RESEARCH ARTICLE | JULY 02 2009

Banach Gelfand Triples for Applications in Physics and Engineering

Hans G. Feichtinger

AIP Conf. Proc. 1146, 189–228 (2009)
https://doi.org/10.1063/1.3183542

Banach Gelfand Triples for Applications in Physics and Engineering

Hans G. Feichtinger

Faculty of Mathematics, NuHAG, University Vienna, AUSTRIA

Abstract.
The principle of extension is widespread within mathematics. Starting from simple objects one constructs more sophisticated ones, with a kind of natural embedding from the set of old objects to the new, enlarged set. Usually a set of operations on the old set can still be carried out, but maybe also some new ones. Done properly one obtains more completed objects of a *similar kind*, with additional useful properties. Let us give a simple example: While multiplication and addition can be done exactly and perfectly in the setting of $\mathbb{Q}$, the rational numbers, the field $\mathbb{R}$ of real numbers has the advantage of being complete (Cauchy sequences have a limit ...) and hence allowing for numbers like π or $\sqrt{2}$. Finally the even "more complicated" field $\mathbb{C}$ of complex numbers allows to find solutions to equations like $z^2 = -1$. The chain of inclusions of fields, $\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$ is a good motivating example in the domain of "numbers".

The main subject of the present survey-type article is a new theory of Banach Gelfand triples (BGTs), providing a similar setting in the context of (generalized) functions. Test functions are the simple objects, elements of the Hilbert space $L^2(\mathbb{R}^d)$ are well suited in order to describe concepts of orthogonality, and they can be approximated to any given precision (in the $\|\cdot\|_2$-norm) by test functions. Finally one needs an even larger (Banach) space of *generalized functions* resp. *distributions*, containing among others pure frequencies and Dirac measures in order to describe various mappings between such Banach Gelfand triples in terms of the most important "elementary building blocks", in a clear analogy to the finite/discrete setting (where Dirac measures correspond to unit vectors).

Our concrete Banach Gelfand triple is based on the Segal algebra $S_0(\mathbb{R}^d)$, which coincides with the *modulation space* $M^1(\mathbb{R}^d) = M_0^{1,1}(\mathbb{R}^d)$, and plays a very important and natural role for time-frequency analysis. We will point out that it provides the appropriate setting for a description of many problems in engineering or physics, including the classical Fourier transform or the Kohn-Nirenberg or Weyl calculus for pseudo-differential operators. Particular emphasis will be given to the concept of w^*-convergence and w^*-continuity of operators which allows to prove conceptual uniqueness results, and to give a correct interpretation to certain formal expressions coming up in various versions of the Dirac formalism.

Keywords: Keywords: Banach Gelfand triples, Fourier transform, Kohn-Nirenberg Symbol, w^*-convergence, spectrogram
PACS: 02.30.Sa Functional analysis, 02.30.Tb Operator theory, 02.30.Nw Fourier analysis, 02.30.Jr Partial differential equations, 03.65.Ca QM Formalism

CP1146, *Modelling of Engineering and Technological Problems*, edited by A. H. Siddiqi, A. K. Gupta, and M. Brokate

MOTIVATION AND INTRODUCTION

Although *Gelfand triples* such as $(\mathscr{S},L^2,\mathscr{S}')$ resp. so-called *rigged Hilbert spaces* (Hilbert spaces endowed with an extra structure of surrounding spaces) have a certain tradition, mostly within theoretical physics, not much systematic mathematical investigation of this concept has been made. It is the purpose of the present paper to bring the advantages of the concept of *Banach Gelfand triples* to the attention of a wider community, to exhibit a concrete, simple and versatile example, coming from time-frequency analysis, and to show how natural it is. The concrete content of these notes is only indicative for the potential, both for the strict derivation of vague but somehow valid claims, but also for teaching purposes, in a context where not the full power of Lebesgue integration or the theory of nuclear topological vector spaces is available. In fact, we even believe that some of the involved mathematical concepts can be replaced by more natural and hence more simple ones.

