac measures are specific linear functionals on $\big(\boldsymbol{C}_0(\mathbb{R}^d),\, \|\cdot\|_\infty\big)$ weak convergence of a sequence (f_n) in $\boldsymbol{C}_0(\mathbb{R}^d)$ implies $f_n(x) = \delta_x(f_n) \to \delta_x(f_0) = f_0(x)$ for any $x \in \mathbb{R}^d$. Conversely, the possibility of approximating a general measure in a bounded way by linear combinations of Dirac measures implies that pointwise convergence indeed implies weak convergence. If one goes into the details of the proof the boundedness of the set of approximating measures as well as the boundedness of the sequence (resp. a net (f_α)) is of importance. $\square$

Remark 40. For *equicontinuous* families one can show that weak (or pointwise) convergence is equivalent to "uniform convergence over compact set". A net (f_α) is weakly convergent if and only if it is UCOCS, etc. dots

Remark 41. How can we characterize w^*-convergence in $\boldsymbol{M}(\mathbb{R}^d)$? (for bounded sets): cf. Bernoulli convergence, resp. *vague convergence* ($= \sigma(\boldsymbol{C}_c(G), \mathcal{R}(G))$) using standard topological constructions in functional analysis), pointwise convergence of the STFT or of the Fourier transform etc...

A neighborhood of $\mathbf{0} \in \mathcal{R}(G)$ is then given by a finite subset $F \subset \boldsymbol{C}_c(G)$ and $\varepsilon > 0$, via

vagneigh1
$$U(F,\varepsilon) := \{\nu \mid |\nu(k)| < \varepsilon, \forall k \in F\} \tag{166}$$

Alternative description of "multipliers" on $\boldsymbol{C}_0(\mathbb{R}^d)$ resp. translation invariant BIBOS (bounded input bounded output systems, with the property of mapping $\boldsymbol{C}_0(\mathbb{R}^d)$ into itself). Any reasonable system mapping *individually* bounded functions into bounded functions will be a bounded linear operator with uniform control of the relative size, due to the Closed Graph Theorem from Functional Analysis.

cterizeHGC0 **Theorem 15.** *Let H be any w^*- total subset*[63] *of $\boldsymbol{M}(\mathbb{R}^d)$, and assume that a bounded linear operator $T \in \mathcal{L}(\boldsymbol{C}_0(\mathbb{R}^d))$ commutes with the action of H, i.e.that the commutators $[C_h, T] \equiv 0$ for all $h \in H$. Then $T \in \mathcal{H}_{\mathbb{R}^d}(\boldsymbol{C}_0(\mathbb{R}^d))$.*

Note that in the original definition the set H was just the set of convolution operators by Dirac measures $\delta_x, x \in \mathbb{R}^d$ (or at least from some dense subset).

UPCOMING MATERIAL:

[63] A subset M in $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ is called total if the linear span of M is dense in $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$.

Embedding of test functions into $\boldsymbol{M}(\mathbb{R}^d)$ (over groups this requires the use of the [invariant] Haar measure, which indeed is a linear functional on $\boldsymbol{C}_c(G)$). Compatibility of operators which are now available on both the functions and the measures (resp. functionals). E.g. we can now do an internal convolution of functions (viewed as bounded measures) or an external action (one is acting as a bounded measure, the other is consider as the $\boldsymbol{C}_0(\mathbb{R}^d)$ element on which the action takes place). Associativity of convolution in the most general situation (also of course **commutativity**, etc.).

$\boldsymbol{C}_b(\mathbb{R}^d)$ is a (closed) subspace of the dual of $\boldsymbol{L}^1(\mathbb{R}^d)$. This is easily verified directly (without using the fact that the full dual space can be identified with $\big(\boldsymbol{L}^\infty(\mathbb{R}^d), \|\cdot\|_\infty\big)$. In fact one just has to verify first that for $h \in \boldsymbol{C}_b(\mathbb{R}^d)$ one finds that $\sigma_h : f \mapsto \int_{\mathbb{R}^d} f(t)h(t)dt$ is a bounded linear functional on $\boldsymbol{L}^1(\mathbb{R}^d)$ and that $\|\sigma_h\|_{\boldsymbol{L}^{1\prime}} = \|h\|_\infty$, and thus consequently $\big(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\big)$ is isometrically embedded into the dual of $\big(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\big)$. Similar statements can be made for general groups, just by replacing the Lebesgue integral by the Haar measure (which may be viewed simply as a translation invariant, continuous linear functional on $\boldsymbol{C}_c(\mathbb{R}^d)$)[64].

Hence it carries a $\sigma(\boldsymbol{C}_{ub}(\mathbb{R}^d), \boldsymbol{L}^1(\mathbb{R}^d))$ topology which can be shown to be equivalent (at least on bounded sets!?) to the uniform convergence over compact sets DETAILS MISSING, SHOULD BE EXPLAINED!

STATEMENT: Every $f \in \boldsymbol{C}_b(\mathbb{R}^d)$ is a limit (in the sense of uniform limit over compact sets) of a *bounded* sequence of functions from $\boldsymbol{C}_0(\mathbb{R}^d)$ resp. even from $\boldsymbol{C}_c(\mathbb{R}^d)$. In fact, on can take the sequence $p_n \cdot f$, where (p_n) is a BAI for $\boldsymbol{C}_0(\mathbb{R}^d)$ consisting of (increasing) plateau functions.

EXTENSION PRINCIPLE. Let (p_n) be as above, and $f \in \boldsymbol{C}_b(\mathbb{R}^d)$ and $\mu \in \boldsymbol{M}(\mathbb{R}^d)$ be given. Then the sequence $\mu(p_n \cdot f)$ is a Cauchy sequence, hence convergent in $\mathbb{C}$. In fact, the limit is the same for any other BAI in $\boldsymbol{C}_0(\mathbb{R}^d)$. Therefore it makes sense to define $\mu(f) := \lim_{n\to\infty} \mu(p_n f)$.

FourStdef5 *Remark* 42. This will be important to define the Fourier Stieltjes transforms for bounded measures, i.e. for $\hat{\mu}(s) = \mu(\chi_{-s})$ later on! We will also have then to verify that this definition is consistent with the definition of the generalized Fourier transform at the level of $\boldsymbol{S}_0'(\mathbb{R}^d)$ or $\boldsymbol{\mathcal{S}}'(\mathbb{R}^d)$.

From the above arguments it is also quite clear that the Fourier Stieltjes transforms for the discrete approximations $D_\Psi\mu$ converge pointwise to the Fourier-Stieltjes transform of μ. We formulate it as lemma.

limFTDpsi **Lemma 45.** *We have for each $s \in \widehat{\mathbb{R}^d}$ and $\mu \in \boldsymbol{M}(\mathbb{R}^d)$:*

$$\lim_{\operatorname{diam}(\Psi)\to 0} \widehat{D_\Psi\mu}(s) = \hat{\mu}(s), \quad \textit{uniformly over compact sets.}$$

Proof. We play the usual trick. In order to show that for a given point $s \in \widehat{\mathbb{R}^d}$ (in the “frequency domain”, so to say) we want to make sure that

siFourconv1
$$|\widehat{D_\Psi\mu}(s) - \hat{\mu}(s)| \le \varepsilon \quad \text{for} \quad \operatorname{diam}(\Psi) < \delta, \tag{167}$$

observation that there exists some plateau function $p \in \boldsymbol{C}_c(\mathbb{R}^d)$ with $\|p\|_\infty = 1$ such that $\|(1-p)\cdot D_\Psi\mu\|_{\boldsymbol{M}} < \varepsilon/2$ for all Ψ, and also (in fact as a consequence) $\|(1-p)\cdot\mu\|_{\boldsymbol{M}} < \varepsilon/3$.

Then observing that $p \cdot \chi_{-s} \in \boldsymbol{C}_c(\mathbb{R}^d)$ we have by the established w^*-convergence the convergence of

$$D_\Psi\mu(p\cdot\chi_{-s}) \to \mu(p\cdot\chi_{-s}),$$

[64] In fact, it automatically extends, due to its translation invariance to the Banach space, in fact the pointwise algebra $\big(\boldsymbol{W}(G), \|\cdot\|_{\boldsymbol{W}(G)}\big)$, the Wiener algebra over G, see Def. 43 below.

or in other words for a well-chosen $\delta > 0$ we have: $\mathrm{diam}(\Psi) < \delta$ implies

$$|(D_\Psi\mu - \mu)(p \cdot \chi_{-s})| \le \varepsilon/3.$$

Combining these estimates we get

$$|\widehat{D_\Psi\mu}(s) - \hat{\mu}(s)| \le |D_\Psi\mu((1 - p \cdot (\chi_{-s})| + ...$$

and consequently the estimate

$$|\widehat{D_\Psi\mu}(s) = \hat{\mu}(s)| \le \|D_\Psi\mu \cdot (1-p)\|_{\boldsymbol{M}}\|\chi_{-s}\|_\infty + |(D_\Psi\mu - \mu)(p \cdot \chi_{-s})| + \|\mu \cdot (p-1)\|_{\boldsymbol{M}}\|\chi_{-s}\|_\infty.$$

Since $\|\chi_r\|_\infty = 1$ for any $r \in \mathbb{R}^d$ we have altogether obtained the desired estimate, whenever $\mathrm{diam}(\Psi) \le \delta$. □

DPsifour2 *Remark* 43. Probably the result could be obtained easier by the fact that the uniform continuity of χ_s, for any given s and in fact for any compact range of parameters $s \in \mathbb{R}^d$ implies that

$$\mathrm{Sp}_\Psi(\chi_{-s}) \to \chi_{-s} \quad \mathrm{in}\big(\boldsymbol{C}_b(\mathbb{R}^d),\, \|\cdot\|_\infty\big).$$

convopcomm **Lemma 46.** *The convolution operators form a (commutative!) Banach algebra of operators. It turns out that the characters can be identified with the joint eigenvectors for this whole class of operators. Indeed, we have $C_\mu(\chi_s) = \hat{\mu}(s)\chi_s$ for any $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ and any character χ_s on $\mathbb{R}^d$.*

Proof. The claim is valid for $\mu = \delta_x$, for any $x \in \mathbb{R}^d$, and hence for finite linear combinations of measures and also immediately for discrete measures (due to a simple approximation argument). However, it is enough to use the fact that one has Bernoulli convergence of $D_\Psi\mu$ to μ (i.e. w^*-convergence by a tight net of discrete, bounded measures with limit μ).

Maybe even more convincing is the following observation. 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

FourConvEig
$$\text{“}\mu * \chi_s(z)\text{''} = C_\mu(\chi_s)(z) = \mu(T_z(\chi_{-s}{}^{\checkmark})) = \mu(\chi_{-s}) \cdot \chi_s(z) = \hat{\mu}(s)\chi_s(z), \tag{168}$$

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

eigenCmu1
$$C_\mu(\chi_s) = \hat{\mu}(s) \cdot \chi_s. \tag{169}$$

So in some sense the Fourier (Stieltjes) transform is the description of a commutative Banach algebra of operators (e.g. on $\boldsymbol{L}^2(\mathbb{R}^d)$, or in our case at the moment $\big(\boldsymbol{C}_0(\mathbb{R}^d),\, \|\cdot\|_\infty\big)$) via their (joint) eigen-vectors. □

Although the following result will be a consequence of basic facts about a more general (distributional) Fourier transform let us claim (and verify) the injectivity of the Fourier Stieltjes transform just defined:

FourStinj **Proposition 5** (Uniqueness of Fourier Stieltjes transform)**.** *Let $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ be given, and assume that $\hat{\mu}(\chi) \equiv 0$. Then $\mu = 0$ in $\boldsymbol{M}_b(\mathbb{R}^d)$. Consequently the bounded linear mapping $\mu \to \hat{\mu}$ from $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ into $\big(\boldsymbol{C}_{ub}(\mathbb{R}^d),\, \|\cdot\|_\infty\big)$ is injective.*

Proof. We use the density of $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ in $\big(\boldsymbol{C}_0(\mathbb{R}^d),\, \|\cdot\|_\infty\big)$ first, which is a consequence of the fact that $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ is a Banach algebra (due to the Convolution Theorem) which is of course closed under conjugation ($h = \hat{f}$ ..) and point separating, hence it is dense as a consequence of the Stone-Weierstrass approximation theorem. Being a subspace of $\boldsymbol{M}_b(\mathbb{R}^d)$ we also know that we can approximate any $g \in \boldsymbol{L}^1(\mathbb{R}^d)$ in the w^*-sense by discrete measures, in such a way that the corresponding Fourier Stieltjes transforms (they are trigonometric polynomials) converge to $\hat{g}$ uniformly over compact sets. □

What is left is to recover (as a w^*-limit) a measure from its Fourier transform. What one can expect is that the same trick as the usual inversion theorem works. Multiply the Fourier transform with summability kernel and let them go to constant 1, e.g. by dilation D_ρ, $\rho \to 0$, for example applied to a Gauss function. The inverse Fourier transform if these objects will converge in the w^*-sense to the original measure, e.g. because it is of the form

$$\mathrm{St}_\rho g_0 * \mu, \quad \rho \to 0. \tag{170}$$

6. Identifying "ordinary functions with functionals"

There is a lot to be said about ordinary functions and generalized functions: above all one has to be aware that most of the things which can be done in one or the other way for ordinary, at least for nice functions, can be done for "generalized functions" in the same way, e.g. they can be translated, multiplied by (decent) continuous functions, they can be dilated, or more generally transformed, e.g. be rotated, and often they have a Fourier transform of the same type. It is possible to determine their support there, and there are also convolutions and other things. In all these situations there should not be any difference whether and when one likes to switch from one consideration to another one. It is like multiplying with rational numbers and/or with real numbers. What is the quotient of πp over πq, i.e. of two irrational numbers, if p, q are non-zero integers. But it is enough to know that we can rewrite it (without doing the actual computation) to

$$\frac{p\pi}{q\pi} = \frac{p}{q}.$$

and if we are asked to compute the inverse of this number we just write q/p (without any computation). OF COURSE we could have evaluated (well: just approximately) the two irrational numbers and then apply a numerical method to compute their quotient, and the result would have been (at least up to the level of precision used) more or less the same. ALSO this mini-example (or ex tempore) should be elaborated a little bit more, cf. my "story of $1/\pi^2$.

There is a natural way to identify "ordinary functions" (say $k \in \boldsymbol{C}_c(\mathbb{R}^d)$) with linear functionals $\mu \in \boldsymbol{M}(\mathbb{R}^d)$, by the following trick: Given

cttomeasure
$$\mu = \mu_k, \qquad \text{resp.} \qquad \mu(f) = \int_{\mathbb{R}^d} f(x)k(x)dx \tag{171}$$

It is important to note that this is indeed an injective (linear) mapping from $\boldsymbol{C}_c(\mathbb{R}^d)$ into $\boldsymbol{M}_b(\mathbb{R}^d)$. As we will see it is isometric, if we impose on $\boldsymbol{C}_c(\mathbb{R}^d)$ the usual $\boldsymbol{L}^1$-norm (but for this one needs only the Riemann integral!).

This is also possible over general locally compact Abelian groups, but requires the existence of the Haar measure (we will not go into this direction, a good explanation is given in Deitmar's book [23]).

isomL1emb **Proposition 6** (Functions to Measures)**.** *There is a natural,* isometric *embedding of* $(\boldsymbol{C}_c(G), \|\cdot\|_1)$ *into* $\boldsymbol{M}_b(G)$*, given by*

functomeas
$$k \mapsto \mu_k : \mu_k(f) = \int_{\mathbb{R}^d} f(x)k(x)dx, f \in \boldsymbol{C}_0(\mathbb{R}^d). \tag{172}$$

Proof. It is clear that each μ_k is a bounded linear functional on $\boldsymbol{C}_0'(G)$ and that the mapping $k \mapsto \mu_k$ is linear and nonexpansive, since evidently for each $f \in \boldsymbol{C}_0(G), k \in \boldsymbol{C}_c(G)$ one has:

$$|\mu_k(f)| = \left| \int_G f(x)k(x)dx \right| \le \int_G |f(x)||k(x)|dx \le \|f\|_\infty \|k\|_1. \tag{173}$$

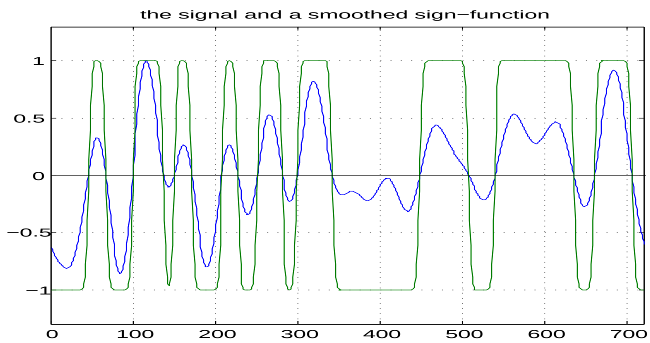


FIGURE 8. signcont1a.pdf

The converse is a bit more involved. In principle one would choose, for the case of a real-valued function k a function $f \in \boldsymbol{C}_0(G)$ which is a minimal (but continuous!) modification of the signum function, which turns (when integrated against k) the negative parts into positive parts, thus turning $\mu_k(f)$ in a good approximation of $\|k\|_1$.

To be more formal let us consider $f \in \boldsymbol{C}_c(G)$, let us consider for any $\eta > 0$ the "essential" support $K_\eta := \{z \in G \,|\, |k(z)| \geq \eta\}$. Then K_η is a compact set and we can find a continuous function [65] h_η with values in $[0,1]$ such that $h_\eta(z) = 1$ on K_η and with support of h_η within (the interior) of $K_{\eta/2}$. The function $f_\eta(x) := h_\eta(x)|k(x)|/k(x)$ is then well defined (because $k(x) \neq 0$ for any point in the support of h_η, and $\|f_\eta\|_\infty \leq 1$). We observe that

$$\mu_k(f) = \int_G f_\eta(x)k(x)dx = \int_G |k(x)|h_\eta(x)dx. \tag{174}$$

It remains to verify that this tends to $\|k\|_1$ for $\eta \to 0$.

[65] The existence of h_η is guaranteed by Tietze's theorem, one of the important theorems concerning locally compact, hence completely regular topological spaces. It helps to avoid the potential problem of a phase discontinuity, i.e. problems with the continuity of $k(x)/|k(x)|$ near the zeros of k.

Writing K_0 for supp(k) this is a direct consequence of the following estimate

xxxyyy

$$\int_G |k(x)(1-h_\eta(x))| \leq \int_{K_0} |k(x)|(1-h_\eta(x))dx \leq Vol(K_0)^{66} \cdot \|k(1-h_\eta)\|_\infty \to 0. \tag{175}$$

□

Notation: $\boldsymbol{M}_{cs}$ = continuous shift.

Li-Def **Definition 36.** We define $\boldsymbol{L}^1(\mathbb{R}^d)$ as the following subspace of $\boldsymbol{M}_b(\mathbb{R}^d)$:

tshiftmeas1

$$\boldsymbol{M}_{cs} := \{\mu \in \boldsymbol{M}_b(\mathbb{R}^d) \,|\, \|T_x\mu - \mu\|_{\boldsymbol{M}_b(\mathbb{R}^d)} \to 0 \quad \text{for} \quad x \to 0.\} \tag{176}$$

It is an easy exercise to verify that

Mcsclosed1 **Lemma 47.** *$\boldsymbol{M}_{cs}(\mathbb{R}^d)$ is a closed ideal within $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ with respect to convolution. It is also invariant under the usual involutions (e.g. $\mu \to \mu^{\checkmark}$.).*

An important result is the fact that for the measure in $\boldsymbol{M}_{cs}(\mathbb{R}^d)$ we can claim that the Fourier transform is not only in $\boldsymbol{C}_{ub}(\mathbb{R}^d)$ but even in $\boldsymbol{C}_0(\mathbb{R}^d)$. This result is known as the *Riemann-Lebesgue Lemma* (it is usually called a lemma, but it is of great use and great importance).

RieLebLemma **Theorem 16.** *For $\mu \in \boldsymbol{M}_{cs}(\mathbb{R}^d) = \boldsymbol{L}^1(\mathbb{R}^d)$ we have $\widehat{\mu} \in \boldsymbol{C}_0(\mathbb{R}^d)$. The range of the FT will be denoted by $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$. It is a dense Banach space inside of $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$, endowed with the norm $\|\widehat{f}\|_{\mathcal{F}\boldsymbol{L}^1} := \|f\|_1$. But it is a proper subspace of $\boldsymbol{C}_0(\mathbb{R}^d)$. Expressed equivalently: the norms $f \to \|\widehat{f}\|_\infty$ and $f \to \|f\|_{\boldsymbol{L}^1}$ are inequivalent.*

Proof. TO BE DONE

The final statement expresses the fact that there exists a sequence $(f_n)_{n\geq 1}$ such that $\|f\|_{\boldsymbol{L}^1} = 1$ and $\lim_{n\to\infty} \|\widehat{f}\|_\infty = 0$. □

This example, which can be written abstractly as a function of the form (for some $\alpha \in \mathbb{N}$, not too small, e.g. $\alpha = 6$ or $\alpha = 10$).

Form

$$h = \sum_{k=-K}^{K} [M_{\alpha k} T_{\alpha k} g_0 + M_{-\alpha k} T_{\alpha k} g_0]$$

or equivalently

$$h_1 = \sum_{k=0}^{K} \cos(2\pi\alpha k t) \cdot (g_0(t-\alpha k) + g_0(t+\alpha k)).$$

This function can be made (adding the conjugate?) Fourier invariant, and then it is clear that the $\mathcal{F}\boldsymbol{L}^1$-norm, which is just the $\boldsymbol{L}^1$-norm of such a function blows up, while the sup-norm is staying fixed, independent of K, the number of terms (roughly). Since $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ is a dense subspace of $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$ the non-equivalence of norms is equivalent to have a proper embedding[67]

Following the Riemann-Lebesgue Lemma we derive:

[66] $Vol(K_0)$ stands for the Haar measure of the set K_0, but the $\|\cdot\|_1$ of a plateau-function $p(x)$ with p with $p(x) \cdot k(x) = k(x)$ would do. In fact, the "measure of K_0, i.e. $Vol(K_0) = \mu(K_0)$ can be shown to be equal to the infimum over all those $\|\cdot\|_1$-norms.

[67] Otherwise the embedding would be surjective, and thus clearly bijective, but then the Theorem of Banach (well known principle from functional analysis) implies that the inverse mapping is also continuous, but that means the norms would have to be equivalent, in contradiction to our observation.

ortest1.jpg

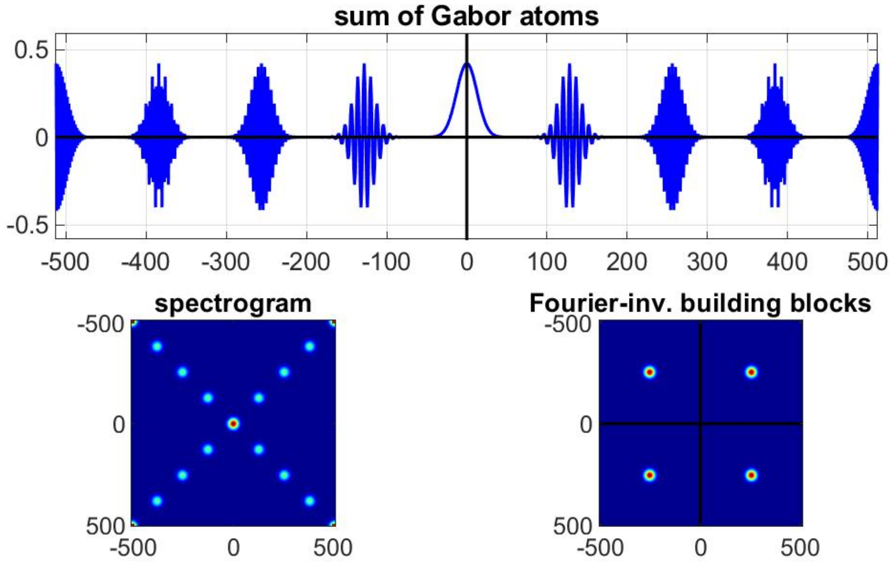


FIGURE 9. FourGabortest1.jpg

ortest2.jpg

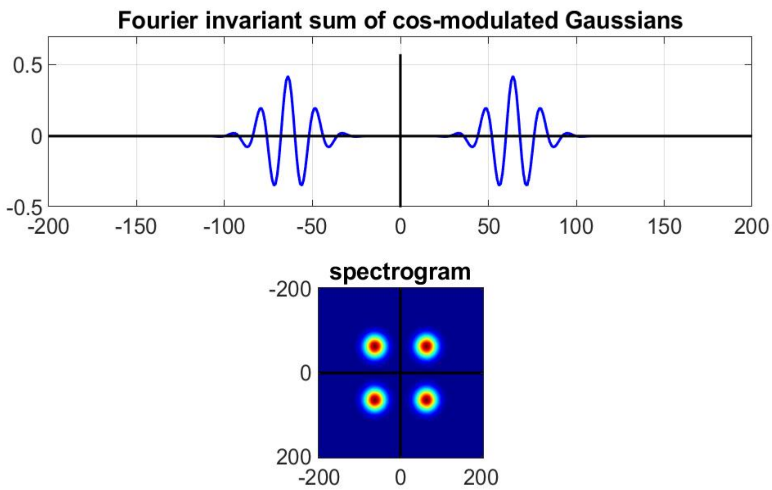


FIGURE 10. FourGabortest2.jpg

Four-fund0 **Theorem 17** (Fundamental Theorem for Fourier Analysis). *Let $f, g \in \boldsymbol{L}^1(\mathbb{R}^d)$ be given. Then*

nd-Fourier0 (177)
$$\int_{\mathbb{R}^d} f(t)\hat{g}(t)dt = \int_{\mathbb{R}^d} \hat{f}(x)g(x)dx$$

Proof. The Fourier transforms $\hat{f}$ and $\hat{g}$ are bounded and continuous, hence both integrands are in $\boldsymbol{L}^1(\mathbb{R}^d)$. The relation (177) then follows via Fubini's theorem (at (*)), since $e^{2\pi i t\cdot x} f(t)g(x) \in$

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

$$\begin{aligned}
\int_{\mathbb{R}^d} f(t)\hat{g}(t)dt &= \int_{\mathbb{R}^d} f(t)\left(\int_{\mathbb{R}^d} e^{-2\pi i x\cdot t} g(x)dx\right) \\
&\underset{(*)}{=} \int_{\mathbb{R}^d} g(x)\left(\int_{\mathbb{R}^d} e^{-2\pi i x\cdot t} f(t)dt\right)dx \qquad (178)\\
&= \int_{\mathbb{R}^d} f(x)\hat{g}(x)dx.
\end{aligned}$$

□

Of course one has to justify this definition, by recalling that the usual definition of $\boldsymbol{L}^1(\mathbb{R}^d)$ based on Lebesgue's integrability criterion provides us with a Banach space (of equivalence classes of measurable functions, identifying two functions if they are equal almost everywhere), which contains $\boldsymbol{C}_c(\mathbb{R}^d)$ as a dense subspace (cf. more or less any book on measure theory for details on this matter: in fact, it is sufficient to approximate - in the $\boldsymbol{L}^1$-norm - indicator functions of parallel-epipeds by continuous functions with compact support, i.e. something like a trapezoidal function sufficiently close to a "box-car"-function in the 1-dimensional case).

It is also of interest to introduce **the concept of a support to measures**, in a way which is compatible with the notion $\mathrm{supp}(k)$ for $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ given above:

eas-support **Definition 37.** A point x does *not* belong to the support of a measure $\mu \in \boldsymbol{M}(\mathbb{R}^d)$ if there exists some $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ with $k(x) = 1$, **but nevertheless** $k \cdot \mu = 0$. (!) The complement of this set is denoted by $\mathrm{supp}(\mu)$.

nsistency11 **Lemma 48.** *.*

- $\mathrm{supp}(\mu)$ *is a closed subset of* $\mathbb{R}^d$*,*
- *there is consistency with the concept already defined for* $k \in \boldsymbol{C}_c(\mathbb{R}^d)$*, in other words:* $\mathrm{supp}(k)$ *(in the old sense) coincides with* $\mathrm{supp}(\mu_k)$*, just defined.*
- *For a discrete measure the support is given by the closure of the union of all points involved, i.e.for* $\mu = \sum_{k=1}^{\infty} c_k \delta_{t_k}$ *we have*[68] $\mathrm{supp}(\mu) = \left(\bigcup_k t_k\right)^{-}$*, where the bar denotes "closure".*
- *The notion of support is compatible with pointwise products: i.e.for any* $h \in \boldsymbol{C}_b(\mathbb{R}^d)$ *on has* $\mathrm{supp}(h\mu) \subseteq \mathrm{supp}(h) \cap \mathrm{supp}(\mu)$ *(as for functions).*
- *consequently every* $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ *can be approximated by compactly supported measures of the form* $p \cdot \mu$*, with* $\mathrm{supp}(p \cdot \mu) \subset \mathrm{supp}(\mu)$*.*

We have the following equivalent description of the $\mathrm{supp}(\mu)$ (which is actually the usual definition):

eas-charact **Lemma 49.** *The following properties are equivalent:*

- $z \in \mathrm{supp}(\mu)$*;*
- *for any* $\varepsilon > 0$ *there exists some* $h \in \boldsymbol{C}_c(\mathbb{R}^d)$ *with* $\mathrm{supp}(h) \subseteq B_\epsilon(z)$ *with* $\mu(h) \neq 0$*.*
- $\mathrm{supp}(\mu)$ *coincides with the intersection of all supports of functions* p *such that* $p \cdot \mu = \mu$*.*

Remark 44. A measure $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ has compact support if and only if for one (or in fact any) BUPU $\Psi = (\psi_i)_{i\in I}$ there exists a finite subset $F \subset I$ such that

$$\mu = \sum_{i\in F} \mu\psi_i = \mu \cdot (\sum_{i\in F} \psi_i) = \mu \cdot \psi_F.$$

[68] Assuming of course the canonical representation of μ, with $t_k \neq t_l$, for $k \neq l$ and $c_k \neq 0$.

Equivalently, μ has compactly support if and only if there exists $p \in \boldsymbol{C}_c(\mathbb{R}^d)$ such that $\mu = \mu \cdot p$. It is a good exercise to check the details of these claims.

The following results indicate the w^*-continuity of the concept of a support.

wstsuppconv **Lemma 50.** *Assume that $\mu_0 = w^* - lim_\alpha \mu_\alpha$, then for any $\alpha_0 \in I$ we have:*

$$\operatorname{supp}(\mu_0) \subseteq \bigcup_{\alpha \succeq \alpha_0} \operatorname{supp}(\mu_\alpha).$$

The notion of support is also compatible with convolution: [69]

HINT: The following result is certainly true for $f \in \boldsymbol{C}_c(\mathbb{R}^d)$ and $\operatorname{supp}(\mu)$ is compact. We are not sure if it is valid in this form in the general situation described next:

onv-support **Lemma 51.** *For $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ and $f \in \boldsymbol{C}_b(\mathbb{R}^d)$ one has*

conv-supp
$$\operatorname{supp}(\mu * f) \subseteq \operatorname{supp}(\mu) + \operatorname{supp}(f). \tag{179}$$

For $\mu_1, \mu_2 \in \boldsymbol{M}_b(\mathbb{R}^d)$ one has:

conv-musupp
$$\operatorname{supp}(\mu_1 * \mu_2) \subseteq \operatorname{supp}(\mu_1) + \operatorname{supp}(\mu_2). \tag{180}$$

Proof. Assume that $Q_1 = \operatorname{supp}(\mu)$ and $Q_2 = \operatorname{supp}(f)$ are two compact sets. Then $Q_1 + Q_2$ is a compact, hence closed subset of $\mathbb{R}^d$ (as the image of the compact set $Q_1 \times Q_2 \subset \mathbb{R}^{2d}$ under the continuous addition). The compactness implies that any $u \notin Q_1 + Q_2$ has a positive distance from $Q_1 + Q_1$, larger than 2γ, for some $\gamma > 0$. Writing $W_l = Q_l + B_\gamma(0), l = 1, 2$ for the enlarged sets we can be sure that there exist plateau-functions $p_l \in \boldsymbol{C}_c(\mathbb{R}^d)$ such that $p_l(t) = 1$ on Q_l and $\operatorname{supp}(p_l) \subset W_l, l = 1, 2$, and thus

$$[\mu * f](u) = \mu(T_u(f^\checkmark)) = (\mu \cdot p_1)(T_u(p_2 \cdot f)^\checkmark) = \mu((p_1 \cdot T_u p_2{}^\checkmark) \cdot f^\checkmark) = \mu(0) = 0.$$

The second claim follows essentially (perhaps also only for measures with compact support?!) from the fact that $\delta_x * \delta_y = \delta_{x+y}$ and the fact that one can approximate μ in the w^*-sense by a bounded sequence of discrete measures supported on $\operatorname{supp}(\mu) + B_\varepsilon(0)$, for any $\varepsilon > 0$. In fact, if Ψ is a BUPU of diameter $\varepsilon > 0$ then it is clear that $\operatorname{supp}(D_\Psi \mu) \subset \operatorname{supp}(\mu) + B_\varepsilon(0)$. ☐

Lemma 52. *Assume that $\mu_0 = lim_{w^*} \mu_\alpha$. Then also for any BUPU Ψ the family $D_\Phi \mu_\alpha$ is w^*-convergent to $D_\Phi \mu_0$. The family $D_\Phi \mu_\alpha$ is uniformly tight and w^*-convergent to μ_0 as $|\Phi| \to 0$.*

Furthermore, the family $D_\Phi(p_K \mu_\alpha)$, where p_K runs through the family of all plateau functions (with $K \to \mathbb{R}^d$), satisfies the same relation. Note that the resulting measures are in fact finite discrete measures.[70]

It is an important and not completely trivial claim that a measure supported within the zero-set of a continuous function the action is zero. More precisely:

upinzeroset **Lemma 53.** $\operatorname{supp}(\mu) \subseteq \{x \,|\, h(x) = 0\}$ *implies* $\mu(h) = 0$.

Proof. It is enough to verify that any function $h \in \boldsymbol{C}_b(\mathbb{R}^d)$ can be approximated in the sup-norm (not so in other norms, like the $\mathcal{F}\boldsymbol{L}^1$-norm!) by other functions which vanish near the zero-set of h. This can be achieved by multiplying h with a plateau-type function which vanishes on $\{x \,|\, |h(x)| \leq \varepsilon/3\}$ and has a plateau on the set $\{x \,|\, |h(x)| \geq 2\varepsilon/3\}$. etc. ...

Alternatively, one may first show that it is no loss of generality to assume that h has compact support and thus that $h \in \boldsymbol{C}_{ub}(\mathbb{R}^d)$. In this case one can replace h by $Sp_\Phi h$, for sufficiently

[69] We probably still have to take care of the notion of support for the case that the measure does not have compact support.

[70] Alternatively one could use only finite subfamilies from the partition of unity, or put p_k on the outside, i.e. write $p_k \cdot D_\Phi \mu_\alpha$. The consequences remain the same for all of these variants!

fine Φ, and then discard all the elements of the form $h(x_j)\phi_j$ with $|h(x_j)| \leq \varepsilon/3$. The effect is practically the same. $\square$

Remark: There are related questions coming up with respect to spectral analysis!

xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx

7. Basic properties of $L^1(Rd)$

We have defined $\boldsymbol{L}^1(\mathbb{R}^d)$ as the closure of $\boldsymbol{C}_c(\mathbb{R}^d)$ (identified via $k \mapsto \mu_k$ with a subspace of $\boldsymbol{M}(\mathbb{R}^d)$) in $(\boldsymbol{M}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}})$. Since this is a Banach space, it is a Banach space itself, and identical with the (abstract) completion of $\boldsymbol{C}_c(\mathbb{R}^d)$ in $\boldsymbol{M}(\mathbb{R}^d)$.

The next theorem gives us more information about the containment of $\boldsymbol{L}^1(\mathbb{R}^d)$ in $\boldsymbol{M}(\mathbb{R}^d)$:

L1-Basic **Theorem 18.** *(Basic properties of $\boldsymbol{L}^1(\mathbb{R}^d)$)*

- $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ *is a closed ideal within* $(\boldsymbol{M}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}})$. *It is a Banach algebra with a BAI (bounded approximate identity, the so-called Dirac sequences).*
- $\boldsymbol{L}^1(\mathbb{R}^d)$ *can be characterized as the closed subspace of with continuous shift, i.e. a bounded measure* μ *is of the form* $\mu = \mu_g$, *resp.* $\mu(f) = \int_{\mathbb{R}^d} f(x)g(x)dx$ *if and only if* $\|T_x\mu - \mu\|_{\boldsymbol{M}} \to 0$ *for* $x \to 0$.
- $\boldsymbol{L}^1(\mathbb{R}^d)$ *is* w^*-*dense in* $\boldsymbol{M}(\mathbb{R}^d)$, *in fact, for every* $\mu \in \boldsymbol{M}(\mathbb{R}^d)$ *there exists a tight (hence bounded) sequence* (f_k) *in* $\boldsymbol{L}^1(\mathbb{R}^d)$, *with* f_k *resp.* $\mu_{f_k} \to \mu$ *in the* w^*-*topology.*

Proof. The main arguments are the identification of the “internal convolution” within $\boldsymbol{M}(\mathbb{R}^d)$ with the usual convolution formula

conf-CcRd (181)
$$f * g(x) = \int_{\mathbb{R}^d} f(x-y)g(y)dy, \quad \text{for} \quad f, g \in \boldsymbol{C}_c(\mathbb{R}^d).$$

but also the external action of $\boldsymbol{M}(\mathbb{R}^d)$ on the homogeneous Banach space

$$\boldsymbol{M}(\mathbb{R}^d)_e = \{\mu \,|\, \|T_x\mu - \mu\|_{\boldsymbol{M}} \to 0 \text{ for } x \to 0.\}$$

The typical bounded approximate units are of the form $(\mathrm{St}_\rho g)_{\rho>0}$, for an arbitrary $g \in \boldsymbol{L}^1(\mathbb{R}^d)$ with $\hat{g}(0) = \int_{\mathbb{R}^d} g(t)dt = 1$.

It is easy to verify that this net is tight and tends to δ_0 in the w^*-topology. In a similar way one can approximate a finite and discrete measure by a linear combination of such Dirac sequences. Since the Dirac measures form a total subset in $\boldsymbol{M}(\mathbb{R}^d)$ with respect to the w^*-topology the w^*-density of $\boldsymbol{L}^1(\mathbb{R}^d)$ in $\boldsymbol{M}(\mathbb{R}^d)$ is established. $\square$

For any $h \in \boldsymbol{C}_c(\mathbb{R}^d)$ the compressed versions of h, i.e. the net $(\mathrm{St}_\rho h)_{h\to 0}$ is convergent to a multiple of δ_0[71]. For most applications h is normalized such that $\int_{\mathbb{R}^d} h(t)dt = 1$ (and for the proof it is enough to work with this normalization).

deltboxdil1 **Lemma 54.** *For any* $h \in \boldsymbol{L}^1(\mathbb{R}^d)$ *one has*

deloconv (182)
$$w^*\text{-}\lim{}_{\rho\to 0}\mathrm{St}_\rho\, \mu_h = \left(\int_{\mathbb{R}^d} h(t)dt\right) \cdot \delta_0.$$

[71] One should rather say, that the induced measures $\mu_{\mathrm{St}_\rho h}$ converge in the w^*-sense!

Proof. It should be enough to verify the claim[72] for $h \in \boldsymbol{C}_c(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d} h(t)dt = \hat{h}(0) = 1$.

Assume that $\mathrm{supp}(h) \subset B_R(0)$, i.e. that $h(x) = 0$ for $|x| \geq R$, for some positive R. Then clearly $\mathrm{St}_\rho h(y) = 0$ for $|y| \geq \rho R$.

Given now $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ and $\varepsilon > 0$ we have to find some $\rho_0 > 0$ such that

Stroest1 (183)
$$|\delta_0(f) - \int_{\mathbb{R}^d} f(x)\mathrm{St}_\rho h(x)dx| < \varepsilon \quad \forall \rho \in (0, \rho_0).$$

We just need the continuity of f at zero to find for that given $\varepsilon > 0$ some $\delta > 0$ such that $|f(0) - f(z)| \leq \varepsilon/\|h\|_1$ whenever $|z| \leq \delta$.

So we can choose $\rho_0 = \delta/R$ and obtain: whenever $0 < \rho < \rho_0$ one can write

$$\delta_0(f) = f(0) = f(0) \cdot 1 = f(0) \int_{\mathbb{R}^d} Strhoh(x)dx$$

but since the support of this last function is small we have

$$|f(0) - \int_{B_\delta(0)} f(z)\mathrm{St}_\rho h(z)dz| \leq \int_{B_\delta(0)} |f(0) - f(z)||\mathrm{St}_\rho h(z)|dz = \varepsilon/\|h\|_1 \cdot \|\mathrm{St}_\rho h\|_1 = \varepsilon,$$

and the proof is complete.

For general $h \in \boldsymbol{W}(\mathbb{R}^d)$ the result follows by further approximation. We leave the proof to the interested reader. Finally by the similar argument one can get it for $h \in \boldsymbol{L}^1(\mathbb{R}^d)$ (see e.g. H. Reiter's introductory chapters in [79]). ☐

Bounded measures operate not only on $\boldsymbol{C}_0(\mathbb{R}^d)$ by convolution but also on **any homogeneous Banach space** $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$. The proof of this fact is essentially based on the idea that it is enough to establish this fact for bounded discrete measures (which is easy) and then by showing that any sequence of discretizations of a given measure (where the diameter of the support of the corresponding BUPUs shrinks to zero) generates a sequence μ_k which is *(uniformly) tight* and bounded, but also produces a Cauchy sequence in $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ of the form $(\mu_k * f)$, for any given $f \in \boldsymbol{B}$. Obviously it makes sense to denote the limit (which does not depend on the choice of discretizations via BUPUs) by $\mu * f$ (although it is formally a new operation, and the "star" just introduced should be distinguished for a little while from the "known" star which denotes convolution within $\boldsymbol{M}(\mathbb{R}^d)$): $\mu * f \in \boldsymbol{B}$.

Of course the natural estimate that one has for discrete measures goes over in the limit to the general case, i.e. one has

nvHomogBan1 (184)
$$\|\mu * f\|_{\boldsymbol{B}} \leq \|\mu\|_{\boldsymbol{M}(\mathbb{R}^d)}\|f\|_{\boldsymbol{B}}, \quad \mu \in \boldsymbol{M}_b(\mathbb{R}^d), f \in \boldsymbol{B}.$$

Interesting and non-trivial examples are provided by the family of Segal algebras (in general, which we do not discuss here specifically), but especially for $\boldsymbol{B} = \boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d)$, or $\boldsymbol{B} = \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$, or $\boldsymbol{B} = \boldsymbol{S}_0(\mathbb{R}^d) = \boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1, \boldsymbol{\ell}^1)(\mathbb{R}^d)$. Let us emphasize that in all these cases we have subspace of $\boldsymbol{C}_0(\mathbb{R}^d)$ which are Banach spaces in their own right, continuously embedded into $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$. Hence the expression $\mu * f$ is the same (pointwise defined) that we have in $\boldsymbol{C}_0(\mathbb{R}^d)$, thanks to Prop.2.

Let us recall that the index set of a "net" is a directed set.

`http://de.wikipedia.org/wiki/Gerichtete_Menge`

Definition 38. A set is called a *directed* set (we use the symbol $\preceq$) if it is a relation on a set X satisfying the following three axioms, namely *reflexivity, transitivity and joint majorant property*:

[72] !?why, there is some approximation to be done, which - if done properly - still requires some "epsilontic", i.e. fiddling around with fractions of ε.

- $x \preceq x$, $\forall x \in X$; (reflexivity);
- $(x \preceq y) \wedge (y \preceq z) \Rightarrow (x \preceq z)$;
- $\forall x, y \in X\, \exists z \in X$ with $x \preceq z$ and $y \preceq z$.

By definition a *net* in set S (in our case typically some vector space) is a mapping from a directed set I into this set S, usually described as $(s_i)_{i \in I}$. Unlike a general *index family* the index has to be directed, but it is still possible that one has repetitions, in contrast to the definition of a set.

Compared to the concept of a sequence, which is a net with the index set being just the natural numbers $\mathbb{N}$ with their natural (total) order, $n \succeq n_0$ if and only if $n \geq n_0$ the concept of a subnet is more flexible:

subnet **Definition 39.** Given a directed set I and a net $(s_i)_{i \in I}$ we call a family $(\tau_j)_{j \in J}$ a *subnet* if $(J, \sqsupseteq)$ is another directed set, and there is some mapping $\eta : I \to J$ such that $\tau_j = s_{\eta(j)}$.

The mapping η does not have to be monotonous, i.e. we do not need the property that $i \succeq i_0$ implies that $\eta(i) \sqsupseteq \eta(i_0)$, but only the weaker condition of *cofinality*

cofinprop1 (185) $$\forall i_0 \in I \exists j_0 \in J \quad \text{such that } j \sqsupseteq j_0 \text{ implies } \eta(j) \succeq i_0.$$

Dirac-Conc **Lemma 55.** *A bounded net $(h_\alpha)_{\alpha \in I}$ within $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ is a BAI for $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ if the following property is satisfied: For every $\varepsilon > 0$ there exists some α_0 such that for $\alpha \succ \alpha_0$ one has:*

Dirac-eps (186) $$\left| \int_{B_\varepsilon(0)} h_\alpha(t)dt - 1 \right| \leq \varepsilon \quad \text{and} \quad \int_{|x| \geq \varepsilon} |h_\alpha(t)|dt \leq \varepsilon.$$

Clearly these conditions are satisfied by any family of the form $\mathrm{St}_\rho g$, for $\rho \to 0$, if $\int_{\mathbb{R}^d} g(x)dx = 1$.

Proof. Argument: Due to the density of $\boldsymbol{C}_c(\mathbb{R}^d)$ in $\boldsymbol{L}^1(\mathbb{R}^d)$ one can reduce the discussion to functions $k \in \boldsymbol{C}_c(\mathbb{R}^d)$, i.e.it is enough to show that $h_\alpha * k \mapsto k$ for any $k \in \boldsymbol{C}_c(\mathbb{R}^d)$. The second condition allows to restrict the attention to a net with common compact support K. Consequently one has $h_\alpha * k(x) \neq 0$ only for $x \in K + \mathrm{supp}(k)$. Furthermore we obtain $h_\alpha * k(x) = \int_{\mathbb{R}^d} h_\alpha(y) k(x-y) dy \mapsto$ and so on...

That the described conditions are met by elements of a family $(\mathrm{St}_\rho g)_{\rho \to 0}$ if $\hat{g}(0) = \int_{\mathbb{R}^d} g(x)dx = 1$ is quite clear, essentially because it is obvious if g has compact support. The rest is the usual approximation trick. $\square$

Remark 45. Of course one can also consider $\boldsymbol{M}(\mathbb{R}^d)$ as a Banach module over the Banach algebra $\boldsymbol{L}^1(\mathbb{R}^d)$ (with respect to convolution). Then the $\boldsymbol{L}^1(\mathbb{R}^d)$-essential part of $\boldsymbol{M}(\mathbb{R}^d)$ is equal to $\boldsymbol{L}^1(\mathbb{R}^d)$ itself.

On the other hand we can consider $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$ as a Banach module over $\boldsymbol{L}^1(\mathbb{R}^d)$ (again with respect to convolution), and then this is an essential Banach module.

8. Convolution for a homogeneous Banach space

So we have the following theorem:

MbConvHomog **Theorem 19.** *Let $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ be a homogeneous Banach spaces of tempered distributions, i.e. a Banach space which is sandwiched between $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ and $\boldsymbol{\mathcal{S}}'(\mathbb{R}^d)$, containing $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ as a dense subspace, and with isometric translations, i.e.*

isomtrans1 (187) $$\|T_z f\|_{\boldsymbol{B}} = \|f\|_{\boldsymbol{B}}, \quad \forall f \in \boldsymbol{B}.$$

Then $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ is a Banach module over $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ with respect to convolution. The bilinear pairing $(\mu, f) \mapsto \mu * f$ is w^*-norm to norm continuous (for bounded, tight subsets of $\boldsymbol{M}_b(\mathbb{R}^d)$), and satisfies

convest07
$$\|\mu * f\|_{\boldsymbol{B}} \leq \|\mu\|_{\boldsymbol{M}_b} \|f\|_{\boldsymbol{B}}, \quad \forall \mu \in \boldsymbol{M}_b(\mathbb{R}^d), f \in \boldsymbol{B}. \tag{188}$$

If viewed as a Banach convolution module over $\big(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\big)$ it is an essential Banach module.

Instead of assuming the embedding $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}}) \hookrightarrow \boldsymbol{\mathcal{S}}'(\mathbb{R}^d)$ one might assume a continuous embedding $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}}) \hookrightarrow \boldsymbol{L}^1_{loc}(\mathbb{R}^d)$, and continuity of the shift, i.e. the condition

contshift04
$$\lim_{z \to 0} \|T_z f - f\|_{\boldsymbol{B}} = 0, \quad \forall f \in \boldsymbol{B}. \tag{189}$$

This corresponds exactly to the concept of *homogeneous Banach spaces* as introduced in the book of Y. Katznelson ([63]). The assumption made there is limiting the attention to BF-spaces, i.e. Banach spaces of locally integrable functions on a LCA group. Equivalently, one could talk about Banach space of absolutely continuous Radon measures, because one assume that for any compact set $K \subset \mathbb{R}^d$ one has a constant, such that

locLiest01
$$\int_K |f(x)| dx = \|f \cdot \mathbf{1}_K\|_{\boldsymbol{L}^1} \leq C_K \|f\|_{\boldsymbol{B}} \quad \forall f \in \boldsymbol{B}. \tag{190}$$

In fact, as a consequence of (187) any such homogeneous Banach space is continuously embedded into $\big(\boldsymbol{W}(\boldsymbol{L}^1, \boldsymbol{\ell}^\infty)(\mathbb{R}^d), \|\cdot|\boldsymbol{W}(\boldsymbol{L}^1, \boldsymbol{\ell}^\infty)(\mathbb{R}^d)\|\big)$ (Exercise). One of the equivalent norms on $\boldsymbol{W}(\boldsymbol{L}^1, \boldsymbol{\ell}^\infty)(\mathbb{R}^d)$ is the following one, for any compact set with non-void interior:

WLilincontn
$$\sup_{x \in \mathbb{R}^d} \int_{x+K} |f(x)| dx < \infty. \tag{191}$$

Proof. We have to show that the family $(D_\Psi \mu *_{\boldsymbol{B}} f)_{|\Psi| \to 0}$ is a Cauchy-net, or more precisely: For $\varepsilon > 0$ there exists $\delta > 0$ such that for $|\Psi| < \delta, |\Phi| < \delta$ implies

diffest11
$$\|D_\Psi \mu *_{\boldsymbol{B}} f - D_\Phi \mu *_{\boldsymbol{B}} f\|_{\boldsymbol{B}} < \varepsilon. \tag{192}$$

To fix notations let us assume that $\mathrm{supp}(\psi_i) \subseteq B_\delta(\xi_i)$ and $\mathrm{supp}(\phi_j) \subseteq B_\delta(\eta_j)$, for some $\delta \leq 1$.

We should think of these expressions as of Riemann sums for $\int_{\mathbb{R}^d} T_x f d\mu(x)$. Hence the closeness of two Riemannian sums works through a "joint refinement", in our case, by taking the "joint BUPU", i.e. the family

$$(\psi_i \cdot \phi_j)_{(i,j) \in I \times J}.$$

Practically speaking many of these terms will be zero because for each $i \in I$ only finitely many indices $j \in J$ are such that $\psi_i \cdot \phi_j \neq 0$. Let us call the set of all *relevant indices* $R \subset I \times J$, so we actually can look at the BUPU

BUPUIJ
$$(\psi_i \cdot \phi_j)_{(i,j) \in R}. \tag{193}$$

In fact,

$$\sum_{(i,j) \in R} \psi_i \phi_j = \sum_{i \in I} \sum_{j \in J} \psi_i \phi_j = 1$$

One the other hand, $(i,j) \in R$ implies that

suppestim1
$$\mathrm{supp}(\psi_i) \cap \mathrm{supp}(\phi_j) \neq \emptyset \quad \Rightarrow \quad B_\delta(\xi_i) \cap B_\delta(\eta_j) \neq \emptyset \quad \Rightarrow \quad |\xi_i - \eta_j| < 2\delta. \tag{194}$$

Let us write $J_i = \{j \in J \,|\, \psi_i \cdot \phi_j \neq 0\}$, and $I_j = \{i \in I \,|\, \psi_i \cdot \phi_j \neq 0\}$, hence

$$\sum_{(i,j) \in R} \psi_i \phi_j = \sum_{i \in I} \sum_{j \in J_i} \psi_i \phi_j = 1 = \sum_{j \in J} \sum_{i \in I_j} \psi_i \phi_j.$$

The support size condition for $\psi_i\phi_j, (i,j) \in R$ is clear since

$$\mathrm{supp}(\psi_i\phi_j) \subset \mathrm{supp}(\psi_i) \cap \mathrm{supp}(\phi_j) \subset B_\delta(\xi_i) \cap B_\delta(\eta_j), \quad \forall (i,j) \in R.$$

Obviously $\mu(\psi_i) = \sum_{j\in I_j} \mu(\psi_i\phi_j)$ and thus we can write

$$D_\Psi\mu = \sum_{i\in I} \mu(\psi_i)\delta_{x_i} = \sum_{i\in I}\sum_{j\in J_i} \mu(\psi_i\phi_j)\delta_{x_i} = \sum_{(i,j)\in R} \mu(\psi_i\phi_j)\delta_{x_i}$$

and correspondingly

$$D_\Phi\mu = \sum_{j\in J}\mu(\phi_j)\delta_{\eta_j} = \sum_{(i,j)\in R} \mu(\psi_i\phi_j)\delta_{\eta_j}.$$

Let us set $c_{i,j} = \mu(\psi_i\phi_j)$, for $(i,j) \in R$. Then we have

$$\sum_{(i,j)\in R} |c_{i,j}| \le \sum_{i\in I} |\mu(\psi_i)| \le \sum_{i\in I} \|\mu\,\psi_i\|_{\boldsymbol{M}_b} = \|\mu\|_{\boldsymbol{M}_b},$$

and

DPsiPhi1
$$D_\Psi\mu = \sum_{(i,j)\in R} c_{i,j}\delta_{x_i}, \quad D_\Phi\mu = \sum_{(i,j)\in R} c_{i,j}\delta_{\eta_j}. \tag{195}$$

Applying the above description of $D_\Psi\mu$ and $D_\Phi\mu$ to the operator level we obtain from (195)

$$D_\Psi\mu * f = \sum_{(i,j)\in R} c_{i,j}T_{\xi_i}f, \quad \text{respectively} \quad D_\Phi\mu * f = \sum_{(i,j)\in R} c_{i,j}T_{\eta_j}f.$$

This enables us to go for an estimate of the difference in $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$:

DPsiPhiD2
$$D_\Psi\mu *_{\boldsymbol{B}} f - D_\Phi\mu *_{\boldsymbol{B}} f = \sum_{(i,j)\in R} c_{i,j}[T_{\xi_i}(f - T_{\eta_j-\xi_i}f)] \tag{196}$$

By choosing $\delta > 0$ such that

shiftest02
$$\|f - T_z f\|_{\boldsymbol{B}} < \varepsilon_1 := \varepsilon/\|\mu\|_{\boldsymbol{M}} \quad \text{for } |z| < 2\delta \tag{197}$$

we obtain the final estimate (using the translation invariance of the norm in $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$):

DPsiPhiest
$$\|D_\Psi\mu *_{\boldsymbol{B}} f - D_\Phi\mu *_{\boldsymbol{B}} f\|_{\boldsymbol{B}} \le \sum_{(i,j)\in R} |c_{i,j}|\|f - T_{\eta_j-\xi_i}\|_{\boldsymbol{B}} \le \varepsilon_1\|\mu\|_{\boldsymbol{M}} = \varepsilon. \tag{198}$$

This shows that the family $(D_\Psi\mu*f)_{(\Psi,(\xi_i))}$ is a Cauchy net inside the Banach space $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$, for every $f \in \boldsymbol{B}$. Recall that the net structure of the index set I is more or less the same as we are used from the theory of the Riemann integral: We say that a BUPU Ψ of size $|\Psi| = \delta_1$ is *more fine* than another BUPU Φ of size $|\Phi| = \delta_2$, and write $\Psi \preceq \Phi$ (in words: Ψ *is more fine than* Φ) if $\delta_1 \le \delta_2$[73]. We have simply shown that two such expression are close to each other whenever both of them are fine enough.

Using the completeness of the Banach space and the fact that every Cauchy-net has a limit one can find that there exists some element in $\boldsymbol{B}$ which arises as the limit of this Cauchy net. It is obvious that we want to call it $\mu * f$, but in order to be careful with notations we write

$$\mu *_{\boldsymbol{B}} f = lim_{\,|\Psi|\to 0} D_\Psi\mu * f, \quad \mu \in \boldsymbol{M}_b(\mathbb{R}^d), f \in \boldsymbol{B}. \tag{199}$$

[73]Again, as with Riemannian sums one should rather say: the index set consists of an indexed family of [well distinguished] pairs $(\psi_i, \xi_i)_{i\in I}$, because the size description of for Ψ is given via the condition $\mathrm{supp}(\psi_i) \subseteq B_\delta(\xi_i)$ which of course depends on the choice of the "quasi-center" ξ_i. This is exactly the situation that we have for the family $(\psi_i\phi_j)_{(i,j)\in R}$, which is viewed as two such BUPUs, one times with the family of centers $(\xi_i)_{i\in I}$ and the other time with centers $(\eta_j)_{j\in J}$.

Note that one has for any Ψ the expression $D_\Psi \mu *_{\boldsymbol{B}} f = \sum_{i \in I} \mu(\psi_i) T_{\xi_i} f$ and thus

convest07

$$\|D_\Psi \mu *_{\boldsymbol{B}} f\|_{\boldsymbol{B}} \leq \sum_{i \in I} |\mu(\psi_i)| \|T_{\xi_i} f\|_{\boldsymbol{B}} = \|f\|_{\boldsymbol{B}} \sum_{i \in I} |\mu(\psi_i)| \leq \|\mu\|_{\boldsymbol{M}_b} \|f\|_{\boldsymbol{B}}. \tag{200}$$

Clearly, this estimate goes over to the limit (in $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$) and thus (184) is valid.

Note that we can follow the same argument as in the case of our action on $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$ in order to verify that one has of course

convDirac

$$\delta_z *_{\boldsymbol{B}} f = T_z f, \quad f \in \boldsymbol{B}, z \in \mathbb{R}^d, \tag{201}$$

in other words, the "new convolution" is really just the *integrated group action* arising from the regular representation of the group $\mathbb{R}^d$ on the homogeneous Banach space, noting that (obviously) the group $(\mathbb{R}^d, +)$ is naturally isomorphic to the group of operators $(T_x)_{x \in \mathbb{R}^d}$ for any Banach space of functions or distributions on $\mathbb{R}^d$ (e.g. for Banach spaces $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ continuously embedded into $\boldsymbol{\mathcal{S}}'(\mathbb{R}^d)$, the space of tempered distributions). □

nsistMBact1 *Remark* 46. It is a natural question to ask, what happens, if we take two different homogeneous Banach spaces of tempered distributions, let us call them $(\boldsymbol{B}^1, \|\cdot\|^{(1)})$ and $(\boldsymbol{B}^2, \|\cdot\|^{(2)})$, and thus we have to consider to formally the expressions $\mu *_{\boldsymbol{B}^1} f$ and $\mu *_{\boldsymbol{B}^2} f$, for any given $f \in \boldsymbol{B}^1 \cap \boldsymbol{B}^2$. Are these elements the same.

The answer to this question is in fact easy and affirmative: Just note that we can endow $\boldsymbol{B}^1 \cap \boldsymbol{B}^2$ with the natural norm

sumnorm2

$$\|f\|_{\boldsymbol{B}^1 \cap \boldsymbol{B}^2} = \|f\|_{\boldsymbol{B}^1} + \|f\|_{\boldsymbol{B}^2}. \tag{202}$$

Then this space is again a Banach space of tempered distributions and consequently $\mu *_{\boldsymbol{B}^1 \cap \boldsymbol{B}^2}$ is well defined. But convergence in this smaller space implies of course that one also has convergence (to the same limit) in each of the given spaces $(\boldsymbol{B}^1, \|\cdot\|^{(1)})$ and $(\boldsymbol{B}^2, \|\cdot\|^{(2)})$, and hence it gives the same element.

In particular, one can take for example $f \in \boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d)$, which belongs to both $\boldsymbol{C}_0(\mathbb{R}^d)$ and $\boldsymbol{L}^1(\mathbb{R}^d)$, but the above reasoning implies that the resulting action can be denoted as $\mu * f$, without specification of the interpretation of the homogeneous Banach spaces chosen.

Remark 47. Having established that $\mu * f$ defines the action of $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ on $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ via convolution (and of course we obtain operators which commute with translations in this way) we still cannot claim at this point that one has associativity, i.e. the important rule

convassoc05

$$\mu_1 *_{\boldsymbol{B}} (\mu_2 *_{\boldsymbol{B}} f) = (\mu_1 *_{\boldsymbol{M}_b} \mu_2) *_{\boldsymbol{B}} f, \quad \forall \mu_1, \mu_2 \in \boldsymbol{M}_b(\mathbb{R}^d), f \in \boldsymbol{B}. \tag{203}$$

Since this rule is obvious for pairs of Dirac measures, hence for finite discrete measures, and (by taking limits) for arbitrary discrete measures in $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ one can take limits and obtain the validity of (203) for any two bounded measures μ_1, μ_2.

Remark 48. Note that the convergence issue, and thus the definition of $\mu *_{\boldsymbol{B}} f$ in $\boldsymbol{B}$ requires only the completeness of $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ it is the embedding into a larger (topological) vector space (like $\boldsymbol{\mathcal{S}}'(\mathbb{R}^d)$ in our case) which makes sure that this definition does *not depend on the interpretation*, whether f belongs to $\boldsymbol{B}^1$ of $\boldsymbol{B}^2$, which is important for applications (it as a matter of *consistency*!

So let us emphasize: Assume that $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ is embedded in any topological vector space (meaning we have a *continuous embedding*, then of course convergence in $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ implies convergence in that bigger space. So instead of working with the intersection $\boldsymbol{B}^1 \cap \boldsymbol{B}^2$ one could also use this embedding of two *compatible Banach spaces* (using the terminology of interpolation space theory, see [8]) in order to find that the meaning of e.g. $\mu * f$ is the same, independent of the choice of $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$, as long as the general circumstances (convergence etc.) are justified.[74]

netssequ *Remark* 49. As in the case of the Riemann integral one has two choices: Either one defines the integral as the limit of general Riemannian sums, more or less in the way we have done it above, or one prefers to avoid the "complex" notion of nets (the explicit description is usually avoided in our analysis courses) and makes use of simple refinements, in the extreme case via bisection, and perhaps even choosing just midpoints of each interval, to really have a concrete Cauchy sequence (in $\mathbb{R}$), and then use the closeness of two Riemannian sums (based on the *uniform continuity* of a continuous function over $[a, b]$) to show that also any other Riemannian sum resorting from a sufficiently fine decomposition of the domain, i.e. the interval $[a, b]$, will give a similar number, so that one has convergence of the Cauchy net of Riemannian sums. Not that in this case one makes use of the indicator functions of the intervals arising from the decomposition, so $\psi_k = \mathbf{1}_{[z_{i-1}, z_i)}$, and thus (for the Lebesgue measure one has $\mu(\psi) = z_i - z_{i-1}$, and the usual points ξ_i correspond to our quasi-centers ξ_i, and thus

$$D_\Psi \mu(f) = \sum_{i \in I} f(\xi_i)(z_i - z_{i-1})$$

are actually the well known Riemannian sums!

ptwae1 *Remark* 50. The above considerations also imply, for example, that for the situation, where pointwise convolution exists (at least in the almost everywhere sense), these (somewhat restricted notions of convolutions) are also *compatible* with the concept introduced above.

As an example. Let $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ be a classical homogeneous Banach space (in the spirit of Y. Katznelson, **ka76 : [63]**), then the ambient topological vector space is $\boldsymbol{L}^1_{loc}(\mathbb{R}^d)$, the space of (equivalence classes of) measurable, locally Lebesgue integrable functions. In the papers by the author (hgfei) such spaces are generally described as *BF-spaces*.

In a similar way one can make claims for

ptwconv1 *Remark* 51. As an example of the above statement one has: Assume that two function $f_1, f_2 \in \boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ are given. Then both of them can be viewed as members of $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ (for any Banach space of tempered distributions), but also as elements of $\boldsymbol{M}_b(\mathbb{R}^d)$ or $\boldsymbol{\mathcal{S}}'(\mathbb{R}^d)$ (via the usual embedding $f \mapsto \mu_f$ resp. σ_f). So the pairing can be taken in many different ways.

Then one has of course

consist008 (204) $$\mu_{f_1} * f_2 = \mu_{f_2} * f_1 = \mu_{f_1} *_{\boldsymbol{C}_0} f_2 = \mu_{f_2} *_{\boldsymbol{C}_0} f_1$$

and so on, and of course it coincides (as a measure or distribution) with the Schwartz function $f_1 * f_2$, given in the usual way

convptw3 (205) $$[f_1 *_{pw} f_2](x) = \int_{\mathbb{R}^d} f_2(x-y) f_1(y)\, dy, \quad x \in \mathbb{R}^d.$$

sh71 : [91], **re71 : [80]**,

Let us summarize our findings:

HomBanMod1 **Theorem 20.** *Given any Banach space of tempered distributions with isometric translation, it is an Banach module over $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ with respect to convolution. It can also be viewed as an essential Banach convolution module over $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$. In particular one has: for any bounded approximate identity $(e_\alpha)_{\alpha \in I}$ in $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ one has*

BAIappr05 (206) $$\lim_\alpha \|e_\alpha * f - f\|_{\boldsymbol{B}} = 0, \quad \forall f \in \boldsymbol{B}.$$

[74] This is one reason why it is so important to look at *families of compatible Banach spaces* (with a common *universe*) and not just an individual Banach space $(\boldsymbol{B}^1, \|\cdot\|^{(1)})$ or $(\boldsymbol{B}^2, \|\cdot\|^{(2)})$!

$$\boldsymbol{L}^1(\mathbb{R}^d) * \boldsymbol{B} = \boldsymbol{B}$$

9. Tight subsets

WARNING: The concept of tightness has been used already on and off!

A given $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ is of course "essentially concentrated" on a compact set (and uniformly small outside a sufficiently large compact set, by definition). We also have shown that a functional $\mu \in \boldsymbol{M}(\mathbb{R}^d)$ is having most of its "mass" sitting within a compact set, while its action outside of this compact set is small. Indeed, since for any BUPU Φ we have $\mu = \sum_{iinI} \phi\mu$ as absolutely convergent sum the tails μ minus a large partial sum is small in the $\boldsymbol{M}(\mathbb{R}^d)$-sense.

Next we want to extend this "concentration over compact sets" concept to general bounded subsets of $\boldsymbol{M}(\mathbb{R}^d)$ (and later other functional spaces):

ght-measdef **Definition 40.** A bounded subset $W \subset \boldsymbol{M}(\mathbb{R}^d)$ is called *(uniformly) tight* if for every $\varepsilon > 0$ there exists $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ such that $\|\mu - k \cdot \mu\|_{\boldsymbol{M}} \leq \varepsilon$ for all $\mu \in W$.

In a similar way we define tightness in $\boldsymbol{C}_0(\mathbb{R}^d)$:

Definition 41. A bounded subset $H \subset \boldsymbol{C}_0(\mathbb{R}^d)$ is called *(uniformly) tight* if for every $\varepsilon > 0$ there exists $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ such that $\|h - k \cdot h\|_\infty \leq \varepsilon$ for all $h \in H$.

Note: For a general $\boldsymbol{C}_0(\mathbb{R}^d)$ module $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ one can define tightness as follows:

Definition 42. A bounded subset $H \subset \boldsymbol{B}$ is called *(uniformly) tight* in $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ if for every $\varepsilon > 0$ there exists $h \in \boldsymbol{C}_c(\mathbb{R}^d)$ such that $\|h - k \cdot h\|_\infty \leq \varepsilon$ for all $h \in H$.

The concept of *tightness* plays a big role in the characterization of *relatively compact subsets* in function spaces and hence is also important for the description of compact operators.

convtight2 **Theorem 21.** *Assume that W is a tight set within $(\boldsymbol{M}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}})$ and that H is a tight subset within $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$. Then*

$$W * H = \{\mu * h \,|\, \mu \in W, h \in H\}$$

is a tight subset in $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$.

cf. the "compactness paper" [32] p.307 (bottom):
http://univie.ac.at/nuhag-php/bibtex/open_files/fe84_compdist.pdf

Proof. Indeed, for any plateau-function τ which satisfies $\tau(x) \equiv 1$ on $\mathrm{supp}(k^1) + \mathrm{supp}(k^2)$, hence the following estimate holds:

$$(1-\tau)(f^1 * f^2) = (1-\tau)(f^1 * f^2 - f^1 k^1 * f^2 k^2);$$

$$(1-\tau)(\mu * f) = (1-\tau)(\mu * f - \mu k^1 * f k^2).$$

Applying norms to both sides and using the triangle inequality we obtain the following estimate in the sup-norm:

$$\|(1-\tau)(\mu * f)\| = \|(1-\tau)(\mu * f - \mu k^1 * f k^2)\| \leq$$

$$\leq \|(1-\tau)\| \|(1-k^1)\mu * f\| + \|k^1 \mu * (1-k^2) f\| \leq$$

$$\leq \|(1-\tau)\| \|(1-k^1)\mu\|_{\boldsymbol{M}} \|f\| + \|k^1\| \|\mu\|_{\boldsymbol{M}} \|(1-k^2) f\|.$$

□

10. The Fourier transform for $LiRd$

The Fourier transform maps $\boldsymbol{M}(\mathbb{R}^d)$ into $\boldsymbol{C}_{ub}(\mathbb{R}^d)$. It will be seen as a non-expansive Banach algebra homomorphism from the closed ideal $\boldsymbol{L}^1(\mathbb{R}^d)$ into the closed ideal $\boldsymbol{C}_0(\mathbb{R}^d)$ of $\boldsymbol{C}_{ub}(\mathbb{R}^d)$ (this result is usually known as Riemann-Lebesgue Lemma).

The range of the Fourier transform is a dense subalgebra (closed under complex conjugation), due to the "locally compact version" of the Stone-Weierstrass theorem.

Recall the standard version of the Stone-Weierstrass theorem:

eierstr-thm **Theorem 22.** *Let $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ be a Banach algebra within $\boldsymbol{C}(X)$, where X is some compact topological space. Then $\boldsymbol{A}$ is dense with respect to the uniform norm if $\boldsymbol{A}$ contains the constant functions, is closed under conjugation, and separates points, i.e., if for any pair of points $x_1, x_2 \in X$, with $x_1 \neq x_2$ there exists some $f \in \boldsymbol{A}$ such that $f(x_1) \neq f(x_2)$.*

Since $\mathcal{F}\boldsymbol{L}^1$ does not have a unit (for pointwise multiplication), due to the fact that $\boldsymbol{L}^1(\mathbb{R}^d)$ does not contain a unit (as the unit with respect to convolution is the Dirac measure δ_0, which *cannot* be approximated in $(\boldsymbol{M}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}})$ from within $\boldsymbol{C}_c(\mathbb{R}^d)$), one has to modify the above result to the locally compact case, by "adding" the constant functions, and replacing $\mathbb{R}^d$ by its Alexandroff (one-point) compactification X of $\mathbb{R}^d$. Indeed, $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ can be identified with a closed subalgebra of all continuous functions vanishing "at infinity". In fact, if $C_0 + h$ in $\boldsymbol{C}(X)$ is approximated by a sequence of the form $C_n + f_n$, then $|C_n - C_0| \to 0$ for $n \to \infty$, so $\|f_n - h\|_\infty$ for $n \to \infty$ (details left to the reader).

FL1-algprop **Lemma 56.** $\left(\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_{\mathcal{F}\boldsymbol{L}^1}\right)$ *is a Banach algebra with respect to the standard norm $\|\hat{f}\|_{\mathcal{F}\boldsymbol{L}^1} := \|f\|_1$ for $f \in \boldsymbol{L}^1(\mathbb{R}^d)$, which is closed under translation, modulations and dilations (in fact, with continuous dependence on the shift resp. dilation parameters). It is also closed under complex conjugation.*

Proof. All these properties follow from the algebraic properties of the Fourier transform (intertwining of modulation and translation operators), and intertwining of St_ρ and D_ρ operators, as well as $f \mapsto f^*$ with $f^*(x) = \overline{f(-x)}$, which intertwines with ordinary conjugation of $\hat{f}$. □

For $\mathbb{R}^d$ one can of course use alternative arguments, e.g. by observing that certain functions, such as the Schwartz functions, belong to the Fourier algebra, so there is enough richness in the function space $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ in order to show the density of $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ in $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$, but the above argument applies to general LCA groups.

FliR-dens1 **Lemma 57.** *Let $k \in \boldsymbol{C}_c(\mathbb{R})$ and Ψ the family of B-splines of any fixed order $s \in \mathbb{N}$ (hence a BUPU of size $s/2$). Then for every $\delta > 0$ the spline-approximation $Sp_{\delta\Psi}k$ is a finite sum of shifted B-spline functions, hence a function having $(s-1)$ continuous derivatives. Also for $s \geq 2$ the B-splines are continuous and belong to $\mathcal{F}\boldsymbol{L}^1(\mathbb{R})$, because their Fourier transforms are of the form sinc^s. In particular, $\boldsymbol{C}^m$ as well as $\mathcal{F}\boldsymbol{L}^1$ are dense subspaces of $\boldsymbol{L}^1(\mathbb{R})$ (and other $\boldsymbol{L}^p$-spaces, for $p < \infty$).*

On of the central statements concerning the Fourier transform is *Plancherel's theorem*, stating that the Fourier transform can be considered as a unitary linear automorphism of the Hilbert space $\boldsymbol{L}^2(\mathbb{R}^d)$ onto itself. This is in complete analogy to the statement that for the case of $\mathbb{C}^n$ the Discrete Fourier transform (often realized in the form of the FFT) is a change of bases from the orthonormal basis of unit vectors to the orthogonal system of pure frequencies. Since the vectors representing the pure frequencies (which are exactly the joint eigenvalues to all the translation operators) are of absolute value one, they are all of norm $\sqrt{n}$ the inverse FFT is essentially the conjugate (transpose) of the Fourier matrix, with a compensating factor of the form $1/n$. The advantage of our normalization in the continuous case (with the factor 2π as

NC1632A.pdf

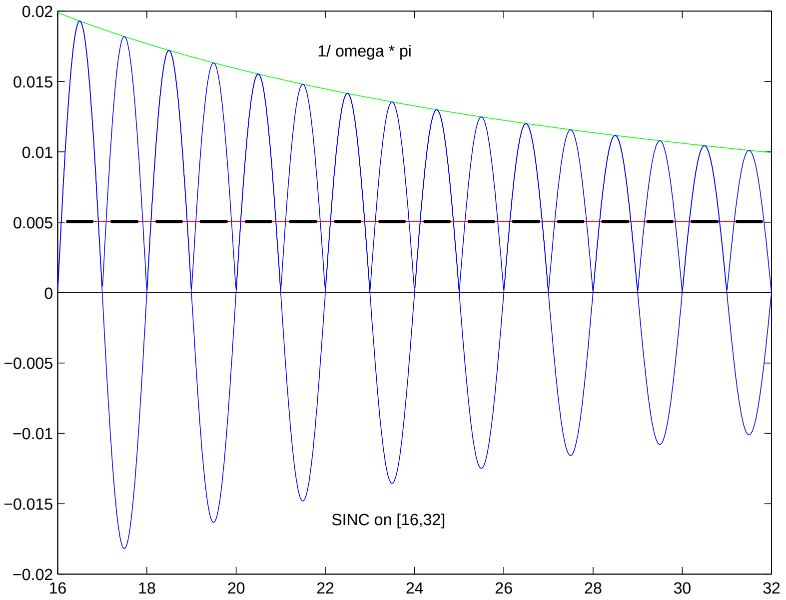


FIGURE 11. SINC1632A.pdf

part of the exponent) has the advantage that the inverse Fourier transform will come in the form

inv-Fourdef

$$h(t) = \int_{\mathbb{R}^d} \hat{f}(s) e^{2\pi i t\cdot s} dt \tag{207}$$

There is a lot to be said about the Fourier inversion. Above all it has to be noted that it is *wrong* to view it as a formula, allowing to recover the function $f \in \boldsymbol{L}^1(\mathbb{R}^d)$ from its Fourier transform pointwise, simply because there is no guarantee that the Fourier transform of such a function should be integrable. The box-car function $\mathbf{1}_{[-1/2,1/2]}$ is a typical example, whose Fourier transform is the so-called SINC-function, defined as

SINCdef1

$$\mathrm{SINC}(t) := \frac{\sin \pi t}{\pi t}, \quad t \in \mathbb{R}. \tag{208}$$

This is where classical analysis is bringing in a lot of Fourier analysis.

As an illustration for the poor decay of the SINC-function (implying SINC $\notin \boldsymbol{L}^1(\mathbb{R})$!), let us have a look at the graph of this function:

Since the integral definition of the FT and its inverse do not apply to general functions $\hat{f} \in \mathcal{F}\boldsymbol{L}^1$, part of the discussion of the **Fourier-Plancherel Theorem** is concerned with technical questions around problems of this kind (how to overcome lack of integrability, e.g. by applying so-called *summability* methods, which are a generalization of the idea of an infinite integral, taken as limit of finite integrals).

RieLemm1

Lemma 58. *It is enough to verify that for some dense subspace $\boldsymbol{B}$ of $\boldsymbol{L}^2(\mathbb{R}^d)$ within $\boldsymbol{L}^1(\mathbb{R}^d) \cap \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) \subset \boldsymbol{C}_0(\mathbb{R}^d)$ one can find that the mapping $f \mapsto \hat{f}$ is well defined and isometric, and with dense range, in order to be able to extend the "classical" Fourier transform and its inverse (given by the integral) to an isometry from $\boldsymbol{L}^2(\mathbb{R}^d)$ onto itself.*

NCplot4.pdf

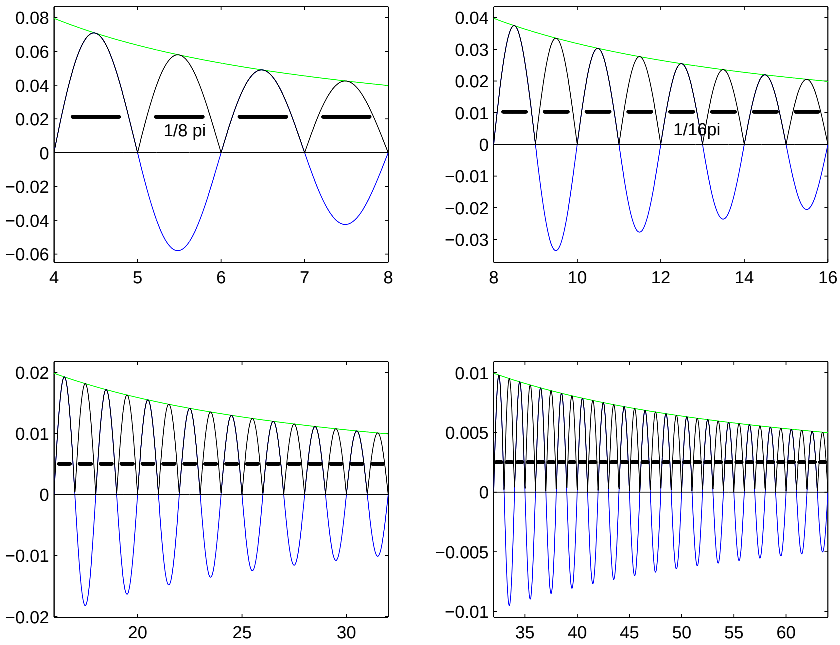


FIGURE 12. SINCplot4.pdf

Proof. Since we assume that $\boldsymbol{B} \subseteq \boldsymbol{L}^1(\mathbb{R}^d) \cap \boldsymbol{C}_0(\mathbb{R}^d) \cap \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ one can claim that the direct and the inverse Fourier transform given via (absolutely convergent) Riemannian integrals is valid. For the rest one only has to show that for an arbitrary $f \in \boldsymbol{L}^2(\mathbb{R}^d)$ and any sequence f_n, with $f_n \in \boldsymbol{B}$ with $\|f - f_n\|_2 \to 0$ for $n \to \infty$ one finds that $\|\hat{f} - \hat{f}_n\|_2 = \|f - f_n\|_2 \to 0$, so by the completeness of $\boldsymbol{L}^2(\mathbb{R}^d)$ the Cauchy sequence $\hat{f}_n$ must have a limit, which may be denoted (by a so-called abuse of language) as the Fourier transform $\hat{f}$ of f. The extended (still isometric) mapping has dense range according to our assumptions, and therefore the same argument can be applied to the inverse Fourier transform in order to realize that the extended mapping (often called the *Fourier-Plancherel* or just *Plancherel* transform defines an isometric automorphism of $\boldsymbol{L}^2(\mathbb{R}^d)$. Due to the polarization identity

polariz2

$$\langle f, g \rangle = \sum_{k=0}^{3} i^k \|f + i^k g\|_2^2. \tag{209}$$

such a mapping also preserves scalar products in general. □

For the proof of Plancherel's theorem one may use $\boldsymbol{B} = \boldsymbol{L}^1(\mathbb{R}^d) \cap \boldsymbol{C}_0(\mathbb{R}^d) \cap \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ or the linear span of all the time-frequency shifted and dilated version of a Gauss function (we do not require any norm on $\boldsymbol{B}$). Ideally one can or should use a space of functions which is invariant under the Fourier transform. Details have been given in the FA (= functional analysis) course WS0506 by HGFei.

Of course the extended Fourier transform is still compatible with convolution resp. pointwise multiplication. In other words, convolution on one side of the FT goes into pointwise product on the other side (and vice versa). As a consequence one obtains a characterization of $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$: A function belongs to $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ if and only if it can be written as a convolution product of two functions in $\boldsymbol{L}^2(\mathbb{R}^d)$, i.e. $h \in \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ if and only if there exist two functions $f, g \in \boldsymbol{L}^2(\mathbb{R}^d)$ such that $h = f * g$. The direct direction (i.e. convolution products are in $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$) is easy, because

IFLICO2.eps

[ht]

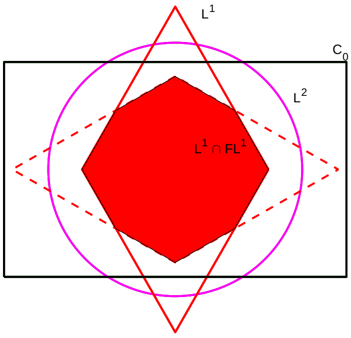


FIGURE 13. LIFLICO2.eps: The red domain represents the domain of the inversion theorem, using Lebesgue integration

their Fourier transforms give a function which is a pointwise product of two $\boldsymbol{L}^2$-functions, and hence according to the Cauchy-Schwarz inequality $\hat{h} = \hat{f} \cdot \hat{g} \in \boldsymbol{L}^2 \cdot \boldsymbol{L}^2 \subseteq \boldsymbol{L}^1$, or equivalently $h \in \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$.

The easiest argument for the converse is again on the Fourier transform side. Write $\hat{h} \in \boldsymbol{L}^1(\mathbb{R}^d)$ as a pointwise product of two $\boldsymbol{L}^2$-functions. If $\hat{h}$ was non-negative there is a natural solution to this problem, just take $\sqrt{\hat{h}}$. If $\hat{h}$ is a complex-valued function, one can apply this trick only to $|\hat{h}|$, and can assign the phase factor to one of the two non-negative square roots (details are left to the reader).

We are deriving (a weak form) of the Fourier inversion theorem from (152).

ier_invers2

Theorem 23. *Assume $f \in \boldsymbol{L}^1 \cap \boldsymbol{C}_0(\mathbb{R}^d)$ with (the additional assumption that) $\hat{f} \in \boldsymbol{L}^1(\mathbb{R}^d)$* [75] *then for every $t \in \mathbb{R}^d$ one has:*

$$f(t) = \int_{\mathbb{R}^d} e^{2i\pi st} \hat{f}(s) ds.$$

Proof. Recall that we have the consistency of the Fourier transform for $\boldsymbol{L}^1(\mathbb{R}^d)$-functions f with the corresponding measure μ_f, since

$$\hat{f}(s) = \int_{\mathbb{R}^d} f(t)\overline{\chi_s(t)} dt$$

In order to retrieve the inversion formula we just have to use (152) with the following choices: □

FTmeasinj

Lemma 59. *$\hat{\mu} = 0$ implies $\mu = 0$.*

Proof. If $\hat{\mu} = 0$, we find from (152) that $\mu(\hat{\nu}) = 0$ for every $\nu \in \boldsymbol{M}_b(\mathbb{R}^d)$, in particular $\mu(\hat{f}) = 0$ for all $f \in \boldsymbol{L}^1(\mathbb{R}^d)$. Since $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ is a dense subspace of $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$ this implies that $\mu = 0$. □

Arguments why $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ is dense in $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$: either using Stone-Weierstrass, or by approximation of any $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ by piecewise linear functions, and the observation that these functions are sums of triangular functions, which in return belong to the *Fourier algebra* because they correspond to convolution squares of rectangular functions and hence their Fourier

[75] Because $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) \subset \boldsymbol{C}_0(\mathbb{R}^d)$ we could express this a bit more symmetric by assuming that $f \in \boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$: Note that this is a Fourier invariant Banach space, and even a Banach algebra, both with respect to pointwise multiplication and convolution!

transforms are (up to dilation and phase factors) just squared SINC-functions, i.e. of the form $\hat{\Delta}(s) = [sin(\pi s)/(\pi s)]^2$. **[Rd-specific!]**

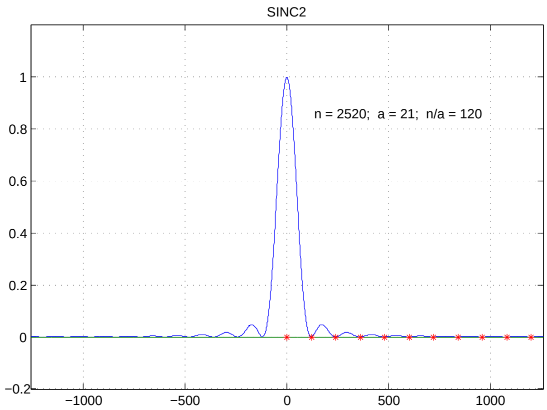


FIGURE 14. The pointwise square of the SINC function, which is the Fourier transform of the standard triangular functions. It belongs to $\boldsymbol{L}^1(\mathbb{R})$, in fact even to Wiener's algebra.

A more direct way of verifying (also relying directly on the Weierstrass approximation theorem) can be formulated as follows

FourStinj3 **Lemma 60.** *The Fourier-Stieltjes transform*[76] $\mu \to \hat{\mu}$, *given by* $\hat{\mu}(s) = \mu(\chi_s)$ *is injective.*

Proof. We start with the observation that $\hat{\mu} = 0$ is equivalent to the statement that $\mu(t) = 0$ for any so-called *trigonometric polynomial*, i.e. any function which is a finite linear combination of pure frequencies: $t(x) = \sum_{k=1}^{n} c_k \chi_{s_k}$. On the other hand, due to the density of $\boldsymbol{C}_c(\mathbb{R}^d)$ in $\boldsymbol{C}_0(\mathbb{R}^d)$

[76]The classical approach to Stieltjes integrals is described at the WIKIPEDIA page concerning the Stieltjes integral, which is intimately related to the F. Riesz representation theorem.

we can assume that a non-zero functional μ would allow some $k \in \boldsymbol{C}_c(G)$ with $|\mu(k)| \geq \gamma > 0$. By choosing a sufficiently large plateau function p with $\|p\|_\infty = 1$ (in principle a sufficiently large partial sum of any given BUPU would do) we can have the following two facts at the same time:

$$k = p \cdot k \quad \text{and} \quad \|\mu - \mu \cdot p\|_{\boldsymbol{M}} < \gamma/(3\|k\|_\infty).$$

Choosing next (based on the Weierstrass approximation theorem) some trigonometric polynomial t such that

$$\|(k - t) \cdot p\|_\infty \leq \gamma/(3\|\mu\|_{\boldsymbol{M}})$$

(meaning: good uniform approximation only over the support of p; k will be zero outside anyway, but the trigonometric polynomial will be periodic and thus repeat with some large period outside, somewhere). Putting things together we reach a contradiction, because the following estimate would have to be valid, in contradiction to our assumption $|\mu(k)| = \gamma$:

$$|\mu(k)| \leq |\mu(k) - mp(k)| + |\mu(p \cdot (k - t))| \leq \|\mu - \mu \cdot p\|_{\boldsymbol{M}} \|k\|_\infty + \|\mu\|_{\boldsymbol{M}} \|p \cdot (k - t)\|_\infty < 2/3\gamma.$$

□

As a consequence of the observation just made (Lemma 60) we can view $\mathcal{F}\boldsymbol{M}_b(\mathbb{R}^d)$ as a Banach space of its own right, with the norm

FStnorm1 (210) $$\|\widehat{\mu}\|_{\mathcal{F}\boldsymbol{M}} := \|\mu\|_{\boldsymbol{M}_b}, \quad \mu \in \boldsymbol{M}_b(\mathbb{R}^d).$$

With this definition in mind we can reformulate Lemma 60, but add some extra information to it.

MbstarFMb **Theorem 24.** *The so-called Fourier-Stieltjes algebra $(\mathcal{F}\boldsymbol{M}_b, \|\cdot\|_{\mathcal{F}\boldsymbol{M}})$ is a Banach algebra with respect to pointwise multiplication, continuously embedded into $\big(\boldsymbol{C}_{ub}(\mathbb{R}^d), \|\cdot\|_\infty\big)$. Moreover translation is isometric and strongly continuous, i.e.*

FTMisomsh1 (211) $$\|T_x\widehat{\mu}\|_{\mathcal{F}\boldsymbol{M}} = \|\widehat{\mu}\|_{\mathcal{F}\boldsymbol{M}} \quad \forall x \in \mathbb{R}^d,$$

and

FTMbcontsh (212) $$\|T_x\widehat{\mu} - \widehat{\mu}\|_{\mathcal{F}\boldsymbol{M}} \to 0 \quad for x \to 0.$$

As a consequence we have

MbstMb1 (213) $$\boldsymbol{M}_b(\mathbb{R}^d) * \mathcal{F}\boldsymbol{M}_b(\mathbb{R}^d) \subseteq \mathcal{F}\boldsymbol{M}_b(\mathbb{R}^d),$$

together with the natural norm estimate

MbstMb1 (214) $$\|\nu * \widehat{\mu}\|_{\mathcal{F}\boldsymbol{M}} \leq \|\nu\|_{\boldsymbol{M}_b} \|\widehat{\mu}\|_{\mathcal{F}\boldsymbol{M}}, \quad \mu, \nu \in \boldsymbol{M}_b.$$

Proof. The first claim is just a reformulation of the statement of Lemma 60 combined with the Convolution Theorem 12.

In order to prove isometric (and continuous) translation on the F-St-side observe that we can define $M_s\mu$ as usual via $M_s\mu(f) := \mu(M_s f)$, and $\mathcal{F}(M_y\mu) = T_y\widehat{\mu}$, which is clearly isometric. But as we know we have $T_y\widehat{\mu} = \delta_y * \widehat{\mu}$. As a consequence (simply be adding up over its discrete constituents) we have for $\nu = \sum_{i \in I} c_i \delta_{x_i} \in \boldsymbol{M}_d(\mathbb{R}^d)$:

$$\|\nu * \widehat{\mu}\|_{\mathcal{F}\boldsymbol{M}} \leq \sum_{i \in I} |c_i| \|\widehat{\mu}\|_{\mathcal{F}\boldsymbol{M}} = \|\nu\|_{\boldsymbol{M}} \|\widehat{\mu}\|_{\mathcal{F}\boldsymbol{M}}.$$

Since every measure can be well approximated by a compactly supported measure we also observe that $\|M_y\mu - \mu\|_{\boldsymbol{M}_b} \to$ for $y \to 0$, giving the continuous shift property on the Fourier transform side.