We address physicists and engineers and mathematicians interested in applications or who have to teach students from the above community. While the applied scientists are often using symbolic expressions and derive in this ways valid identities the more strict mathematical view-point requires to have solid mathematical definitions, clear rules and valid logical concatenations of arguments, step by step. By suggesting the concept of Banach Gelfand triples (BGTs), which somehow extend the idea of rigged Hilbert spaces, we hope to offer a quite natural but very powerful tool, which allows to validate some of these heuristic ideas. One of the specific points emphasized is the relevance of w^*-convergence of sequences of generalized functions and w^*-w^*-continuity of operators. Intuitively this can be explained to an audio engineer as follows: A sequence σ_n of distributions converges to σ_0 in the w^*-sense if (and only if) the spectrum (the short-time Fourier transform) of σ_n with respect to any reasonable (say Gaussian) window is going to look more and more like the spectrum of σ_0 over larger and larger parts of the time-frequency plane.

Let us mention that this is a written realization of explanations and statements given at various occasions in talks on this subject during the last four years[1]. The material will be covered in much more detail in a forth-coming book publication by the author (jointly with G. Zimmermann, for Birkhäuser's NAHA series).

We also view this as a part of a series of publications, showing how to get from basic linear algebra concepts to time-frequency analysis, in particular to Gabor analysis ([65], the discretized and computationally relevant version of time-frequency analysis). It starts with the "Guided Tour from Linear Algebra to the Foundations of Gabor Analysis" ([56]), where the basic algebraic principles are explained using the standard concepts of *linear independence* and *generating systems of vectors*. It uses linear algebra terminology, and works in the setting of finite dimensional vector spaces (cf. e.g. [108]). In fact, finite vectors are understood as functions on the cyclic group $\mathbb{Z}_N$ of unit roots of order N, and the properties of involved matrices (e.g. PINV-matrices) show how to

[1] Most of them are downloadable from NuHAG Talk server
`http://www.univie.ac.at/nuhag-php/program/talks.php`

obtain implementations in an efficient way[2]. The algebraic theory is then pursued in [54] in the setting of general finite Abelian groups. Based on this the papers [55, 24] provide a refined view on the tools needed to handle the continuous case. Basic facts had been already presented in [53, 61] and above all in the book "Foundations of Time-Frequency Analysis" by K. Gröchenig ([69]).

Linear Algebra and Matrix Analysis

Coming from linear algebra we have learned to focus on *bases*, i.e. *coordinate systems* which allow to express *any vector* in a *unique way* as a *finite linear combination* of the elements of a basis. In matrix terminology this boils down to concentrate on *invertible* matrices A, which have the pleasant property of allowing for *every right hand side* b a *unique solution* x of the linear equation expressed[3] as $A * x = b$, or equivalently write b as a linear combination of column vectors of A. Solving for x is then possible in various ways, e.g. using Gauss elimination, but in MATLAB^{TM} one could simply use the command $x = inv(A) * b$. If one makes use of a scalar product on $\mathbb{R}^d$ or $\mathbb{C}^d$ one finds that some bases are much more convenient than others, because they allow for an effortless calculation of the coefficients of a vector, by calculating scalar products. Let $(u_k)_{k=1}^d$ in $\mathbb{C}^d$ be such an *orthonormal basis*, then we can form a matrix U, with these vectors as column vectors. The fact that $x = \sum_{k=1}^d \langle x, u_k \rangle u_k$ for all $x \in \mathbb{C}^d$ is then equivalent to the fact that [4] $U * U' = Id$, the unit matrix of size $n \times n$, or equivalently:

$$Id = \sum_{k=1}^{d} P_k, \quad \text{where} \quad P_k = u_k * u_k'. \tag{1}$$

Since for the case of square matrices any right inverse matrix is also a left inverse matrix this good property is indeed equivalent to $U' * U = Id$. This is compactly expressing the fact that the columns of U (and hence in fact also the rows) form an orthonormal set, or in terms of the individual elements of the Gramian matrix $G = U' * U$ and using Kronecker's δ-symbol:

$$\langle u_k, u_j \rangle = \delta_{k,j}. \tag{2}$$

Much of this *spirit of doing linear algebra*, i.e. to work in the setting of finite dimensional vector spaces, using bases to expand vectors, or matrices in order to describe linear mappings, is simulated in the *bra-ket formalism* going back to Paul Dirac. This allows for continuous integrals instead of (finite or infinite) sums, keeping in mind the dual use of vectors, either as building blocks for *synthesis* (as with matrix multiplication $A \mapsto A * x$, building linear combinations of the column vectors of A) or *analysis*, taking scalar products with the same set of vectors, by forming $y \mapsto A' * y$. Unfortunately this freedom makes things occasionally quite vague, due to a couple of new problems:

---

[2] E.g. within the MATLAB^{TM} software.

[3] Here $*$ denotes matrix multiplication, following the convention used by MATLAB.

[4] We use the MATLAB convention of writing U' for the transpose, conjugate matrix of U.

1. vectors and operators are expanded as a *continuous superposition* (in terms of integrals) of certain building blocks instead of a series or sum;
2. the meaning of these integrals is not obvious (Riemann, Lebesgue);
3. the building blocks may not belong to the Hilbert space anymore;
4. hence scalar products between two such elements are not meaningful a priori;
5. one may even have problems with the domain of the rank-one operators;
6. as in the finite-dimensional case, one may have orthonormality *without completeness* (and vice versa); however, in the infinite dimensional setting one cannot argue with dimensions.

Frames and Riesz Bases in Hilbert Spaces

Let us therefore describe an intermediate step, where we have collections of vectors in a Hilbert space $\mathscr{H}$, for which the synthesis and/or the analysis mapping make sense, as bounded linear mappings between $\mathscr{H}$ and $\ell^2 = \ell^2(I)$ for some (countable) index sequence I. We will see concrete examples (Gabor families) in a moment.

Definition 1. A family $(g_i)_{i\in I}$ in a Hilbert space $\mathscr{H}$ is called a *Bessel family* if the *analysis* mapping $\mathbf{C} : f \mapsto (\langle f, g_i\rangle)_{i\in I}$ is bounded from $\mathscr{H}$ into $\ell^2(I)$, i.e. if and only if there exists some positive constant $B > 0$ such that

$$\|\mathbf{C}f\|^2_{\ell^2(I)} = \sum_{i\in I} |\langle f, g_i\rangle|^2 \leq B\|f\|^2_{\mathscr{H}} \quad \text{for all } f \in \mathscr{H}. \tag{3}$$

By adjointness this is the case if and only if the corresponding *synthesis mapping* $\mathbf{R} : \mathbf{c} = (c_i)_{i\in I} \mapsto \sum_{i\in I} c_i g_i$ is bounded. Using standard terminology known from O. Christensen's book ([19]) one defines:

Definition 2. A family $(g_i)_{i\in I}$ in a Hilbert space $\mathscr{H}$ is called a *frame* if there exist constants $A, B > 0$ such that for all $f \in \mathscr{H}$

$$A\|f\|^2 \leq \sum_{i\in I} |\langle f, g_i\rangle|^2 \leq B\|f\|^2. \tag{4}$$

Definition 3. A family $(g_i)_{i\in I}$ in $\mathscr{H}$ is called a *Riesz (basic) sequence* if $\sum_{i\in I} c_i g_i$ has a Hilbert space norm equivalent to the ℓ^2-norm of the sequence $(c_i)_{i\in I}$, i.e. if there exist constants $C, D > 0$ such that

$$C \cdot \|\mathbf{c}\|^2_{\ell^2} \leq \|\sum_{i\in I} c_i g_i\|^2_{\mathscr{H}} \leq D\|\mathbf{c}\|^2_{\ell^2} \qquad \text{for all } \mathbf{c} \in \ell^2. \tag{5}$$