Give $\nu \in \boldsymbol{M}_b(\mathbb{R}^d)$ and $h \in \mathcal{F}\boldsymbol{M}_b(\mathbb{R}^d) \subset \boldsymbol{C}_b(\mathbb{R}^d)$ we can clearly form $\nu * h \in \boldsymbol{C}_b(\mathbb{R}^d)$. In order to verify that this function belongs to $\mathcal{F}\boldsymbol{M}_b(\mathbb{R}^d)$ (together with the corresponding norm estimate) we use the standard argument: It can be viewed as the uniform limit of the convolution products of the form $\mathrm{D}_\Psi(\nu) * h$, for $\mathrm{diam}(\Psi) \to 0$.

Hence it is enough to check that these approximating functions form a Cauchy-sequence in $(\mathcal{F}\boldsymbol{M}_b, \|\cdot\|_{\mathcal{F}\boldsymbol{M}})$. To this end one only has to compare two such approximations, say .. $\square$

algebraSecR

11. Wiener's algebra $W(R^d)$

Because it can be defined without the existence of a Haar measure the following space plays an important role within Harmonic Analysis. We define the so-called *Wiener algebra* $\boldsymbol{W}(G)$ as follows (in a way which makes sense for general LC groups!):

WCdefphi **Definition 43.** Let φ be any non-zero, non-negative function in $\boldsymbol{C}_c(\mathbb{R}^d)$.

WCdefphi1
$$\boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{L}^1) = \{\, f \in \boldsymbol{C}_0(\mathbb{R}^d) \,|\, \exists (c_k)_{k\in\mathbb{N}} \in \boldsymbol{\ell}^1, (x_k)_{k\in\mathbb{N}} \,\mathrm{in}\, \mathbb{R}^d, |f(x)| \le \sum_{k\in\mathbb{N}} c_k \varphi(x - x_k) \,\} \tag{215}$$

We define

WGnorm2
$$\|f\|_{\boldsymbol{W}(\mathbb{R}^d)} = \|f\|_{\boldsymbol{W}} := \inf \{ \|\mathbf{c}\|_{\boldsymbol{\ell}^1} = \sum_{k\in\mathbb{N}} |c_k| \,\} \tag{216}$$

where the infimum is taken over all "admissible dominations" of f as in (215).

It is obvious that $\boldsymbol{W}(\mathbb{R}^d)$ is continuously embedded into $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$ [77] since

WRintoC01
$$\|f\|_\infty \le (\sum_{k\in\mathbb{N}} |c_k|)\|\varphi\|_\infty. \tag{217}$$

By a similar argument we have a continuous embedding of
$\big(\boldsymbol{W}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{W}}\big)$ into $\big(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\big)$ or in any other $\boldsymbol{L}^p$-space, for $1 \le p \le \infty$.

A convenient characterization of Wiener's algebra over $\mathbb{R}^d$ can be given using (arbitrary, non-negative, regular) BUPUs (see Def. 16).

iener-BUPU2 **Lemma 61.** *Given a regular BUPU Ψ. Then a continuous function belongs to $\boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d)$ if and only if the decomposition $f = \sum_{\lambda\in\Lambda} \psi_\lambda \cdot f$ is absolutely convergent, i.e. if and only if $\|f\|_{\boldsymbol{W},\Psi} = \sum_{\lambda\in\Lambda} \|\psi_\lambda \cdot f\|_\infty < \infty$. Moreover, every such norm $\|f\|_{\boldsymbol{W},\Psi}$ is equivalent to the one given above (in the definition).*

The proof is left as an exercise to the reader. Of course any continuous function f with the expression $\|f\|_{\boldsymbol{W},\Psi}$ finite only countably many non-zero contributions for $\lambda_n, n \ge 1$ and thus essentially has a representation as in Def. 43, with $f_n = T_{\lambda_n} f \psi_{\lambda_n}$.

As a consequence of this observation (fix the BUPU and read the proof with two different functions φ_1 resp. φ_2 used in the definition of the space) one obtains that the space is independent from the chosen (non-zero) function $\varphi \in \boldsymbol{C}_c(\mathbb{R}^d)$. This could have been verified also directly, because for every two such functions it is always possible to dominate one by a finite sum of translates of the other (based on a simple compactness argument: any such function is obviously bounded away from zero on some open set).

[77] by the same argument $\boldsymbol{W}(\mathbb{R}^d)$ is also contained in many other Banach spaces of functions with the property that translations are isometric and that the space is solid.

There are further equivalent descriptions of the space. The first one is a kind of *atomic characterization* of $\boldsymbol{W}(\mathbb{R}^d)$:

wienatoms2 **Lemma 62.** *A function $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ belongs to $\boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d)$ if and only if it can be written as an absolutely convergent some of shifted atoms, where we consider (for now) as atoms the collection*

$$\{k \in \boldsymbol{C}_c(\mathbb{R}^d), \mathrm{supp}(k) \subset B_1(0), \|k\|_\infty \leq 1\}^{78}:$$

wienatoms2b (218)
$$f = \sum_{n=1}^{\infty} T_{x_n} k_n, \quad \textit{with} \quad \mathrm{supp}(k_n) \subseteq B_1(0), \ \textit{and} \ \sum_{n=1}^{\infty} \|k_n\|_\infty < \infty.$$

Finally we would like to mention another (also characteristic) norm, which is a kind of "continuous version" of the BUPU-norm: Given any non-zero (typically non-negative and symmetric) function $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ we define the *control function* of $f \in \boldsymbol{C}_b(\mathbb{R}^d)$ or $\boldsymbol{C}_0(\mathbb{R}^d)$ with respect to k by

elcontrfct1 (219)
$$\kappa(f,k)(x) = \|T_x k \cdot f\|_\infty.$$

With the help of this continuous functions one can show again (cf. [31]):

encontnorm3 **Lemma 63.** *For any $0 \neq k \in \boldsymbol{C}_c(\mathbb{R}^d)$ the sup-norm $\|\kappa(f,k)\|_{\boldsymbol{L}^1}$ defines an equivalent norm on $\boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d)$[79] and a continuous function belongs to $\boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d)$ if and only if this norm is finite.*

Note: We refer to [31] for details, and plan to provide the details later on. However, let us mention here that the above lemma also implies that the definition of $\boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d)$ is independent from the BUPU used. In other words, any two such BUPUs define the same space with equivalent norms. Of course one can establish this equivalence directly from the definition, using the fact that the elements of a BUPU have uniform size and that their centers are well spread (e.g. along a lattice).

testremark *Remark* 52. Of course one can define Wiener-type space, are better *Wiener amalgam spaces* of the form $\boldsymbol{W}(\boldsymbol{B}, \boldsymbol{\ell}^q)$ (or weighted versions thereof, cf. [57]) in a similar way. Again we refer to [31] for the original work, and [39] for results on duality and pointwise multipliers, although such results are not difficult to realize.

There are many good reasons to introduce the Wiener algebra, which in our case will be the Wiener amalgam space $\boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d)$. Let us first illustrate it with a picture, showing how the situation between dual pairs improves when moving from $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$ to $\boldsymbol{W}(\mathbb{R}^d)$.

The classical approach to the Wiener algebra uses the BUPU consisting of shifted indicator functions of the form $(\mathbf{1}_{[\mathbf{n},\mathbf{n+1})})_{\mathbf{n}\in\mathbb{Z}}$. The resulting sum of local maxima is then equal to the upper Riemannian sum for the integral for any $k \in \boldsymbol{C}_c(\mathbb{R})$. This can be illustrated as follows.

Wiener-Conv **Lemma 64.** $\big(\boldsymbol{W}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{W}}\big)$ *is a Banach algebra with respect to convolution. In fact it is a Banach ideal (and actually a so-called Segal algebra in the sense of H. Reiter) within* $\big(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\big)$.

Proof. Since every element in $\big(\boldsymbol{W}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{W}}\big)$ is an absolutely convergent sum of elements which are (up to translation) all bounded and have common compact support (within $\mathrm{supp}(\psi)$)

[78] Or any other fixed compact set with non-void interior in $\mathbb{R}^d$.

[79] This justifies the common use of the symbol $\boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{L}^1)(\mathbb{R}^d)$ for the Wiener algebra and similar spaces. We prefer to stay with the discrete setting because that reflects better the properties of duality and inclusions results!

BWRWRD1.jpg

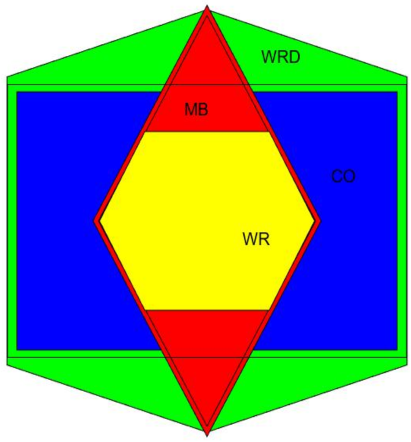


FIGURE 15. Showing that $\boldsymbol{C}_0(\mathbb{R}^d)$ and $\boldsymbol{M}_b(\mathbb{R}^d)$ are not contained in each other, while $\boldsymbol{W}(\mathbb{R}^d)$ is clearly a subspace of its dual (green symbol)

nnorm1a.pdf

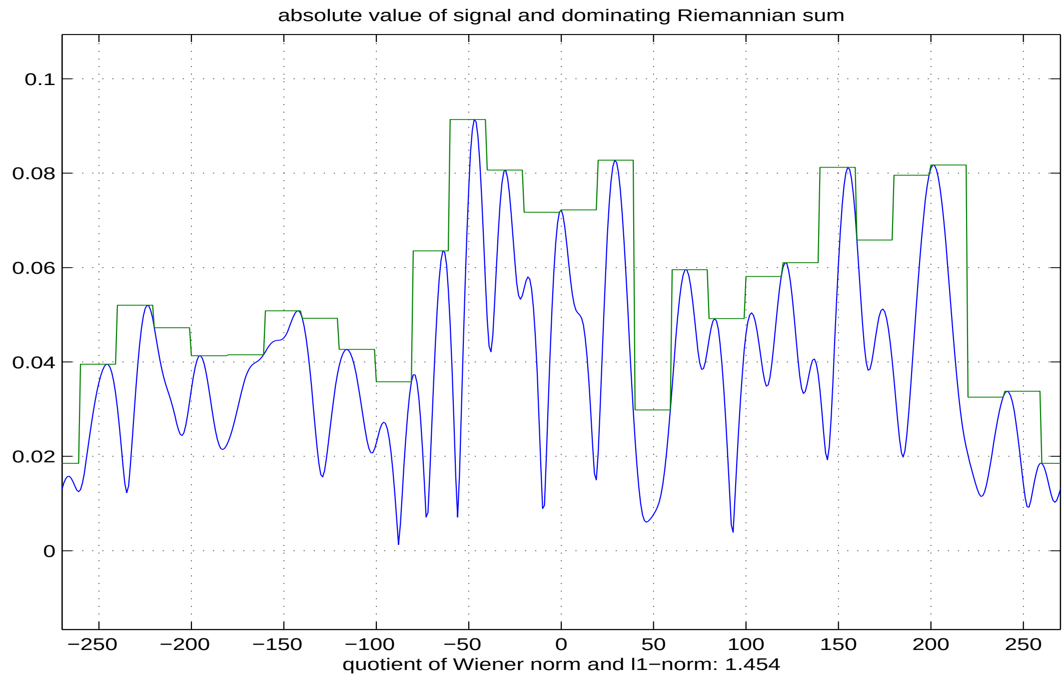


FIGURE 16. wiennorm1a.pdf: looking at the Wiener norm

it is enough split a convolution product of $g \in \boldsymbol{L}^1(\mathbb{R}^d) \subseteq \boldsymbol{M}_b(\mathbb{R}^d)$ with $f \in \boldsymbol{W}(\mathbb{R}^d)$ into little blocks and estimate

eas-WRdconv

$$\|\mu\psi_k * f\psi_n\|_\infty \le \|\mu\psi_k\|_{\boldsymbol{M}} \cdot \|f\psi_n\|_\infty . \tag{220}$$

Assuming that $\operatorname{diam}(\Psi) \leq \gamma$ for some $\gamma > 0$ we see that the support of these convolution products belongs are of size at most 2γ. Thus for any non-negative $\varphi \in \boldsymbol{C}_c(\mathbb{R}^d)$ with $\varphi(x) \geq 1$ on $B_{2\gamma}(0)$ we have

omin-atoms1 (221)
$$|\mu\psi_k * f\psi_n(x)| \leq \|\mu\psi_k\|_{\boldsymbol{M}} \|f\psi_n\|_\infty T_{x_{n,k}}\varphi(x).$$

and hence for some constant $C > 0$ (by Theorem 4 and Lemma 61) we have

$$\|\mu * f\|_{\boldsymbol{W}} \leq \sum_k \sum_n \|\mu\psi_k\|_{\boldsymbol{M}} \|f\psi_n\|_\infty \leq C\|\mu\|_{\boldsymbol{M}} \|f\|_{\boldsymbol{W}}.$$

□

Remark 53. An alternative method to proof the $\big(\boldsymbol{W}(\mathbb{R}^d),\, \|\cdot\|_{\boldsymbol{W}}\big)$ is a *Banach convolution module*, i.e. a Banach module over $\big(\boldsymbol{L}^1(\mathbb{R}^d),\, \|\cdot\|_1\big)$ with respect to convolution as abstract multiplication is to verify that it is a *homogeneous Banach space*. Any such homogeneous Banach space is a Banach convolution module over the Banach algebra $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ (with convolution), as well as an *essential Banach convolution module* over $\big(\boldsymbol{L}^1(\mathbb{R}^d),\, \|\cdot\|_1\big)$.

Remark 54. Yet another method to control convolution products of Wiener amalgam spaces is provided by the paper [15]. Essentially, instead of using the domination property using a sufficiently wide plateau-function in the last step of the proof of Lemma 64 above they use the fact that one can split the convolution product into a controlled number of building blocks and thus only has to rearrange the sums. This method would be easy in the current situation but requires and extra argument for more general Wiener amalgam spaces.

Similar statements can be made for other function spaces, in particular for the pointwise algebra $(\boldsymbol{A},\, \|\cdot\|_{\boldsymbol{A}}) = \big(\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d),\, \|\cdot\|_{\mathcal{F}\boldsymbol{L}^1}\big)$. This is a special case of a Segal algebra (see next section). It has been introduced around 1979 as $\big(\boldsymbol{S}_0(G), \|\cdot\|_{\boldsymbol{S}_0}\big)$ in [30] (the zero in the subscript stands for the *minimality* of this space). Meanwhile (see [81]) this space is called *Feichtinger's algebra*.

Within the family of Wiener amalgam spaces (defined via BUPUs, which are bounded in the Fourier Algebra $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$) we can give the following definition:

SO-WFLili **Definition 44.** Let $\Phi = (\varphi_\lambda)_{\lambda\in\Lambda}$ be any bounded BUPU in $\big(\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d),\, \|\cdot\|_{\mathcal{F}\boldsymbol{L}^1}\big)$.

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

Just as a preview about inclusions between the various spaces let us mention right away that we will obtain the following scenario in terms of inclusions with established spaces (here `SO` stands for $\boldsymbol{S}_0(\mathbb{R}^d)$ and `SOP` for $\boldsymbol{S}_0'(\mathbb{R}^d)$).

One of the first and most important properties is the invariance of the space under the Fourier transform! We derive it from the convolution properties of Wiener amalgam spaces.

CHECK CONSISTENCY OF ARGUMENT!

SO-FOURinv **Theorem 25.**

SOFTSO (222)
$$\mathcal{F}[\boldsymbol{S}_0(\mathbb{R}^d)] = \boldsymbol{S}_0(\mathbb{R}^d),$$

with equivalence of norms[80]. *Moreover, both the time-shift operators and the modulation operators act uniformly bounded on* $\big(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\big)$, *and furthermore* $\|T_x f - f\|_{\boldsymbol{S}_0(\mathbb{R}^d)} \to 0$ *for* $x \to 0$ *as well as* $\|M_s f - f\|_{\boldsymbol{S}_0(\mathbb{R}^d)} \to 0$ *for* $s \to 0$, *for each* $f \in \boldsymbol{S}_0(\mathbb{R}^d)$.

As a consequence one has $\boldsymbol{L}^1(\mathbb{R}^d) * \boldsymbol{S}_0(\mathbb{R}^d) \subset \boldsymbol{S}_0(\mathbb{R}^d)$ *and* $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) \cdot \boldsymbol{S}_0(\mathbb{R}^d) \subset \boldsymbol{S}_0(\mathbb{R}^d)$

[80] Later on we will see that with a specific norm one can make the Fourier transform an isometric automorphism on $\big(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\big)$.

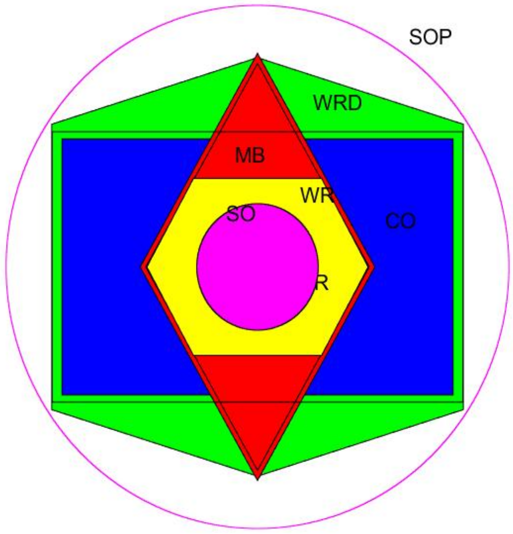


FIGURE 17. The Wiener algebra $\left(\boldsymbol{W}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{W}}\right)$ (yellow) inside of $\boldsymbol{L}^1(\mathbb{R}^d)$, and of course also within $\boldsymbol{L}^2(\mathbb{R}^d)$ and $\boldsymbol{C}_0(\mathbb{R}^d) \subset \boldsymbol{L}^\infty(\mathbb{R}^d)$ (here blue), but containing (as a proper subspace [magenta]) the Segal algebra $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$. The dual space of the Wiener algebra is marked green in this plot. FILE: COMBWR-WRD1SOSOP.jpg

Proof. First the uniform boundedness of both time and frequency shift operators has to be considered (should be considered an exercise). In fact, it is enough to verify it on atoms (i.e. on function of the form $f\varphi_n$), and there the properties of a BUPU in combination with the algebra properties come into play.

Since obviously the compactly supported (partial sums) elements are dense in $\boldsymbol{S}_0(\mathbb{R}^d)$. But over compact sets the $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$-norm and the $\boldsymbol{S}_0(\mathbb{R}^d)$-norms are equivalent, and hence to continuous dependencies claimed above are valid.

As a further consequence by (vector-valued) integration we get $\boldsymbol{M}_b(\mathbb{R}^d) * \boldsymbol{S}_0(\mathbb{R}^d) \subset \boldsymbol{S}_0(\mathbb{R}^d)$, and hence in particular $\boldsymbol{L}^1(\mathbb{R}^d) * \boldsymbol{S}_0(\mathbb{R}^d)$, with $\|g * f\|_{\boldsymbol{S}_0(\mathbb{R}^d)} \leq \|g\|_1 \|f\|_{\boldsymbol{S}_0}$, for all $g \in \boldsymbol{L}^1$, $f \in \boldsymbol{S}_0$.

Since by the definition the decomposition $f = \sum_{\lambda\in\Lambda} \psi_\lambda \cdot f$ is an absolutely convergent series in $\left(\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_{\mathcal{F}\boldsymbol{L}^1}\right)$ it follows that $\hat{f} = \sum_{\lambda\in\Lambda} \widehat{\psi_\lambda \cdot f}$ is absolutely convergent in $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$. But we want more. The series should also be convergent in the (much smaller) Banach space $\boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1, \boldsymbol{\ell}^1)(\mathbb{R}^d)$. In order to do this we take *any* function in $\boldsymbol{S}_0(\mathbb{R}^d)$ (e.g. $SINC.^2$) such that $\hat{g}$ belongs to $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ and has compact support (e.g. the Vallee DePoussin kernel), and such that $\psi \cdot \hat{g} = \psi$, hence

$$T_\lambda(\psi \cdot \hat{g}) = T_\lambda\psi \cdot T_\lambda\hat{g} = T_\lambda\psi \cdot \widehat{M_\lambda g}$$

and consequently we can derive

$$\|\widehat{\psi_\lambda \cdot f}\|_{\boldsymbol{S}_0(\mathbb{R}^d)} = \|T_\lambda\hat{g}\cdot\widehat{\psi_\lambda \cdot f}\|_{\boldsymbol{S}_0(\mathbb{R}^d)} \leq \|M_\lambda g\|_{\boldsymbol{S}_0(\mathbb{R}^d)} \|\widehat{\psi_\lambda \cdot f}\|_{\boldsymbol{L}^1(\mathbb{R}^d)} = \|g\|_{\boldsymbol{S}_0(\mathbb{R}^d)} \|\widehat{\psi_\lambda \cdot f}\|_{\boldsymbol{L}^1(\mathbb{R}^d)}.$$

and consequently

$$\|\hat{f}\|_{\boldsymbol{S}_0(\mathbb{R}^d)} \leq \sum_{\lambda\in\Lambda} \|\widehat{\psi_\lambda \cdot f}\|_{\boldsymbol{S}_0(\mathbb{R}^d)} \leq \|g\|_{\boldsymbol{S}_0(\mathbb{R}^d)} \sum_{\lambda\in\Lambda} \|\widehat{\psi_\lambda \cdot f}\|_{\boldsymbol{L}^1(\mathbb{R}^d)} \leq \|g\|_{\boldsymbol{S}_0(\mathbb{R}^d)} \cdot \|f\|_{\boldsymbol{S}_0(\mathbb{R}^d)}.$$

□

12. The Segal algebra $S_O(Rd)$ and Banach Gelfand triples

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.

We have now a Banach space $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ which is Fourier invariant, contained (continuously embedded, and densely) in $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$. Together with the dual space $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$ we will have a convenient setting for a theory of generalized Fourier transforms. The space $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ is also isometrically invariant under TF-shifts, as well as the various involutions (such as $f \mapsto \overline{f}$ or $f \mapsto f^*$ or $f \mapsto f^{\checkmark}$), but also (not-isometrically[81]) invariant under the *dilation operators* (of course both D_ρ and St_ρ), but also more general automorphism of $\mathbb{R}^d$.

What we are going for is a simplified situation with just one Banach space of test-function, embedded into the Hilbert space $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$ and this in turn embedded into the dual space, the space of "distributions" (as we call them by only a slight "abuse de language").

13. The inverse STFT

Part of this material is from Severin Bannert's master thesis (ba10-2 5). An upcoming article in the Notices of the AMS ([45]) will provide a good summary of this part of Harmonic Analysis. Already now [19] can serve as a good introduction to Banach Gelfand triples.

[81]Asymptotic behavior of the dilation operators on $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ has been studied by Cordero et. al., see [20, 94].

soplot.pdf

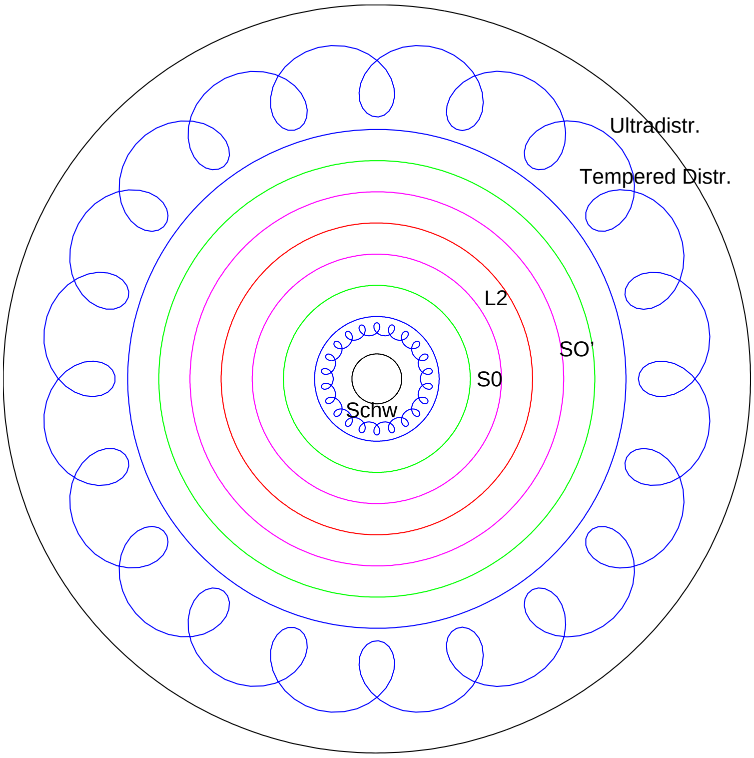


FIGURE 18. Spaces of *ultra-distributions*: just for illustration: Inside of the Schwartz space there are various Banach spaces of test functions, we just plot a possible one, e.g. a Shubin-type space, with a radial symmetric weight of sub-exponential type on the TF-plane. One can find quite a number of such spaces which are algebras and also Fourier invariant. Further more the theory of Gabor representations based on *coorbit theory* (see [40]) allows to derive atomic decomposition for such spaces. The corresponding dual space is then a Banach space of *ultra-distributions*.

Definition 45. The *Short Time Fourier Transform (STFT)* of a function $f \in \boldsymbol{L}^2(\mathbb{R}^d)$ with respect to a window[82] $g \in \boldsymbol{L}^2(\mathbb{R}^d)$ is defined for $(x, \omega) \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d$ via

$$\begin{aligned} V_g f(x,\omega) &= \int_{\mathbb{R}^d} f(t)\overline{g(t-x)}e^{-2\pi i t\cdot\omega}dt \\ &= \int_{\mathbb{R}^d} f(t)\overline{M_\omega T_x g(t)}dt \,=\, \langle f, M_\omega T_x g\rangle. \end{aligned}$$

In other words: the STFT describes the similarity (correlation) between the given signal (which is to be analyzed) with all the possible time-frequency shifted copies of the template signal g. There are many good reasons (mostly practical ones) to think of g as a symmetric and smooth, real valued function which is well concentrated around the origin.

[82]The word *"window"* is used with the understanding that the multiplication of the signal f to be analyzed by the typically compactly supported function g localizes the function. Often g is assumed to be symmetric and real-valued, centered at zero, maybe a plateau-function, or a B-spline. For this reason the STFT is also often called the *sliding window Fourier transform*.

soplot2.pdf

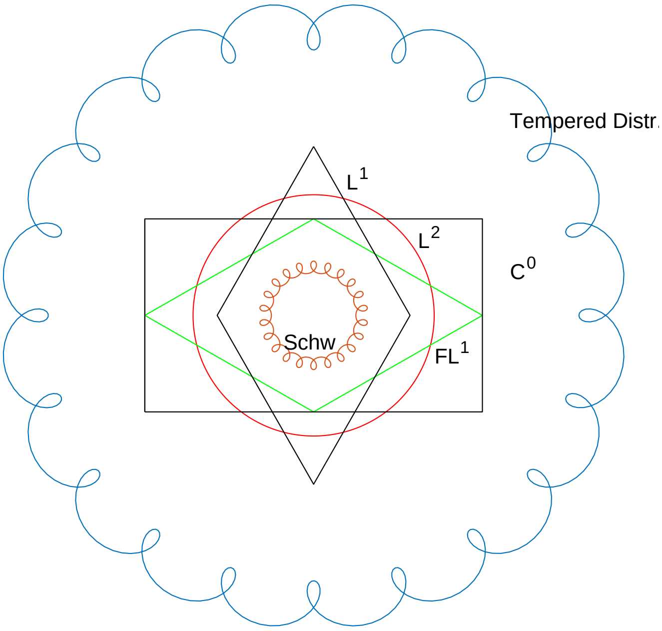


FIGURE 19. The usual space for the Fourier transform, showing innermost the Schwartz space $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$, contained in the space $\boldsymbol{L}^1(\mathbb{R}^d)$ (black), $\boldsymbol{L}^2(\mathbb{R}^d)$ (red circle) and $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ (green). All the space are in turn contained within the space $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ of *tempered distributions*. The curly structure of these plots - in fact plots of trigonometric polynomials in the complex plane! - indicate that $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ has many seminorms, and the dual space arises essentially by reflection of this structure across the circle representing the Hilbert space $\boldsymbol{L}^2(\mathbb{R}^d)$: PDF: SO-PLOT2

OrthSTFT **Theorem 26.** *(Orthogonality relations of the STFT)*
Let $f_1, f_2, g_1, g_2 \in \boldsymbol{L}^2(\mathbb{R}^d)$, then

th-rel-stft
$$\langle V_{g_1} f_1, V_{g_2} f_2 \rangle = \langle f_1, f_2 \rangle \overline{\langle g_1, g_2 \rangle}. \tag{223}$$

In particular $V_{g_k} f_k \in \boldsymbol{L}^2(\mathbb{R}^{2d})$ for $k = 1, 2$.

Proof. See [53], p. 42. □

The orthogonality relations immediately lead to the following corollary.

Corollary 6. *(Moyals Formula) Let $f, g \in \boldsymbol{L}^2(\mathbb{R}^d)$, then*

eq:moyals
$$\|V_g f\|_{\boldsymbol{L}^2(\mathbb{R}^d \times \widehat{\mathbb{R}}^d)} = \|g\|_{\boldsymbol{L}^2(\mathbb{R}^d)} \|f\|_{\boldsymbol{L}^2(\mathbb{R}^d)} \tag{224}$$

Remark 55. If in particular $\|g\|_2 = 1$ (for example if g is the normed Gaussian), then (224) says that the STFT is an isometry, $V_g : \boldsymbol{L}^2(\mathbb{R}^d) \to \boldsymbol{L}^2(\mathbb{R}^{2d})$, and thus injective, i.e each $f \in \boldsymbol{L}^2(\mathbb{R}^d)$ is uniquely determined by its STFT. Next we will find a way to reconstruct f from its STFT.

With the previous results we have the right tools at hand to formulate and proof the inversion formula of the STFT.

Theorem 27. *(The Inversion formula of the STFT)*
Let $g \in \boldsymbol{L}^2(\mathbb{R}^d)$ with $\|g\|_{\boldsymbol{L}^2} = 1$, then for $f \in \boldsymbol{L}^2(\mathbb{R}^d)$ we have the following identity (to be

elfill1.jpg

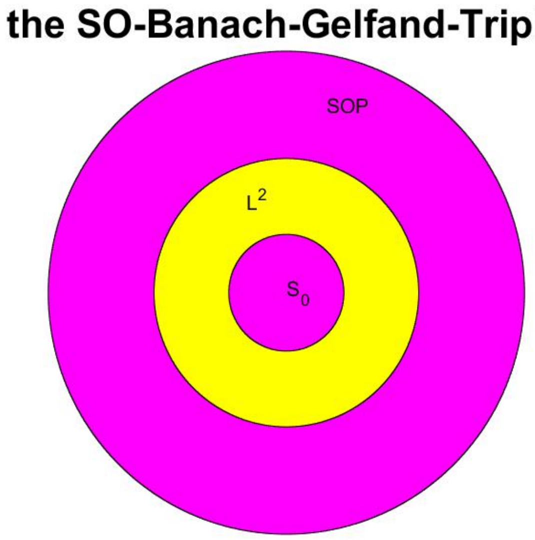


FIGURE 20. The simple situation representing the Banach Gelfand Triple $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$. Innermost the Segal algebra $\big(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\big)$, serving as algebra of test functions, which is Fourier invariant, in the middle the Hilbert space $\big(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\big)$ and everything surrounded by the space $\big(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\big)$ (also endowed with the w^*-convergence), containing Dirac measures, pure frequencies or Dirac combs. PLOT: sogelfill1.jpg

understood in the weak sense):

eq:inv-stft

$$f = \int_{\mathbb{R}^d \times \widehat{\mathbb{R}}^d} V_g f(x,\omega) M_\omega T_x g \, d\omega dx \tag{225}$$

Proof. Since $\|g\|_2 = 1$, (223) implies that

$$\langle V_g f_1, V_g f_2 \rangle = \langle f_1, f_2 \rangle \qquad \forall f_1, f_2 \in \boldsymbol{L}^2(\mathbb{R}^d),$$

which leads to

:Vg*Vgf1-f2

$$\langle V_g^* V_g f_1, f_2 \rangle = \langle f_1, f_2 \rangle, \tag{226}$$

which in turn implies $V_g^* V_g = \mathrm{id}$, where V_g^* denotes the adjoint operator of V_g. What we have to show now is that[83]

*F-inv-stft

$$V_g^* F = \int_{\mathbb{R}^d \times \widehat{\mathbb{R}}^d} F(x,\omega) M_\omega T_x g \, dx d\omega, \quad \text{for} \quad F \in \boldsymbol{L}^2(\mathbb{R}^{2d}). \tag{227}$$

[83]Recall that V_g^* is the adjoint of the operator $V_g : \boldsymbol{L}^2(\mathbb{R}^d) \to \boldsymbol{L}^2(\mathbb{R}^{2d})$.

This is done by the following computations:

$$\begin{aligned}\langle f, V_g^* F\rangle = \langle V_g f, F\rangle &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} V_g f(x,\omega)\overline{F(x,\omega)}\,dxd\omega\\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} f(t)\overline{T_x g(t)}e^{-2\pi i t\omega}dt\right)\overline{F(x,\omega)}\,dxd\omega\\ &= \int_{\mathbb{R}^d} f(t)\overline{\left(\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} F(x,\omega)M_\omega T_x g(t)\,dxd\omega\right)}dt.\end{aligned}$$

Hence, by setting $F = V_g f$ and with the help of (226) the proof is complete. □

Remark 56. We can omit the assumption that $\|g\|_{\boldsymbol{L}^2} = 1$ in the previous theorem. The inversion formula then reads

$$f = \frac{1}{\|g\|_2^2}\int_{\mathbb{R}^d\times\widehat{\mathbb{R}}^d} V_g f(x,\omega)M_\omega T_x g\,dxd\omega.$$

About the actual convergence of the integral (describing the adjoint linear mapping, applied to F) one has to say: it is only defined in the "weak sense" (i.e. it is understood in a kind of symbolic sense in general). For $F\in\boldsymbol{L}^1(\mathbb{R}^{2d})$, specifically for $F = V_g(f)$ for some $f\in\boldsymbol{S}_0(\mathbb{R}^d)$ the convergence in the spirit of a Riemannian sum is taking place in the $\boldsymbol{S}_0$-sense, hence also uniformly, in $\big(\boldsymbol{L}^1(\mathbb{R}^d),\, \|\cdot\|_1\big)$ as well as pointwise (cf. the work of F. Weisz: [102]).

$$f(t) = \frac{1}{\|g\|_2^2}\int_{\mathbb{R}^d\times\widehat{\mathbb{R}}^d} V_g f(x,\omega)M_\omega T_x g(t)\,dxd\omega,$$

where in fact it would be enough to use an absolutely convergent (improper) Riemannian integral with respect to $\mathbb{R}^d\times\widehat{\mathbb{R}}^d$. For *good windows* $g\in\boldsymbol{S}_0(\mathbb{R}^d)$ one has norm convergence (but to my knowledge not necessarily pointwise convergence) for $f\in\boldsymbol{L}^2(\mathbb{R}^d)$.

The next object to be introduced here is the dual space $(\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$, consisting of all bounded linear functionals on the Banach space $\big(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\big)$.

14. The distribution space $S_0'(R^d)$

In the same way as we extend, as needed, the field $\mathbb{Q}$ of rational numbers to the field of real numbers (in order to gain *completeness* in the metric sense, because within $\mathbb{R}$ every *Cauchy sequence* is *convergent* in $\mathbb{R}$) and afterwards embed $\mathbb{R}$ into the even larger field of complex numbers $\mathbb{C}$, in order to gain more freedom in computing object (e.g. write every polynomial in a unique way as a product of linear factors!) we are going to *enlarge the vector space objects* which we are able to handle from the collection of ordinary functions (such as $\boldsymbol{C}_c(G), \boldsymbol{L}^2(G), \boldsymbol{L}^1(G)$, in fact up to equivalence, identifying two measurable functions if they are equal almost everywhere) to a space of "generalized functions", also called distributions[84].

There are certainly different choices (how large the space of generalized functions should be) and how general, perhaps complicated the corresponding theory of generalized functions should be chosen. From a didactical point of view one will certainly choose the setting in accordance with the background of the readers or students attending such a course, but also what the purpose is.

For example, it is save to say that in the context of PDE (partial differential equations, where one needs the concept of a "generalized or weak solution") the space of tempered distributions $\boldsymbol{\mathcal{S}}'(\mathbb{R}^d)$ is the right setting, which is the dual of L. Schwartz's space $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ of rapidly decreasing functions. As we will see, the dual space of $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$ is well suited for practically all the questions arising in standard engineering applications as well as for the treatment of questions of *abstract* (and as the author tries to promote) *conceptual harmonic analysis* (see [37]).

Of course one could go on and use even larger spaces, such as the space of *ultra-distributions* introduced by Björck ([9]).

There are different ways to generate such larger objects. The more intuitive one (comparable with the idea of *completion of metric spaces* which can be used to obtain, i.e. somehow *construct* $\mathbb{R}$ from $\mathbb{Q}$ by adding *limits of non-convergent Cauchy-sequences* is unfortunately the more cumbersome, and despite its simple "start" requiring rather elaborate and lengthly (see the books of Lighthill ([68]) or Jones ([61] for such an approach). Obviously this approach has not followed many followers, because it brings some restrictions and even the question whether the objects obtained in this way are the same as the ones described in other books and papers concerning distributions has to be addressed if one tries to make use of this theory.

So in effect the established way to create various spaces of *generalized functions* is to define them as dual spaces, i.e. as a space of linear functionals on some topological vector space of test functions. In other words, one starts by describing a vector space of "nice" functions (typically well concentrated, maybe with some smoothness), with ordinary pointwise addition and scalar multiplication, but also a natural notion of convergence (i.e. a topology, compatible with the vector space structure, maybe even metrizable, turning the vector space into a complete metric vector spaces, as in the case of the Schwartz space $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$). At the beginning of these notes we have chosen to work with $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$ (which is a Banach algebra), and in some sense

[84]This wording should not be confused with the same vocabulary arising in probability theory. A probability measure is certainly describing the distribution of values of a random variable, and is at the same time a bounded measure, hence a distribution (or order zero), but this duplication of terminology is more or less pure coincidence and should not confuse the reader! There are also authors restricting the use of the word "distribution" to dual spaces of spaces of test functions which are infinitely differentiably, while the more general situation of spaces which are dual to more simple spaces of test functions are just called *generalized functions*. Again, we will not narrow down our terminology in this respect. For us the words "distributions" and "generalized" functions will have the same meaning. The only crucial aspect is that a space of generalized functions has to contain (certain!) ordinary functions and that operations easy to define on ordinary functions, such as translation, convolution, etc. , extend in a natural way to those generalized functions.

$\boldsymbol{M}_b(\mathbb{R}^d)$ is the first space of "generalized functions" that we have met so far. The disadvantage of this situation is the fact that $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$ cannot embedded into $\boldsymbol{M}_b(\mathbb{R}^d)$ in a natural way.

Another important point in this context it the *ability to consider ordinary functions as generalized functions*, i.e. the possibility of creating (a natural) identification of "ordinary" functions with generalized functions. Such a situation is well known from probability: Whenever we have a continuous real-valued random variable (e.g. the normal distribution) with a density $g(t)$ we identify this function with the corresponding *probability distribution* or probability measure (as we prefer to call it here), with

$$\mu_g(A) = \int_A g(t)dt = \int_{\mathbb{R}} \mathbf{1}_A(t) \cdot g(t)dt,$$

or viewed as a functional on $\boldsymbol{C}_0(\mathbb{R})$:

natemb03 (228)
$$\mu_g(f) = \int_{\mathbb{R}} f(t) \cdot g(t)dt,$$

which is the same as **integration against** $g(t)$ (or take g as a density).

So in the same way as $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ is *embedded* (in fact even isometrically) into $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ in this way, we want to think of a natural embedding (in fact the same) as a space of test functions always being embedded into its dual space.

There are other considerations which may influence the choice of spaces: Whenever the space of test functions is sufficiently rich (one has enough compactly supported functions) one can define the support of a distribution (supp(σ)), which is of course compatible with the usual operations (e.g. shift or dilation). So, even if it is *typically NOT possible* to talk about the function values of a distribution, there it is still possible to talk about the subset of the domain on which it is defined where its "action takes place" (namely the complement of the open domain on which the linear functional acts in a *trivial way*, i.e. just like the zero functional).

Thus the question is: How large should the space of test functions be chosen and what kind of topology should one take. It is plausible that a small space will go along with a fine topology, hence the collection of convergent sequences will be reduced more and more, the smaller the space of test functions is and the finer the topology is. This goes hand in hand, especially for Banach spaces of test functions one can show that the natural embedding of a small space of test functions is always a bounded linear mapping (at least of convergence implies local $\boldsymbol{L}^1$-convergence, by the closed graph theorem). We also find it preferable to have a simple (single) norm and thus a Banach space of test functions (namely $\left(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\right)$) and not just a topology arising from a countable collection of seminorms, simply because it makes it more difficult and cumbersome to describe convergence.

Thus it seems that one should take the space of test functions rather small, but there are limitations to this. For example, it is a bad idea to consider spaces of test functions consisting of (restrictions from $\mathbb{C}$ to $\mathbb{R}$) of analytic functions, because none of them will have compact support.

Also one of the evident advantages of $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ (but also of $\boldsymbol{S}_0(\mathbb{R}^d)$ or $\boldsymbol{S}_0(G)$) is its natural behaviour under the Fourier transform. With the definition of $\boldsymbol{S}_0(G)$ over LCA groups one can show that $\mathcal{F}(\boldsymbol{S}_0(G) = \boldsymbol{S}_0(\widehat{G})$.

Depending on the approach taken (either via a kind of completion process or by duality methods) the definitions of extended operations look a bit different, but one can certainly argue, that a careful realization of any such process leads to the same concept (e.g. the generalization of translation, dilation, convolution, the Fourier transform, etc.), i.e. the *only natural* extension one should consider. In fact, it will be always a kind of w^*-continuous extension of procedures known to work well on the space of test functions.

There are different ways of doing this. If one thinks of the "large" space of "generalized objects" as the set of abstract elements giving limits to (kind of) Cauchy-sequences then one has to show that the linear operators in mind are compatible with this procedure of completion. Given such a sequence (or in fact a net) representing such an abstract element one has to show that after applying the operator (defined on the space of test functions) one has again a well defined sequence of test function, which is showing the same Cauchy-type behaviour, thus defining another element (the target for the extended operator). Of course one has to verify that different representative sequences (or nets) in the domain of the operator result in equivalent sequences in the target domain, i.e. that the mapping is well defined (and preserves equivalence classes, making up the elements of the abstract completion). While such an approach appears to be very natural and elementary turns out to be rather cumbersome in practice and overall the burden of book-keeping outweighes the perceived simplicity of the approach at first sight.

For this reason the much more elegant (but at the beginning more abstract looking) approach *using dual pairs of Banach spaces* together with the w^*-convergence of sequences (and/or nets) is preferred in most cases. We also will follow this method (but will try a bit later to indicate that the two approaches are in fact equivalent and lead to the same extensions).

In fact, a combination of such approaches is perhaps most intuitive: Given the bounded linear operator T (say the Fourier transform) on $\boldsymbol{S}_0(\mathbb{R}^d)$ we can extend it by duality to all of $\boldsymbol{S}_0'(\mathbb{R}^d)$ (with the usual formula XXX), and then check easily that this extended mapping is w^*-w^*-- continuous . Hence the action on a given distribution $\sigma \in \boldsymbol{S}_0'(\mathbb{R}^d)$ can be realized by saying: Whatever regularizing net/sequence of operators is applied to σ (with $\sigma = w^*\text{-}\lim_\alpha A_\alpha(\sigma)$), then we will have that

opbyreg00
$$T(\sigma) = w^*\text{-}\lim_\alpha T(A_\alpha(\sigma)). \tag{229}$$

Of course the choice of this regularizing net of operators may depend on the given σ, i.e. the observed deficiencies of σ (with respect to the property of belong to the space of test function). It may be sufficient to add decay, if it is a smooth function (the typical situation when summability methods are applied) or smoothness, if it is a rough object (not continuous, or maybe even a discrete measure), or both deficiencies, usually with the help of product-convolution operators.

In our setting one can say, whenever it comes to the inversion of the Fourier transform, defined on $\boldsymbol{L}^1(\mathbb{R}^d)$: The range contains only functions within $\boldsymbol{C}_0(\mathbb{R}^d)$ (according to the Riemann Lebesgue Lemma), and in fact the local structure of $\mathcal{F}\boldsymbol{L}^1$ and $\boldsymbol{S}_0$ coincide, and $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ is in the pointwise multiplier space of $\boldsymbol{S}_0(\mathbb{R}^d)$, *but not every function* from $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ is also in $\boldsymbol{L}^1(\mathbb{R}^d)$, leave alone within $\boldsymbol{S}_0(\mathbb{R}^d)$. Hence one has to just pull functions from $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ into $\boldsymbol{S}_0(\mathbb{R}^d)$ which is easily done by means of summability kernels, all of which belong to $\boldsymbol{S}_0(\mathbb{R}^d)$ (cf. papers by F. Weisz) and which converge to the constant 1, uniformly over compact sets, which is enough to guarantee w^*-convergence (within the dual pairing $(\boldsymbol{S}_0'(\mathbb{R}^d), \boldsymbol{S}_0(\mathbb{R}^d))$)).

Exercise: Given a bounded sequence (h_n) in $\big(\boldsymbol{C}_b(G), \|\cdot\|_\infty\big)$. If $h_n(x) \to h_0(x)$ uniformly over compact subsets then $h_0 = w^*\text{-}\lim h_n$ in $\boldsymbol{S}_0'(\mathbb{R}^d)$ (viewed as distributions). In fact a partial converse is true. A bounded sequence $(\sigma_n)_{n=1}^\infty$ is w^*-convergent in $\boldsymbol{S}_0'(\mathbb{R}^d)$ if and only if every convolution with some fixed $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ results in a net $(\sigma_n * g)_{n=1}^\infty$ which has this property: It converges uniformly over compact sets to the limit $\sigma_0 * g \in \boldsymbol{C}_b(\mathbb{R}^d)$.

For functions $f \in \boldsymbol{L}^1(\mathbb{R}^d)$, in fact for bounded measures $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ one has already good enough decay, so that only smoothing is required. Given any approximate unit (bounded with respect to the $\boldsymbol{L}^1(\mathbb{R}^d)$-norm) inside of $\boldsymbol{S}_0(\mathbb{R}^d)$ (e.g. the net $(\mathrm{St}_\rho g_0)_{\rho\to 0}$, with $\int_{\mathbb{R}^d} g_0(t)dt = 1$) one has a w^*-approximation of μ (or in fact an $\boldsymbol{L}^1$-approximation of f in case $\mu = \mu_f$, i.e. for absolutely continuous measures on the group). For $h \in \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ or even $h = \widehat{\mu}$ for some $\mu \in \boldsymbol{M}(\mathbb{R}^d)$ one has good enough local properties (in fact, they are within the pointwise multiplier algebra of $\boldsymbol{S}_0(\mathbb{R}^d)$) so that any *pointwise approximate unit* in $\boldsymbol{S}_0(\mathbb{R}^d)$ will serve as *regularizing operator*. The classical cases where such a situation arises is the Fourier inversion problem. Due to the problem that in many cases (and certainly in the case of Fourier Stieltjes transforms).

Note that one has - for almost trivial reasons, bases on the symmetry of the concept of the Banach Gelfand triple - with respect to the Fourier transform: A family is a pointwise regularizer (in the sense of summability kernels) if and only if its Fourier transformed version is a Dirac-like smoothing operator. One just has to recall that in the current context the approximate units act via pointwise convolution or via convolution respectively, and that the Fourier transforms intertwines these two operators.

In fact, one has a similar property with respect to the Shah-distribution (or Dirac comb). Pointwise multiplication of a bounded function h gives a translation bounded measure $\sum_{\lambda \in \Lambda} h(\lambda)\delta_\lambda$, which obviously belongs to $\boldsymbol{S}_0'(\mathbb{R}^d)$, which has a Fourier transform. Since this is a discrete measure supported by the lattice Λ it is clear that its Fourier transform is a distribution in $\boldsymbol{S}_0'(\mathbb{R}^d)$ which is periodic with respect to the dual (or orthogonal) lattice $\Lambda^\perp$.

On the other hand, viewed as a distribution $h \in \boldsymbol{C}_b(\mathbb{R}^d) \subset \boldsymbol{S}_0'(\mathbb{R}^d)$ has a Fourier transform $\hat{h}$, and one may expect that this Fourier transform just has to be periodized with respect to $\Lambda^\perp$, i.e. it is supposed to be of the form $\sum_{\Lambda^\perp} T_{\lambda^\perp}\hat{h}$. But in which sense is this periodization possible. Since both factors in this in this convolution product, namely $⧢_{\Lambda^\perp}$ and $\hat{h}$ may be "real distributions" their convolution product may not be defined in the usual way, and one has to verify in which sense the convolution product can be taken. And again: If $\hat{h}$ was in $\boldsymbol{L}^1(\mathbb{R}^d)$ (or more general a bounded measure), then one could view itself as an absolutely convergent sum of its pieces $\mu_l amp$, by just using a regular BUPU over $\Lambda^\perp$ in order to cut the measure into pieces, and use the fact that $\sum_{\lambda^\perp \in \Lambda^\perp} \psi_{\lambda^\perp}\mu$ is absolutely convergent with

$$\|\mu\| = \sum_{\lambda^\perp \in \Lambda^\perp} \|\mu \cdot \psi_{\lambda^\perp}\|_{\boldsymbol{M}}.$$

(cf. Thm.4).

But in fact, $\hat{h}$ might be a measurable function which is compactly supported but *not locally integrable* (at one single point) and in this case the periodization would make perfect sense, using the pointwise addition.

On the other hand it may be quite possible to find an example where the function h is made in such a way that despite the fact that the pseudo-measure is defined by a pointwise function (in this case with large support, so that pointwise infinite sums have to be considered) one has a problem in defining the periodization in the usual pointwise sense. The chirp signals $ch(t) = exp(i\alpha t^2)$ or discrete versions (to be discussed later on) are examples of this type.

NOTE: Finally these remarks have become far to long and to "exhaustive" (for the poor readers).

Material for course October 2006, Double modules

Given a Banach module $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ over a Banach algebra $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ with bounded approximate units we define the *essential part* $\boldsymbol{B_A}$ and the *relative completion* of $\boldsymbol{B}$ with respect to $\boldsymbol{A}$, which is given as $\boldsymbol{B^A} := \mathcal{H}_{\boldsymbol{A}}(\boldsymbol{A}, \boldsymbol{B})$. Note that this is again in a natural way a Banach module (with respect to the operator norm) over the original Banach algebra $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$, and that $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ can be mapped into this Banach module in a natural way as a closed subspace (at least of $\boldsymbol{B} = \boldsymbol{B_A}$ and if $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ has bd. approx. units).

Thus we have altogether the following chain of embeddings of closed subspaces:

essrelemb1

$$\boldsymbol{B_A} \subseteq \boldsymbol{B} \subseteq \boldsymbol{B^A}. \tag{230}$$

Just as one non-trivial example we have for $\boldsymbol{B} = \left(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\right)$ with the Banach algebra $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ acting via convolution: The essential part of $\boldsymbol{L}^\infty(\mathbb{R}^d)$ is $\boldsymbol{C}_{ub}(\mathbb{R}^d)$, the closed subalgebra of uniformly continuous, bounded functions:

essrelemb2

$$\boldsymbol{C}_{ub}(\mathbb{R}^d) \subsetneq \boldsymbol{C}_b \subsetneq \boldsymbol{L}^\infty. \tag{231}$$

Proof. Since $\left(\boldsymbol{L}^\infty(\mathbb{R}^d), \|\cdot\|_\infty\right)$ is a dual space (namely the dual space to $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$) the relative completion of $\left(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\right)$ cannot be larger than $\boldsymbol{L}^\infty(\mathbb{R}^d)$. On the other hand, using the fact the convolution commutes with translation and the fact that translation is continuous in $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ we get for $h = g * f$, with $g \in \boldsymbol{L}^1(\mathbb{R}^d)$ and $f \in \boldsymbol{L}^\infty(\mathbb{R}^d)$:

LiconfLinf

$$\|T_x h - h\|_\infty \le \|T_x g - g\|_1 \|f\|_\infty \to 0, \quad \text{for } x \to 0. \tag{232}$$

Thus we can apply Lemma 12 in order to come to the conclusion that $\boldsymbol{B_A} \subset \boldsymbol{C}_{ub}(\mathbb{R}^d)$ in this case. But clearly approximate units from $\boldsymbol{L}^1(\mathbb{R}^d)$ (e.g. of the form $(\mathrm{St}_\rho k)_{\rho\to 0}$ for some $k \in \boldsymbol{C}_c(\mathbb{R}^d)$) act properly, hence we have the claimed equality.

□

ARGUMENT: One has to identify each element $b \in \boldsymbol{B}$ with the operator $T_b \in \mathcal{H}_{\boldsymbol{A}}(\boldsymbol{B})$ obtained by something like the "right regular representation", i.e.the operator $T_b : a \mapsto a \bullet b$. It is always true that $\|T_b\|_{op} \le \|b\|_{\boldsymbol{B}}$, and by applying T_b to the elements of some bounded approximate unit in $\boldsymbol{A}$ one finds the converse estimate, i.e.$(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ can be identified with a closed subspace of all bounded linear operators from $\boldsymbol{A}$ to $\boldsymbol{B}$.

Note that one should not forget that one has to impose the "natural" $\boldsymbol{A}$-module structure on $\mathcal{H}_{\boldsymbol{A}}(\boldsymbol{B}_1, \boldsymbol{B}_2)$, before making the claim that the $\boldsymbol{A}$-module $\boldsymbol{B}$ can be embedded via an $\boldsymbol{A}$-module homomorphism (embedding) into the larger $\boldsymbol{A}$-module $\boldsymbol{B^A}$.

Exercise: For the case of the pointwise algebra $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}}) = \left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$ one finds that $\mathcal{H}_{\boldsymbol{A}}(\boldsymbol{A}, \boldsymbol{A}) = \left(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\right)$ in a "natural way. Note that the "identity operator always belongs to $\mathcal{H}_{\boldsymbol{A}}(\boldsymbol{A}, \boldsymbol{A})$ (obviously it commutes with any other operator), and therefore the "enlargement" from $\boldsymbol{A}$ to $\mathcal{H}_{\boldsymbol{A}}(\boldsymbol{A})$ also implies the adjunction of a unit element to the Banach algebra $\boldsymbol{A}$, but typically much more than this. So in a way the (in this case isometric) embedding of $\boldsymbol{C}_0(\mathbb{R}^d)$ into $\boldsymbol{C}_b(\mathbb{R}^d)$ (both with the sup-norm) can be seen as an embedding of $\boldsymbol{A} = \boldsymbol{C}_0(\mathbb{R}^d)$ into the maximal algebra "with the same norm" (and a unit).

Plancherel's theorem can be used (we skip those details) to identify $\mathcal{H}_{\boldsymbol{L}^1}(\boldsymbol{L}^2, \boldsymbol{L}^2)$ (or equivalently the set of all bounded linear operators from $\boldsymbol{L}^2(\mathbb{R}^d)$ into $\boldsymbol{L}^2(\mathbb{R}^d)$ which commute with translation) with $\mathcal{H}_{\boldsymbol{A}}(\boldsymbol{L}^2, \boldsymbol{L}^2)$ (for $\boldsymbol{A} = \mathcal{F}\boldsymbol{L}^1$, i.e.the operators from $\boldsymbol{L}^2(\mathbb{R}^d)$ into $\boldsymbol{L}^2(\mathbb{R}^d)$ which commute with pointwise multiplications with elements from $\boldsymbol{A} = \mathcal{F}\boldsymbol{L}^1$, or equivalently, just with the multiplication with pure frequencies resp. characters on $\mathbb{R}^d$, which are the functions $x \mapsto e^{2\pi i s \cdot x}$. These are again pointwise multiplication operators, and it is not hard to find out that a pointwise multiplier h of $\boldsymbol{L}^2(\mathbb{R}^d)$ is has to be a measurable function which is essentially bounded, i.e.$h \in \boldsymbol{L}^\infty(\mathbb{R}^d)$.

Let us just sketch the basic idea behind this fact:

Lemma 65. *Assume that $\boldsymbol{B}$ is a Banach module with respect to a pointwise Banach algebra $\boldsymbol{A}$ (and assume that $\boldsymbol{A}_c(\mathbb{R}^d) = \boldsymbol{C}_c(\mathbb{R}^d) \cap \boldsymbol{A}$ is dense in $\boldsymbol{A}$), and assume that $\boldsymbol{A} \cap \boldsymbol{B}$ contains arbitrary large plateau-functions, i.e.with the property, that for each compact set $K \subseteq \mathbb{R}^d$ there exists some $q \in \boldsymbol{B} \cap \boldsymbol{A}$ such that $q(x) \equiv 1$ on K. Then the elements in $\mathcal{H}_{\boldsymbol{A}}(\boldsymbol{B}, \boldsymbol{B})$ are pointwise multipliers with suitable functions h which belong locally to $\boldsymbol{A}$.*

Proof. TO BE GIVEN LATER ON! □

ptwmultalg1 **Lemma 66.** *Assume that an operator S on some function space $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ satisfies the property $S(h \cdot f) = h \cdot S(f)$ for a sufficiently rich family of pointwise multipliers h on $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$. Then $S(f) = h_0 \cdot f$ for some function h_0.*

Proof. Note that in the case $G = \mathbf{Z}_N$ the claim is simply: Assume a linear mapping from $\mathbb{C}^N$ into itself, represented by a matrix, commutes with pointwise multiplication with unit vectors then it must be a diagonal matrix. In fact, such linear mappings do not increase the support of a function. In fact, assume that some coordinates of f are zero, then f does not change of it is multiplied with the sequence h taking the value one on those specific coordinates. On the other hand after pulling out of the argument one finds that also $h \cdot S(f) = S(h\dot{f}) = S(f)$ must be zero at the same coordinated, which easily implies that $N \times N$ matrix representing the linear mapping $f \mapsto S(f)$ must be a diagonal matrix, resp. S must be a multiplication operator (with $h_0 = \mathrm{diag}(S)$).

Let us now do the continuous version of the argument. For simplicity we assume that $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ allows pointwise multiplications by the elements of some nice, regular Banach algebra (such as $\left(\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_{\mathcal{F}\boldsymbol{L}^1}\right)$), and that we can make use of BUPUs, i.e. we have a bounded family (of pointwise multipliers on $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$) such that $\mathbf{1} = \sum_{\lambda\in\Lambda} \psi_\lambda$. Using (as usual now) we have $\psi = \psi \cdot \psi^*$, hence also $f = \sum_{\lambda\in\Lambda} \psi_\lambda \cdot \psi_\lambda^* f$, hence (assuming now that the series is norm convergent in $\boldsymbol{L}^2$ or $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$), for example, and setting $h_0 := \sum_{\lambda\in\Lambda} S(\psi_\lambda)$. Note that the sequence is pointwise well convergent, since $\psi_\lambda^* \cdot \psi_\lambda = \psi_\lambda$ implies the $\mathrm{supp}(S\psi_\lambda) \subseteq \mathrm{supp}(\psi_\lambda^*)$, but only finitely many of them overlap! On the other hand one has

ointw-mult1
$$Sf = S\left(\sum_{\lambda\in\Lambda} \psi_\lambda \cdot f\right) = \sum_{\lambda\in\Lambda} S(f \cdot \psi_\lambda) = f \cdot \sum_{\lambda\in\Lambda} S(\psi). \tag{233}$$

□

For us an important Banach algebra is $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$, endowed with convolution as multiplication (which is commutative, due to the commutativity of addition in $\mathbb{R}^d$). It does not have any units, but we have shown earlier (?) that $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ has bounded approximate units. Typically such a family is obtained by taking any sequence f_n (or net) of functions, e.g. in $\boldsymbol{C}_c(\mathbb{R}^d)$, with $\int_{\mathbb{R}^d} f_n(x)dx = 1$ and “shrinking support”. This can be obtained by choosing the shape of f_n arbitrarily, but assuming that $f_n(x) = 0$ for $|x| \geq \delta_n$ for some null-sequence $\delta_n \to 0$ for $n \to \infty$. Alternatively one compresses a given function $f \in \boldsymbol{C}_c(\mathbb{R}^d)$ or more generally in $\boldsymbol{L}^1(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d} f_n(x)dx = 1$, and chooses $f_n = St_{\rho_n} f$, for some sequence $\rho_n \to 0$ for $n \to \infty$.

The choice $f(x) = e^{-\pi x^2}$ is a popular choice (which also shows that it is not important for f to be compactly supported).

We will see shortly that $\mathcal{H}_{\boldsymbol{L}^1}(\boldsymbol{L}^1) = \mathcal{H}_{\boldsymbol{L}^1}(\boldsymbol{L}^1, \boldsymbol{L}^1)$ can be identified with the space $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ of all bounded measures (resp. with $\big(\boldsymbol{C}_0{}'(\mathbb{R}^d),\, \|\cdot\|_{\boldsymbol{M}}\big)$). This result is called "Wendel's theorem" ([67,103]). It has of course two parts: First of all one has to show that convolution operators induced by elements from $\boldsymbol{M}_b(\mathbb{R}^d)$ leave the closed subspace $\boldsymbol{L}^1(\mathbb{R}^d)$ invariant. In the second part one has to verify that any abstract (bounded and linear) operator on $\big(\boldsymbol{L}^1(\mathbb{R}^d),\, \|\cdot\|_1\big)$ commuting with all the translation must be such a convolution operator.

Wendel-thm1 **Theorem 28.** *The space of $\mathcal{H}_{\boldsymbol{L}^1}(\boldsymbol{L}^1, \boldsymbol{L}^1)$ all bounded linear operators on $\boldsymbol{L}^1(G)$ which commute with translations (or equivalently: with convolutions) is naturally and isometrically identified with $(\boldsymbol{M}_b(G), \|\cdot\|_{\boldsymbol{M}_b})$.*

Proof. For the first part we have to verify that $\boldsymbol{L}^1(\mathbb{R}^d)$ is a closed ideal of $\boldsymbol{M}_b(\mathbb{R}^d)$, i.e.that $\boldsymbol{M}_b(\mathbb{R}^d) * \boldsymbol{L}^1(\mathbb{R}^d) \subseteq \boldsymbol{L}^1(\mathbb{R}^d)$. One way to do that is to check that $\big(\boldsymbol{L}^1(\mathbb{R}^d),\, \|\cdot\|_1\big)$ is a *homogeneous Banach space*, i.e.to show that the group $\mathbb{R}^d$ acts in a continuous and isometric way on $\boldsymbol{L}^1(\mathbb{R}^d)$. This means that we have to verify the following conditions: $\|T_x f\|_1 = \|f\|_1$ for all $x \in \mathbb{R}^d$ and all $f \in \boldsymbol{C}_c(\mathbb{R}^d)$ (hence all $f \in \boldsymbol{L}^1(\mathbb{R}^d)$) and also

ontLi-shift
$$\lim_{x\to 0} \|T_x f - f\|_1 = 0 \quad \text{for} \quad x \to 0, \tag{234}$$

for any $f \in \boldsymbol{C}_c(\mathbb{R}^d)$, hence (by approximation for any $f \in \boldsymbol{L}^1(\mathbb{R}^d)$).

That $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$ (viewed as a Banach algebra with respect to convolution) is acting boundedly on *any homogeneous Banach space* will be discussed separately.

Alternatively one can describe $\boldsymbol{L}^1(\mathbb{R}^d)$ as the subset of all bounded measures which have continuous translation, in other words, one can show (see a classical paper by Plessner, 1929) that $\|T_x\mu - \mu\|_{\boldsymbol{M}_b} \to 0$ for $x \to 0$ implies that μ is an "absolutely continuous" measure, i.e.belongs to $\boldsymbol{L}^1(\mathbb{R}^d)$.

The argument for this result typically relies on a compactness argument (w^*- compactness of the unit ball in the dual Banach space $\boldsymbol{M}_b(\mathbb{R}^d) = \big(\boldsymbol{C}_0{}'(\mathbb{R}^d),\, \|\cdot\|_{\boldsymbol{M}}\big)$). One applies the given operator $T \in \mathcal{H}_{\boldsymbol{L}^1}(\boldsymbol{L}^1)$ to any *Dirac- sequence* f_n which forms a bounded approximate unit in $\boldsymbol{L}^1(\mathbb{R}^d)$. By the boundedness of (f_n) and the operator T the image sequence $T(f_n)$ is also bounded in $\boldsymbol{L}^1(\mathbb{R}^d)$, hence in the larger (dual) space $(\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b}))$. By the w^*-compactness of bounded balls in this space we obtain a w^*-convergent subnet, with some limit $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$. It remains to show that this limit is inducing the operator, i.e.one has to verify that $T = C_\mu$. $\square$

Corollary 7. *In the situation described in Wendel's theorem we have: Every multiplier of $\big(\boldsymbol{L}^1(G),\, \|\cdot\|_1\big)$ has the property that the restriction to $\boldsymbol{L}^1 \cap \boldsymbol{C}_0(G)$ (endowed with the natural norm $\|f\|_S := \|f\|_1 + \|f\|_\infty$) is also a multiplier on that space, which in turn is dense both $\big(\boldsymbol{L}^1(G),\, \|\cdot\|_1\big)$ as well as $\big(\boldsymbol{C}_0(G),\, \|\cdot\|_\infty\big)$. The same is true for multipliers on $\big(\boldsymbol{C}_0(G),\, \|\cdot\|_\infty\big)$.*

Also the *converse* is true, at least for $G = \mathbb{R}^d$: every multiplier on the Segal algebra $\boldsymbol{L}^1 \cap \boldsymbol{C}_0(G)$. A proof should be found in Larsen's book [67], Corollary 3.5.2.

Material of Nov. 9-th (! given below) is partially covered by the paper "Banach spaces of distributions having two module structures J. Funct. Anal. (1983)" ([13]). The main result of this paper is a "chemical diagram" that can be attached to each of the spaces in >>> standard situation:

Another Segal algebra is $\boldsymbol{L}^1 \cap \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$ with the natural norm $\|f\|_S := \|f\|_1 + \|\hat{f}\|_1$.

Some comments on the classical Riemann-Stieltjes integral (in German)

`http://de.wikipedia.org/wiki/Stieltjes-Integral`
`http://de.wikipedia.org/wiki/Beschr%C3%A4nkte_Variation`
`http://de.wikipedia.org/wiki/Absolut_stetig`
`http://de.wikipedia.org/wiki/Satz_von_Radon-Nikodym`
`http://de.wikipedia.org/wiki/Lebesgue-Integral`

Some of the material in this course has been already given in courses in Heidelberg (1980), Maryland (1989/90) or at the university of Vienna in the last 20 years.

The material concerning the Segal algebra $\boldsymbol{S}_0(G)$ is going back to various original publications by the author, see for example [30], where this particular Segal algebra has been introduced and where it is shown that it is the *minimal* TF-isometric homogeneous Banach space (and many other properties). The role of the dual space has been described already in [29] (both papers downloadable from the NuHAG site). The *double module view-point* is described in much detail in [13] (Banach spaces of distributions having two module structures, J. Funct. Anal.). A detailed account of notions of *compactness* (and also a clean description of tightness, etc.) is given in [32]. The first atomic characterization of modulation spaces ($\boldsymbol{S}_0(\mathbb{R}^d)$ is among them) has been given at a conference in Edmonton in 1986 (published then in [33]).

There are many places where especially the role of the Segal algebra $\boldsymbol{S}_0(G)$ for the discussion of basic questions in Gabor Analysis has been described. The very first systematic discussion as been probably given in the Chapter by Feichtinger and Zimmermann in the first Gabor book of 1998 ([47]). Another relevant paper is the one by Feichtinger and Kaiblinger ([42]) where it is shown, that (in the $\boldsymbol{S}_0(\mathbb{R}^d)$ context) the dual window is depending continuously on the lattice constants in the case of Gabor frames resp. Gabor Riesz bases.

Preview: In order read about Gabor multipliers the best source is probably the survey article in the second (blue) Gabor book, "Advances in Gabor Analysis", by Feichtinger and Nowak ([46]).

General references are: Hans Reiter's book on Harmonic Analysis (including a very fine and compact introduction to Integration Theory over Locally Compact Groups, but without the proof of the existence of the Haar measure) [79]. An updated version (edited by his former PhD student Ian Stegeman is [81]). The book describes (see also [80]) the concepts of Segal algebras (such as $S_0(G)$), and Beurling algebras $\boldsymbol{L}^1_w(G)$ (with respect to multiplicative weights). Both books are available in the NuHAG library.

A nice introduction into "abstract harmonic analysis" is that of Deitmar ([23]), and of course always Katznelson (starting with classical Fourier series, but also talking about the Gelfand theory for commutative C^*-algebras) is [64]. It also contains the first "propagation" of the concept of *homogeneous Banach spaces*. A similar unifying viewpoint is taken in the book of Butzer and Nessel ([16]) and of course several of the books of Hans Triebel (see BIBTEX collection and book-list). These two mathematicians can also be seen as pioneers of interpolation theory and the so-called theory of function spaces. (Abstract) Homogeneous Banach spaces are also treated in the Lecture Notes by H. S. Shapiro ([91]), having approximation theoretic questions in mind (he makes the associativity of the action of bounded measures an axiom, obviously because he could proof it only in concrete cases).

Solid Banach spaces of function (under the name of *Banach function spaces*) appear in the work of Zaanen: [106] (or [107]), both books should be available in the NuHAG library (Vienna, Oskar-Morgenstern-Platz 1, Level 5, Room 1.531, ca. 900 books!)

A very good source to learn about Besov spaces is Jaak Peetre's book entitled "New Thoughts on Besov Spaces" ([75]).

Banach modules: Rieffel's work [83].

A NEW book (2019) on Numerical Fourier analysis is [77] (Plonka, Potts, Steidl und Tusche).

Generally interesting reference on "mathematics and signal processing": R.Holmes: [60]

What are standard spaces?

Banach spaces of functions or distributions which are useful for "daily use", and which have properties typically available on concrete situations. In fact most often results are presented in the literature for $\boldsymbol{L}^p$-space, although they apply to much wider classes of function spaces.

But what could be a sufficiently large family (in order to capture a large variety of concrete examples), which still allows to prove an extensive list of interesting properties. Well, we think that such space should (among others) allow sufficiently many *regularization operators*, by convolution) but also *localization* (by pointwise multiplication). Which kind of objects do we want? Banach spaces of continuous functions? Banach spaces of locally (Lebesgue-) integrable functions? Banach spaces of Radon measures, or (tempered?) distributions? Should we allow even ultra-distributions? Wishes: We would like to be able to do functional analysis, so with each spaces we would like to have the dual space in the same family (as long as it can be viewed as a Banach space of distributions, hence only if it can be completely characterized by the sum of the local actions).

With each space the Fourier transform of the space should be in the same family, etc. etc.

Formal suggestion in this direction is given by the following definition, which describes at "reasonable" generality a family of Banach spaces which is *not* restricted to Banach spaces of functions, because such a family will typically *not* be closed under duality (an exception is the family of $\boldsymbol{L}^p$-classes, but already the dual space of $\boldsymbol{C}_0(\mathbb{R}^d)$ contains discrete measures which are not represented by (integrals against measurable) functions.

ardspacdef0 **Definition 46.** A Banach space $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ is called a *(restricted) standard space* if

(1) $\big(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\big) \hookrightarrow (\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}}) \hookrightarrow (\boldsymbol{S}_0'(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0'})$ (continuous embeddings);

(2) $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d) \cdot \boldsymbol{B} \subseteq \boldsymbol{B}$, with $\|h \cdot f\|_{\boldsymbol{B}} \leq \|h\|_{\mathcal{F}\boldsymbol{L}^1} \|f\|_{\boldsymbol{B}}$ for $h \in \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d), f \in \boldsymbol{B}$;

(3) $\boldsymbol{L}^1(\mathbb{R}^d) * \boldsymbol{B} \subseteq \boldsymbol{B}$ with $\|g * f\|_{\boldsymbol{B}} \leq \|g\|_{\boldsymbol{L}^1} \|f\|_{\boldsymbol{B}}$; for $g \in \boldsymbol{L}^1(\mathbb{R}^d), f \in \boldsymbol{B}$.

Remark 57. The main idea behind this specific definition (also the reason why it is called for a while the "*restricted standard situation*" is the fact that the fact that $\boldsymbol{S}_0(\mathbb{R}^d)$ and its dual $\boldsymbol{S}_0'(\mathbb{R}^d)$, or that the whole Banach algebra $\boldsymbol{L}^1(\mathbb{R}^d)$ is acting on $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ via convolution, can be seen as a matter of convenience. In this way we avoid a number of technical conditions involving weights and still have a fairly large collection of examples available. We will be able to demonstrate the roles of pointwise multiplication and convolutive action in the present context, and it will be easy for the reader to generalize the observations to more general situations.

It is clear that almost all the spaces used "normally" in Fourier analysis are such "standard spaces" (to be defined properly next). It is sufficient that a space (of locally integrable functions or Radon measures) is isometrically invariant under the time-frequency shifts $\pi(\lambda) = M_\omega T_t$ for $\lambda = (t, \omega) \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d$ and that e.g. the Schwartz space $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ is contained in $\boldsymbol{B}$ as a dense subspace, to ensure that the above conditions are satisfied. Let us formulate this claim as a lemma:

andardonRd1 **Lemma 67.** *Assume that $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ is a Banach space of locally integrable functions on $\mathbb{R}^d$ such that $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ is contained in $\boldsymbol{B}$ as a dense subspace and that $\|M_\omega T_t f\|_{\boldsymbol{B}} = \|f\|_{\boldsymbol{B}}$ for all $\lambda = (t, \omega) \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d$. Then it is a standard space.*

A typical alternative view in the context of $G = \mathbb{R}^d$ is the following setting:

tempBanSp **Definition 47.** [convenient description] .
A Banach space $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ is called a *tempered standard space* on $\mathbb{R}^d$ if

(1) $\boldsymbol{\mathcal{S}}(\mathbb{R}^d) \hookrightarrow (\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}}) \hookrightarrow \boldsymbol{\mathcal{S}}'(\mathbb{R}^d)$ (continuous embeddings);
(2) $\boldsymbol{\mathcal{S}}(\mathbb{R}^d) \cdot \boldsymbol{B} \subseteq \boldsymbol{B}$
(3) $\boldsymbol{\mathcal{S}}(\mathbb{R}^d) * \boldsymbol{B} \subseteq \boldsymbol{B}$

Aside from the fact that one needs some functional analytic argument in order to establish the equivalence between this "convenient" and another more technical one (which is however what one needs in order to make those concepts useful). It will be convenient for this purpose to make use of polynomial (submultiplicative) weights w_s, given by $w_x : x \mapsto (1+|z|^2)^{s/2}$:

empStandDef **Definition 48.** [technical definition] A Banach space $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ is called a *tempered standard space* on $\mathbb{R}^d$ if

(1) $\boldsymbol{\mathcal{S}}(\mathbb{R}^d) \hookrightarrow (\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}}) \hookrightarrow \boldsymbol{\mathcal{S}}'(\mathbb{R}^d)$ (continuous embeddings);
(2) There $s \geq 0$ such that $\boldsymbol{L}^1_{w_s}(\mathbb{R}^d)$ acts on $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ by convolution and $\|g * f\|_{\boldsymbol{B}} \leq C_s \|g\|_{1,w_s} \|f\|_{\boldsymbol{B}};$ for $g \in \boldsymbol{L}^1_{w_s}, f \in \boldsymbol{B}$, for some $C_s > 0$;
(3) $\boldsymbol{\mathcal{S}}(\mathbb{R}^d) \cdot \boldsymbol{B} \subseteq \boldsymbol{B}$ and there exists some constant $\|h \cdot f\|_{\boldsymbol{B}} \leq \|\widehat{h}\|_{1,w_r} \|f\|_{\boldsymbol{B}}$;

Remark 58. The typical examples of *reduced standard spaces* arise from Banach spaces of say tempered distributions which are isometrically invariant under TF-shifts, i.e.which satisfy

$$\|\pi(t,\omega)f\|_{\boldsymbol{B}} = \|f\|_{\boldsymbol{B}} \quad \forall f \in \boldsymbol{B}.$$

and which contain $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ or $\boldsymbol{S}_0(\mathbb{R}^d)$ as a dense subspace. In fact, in such a case on can argue that the isometric invariance of the space implies that the continuity of the mapping (t,ω) into $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ resp. $\boldsymbol{S}_0(\mathbb{R}^d)$ implies that one can extend the strong continuity to all of $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$, in other words, one obtains a so-called *time-frequency homogeneous Banach space* $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ in this case, and the mapping (t,ω) to $\pi(t,\omega)(f)$ is continuous for *every* $f \in \boldsymbol{B}$.

One can also discuss from a technical side the need of assuming that the embedding from $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ into $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ should be a continuous one with respect to the occurring natural topologies. In fact, it should be enough, for example, to verify that $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ itself is continuously embedded into the space of all locally integrable functions and that for each norm convergent sequence (f_n) in $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ there exists a subsequence (f_{n_k}) which is pointwise convergent almost everywhere. etc. ...

Standard spaces are also described within the TEXblock system (part of the NuHAG DB system):

`http://www.univie.ac.at/nuhag-php/tex-blocks/search.php?id=19`
`Talk Server: TO BE DONE`

Starting from the observation that for any of the module actions, arising from the pointwise algebra $\boldsymbol{A} = \mathcal{F}\boldsymbol{L}^1$ and the convolutional Banach module structure over $\boldsymbol{L}^1$ one can build two types of completions and also two typos of "essential parts". We will write $\boldsymbol{B}^{\boldsymbol{A}}$ for the $\boldsymbol{A}-$completion of $\boldsymbol{B}$, and $\boldsymbol{B}_{\boldsymbol{A}}$ for the essential part with respect to the pointwise module action. It is easy to verify that an element is in $\boldsymbol{B}_{\boldsymbol{A}}$ if and only if it can be approximated by elements with compact support, or if and only if any bounded sequence of plateau-like functions (forming a bounded approximate unit in $\mathcal{F}\boldsymbol{L}^1$) acts as approximation to the identity operator on the given element.

Analogously we define the completion and the essential part with respect to the Banach algebra $\boldsymbol{L}^1$. Since the action of this algebra usually comes from the group action (by translation), we will use the symbols $\boldsymbol{B}^G$ and $\boldsymbol{B}_G$.

Combining those four operations in a serial way we can come up with a large number of new spaces, derived from any of those spaces. Since the operations of completion and essential part with respect to the *same* algebra action are canceling each other (similar to the operation of taking a closure resp. the interior of a "nice set") we can concentrate in our discussion on "mixed series", such as: ${\boldsymbol{B}_G}^{\boldsymbol{A}}{}_G$, or even longer chains of operations of a similar kind.

The result that has been derived in [13] can be summarized in the following way: JUST the LAST OCCURRENCE of each algebra operation counts, i.e.the last occurrence of the symbol G and the last occurrence if $\boldsymbol{A}$. So we have ${\boldsymbol{B}_G}^{\boldsymbol{A}}{}_G = \boldsymbol{B}^{\boldsymbol{A}}{}_G$ or ${\boldsymbol{B}_G}^{\boldsymbol{A}}{}_{G\boldsymbol{A}} = \boldsymbol{B}_{G\boldsymbol{A}}$.

The most important spaces in this family are the minimal space, which turns out to be the "double essential part", where the order of algebra operations does *not* play a role anymore. It coincides with the closure of the test functions in the standard spaces. So we have

doubesspart
$$\boldsymbol{B}_{\boldsymbol{A}G} = \overline{\boldsymbol{S}_0}^{\boldsymbol{B}} = \boldsymbol{B}_{G\boldsymbol{A}} \tag{235}$$

On the other hand here is a (single) double completion, which coincides with the w^*-relative completion of $\boldsymbol{B}$ within $\boldsymbol{S}_0'$ (details to be presented at another time).

doublcompl
$$\boldsymbol{B}^{\boldsymbol{A}G} = \widetilde{\boldsymbol{B}} = \boldsymbol{B}^{G\boldsymbol{A}} \tag{236}$$

where we define the vague or w^*-relative completion of $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ in $\boldsymbol{S}_0'$ as follows:

rel-complet
$$\widetilde{\boldsymbol{B}} = \{\sigma \in \boldsymbol{S}_0' \mid \sigma = w^* - lim_\alpha f_\alpha, \sup_\alpha \|f_\alpha\|_{\boldsymbol{B}} < \infty\} \tag{237}$$

The infimum over all the bounds $\sup_\alpha \|f_\alpha\|_{\boldsymbol{B}}$ makes $\widetilde{\boldsymbol{B}}$ into a standard space, which contains $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ as a closed subspace.

Example:
Starting from $\boldsymbol{C}_0(\mathbb{R}^d)$ one can find that is relative completion is just $\big(\boldsymbol{L}^\infty(\mathbb{R}^d), \|\cdot\|_\infty\big)$.

A WIKIPEDIA contribution:

```
http://en.wikipedia.org/wiki/Harmonic_analysis
http://en.wikipedia.org/wiki/Tempered_distribution
http://en.wikipedia.org/wiki/Fourier_analysis
http://en.wikipedia.org/wiki/Discrete-time_Fourier_transform
http://en.wikipedia.org/wiki/Colombeau_algebra
```

15. Integrated Group Action: isometric case

Next we will show that the convolution action of bounded discrete measures on a *homogenous Banach space* can be extended to all of the measures in order to generate an action of $(\boldsymbol{M}_b(G), \|\cdot\|_{\boldsymbol{M}_b})$ on such a Banach space $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$.

LATER on we will relabel the symbols and write $\mu \bullet_\pi f$ instead of $\mu \bullet_\rho f$!

meas-homBR0 **Theorem 29** (Bounded Measures act on Homogeneous Banach Spaces)**.** *Assume that $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ is a homogeneous Banach-space with respect to some "abstract" group action ρ, i.e. we assume that $x \to \rho(x)f$ is continuous from G into $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$, isometric in the sense that $\|\rho(x)f\|_{\boldsymbol{B}} = \|f\|_{\boldsymbol{B}}$ for all $x \in G$ and $f \in \boldsymbol{B}$, and that $\rho(x_1x_2) = \rho(x_1)\rho(x_2)$. Of course we can define $\mu \bullet_\rho f$ for any discrete measure, hence for the family $D_\Psi \mu$ as the (unique) limit of this Cauchy net in $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$. Then one has*

meas-onB1
$$\|\mu \bullet_\rho f\|_{\boldsymbol{B}} \le \|\mu\|_{\boldsymbol{M}} \|f\|_{\boldsymbol{B}}, \quad \text{for all } f \in \boldsymbol{B}. \tag{238}$$

In fact, $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ becomes a Banach module over $(\boldsymbol{M}_b(G), \|\cdot\|_{\boldsymbol{M}_b})$ in this way.

Proof. The starting point is of course to translate the action of the group G into an action of the Dirac measures, i.e.to set

rhobuldelta
$$\delta_x \bullet_\rho f =: \rho(x)f, \quad f \in \boldsymbol{B}. \tag{239}$$

Since discrete measures are absolutely convergent sums of Dirac measures it is then clear that we have for a discrete measure (according to definition 50) $\mu = \sum_{n=1}^{\infty} c_k \delta_{x_k}$ simply

$$\mu \bullet_\rho f = \sum_{n=1}^{\infty} c_k \rho(x_k) f,$$

the sum being absolutely convergent for each f and $\mu \in \boldsymbol{M}_d(\mathbb{R}^d)$. Observe also that the assumption about ρ imply that this action of $\boldsymbol{M}_d(\mathbb{R}^d)$ is not just an individual action (given for each $\mu \in \boldsymbol{M}_d(\mathbb{R}^d)$) but it defines in fact a representation of the Banach convolution algebra $(\boldsymbol{M}_d(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b})$, since we have (as an exercise for the reader) of course:

rhobulass1
$$(\mu_2 * \mu_1) \bullet_\rho f = \mu_2 \bullet_\rho (\mu_1 \bullet_\rho f), \quad \mu_1, \mu_2 \in \boldsymbol{M}_d(\mathbb{R}^d),\ f \in \boldsymbol{B}. \tag{240}$$

Consequently, for any given $\mu \in \boldsymbol{M}_b(G)$ and some $f \in \boldsymbol{B}$ we have an clear interpretation of the (bounded) net

$$D_\Psi \mu \bullet_\rho f = \sum_{i \in I} \mu(\psi_i) \rho(x_i) f.$$

We will show next that it is convergent, as $|\Psi| \to 0$ or $\mathrm{diam}(\Psi) \to 0$.

Motivation: One should consider the action of $D_\Psi \mu$ on f as a Riemann-type sum for the Banach-space valued integral of $x \to \rho(x)f$, usually written as $\int_G \rho(x) f(x) d\mu(x)$. Therefore it is natural to check that the action of bounded discrete measures is OK (this is an easy consequence of the assumptions) and then to compare two such expressions, namely $DPsimu \bullet_\rho f$ and $D_\Phi \mu \bullet_\rho f$ by making use of their joint refinement, constituted by the (double indexed family) $(\psi_i \phi_j)$.

Let us first estimate the norm of $D_\Psi \mu \bullet_\rho f$. Using the isometry of the action of ρ on $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ one has, independently from Ψ:

crmeasconv1
$$\|D_\Psi \mu \bullet_\rho f\|_{\boldsymbol{B}} \le \sum_{i \in I} |\mu(\psi_i)| \|\rho(x_i) f\|_{\boldsymbol{B}} \le \|f\|_{\boldsymbol{B}} \sum_{i \in I} |\mu(\psi_i)| \le \|\mu\|_{\boldsymbol{M}} \|f\|_{\boldsymbol{B}}. \tag{241}$$

Assume next that there are two families $\Psi = (\psi_i)_{i\in I}$ and $\Phi = (\phi_j)_{j\in J}$ are given, with central points $(x_i)_{i\in I}$ and $(y_j)_{j\in}$. Then we can define the *joint refinement* $\Psi - \Phi$ as the family $(\psi_i\phi_j)_{(i,j)\in I\diamond J}$, where we can agree to call $I \diamond J$ the family of all index pairs such that $\psi_i \cdot \phi_j \neq 0$ (because all the other products are trivial and should be neglected). In fact, if both Ψ and Φ are sufficiently "fine" BUPUs one has: [85]

mann-estim1
$$\|D_\Psi\mu\bullet_\rho f - D_\Phi\mu\bullet_\rho f\|_{\boldsymbol{B}} = \sum_{(i,j)\in I\diamond J} \|\rho(x_i)f - \rho(y_j)f\|_{\boldsymbol{B}}|\mu(\psi_i\phi_j)| \leq \tag{242}$$
$$\sup_{(i,j)\in I\diamond J} \|\rho(x_i)[f - \rho(y_j - x_i)f]\|_{\boldsymbol{B}} \sum_{(i,j)\in I\diamond J} \|(\psi_i\phi_j)\mu\|_{\boldsymbol{M}} \leq \varepsilon\|\mu\|_{\boldsymbol{M}},$$

if only Ψ resp. Φ are fine enough. Due to the completeness of $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ one finds that there is a uniquely determined limit, which we will call $\mu\bullet_\rho f$. It is then obvious that

astf-estim1
$$\|\mu\bullet_\rho f\|_{\boldsymbol{B}} = \lim_{|\Psi|\to 0} \|D_\Psi\mu\bullet_\rho f\|_{\boldsymbol{B}} \leq \limsup \|D_\Psi\mu\|_{\boldsymbol{M}}\|f\|_{\boldsymbol{B}} = \|\mu\|_{\boldsymbol{M}}\|f\|_{\boldsymbol{B}}. \tag{243}$$

Of course it remains to show that the action defined in this way is associative, i.e. that

assocat
$$(\mu * \mu_2)\bullet_\rho f = (\mu_1)\bullet_\rho(\mu_2\bullet_\rho f), \quad \forall \mu_1, \mu_2 \in \boldsymbol{M}_b(\mathbb{R}^d), f \in \boldsymbol{B}, \tag{244}$$

but his is clear because the associativity is valid for the discrete measures $D_\Psi\mu$ and $D_\Phi\mu$. [86] □

Remark 59. In the derivation above we have used the isometric property and the fact that $\rho(x_1x_2) = \rho(x_1) \circ \rho(x_2)$. It would have been no problem if this identity was only true "up to some constant of absolute value one", i.e. if one has a *projective representation* of G only, such as the mapping $\lambda = (t,\omega) \mapsto \pi(\lambda) = M_\omega T_t$ from $\mathbb{R}^d \times \widehat{\mathbb{R}}^d$ into the unitary operators on the Hilbert space $\big(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\big)$, which is one of the key players in *time-frequency analysis.*

Remark 60. The following remark may be viewed almost as a triviality (but I think it will be useful later on): Given two Banach spaces as described in the above theorem (Thm. **??**), with $(\boldsymbol{B}^1, \|\cdot\|^{(1)}) \hookrightarrow (\boldsymbol{B}^2, \|\cdot\|^{(2)})$. Then - as we assume - there is an action of ρ on $\boldsymbol{B}^2$ whose restriction to $\boldsymbol{B}^1$ leaves that subspace invariant. Then it obvious that for the symbol $\mu\bullet_\rho f$ is justified without distinction of the corresponding homogeneous Banach spaces. In fact $f \mapsto \mu\bullet_\rho f$ considered as a mapping on $(\boldsymbol{B}^1, \|\cdot\|^{(1)})$ can be viewed as the restriction of $f \mapsto \mu\bullet_\rho f$ realized as a bounded linear mapping on $(\boldsymbol{B}^1, \|\cdot\|^{(1)})$.

15.1. **Isometric mappings and partial inversion on Hilbert spaces.** For the next step we need a simple observation from abstract Hilbert space theory.

Lemma 68. *Assume that a (complex) linear mapping between two Hilbert space over the complex numbers, $\mathscr{H}_1 \to \mathscr{H}_2$ is isometric, i.e. satisfies*

isom-embed1
$$\|T(h)\|_{\mathscr{H}_2} = \|h\|_{\mathscr{H}_1} \quad \forall h_1 \in \mathscr{H}_1. \tag{245}$$

Then the adjoint mapping $T' : \mathscr{H}_2 \to \mathscr{H}_1$ is the inverse on the range, i.e. one has

adjinvers
$$T'(Tf) = f \quad \forall h_1 \in \mathscr{H}_1. \tag{246}$$

[85]Using that $\psi_i = \sum_{j\in j} \psi_i\phi_j$, hence $\sum_{(i,j)\in I\diamond J} \psi_i\phi_j \equiv 1$ and $\sum_{(i,j)\in I\diamond J} \|(\psi_i\phi_j)\mu\|_{\boldsymbol{M}} = \|\mu\|_{\boldsymbol{M}}$.

[86]Note that H.S.Shapiro (cf. [91]) is making this associativity an extra axiom, apparently because he could not proof it directly for technical reasons, because he defines the action of the bounded measures on an "abstract homogeneous Banach space". H.C. Wang exhibits in [101] an example of what he calls a *semi-homogeneous Banach space* (without strong continuity of the action of G on $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$, which does not allow the extension to all of the bounded measures. Indeed, it is a Banach space of measurable and bounded functions on $\mathbb{R}$ which is non-trivial, but which does not contain any non-zero *continuous* function! The example was suggested to him in a correspondence by the author of these notes.

Proof. The claim follows from the fact that an isometric embedding also preserves scalar products, as a consequence of the polarization identity

larization

$$\langle f, g\rangle = \frac{1}{4}\sum_{k=0}^{3} i^k \langle f + {}^k g, f + i^k g\rangle = \frac{1}{4}\sum_{k=0}^{3} i^k \|f + i^k g\|_2^2. \tag{247}$$

Hence

isom-adjinv

$$\langle T'(Tf), g\rangle = \langle Tf, Tg\rangle = \langle f, g\rangle \quad \forall f, g \in \mathscr{H}_1. \tag{248}$$

Since this is true for every $g \in \mathscr{H}_1$ the required claim is valid. Usually one says that $T'(h_2)$ is defined in the weak sense for $h_2 \in \mathscr{H}_2$, through the identity

weak-Tadj

$$\langle T'h_2, h_1\rangle_{\mathscr{H}_1} = \langle h_2, T(h_1)\rangle_{\mathscr{H}_2}, \quad h_1 \in \mathscr{H}_1, h_2 \in \mathscr{H}_2. \tag{249}$$

□

Application: $(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0})$ is defined via its STFT: $f \in \boldsymbol{L}^2(\mathbb{R}^d)$ belongs to $\boldsymbol{S}_0(\mathbb{R}^d)$ if and only if $V_{g_0} f \in \boldsymbol{L}^1(\mathbb{R}^{2d})$, where g_0 is the Gauss-function (or any other nonzero Schwartz-function). Since $f \mapsto V_g f$ is isometric (assuming that $\|g_0\|_2 = 1$) we have according to the above lemma the (weak) reconstruction formula

$$f = \int_{\mathbb{R}^d \times \widehat{\mathbb{R}}^d} V_{g_0} f(\lambda)\pi(\lambda) g_0 \, d\lambda,$$

but if $V_{g_0} f \in \boldsymbol{L}^1(\mathbb{R}^{2d}) \subset \boldsymbol{M}(\mathbb{R}^{2d})$ then we have

$$f = V_{g_0} f \bullet_\pi g_0$$

in the spirit of the above abstract statement (for $\rho = \pi$). It follows that one has for every TF-homogeneous Banach space $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$, i.e. for every Banach space $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ such that $\|\pi(\lambda) f\|_{\boldsymbol{B}} = \|f\|_{\boldsymbol{B}}$ and $\|\pi(\lambda) f - f\|_{\boldsymbol{B}} \to 0$ for $\lambda \to 0$:

so-minimal1

$$\|f\|_{\boldsymbol{B}} = \|V_{g_0} f \bullet_\pi g_0\|_{\boldsymbol{B}} \le \|V_{g_0} f\|_{\boldsymbol{L}^1(\mathbb{R}^{2d})} \|g_0\|_{\boldsymbol{B}} = \|f\|_{\boldsymbol{S}_0} \|g_0\|_{\boldsymbol{B}}. \tag{250}$$

From now on we may use $\bullet_\pi$ instead of $\bullet_\rho$. Adjustment will be realized later on!