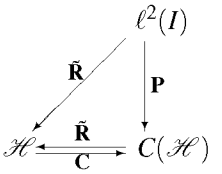
If $(g_i)_{i\in I}$ is a frame respectively Riesz sequence then the analysis mapping $\mathbf{C}$ respectively synthesis mapping $\mathbf{R}$ establishes an isomorphism between its domain Hilbert space and its closed(!) range within its target Hilbert space.

One easily shows that a family $(g_i)_{i\in I}$ in a Hilbert space $\mathscr{H}$ is a frame if and only if the so-called *frame operator* $\mathbf{S} := \mathbf{R} \circ \mathbf{C}$ is bounded and invertible, with bounded inverse.

Analogous results apply for Riesz basic sequences, with $\mathbf{C} \circ \mathbf{R}$ instead of $\mathbf{R} \circ \mathbf{C}$. Note that the fact that the composition $\mathbf{R} \circ \mathbf{C}$ (in the case of a frame) or $\mathbf{C} \circ \mathbf{R}$ (for a Riesz basic sequence) is invertible does *not* imply that $\mathbf{C}$ or $\mathbf{R}$ is invertible. However, if this is indeed the case we have: A family $(g_i)_{i\in I}$ is called a *Riesz basis* for $\mathscr{H}$ if it is both a frame and a Riesz sequence. In that case of course both $\mathbf{C}$ and $\mathbf{R}$ establish isomorphisms between $\mathscr{H}$ and $\ell^2(I)$.

It is not surprising that many of the concepts known from linear algebra extend first in a very natural way to (separable) Hilbert spaces $\mathscr{H}$ such as $L^2(\mathbb{R}^d)$. Instead of finite sequences of vectors (resp. functionals) one deals with infinite sequences and makes corresponding boundedness assumptions, which allow to establish a rather complete analogy to the *four spaces* concept proposed by G. Strang ([108]): The closed range of an operator and its adjoint together with the corresponding null-spaces can be used to have a full geometric understanding of the situation.

For the case of a frame, the situation is best described by the following diagram, where $\mathbf{C}$ is the analysis mapping $f \mapsto (\langle f, g_i\rangle)$, the operator $\tilde{\mathbf{R}}$ is defined as $\tilde{\mathbf{R}} := \mathbf{S}^{-1} \circ \mathbf{R}$, and $\mathbf{P} := \mathbf{C} \circ \tilde{\mathbf{R}}$ is the orthogonal projection from $\ell^2(I)$ onto the range of $\mathbf{C}$.



Hence $\tilde{\mathbf{R}} \circ \mathbf{C} = Id_{\mathscr{H}}$, i.e. $\tilde{\mathbf{R}}$ is a left-inverse to $\mathbf{C}$, the so called *Moore-Penrose inverse* to $\mathbf{C}$ (realized as PINV in MATLAB). Explicitly one finds for $f \in \mathscr{H}$

$$f = Id_{\mathscr{H}} f = (\tilde{\mathbf{R}} \circ \mathbf{C}) f = (\mathbf{S}^{-1} \circ \mathbf{R} \circ \mathbf{C}) f \tag{6}$$

$$= \mathbf{S}^{-1}\Big(\sum_{i\in I} \langle f, g_i\rangle g_i\Big) = \sum_{i\in I} \langle f, g_i\rangle \mathbf{S}^{-1} g_i. \tag{7}$$

This motivates the definition of the so-called *dual frame* $(\tilde{g}_i)_{i\in I}$ by $\tilde{g}_i := \mathbf{S}^{-1} g_i$. Using the dual frame one can reconstruct any $f \in \mathscr{H}$ from its coefficients $\mathbf{C} f = (\langle f, g_i\rangle)$ as

$$f = \sum_{i\in I} \langle f, g_i\rangle \tilde{g}_i, \tag{8}$$

i.e. $\tilde{\mathbf{R}}$ is just the synthesis operator with respect to the dual frame $(\tilde{g}_i)_{i\in I}$.

For details concerning *frames* resp. *Riesz basic sequences* we refer to O. Christensen's book ([19]) or [77]. The definition of a frame can be generalized to also cover *continuous frames*, e.g. *coherent frames* obtained by the action of a continuous group on some reference vector. Instead of a discrete (typically countable) index set a measure space Ω is used, the mapping $\mathbf{C}$ is now an injective mapping into $L^2(\Omega)$ with closed range, and hence the same kind of diagram is still valid. This concept has made early appearence in the work of G. Kaiser ([86, 85]), and S.T. Ali, J.P. Antoine and J.P. Gazeau ([3, 4]). There are more recent papers on this subject by J.P. Gabardo and D. Han [64] or M. Fornasier and H. Rauhut in [63], discussing the transition from a (redudant)

continuous to a (typically still redundant) discrete frame. Their work has been certainly inspired by the papers on coorbit theory by Feichtinger/Gröchenig ([48, 49]), which are also the basis for the first appearance of Banach frames in [67]. In this setting (the so-called *coorbit spaces*) concrete continuous frames appear in the context of irreducible, square-integrable group representations. Further generalizations of coorbit theory and continuous frames are treated in the work of S. Dahlke and his coauthors, [31, 30, 29].

Going beyond Frame Theory, towards Dirac

The Fourier transform is an important tool for both physics and engineering, making use of the "pure frequencies". What makes them so important is the fact that they are *eigenvectors* for the translation operators. Mathematicians like to consider the functions $\chi_s(t) = e^{2\pi i s\cdot t}$ as characters of the group $\mathbb{R}^d$, viewed as a LCA (= locally compact Abelian) group, with respect to addition of vectors. The exponential law implies that $\chi_s(x+y) = \chi_s(x)\cdot\chi_s(y),\ x,y\in\mathbb{R}^d$. Since we have the pointwise relation $\chi_r\cdot\chi_s = \chi_{r+s}$ we find that the *dual group*, or *frequency domain* is just

$$\widehat{\mathbb{R}^d} = \{\chi_s \,|\, s\in\mathbb{R}^d\}. \tag{9}$$

Spectral synthesis and *spectral analysis* (or Fourier analysis, or harmonic analysis in more general terms, see [102, 11]) address the question whether one can *compose* any signal, function, distribution f from this (continuous) family of "elementary building blocks" by superposition (since we have a continuous parameter it is natural to think of an integral representation), and on the other hand, wether and how one can identify the required coefficients (amplitudes/spectral components) from the signal f.

As in linear algebra, one has to settle the problem whether *every function*, or more precisely, every element from a given (topological) vector space can be represented, and secondly whether the representation is unique. As we will see, the setting of BGTs will also allow to differentiate and decide which one of the objects (of different complexity) can be composed or decomposed in which concrete way, e.g. through integral representation, in the weak sense in the case of the Hilbert space $\mathscr{H} = L^2(\mathbb{R}^d)$ or in the w^*-sense within the dual space $S_0'(\mathbb{R}^d)$.

Although it would be more natural from the linear algebra view-point described above to start with the synthesis problem, we find it more natural (in the Fourier context) to start with the analysis part. After all, according to our philosophy the two operations are mutually adjoint to each other.

Following the usual path the Fourier transform $\mathscr{F}$ is defined on $L^1(\mathbb{R}^d)$, the Banach space of all absolutely Lebesgue-integrable functions (modulo null-functions) as an integral transform as follows:

$$(\mathscr{F}f)(s) \equiv \hat{f}(s) = \int_{\mathbb{R}^d} f(t)\overline{\chi_s(t)}\,\mathrm{d}t = \langle f, \chi_s\rangle. \tag{10}$$

We will see later that it suffices to know it on some smaller spaces (such as the Schwartz space $\mathscr{S}(\mathbb{R}^d)$ or $S_0(\mathbb{R}^d)$, where it is enough to use the ordinary Riemannian integral).