Proof. The proof relies on the fact that for any net of convergent nets Ψ_β (all sufficiently "fine") the overall family $(D_{\Psi_\beta}\mu_\alpha)_{\alpha,\beta}$ is bounded and uniformly tight! [87] Moreover it is clear that for each fixed β the net $D_{\Psi_\beta}\mu_\alpha$ is w^*-convergent to $D_{\Psi_\beta}\mu_0$. Given the tightness of the family only a finite number of indices of the family $(\psi_i^\beta)_{i\in I}$ is relevant for the convergence, hence $\mu_\alpha(\psi_i^\beta) \to \mu_0(\psi_i^\beta)$ implies that

$$D_{\Psi_\beta}\mu_\alpha \bullet_\pi f \to D_{\Psi_\beta}\mu_0 \bullet_\pi f.$$

aussdensapp

Lemma 69. *The collection of all dilated and shifted Gaussians form a total subset in most of the function spaces, including* $(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty)$, $(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1)$, $(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2)$, *but also in Wiener's algebra* $(\boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{W}})$.

Proof. The proof is quite easy observing first that we have $\|\mathrm{St}_\rho g * f - f\|_p \to 0$ for every $f \in (\boldsymbol{L}^p(\mathbb{R}^d), \|\cdot\|_p)$ resp. $f \in \boldsymbol{C}_0(\mathbb{R}^d)$, endowed with the sup-norm $\|f\|_\infty$. But for any fixed $\rho > 0$ the convolution product itself can be approximated by expressions of the form $\mathrm{St}_\rho g * D_\Psi(f)$. In any case we can start our considerations (without loss of generality) from some $f \in \boldsymbol{C}_c(\mathbb{R}^d)$.

[87] It is a good exercise to check the technical details yourself!

Hence by first choosing $\rho_0 > 0$ small enough, and subsequently (depending on the smoothness of $\mathrm{St}_\rho g$) a BUPU which is fine enough (suitable choice of $\delta_0 > 0$) such that one has altogether

approxgaus
$$\|f - \mathrm{St}_\rho g * D_\Psi f\|_\infty \leq \|f - \mathrm{St}_\rho g * f\|_\infty + \|\mathrm{St}_\rho g * (f - D_\Psi f)\|_\infty. \tag{251}$$

Hence by *first* choosing $\rho > 0$ and then (for fixed $\rho > 0$ the appropriate size of Ψ, i.e. $\delta > 0$ one can make the difference arbitrarily small. □

Since the finite linear combinations of such Gaussians are dense in all the space mentioned one can use them. For example, it is quite easy to derive the validity of the Fourier inversion formula on these functions. In fact it is better to use not only shifted and dilated, but also the modulated version of the Gauss function. Since the Gauss-function $g_0(t) := e^{\pi |t|^2}$ is invariant under the Fourier transform (cf. [108]) the rules for the intertwining properties between dilation, modulation and shifting implies that the Fourier inversion is valid for all linear combinations of such functions, and in fact certain infinite sums (as long as they are convergent). Such functions make up a so-called exotic Banach space ([48] gives details).

DPsi-tight1 **Lemma 70.** *Let Ψ run through the family of all BUPUs of a given maximal size Q, e.g. with all the function (ψ_i) constituting Ψ (whatever Ψ is used) satisfying with $supp(\psi_i) \subseteq x_i + Q$, for any such ψ and suitably chosen points (x_i), for some fixed compact set Q. Then for any $\mu \in \boldsymbol{M}_b(G)$ on has: $(D_\Psi \mu)_\Psi$ is a tight set.*[88]

Proof. We can start from any fixed BUPU Φ, and recall that $\sum_j \|\phi_j \cdot \mu\|_{\boldsymbol{M}} \leq \|\mu\|_{\boldsymbol{M}}$.

We can use this to approximate μ by a partial sum, i.e. by choosing for $\varepsilon > 0$ a finite set $F \subset J$ can be found such that

$$\|\sum_{j \in F} \phi_j \cdot \mu - \mu\|_{\boldsymbol{M}} = \sum_{j \notin F} \|\phi_j \mu\|_{\boldsymbol{M}} \leq \varepsilon.$$

or equivalently, writing $\Phi_F := \sum_{j \in F} \phi_j$:

$$\|\Phi_F \cdot \mu - \mu\|_{\boldsymbol{M}} \leq \varepsilon.$$

Let now $h \in \boldsymbol{C}_c(\mathbb{R}^d)$ by any plateau-function such that $h(x) \equiv 1$ on some neighborhood of $\mathrm{supp}(\Phi_F)$. More precisely, given the uniform size of the BUPUs Ψ to be considered, we can have $h \cdot \psi = \psi$ for all indices $i \in I$ such that $p \cdot \psi \neq 0$.

Then for any BUPU $\Psi = (\psi_i)_{i \in I}$ of size Q $\|h(D_\Psi \mu) - D_\Psi \mu\|_{\boldsymbol{M}}$ can be controlled in a uniform way[89].

ARGUMENTS:

ss-estimat1
$$|\mu(\phi)| \leq \|\phi \cdot \mu\|_{\boldsymbol{M}}, \quad \forall \phi \in \boldsymbol{C}_b(\mathbb{R}^d), \mu \in \boldsymbol{M}(\mathbb{R}^d), \tag{252}$$

because in case $\phi \in \boldsymbol{C}_c(\mathbb{R}^d)$ one can choose any other function $\phi^* \in \boldsymbol{C}_c(\mathbb{R}^d)$ such that $\|\phi^*\|_\infty = 1$ $\phi^*(y) = 1$ on $\mathrm{supp}(\phi)$ and hence

abs-estim2
$$\phi^* \cdot \phi = \phi \Rightarrow |\mu(\phi)| = |\mu(\phi^* \phi)| \leq |[\phi \cdot \mu](\phi^*)| \leq \|\phi \cdot \mu\|_{\boldsymbol{M}}, \tag{253}$$

thus completing the argument. The general case follows from this using the standard reduction to compactly supported measures[90].

[88] In the Euclidean case one will of course talk of BUPUs of size $R > 0$ if the set Q can be chosen to be the closed ball of radius $B_R(0)$ around zero. Hence the assumption is simply that $\mathrm{supp}(\psi_i) \subseteq B_R(x_i)$ for some $x_i \in \mathbb{R}^d$.

[89] Despite the fact that $p \cdot D_\Psi \mu$ is not equal to $\mathrm{D}_\Psi(p\mu)$

[90] to carry out the details is in fact a fairly good exercise, showing how well the technical details have been understood!

Given F we can find some $h \in \boldsymbol{C}_c(\mathbb{R}^d)$ with $h(x) \in [0,1]$, hence with $\|h\|_\infty = 1$, such that $h(x) \equiv 1$ on $\bigcup_{j \in F} \operatorname{supp}(\phi_j)$. (due to the limited size of the support this is a compact set which can be chosen independently from the concrete choice of Ψ). Consequently

Dpsi-estim3
$$\|h \cdot D_\Psi \mu - D_\Psi \mu\|_{\boldsymbol{M}} \leq \left\| \sum_{\{i \,|\, \psi_i \cdot p = 0\}} \mu(\psi_i)\delta_{x_i} \right\|_{\boldsymbol{M}} = \sum_{\{i \,|\, \psi_i \cdot p = 0\}} |[\mu - p\mu](\psi_i)| \leq \|\mu - p\mu\|_{\boldsymbol{M}} \leq \varepsilon. \tag{254}$$

□

Based on this estimate we observe that one has for any $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$ one has:

$$\|\mu \bullet_\rho f - D_\Psi \mu \bullet_\rho f\|_{\boldsymbol{B}} \leq \varepsilon \|\mu\|_{\boldsymbol{M}}$$

depending only on the element $f \in \boldsymbol{B}$ and the level of "refinement" of Ψ but NOT on the individual choice of μ. □

Next we want to show that there is an important form of continuity from in this action from $\boldsymbol{M}(G) \times \boldsymbol{B} \to \boldsymbol{B}$, with respect to the w^*-topology on $C: /NuHAGALL/NuHAGPDFs/$.

wst-to-norm
Theorem 30. *Assume that $(\mu_\alpha)_{\alpha \in I}$ is a bounded and tight net of bounded measures, which is w^*-convergent to some limit measure $\mu_0 \in \boldsymbol{M}(\mathbb{R}^d)$. Then for every $f \in \boldsymbol{B}$:*

st-to-norm1
$$\|\mu_\alpha \bullet_\rho f - \mu_0 \bullet_\rho f\|_{\boldsymbol{B}} \to 0 \quad \textit{for} \quad \alpha \to \infty. \tag{255}$$

Proof. One can verify this relation by observing that the family $D_\Psi \mu_\alpha$ running through any net of BUPUs with $|\Psi| \leq 1$ and $\alpha \in I$ is a tight family, hence up to some $\varepsilon > 0$ one can replace the given net (μ_α) by a family of measures with joint compact support. Hence all the BUPUs will consist of finitely many discrete measures, and in particular $\|D_\Psi \mu_\alpha \bullet_\rho f - \mathrm{D}_\Psi \mu_0 \bullet_\rho f\|_{\boldsymbol{B}} \to 0$ for any $f \in \boldsymbol{B}$, for fixed Ψ. □

Corollary 8. *For every homogeneous Banach space $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ one has: For every $\mu \in \boldsymbol{M}(\mathbb{R}^d)$ and every $f \in \boldsymbol{B}$: $\mu * f$ is the limit finite linear combinations translates of f. In particular one has: Given μ and f and $\varepsilon > 0$ there exists a finite sequence $(x_i)_{i \in F}$ and a finite sequence of complex coefficients $(c_i)_{i \in F}$ such that*

approxtrans
$$\left\| \mu * f - \sum_{i \in F} c_i T_{x_i} f \right\|_{\boldsymbol{B}} \leq \varepsilon. \tag{256}$$

Proof. We just have to remember that the discrete bounded measures $D_\Psi \mu$ resp. their partial sums form a tight family of measures, which is w^*-convergent towards μ. Since each of the approximating measures $D_\Psi \mu$ is of the form $\sum_{i \in I} \mu(\psi_i)\delta_{x_i}$ and can be approximated (even in the norm of $(\boldsymbol{M}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}})$) by finite sums we just have to put $c_i = \mu(\psi_i)$ and observe that

$$\left(\sum_{i \in F} c_i \delta_{x_i} \right) * f = \sum_{i \in F} c_i T_{x_i} f.$$

□

Remark 61. A more precise way of expressing what is going on is the use of suitable indexing, allowing to express that it is enough that "sufficiently many" elements from a "sufficiently fine" BUPU Ψ have to be used. Let us choose as index set pairs of the form (K, δ), with $\delta > 0$ and K a compact subset of $\mathbb{R}^d$. They have a partial order, with $(K_1, \delta_1) \succeq (K_2, \delta_2)$ if $K_2 \supseteq K_1$ and $\delta_2 \leq \delta_1$. Then we say that $\sum_{i \in F} c_i \delta_{x_i}$ has the index (K, δ) if Ψ is a δ-BUPU, and $F \supseteq \{i \in I \,|\, x_i \in K\}$. One could also talk of a local BUPU, by assuming (slightly differently, but technically equivalent) that $\sum_{i \in F} \psi(x) \equiv 1$ on K.

In this setting one can say: For every $\mu, f, \varepsilon > 0$ there exists a pair (K_o, δ_0) such that for all local BUPUs which satisfy at least $\sum_{i\in F}\psi(x) \equiv 1$ on K and are at least δ_0-fine, one has

nvtransapp2

$$\|\mu * f - \sum_{i\in F} \mu(\psi_i) T_{x_i} f\|_{\boldsymbol{B}} < \varepsilon. \tag{257}$$

The point of the last remark is that the user may even choose the points by himself, and then suitable coefficients can be found (by making use of an BUPU adapted to the given set), cf. [35].

A detailed discussion of similar arguments for the family of *translation and modulation invariant Banach spaces* (of ultra-distributions is given in the recent paper [41].

16. Sobolev spaces, derivatives in $L^2(Rd)$

Sobolev spaces and the relation with the Fourier transform: Relate the classical concept of derivatives or multiple derivatives (as done originally by Sobolev) to the properties of their Fourier transform.

iffer-Four1 **Lemma 71.** *Assume that $f \in \boldsymbol{C}_c(\mathbb{R}^d)$ has continuous partial derivatives. Then their Fourier transforms coincide with $\hat{f}$, multiplied with essentially the coordinate functions.*

Proof. We can restrict the attention to the case $d = 1$. Considering the fact that $f' \in \boldsymbol{C}_c(\mathbb{R}^d)$ by assumption we can guarantee that the difference quotients $\frac{1}{h}(T_h f - f)(x)$ converge *uniformly* to f', hence in the $\boldsymbol{L}^1$-norm.

Consequently their Fourier transforms are uniformly convergent to the Fourier transforms of f', using the mean-value theorem.

diff-four2
$$f(x+h) - f(x) = f'(\xi) \cdot h, \quad \text{for some} \quad \xi \in (x, x+h). \tag{258}$$

Since translation by h goes into multiplication with characters the rule follows by going to the limit

ff-fourier1
$$\frac{1}{h}(e^{2\pi ish} - 1) \to 2\pi is \cdot e^{2\pi ist}, \tag{259}$$

□

Definition 49. A strictly positive and (without loss of generality) continuous function w is called *submultiplicative* (resp. a *Beurling weight*) if it satisfies

Beurlingdef
$$w(x+y) \le w(x) \cdot w(y) \quad \forall x, y \in G. \tag{260}$$

Examples: $w_s(x) = (1 + |x|)^s$ for $s \ge 0$. It is trivial (Ex!) to prove it for $s = 1$ and then by taking the $s - th$ power it is still valid! The corresponding weighted $\boldsymbol{L}^2$ on $\mathbb{R}^d$ is denoted by $\boldsymbol{L}^2_{w_s}(\mathbb{R}^d)$.

It is easy to verify that this implies a pointwise estimate (using $w(x) \le w(x-y) \cdot w(y)$ in the convolution integral):

conv-weight
$$|(f * g) \cdot w| \le |f|w \, * \, |g|w. \tag{261}$$

In fact, we have for every fixed $x \in G$:

convwest1
$$f * g(x)w(x) \le \int_G |g(x-y)||f(y)|w(x)dy \le \int_G |g(x-y)|w(x-y)|f(y)|w(y)dy. \tag{262}$$

As an immediate consequence the corresponding weighted $\boldsymbol{L}^1$-space $\boldsymbol{L}^1_w$ is a Banach algebra with respect to convolution, a so-called *Beurling algebra* (cf. [81]).

Another important class are the so-called WSA weights:

WSA-wgt **Definition 50.** A strictly positive and continuous weight w is called WSA (*weakly subadditive*) if there is some constant $C > 0$ such that for all x, y:

WSA-def
$$w(x+y) \le C(w(x) + w(y)). \tag{263}$$

It implies in a completely similar manner another useful pointwise estimate:

conv-WSA
$$|(f * g) \cdot w| \le C(|f|w * |g| + |f| * |g|w) \tag{264}$$

recall that $\boldsymbol{L}^2_w := \{f \mid fw \in \boldsymbol{L}^2\}$, with the natural norm $\|f\|_{2,w} := \|fw\|_2$.

Using equation (16) and the fact that $\boldsymbol{L}^1 * \boldsymbol{L}^2 \subseteq \boldsymbol{L}^2$ (with corresponding norm inequalities)

$$\|g * f\|_2 = \|\hat{g} \cdot \hat{f}\|_2 \le \|\hat{g}\|_\infty \cdot \|\hat{f}\|_2.$$

Theorem 31. *Let w be a WSA weight. $\boldsymbol{L}^1 \cap \boldsymbol{L}^2_w$ is a Banach algebra with respect to convolution. In particular, $\boldsymbol{L}^2_w$ is a Banach algebra with respect to convolution if $1/w \in \boldsymbol{L}^2$.*

Proof. The $\boldsymbol{L}^1$-norm is no problem, so we only have to find for $f, g \in \boldsymbol{L}^1 \cap \boldsymbol{L}^2_w$ an estimate for
conv-WSA $\|(f * g)w\|_2$, which can be obtained from the estimate (). □

Argument: then $\boldsymbol{L}^2_w = \boldsymbol{L}^1 \cap \boldsymbol{L}^2_w$ with equivalence of the natural norms associated with these spaces!

Since the Fourier transform maps $\boldsymbol{L}^1$ into $\boldsymbol{C}_0$ (according to the Riemann Lebesgue Lemma) it follows that under the same conditions a Sobolev space is embedded into $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$. Let us write $\mathcal{H}_s(\mathbb{R}^d)$ for $\mathcal{F}^{-1}\left(\boldsymbol{L}^2_w(\mathbb{R}^d), \|\cdot\|_{2,w}\right)$, with norm $\|f\|_{\mathcal{H}_s} := \|\hat{f} \cdot w\|_2$.

For integer values of $s \geq 0$ one can show that this space coincides with the space of all $\boldsymbol{L}^2-$ elements which have (in the sense of distributions!) derivatives of order up to s, i.e. the (within the theory of distributions well defined objects f, f', f" etc. up to order s are regular distributions which can be represented by $\boldsymbol{L}^2$-functions, or equivalently: smoothing the distribution f one has uniform control of those derivatives in the classical sense independent of the order of regularization, e.g. by convolution with a very narrow Gauss function);

obolev-emb0 **Corollary 9.** *For $s > d/2$ the Sobolev space $\mathcal{H}_s$ is continuously embedded into $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$, hence also into $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$.*

Proof. Observing that $1/w_s \in \boldsymbol{L}^2(\mathbb{R}^d)$ if (and only if) $w > d/2$ on finds that $\boldsymbol{L}^2_w(\mathbb{R}^d) \hookrightarrow \boldsymbol{L}^1(\mathbb{R}^d)$ via Cauchy-Schwarz, writing $h \in \boldsymbol{L}^2_w(\mathbb{R}^d)$ as a product of hw with $1/w$:

LtwinLi (265)
$$\|h\|_1 = \|hw \cdot 1/w\|_1 \leq \|hw\|_2 \cdot \|1/w\|_2.$$

Applying the Fourier transform on both sides we obtain that $\mathcal{H}_s(\mathbb{R}^d) \hookrightarrow \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$, 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 Schwarz in the last inclusion), which by a variant of the Hausdorff-Young inequality for Wiener amalgam spaces ([34]) 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 spaces, because the norm obviously comes from the scalar product

HS-scalprod (266)
$$\langle f, g\rangle_{\mathcal{H}_s} := \langle \hat{f}w, \hat{g}w\rangle_{\boldsymbol{L}^2} = \int_{\mathbb{R}^d} \hat{f}(s)\,\overline{\hat{g}(s)}\; w^2(s)ds.$$

Hs-RKHS **Corollary 10.** *$\mathcal{H}_s(\mathbb{R}^d)$ 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*

Hs-RKHSphi (267)
$$f(x) = \langle f, k_x\rangle_{\mathcal{H}_s} \quad \textit{for all} \quad x \in \mathbb{R}^d, f \in \mathcal{H}_s$$

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 $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$ 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 $\hat{f} \in \boldsymbol{L}^1(\mathbb{R}^d)$).

s-delta0rep
$$f(0) = \int_{\mathbb{R}^d} \hat{f}(s)ds = \int_{\mathbb{R}^d} \hat{f}(s)w(s) \cdot \frac{1}{w^2}(s)w(s)ds =: \int_{\mathbb{R}^d} \hat{f}(s)\,\overline{\hat{\varphi}(s)}\; w^2(s)ds, \tag{268}$$

if we put $\varphi = \mathcal{F}^{-1}(1/w^2)$.

Now we can apply the shift invariance of the scalar product (exercise):

S-scalshift
$$\langle T_x f, T_x h\rangle_{\mathcal{H}_s} = \langle f, h\rangle_{\mathcal{H}_s} \quad \text{for all} \quad f, h \in \mathcal{H}_s(\mathbb{R}^d), \tag{269}$$

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. □

Another interesting embedding reads as follows:

ob-decay-S0 **Theorem 32.** *For $s > d$ the intersection of $\mathcal{H}_s(\mathbb{R}^d)$ with $\boldsymbol{L}^2_{w_s}(\mathbb{R}^d)$ with its natural norm (sum of the two norms) is continuously (and densely) embedded into $\big(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\big)$.*

Remark 62. Cf. the work of K. Gröchenig: there are also non-symmetric conditions and even conditions with weighted $\boldsymbol{L}^p-$norms on the time side and other weighted $\boldsymbol{L}^q$-conditions on the Fourier transform side which can be used to show that function with a certain amount of both time- and frequency concentration (as expressed by these conditions) are necessarily lying within $\boldsymbol{S}_0(\mathbb{R}^d)$. Of course one can easily adapt those conditions to conditions which are sufficient conditions for other modulation spaces, especially to be within Shubin classes $\boldsymbol{Q}(\mathbb{R}^d)$, or modulation spaces $\boldsymbol{M}^1_w(\mathbb{R}^d)$.

details are in [28], and about WSA functions in the work of Brandenburg [12], online at
`http://univie.ac.at/nuhag-php/bibtex/open_files/1407_9550001.pdf`

For other purposes we mention also the concept of *moderate weights*
COMES TWICE!

A weight function v on $\mathbb{R}^{2d}$ is called *submultiplicative*, if

submultdef1
$$v(z_1 + z_2) \le v(z_1)v(z_2), \tag{270}$$

for all $z_1 = (x_1, \xi_1), z_2 = (x_2, \xi_2) \in \mathbb{R}^{2d}$.

Remark 63. Weight functions are very natural if one does not want to consider points of a function to be of equal importance. For strictly positive weights $w(x), x \in G$ with *growing weights* (e.g. $\lim_{x\to\infty} w(x) = +\infty$) one can express decay at infinity by membership of f in some of the weighted spaces.

FURTHER COMMENTS ON WEIGHTED SPACES! IN PARTICULAR WEIGHTED L^p0 SPACES A weight function w on $\mathbb{R}^{2d}$ is *v-moderate* if

For applications in the context of Harmonic Analysis translation-invariant function spaces play an important role. This is where the notion of *moderateness* of weights comes in (see e.g. [28], an English summary is found at [27], or the more recent article [55]). In the literature on function spaces these concepts, which go back to papers of R.E. Edwards and G. Gaudry (see [25] and [51]) and of course the work of A. Beurling (see [79]) and Y. Domar ([24]).

efmoderate1

$$w(z_1 + z_2) \leq cv(z_1)w(z_2)(*) \tag{271}$$

for all $z_1 = (x_1, \xi_1), z_2 = (x_2, \xi_2) \in \mathbb{R}^{2d}$.

Two weights w_1 and w_2 are called *equivalent*, written $w_1 \asymp w_2$, if

defequwgt1

$$C^{-1}w_1(z) \leq w_2(z) \leq Cw_1(z) \tag{272}$$

for all $z = (x, \xi) \in \mathbb{R}^{2d}$ and some positive constant C.

17. Some pointwise estimates

Pointwise estimates: Convolution preserves monotonicity We have to define[91] $|\mu|$ for a given measure, and have to show that $\| \, |\mu| \, \|_{\boldsymbol{M}} = \|\mu\|_{\boldsymbol{M}}$ for each $\mu \in \boldsymbol{M}_b(\mathbb{R}^d)$.

point-conv
$$|\mu * f| \leq |\mu| * |f|, \quad \forall \mu \in \boldsymbol{M}_b(\mathbb{R}^d), f \in \boldsymbol{C}_b(\mathbb{R}^d). \tag{273}$$

point-conv1
$$|f| \leq |g| \Rightarrow |\mu| * |f| \leq |\mu| * |g| \tag{274}$$

osc-estim1a
$$[(\mathrm{D}_\Psi f - f) * g](x) \;\leq\; [|f| * \mathrm{osc}_\delta(g)](x), \, \forall x \in \mathbb{R}^d. \tag{275}$$

or in short, a pointwise estimate of the form

sc-estim1ba
$$[(\mathrm{D}_\Psi f - f) * g] \;\leq\; [|f| * \mathrm{osc}_\delta(g)]. \tag{276}$$

MAIN ESTIMATE

ain-estim11
$$|D_\Psi \mu * f - \mu * f| \leq |\mu| * |\, \mathrm{Sp}_\Psi f - f| \leq |\mu| * \mathrm{osc}_\delta f \tag{277}$$

if $\mathrm{diam}(\Psi) \leq \delta$.

It relies on a couple of "simple" estimates, such as

osccheck
$$(\mathrm{osc}_\delta f)^{\checkmark} = \mathrm{osc}_\delta(f^{\checkmark}) \tag{278}$$

and

osctransl
$$\mathrm{osc}_\delta(T_x f) = T_x(\mathrm{osc}_\delta f). \tag{279}$$

Obviously

splinapprx
$$|\, \mathrm{Sp}_\Psi f(x) - f(x)| \;\leq\; \mathrm{osc}_\delta f(x), \, \forall x \in \mathbb{R}^d. \tag{280}$$

Moreover the fact that the discretization operator $\mathrm{D}_\Psi : \boldsymbol{M}_b(\mathbb{R}^d) \mapsto \boldsymbol{M}_b(\mathbb{R}^d)$ is the adjoint of the spline operator $\mathrm{Sp}_\Psi \colon \boldsymbol{C}_0(\mathbb{R}^d) \mapsto \boldsymbol{C}_0(\mathbb{R}^d)$, implies also that we have:

conv-Dpsi
$$D_\Psi \mu * f = \mu * \mathrm{Sp}_\Psi f, \qquad \mu \in \boldsymbol{M}_b(\mathbb{R}^d), \; f \in \boldsymbol{C}_0(\mathbb{R}^d). \tag{281}$$

Lemma 72. *If $f \in \boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^p)$ then also $\mathrm{osc}_\delta f \in \boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^p)$.*

Proof. The estimate makes essentially use of the density of $\boldsymbol{C}_c(\mathbb{R}^d)$ in $\big(\boldsymbol{W}(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{W}}\big)$ and the fact, that it follows from the uniform continuity of functions with compact support. The simple estimate (*sub-additivity*):

splitoscd
$$\mathrm{osc}_\delta(\alpha f_1 + \beta f_2) \leq |\alpha| \, \mathrm{osc}_\delta(f_1) + |\beta| \, \mathrm{osc}_\delta(f_2), \tag{282}$$

is part of the argument. The details are left to the reader. □

[91] This is absolutely a non-trivial task and has to be done carefully. We cannot give a reference to the literature so far but are convinced that this is the way to go. One technical way, which does not require the use of measure theoretical arguments, is to define $|D_\Psi \mu| := \sum_{i \in I} |\mu(\psi_i)| \delta_{x_i}$ and verify than (this is the difficult part) that this tends to a limit in the w^*-sense!

oscWienerp **Lemma 73.** • *A function $f \in C_b(\mathbb{R}^d)$ belongs to $W(C_0, \ell^p)$ if and only if $f \in L^p(\mathbb{R}^d)$ and $\mathrm{osc}_\delta f \in L^p(\mathbb{R}^d)$.*

- *A function belongs to $W(C_0, \ell^p)$ if and only if $f^\#$, the local maximal function belongs to $W(C_0, \ell^p)$.*

Recall: $f^\#(x) = max_{|z-x|\leq 1}\{|f(z)|\}$.

18. Discretization and the Fourier transform

We shall define here $⊔⊔_a := \sum_k \delta_{ak}$ and $⊔⊔^a = \frac{1}{a}\sum_n \delta_{\frac{n}{a}}$. Then $\mathcal{F}⊔⊔_a = ⊔⊔^a$. In fact, one has for $a = 1$ according to Poisson's formula $\mathcal{F}⊔⊔_1 = ⊔⊔^1$, and the general formula follows from this by a standard dilation argument: Mass preserving compression St_ρ is converted into "value-preserving" dilation D_ρ on the Fourier transform side, and $\mathrm{D}_\rho ⊔⊔^1 = ⊔⊔^{1/\rho}$.

Let us put a few observations of importance at the beginning of this section:

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

frepper1 (283) $$\mathcal{F}\left[⊔⊔_{aN} * (⊔⊔^a \cdot f)\right] = ⊔⊔^{Nb} \cdot (⊔⊔_b * \hat{f}) = ⊔⊔_b * (⊔⊔^{Nb} \cdot \hat{f})$$

The last step in the proof of formula (283) is easily verified directly: sampling and periodization commute if (and only if) the periodization constant (bN in our case) is a multiple of the sampling period (in our case b).

The question of approximately obtaining the continuous Fourier transform $\hat{f}$ of a "nice function" f from the FFT of it's sampled version can be derived from this fact. Essentially it is based on the fact that good decay ensures that the effect of periodization disappears for large periods, while smoothness combined with a high sampling rate guarantees good local representation of the function based on (regular) samples.

Given $h > 0$ and some prescribed function ψ on $\mathbb{R}^d$, such as a cubic B-spline, the quasi-interpolation $Q_h f = Q_h^\psi f$ of a continuous function f on $\mathbb{R}^d$ is defined by

eqdefQh (284) $$Q_h f(x) = \sum_{k\in\mathbb{Z}} f(hk)\psi(x/h - k), \qquad x \in \mathbb{R}^d.$$

For suitable ψ, this formula describes an approximation to f from its samples on the fine grid $h\mathbb{Z}^d \subset \mathbb{R}^d$.

theoS0 **Theorem 33.** *Assume that $\psi \in S_0(\mathbb{R}^d)$ satisfies $\sum_{k\in\mathbb{Z}^d} \psi(x-k) \equiv 1$, i.e. that ψ generates a partition of unity. Then one has for any $f \in S_0(\mathbb{R}^d)$: $\|Q_h f - f\|_{S_0} \to 0$ as $h \to 0$.*

Note, that under the same restrictions on ψ one also has convergence of the quasi-interpolation scheme in the Fourier algebra $\mathcal{F}L^1$ for all $f \in \mathcal{F}L^1(\mathbb{R}^d)$.

Consequently one has $Q_h^*\sigma \to \sigma$ in the weak*-sense for each $\sigma \in S_0'(\mathbb{R}^d)$. But $Q_h^*\sigma = \sum_k \sigma(\psi_k)\delta_{hk}$. Hence the discrete measures are w^*-dense in $S_0'(\mathbb{R}^d)$.

Comment on the consistency of distributional Fourier transform with the classical one (defined on $\boldsymbol{L}^1(\mathbb{R}^d)$, using the Lebesgue integration formula).

For any Fourier invariant space of test functions (such as $\big(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\big)$ or the Schwartz space of *rapidly decreasing functions* $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$) one defines a *generalized Fourier transform* by the formula

Definition 51.

Four-def2

$$\mathcal{F}\sigma = \hat{\sigma} : \quad \hat{\sigma}(f) := \sigma(\hat{f}) \quad \forall f \in \boldsymbol{S}_0(\mathbb{R}^d). \tag{285}$$

The above formula can then be interpreted as a consistency relation. Write σ_g and $\sigma_{\hat{g}}$ for the distributions generated by the functions $g \in \boldsymbol{L}^1(\mathbb{R}^d)$ and $\hat{g} \in \boldsymbol{C}_0(\mathbb{R}^d)$ respectively [92] then the formula (177) tell us that

$$\widehat{\sigma_g} = \sigma_{\hat{g}}.$$

see e.g. [50], Lesson 17, page 156.

Material inserted 03.06.2017, TUM

$$\sqcup\!\sqcup\!\sqcup(f) = \sqcup\!\sqcup * (\sqcup\!\sqcup \cdot f) = \sqcup\!\sqcup \cdot (\sqcup\!\sqcup * f)$$

with the important property that

FTshahFT1

$$\mathcal{F}(\sqcup\!\sqcup\!\sqcup(f)) = \sqcup\!\sqcup\!\sqcup(\mathcal{F}f), \quad \forall f \in \boldsymbol{S}_0(\mathbb{R}^d). \tag{286}$$

or equivalently

FTshahFT

$$\mathcal{F} \circ \sqcup\!\sqcup\!\sqcup = \sqcup\!\sqcup\!\sqcup \circ \mathcal{F}. \tag{287}$$

In our code for `gaussnk.m` we use exactly this idea:

```
g=zeros(n,1); t=(0:n-1).';
p=sqrt(n); s = 1/sqrt(n);
for j=-3:3;  % rough periodization and
  g=g+exp(-pi*(t*s+j*p).^2);  % sampling at rate s
end;
g=g(:)'/norm(g);  %normalization norm(g) == 1
```

We also need the dilated version, which I would like to call $\sqcup\!\sqcup\!\sqcup_\rho$: The definition goes as follows (using the two dilation operators that I have in my lecture notes):

$$\sqcup\!\sqcup\!\sqcup_\rho(f) := \mathrm{St}_\rho \sqcup\!\sqcup * (\mathrm{D}_\rho \sqcup\!\sqcup \cdot f).$$

We will be interested in $\rho \to 0$, e.g. $\rho = 1/\sqrt{N}$ (if we want to obtain finite signals on $\mathbb{Z}_N$) or $\rho = 1/N^2$ if we want to have a sampling rate of $1/N$ and periodization by N, and thus obtain N^2 discrete values. So in principle if $(1/\rho)^2 \in \mathbb{N}$ we also have

ampperiod00

$$\mathrm{St}_\rho \sqcup\!\sqcup * (\mathrm{D}_\rho \sqcup\!\sqcup \cdot f) = \mathrm{D}_\rho \sqcup\!\sqcup \cdot (\mathrm{St}_\rho \sqcup\!\sqcup * f) \tag{288}$$

and consequently we have again

FTshahFTrho

$$\sqcup\!\sqcup\!\sqcup_\rho \circ \mathcal{F} = \mathcal{F} \circ \sqcup\!\sqcup\!\sqcup_\rho \tag{289}$$

In other words, we can apply the ordinary Fourier transform first on a function, OR alternatively apply the DFT/FFT on the finite sequence describing its periodized and sampled version. We consider it a **remarkable fact** that these operators intertwine. In other words, we have to

[92] so the space of test functions must be contained in $\boldsymbol{L}^1(\mathbb{R}^d)$ to ensure that $\boldsymbol{C}_0(\mathbb{R}^d)$-functions define continuous linear functionals.

prove that finite sequence describing the periodic and discrete signal can be mapped onto the corresponding sequence *of equal length* (!!) describing the Fourier transform of such a sequence by just applying [perhaps up to some normalization!] the Discrete Fourier Transform (DFT), and practically speaking the FFT (Fast Fourier Transform). Furthermore, we will show that any periodic and discrete signal is of this form.

Spline-type quasiinterpolation is then what you call regaining the continuous function from a discrete sequence, which gives a continuous and periodic function which then by windowing has to be localized, but all the can be described well with these symbols.

nocite: feka07, fe92 [43], [36]

WHAT IS MISSING NOW IS THE FACT THAT ONE CAN COMPUTE THE FOURIER TRANSFORM OF THE DISCRETE AND PERIODIC SEQUENCE IN $\boldsymbol{S}_0(\mathbb{R}^d)$ OBTAINED IN AN EXACT WAY BY JUST APPLYING THE FFT (THE FAST REALIZATION OF THE DFT, THE (*Discrete Fourier Transform).*

19. Quasi-Interpolation

The piecewise linear interpolation operator for data available on the lattice of integers $\mathbb{Z}$, say $(c_k)_{k\in\mathbb{Z}}$, can be described as a sum of shifted triangular functions $\Delta(0) = 1, \Delta(k) = 0$ for $k \notin \mathbb{Z}$. Hence it can be written as a convolution product of the form

$$\big(\sum_{k\in\mathbb{Z}} c_k\delta_k\big) * \Delta.$$

It is easy to show that the resulting sum (the interpolant) belongs to $\boldsymbol{L}^p(\mathbb{R})$ if the sequence $\mathbf{c}$ is from $\boldsymbol{\ell}^p(\mathbb{Z})$. But this is true for much more general functions than then triangular function. It suffices to have $\varphi \in \boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R})$ in order to find out that $\sum_{k\in\mathbb{Z}} T_k\varphi$ belongs to $\boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^p)(\mathbb{R})$ for $\mathbf{c} \in \boldsymbol{\ell}^p(\mathbb{Z})$. In fact, this assumption implies $\sum_k c_k\delta_k \in \boldsymbol{W}(\boldsymbol{M}, \boldsymbol{\ell}^p)$ and hence the convolution relations for Wiener amalgam spaces imply:

$$f = \sum_{k\in\mathbb{Z}} T_k\varphi = \big(\sum_{k\in\mathbb{Z}} c_k\delta_k\big) * \varphi \in \boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^p)(\mathbb{R}).$$

As a consequence f is a continuous function and can be sampled, e.g. over the integers, but in most cases $f(k)$ will be perhaps close to, but different from the original sequence $(c_k)_{k\in\mathbb{Z}}$, hence the name *quasi-interpolation.* [93]

The so-called quasi-interpolation operators make sense for functions from $\boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^p)(\mathbb{R}^d)$, to choose the appropriate generality from now on. For those functions one can guarantee that for some $C > 0$ and all $p \in [1, \infty]$ one has:

$$\|(f(k))_{k\in\mathbb{Z}^d}\|_p \le C\|f \,|\, \boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^p)\| \quad \forall f \in \boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^p)(\mathbb{R}^d).$$

The same is true for any other lattice $\Lambda \lhd \mathbb{R}^d$, with

$$\|(f(\lambda))_{\lambda\in\Lambda}\|_p \le C_\Lambda\|f \,|\, \boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^p)\| \quad \forall f \in \boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^p)(\mathbb{R}^d).$$

Hence the operator

$$f \mapsto \sum_{\lambda\in\Lambda} f(\lambda)T_\lambda\varphi$$

[93] Note that SINC is not covered by this example, although for $p \in (1, \infty)$ it shares more or less all the properties described above.

is a well defined operator on $\boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^p)(\mathbb{R}^d)$ (even uniformly bounded with respect to the range $p \in [1, \infty]$. We will call such an operator the *quasi-interpolation operator* with respect to the pair (Λ, φ).

Among the quasi-interpolation operators those which arise from BUPUs, i.e.from functions $\varphi \in \boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d)$ satisfying

$$\sum_{\lambda \in \Lambda} T_\lambda \varphi(x) \equiv 1$$

are the most important ones. We are going to show that quasi-interpolation operators with respect to "fine lattices" Λ are good approximation operators.

The interesting phenomenon is the behaviour of piecewise linear interpolation over lattices of the form $\alpha\mathbb{Z}^d$, for $\alpha \to 0$.

Let us recall that $(T_k\varphi)_{k\in\Lambda}$ is a BUPU for some $\varphi \in \boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d)$ if and only if $\hat{\varphi}$ is a Lagrange interpolator over the orthogonal lattice $\Lambda^\perp = \{\chi \,|\, \langle \chi, \lambda \rangle \equiv 1 \quad \forall \lambda \in \Lambda\}$, i.e.that

Lagr-intp1 (290)
$$\widehat{\varphi}(\lambda') = \delta_{0,\lambda'} \quad \forall \lambda' \in \Lambda^\perp.$$

Proof. We can reinterpret the BUPU condition as $\sqcup\!\sqcup_H * \varphi \equiv 1$, which turns into

$$\mathcal{F}(\sqcup\!\sqcup_H) \cdot \hat{\varphi} = \mathcal{F}(1) = \delta_0.$$

Since $\mathcal{F}(\sqcup\!\sqcup_H) = C_H \sqcup\!\sqcup_{H^\perp}$ this condition reduces to (using $f \cdot \delta_x = f(x)\delta_x$):

$$C_H \sqcup\!\sqcup_{H^\perp} \cdot \hat{\varphi} = \sum_{h' \in H^\perp} \hat{\varphi}(h')\delta_{h'} = \delta_0,$$

which in turn is true if and only if $\hat{\varphi}(h') = 0$ for $h' \neq 0$ for all $h' \in H^\perp$.

□

Remark 64. The condition described above is invariant with respect to pointwise powers on the Fourier transform side, i.e.$\hat{\varphi}$ satisfies (290) then the same is true for $\hat{\varphi}^2 = \widehat{\varphi * \varphi}$.

The quasi-interpolation operator $Q_{\Lambda,\varphi}$ can thus be described as the mapping

quasint (291)
$$f \mapsto (\sqcup\!\sqcup_H \cdot f) * \varphi.$$

Note that this operators is bounded on $\boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^p)(\mathbb{R}^d)$ because $\sqcup\!\sqcup_H \cdot f \in \boldsymbol{W}(\boldsymbol{M}, \boldsymbol{\ell}^\infty)(\mathbb{R}^d) \cdot \boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^p)(\mathbb{R}^d) \subseteq \boldsymbol{W}(\boldsymbol{M}, \boldsymbol{\ell}^p)(\mathbb{R}^d)$, hence

$$(f \cdot \sqcup\!\sqcup_H) * \varphi \subseteq \boldsymbol{W}(\boldsymbol{M}, \boldsymbol{\ell}^p) * \boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1) \subseteq \boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^p)(\mathbb{R}^d).$$

It is of interest to check the behaviour of quasi-interpolation for the lattices $h\mathbb{Z}^d$, with $h \to 0$:
Given $h > 0$ and some prescribed function ψ on $\mathbb{R}^d$, such as a B-spline, the quasi-interpolation $Q_h f = Q_h^\psi f$ of a continuous function f on $\mathbb{R}^d$ is defined by

eqdefQh (292)
$$Q_h f(x) = \sum_{k\in\mathbb{Z}} f(hk)\psi(x/h - k), \qquad x \in \mathbb{R}^d.$$

For suitable ψ, this formula describes an approximation to f from its samples on the fine grid $h\mathbb{Z}^d \subset \mathbb{R}^d$.

theoS0 **Theorem 34.** *Assume that $\psi \in \boldsymbol{S}_0(\mathbb{R}^d)$ satisfies $\sum_{k\in\mathbb{Z}^d} \psi(x-k) \equiv 1$, i.e.that the family $(T_k\psi)_{k\in\mathbb{Z}^d}$ forms a partition of unity. Then for all $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ we have $\|Q_h f - f\|_{\boldsymbol{S}_0} \to 0$ as $h \to 0$.*

Note, that under the same restrictions on ψ one also has convergence of the quasi-interpolation scheme in the Fourier algebra $\mathcal{F}\boldsymbol{L}^1$ for all $f \in \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$.

20. Pseudo-measures and other auxiliary terms

The following definition requires to have already a "generalized Fourier transform", which allows to define the Fourier transform for $\boldsymbol{L}^\infty(\mathbb{R}^d)$.

pseudomeas1 **Definition 52.** The space $\mathcal{F}\boldsymbol{L}^\infty$ is defined as the (inverse) image under the (generalized) Fourier transform of $\boldsymbol{L}^\infty(\mathbb{R}^d) = [(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1)]'$.

An alternative definition of pseudo-measures (used e.g. in the work of G. Gaudry or R. Larsen) is based on the following characterization:

ourLinfdual **Lemma 74.** *There is a natural identification of* $\mathcal{F}\boldsymbol{L}^\infty(\mathbb{R}^d)$ *with the* $[(\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_{\mathcal{F}\boldsymbol{L}^1})]'$*, via* $\sigma(\hat{h}) = \sigma_h(\widehat{\sigma}) = \int_{\mathbb{R}^d} \widehat{\sigma}(s)h(s)ds$.

Even in the context of non-Abelian Groups G this makes sense, if one uses Eymard's Fourier algebra $\boldsymbol{A}(G)$ (ey64 [26]).

21. Advantages of a distributional Fourier Transform

Whereas most books in the field of Fourier analysis describe the Fourier transform at various levels, typically starting from the classical case of the Fourier transform for periodic functions [23, 63]. Sometimes the need of a generalized Fourier transform is motivated by the fact, that certain objects (like the "pure frequencies") do not have a Fourier transform in the usual sense, because first of all the classical Fourier transform is bound to diverge, while on the other hand the Fourier transform (which is in the generalized calculus a Dirac measure) is not an ordinary function, but has to be a kind of generalized function (in fact a bounded measure in that case), cf. [6].

In this section we want to emphasize (by demonstrating the situation, valid even for locally compact groups in full generality, through the example $\mathcal{G} = \mathbb{R}^d$).

The distributional Fourier transform (and it suffices to know the $\boldsymbol{S}_0'$-theory for this purpose) allows to define the Fourier transform even for objects which are not decaying at infinity (like functions in any of the $\boldsymbol{L}^p$-spaces), but also for periodic objects (such as periodic functions belonging locally to $\boldsymbol{L}^p$), even with different periods (which brings us already close to the discussion of almost periodic functions).

As a central topic let us therefore discuss the Fourier transform of periodic functions (or measures, or distributions) as the "infinite limit" of it's periodic repetitions. First of let us recall that it is easy to find for each lattice $\Lambda = \boldsymbol{A} * \mathbb{Z}^d$, for some non-singular $d \times d$-matrix $\boldsymbol{A}$ a fundamental domain (equal to $Q = \boldsymbol{A} * [0,1)^d$) and also bounded partitions of unity of the form $(\phi_\lambda)_{\lambda\in\Lambda} = (T_\lambda\varphi)_{\lambda\in\Lambda}$, with $\varphi \in \boldsymbol{C}_c(\mathbb{R}^d)$ or even in $\mathcal{S}(\mathbb{R}^d)$. For the next lemma we need $\varphi \in \boldsymbol{S}_0(\mathbb{R}^d)$, or better, in $\mathcal{S}(\mathbb{R}^d)$.

periodiz02 **Lemma 75.** *A function* f *(or distribution in* $\boldsymbol{S}_0'(\mathbb{R}^d)$*)* [94] *is periodic with respect to* $\Lambda \lhd \mathbb{R}^d$ *if and only if it is of the form*

riodizefct2 (293)
$$f = \sum_{\lambda\in\Lambda} T_\lambda f^\circ_\lambda,$$

for some compactly supported pseudo-measure $f^\circ \in \mathcal{F}\boldsymbol{L}^\infty(\mathbb{R}^d)$.