Together with $\mathscr{F}$ we have to consider the adjoint mapping, i.e. *Fourier synthesis*, which in analogy to the situation in matrix analysis is given, at least for nice functions h, by

$$\mathscr{F}^* h = \int_{\mathbb{R}^d} h(s)\chi_s \, \mathrm{d}s. \tag{11}$$

This integral can be understood in the following sense: Each χ_s is a bounded and continuous function with $|\chi_s(t)| = 1$ for $t, s \in \mathbb{R}^d$, hence we have a pointwise well-defined function $(\mathscr{F}^* h)(t)$ if h is Riemann-integrable.

Although it is well known how to extend[5] to a unitary automorphism of $\left(L^2(\mathbb{R}^d), \|\cdot\|_2\right)$, thanks to the fundamental identity of the Fourier transform

$$\int_{\mathbb{R}^d} \hat{g}(y) f(y) \, \mathrm{d}y = \int_{\mathbb{R}^d} \hat{f}(s) g(s) \, \mathrm{d}s, \tag{12}$$

it is clear that one has to expect a lot of trouble with the domains of $\mathscr{F}$ and $\mathscr{F}^*$, because the different domains do not fit (a typical element $\hat{f} \in \mathscr{F}L^1(\mathbb{R}^d)$ may not be integrable itself, e.g. if f has discontinuities) and because the elementary building blocks, the pure frequencies $(\chi_s)_{s \in \mathbb{R}^d}$, do *not belong* to the Hilbert space $L^2(\mathbb{R}^d)$. On the other hand it is tempting to describe this continuous family (as Dirac did in some sense) as a "continuous coordinate system", satisfying a kind of (distributional) orthogonality relation as well as a decomposition of the identity operator as a continuous integral of rank-one operators comparable to the pair (2) and (1). We will provide arguments towards a meaningful interpretation of such claims in the context of *Banach Gelfand triples*.

THE BANACH GELFAND TRIPLE (S_0, L^2, S_0')

The above observation already calls for a unified treatment of the Fourier transform in the finite as well as in the Euclidean setting, or even (according to A. Weil) in the setting of LCA (locally compact Abelian) groups, including the field of p-adic numbers (see [116, 83, 102]). It is also clear that one cannot - despite its importance - stay within the Hilbert space $L^2(\mathbb{R}^d)$ anymore. We will try to convince the reader that the concept of BGTs (= Banach Gelfand triples) is a good way out of this problem.

Banach Gelfand Triples and their Morphisms

Recall the famous formula $e^{2\pi i} = 1$. It would not have made sense to a Greek mathematician, even if he had perfect knowledge of the field $\mathbb{Q}$ of rationals. One has to be able to create irrational numbers such as π beforehand, and one also has to be able to extend addition and multiplication to the larger domain of complex numbers. Finally one has to have a canonical way to give a meaning to the power series expression (adding up

[5] By arguments quite similar to those used in the extension of multiplication of rational numbers to the domain $\mathbb{R}$.

infinitely many of those numbers) to reach the perfect calculus of exponential functions in the complex domain. We will follow a similar path, with the Hilbert space $\mathscr{H}$ (typically $L^2(\mathbb{R}^d)$) playing the role of $\mathbb{R}^d$ and the generalized functions being the analogue of the complex numbers. Before going to the concrete BGT (S_0, L^2, S_0') let us introduce the concept of Banach Gelfand triples in full generality (for the sake of simplicity we restrict our discussions to the case of *separable* Banach spaces).

Definition 4. We call a triple of vector spaces $(B, \mathscr{H}, B')$ a *Banach Gelfand triple* if $(B, \|\cdot\|_B)$ is Banach space, which is dense in some Hilbert space $\mathscr{H}$, and which in turn is contained in B', the dual of B.

There are many examples, and the basic fact is a natural embedding of the elements of B (usually the space of *test functions*) into its dual space B', the space of *generalized functions* or *distributions*[6].

Although the idea of *rigged Hilbert spaces* ([103, 6, 35, 120, 36, 2, 7]) is very close to our BGT concept there are two important differences First of all it is clear that we allow for Banach spaces instead of a Hilbert spaces of dual Hilbert spaces "surrounding" the central Hilbert space, nor any *nuclear topological* vector space, such as $\mathscr{S}(\mathbb{R}^d)$. The concrete example, starting from the space $S_0(\mathbb{R}^d)$ allows to obtain nevertheless a kernel theorem. We are not aware of any kernel theorem for rigged Hilbert spaces other than those using nuclear (hence not Banach or Hilbert) spaces. One can trace the validity of the kernel theorem back to the tensor product factorization property (Lemma 4), which in turn has to do with the "separation of variables" property in the Fourier algebra, which has been historically one of the highlights of J.B. Fourier's concept.

It has been expressed by several authors (cf. [36, 37, 66**?**]) that rigged Hilbert spaces (a triple of Hilbert spaces, forming a BGT in our sense) allows to describe valid identities which cannot be formulated in the Hilbert space setting alone[7] We take the same view-point, but emphasize the close connection between the inner Banach space and its dual by working with *four topologies*, i.e. by giving the (natural) w^*-topology on B' a prominent role. Note that the dual space for S_0' endowed with the w^*-topology (often denoted as the weak $\sigma(B', B)$ topology) is just B itself, and hence one has a kind of Riesz-representation theory for BGTs in the background. Furthermore it is helpful to recall that bounded (closed) subsets in B' are compact in this topology according to the theorem of Banach-Alaoglou ([106], section 3.15).

The *prototype* of a Banach Gelfand triple is $(\ell^1, \ell^2, \ell^\infty)(\mathbb{Z})$, where the w^*-topology describes coordinate-wise convergence, i.e. views ℓ^∞ as subset of $\mathbb{R}^{\mathbb{Z}}$ with the product topology in the sense of Tychonoff.

In fact, one may view Banach Gelfand triples as a new *category* in the spirit of MacLane ([96]), where the *morphisms* are the "structure preserving mappings", i.e.

[6] While smaller spaces of test functions give larger space of bounded linear functionals on them, one has to keep in mind that B is not degenerating, because then this construction breaks down. So for our purpose one may think of a situation where $\mathscr{S}(\mathbb{R}^d) \subset B$.

[7] For example, point evaluations do not make sense on $L^2(\mathbb{R}^d)$ while they make perfect sense on a Sobolev space, once the smoothness parameter satisfies $s > d/2$, according to Sobolev's embedding theorem.

linear mappings which are continuous with respect to each of the *four* (!) topologies.

Definition 5. If $(B_1, \mathscr{H}_1, B_1')$ and $(B_2, \mathscr{H}_2, B_2')$ are Gelfand triples then a linear operator T is called a [unitary] Gelfand triple isomorphism if

1. A is an isomorphism between B_1 and B_2.
2. A is [a unitary operator resp.] an isomorphism between $\mathscr{H}_1$ and $\mathscr{H}_2$.
3. A extends to a weak* isomorphism as well as a norm-to-norm continuous isomorphism between B_1' and B_2'.

In principle every ONB (= *orthonormal basis*) $\Psi = (\psi_i)_{i \in I}$ for a given Hilbert space $\mathscr{H}$ can be used to establish such a unitary isomorphism, by choosing as B the space of elements within $\mathscr{H}$ which have an absolutely convergent expansion, i.e. satisfy $\sum_{i \in I} |\langle x, \psi_i \rangle| < \infty$. Of course, this space, which deserves perhaps the symbol A_Ψ, depends on the choice of the orthonormal basis Ψ, but of course one has many *equivalent* bases describing the same space.