Proof. If f is Λ-periodic, i.e.if $T_\lambda f = f$ for all $\lambda\in\Lambda$, then we can choose $f^\circ = f\varphi$, for a function φ with compact support, generating a Λ-BUPU as described above, because

$$f = \sum_{\lambda\in\Lambda} T_\lambda(T_{-\lambda}(f\varphi_\lambda)) = \sum_{\lambda\in\Lambda} T_\lambda f^\lambda = \sum_{\lambda\in\Lambda} T_\lambda f^\circ.$$

because $f^\lambda = T_{-\lambda}(f\cdot\varphi_\lambda) = T_{-\lambda}f \cdot T_{-\lambda}\varphi_\lambda = f\cdot\varphi =: f^\circ$ by the periodicity of f.

Conversely, let f° a compactly supported pseudo-measure or even in $\boldsymbol{W}(\mathcal{F}\boldsymbol{L}^\infty, \boldsymbol{\ell}^1)$. Since $\boldsymbol{W}(\boldsymbol{M}, \boldsymbol{\ell}^\infty) * \boldsymbol{W}(\mathcal{F}\boldsymbol{L}^\infty, \boldsymbol{\ell}^1) \subset \boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1, \boldsymbol{\ell}^\infty) = \boldsymbol{S}_0'$ the partial sums of the periodization are both

[94] A distribution from $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ which is $\Lambda-$periodic for any co-compact lattice Λ in $\mathbb{R}^d$ does in fact belong automatically to $\boldsymbol{S}_0'(\mathbb{R}^d)$!

uniformly in S_0' as well as w^*-convergent, as can been seen from the interpretation

$$\lim_{F\to\Lambda}\sum_{\lambda\in F} T_\lambda f^\circ = \lim_{F\to\Lambda}\left((\sum_{\lambda\in F}\delta_\lambda) * f^\circ \right).$$

In fact, one can reduce the general case of w^*-convergence to the case where (first of all) f° as compact support, but in testing that the action on arbitrary test functions $h \in S_0(\mathbb{R}^d)$ those in $(S_0)_c = A_c(\mathbb{R}^d)$ are sufficient.

$\square$

We remark that a compactly supported pseudo-measure has the property that its Fourier transform is indeed a bounded and continuous function. Indeed, we can find some $\varphi \in S_0(\mathbb{R}^d)$ such that $f^\circ \cdot \varphi$, hence $\mathcal{F}(f^\circ \cdot \varphi) = \widehat{f^\circ} * \hat{\varphi} \in L^\infty * S_0 \subseteq C_b(\mathbb{R}^d)$. One can also show that the Fourier coefficients of the periodic version are just the values $\widehat{f^\circ}$ over the orthogonal lattice $\Lambda^\perp$. Let us describe this in detail:

Using now the fact that the distributional FT of $\sqcup\!\sqcup_\Lambda$ coincides with a multiple of $\sqcup\!\sqcup_{\Lambda^\perp}$, i.e.$\mathcal{F}\sqcup\!\sqcup_\Lambda = C_\Lambda \sqcup\!\sqcup_{\Lambda^\perp}$ in the S_0'-sense, we find that for any periodic function or distribution f we have

FTFourS1

$$\mathcal{F}f = \mathcal{F}(\sqcup\!\sqcup_\Lambda * f^\circ) = \mathcal{F}(\sqcup\!\sqcup_\Lambda) \cdot \hat{f}^\circ = C_\Lambda \sqcup\!\sqcup_{\Lambda^\perp} \cdot \hat{f}^\circ. \tag{294}$$

Consequently we have $\mathrm{supp}(\hat{f}) \subseteq \mathrm{supp}(\sqcup\!\sqcup_{\Lambda^\perp}) = \Lambda^\perp$. In fact, we can give a more explicit description of $\hat{f}$: by carrying out the pointwise multiplication $\sqcup\!\sqcup_{\Lambda^\perp} \cdot \hat{f}^\circ$ we find that

$$\hat{f} = \sum_{\lambda^\perp\in\Lambda^\perp} \hat{f}^\circ(\lambda^\perp)\,\delta_{\lambda^\perp}.$$

But for $f^\circ \in W(\mathcal{F}L^\infty, \ell^1)$ (by the Hausdorff-Young principle for generalized amalgams) we know that $\hat{f}^\circ \in W(\mathcal{F}L^1, \ell^\infty) \subset W(C_0, \ell^\infty) \subset C_b(\mathbb{R}^d)$. Hence we can even claim that $\hat{f}$ is a sum of Dirac measures located at the points of $\Lambda^\perp$, with a bounded sequence of coefficients in $\ell^\infty(\Lambda)$.

In fact (this has to be shown separately), one can be shown that this is a complete *characterization* all the tempered distributions $\sigma \in S_0'(\mathbb{R}^d)$ with $\mathrm{supp}(\sigma) \subseteq \Lambda^\perp$.

ardiscsupp1 **Lemma 76.** *(Characterization of distributions supported on discrete subgroups)* *A distribution* $\sigma \in S_0'(\mathbb{R}^d)$ *satisfies* $\mathrm{supp}(\sigma) \subseteq \Lambda$ *if and only if it is of the form*

$$\sigma = \sum_{\lambda\in\Lambda} c_\lambda \delta_\lambda$$

for some sequence $\mathbf{c} = (c_\lambda)_{\lambda\in\Lambda} \in \ell^\infty(\Lambda)$.

Summarizing we have a mapping from the periodic elements (in $S_0'(\mathbb{R}^d)$) into $\ell^\infty(\Lambda)$, of the form $f \to (\hat{f}^\circ(\lambda))_{\lambda\in\Lambda}$ (which is of course independent of the choice of f°). Assume we have a regular distribution, "coming from" some $f \in L^1(\mathbb{T})$. Let us discuss (for simplicity) the situation for $d = 1$ and $\Lambda = \mathbb{Z}$. Then we have $Q = [0, 1)$ and we can choose $f^\circ = f \cdot \mathbf{1}_Q \in L^1(\mathbb{R})$. Since $\mathbb{Z}^\perp = \mathbb{Z}$ the Fourier transform of a periodic (local) L^1-function are a sequence on $\mathbb{Z}$ again: $\hat{f}^\circ(n) = \int_0^1 f(t)e^{-2\pi i t}dt$ which coincides with the usual ("classical") definition of Fourier coefficients for locally integrable, $\mathbb{Z}$-periodic functions.

Note that in our context local square (or generally p-) integrability corresponds to additional properties on f° which in such a case will belong locally to $L^2(\mathbb{R})$ (resp. $L^p(\mathbb{R})$), resp. to $L^2(\mathbb{T})$ or $L^p(\mathbb{T})$. In our terminology $f^\circ \in W(L^2, \ell^1)(\mathbb{R})$ resp. $W(L^p, \ell^1)(\mathbb{R})$, and again by the

Hausdorff-Young principle we obtain that $\hat{f}^\circ \in \boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1, \boldsymbol{\ell}^\infty)(\mathbb{R})$, or the Fourier coefficients of f resp. the values of $(\hat{f}^\circ(n))_{n\in\mathbb{Z}}$ belong to $\boldsymbol{\ell}^{p'}(\mathbb{Z})$.

An important subspace of periodic functions is of course $\big(\boldsymbol{A}(\mathbb{T}), \|\cdot\|_{\boldsymbol{A}}\big)$, the (again introduced by N. Wiener) Banach space of all continuous functions with absolutely convergent Fourier series, i.e. with

ATnorm1 (295)
$$\|f\|_{\boldsymbol{A}^2} := \|\widehat{f}\|_1 := \sum_{k\in\mathbb{Z}} |\widehat{f}(k)| < \infty.$$

Clearly this is the subspace of $\big(\boldsymbol{L}^2(\mathbb{T}), \|\cdot\|_2\big)$, which is mapped under the Fourier transform onto $\boldsymbol{\ell}^1(\mathbb{Z})$, in fact isometrically.

ATandSO1 **Lemma 77.** *Given any regular partition of unity as used in Lemma 75 the space of periodic functions obtained by restricting the periodization mapping from* $\boldsymbol{PM}(\mathbb{R})$ *to* $\big(\boldsymbol{S}_0(\mathbb{R}), \|\cdot\|_{\boldsymbol{S}_0}\big)$ *is exactly* $\big(\boldsymbol{A}(\mathbb{T}), \|\cdot\|_{\boldsymbol{A}}\big)$.

Proof. We have seen that the periodization mapping $f \mapsto ⊔⊔ * f$ has the property that the Fourier coefficients of f_{per} are just the samples of the Fourier transform of f. Since f has compact support its Fourier transform is in fact in $\boldsymbol{C}_b(\mathbb{R})$, so pointwise evaluation is meaningful.

As we know $f \mapsto \widehat{f}$ maps $\big(\boldsymbol{S}_0(\mathbb{R}), \|\cdot\|_{\boldsymbol{S}_0}\big)$ into itself, and furthermore the restriction of functions (like $\widehat{f}$) in $\boldsymbol{S}_0(\mathbb{R})$ to any lattice are in $\big(\boldsymbol{\ell}^1, \|\cdot\|_1\big)$, with $\|(\widehat{f}(k))_{k\in\mathbb{Z}} \leq C\|f\|_{\boldsymbol{S}_0}$, for all $f \in \boldsymbol{S}_0(\mathbb{R})$.

Conversely we have to show that for any periodic function $g \in \boldsymbol{A}(\mathbb{T})$ there exists some $f \in \boldsymbol{S}_0(\mathbb{R})$ with $\widehat{f}(k) = \widehat{g}(n)$ (the given Fourier coefficients). This is easy to obtain. Simply use the triangular function and form (on the Fourier transform side) the function

$$h = \sum_{k\in\mathbb{Z}} \widehat{g}(k) T_k \Delta.$$

Since $\mu = \sum_{k\in\mathbb{Z}} \widehat{g}(k)\delta_k$ is a bounded and discrete measure it is clear that

$$\|\mu * \Delta\|_{\boldsymbol{S}_0} \leq C' \cdot \|\mu\|_{\boldsymbol{M}} \|\Delta\|_{\boldsymbol{S}_0} = C' \cdot \left(\sum_{k\in\mathbb{Z}} |\widehat{g}(k)|\right) \|\Delta\|_{\boldsymbol{S}_0}.$$

In other words, all we have to do is to *apply the piecewise linear interpolation operator* to the given data sequence $(\widehat{g}(k)) \in \boldsymbol{\ell}^1(\mathbb{Z})$. □

Remark 65. Obviously this principle of proof applies to general lattices in $\mathbb{R}^d$, where periodicity with respect to the lattice Λ corresponds to support of the Fourier transforms on the orthogonal lattice. For the simple case of $\mathbb{Z}^d \lhd \mathbb{R}^d$ the tensor products of BUPUs in each direction makes it easy to produce an analogous situation. For general lattice of the form $\Lambda = \boldsymbol{A} * \mathbb{Z}^d$ one obtains suitable interpolations by deformation of the simple ones.

The periodization mapping has as a possible natural domain the Wiener amalgam space $\boldsymbol{W}(\mathcal{F}\boldsymbol{L}^\infty, \boldsymbol{\ell}^1)(\mathbb{R}^d)$, which is the Fourier image of $\boldsymbol{M}^{\infty,1}(\mathbb{R}^d)$, also known as Sjöstrand class in the literature on pseudo-differential operators ([54, 56]).

A topic of interest in connection with standard spaces is the following one, which is based on the fact that for any (restricted) standard space $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ we have the following chain of continuous inclusion:

WB-incl0 (296)
$$\boldsymbol{W}(\boldsymbol{B}, \boldsymbol{\ell}^1) \hookrightarrow (\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}}) \hookrightarrow \boldsymbol{W}(\boldsymbol{B}, \boldsymbol{\ell}^\infty)$$

and the fact that we have

WB-incl2 (297)
$$\boldsymbol{W}(\boldsymbol{B},\boldsymbol{\ell}^p) \hookrightarrow \boldsymbol{W}(\boldsymbol{B},\boldsymbol{\ell}^q) \quad \text{if} \quad p \leq q.$$

ow-uppindex **Definition 53.** Given a (restricted) standard space $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ we define the *lower* resp. *upper index* as follows:

low-index (298)
$$\text{lowind}(\boldsymbol{B}) := \sup\{\, p \,|\, \boldsymbol{W}(\boldsymbol{B},\boldsymbol{\ell}^p) \hookrightarrow \boldsymbol{B}\}$$

and

upp-index (299)
$$\text{uppind}(\boldsymbol{B}) := \inf\{\, q \,|\, \boldsymbol{B} \hookrightarrow \boldsymbol{W}(\boldsymbol{B},\boldsymbol{\ell}^q)\}$$

In most cases the supremum resp. infimum will *not be attained.* However, we will have in any case

lowlequpp (300)
$$\text{lowind}_{\boldsymbol{B}} \leq \text{uppind}_{\boldsymbol{B}}$$

The following condition is in general slightly stronger than the case of equality of indices:

Definition 54. A Banach space is called to be of *global type p* if one has

deftypep (301)
$$\boldsymbol{B} = \boldsymbol{W}(\boldsymbol{B},\boldsymbol{\ell}^p).$$

Aside from the trivial facts that $\boldsymbol{L}^p(\mathbb{R}^d)$ is of course of type p for any p, one can check that $\boldsymbol{M}_b(\mathbb{R}^d)$ is of type 1, while the usual $\boldsymbol{L}^2$-Sobolev spaces are of type 2. They have therefore been called $\boldsymbol{\ell}^2$-puzzles by P. Tchamitchian in [95, 96] (?)

One of the interesting and non-trivial facts (no proof is given here) is that one has (the most interesting perhaps being the case $p = 1$), see [52]:

Lemma 78. *For* $1 \leq p \leq 2$ $\text{lowind}(\mathcal{F}\boldsymbol{L}^p) = p = \text{lowind}(\mathcal{F}\boldsymbol{L}^{p'})$ *while* $\text{uppind}(\mathcal{F}\boldsymbol{L}^p) = p' = \text{uppind}(\mathcal{F}\boldsymbol{L}^{p'})$. *Hence, except for the case $p = 2$ the space $\mathcal{F}\boldsymbol{L}^p$ is never of a particular type.*

Another interesting family of spaces is the set of all "multipliers" on $\boldsymbol{L}^p$, for say $1 \leq p < \infty$, which we denote by $H_{\boldsymbol{L}^1}(\boldsymbol{L}^p,\boldsymbol{L}^p)$, which is indeed another (restricted) standard space. It is well known that it coincides (by duality) with $H_{\boldsymbol{L}^1}(\boldsymbol{L}^{p'},\boldsymbol{L}^{p'})$, that $\mathcal{H}_{\boldsymbol{L}^1}(\boldsymbol{L}^1) = \boldsymbol{M}_b(\mathbb{R}^d)$ and $\mathcal{H}_{\boldsymbol{L}^1}(\boldsymbol{L}^2,\boldsymbol{L}^2) = \mathcal{F}\boldsymbol{L}^\infty(\mathbb{R}^d)$. Hence one may *conjecture* that lowind and uppind of these spaces equals p and p', but to my knowledge nothing is known about it for $1 < p < 2$.

Another interesting question could be the upper and lower index for modulation spaces (which can be shown to be of local type $\mathcal{F}\boldsymbol{L}^q$). Since the space $\boldsymbol{M}^p(\mathbb{R}^d) = \boldsymbol{M}^s_{p,q}(\mathbb{R}^d)$ for $p = q$ are Fourier invariant an coincide with $\boldsymbol{W}(\mathcal{F}\boldsymbol{L}^p,\boldsymbol{\ell}^p)$ they are clearly of *type* p.

Lower and Upper Index of a Function Space

TEST:

$$\text{upp}_{\boldsymbol{M}}(\boldsymbol{B}), \text{low}_{\boldsymbol{M}}(\boldsymbol{B}), \text{upp}_{\boldsymbol{W}}(\boldsymbol{B}), \text{low}_{\boldsymbol{W}}(\boldsymbol{B})$$

$$(TEST)(\boldsymbol{B})$$

There is another index which can be defined on the basis of the observation that there is a *scale of Banach spaces*, connecting $\boldsymbol{S}_0(\mathbb{R}^d) = \boldsymbol{M}^{1,1}_0(\mathbb{R}^d)$ with $\boldsymbol{S}_0'(\mathbb{R}^d) = \boldsymbol{M}^\infty(\mathbb{R}^d) = \boldsymbol{M}^{\infty,\infty}(\mathbb{R}^d)$ via complex interpolation. In other words we have

Mppinterpol (302)
$$\boldsymbol{M}^p(\mathbb{R}^d) = [\boldsymbol{M}^1(\mathbb{R}^d), \boldsymbol{M}^\infty(\mathbb{R}^d)]_{[1/p]}, \quad 1 \leq p \leq \infty.$$

Not surprisingly we have thus $\boldsymbol{L}^2(\mathbb{R}^d) = \boldsymbol{M}^2(\mathbb{R}^d)$ arises as the "midpoint" in the interpolation scheme between the extreme points $\boldsymbol{M}^1(\mathbb{R}^d)$ and $\boldsymbol{M}^\infty(\mathbb{R}^d)$.

Furthermore, by the reiteration results valid for *complex interpolation* of Banach spaces (see [8]) one has

pinterpolc1 **Lemma 79.** *For $1 \leq p_1, p_2 \leq \infty$ and $\theta \in [0,1]$:*

Mpinterpolc
$$M^p(\mathbb{R}^d) = [M^{p_1}(\mathbb{R}^d), M^{p_2}(\mathbb{R}^d)]_{[\theta]}, \quad \text{for } 1/p = (1-\theta)/p_1 + \theta/p_2. \tag{303}$$

These space are naturally ordered by (strict) inclusion.

Mpinclus1 **Lemma 80.** *For $1 < p_1, p_2 < \infty$ ones has*

Mpinclus
$$M^1(\mathbb{R}^d) \subset M^{p_1}(\mathbb{R}^d) \subset M^{p_2}(\mathbb{R}^d) \subset M^\infty(\mathbb{R}^d). \tag{304}$$

The inclusions are proper and dense (except for the last one).

Gabor atomic characterizations.

$$f \in M^p(\mathbb{R}^d) \Leftrightarrow f = \sum_{\lambda \in \Lambda} c_\lambda g_\lambda \, \text{with} \, \mathbf{c} \in \ell^p(\Lambda).$$

Since the scale $(M^p(\mathbb{R}^d))_{1 \leq p \leq \infty}$ is showing the same inclusion relations as the scale of Wiener amalgam spaces $W(B, \ell^q)$ with $1 \leq q \leq \infty$ (for fixed $(B, \|\cdot\|_B)$) we can define in analogy to the **AMALGAM**-*index* given in Def. 53 the *lower and upper* M^p**-index.**

wupModIndex **Definition 55.** Given a Fourier standard space $(B, \|\cdot\|_B)$ on $\mathbb{R}^d$ we define the *lower* resp. *upper* M^p*-index* as follows:

low-index
$$\mathrm{low}_M(B) := \sup\{\, p \mid M^p(\mathbb{R}^d) \hookrightarrow B\} \tag{305}$$

and

upp-index
$$\mathrm{upp}_M(B) := \inf\{\, q \mid B \hookrightarrow M^q(\mathbb{R}^d)\} \tag{306}$$

Since by assumption $M^1(\mathbb{R}^d) \hookrightarrow B \hookrightarrow M^\infty(\mathbb{R}^d)$ the corresponding parameter sets are non-empty.

Exercise: For the ordinary spaces $\big(L^p(\mathbb{R}^d), \|\cdot\|_p\big)$, with $1 \leq p \leq 2$ one has (more or less the amalgam indices for $\big(\mathcal{F}L^p(\mathbb{R}^d), \|\cdot\|_F Lpsp\big)$):

LpModInds
$$\mathrm{low}_M(L^p(\mathbb{R}^d)) = p \quad \text{and} \quad \mathrm{upp}_M(L^p(\mathbb{R}^d)) = p', \quad \text{with} \quad 1/p + 1/p' = 1, \tag{307}$$

in contrast to the obvious fact that $L^p(\mathbb{R}^d) = W(L^p, \ell^p)$ resulting in the identity

LpWInds
$$\mathrm{low}_W(L^p(\mathbb{R}^d)) = p = \mathrm{upp}_W(L^p(\mathbb{R}^d)), \quad 1 \leq p \leq \infty. \tag{308}$$

The following lemma is an easy consequence of duality considerations.

ndexduality **Lemma 81.** *Assume that $(B, \|\cdot\|_B)$ is a minimal FouSS, i.e. that $S_0(\mathbb{R}^d)$ is dense in $(B, \|\cdot\|_B)$.*

The $(B', \|\cdot\|_{B'})$ is also a Fourier standard space and the upper and lower indices of $(B, \|\cdot\|_B)$ and $(B', \|\cdot\|_{B'})$ are related by the usual conjugation property

$$1/\alpha + 1/beta = 1.$$

- $\alpha = \mathrm{upp}_M(B) \quad \Leftrightarrow \quad \beta = \mathrm{low}_M(B')$;
- $\alpha = \mathrm{low}_M(B) \quad \Leftrightarrow \quad \beta = \mathrm{upp}_M(B')$;
- $\alpha = \mathrm{upp}_W(B) \quad \Leftrightarrow \quad \beta = \mathrm{low}_W(B')$;
- $\alpha = \mathrm{low}_W(B) \quad \Leftrightarrow \quad \beta = \mathrm{upp}_W(B')$.

For Fourier Standard Spaces one can also define the analogue of the Koethe α-dual. Let us recall that for a *BF-space*, i.e. a Banach space which is continuously embedded into $\boldsymbol{L}^1_{loc}(\mathbb{R}^d)$, meaning that for any compact set $K \subset \mathbb{R}^d$ there exists $C_K > 0$ such that

BFestim1
$$\int_K |f(x)|dx = \|f \cdot \mathbf{1}_K\|_1 \leq C_K \|f\|_{\boldsymbol{B}}, \quad \forall f \in \boldsymbol{B}, \tag{309}$$

the Koethe-dual is defined as follows:

Definition 56.

lphdualdef2
$$\boldsymbol{B}^\alpha := \{h \in \boldsymbol{L}^1_{loc}(\mathbb{R}^d) \,|\, h \cdot f \in \boldsymbol{L}^1(\mathbb{R}^d) \;\forall f \in \boldsymbol{B}\} \tag{310}$$

The theory of *Koethe α-dual* spaces was developed systematically in the context of *Banach function spaces* in the sense of Luxemburg-Zaanen (see BOOK of Zaanen za67 [107], [66], etc.). In our setting we describe them as *solid BF-spaces*, because they satisfy the *solidity property*[95]

lphdubasic1
Lemma 82. (1) *If $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ is a minimal Banach function space, than $\boldsymbol{B}^\alpha = \boldsymbol{B}'$;*
(2) *For general Banach function spaces $\boldsymbol{B}^\alpha$ can be identified with the dual space of $\boldsymbol{B}_0$, the closure of the test functions in $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$.*
(3) *Thus the α-dual does not change when passing from $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ to $\boldsymbol{B}_0$!!*
(4) *(citation?) We have $\boldsymbol{B}^{\alpha\alpha} = \boldsymbol{B}$ if and only if $\boldsymbol{B}$ is a dual Banach function space.*

This is an inline comment in green!30.

H.S.Shapiro, 1971, in [91] introduces *ABSTRACT Banach modules* ... He is assuming the associativity in the axiomatic description and verifies that it is valid essentially for BF-spaces. (more precise reference to come, one of the last chapters of that book).

22. 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:

cospdef01
Definition 57. A distribution or measure σ is acting trivially at a given point $t \in \mathbb{R}^d$ (generally: $t \in G$) if there exists some $k \in \boldsymbol{C}_0(\mathbb{R}^d)$ (resp. $k \in \boldsymbol{S}_0(\mathbb{R}^d)$), with $k(t) = 1$ but $\sigma \cdot k = 0$, or equivalently $\sigma(k \cdot f) = 0, \forall f \in \boldsymbol{C}_0(\mathbb{R}^d)$ (resp. $\boldsymbol{S}_0(\mathbb{R}^d)$).

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 83. *A point t does not belong to the spectrum 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 $f \in \boldsymbol{C}_c(\mathbb{R}^d)$ with $\operatorname{supp}(f) \subset U$.*

[95] So $\left(\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_{\mathcal{F}\boldsymbol{L}^1}\right)$ is not a Banach function spaces, but of course a Banach space of continuous functions. So it is a function space in the sense of Triebel, who would also view the dual space, usually described as space of *pseudo-measures* as a function space, because it is a natural Banach space of tempered distributions on $\mathbb{R}^d$. We use the symbol $\mathcal{F}\boldsymbol{L}^\infty(\mathbb{R}^d)$ for this space, or define it simple as the dual space of $\mathcal{F}\boldsymbol{L}^\infty(\mathbb{R}^d)$. NORMED VERSION!

Proof. Assume that $k \cdot \sigma = 0$, then it is clear that $0 = (k \cdot \sigma)(h) = \sigma(k \cdot h)$ for any $h \in \boldsymbol{C}_0(\mathbb{R}^d)$. 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 [81]) one can find some function $\phi \in \boldsymbol{C}_c(\mathbb{R}^d)$ with $k(t) \cdot \phi(t) \equiv 1$ on some small 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". I don't think it is easy to describe without BUPUs! 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*

$$\psi_1 \cdot \sigma(f) = \sigma(\psi_1 \cdot f) = 0,$$

by the assumption, since

$$\operatorname{supp}(\psi_1 \cdot f) \subseteq \operatorname{supp}(\psi) \subseteq U.$$

This argument completes the proof. □

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).

pentrivact1 **Lemma 84.** *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. 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 $\sigma \cdot k = 0$. □

Now it is time to define the support of a distribution.

suppSOPRd **Definition 58.** The **support** $\operatorname{supp}(\sigma)$ 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 $\operatorname{supp}(\sigma)$ 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:

uppfctdist1 **Lemma 85.** *Given $k \in \boldsymbol{C}_c(\mathbb{R}^d)$ we find that the classical support* $\operatorname{supp}(k) := \{t \,|\, k(t) \neq 0\}^-$ *and the distributional support* $\operatorname{supp}(\sigma_k)$ *coincide. Furthermore, the notion of support does not depend on the algebra of test functions used.*

Proof.
- Assume that $t \notin \operatorname{supp}(k)$ (classical). Since $\operatorname{supp}(k)$ 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 $\operatorname{supp}(k_0) \subseteq U$, hence $k \cdot k_0 = 0$, as a function and consequently as a generalized function.
- Conversely, assume that $t \notin \operatorname{supp}(k)$, 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 ...

□

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 $Q := \operatorname{supp}(k)$ 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 this way the closure of $\{x \mid k(x) \neq 0\}$ 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 "nearby 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 59.** 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:

arspectrum1 **Lemma 86.** *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 $\boldsymbol{S}_0(\mathbb{R}^d) * \sigma \subset \boldsymbol{C}_{ub}(\mathbb{R}^d)$, 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 $g_\alpha * \sigma \to \chi_s$ 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 $\nu * \sigma$, 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 $\boldsymbol{L}^1(\mathbb{R}^d) * \sigma$, or even just the w^*-closure of the more decent (because continuous and bounded) functions of the form $\boldsymbol{S}_0(\mathbb{R}^d) * \sigma$ 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.

23. Ideas on BUPUS, Wiener Amalgam and Spline-Type Spaces

BUPUs are a universal tool for many questions in the theory of function spaces, they are quite useful in order to develop concepts for (conceptual) harmonic analysis, and they are crucial for the definition of Wiener amalgam spaces.

See a talk concerning the importance of BUPUs (Oslo, Feb. 2013)

`http://univie.ac.at/nuhag-php/dateien/talks/2484\_oslo13A.pdf`

Here is a short list of properties of these families that make them so important:

- They can be defined over arbitrary locally compact groups G. In fact, there use is implicit in the construction of the Haar measure following Cartan resp. A. Weil;
- By means of BUPUs it is easy to show that the discrete measures are w^*-dense in the space of bounded measures $\boldsymbol{M}_b(G)$.
- Obviously they are extremely useful in defining Wiener amalgam spaces (the discrete description is much more general, at least in order to introduce the spaces, than the "continuous" description);
- Spline-Type are quite important as well; they are obtained as "closed linear span" of a set of function and their translates, within some larger function space, say $\big(\boldsymbol{L}^p(\mathbb{R}^d), \|\cdot\|_p\big)$. If we find a *Riesz projection basis* for such a space than typically the discrete $\boldsymbol{\ell}^p$-norm on the coefficients, and the $\boldsymbol{L}^p(\mathbb{R}^d)$ or also the $\boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^p)(\mathbb{R}^d)$-norm are equivalent on the corresponding spline-type space.

For most purposes any result that is valid for regular BUPUs on $\mathbb{R}^d$ can be easily transferred to general statements about arbitrary BUPUs over locally compact, or at least locally compact Abelian groups. Among them the so-called *elementary* LCA groups are the most simple cases (cf. [44]).

24. Banach Gelfand Triples and Regularization Operators

The Banach Gelfand Triple $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$ (and in a similar war $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(G)$ over any given LCA group)

25. Fourier Standard Spaces

HINT: The idea of Fourier Standard Spaces has been developed in 2017/2018. So far there is no publication on this subject available, although the author of these notes has talked about the subject at different occasions, in particular first in the NuHAG seminar, and secondly at in invited talk during the conference on Fourier Analysis in Pecs (Hungary), in August 2017. The material presented below is more or less distilled from that talk, which is accessible at the NuHAG talk server at the address given below, see also other talks in the last few years found at the NuHAG server.

`https://www.univie.ac.at/nuhag-php/dateien/talks/3368\_PecsFeiAug17A.pdf`

Characterizing $\boldsymbol{S}_0(\mathbb{R}^d)$ via the STFT

While $V_g(f)$ is a continuous and square integrable function for any pair $g, f \in \boldsymbol{L}^2(\mathbb{R}^d)$, the choice of $g = g_0$, with $g_0(t) = e^{-\pi|t|^2}$, the usual Gaussian function is particularly pleasing, in fact in this case $V_{g_0}(f)$ is even an analytic function (>> Fock spaces).

For certain $\boldsymbol{L}^1(\mathbb{R})$-functions, like step functions f, their STFT $V_{g_0}(f)$ is *not integrable*, same for the SINC-function (bad decay at infinity), *but if* $V_{g_0}(f) \in \boldsymbol{L}^1(\mathbb{R}^{2d})$ it means that the STFT is reasonably well concentrated within phase space.

charS0Rd1 **Theorem 35.** *$f \in \boldsymbol{L}^2(\mathbb{R}^d)$ belongs to $\boldsymbol{S}_0(\mathbb{R}^d)$ if and only if $V_{g_0}(f) \in \boldsymbol{L}^1(\mathbb{R}^{2d})$. Moreover $\|V_{g_0} f\|_{\boldsymbol{L}^1}$ is an equivalent norm on $\boldsymbol{S}_0(\mathbb{R}^d)$. $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ is a dense subspace of the Banach space $\big(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\big)$.*

Characterization of $\boldsymbol{S}_0'(\mathbb{R}^d)$ and w^*-convergence

A tempered distribution $\sigma \in \boldsymbol{\mathcal{S}}'(\mathbb{R}^d)$ belongs to $\boldsymbol{S}_0'(\mathbb{R}^d)$ if and only if its (continuous) STFT is a *bounded* function. Furthermore convergence corresponds to uniform convergence of the spectrogram (different windows give equivalent norms!).

We can also extend the Fourier transform form $\boldsymbol{S}_0(\mathbb{R}^d)$ to $\boldsymbol{S}_0'(\mathbb{R}^d)$ via the usual formula $\hat{\sigma}(f) := \sigma(\hat{f})$.

The weaker convergence, arising from the functional analytic concept of w^*-convergence has the following very natural characterization: A (bounded) sequence σ_n is w^*- convergence to σ_0 if and only if for one (resp. every) $\boldsymbol{S}_0(\mathbb{R}^d)$-window g one has

$$V_g(\sigma_n)(\lambda) \to V_g(\sigma_0)(\lambda) \quad \text{for} n \to \infty,$$

uniformly over compact subsets of phase-space.

$\boldsymbol{L}^1(\mathbb{R}^d)$ and the Fourier Algebra $\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$

Fourier Standard Spaces, the Idea

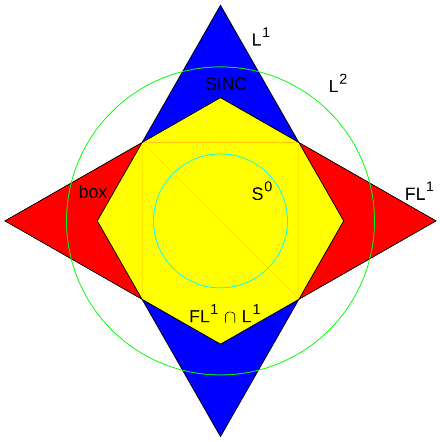


FIGURE 21. soplLIFLI3.jpg generated by soplot5.m

rStandDef00 **Definition 60.** A Banach space $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$, continuously embedded between $\boldsymbol{S}_0(G)$ and $\big(\boldsymbol{S}_0'(G), \|\cdot\|_{\boldsymbol{S}_0'}\big)$, i.e. with

sandwhich00 (311)
$$\big(\boldsymbol{S}_0(G), \|\cdot\|_{\boldsymbol{S}_0}\big) \hookrightarrow (\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}}) \hookrightarrow \big(\boldsymbol{S}_0'(G), \|\cdot\|_{\boldsymbol{S}_0'}\big)$$

is called a **Fourier Standard Space** on G (a FouSS) if it has a double module structure over $(\boldsymbol{M}_b(G), \|\cdot\|_{\boldsymbol{M}_b})$ with respect to convolution and over the (Fourier-Stieltjes algebra) $\mathcal{F}(\boldsymbol{M}_b(\widehat{G}))$ with respect to pointwise multiplication.

Typically we just require that in addition to (311) one has:

oubmodcond1 (312)
$$\boldsymbol{L}^1 * \boldsymbol{B} \subseteq \boldsymbol{B} \quad \text{and} \quad \mathcal{F}\boldsymbol{L}^1 \cdot \boldsymbol{B} \subseteq \boldsymbol{B}.$$

Constructions within the FouSS Family

(1) Taking **Fourier transforms**;
(2) Conditional dual spaces, i.e. the **dual space** of the closure of $\boldsymbol{S}_0(G)$ within $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$;
(3) With two spaces $\boldsymbol{B}^1, \boldsymbol{B}^2$: take **intersection or sum**
(4) forming **amalgam spaces** $\boldsymbol{W}(\boldsymbol{B}, \boldsymbol{\ell}^q)$; e.g. $\boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1, \boldsymbol{\ell}^1)$;
(5) defining pointwise or convolution **multipliers**;
(6) using complex (or real) **interpolation methods**, so that we get the spaces $\boldsymbol{M}^{p,p} = \boldsymbol{W}(\mathcal{F}\boldsymbol{L}^p, \boldsymbol{\ell}^p)$ (all Fourier invariant);
(7) any **metaplectic** image of such a space, e.g. the **fractional Fourier transform**.

Recalling the Wiener Amalgam Concept

We recall the concept of BUPUs, ideally as translates along a lattice $(T_|lambda\varphi)$, with compact support and a certain amount of smoothness, perhaps cubic B-splines.

The Wiener amalgam space $\boldsymbol{W}(\boldsymbol{B}, \boldsymbol{\ell}^q)$ is defined as the set

$$\{f \in \boldsymbol{B}_{loc} \,|\, \|f|\boldsymbol{W}(\boldsymbol{B}, \boldsymbol{\ell}^q)\| := (\sum_{\lambda\in\Lambda} \|f \cdot T_\lambda\varphi\|_{\boldsymbol{B}}^q)^{1/q}\}$$

There are many "natural results" concerning Wiener amalgam spaces ([31]), namely coordinatewise action, e.g.

- duality (if test functions are dense and $q < \infty$;
- convolution and pointwise multiplication;

- interpolation (real or complex).

FOURIER STANDARD SPACES: II

The spaces in this family are useful for a discussion of questions in Gabor Analysis, which is an important branch of time-frequency analysis, but also for problems of classical Fourier Analysis, such as the discussion of Fourier multipliers, Fourier inversion questions and so on. Thus among others the space $\boldsymbol{L}^1(\mathbb{R}^d) \cap \mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d)$.

Within the family there are two subfamilies, namely the **Wiener amalgam spaces** and the so-called **modulation spaces**, among them the Segal algebra $\big(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\big)$ or (what we call) **Wiener's algebra** $\big(\boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{W}}\big)$.

TF-homogeneous Banach Spaces

TFhomogdef1 **Definition 61.** A Banach space $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ with

$$\boldsymbol{\mathcal{S}}(\mathbb{R}^d) \hookrightarrow (\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}}) \hookrightarrow \boldsymbol{\mathcal{S}}'(\mathbb{R}^d)$$

is called a **TF-homogeneous Banach space** if $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ is dense in $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ and TF-shifts act isometrically on $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$, i.e. if

$$\|\pi(\lambda)f\|_{\boldsymbol{B}} = \|f\|_{\boldsymbol{B}}\,, \quad \forall \lambda \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d,\ f \in \boldsymbol{B}. \tag{313}$$

For such spaces the mapping $\lambda \to \pi(\lambda)f$ is continuous from $\mathbb{R}^d \times \widehat{\mathbb{R}}^d$ to $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$. If it is not continuous one often has the *adjoint action* on the dual space of such TF-homogeneous Banach spaces (e.g. $\big(\boldsymbol{L}^\infty(\mathbb{R}^d), \|\cdot\|_\infty\big)$).

TF-homogeneous Banach Spaces II

An important fact concerning this family is the minimality property of the Segal algebra $\big(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\big)$.

S0min1 **Theorem 36.** *There is a smallest member in the family of all TF-homogeneous Banach spaces, namely the Segal algebra* $\big(\boldsymbol{S}_0(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{S}_0}\big) = \boldsymbol{W}(\mathcal{F}\boldsymbol{L}^1, \boldsymbol{\ell}^1)(\mathbb{R}^d)$.

Justifying the properties of the family

There is a large number of results concerning $\boldsymbol{S}_0(\mathbb{R}^d)$ (defined for any dimension, but in fact for any LCA group):

$$\mathcal{F}_G \boldsymbol{S}_0(G) = \boldsymbol{S}_0(\widehat{G}).$$

There is a tensor product property, namely

$$\boldsymbol{S}_0(\mathbb{R}^{2d}) = \boldsymbol{S}_0(\mathbb{R}^d) \widehat{\otimes} \boldsymbol{S}_0(\mathbb{R}^d).$$

Multipliers are easily characterized:

multS0S0P1 **Theorem 37.** *The continuous linear operators from* $\boldsymbol{S}_0(\mathbb{R}^d)$ *to* $\boldsymbol{S}_0'(\mathbb{R}^d)$ *are exactly the convolution operators by "kernels"* $\sigma \in \boldsymbol{S}_0'(\mathbb{R}^d)$, *with equivalence of norms (the operator norm of the convolution operator, resp. translation invariant linear system) and the* $\boldsymbol{S}_0'$*-norm of the corresponding convolution kernel* $\|\sigma\|_{\boldsymbol{S}_0'}$.

The Kernel Theorem

It is clear that such operators between functions on $\mathbb{R}^d$ cannot all be represented by integral kernels using locally integrable $K(x,y)$ in the form

kernelth002

$$Tf(x) = \int_{\mathbb{R}^d} K(x,y)f(y)dy, \quad x,y \in \mathbb{R}^d, \tag{314}$$

because clearly multiplication operators should have their support on the main diagonal, but $\{(x,x) \mid x \in \mathbb{R}^d\}$ is just a set of measure zero in $\mathbb{R}^d \times \mathbb{R}^d$!

Also the expected "rule" to find the kernel, namely

entkernel01

$$K(x,y) = T(\delta_y)(x) = \delta_x(T(\delta_y)) \tag{315}$$

might not be meaningful at all.

The Hilbert Schmidt Version

There are two ways out of this problem

- restrict the class of operators
- enlarge the class of possible kernels

The first one is a classical result, i.e. the characterization of the class $\mathcal{HS}$ of Hilbert Schmidt operators.

lbertSchm01 **Theorem 38.** *A linear operator T on $\left(\boldsymbol{L}^2(\mathbb{R}^d), \|\cdot\|_2\right)$ is a Hilbert-Schmidt operator, i.e. is a compact operator with the sequence of singular values in $\boldsymbol{\ell}^2$ if and only if it is an integral operator of the form (314) with $K \in \boldsymbol{L}^2(\mathbb{R}^d \times \mathbb{R}^d)$. In fact, we have a unitary mapping $T \to K(x,y)$, where $\mathcal{HS}$ is endowed with the Hilbert-Schmidt scalar product $\langle T, S\rangle_{\mathcal{HS}} := \mathrm{trace}(T \circ S^*)$.*

The Schwartz Kernel Theorem

The other well known version of the kernel theorem makes use of the *nuclearity* of the *Frechet space* $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ (so to say the complicated topological properties of the system of seminorms defining the topology on $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$).

Note that the description cannot be given anymore in the form (314) but has to replaced by a "weak description". This is part of the following well-known result due to L. Schwartz.

artzKern001 **Theorem 39.** *There is a natural isomorphism between the vector space of all linear operators from $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ into $\boldsymbol{\mathcal{S}}'(\mathbb{R}^d)$, i.e. $\mathcal{L}(\boldsymbol{\mathcal{S}}, \boldsymbol{\mathcal{S}}')$, and the elements of $\boldsymbol{\mathcal{S}}'(\mathbb{R}^{2d})$, via $\langle Tf, g\rangle = \langle K, f \otimes g\rangle$, for $f, g \in \boldsymbol{\mathcal{S}}(\mathbb{R}^d)$.*

The S_0-KERNEL THEOREM

In the current setting we can describe the kernel theorem as a unitary Banach Gelfand Triple isomorphism, between operator and their (distributional) kernels, extending the classical Hilbert Schmidt version.

First we observe that $\boldsymbol{S}_0$-kernels can be identified with $\mathcal{L}(\boldsymbol{S}_0, \boldsymbol{S}_0')$, i.e. the *regularizing operators* from $\boldsymbol{S}_0'(\mathbb{R}^d)$ to $\boldsymbol{S}_0(\mathbb{R}^d)$, even mapping bounded and w^*- convergent nets into norm convergent sets. For those kernels also the recovery formula (315) is valid.

kernelthm01 **Theorem 40.** *The unitary Hilbert-Schmidt kernel isomorphisms extends in a unique way to a Banach Gelfand Triple isomorphism between $(\mathcal{L}(\boldsymbol{S}_0', \boldsymbol{S}_0), \mathcal{HS}, \mathcal{L}(\boldsymbol{S}_0, \boldsymbol{S}_0'))$ and $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d \times \mathbb{R}^d)$.*

The Spreading Representation

The **spreading representation** of operators interpretation. In some sense it can be viewed as a kind of Fourier Transform for operators. For the case of $G = \mathbb{Z}_N$ we have N^2 time-frequency shift operators (cyclic shifts combined with pointwise multiplication by pure frequencies), and in fact they form an orthonormal basis for the (Euclidean) space of $N \times N$-matrices (linear operators on $\mathbb{C}^N$), with the Erogenous scalar product.

readRepr001 **Theorem 41.** *There is a unique (unitary) Banach Gelfand triple isomorphism between* $(\mathcal{L}(\boldsymbol{S}_0', \boldsymbol{S}_0), \mathcal{H}S, \mathcal{L}(\boldsymbol{S}_0, \boldsymbol{S}_0'$ *and* $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d \times \widehat{\mathbb{R}}^d)$*, which maps the time frequency shift operators* $\pi(\lambda) := M_\omega T_t$ *to the Dirac measures* $\delta_{t,\omega} \in \boldsymbol{S}_0'(\mathbb{R}^d \times \widehat{\mathbb{R}}^d)$*.*

Spreading Representation II

This also tells us, that an operator $T \in \mathcal{L}(\boldsymbol{S}_0', \boldsymbol{S}_0)$ is regularizing if it can be written as an operator-valued Riemannian integral

SOspread01

$$T = \int_{\mathbb{R}^d \times \widehat{\mathbb{R}}^d} \eta(\lambda) \pi(\lambda) d\lambda. \tag{316}$$

Of course one can also write explicit formulas (involving various transformations and partial Fourier transform) for the transition between the kernel of an operator T and its spreading "function" $\eta(T)$ (cf. [44]) which are valid in the pointwise sense (using standard integration theory), while one has to extend it to the Hilbert space case by continuity (like the usual proofs of Plancherel's theorem) and then extend it to the *outer layer* via duality (or w^*-continuity).

See also [19]

The Kohn-Nirenberg Symbol

For various applications in the area of *pseudo-differential operators* and for applications in Gabor Analysis also the so-called **Kohn-Nirenberg Symbol** $\sigma(T)$ of an operator T is of interest. It is obtained from the spreading representation via the so-called *symplectic Fourier transform.*

The (unitary)KNS Banach Gelfand triple isomorphism between $(\mathcal{L}(\boldsymbol{S}_0', \boldsymbol{S}_0), \mathcal{H}S, \mathcal{L}(\boldsymbol{S}_0, \boldsymbol{S}_0'))$ and $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d \times \widehat{\mathbb{R}}^d)$, $T \to \sigma(T)$ has the following covariance property:

$$\sigma[\pi(\lambda) \circ T \circ \pi(\lambda)'] = T_{t,\omega} \sigma(T).$$

Fourier Standard spaces on $\mathbb{R}^{2d}$

The kernel theorem provides a variety of new possibilities to define Fourier Standard spaces on product domains, e.g. (for simplicity) on $\mathbb{R}^d \times \mathbb{R}^d$. Among others we have

(1) Given two Fourier standard spaces on $\mathbb{R}^d$ the kernels (or the Kohn-Nirenberg symbols, etc.) of the space of all linear operators $\boldsymbol{L}(\boldsymbol{B}^1, \boldsymbol{B}^2)$ is a FouSS;
(2) Given any **operator ideal** within $\mathcal{L}(\mathcal{H})$ defines a corresponding FouSS of kernels (e.g. Schatten S_p-classes, various types of nuclear operators etc., see book of Pietsch);
(3) some of them can be periodized (or restricted) along certain subgroups, e.g. die diagonal $\{(x,x) \,|\, , x \in \mathbb{R}\}$. The resulting spaces are then comparable with the **Herz algebras** $\boldsymbol{A}_p(\mathbb{R}^d)$, characterizing the pre-dual of all $\boldsymbol{L}^p$-multipliers.

What kind of questions can we ask?

There are at least two major type of questions which one can ask, related to the possibility of creating new spaces within the family. The key constructions have to do with

(1) Wiener amalgam spaces of the form $\boldsymbol{W}(\boldsymbol{B}, \boldsymbol{\ell}^q)$;
(2) the double module structure on these spaces.

Let us recall that for the construction of Wiener amalgam spaces we only need the possibility of applying a BUPU (a bounded partition of unity), in our case boundedness refers to $\big(\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d), \, \|\cdot\|_{\mathcal{F}\boldsymbol{L}^1}\big)$.

Moreover there is a clear chain of proper inclusions then, with

$$\boldsymbol{W}(\boldsymbol{B},\boldsymbol{\ell}^1) \subset \boldsymbol{W}(\boldsymbol{B},\ell^{p_1})$$

Lower and Upper Index of a Function Space

We define (compare [39]) the *lower and upper index* of such a space is called

upplowind1 **Definition 62.**

$$\mathrm{low}(\boldsymbol{B}) := \sup\{r \,|\, W(\boldsymbol{B},\ell^r) \subseteq \boldsymbol{B}\}.$$

upplowind2 **Definition 63.** The *upper index* of $\boldsymbol{B}$ is defined as follows:

$$\mathrm{upp}(\boldsymbol{B}) := \inf\{s \,|\, \boldsymbol{B} \subseteq W(\boldsymbol{B},\ell^s)\}.$$

Indices of concrete function spaces

LpFLpIndex **Lemma 87.**
- *Clearly* $\mathrm{low}(\boldsymbol{L}^p) = p = \mathrm{upp}(\boldsymbol{L}^p)$, *for* $1 \le p \le \infty$; *in fact* $\boldsymbol{W}(\boldsymbol{L}^p,\boldsymbol{\ell}^p) = \boldsymbol{L}^p(\mathbb{R}^d)$ *with norm equivalence;*
- *For* $1 \le p \le 2$ *one has*

$$\mathrm{low}(\mathcal{F}\boldsymbol{L}^p) = p, \quad \mathrm{upp}(\mathcal{F}\boldsymbol{L}^p) = p'.$$

For the space of multipliers we only know for sure: Since

$$\boldsymbol{B}^1 = \mathcal{H}_{\mathbb{R}^d}(\boldsymbol{L}^1(\mathbb{R}^d)) = \boldsymbol{M}_b(\mathbb{R}^d) = \boldsymbol{W}(\boldsymbol{M}_b(\mathbb{R}^d),\boldsymbol{\ell}^1)$$

$$\mathrm{low}(\boldsymbol{B}^1) = 1 = \mathrm{upp}(\boldsymbol{B}^1)$$

while in contrast for $p = 2$ we have $\boldsymbol{B}^2 = \mathcal{H}_{\mathbb{R}^d}(\boldsymbol{L}^2(\mathbb{R}^d)) = \mathcal{F}\boldsymbol{L}^\infty(\mathbb{R}^d)$ and hence

$$\mathrm{low}(\boldsymbol{B}^2) = 1, \quad \mathrm{upp}(\boldsymbol{B}^2) = \infty.$$

Lower and Upper $\boldsymbol{M}^p$-index: NEW 2018!

There is another index which can be defined on the basis of the observation that there is a *scale of Banach spaces*, connecting $\boldsymbol{S}_0(\mathbb{R}^d) = \boldsymbol{M}^{1,1}_0(\mathbb{R}^d)$ with $\boldsymbol{S}_0'(\mathbb{R}^d) = \boldsymbol{M}^\infty(\mathbb{R}^d)$

Banach Module Terminology

Recall the definition 25 which we repeat here:

BanMod2 **Definition 64.** A Banach space $(\boldsymbol{B}, \, \|\cdot\|_{\boldsymbol{B}})$ is a *Banach module* over a Banach algebra $(\boldsymbol{A}, \cdot, \, \|\cdot\|_{\boldsymbol{A}})$ if one has a bilinear mapping $(a,b) \mapsto a \bullet b$, from $\boldsymbol{A} \times \boldsymbol{B}$ into $\boldsymbol{B}$ bilinear and associative, such that

BanModEst (317)
$$\|a \bullet b\|_{\boldsymbol{B}} \le \|a\|_{\boldsymbol{A}} \|b\|_{\boldsymbol{B}} \quad \forall\, a \in \boldsymbol{A}, b \in \boldsymbol{B},$$

BanModAss (318)
$$a_1 \bullet (a_2 \bullet b) = (a_1 \cdot a_2) \bullet b \quad \forall a_1, a_2 \in \boldsymbol{A}, b \in \boldsymbol{B}.$$

BanIdDef **Definition 65.** A Banach space $(\boldsymbol{B}, \, \|\cdot\|_{\boldsymbol{B}})$ is a *Banach ideal* in (or within, or of) a Banach algebra $(\boldsymbol{A}, \cdot, \, \|\cdot\|_{\boldsymbol{A}})$ if $(\boldsymbol{B}, \, \|\cdot\|_{\boldsymbol{B}})$ is continuously embedded into $(\boldsymbol{A}, \cdot, \, \|\cdot\|_{\boldsymbol{A}})$, and if in addition (317) is valid with respect to the internal multiplication inherited from $\boldsymbol{A}$.

Wendel's Theorem

Wendel-thm1 **Theorem 42.** *The space of $\mathcal{H}_{\boldsymbol{L}^1}(\boldsymbol{L}^1, \boldsymbol{L}^1)$ all bounded linear operators on $\boldsymbol{L}^1(G)$ which commute with translations (or equivalently: with convolutions) is naturally and isometrically identified with $(\boldsymbol{M}_b(G), \|\cdot\|_{\boldsymbol{M}_b})$. In terms of our formulas this means*

$$\mathcal{H}_{\boldsymbol{L}^1}(\boldsymbol{L}^1, \boldsymbol{L}^1)(\mathbb{R}^d) \simeq (\boldsymbol{M}_b(\mathbb{R}^d), \|\cdot\|_{\boldsymbol{M}_b}),$$

$$\textit{via } T \simeq C_\mu : f \mapsto \mu * f, \quad f \in \boldsymbol{L}^1, \mu \in \boldsymbol{M}_b(\mathbb{R}^d).$$

For a given Banach space of distributions which is an $\boldsymbol{L}^1(\mathbb{R}^d)$ Banach convolution module we have in a quite general setting:

BsupLisp **Lemma 88.**

$$B_{\boldsymbol{L}^1} = \{f \in \boldsymbol{B} \,|\, \|T_x f - f\|_{\boldsymbol{B}} \to 0, \ for\ x \to 0\}.$$

Consequently we have $(\boldsymbol{M}_b(\mathbb{R}^d))_{\boldsymbol{L}^1} = \boldsymbol{L}^1(\mathbb{R}^d)$, the closed ideal of absolutely continuous bounded measures on $\mathbb{R}^d$.

Pointwise Multipliers

Via the Fourier transform we have similar statements for the Fourier algebra, involving the *Fourier Stieltjes algebra.*

HFLiFli (319)
$$\mathcal{H}_{\mathcal{F}\boldsymbol{L}^1}(\mathcal{F}\boldsymbol{L}^1, \mathcal{F}\boldsymbol{L}^1) = \mathcal{F}(\boldsymbol{M}_b(\mathbb{R}^d)), \quad \mathcal{F}(\boldsymbol{M}_b(\mathbb{R}^d))_{\mathcal{F}\boldsymbol{L}^1} = \mathcal{F}\boldsymbol{L}^1.$$

HCOCOCO **Theorem 43.** *The completion of $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$ (viewed as a Banach algebra and module over itself) is given by*

$$\mathcal{H}_{\boldsymbol{C}_0}(\boldsymbol{C}_0, \boldsymbol{C}_0) = \big(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\big).$$

On the other hand we have $(\boldsymbol{C}_b(\mathbb{R}^d))_{\boldsymbol{C}_0} = \boldsymbol{C}_0(\mathbb{R}^d)$.

Essential part and closure

In the sequel we assume that $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ is a Banach algebra with bounded approximate units, such as $\big(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\big)$ (with convolution), or $\big(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\big)$ or $\big(\mathcal{F}\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_{\mathcal{F}\boldsymbol{L}^1}\big)$ with pointwise multiplication.

ModClosHull **Theorem 44.** *Let $\boldsymbol{A}$ be a Banach algebra with bounded approximate units, and $\boldsymbol{B}$ a Banach module over $\boldsymbol{A}$. Then we have the following general identifications:*

BClosHull (320)
$$(\boldsymbol{B}_{\boldsymbol{A}})_{\boldsymbol{A}} = \boldsymbol{B}_{\boldsymbol{A}}, \quad (\boldsymbol{B}^{\boldsymbol{A}})_{\boldsymbol{A}} = \boldsymbol{B}_{\boldsymbol{A}}, \quad (\boldsymbol{B}_{\boldsymbol{A}})^{\boldsymbol{A}} = \boldsymbol{B}^{\boldsymbol{A}}, \quad (\boldsymbol{B}^{\boldsymbol{A}})^{\boldsymbol{A}} = \boldsymbol{B}^{\boldsymbol{A}}.$$

or in a slightly more compact form:

BClosHullx (321)
$$\boldsymbol{B}_{\boldsymbol{A}\boldsymbol{A}} = \boldsymbol{B}_{\boldsymbol{A}}, \quad \boldsymbol{B}^{\boldsymbol{A}}{}_{\boldsymbol{A}} = \boldsymbol{B}_{\boldsymbol{A}}, \quad \boldsymbol{B}_{\boldsymbol{A}}{}^{\boldsymbol{A}} = \boldsymbol{B}^{\boldsymbol{A}}, \quad \boldsymbol{B}^{\boldsymbol{A}\boldsymbol{A}} = \boldsymbol{B}^{\boldsymbol{A}}.$$

Essential Banach modules and BAIs

The usual way to define the *essential part* $\boldsymbol{B}_{\boldsymbol{A}}$ resp. $\boldsymbol{B}_e$ of a Banach module $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ with respect to some Banach algebra action $(\mathbf{a}, \mathbf{b}) \mapsto \mathbf{a} \bullet \mathbf{b}$ is defined as the closed linear span of $\boldsymbol{A} \bullet \boldsymbol{B}$ within $\big(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}}\big)$. This subspace has other nice characterizations using BAIs (bounded approximate units (BAI) in $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$):

sspartchar1 **Lemma 89.** *For any BAI $(\mathbf{e}_\alpha)_{\alpha \in I}$ in $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ one has:*

esspart111 (322)
$$\boldsymbol{B}_{\boldsymbol{A}} = \{\mathbf{b} \in \boldsymbol{B} \,|\, \lim_\alpha \mathbf{e}_\alpha \bullet \mathbf{b} = \mathbf{b}\}$$

The Cohen-Hewitt Factorization Theorem

In particular one has: Let $(\mathbf{e}_\alpha)_{\alpha\in I}$ and $(\mathbf{u}_b)_{\beta\in J}$ be two bounded approximate units (i.e. bounded nets within $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ acting in the limit like an identity in the Banach algebra $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$. Then

BAIequiv2b
$$\lim_\alpha \mathbf{e}_\alpha \bullet \mathbf{b} = \mathbf{b} \quad \Leftrightarrow \quad \lim_\beta \mathbf{u}_\beta \bullet \mathbf{b} = \mathbf{b}. \tag{323}$$

CohHew **Theorem 45.** *(The Cohen-Hewitt factorization theorem, without proof, see [58])*
Let $(\boldsymbol{A}, \|\cdot\|_{\boldsymbol{A}})$ be a Banach algebra with some BAI of size $C > 0$, then the algebra factorizes, which means that for every $a \in \boldsymbol{A}$ there exists a pair $a', h' \in \boldsymbol{A}$ such that $a = h' \cdot a'$, in short: $\boldsymbol{A} = \boldsymbol{A} \cdot \boldsymbol{A}$. In fact, one can even choose $\|a - a'\| \leq \varepsilon$ and $\|h'\| \leq C$.

Essential part and closure II

Having now Banach spaces of distributions which have two module structures, we have to use corresponding symbols.
FROM NOW ON we will use the letter $\boldsymbol{A}$ mostly for pointwise Banach algebras and thus for the $\mathcal{F}\boldsymbol{L}^1$-action on $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$, and we will use the symbol G (because convolution is coming from the integrated group action!) for the $\boldsymbol{L}^1$ convolution structure. We thus have

BClosHullp
$$\boldsymbol{B}_{\boldsymbol{GG}} = \boldsymbol{B}_{\boldsymbol{G}}, \quad {\boldsymbol{B}^{\boldsymbol{G}}}_{\boldsymbol{G}} = \boldsymbol{B}_{\boldsymbol{G}}, \quad \boldsymbol{B}_{\boldsymbol{G}}^{\boldsymbol{G}} = \boldsymbol{B}^{\boldsymbol{G}}, \quad \boldsymbol{B}^{\boldsymbol{GG}} = \boldsymbol{B}^{\boldsymbol{G}}. \tag{324}$$

In this way we can combine the two operators (in view of the above formulas we can call them interior and closure operation) with respect to the two module actions and form spaces such as

$$ {\boldsymbol{B}^{\boldsymbol{G}}}_{\boldsymbol{A}}, \quad {\boldsymbol{B}_{\boldsymbol{A}}}^{\boldsymbol{G}}{}_{\boldsymbol{A}}, \quad {{\boldsymbol{B}^{\boldsymbol{G}}}_{\boldsymbol{A}}}^{\boldsymbol{G}}{}_{\boldsymbol{A}} \cdots $$

or changes of arbitrary length, as long as the symbols $\boldsymbol{A}$ and
$\boldsymbol{G}$ appear in alternating form (at any position, upper or lower).

Combining the two module structures

Fortunately one can verify (paper with W.Braun from 1983, J.Funct.Anal.) that any "long" chain can be reduced to a chain of at most two symbols, the *last occurrence of each of the two symbols being the relevant one!* So in fact all the three symbols in the above chain describe the same space of distributions.

But still we are left with the following collection of altogether eight two-letter symbols:

twoletter
$$\boldsymbol{B}_{\boldsymbol{GA}}, \boldsymbol{B}_{\boldsymbol{AG}}, {\boldsymbol{B}_{\boldsymbol{A}}}^{\boldsymbol{G}}, {\boldsymbol{B}^{\boldsymbol{G}}}_{\boldsymbol{A}}, {\boldsymbol{B}_{\boldsymbol{G}}}^{\boldsymbol{A}}, {\boldsymbol{B}^{\boldsymbol{A}}}_{\boldsymbol{G}}, \boldsymbol{B}^{\boldsymbol{AG}}, \boldsymbol{B}^{\boldsymbol{GA}} \tag{325}$$

and of course the four one-symbol objects

twoletter
$$\boldsymbol{B}_{\boldsymbol{A}}, \boldsymbol{B}_{\boldsymbol{G}}, \boldsymbol{B}^{\boldsymbol{A}}, \boldsymbol{B}^{\boldsymbol{G}} \tag{326}$$

Some structures, simple facts

There are other, quite simple and useful facts, such as

$$\mathcal{H}_{\boldsymbol{A}}({\boldsymbol{B}^1}_{\boldsymbol{A}}, \boldsymbol{B}^2) = \mathcal{H}_{\boldsymbol{A}}({\boldsymbol{B}^1}_{\boldsymbol{A}}, {\boldsymbol{B}^2}_{\boldsymbol{A}}) \tag{327}$$

which can easily be verified if ${\boldsymbol{B}^1}_{\boldsymbol{A}} = \boldsymbol{A}\bullet\boldsymbol{B}^1$, since then $T \in \mathcal{H}_{\boldsymbol{A}}({\boldsymbol{B}^1}_{\boldsymbol{A}}, \boldsymbol{B}^2)$ applied to $\mathbf{b}^1 = \mathbf{a}\bullet\mathbf{b}^{1'}$ gives

$$T(\mathbf{b}^1) = T(\mathbf{a} \bullet \mathbf{b}^{1'}) = \mathbf{a} \bullet T(\mathbf{b}^{1'}) \in {\boldsymbol{B}^2}_{\boldsymbol{A}}.$$

The Main Diagram

This diagram is taken from [13].

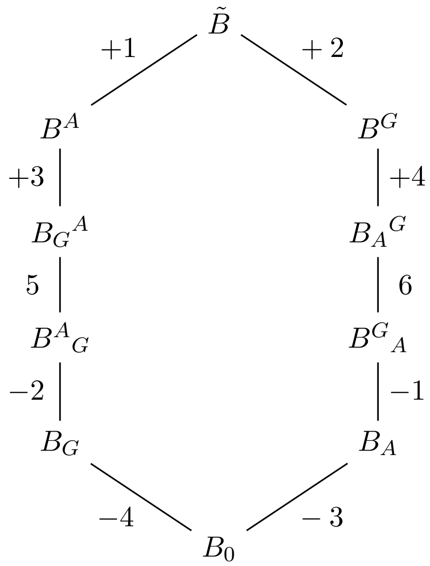


FIGURE 22. Figure caption fig:dummy

26. For $\left(\boldsymbol{C}_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$

In this case we get 6 different spaces (to be identified separately).

From the point of view of the main diagram it is clear that $C_0(R^d)$ is minimal, i.e. $\boldsymbol{S}_0(\mathbb{R}^d)$ is dense in $\boldsymbol{B}$, hence the connections labeled as -4 and -3 become identities, or in other words:

$$\boldsymbol{B}_A = \boldsymbol{B}_{AG} = \boldsymbol{B}_0 = \boldsymbol{B}_{GA} = \boldsymbol{B}_G.$$

Consequently the inclusions $+3$ and $+4$ are just identities as well, *but all the other six inclusions are proper inclusions!*

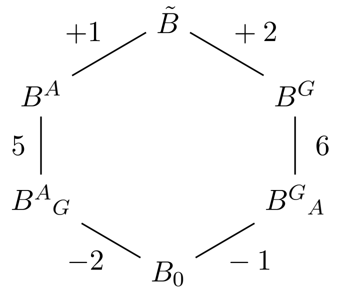


FIGURE 23. Figure caption fig:dummy

27. For $L^\infty(R^d)$

For $\boldsymbol{L}^\infty(\mathbb{R}^d)$ we have the same collection of spaces but with different labeling. Since $\boldsymbol{L}^\infty(\mathbb{R}^d)$ is clearly a dual space the space is a *maximal space* and hence the inclusions labeled with $+/-1$ and $+/-2$ are identities, while 3 and 4 then proper inclusions.

So let us list the spaces which appear:

Starting from the dual space it is clear that $\boldsymbol{B}_A$ is the (proper) subspace of all functions in this space "vanishing at infinity", while $\boldsymbol{B}_A^G$ is the space of all functions which, whenever convolved with an $\boldsymbol{L}^1$-function will tend to zero at infinity (and in fact will convolve all of $\boldsymbol{L}^1$ into $\boldsymbol{C}_0$. This space contains functions with narrow spikes, which do NOT tend to zero in its amplitude but which have smaller and smaller support, so that they tend to zero after convolution with an $\boldsymbol{L}^1$-function.

Minimal and Maximal Spaces

More or less the diagram shows that for every FouSS norm there exists a whole family of function spaces *with the SAME norm*, i.e. closed within each other.

A FouSS space $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ is called **maximal** if $\boldsymbol{B} = \widetilde{\boldsymbol{B}}$, or equivalently the following is true: Assume that σ_α is a bounded net of elements in $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$, with $\sigma_0 = w^* - lim_\alpha \sigma_a$ (in $\boldsymbol{S}_0'(\mathbb{R}^d)$) implies that $\sigma_0 \in \boldsymbol{B}$.

A FouSS space $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ is called **minimal** if $\boldsymbol{B} = \boldsymbol{B}_{AG}$, resp. if $\boldsymbol{S}_0(\mathbb{R}^d)$ is dense in $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$.

Characterization

dualmax1 **Theorem 46.** *A Banach space $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ is maximal if and only if it is the dual space of some other minimal Fourier standard space. The predual can be determined as*

$$\boldsymbol{B} = (\boldsymbol{B}_o)_o'.$$

Example: $\boldsymbol{B} = \boldsymbol{L}^\infty(\mathbb{R}^d)$ is the dual space. $\boldsymbol{B}_o = \boldsymbol{C}_0(\mathbb{R}^d)$, hence $\boldsymbol{B}_o' = \boldsymbol{M}_b(\mathbb{R}^d)$, and consequently $(\boldsymbol{B}_o')_o = \boldsymbol{L}^1(\mathbb{R}^d)$ shows up as the pre-dual of $\boldsymbol{L}^\infty(\mathbb{R}^d)$.

reflex01 **Theorem 47.** *A FouSS $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ is reflexive if and only if $\boldsymbol{B}$ and $\boldsymbol{B}'$ are minimal and maximal.*

Outlook

COMMENT given at the presentation of this material at a conference on Fourier Analysis in Pecs (Hungary, Summer 2017), see:

`http://www.univie.ac.at/nuhag-php/program/talks_details.php?id=3368`

I hope that I have shown that this view on (classical and modern) function spaces provide still a large number of *interesting and open questions.* So Fourier Analysis is not at all an outdated subject area within mathematical analysis.

We have not even discussed *Time-Frequency Analysis*, Gabor Analysis, wavelet theory, shearlet theory, or the theory of pseudo-differential operators.

Optimal extension of the Hausdorff-Young inequality

```
MR2427981 (2009j:46071) Mockenhaupt, G. ; Ricker, W. J.
Optimal extension of the Hausdorff-Young inequality.
J. Reine Angew. Math. 620 (2008), 195211.
```

Main result: For the typical Hausdorff-Young setting, i.e. for $p \in [1,2)$ there is a solid (meaning Banach lattice) FouSS $(\boldsymbol{B}, \|\cdot\|_{\boldsymbol{B}})$ strictly larger than $\boldsymbol{L}^p$ such that still $\mathcal{F}(\boldsymbol{B}) \subset \boldsymbol{L}^q$.

This space can be described as the pointwise multipliers from $\boldsymbol{L}^\infty$ to $\mathcal{F}^{-1}(\boldsymbol{L}^p)$, so it is obviously solid (because pointwise multiplication with $\boldsymbol{L}^\infty$ is a bounded operation) and has the right Fourier property. The strict inclusion requires analysis!

Constructions within the FouSS Family IV

There is a small body of literature (mostly papers by Kelly McKennon, a former PhD student of Edwin Hewitt) concerning spaces of "tempered elements". He has done the case starting $\boldsymbol{B} = \boldsymbol{L}^p(G)$, over general LC groups, but the construction makes sense if (and only if) one has a nice invariant space which happens not to be a convolution (or pointwise) algebra.

By *intersecting* the space with its own "multiplier algebra" one obtains an (abstract) Banach algebra, and often the Banach algebra homomorphism of this new algebra "are" just the translation invariant operators on the original spaces.

For the case of $\boldsymbol{B} = \big(\boldsymbol{L}^p(\mathbb{R}^d), \|\cdot\|_p\big)$ one would define

$$\boldsymbol{L}_p^t := \boldsymbol{L}^p \cap \mathcal{H}_G(\boldsymbol{L}^p, \boldsymbol{L}^p).$$

Tempered elements in $\boldsymbol{L}^p$-spaces, understood as the intersection of two FouSSs, with the natural norm, which is the sum of the $\boldsymbol{L}^p$-norm of f plus the operator norm of the convolution operator.

For $p > 2$ one has to be careful and has to define that operator norm only by looking at the action of $k \to k * f$ on $\boldsymbol{C}_c(\mathbb{R}^d)$! (convolution in the pointwise sense might fail to exist, on more than just a null-set!).

However it is not a problem to approximate every element (in norm or even just in the w^*-sense) by test-functions in $\boldsymbol{S}_0(\mathbb{R}^d)$ and then take the limit of the convolution products of the regularized expressions

28. Riesz Bases and Banach Frames

In the terminology introduced in XX this means that $\mathcal{R}$ is injective, but not surjective, and $\mathcal{C}$ is a left inverse of $\mathcal{R}$. Thus we have the following commutative diagram.

$$\begin{tikzcd}
\boldsymbol{X} \arrow[d, "\mathbf{P}"'] \arrow[dr, "\mathcal{C}"] & \\
\boldsymbol{X}_0 \arrow[r, shift left, "\mathcal{C}"] & \boldsymbol{Y} \arrow[l, shift left, "\mathcal{R}"]
\end{tikzcd}$$

29. Historical Notes

This is of course a very subjective area. From the point of view taken in these notes functional analysis can be used (practically speaking: jus the theory of Banach and Hilbert spaces is used, together with the main principles of the theory of Banach spaces, operators, completions and the like, such as the w^*-compactness of the closed unit ball in a dual Banach space $(\boldsymbol{B}', \|\cdot\|_{\boldsymbol{B}'})$).

30. Open Questions, things to do, state May 2017

We have started by defining the internal convolution in $(\boldsymbol{M}_b(G), \|\cdot\|_{\boldsymbol{M}_b})$ via the action of bounded measures by convolution (resp. moving averages) on $\big(\boldsymbol{C}_0(G), \|\cdot\|_\infty\big)$, and have shown that this gives an internal, commutative and of course associative (immediately from the setup) "multiplication". But there is also the adjoint action of this bilinear $(\boldsymbol{M}_b(G), \boldsymbol{C}_0(G))$ mapping, which also maps $\boldsymbol{M}_b(G)$ into $\boldsymbol{M}_b(G)$ and one has to show that both these two actions are the same. On the other hand the internal convolution in $(\boldsymbol{M}_b(G), \|\cdot\|_{\boldsymbol{M}_b})$ restricts to the subspace which we call $\big(\boldsymbol{L}^1(G), \|\cdot\|_1\big)$, and thus there is a dual action on the dual space of $\big(\boldsymbol{L}^1(G), \|\cdot\|_1\big)$, which is $\big(\boldsymbol{L}^1(G), \|\cdot\|_1\big)$, which in turn contains $\big(\boldsymbol{C}_0(G), \|\cdot\|_\infty\big)$ as a closed, translation invariant subspace.

Our concern is: are all these possible interpretations compatible?

The answer will be, "YES, of course, and the argument will essentially rely on continuity considerations.

Later on similar questions can be raised for general homogeneous Banach spaces and their duals.

Let us write $*'$ for the adjoint action, thus

$$\mu *' \nu(f) := \nu(\mu * f), \quad \mu, \nu \in \boldsymbol{M}_b(G), f \in \boldsymbol{C}_0(G).$$

We will verify that $\mu *' \nu = \mu^{\checkmark} * \nu$.

Proof: It will be enough to verify that this claim is valid for the point measures, because in both sides of the claimed identity the output depends continuously (in the w^*-sense) on the input. So let us assume that $\mu = \delta_x$ and $\nu = \delta_y$.

Then

$$\delta_{-x} *' \delta_y(f) \overbrace{adj.action} = \delta_{-x}(C_{\delta_y}(f)) = \delta_{-x}(T_y f) = f(-x-y) = \delta_{-x-y}(f).$$

OLD MATERIAL in words: The argument is essentially of the following nature. First we verify that $\boldsymbol{M} * \boldsymbol{L}^1(\mathbb{R}^d) \subset \boldsymbol{L}^1(\mathbb{R}^d)$ (either because $\boldsymbol{L}^1(\mathbb{R}^d)$ is the set of bounded measures with continuous shifts, which is preserved under convolution given as the dual action from the action on $\boldsymbol{C}_0(\mathbb{R}^d)$, lifted to the measures, or just by using the fact that $\boldsymbol{L}^1(\mathbb{R}^d)$ is a homogeneous Banach space, hence $\boldsymbol{M}(\mathbb{R}^d) \bullet \boldsymbol{L}^1(\mathbb{R}^d) \subset \boldsymbol{L}^1(\mathbb{R}^d)$, where the abstract action is just convolution. This action is again by a second adjointness relation lifted to $\big(\boldsymbol{L}^1, \|\cdot\|_1\big)$ (which is just the natural continuation of the convolution of ordinary functions to the *dual space* of $\big(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\big)$). The elements satisfying

$$\lim_{x \to 0} \|T_x h - h\|_\infty = 0$$

can be approximated in the norm by elements of the form $g * h \in \boldsymbol{L}^1(\mathbb{R}^d) * \boldsymbol{L}^\infty(\mathbb{R}^d) \subset \boldsymbol{C}_{ub}(\mathbb{R}^d)$. Since this is a closed subspace of $\boldsymbol{L}^\infty(\mathbb{R}^d)$ (cf. lemma below) the limit, i.e. h must belong (more precisely: must be represented) by a function in $\boldsymbol{C}_{ub}(\mathbb{R}^d)$.

CbintoLinf1 **Lemma 90.** *The embedding $h \to \sigma_h$: given by*

CbtoLinf (328)
$$\sigma_h(f) = \int_{\mathbb{R}^d} f(x)\, h(x) dx$$

from $\big(\boldsymbol{C}_b(\mathbb{R}^d), \|\cdot\|_\infty\big)$ into the dual space of $\big(\boldsymbol{L}^1, \|\cdot\|_1\big)$ (with its natural norm) is isometric:

CbLinfisom (329)
$$\|h\|_\infty = \sup_{\|g\|_1 \le 1} \left| \int_{\mathbb{R}^d} g(x) h(x) dx \right|.$$

For non-negative functions this is an easy argument ef_abs_cont **Definition 66** (absolutely continuous, [59], 18.10)**.** Let f be a complex-valued function defined on a subinterval J of $\mathbb{R}$. Suppose that for every $\epsilon > 0$, there is a $\delta > 0$ such that

$$(i) \quad \sum_{k=1}^{n} |f(d_k) - f(c_k)| < \epsilon$$

for every finite, pairwise disjoint, family $\{]c_k, d_k[\}_{k=1}^n$ of open subintervals of J for which

$$(ii) \quad \sum_{k=1}^{n} (d_k - c_k) < \delta.$$

Then f is said to be absolutely continuous on J.

e_variation **Theorem 48** ([59], 18.12)**.** *Any complex-valued absolutely continuous function f defined on $[a, b]$ has finite variation on $[a, b]$*

_meas_deriv **Theorem 49** ([59], 18.14)**.** *If f is a real-valued, nondecreasing function on $[a, b]$, then f' is Lebesgue measurable and*

$$\int_a^b f'(x)dx \leq f(b) - f(a).$$

If g is a complex-valued function of finite variation on $[a, b]$, then $g' \in \boldsymbol{L}^1([a, b])$.

_derivative **Theorem 50** ([59], 18.15)**.** *Let f be an absolutely continuous complex-valued function on $[a, b]$ and suppose that $f'(x) = 0$ almost everywhere in $]a, b[$. Then f is a constant.*

al_calculus **Theorem 51** (Fundamental theorem of the integral calculus for Lebesgue integrals, [59], 18.16)**.** *Let f be a complex-valued, absolutely continuous function on $[a, b]$. Then $f' \in \boldsymbol{L}^1([a, b])$ and*

tal_theorem
$$f(x) = f(a) + \int_a^x f'(t)dt \tag{330}$$

for every $x \in [a, b]$. The converse is true as well.

ef_integral **Theorem 52** ([59], 18.17)**.** *A function f on $[a, b]$ has the form*

$$f(x) = f(a) + \int_a^x \varphi(t)dt$$

for some $\varphi \in \boldsymbol{L}^1([a, b])$ if and only if f is absolutely continuous on $[a, b]$. In this case we have $\varphi(x) = f'(x)$ almost everywhere on $]a, b[$.

comments on absolutely continuous functions, classically

integral_v1 **Theorem 53** ([59], 18.18)**.** *A function f on $\mathbb{R}$ has the form*

$$f(x) = \int_{-\infty}^x \varphi(t)dt$$

for some $\varphi \in \boldsymbol{L}^1(\mathbb{R})$ if and only if f is absolutely continuous on $[-A, A]$ for all $A > 0$, $V_{-\infty}^{\infty} f$ is finite, and $\lim_{x \to -\infty} f(x) = 0$.

hm_part_int **Theorem 54** ([59], 18.19). *Let f, g be functions in $\boldsymbol{L}^1([a,b])$, let*

$$F(x) = \alpha + \int\limits_a^x f(t)dt,$$

and let

$$G(x) = \beta + \int\limits_a^x g(t)dt.$$

Then

$$\int\limits_a^b G(t)f(t)dt + \int\limits_a^b g(t)F(t)dt = F(b)G(b) - F(a)G(a).$$

or_part_int **Corollary 11** ([59], 18.20). *Let f and g be absolutely continuous functions on $[a,b]$. Then*

$$\int\limits_a^b f(t)g'(t)dt + \int_a^b f'(t)g(t)dt = f(b)g(b) - f(a)g(a)$$

splin-abs1
$$||\mathrm{Sp}_\Psi(f)| - |f||(x) \leq |(\mathrm{Sp}_\Psi(f) - f)(x)| \to 0 \quad \text{for} \quad |\Psi| \to 0. \tag{331}$$

It is of course easy to find some function $h \in \boldsymbol{C}_c(\mathbb{R}^d)$ such that $\|h\|_\infty = 1$ and $h(x_i) = |k(x_i)|/k(x_i)$ at all the points where (x_i) is the family of centers of the BUPU Φ (and the condition is of course only applied when $k(x_i) \neq 0$. Then $\mathrm{Sp}_\Psi(|k|) = \mathrm{Sp}_\Psi(hk)$. Alternative: LOOK at the function (to be embedded into $\boldsymbol{C}_0'(\mathbb{R}^d)$) by considering the set where it is strictly positive and strictly negative, with some margin.

ANOTHER strategy (comparable with this one) is to show that $|k|$ can be norm-approximated (uniformly, with fixed compact support) by functions of the form $h \cdot k$, where h is a continuous version of the $sign(k(x))$-function. (can we do it using find BUPUS?, or does one need Tietze-Urysohn?).

LEFT OVER MATERIAL

Kantorovich
Theorem 55. `[Kantorowich Lemma]` *Let T_α be a strongly convergent and bounded sequence of invertible operators between Banach spaces, with limit T_0, which is assumed to be invertible itself. Then the inverse operators are strongly convergent as well if the inverse operators are uniformly bounded. If we consider only sequences this is a criterion, because then the strong convergence of $T_n^{-1}(y) \to T_0^{-1}(y)$ for every y implies uniform boundedness of the sequence T_n^{-1}.*

QUESTION: For tight nets of bounded measures w^*-convergence implies pointwise and uniform over compact convergence of their FTs. But also the converse is true!!!! (Exercise). [for non-tight families this is not true, just think of the case $\delta_n, n \to \infty$!

An elementary proof showing that the Gauss function $g_0(t) = e^{-\pi|t|^2}$ is mapped itself by the Fourier transform has been given by Georg Zimmermann, see

`http://www.univie.ac.at/NuHAG/FEICOURS/ws0607/efoft.pdf`

31. APPENDIX

Definition 67. A net $\{f_\alpha\}_{\alpha\in A}$ in a Banach space $\mathscr{X}$ is said to be a Cauchy net if for every $\epsilon > 0$, there is a α_0 in A such that $\alpha_1, \alpha_2 \geq \alpha_0$ implies $||f_{\alpha_1} - f_{\alpha_2}|| < \epsilon$.

Proposition 7. *In a Banach space each Cauchy net is convergent.*

Proof. Let $\{f_\alpha\}_{\alpha\in A}$ be a Cauchy net in the Banach space $\mathscr{X}$. Choose α_1 such that $\alpha \geq \alpha_1$ implies $||f_\alpha - f_{\alpha_1}|| < 1$. Having chosen $\{\alpha_k\}_{k=1}^n$ in A, choose $\alpha_{n+1} \geq \alpha_n$ such that $\alpha \geq \alpha_{n+1}$ implies

$$||f_\alpha - f_{\alpha_{n+1}}|| < \frac{1}{n+1}.$$

The sequence $\{f_{\alpha_n}\}_{n=1}^\infty$ is clearly Cauchy and, since $\mathscr{H}$ is complete, there exists f in $\mathscr{H}$ such that $\lim_{n\to\infty} f_{\alpha_n} = f$.
It remains to prove that $\lim_{\alpha\in A} f_\alpha = f$. Given $\epsilon > 0$, choose n such that $\frac{1}{n} < \frac{\epsilon}{2}$ and $\|f_{\alpha_n} - f\| < \frac{\epsilon}{2}$. Then for $\alpha \geq \alpha_n$ we have

$$||f_\alpha - f|| \leq ||f_\alpha - f_{\alpha_n}|| + ||f_{\alpha_n} - f|| < \frac{1}{n} + \frac{\epsilon}{2} < \epsilon.$$

$\square$

ef_conv_sum **Definition 68.** Let $\{f_\alpha\}_{\alpha\in A}$ be a set of vectors in the Banach space $\mathscr{X}$. Let $\mathscr{F} = \{F \subset A : F \text{ finite}\}$. If we define $F_1 \leq F_2$ for $F_1 \subset F_2$, then $\mathscr{F}$ is a directed set. For each F in $\mathscr{F}$, let $g_F = \sum\limits_{\alpha\in F} f_\alpha$. If the net $\{g_F\}_{F\in\mathscr{F}}$ converges to some g in $\mathscr{H}$, then the sum $\sum\limits_{\alpha\in A} f_\alpha$ is said to converge and we write $g = \sum\limits_{\alpha\in A} f_\alpha$.

Proposition 8. *If $\{f_\alpha\}_{\alpha\in A}$ is a set of vectors in the Banach space $\mathscr{X}$ such that $\sum\limits_{\alpha\in A} ||f_\alpha||$ converges in the real line $\mathbb{R}$, then $\sum\limits_{\alpha\in A} f_\alpha$ converges in $\mathscr{X}$ (abs. convergence implies convergence!)*

Proof. It suffices to show, in the notation of Definition 68, that the net $\{g_F\}_{F\in\mathscr{F}}$ is Cauchy. Since $\sum\limits_{\alpha\in A} ||f_\alpha||$ converges, for $\epsilon > 0$, there exists F_0 in $\mathscr{F}$ such that $F \geq F_0$ implies

$$\sum_{\alpha\in F} ||f_\alpha|| - \sum_{\alpha\in F_0} ||f_\alpha|| < \epsilon.$$

Thus for $F_1, F_2 \geq F_0$ we have

$$\begin{aligned}
||g_{F_1} - g_{F_2}|| &= \left\| \sum_{\alpha\in F_1} f_\alpha - \sum_{\alpha\in F_2} f_\alpha \right\| \\
&= \left\| \sum_{\alpha\in F_1\setminus F_2} f_\alpha - \sum_{\alpha\in F_2\setminus F_1} f_\alpha \right\| \\
&\leq \sum_{\alpha\in F_1\setminus F_2} ||f_\alpha|| + \sum_{\alpha\in F_2\setminus F_1} ||f_\alpha|| \\
&\leq \sum_{\alpha\in F_1\cup F_2} ||f_\alpha|| - \sum_{\alpha\in F_0} ||f_\alpha|| < \epsilon.
\end{aligned}$$

Therefore, $\{g_F\}_{F\in\mathscr{F}}$ is a Cauchy net and $\sum\limits_{\alpha\in A} f_\alpha$ converges by definition. $\square$

charBan **Corollary 12.** *A normed linear space $\mathscr{X}$ is a Banach space if and only if for every sequence $\{f_n\}_{n=1}^{\infty}$ of vectors in $\mathscr{X}$ the condition $\sum_{n=1}^{\infty} ||f_n|| < \infty$ implies the convergence of $\sum_{n=1}^{\infty} f_n$.*

Proof. If $\mathscr{X}$ is a Banach space, then the conclusion follows from the preceding proposition. Therefore, assume that $\{g_n\}_{n=1}^{\infty}$ is a Cauchy sequence in a normed linear space $\mathscr{X}$ in which the series hypothesis is valid. Then we may choose a subsequence $\{g_{n_k}\}_{k=1}^{\infty}$ such that $\sum_{k=1}^{\infty} ||g_{n_{k+1}} - g_{n_k}|| < \infty$ as follows: Choose n_1 such that for $i, j \geq n_1$ we have $||g_i - g_j|| < 1$; having chosen $\{n_k\}_{k=1}^{N}$ choose $n_{N+1} > n_N$ such that $i, j > n_{N+1}$ implies $||g_i - g_j|| < 2^{-N}$. If we set $f_k = g_{n_k} - g_{n_{k-1}}$ for $k > 1$ and $f_1 = g_{n_1}$, then $\sum_{k=1}^{\infty} ||f_k|| < \infty$, and the hypothesis implies that the series $\sum_{k=1}^{\infty} f_k$ converges. It follows from the definition of convergence that the sequence $\{g_{n_k}\}_{k=1}^{\infty}$ converges in $\mathscr{X}$ and hence so also does $\{g_n\}_{n=1}^{\infty}$. Thus $\mathscr{X}$ is complete and hence a Banach space. $\square$

32. Fourier Analysis over finite groups approximating LCA groups

On the on hand there is the following *structure theorem* for LCA (*locally compact Abelian*) groups G. Probably details can be found in Hewitt-Ross ([58]).

Astructure1 **Theorem 56.** *Every LCA group can be approximated by elementary LCA groups, which are of the form* $\mathbb{Z}^k \times \mathbb{R}^d \times \cdots$. *More precisely*

For every LCA group G there exist arbitrary small compact subgroups $K \lhd G$ and open subgroups $H \lhd G$ such that H/K is an elementary group.

This can be used to define the *Schwartz-Bruhat* space (the natural generalization of $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$ to LCA groups, see [14]), a much more convenient characterization is given by M.S. Osborne in [71].

It is heavily used in [82]. D. Poguntke has shown that $\boldsymbol{\mathcal{S}}(G) \subset \boldsymbol{S}_0(G)$ for general LCA groups.

33. Wiener's algebra $W(G)$ over LC groups

algebraSecG

Already the construction of the Haar measure (cf. Cartier's proof the existence of an invariant Haar measure, see [100], or [23]) uses this space implicitly (without mentioning this terminology).

The usual definition of Wiener's algebra (as given e.g. in the book of H. Reiter [79]) involves the use of a lattice, i.e. a discrete and cocompact lattice Λ in the group (typically $\mathbb{Z}^d \lhd \mathbb{R}^d$, with $\mathbb{R}^d = \bigcup_{k \in \mathbb{Z}^d} k + Q$ for some (relatively) compact set $Q \subset \mathbb{R}^d$). But it is possible to describe it equivalently in a different way (in fact without using BUPUs). To justify the general definition we first show that it gives the same space in the well-known situation:

WienerCharG **Lemma 91.** *Given any non-zero, non-negative (bump-) function $k_0 \in \boldsymbol{C}_c(\mathbb{R}^d)$ one has: $f \in \boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(\mathbb{R}^d)$ if and only if there exists a sequence $(x_n)_{n \geq 1}$ in $\mathbb{R}^d$ and a sequence of complex numbers $(c_n)_{n \geq 1} \in \boldsymbol{\ell}^1(\mathbb{N})$ such that*

WienerCsum1

$$|f(x)| \leq \sum_{n \geq 1} c_n T_{x_n} k_0(x), \forall x \in \mathbb{R}^d. \tag{332}$$

34. Segal algebras and the Ideal Theorem

Segal-idthm **Theorem 57.** *For any Segal algebra $(\boldsymbol{S}, \|\cdot\|_{\boldsymbol{S}})$ over a LCA group G the mappings*

ideal-map

$$I \mapsto I_{\boldsymbol{S}} := I \cap \boldsymbol{S} \qquad I_{\boldsymbol{S}} \mapsto (I_{\boldsymbol{S}})^{\boldsymbol{L}^1} \tag{333}$$

are inverse to each other, i.e.they establish a bijection between the closed ideals $I \subseteq \boldsymbol{L}^1(G)$ and the closed ideals $I_{\boldsymbol{S}}$ of $\boldsymbol{S}$.

Segal algebras (or as I call them *Reiter's Segal algebras*) have been introduced by Hans Reiter in order to study questions of spectral analysis in a unified way. Due the ideal theorem above one can study the existence of sets of spectral synthesis as via Segal algebras, see e.g. [79, 81] or his Lecture Notes on Segal algebras ([80]).

Recall that for every ideal there is a *cospectrum*, essentially the complement of all spectra of elements from the closed ideal, better directly described as the (closed) set of common zeros (as the intersection of a collection of closed sets it is closed itself).

Definition 69. Given a closed ideal in $\left(\boldsymbol{L}^1(G), \|\cdot\|_1\right)$ of some Segal algebra $(\boldsymbol{S}, \|\cdot\|_{\boldsymbol{S}})$ its *cospectrum* is defined by

$$\operatorname{cosp}(I) := \{s \in \widehat{G}, \widehat{f}(s) = 0 \ \forall f \in I\}.$$

However, it may happen that an ideal is not completely characterized by its cospectrum ($\mathrm{cosp}(I)$), let us call it E. In this (very special) situation the minimal closed ideal, which can be shown to be the closure of the subspace of all $\boldsymbol{L}^1(\mathbb{R}^d)$-functions having compact support disjoint from E (since they have compact support they also belong to any Segal algebra $\boldsymbol{S}(G)$), and the maximal ideal, which consists simply of all $\boldsymbol{L}^1(G)$-functions vanishing on the set E, will NOT coincide. In this case E is said to be a closed set for which *spectral synthesis fails*, or a set of non-synthesis.

For a long time the existence of such sets of non-spectral synthesis was open, until it was solved by P. Malliavin ([69]). He was able to show that the surface of the unit-ball in $\mathbb{R}^3$, although a smooth set, is not of spectral synthesis. Later on it was shown that essentially in all non-trivial cases sets of non-synthesis exist within LCA groups.

In it's dual form it is even more interesting. Since the orthogonal complement of the maximal ideal mention above is the set of all point evaluations of $\hat{f}$ on $\hat{G}\setminus E$ spectral analysis *is possible* if for every $\sigma \in \boldsymbol{S}_0'(G)$ with $\mathrm{supp}(\sigma) \subseteq E$ on can find a net of discrete measures, **also supported on the !same! set** E, which is w^*-convergent to the given measure.

See also H. Reiter's book: [79] (or the new version [81]) for this subject (Chap. 7!?).

`http://en.wikipedia.org/wiki/Irving_Segal`

35. Further reading and articles by the author

Here we point out articles which have a survey character and have been written in recent years.

E.g. the article "Choosing Function Spaces in Harmonic Analysis" ([37]). It tries to discuss the standard approach to Fourier Analysis from a critical view-point, emphasizing that not only does the community of analysts dispose currently over a huge collection of function spaces, but also should start asking itself, which spaces are good for which purpose. So while on the one hand giving a long list of construction principles (certainly not an exhaustive one) the article also encourages the discussion of the "information content" of a function space. Of course there will be more examples in the (near?) future and concrete cases, aside from the Banach Gelfand Triple $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$.

There is another article in a similar spirit ([49]) which reinforces the concept of "Conceptual Harmonic Analysis" as the meeting point of application oriented and abstract harmonic analysis. It discusses questions such as "How does the finite Fourier analysis" (e.g. based on the FFT available within MATLAB or other mathematical software packages) allow us to "approximate the continuous reality". Which function spaces and which algorithms do we need in order to come up with good solutions to the problems arising in applications? How can we help engineers to use the proper form of mathematics (and is the currently practised form suitable for this purpose: one may have doubts about this!)?

There is also the hint to the NuHAG talk-server, where a long list of talks on various subjects can be found.

`http://www.univie.ac.at/nuhag-php/nuhag_talks/`
`http://www.univie.ac.at/nuhag-php/home/skripten.php`

36. Related efforts

There are various papers and notes which try to help engineers and physicists to find an easier approach to distribution theory.

An early example is (!to be checked) an Appendix to the early book of A. Papoulis on the *Fourier Transforms and its Applications*, [72],

37. The usual theory of Lp-spaces

In the present description **on purpose** the profile of the usual Lebesgue spaces, even of the certainly important spaces $\boldsymbol{L}^1(\mathbb{R}^d), \boldsymbol{L}^2(\mathbb{R}^d)$ or $\boldsymbol{L}^\infty(\mathbb{R}^d)$, has been kept low. In fact, we are convinced that the corresponding *norms* are highly relevant for *any* discussion of aspects of Fourier analysis, while the concrete realization is of much lower importance, essentially due to the fact that any two (metric) completions of a given normed space (e.g. of $\boldsymbol{C}_c(G)$ with the norm inherited from the natural embedding of $\boldsymbol{C}_c(G)$ into $\boldsymbol{M}_b(G)$ via the Haar measure) are isometrically isomorphic anyway, so that claims verified in one setting (and the corresponding technology) is valid for any other realization in exactly the same way.[96]

So we emphasize here that our aim was more to show that one can do all the essential parts of Fourier Analysis without using Lebesgue integration (just the existence of the Haar measure is needed!), although admittedly they are quite prominent in all standard text books. Of course we admit, that it is a fact that both for a proper discussion of convolutions as integrals (existing a.e.) and the definition of the Fourier transform, viewed as an *integral transform*, the *Lebesgue integral* is the right integral. But is the Fourier transform an integral transform by definition or by any other principle, or is it just convenient to express it via an integral, if the object to be transformed happens to be an integrable function (this is of course just our viewpoint!).

We would also like to mention a (negative?) side-effect of having this "perfect integral". It is sometimes (at least psychologically) an obstacle to take a wider perspective. It does not help us to recognize that the Fourier transform is mapping which has the property (if extended properly) that it assigns to a pure frequency the corresponding Dirac measure in the Fourier domain, indicating exactly the frequency and the amplitude of this pure frequency (resp. plane wave for $d = 2$, etc.). So the wider and perhaps more natural perspective on the subject is this one (aside from the technical aspect which involves integrals): How can one *describe* the amount of pure frequency in a given signal (*Fourier Analysis*) and can one resynthesize (using the result of the analysis step) the signal or distribution from the information content? What is the most suitable setting for turning such vague wishes into concrete mathematics? What is the level of technicality that one has to go through in order to have both a mathematically correct description of some model, which ideally should be intimately connected with the physical reality to be described. Finally (and this is even the least developed aspect of Fourier Analysis so far) it should be possible to simulate at least parts of on a computer, which of course is only able to handle finite length vectors.

Having found some indication about some mathematical question formulated in the context of time-frequency or Fourier analysis through computation, we come to the question of "how close is that computed result from the ground-truth" (i.e. the corresponding numbers that show up when the continuous problem under investigation would have been solved exactly, e.g. by analytic methods)? Again we come to the question of norms for function spaces (in which norm should one measure the deviation) and again one has to hope that functional analysis provides a suitable collection of Banach spaces which allows to describe in a good sense the error due to the reduction to the finite setting in a given situation.

Let us come back to our approach to $\boldsymbol{L}^1(G)$:

In our context $\left(\boldsymbol{L}^1(\mathbb{R}^d), \|\cdot\|_1\right)$ (resp. $\left(\boldsymbol{L}^1(G), \|\cdot\|_1\right)$) appears as $\boldsymbol{M}_c(G)$, i.e. as a closed subspace of $(\boldsymbol{M}_b(G), \|\cdot\|_{\boldsymbol{M}_b})$, which (once the Haar integral is established as a functional on

[96]The argument is along the idea, that the representation of the real number system, the completion of the field of rationals $\mathbb{Q}$, endowed with the Euclidean distance, is conveniently identified with the infinite decimal expressions, but Dedekind sections of equivalence classes of Cauchy sequences in $\mathbb{Q}$ are equally valid interpretations.

$\big(\boldsymbol{W}(\boldsymbol{C}_0, \boldsymbol{\ell}^1)(G), \|\cdot\|_{\boldsymbol{W}}\big)$) obtained as the closure of $\boldsymbol{C}_c(G)$ (viewed as space of generalized functions, via the identification of $k \in \boldsymbol{C}_c(G)$ with $\sigma_k \in \boldsymbol{M}_b(G)$). In this sense there is also a dual space (which one may call $\big(\boldsymbol{L}^\infty(G), \|\cdot\|_\infty\big)$, but it should be rather identified with a subspace of $\boldsymbol{S}_0'(G)$.

The discussion of the Hilbert space $\big(\boldsymbol{L}^2(G), \|\cdot\|_2\big)$, of course to be identified with the completion of $\boldsymbol{C}_c(G)$, endowed with the usual scalar product arising from the Haar integral (we write $d\mu(x)$, i.e.

scalprodG (334)
$$\langle f, g\rangle := \int_G f(x)\overline{g(x)} d\mu(x).$$

has to be discussed and developed separately, but should be also done within the reality of $\big(\boldsymbol{S}_0'(G), \|\cdot\|_{\boldsymbol{S}_0'}\big)$.

It should be possible to define the $\boldsymbol{L}^p$-spaces as completions of $\boldsymbol{C}_c(G)$ within $\boldsymbol{S}_0'(G)$. In fact, we would not need any Lebesgue integration theory in that form. Certain "objects" would constitute the elements of that *a priori abstract* completion, and the STFT would be extended (of course for $\boldsymbol{S}_0(G)$-windows g only) to these spaces. The details of such an approach have to be written up in the near future (summer 2017, hgfei).

It is also no doubt that Banach spaces, Banach algebras, but also Banach modules and approximate identities provide a set of tools which helps to organize the facts arising in the context of Fourier Analysis. This is way devoted some space to their description in these notes.

38. Further general functional analytic considerations

this is to be treated like an APPENDIX, collecting the needed material from general functional analysis, as found in most books on the subject, and as needed anyway, also for other parts of analysis.

E.g. bounded linear operators, dual operators, adjoint operators between Hilbert spaces, Riesz representation theorem, Stone-von-Neumann theorem, Gelfand theory of commutative C^*-algebras etc.. Such material is collection in my course notes of my functional analysis course (SW 13/14), see for example
`http://www.univie.ac.at/nuhag-php/login/skripten/data/FAws1314Fei.pdf`

39. Critical comment on some "BAD" applied literature

39.1. **Foundations of Signal Processing by Vetterli and Co.** Before going into a critical comments let me point out that this is a great book with a lot of details (**gokove14 : [98]**), which is also found (as PDF file) on the internet.

See also Prandoni/Vetterli **prve08 : [78]** (Signal Processing for Communications). Page 94 talks about the *quirky mathematical entity* called the *Dirac delta functional*; this is defined in an implicit way by the following formula

$$\int_{-\infty}^{\infty} \delta(t-\tau) f(\tau) d\tau = f(\tau),$$

where $f(t)$ is an arbitrary **integrable** (??) function on the real line. In particular

$$\int_{-\infty}^{\infty} delta(t) f(t) dt = f(0).$$

It is understood as the limit of rectangular functions $k\,\mathrm{rect}(kt)$. Later on the *pulse train* is introduced.

39.2. **Introduction to the subject.** The recent article [97] also provides further explicit examples, emphasizing the fact that it is NOT true that any BIBOS system has to come from an integrable function, i.e. our measure μ representing $T(f) = \mu * f$, for $f \in \boldsymbol{C}_0(\mathbb{R}^d)$ does not have to belong to the subspace $\boldsymbol{L}^1(\mathbb{R}^d) \subset \boldsymbol{M}_b(\mathbb{R}^d)$ (with the usual identification), because it can be (for example) a convolution by some discrete bounded measure $\mu \in \boldsymbol{M}_d(\mathbb{R}^d)$.

At least from a purely mathematical point various books for engineers provide explanations at an acceptable or sometimes of a questionable level.

The current notes are an attempt to provide ways to explain the situation faced by applied sciences, e.g. when they have to do signal processing and make somehow use of the Fourier transform (using the FFT) a bit more transparent, also by providing alternative views, establishing connections with ideas from linear algebra and a backup from functional analysis, i.e.the theory of infinite dimensional signal spaces (arising naturally one we deal with continuous variables or non-periodic functions).

Among the classical books in the field one can mention the books of Athanasios Papoulis, notably [72, 73] and [74]. Especially Appendix I there, describing the *impulse function* as a *distribution* is mathematically correct but still connecting well with the approach taken by applied people.

On page 275 (file page 284) he also correctly mentions that $\delta^2(t)$ is undefined, however convolution can be given, and due to the usual symbols he has to write the relation

$$\delta_{y_1} * \delta_{y_2} = \delta_{y_1+y_2}$$

which we have used a lot at the beginning of this not as a *quasi-mysterious* integral equation of the form

$$\delta(t-t_1) * \delta(t-t_2) = \int_{-\infty}^{\infty} \delta(\tau - t_1) * \delta(t-\tau-t_2) = \delta[t-(t_1+t_2)] \tag{335}$$

The Review MR0142981 (26 #548) J. L. Griffith starts with the following statement: *The basic theorems of the Fourier integral, as they appear in standard text-books, are expressed in terms of the Lebesgue integral and are thus unintelligible to a large number of engineering students. On the other hand, engineering and applied mathematics have created a vast number of formal results which, though very useful, are completely unknown to the pure mathematician. The author has made an admirable attempt to link these two sections together. ....*

39.3. **Pitfalls and Artifacts.** The book [3], Mastering the Discrete Fourier Transform in One, Two or Several Dimensions. Provides some problems.

David Voelz has published [99] (vo11): Computational Fourier Optics. A MATLAB tutorial.

39.4. **Comments on the book by Kanwal on Distributions.**

`Material already inserted as it is here into TBL 620 (cf. 640), TROUBLE!`

The book [62] by Kanwal (a physicist) was published in 2004:

page 10: *Example 3.* says the following (citation):

Various books in physics leave the reader with the impression that $\delta(x) = +\infty$ at $x = 0$. The following sequence illustrates that this is **not always true:** *.... This sequence is shown in Fig. 1.8, (p.11), which shows that* $\lim_m s_m(x) = -\infty$, *although it yields the unit positive charge for large m; indeed, it is a delta-sequence, as will now be proved. [The proof actually takes !2 pages].*

In addition it suggest wrongly that one should expect to have $|\delta(0)| = +\infty$, which is non-sense. The point evaluation, well defined on $\boldsymbol{C}_b(\mathbb{R}^d)$ (may be viewed as subspace of $\boldsymbol{S}_0'(\mathbb{R}^d)$) to all of $\boldsymbol{M}_b(\mathbb{R}^d)$ (so that it might be meaningful for $\delta \in \boldsymbol{M}_b(\mathbb{R}^d)$).

As with Example 4. one just has to use the stretching operator!

p.14: The **sifting property**,

`sift02` (336)
$$\int_{-\infty}^{\infty} f(x)\delta(x-\xi)dx = f(\xi),$$

is interpreted as the action of the generalized function $\delta(x-\xi)$ on $f(x)$: that is, when $\delta(x-\xi)$ acts on a suitably smooth function $f(x)$, it sifts out the value $f(\xi)$ at $x = \xi$. Since $\delta(x) = \delta(-x)$, or $\delta^{\checkmark} = \delta$, this is the same as $\delta_x(f) = f(x)$.

p.28: *Alternative terminology for*

Dirac comb: The *sampling function* or *replicating function* because

$\sqcup\!\sqcup(x)f(x) = \sum_{n=-\infty}^{\infty} f(x)\delta(x-n)$, but should in fact be $\sqcup\!\sqcup(x)f(x) = \sum_{n=-\infty}^{\infty} \mathbf{f}(\mathbf{n})\delta(x-n)$.

see also the comment on the book of Debnath

39.5. **Debnath's article on the genesis of convolution.** Note that Zygmund's books almost does not contain any relevant hint to convolution, while convolution and the Banach algebra $\big(\boldsymbol{L}^1(G), \|\cdot\|_1\big)$ (convolution) is seen as the central object of Harmonic Analysis in Reiter's book ([79]).

This are some personal comments by HGFei (31.05.2017) concerning the article [21] which appeared in ANALYSIS in 2012.

The manuscript is more written for engineers than for mathematicians. For example it uses for the description of the Fourier transform this formula

`DebFT1` (337)
$$\mathcal{F}\{f(x)\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx} f(x)dx, \quad k \in \mathbb{R}.$$

Convolution is defined (3.11), p.48 as

`DebConv1` (338)
$$(f * g)(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x-\xi)g(\xi)d\xi, \quad x \in \mathbb{R}.$$

Using mostly Fubini's Theorem various properties, including the convolution theorem and the dual convolution theorem (pointwise multiplication goes to convolution, if the convolution side has two integrable functions!), also *associativity* of convolution (p.51). Explicitly this is presented in [98], p.404.. JJJJJ

PROBLEM: With this definition the product of two probability measures is not a probability measure anymore!! PLEASE CHECK! There IS still a convolution theorem, see page 49 of [21], Most likely the motivation for the convention (337) is the symmetry of forward and inverse Fourier-Plancherel transform despite the (not recommended!) convention of using only e^{ikx} in the integral instead of $e^{2\pi ikx}$! is to "adapt the convolution definition" in order to have a convolution theorem. But this destroys more than it helps to get out of the problem (*private comment by HGFei*, Jan. 2018).

There is also another article on the *Developments of the theory of generalized functions or distributionsA vision of Paul Dirac* in [22].

39.5.1. *More serious issues.*

(1) In the definition of $\boldsymbol{L}^1(\mathbb{R})$ (p.46) no *measurability* assumption is made, just integrability of $|f(x)|$;

(2) **p.52: In order to prove the Fundamental Relation for the Fourier transform, here called in Thm.3.3 the General Parseval Relation, a formal step is taken.** The identity

xpintdelta1

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{i(k-l)x}dx = \delta(k-l) \tag{339}$$

This is later explained (p.63) also via the integral representation of the Dirac delta **FUNCTION (!)**

DebDirac

$$\delta(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{i(k-l)x}dx \tag{340}$$

Intuitively it would be OK to think in this way, because of course the FT of δ_0 is just the constant 1.

(3) The estimate $\|f*g\|_1 \leq \|f\|_1\|g\|_1$ is explained as the fact that $(f*g)(x)$ is at least as smooth as the smoother of the two functions f and g. Derivatives are only discussed later!

(4) Reiter's fundamental relation is only proved afterwards (item 9, p.55)

(5) p.63 Formula (340) is also claimed to provide a *simple proof (?)* of the Fourier inverse transform!

(6) Another strange argument is the proof of (3.54), p.58: It makes use of an interchange of limits and integration (without justification), and in addition the claim:

ebDiracLimG

$$\lim_{t\to 0+} G_t(x-\xi) = \delta(x-\xi). \tag{341}$$

Here the Gauss kernel (3.53) is used, in (3.56) later on the Abel kernel is used, with a similar argument.

40. Comments on mathematical books

40.1. **Eric Stade: Fourier Analysis.** In section 6 of **st05-3: [92]** the Fourier transform pair is first introduced "informally". Starting form P-periodic functions the author asks: Let's now consider, quite informally for the moment, what happens to (6.2) in the *limit* $P \to \infty$. What does happen? ... pointing also the analogy to orthogonal decomposition (p.298), but with the difficulty that e_s (our χ_s) is NOT in $\boldsymbol{L}^2(\mathbb{R}^d)$.

Example 6.1.1 demonstrates what the Fourier transform is for $\mathbf{1}_{[-b,b]}$, namely $\sin(2\pi bs)/\pi s$.

Example 1.6.2 also does the computation for $f(x) = e^{-2\pi|x|}$, namely $1/(1+\pi x^2)$.

Then he goes on to derive Plancherel's Theorem and the Convolution Theorem (for $\boldsymbol{L}^1(\mathbb{R})$).

The demonstration that $\widehat{\delta} = 1$ (formula (6.45)) is commented as follows: *One could have a field day finding holes in the arguments of the above example (he is taking limits there without justification..). But again, it's all in fun (!).*

Specifically Lemma 6.2.1 (.307) demonstrates that

$$\int_{-\infty}^{\infty} e^{-\pi x^2} dx = 1.$$

(i.e. that the Gauss function defines a continuous probability distribution). As usual he makes use of the trick of doing it in the 2D-case, using polar coordinates in $\mathbb{R}^2$.

Examples 6.2.3: The Fourier transform of the Gauss function is the Gauss function itself:

$$G(x) = \widehat{G}(x) = e^{-\pi x^2}.$$

CITE: This is a very cool property of G.

Section 6.4. goes on to prove the Fourier Inversion Theorem (p.316):

The first, *intuitive* proof is commented as: *Well no, not really. Our "proof" isn't a proof; it involves* a number of illegitimate steps.

Thm.6.4.2.: $\mathcal{F}(\mathcal{F}(f)) = f^{\checkmark}$.

This is followed by the $\boldsymbol{L}^2(\mathbb{R})$-theory (Plancherel's theorem).

The whole book is mostly restricted to the 1D-case!

The multi-dimensional case is only treated in Sec.6.9., followed by a section on tempered distributions.

doubnetlim **Lemma 92.** *A statement about iterated bounded, strongly convergent nets of operators.*

Assume that two bounded nets of operators between normed spaces, $(T_\alpha)_{\alpha\in I}$ and $(S_\beta)_{\beta\in J}$ are strongly convergent to some limit operators T_0 and S_0 respectively. Then the iterated limit of **any order** *exists and the two limits are equal.*

In fact, the index set $(\alpha,\beta)\in I\times J$ with the natural order [97] *is turning $(S_\beta\circ T_\alpha)$ into a strongly convergent net.*

Proof. . . . Without loss of generality we may assume $T_0 = 0$ and $S_0 = 0$ (otherwise treat $T_\alpha - T_0$ etc.). THE REST OF THE PROOF is LEFT TO THE READER. HINT: Cf. Proposition 6. □

41. Appendix on Functional Analysis

Various terms which are not explained in detail here are given my be found in any of the standard text books on functional analysis, e.g. in Conway's book [18] or probably Rudin's book [88].

We consider these books as interesting further source concerning basic facts of functional analysis:
[Bowers, Adam;Kalton, Nigel, J.] An Introductory Course in Functional Analysis [11], also with respect to applications: [Yamamoto, Yutaka] From Vector Spaces to Function Spaces: Introduction to Functional Analysis with Applications: [105], [17] about a Unified approach to Signal Processing;

German books are: [Werner, Dirk] Functional Analysis. (Funktionalanalysis) 5., Erweiterte Aufl. [104] or [Alt, Hans Wilhelm] Lineare Funktionalanalysis. Eine Anwendungsorientierte Einfhrung. (Linear Functional Analysis. An Application Oriented Introduction). [1]

We assume that the reader of this note is familiar with the foundations of linear functional analysis, i.e. with normed spaces (linear spaces endowed with a *norm*, which allows to measure in some sense the *size of an element of this spaces*, often called an (abstract) *vector* or *signal*.) Banach spaces are those normed spaces which are complete with respect to the *metric* induced by the norm: Any *Cauchy-sequence* has a limit, or equivalently: every absolutely convergent series has a limit, meaning this: Whenever $\sum_{n=1}^{\infty}\|\mathbf{b}_n\|_{\boldsymbol{B}}<\infty$ the sequence of partial sums will have a limit, which will be denoted by

$$\sum_{n=1}^{\infty}\mathbf{b}_n.$$

We have the notation of *equivalence of norms*.

normequ1 **Definition 70.** *Formulation* Two norms are equivalent, whenever the $\boldsymbol{B}^1 = \boldsymbol{B} = \boldsymbol{B}^2$ the identity map establishes a norm-isomorphism (not necessarily isometric) between $(\boldsymbol{B}^1, \|\cdot\|^{(1)})$ and $(\boldsymbol{B}^2, \|\cdot\|^{(2)})$, or equivalently: There exists $C>0$ such that

normequ2
$$C^{-1}\|\mathbf{b}\|^{(1)} \le \|\mathbf{b}\|^{(2)} \le C\|\mathbf{b}\|^{(1)}, \quad \forall \mathbf{b}\in\boldsymbol{B}. \tag{342}$$

This means of course that convergence with respect to one norm applies if and only if applies to the other norm. We have the same convergent sequences, hence the same collection of closed or open subsets, the same closed or compact sets, in short, we have the same topology!

[97] explanation: the natural ordering is given by the fact that $(\alpha,\beta)\succeq(\alpha_0,\beta_0)$ if $\alpha\succeq\alpha_0$ and $\beta\succeq\beta_0$.

41.0.1. *Adjoint operators.* For any bounded operator $T : \boldsymbol{B}^1 \to \boldsymbol{B}^2$ we have an adjoint operator $T' : \boldsymbol{B}_1' \to \boldsymbol{B}_2'$, given by $T'(\mathbf{y}')(\mathbf{x}) = \mathbf{y}'(T\mathbf{x})$, $\mathbf{x} \in \boldsymbol{B}^1, \mathbf{y}' \in \boldsymbol{B}_2'$. This operator has the same operator norm as the original operator, i.e. $|||T||| = |||T'|||$, or more precisely

$$|||T|||_{\boldsymbol{B}^1 \to \boldsymbol{B}^2} = |||T'|||_{\boldsymbol{B}_2' \to \boldsymbol{B}_1'} .$$

There is also another, quite similar convention making use of the scalar product in an abstract (or of course concrete) Hilbert space:

It can be shown that $T^{**} = T$ and therefore $|||T||| = |||T^*|||$. The estimate $|||T^*||| \leq |||T|||$ is an easy exercise, but then

$$|||T||| \leq |||T^{**}||| \leq |||T^*||| \leq |||T||| .$$

Next we want to see what happens for the case that $\boldsymbol{B}^1 = \mathcal{H}_1 = \mathbb{C}^n$ and $\boldsymbol{B}^2 = \mathcal{H}_2 = \mathbb{C}^m$, i.e.we are looking at the familiar situation that we have a standard matrix representation for all the operators involved, either as an $m \times n$-matrix $\boldsymbol{A}$ or as some $m \times n$-matrix $\boldsymbol{B}$, depending on the dimension of domain and target space.

42. Appendix on Group Theory

We need only very little group theory, more or less the basic definitions. Our prototypes of groups will be the real line $\mathbb{R}$ with addition, or the unit roots of order N, denoted by $\mathbb{Z}_N$ or $\mathbb{U}_N$ (with multiplication of complex numbers).

We also have various groups if operators. The prototype being the group $\mathrm{GL}_n(\mathbb{R})$ of all non-singular matrices $n \times n$-matrices (i.e. with $\det \boldsymbol{A} \neq 0$), or $\mathrm{SL}(n, \mathbb{R})$, the special linear group of all $n \times n$-matrices with $\det(\boldsymbol{A}) = 1$.

In each case we have (this is the list of axioms one has to check) a composition on some (a priori abstract set) G, which assigns to each pair $(x, y), x, y \in G$ a new element, usually written as $x \cdot y$ or just $xy \in G$.

This mapping is supposed to satisfy the *associativity law*

assoccG
$$(x \cdot y) \cdot z = x \cdot (y \cdot z), \quad x, y, z \in G. \tag{343}$$

With the validity of (343) we have a *semigroup*. It is called *commutative* (or often *Abelian*) if one also has

commutcG
$$x \cdot y = y \cdot x, \quad \forall x, y \in G. \tag{344}$$

Any group has to have a *neutral* or *zero* or *unit* element $\mathbf{e}$ (which can be shown to be, once it exists, to be uniquely determined, based on its properties, satisfying

unitelemcG
$$\mathbf{e} \cdot x = x \cdot x, \quad \forall x \in G. \tag{345}$$

Moreover (finally) every element $x \in G$ must have an inverse element $z \in G$ (also uniquely determined for algebraic reasons, and then denoted by $-x$ for additively written groups, or x^{-1} or $1/x$ in the multiplicative setting) such that

inversecG
$$z \cdot x = \mathbf{e} = x \cdot z. \tag{346}$$

Remark 66. For commutative groups often the symbol $+$ is preferred, (resp. whenever a group is written additively it can safely be assumed that it is meant to be a commutative group), while the use of a multiplicative symbol may or may not indicate the lack of commutativity. For the space $\mathcal{M}_{n,n}(\mathbb{R})$ of all $n \times n$-matrices one can view it as a vector space which is an *additive group* under coordinate-wise (if you want pixel-wise) addition. Clearly $(\mathcal{M}_{n,n}(\mathbb{R}), +)$ is a commutative group, while $(\mathrm{GL}(n, \mathbb{R}), *)$ or $(\mathrm{SL}(n, \mathbb{R}), *)$ (with matrix multiplication) are non-commutative groups for $n \geq 2$.

Clearly $(\mathbb{U}_N, \cdot)$ is a commutative group with respect to multiplication of complex numbers. The exponential mapping is an isomorphism from $(\mathbb{R}, +)$ to $(\mathbb{R}_+^*, \cdot)$.

omomorphism **Definition 71.** A mapping $\alpha : G_1 \to G_2$ from one group to another is called a (semi-group) *homomorphism* if it preserves the group operations, i.e. if it does not matter whether the composition of group elements is applied before or after the mapping, i.e. if (and only if)

homomdef01 (347) $$\alpha(x \cdot_1 y) = \alpha(x) \cdot_2 \alpha(y), \quad \forall x, y \in G_1$$ [98]

basichomocG **Lemma 93.** *For any group homomorphism one has:*

(1) $\alpha(\mathbf{e}_1) = \mathbf{e}_2$ *(i.e. neutral elements go to neutral elements);*
(2) $\alpha(x^{-1}) = (\alpha(x))^{-1}$ *(i.e. inverse elements go to inverse elements).*

Remark 67. As one shows in linear algebra that the inverse mapping for a bijective linear mapping is linear as well one also shows easily (by purely algebraic arguments) that for any group homomorphism $\alpha : G_1 \to G_2$ also the *inverse mapping* $\alpha^{-1} : G_2 \to G_1$ is a group homomorphism as well.

automorphcG **Definition 72.** A bijective homomorphism $\alpha : G_1 \to G_2$ is called a *group isomorphism.* The corresponding groups are called *isomorphic* (via α).

A group isomorphism $\alpha : G \to G$ is called an *automorphism* of the group G.

Clearly $(\mathbb{U}_N, \cdot)$ is a commutative group with respect to multiplication of complex numbers. The exponential mapping is an isomorphism from $(\mathbb{R}, +)$ to $(\mathbb{R}_+^*, \cdot)$ (multiplication of non-zero real numbers). The inverse mapping (logarithm) helps to turn the (in the old days) cumbersome multiplication or division into addition resp. subtraction.

The group $(\mathbb{U}_N, \cdot)$ is isomorphic to the quotient group $\mathbb{Z}/N\mathbb{Z}$ ($\mathbb{Z}$ modulo N, i.e. addition modulo N).

Any linear mapping $\mathbb{R}^n \to \mathbb{R}^n$ is given as matrix multiplication: $\mathbf{x} \to \boldsymbol{A} * \mathbf{x}$. It is *invertible*, i.e. i.e. an automorphism of $(\mathbb{R}, +)$ if and only if $\det(\boldsymbol{A}) \neq 0$.

For a finite Abelian group the *dual group* is the group of all homomorphisms from $(G, +)$ into $(\mathbb{U}, \cdot)$ (the torus group, or *unit circle*, or $S_1 \subset \mathbb{C}$ (one-dimensional sphere), with complex multiplication). Usually its elements are called *characters.* We write

$$\widehat{G} := \{\chi, \chi : G \to \mathbb{U}...\}.$$

It turns out to be again itself a group, with pointwise multiplication, i.e. with

$$\chi_1 \cdot \chi_2(x) = \chi_1(x) \cdot \chi_2(x), \quad x \in G.$$

This is of course again a commutative group, with neutral element begin the *trivial character* $\chi_0(x) = 0 \forall x \in G$ and inverse element $x \to \overline{\chi(x)} = 1/\chi(x)$.

WE ALSO HAVE TO DISCUSS *topological groups* AND *continuous homomorphisms* AND THEN *isomorphisms of topological groups*

Also the term: *local compactness* has to be discussed.

[98] Here of course the symbols $\cdot_1$ and $\cdot_2$ are used to indicate the composition law in G_1 and G_2 respectively.

References

al99 [1] H. W. Alt. *Lineare Funktionalanalysis. Eine Anwendungsorientierte Einführung. (Linear Functional Analysis. An Application Oriented Introduction).* Springer, Berlin, 1999.

amfupa00 [2] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of Bounded Variation and Free Discontinuity Problems.* Clarendon Press, 2000.

am13 [3] I. Amidror. *Mastering the Discrete Fourier Transform in One, Two or Several Dimensions. Pitfalls and Artifacts*, volume 43. London: Springer, 2013.

anmisi73 [4] P. Antosik, J. Mikusinski, and R. Sikorski. *Theory of Distributions. The Sequential Approach. Reprints.* Elsevier Scientific Publishing Company, 1973.

ba10-2 [5] S. Bannert. Banach-Gelfand Triples and Applications in Time-Frequency Analysis. Master's thesis, University of Vienna, 2010.

be96 [6] J. J. Benedetto. *Harmonic Analysis and Applications.* Stud. Adv. Math. CRC Press, Boca Raton, FL, 1996.

becz09 [7] J. J. Benedetto and W. Czaja. *Integration and Modern Analysis.* Birkhäuser, 2009.

belo76 [8] J. Bergh and J. Löfström. *Interpolation Spaces. An Introduction.* Number 223 in Grundlehren Math. Wiss. Springer, Berlin, 1976.

bj71 [9] G. Björck. Beurling distributions and linear partial differential equations. In *Symposia Mathematica, Vol. VII (Convegno sulle Equazioni Ipoellittiche e Spazi Funzionali, INDAM, Rome, 1970)*, pages 367–379. Academic Press, London, 1971.

bo55-1 [10] S. Bochner. *Harmonic Analysis and the Theory of Probability.* University of California Press, 1955.

boka14-2 [11] A. Bowers and N. Kalton. *An Introductory Course in Functional Analysis.* New York, NY: Springer, 2014.

br75 [12] L. H. Brandenburg. On identifying the maximal ideals in Banach algebras. *J. Math. Anal. Appl.*, 50:489–510, 1975.

brfe83 [13] W. Braun and H. G. Feichtinger. Banach spaces of distributions having two module structures. *J. Funct. Anal.*, 51:174–212, 1983.

br61 [14] F. Bruhat. Distributions sur un groupe localement compact et applications l etude des représentations des groupes p-adiques. *Bull. Soc. Math. France*, 89:43–75, 1961.

busm81 [15] R. C. Busby and H. A. Smith. Product-convolution operators and mixed-norm spaces. *Trans. Amer. Math. Soc.*, 263:309–341, 1981.

bune71 [16] P. L. Butzer and R. J. Nessel. *Fourier Analysis and Approximation. Vol. 1: One-dimensional Theory.* Birkhäuser, Stuttgart, 1971.

ca11 [17] G. Cariolaro. *Unified Signal Theory.* Springer, London, 2011.

co90 [18] J. B. Conway. *A Course in Functional Analysis. 2nd ed.* Springer, New York, 1990.

cofelu08 [19] 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.

cook12 [20] E. Cordero and K. A. Okoudjou. Dilation properties for weighted modulation spaces. *J. Funct. Spaces Appl.*, pages Art. ID 145491, 29, 2012.

de12-5 [21] L. Debnath. The genesis of convolution in Fourier analysis with applications. *Analysis (Munich)*, 32(1):39–66, 2012.

de13-2 [22] L. Debnath. Developments of the theory of generalized functions or distributions—a vision of Paul Dirac. *Analysis (Munich)*, 33(1):57–99, 2013.

de02 [23] A. Deitmar. *A First Course in Harmonic Analysis.* Universitext. Springer, New York, NY, 2002.

do56 [24] Y. Domar. Harmonic analysis based on certain commutative Banach algebras. *Acta Math.*, 96:1–66, 1956.

ed59 [25] R. E. Edwards. The stability of weighted Lebesgue spaces. *Trans. Amer. Math. Soc.*, 93:369–394, 1959.

ey64 [26] P. Eymard. L´algebre de Fourier d´un groupe localement compact. *Bull. Soc. Math. France*, 92:181–236, 1964.

fe79-2 [27] H. G. Feichtinger. English translation of: Gewichtsfunktionen auf lokalkompakten Gruppen (download at www.nuhag.eu/bibtex). *Sitzungsber.d.österr. Akad.Wiss.*, 188, 1979.

fe79 [28] H. G. Feichtinger. Gewichtsfunktionen auf lokalkompakten Gruppen. *Sitzber. d. österr. Akad. Wiss.*, 188:451–471, 1979.

fe80 [29] H. G. Feichtinger. Un espace de Banach de distributions tempérées sur les groupes localement compacts abéliens. *C. R. Acad. Sci. Paris S'er. A-B*, 290(17):791–794, 1980.

fe81-2 [30] H. G. Feichtinger. On a new Segal algebra. *Monatsh. Math.*, 92:269–289, 1981.

fe83 [31] 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.

fe84 [32] H. G. Feichtinger. Compactness in translation invariant Banach spaces of distributions and compact multipliers. *J. Math. Anal. Appl.*, 102:289–327, 1984.

fe89-1 [33] 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.

fe90 [34] H. G. Feichtinger. Generalized amalgams, with applications to Fourier transform. *Canad. J. Math.*, 42(3):395–409, 1990.

fe91-1 [35] H. G. Feichtinger. Discretization of convolution and reconstruction of band-limited functions from irregular sampling. pages 333–345. Academic Press, Boston, MA, 1991.

fe92 [36] H. G. Feichtinger. New results on regular and irregular sampling based on Wiener amalgams. In K. Jarosz, editor, *Function Spaces, Proc Conf, Edwardsville/IL (USA) 1990*, volume 136 of *Lect. Notes Pure Appl. Math.*, pages 107–121, New York, 1992. Marcel Dekker.

fe15 [37] 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.

fe16 [38] 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.

fegr85 [39] H. G. Feichtinger and P. Gröbner. Banach spaces of distributions defined by decomposition methods. I. *Math. Nachr.*, 123:97–120, 1985.

fegr89 [40] 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.

fegu20 [41] H. G. Feichtinger and A. Gumber. Completeness of shifted dilates in invariant Banach spaces of tempered distributions. *submitted*, page 12, 08 2020.

feka04 [42] H. G. Feichtinger and N. Kaiblinger. Varying the time-frequency lattice of Gabor frames. *Trans. Amer. Math. Soc.*, 356(5):2001–2023, 2004.

feka07 [43] H. G. Feichtinger and N. Kaiblinger. Quasi-interpolation in the Fourier algebra. *J. Approx. Theory*, 144(1):103–118, 2007.

feko98 [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.

felu15 [45] H. G. Feichtinger and F. Luef. Gabor Analysis and Algorithms. In B. Engquist, editor, *Encyclopedia of Applied and Computational Mathematics*, pages 575–579. Springer Berlin Heidelberg, 2015.

feno03 [46] H. G. Feichtinger and K. Nowak. A first survey of Gabor multipliers. In H. G. Feichtinger and T. Strohmer, editors, *Advances in Gabor Analysis*, Appl. Numer. Harmon. Anal., pages 99–128. Birkhäuser, 2003.

fezi98 [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.

fezi02 [48] H. G. Feichtinger and G. Zimmermann. An exotic minimal Banach space of functions. *Math. Nachr.*, 239-240:42–61, 2002.

fe15-2 [49] C. Fernandez Granda. Super-resolution of point sources via convex programming. *ArXiv e-prints*, jul 2015.

gawi99 [50] C. Gasquet and P. Witomski. *Fourier Analysis and Applications Filtering, Numerical Computation, Wavelets Transl From the French by R Ryan.* Springer, 1999.

ga69 [51] G. I. Gaudry. Multipliers of weighted Lebesgue and measure spaces. *Proc. Lond. Math. Soc., III. Ser.*, 19:327–340, 1969.

gr92-2 [52] P. Gröbner. *Banachräume glatter Funktionen und Zerlegungsmethoden.* PhD thesis, University Vienna, 1992.

gr01 [53] K. Gröchenig. *Foundations of Time-Frequency Analysis.* Appl. Numer. Harmon. Anal. Birkhäuser, Boston, MA, 2001.

gr06 [54] K. Gröchenig. Time-Frequency analysis of Sjöstrand´s class. *Rev. Mat. Iberoam.*, 22(2):703–724, 2006.

gr07 [55] 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.

grst07 [56] K. Gröchenig and T. Strohmer. Pseudodifferential operators on locally compact abelian groups and Sjöstrand's symbol class. *J. Reine Angew. Math.*, 613:121–146, 2007.

he03 [57] C. Heil. An introduction to weighted Wiener amalgams. In M. Krishna, R. Radha, and S. Thangavelu, editors, *Wavelets and their Applications (Chennai, January 2002)*, pages 183–216. Allied Publishers, New Delhi, 2003.

hero70 [58] 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.

hest69 [59] E. Hewitt and K. Stromberg. *Real and Abstract Analysis. A Modern Treatment of the Theory of Functions of a Real Variable. 2nd Printing corrected.* Springer, Berlin, Heidelberg, New York, 1969.

ho79-4 [60] R. B. Holmes. Mathematical foundations of signal processing. *SIAM Rev.*, 21(3):361–388, jul 1979.

jo66-1 [61] D. S. Jones. *Generalised Functions.* McGraw-Hill Book Co., New York-Toronto, 1966.

ka04 [62] R. P. Kanwal. *Generalized Functions. Theory and Applications. 3rd Revised ed.* Birkhäuser, MA, 2004.

ka76 [63] Y. Katznelson. *An Introduction to Harmonic Analysis. 2nd corr. ed.* Dover Publications Inc., New York, 1976.

ka04-1 [64] Y. Katznelson. *An Introduction to Harmonic Analysis. 3rd Corr. ed.* Cambridge University Press, 2004.

ke75 [65] J. L. Kelley. *General Topology. 2nd ed.* Springer, Berlin, Heidelberg, New York, 1975.

ko69 [66] G. Köthe. *Topological vector spaces. I.* Translated from the German by D. J. H. Garling. Die Grundlehren der mathematischen Wissenschaften, Band 159. Springer-Verlag New York Inc., New York, 1969.

la71 [67] R. Larsen. *An Introduction to the Theory of Multipliers.* Springer-Verlag, New York-Heidelberg, 1971.

li66 [68] M. J. Lighthill. *Einführung in die Theorie der Fourieranalysis und der Verallgemeinerten Funktionen.* Bibliographisches Institut, Mannheim, 1966.

ma59-1 [69] P. Malliavin. Sur l'impossibilite de la synthese spectrale sur la droite. *C. R. Math. Acad. Sci. Paris*, 248(2):2155–2157, 1959.

luoh06 [70] J.-R. Ohm and H. D. Lueke. *Signalübertragung: Grundlagen der Digitalen und Analogen Nachrichtenübertragungssysteme.* Springer, 2006.

os75 [71] M. S. Osborne. On the Schwartz-Bruhat space and the Paley-Wiener theorem for locally compact Abelian groups. *J. Funct. Anal.*, 19:40–49, 1975.

pa62 [72] A. Papoulis. *The Fourier Integral and Its Applications.* McGraw-Hill Book Co., Inc., 1962.

pa65 [73] A. Papoulis. *Probability, Random Variables, and Stochastic Processes.* McGraw-Hill Book Company, New York, 1965.

pa77 [74] A. Papoulis. *Signal Analysis.* McGraw-Hill Book Company, New York, 1977.

pe76 [75] J. Peetre. *New Thoughts on Besov Spaces.* Duke University Mathematics Series, No. 1. Mathematics Department, Duke University, 1976.

pl29 [76] A. Plessner. Eine Kennzeichnung der totalstetigen Funktionen. *J. Reine Angew. Math.*, 160:26–32, 1929.

plpostta19 [77] G. Plonka, D. Potts, G. Steidl, and M. Tasche. *Fourier Transforms.* Birkhäuser Basel, 2019.

prve08 [78] P. Prandoni and M. Vetterli. *Signal Processing for Communications.* CRC Press, 2008.

re68 [79] H. Reiter. *Classical Harmonic Analysis and Locally Compact Groups.* Clarendon Press, Oxford, 1968.

re71 [80] H. Reiter. *L^1-algebras and Segal Algebras.* Springer, Berlin, Heidelberg, New York, 1971.

rest00 [81] H. Reiter and J. D. Stegeman. *Classical Harmonic Analysis and Locally Compact Groups. 2nd ed.* Clarendon Press, Oxford, 2000.

re78 [82] H. J. Reiter. Über den Satz von Weil–Cartier. *Monatsh. Math.*, 86:13–62, 1978.

ri69-1 [83] M. A. Rieffel. Multipliers and tensor products of L^p-spaces of locally compact groups. *Studia Math.*, 33:71–82, 1969.

ro76-3 [84] A. Robert. A short proof of the Fourier inversion formula. *Proc. Amer. Math. Soc.*, 59:287–288, 1976.

ro83 [85] A. Robert. *Introduction to the Representation Theory of Compact and Locally Compact Groups.* Cambridge University Press, Cambridge, 1983.

ro00 [86] A. Robert. *A course in p-adic analysis.* Graduate Texts in Mathematics. 198. New York, NY: Springer., 2000.

ru62 [87] W. Rudin. *Fourier Analysis on Groups.* Interscience Publishers, New York, London, 1962.

ru91 [88] W. Rudin. *Functional Analysis 2nd ed.* International Series in Pure and Applied Mathematics. McGraw-Hill, New York, 1991.

sa98 [89] I. W. Sandberg. The superposition scandal. *Circuits Syst. Signal Process.*, 17(6):733–735, 1998.

sa04-1 [90] I. W. Sandberg. Continuous multidimensional systems and the impulse response scandal. *Multidimensional Syst. Signal Process.*, 15(3):295–299, 2004.

sh71 [91] H. S. Shapiro. *Topics in Approximation Theory*, volume 187 of *Lecture Notes in Mathematics.* Springer-Verlag, Berlin, 1971.

st05-3 [92] E. Stade. *Fourier analysis.* Pure and Applied Mathematics. A Wiley-Interscience Series of Texts, Monographs, 2005.

st70 [93] E. M. Stein. *Singular Integrals and Differentiability Properties of Functions.* Princeton University Press, Princeton, N.J., 1970.

suto07 [94] M. Sugimoto and N. Tomita. The dilation property of modulation spaces and their inclusion relation with Besov spaces. *J. Funct. Anal.*, 248(1):79–106, 2007.

tc84 [95] P. Tchamitchian. Généralisation des algèbres de Beurling. *Ann. Inst. Fourier (Grenoble)*, 34(4):151–168, 1984.

tc87 [96] P. Tchamitchian. tude dans un cadre hilbertien des algèbres de Beurling munies d'un poids radial croissance rapide. (Study in a Hilbertian framework of Beurling algebras equipped with a radial weight having rapid growth). *Ark. Mat.*, 25:295–312, 1987.

un20 [97] M. Unser. A note on BIBO stability. *arXiv preprint arXiv:2005.14428*, 2020.

gokove14 [98] M. Vetterli, V. Goyal, and J. Kovacevic. *Foundations of Signal Processing.* Cambridge Univ. Press, 2014.

vo11 [99] D. G. Voelz. *Computational Fourier Optics. A MATLAB Tutorial.* Tutorial text. Vol. 89. SPIE, 2011.

vo99 [100] J. von Neumann. *Invariant Measures.* Providence, RI: American Mathematical Society (AMS). xiv, 1999.

wa77 [101] H.-C. Wang. *Homogeneous Banach Algebras.* Marcel Dekker, New York, Basel, 1977.

we07-1 [102] F. Weisz. Inversion of the short-time Fourier transform using Riemannian sums. *J. Fourier Anal. Appl.*, 13(3):357–368, 2007.

we52 [103] J. G. Wendel. Left centralizers and isomorphisms of group algebras. *Pacific J. Math.*, 2:251–261, 1952.

we02-2 [104] D. Werner. *Functional Analysis. (Funktionalanalysis) 5., Erweiterte Aufl.* Springer-Verlag, Berlin, 2002.

ya12 [105] Y. Yamamoto. *From Vector Spaces to Function Spaces: Introduction to Functional Analysis with Applications.* SIAM, 2012.

za53 [106] A. C. Zaanen. *Linear Analysis Measure and Integral, Banach and Hilbert Space, Linear Integral Equations*, volume 2 of *Bibliotheca mathematica; a series of monographs on pure and applied mathematics.* North-Holland Publishing Company, Amsterdam, 1953.

za67 [107] A. C. Zaanen. *Integration.* North–Holland, Amsterdam, Completely revised edition of: An Introduction to the Theory of Integration edition, 1967.

zi98-1 [108] G. Zimmermann. Eigenfunctions of the Fourier Transform. Technical report, 1998.

Faculty of Mathematics, NuHAG,
Oskar-Morgenstern-Platz 1, 1090 Vienna, AUSTRIA
Present Course: Sept. to Dec. 2020 to Jan. at ETH Zuerich, see www.nuhag.eu/ETH20 for course instructions, previously autumn 2018: Charles Univ. Prague and Spring 2017 and Spring 2018, TU Muenich)

Email address: hans.feichtinger@univie.ac.at