For the case of the perhaps most important ONB for $\mathscr{H} = L^2([0,1])$, i.e. for the trigonometric system, the corresponding definition is already around since the times of N. Wiener, who suggested to consider specifically $A(\mathbb{T})$, the space of absolutely continuous Fourier series, because it has very good and useful properties (compared to the Lebesgue space $\left(L^1(\mathbb{T}), \|\cdot\|_1\right)$, where e.g. the Fourier inversion is a non-trivial matter). It is also not surprising in retrospect to see that in the discussion the dual space $PM(\mathbb{T}) = A(\mathbb{T})'$ came up, the space of *pseudo-measures*. One can extend the Fourier transform to this space, and in fact interpret this extended mapping, in conjunction with the classical Plancherel theorem as the first unitary Banach Gelfand triple isomorphism, between $(A, L^2, PM)(\mathbb{T})$ and $(\ell^1, \ell^2, \ell^\infty)(\mathbb{Z})$. [8]

It is the main goal of this article to show how the use of the Banach algebra $S_0(\mathbb{R}^d)$ allows to have a similar interpretation of the Fourier transform (and many other mappings relevant for physics, engineering, or mathematical considerations in time-frequency analysis), how to make use of the w^*-concept and how to re-interpret the Dirac formalism in this context.

Having expounded the general theory of Banach Gelfand triples, we are now ready to introduce the constituents of a particularly useful example, namely the Banach Gelfand triple (S_0, L^2, S_0').

Modulation Spaces

The Banach space $\left(S_0(\mathbb{R}^d), \|\cdot\|_{S_0}\right)$ of test functions to be used in the following is a particular instance of a class of function spaces studied in time-frequency analysis (TF-analysis), called *modulation spaces*. In order to define these spaces we have to recall some concepts from that field. The basic tools in TF-analysis are *time-* and *frequency shifts* (TF-shift) given by $T_x f(t) = f(t-x)$ and $M_\omega f(t) = e^{2\pi i \omega \cdot t} f(t)$, for functions f

[8] The Segal algebra $S_0(G)$, defined for general LCA (= locally compact Abelian) groups is in fact a generalization of this construction, i.e. $S_0(\mathbb{T}) = A(\mathbb{T})$.

on $\mathbb{R}^d$. They are combined to (unitary) *time-frequency shift* operators

$$\pi(\lambda) = \pi(x,\omega) = M_\omega T_x, \quad \text{for } \lambda = (x,\omega) \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d. \tag{13}$$

Using these operators one defines (e.g. on $L^2(\mathbb{R}^d)$ or for continuous and absolutely Riemann-integrable functions) the *Short-Time Fourier Transform* as a function on the *time-frequency plane* ([69]) resp. *phase space* ([62]) in the following way:

$$V_g f(\lambda) = V_g f(x,\omega) = \langle f, M_\omega T_x g\rangle = \langle f, \pi(\lambda) g\rangle \text{ for } \lambda = (x,\omega) \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d. \tag{14}$$

Modulation spaces occur in the study of the concentration of a function in the time-frequency plane, described in terms of function spaces over $\mathbb{R}^d \times \widehat{\mathbb{R}}^d$. The classical ones are defined as follows: Let $g \in \mathscr{S}(\mathbb{R}^d)$ be a Schwartz function, $1 \le p,q < \infty$, $s \in \mathbb{R}$, then

$$M_s^{p,q}(\mathbb{R}^d) = \{f \in \mathscr{S}'(\mathbb{R}^d) : \|f\|_{M_s^{p,q}} < \infty\}, \tag{15}$$

where the norm $\|f\|_{M_s^{p,q}}$ on $M_s^{p,q}(\mathbb{R}^d)$ is given as

$$\|f\|_{M_s^{p,q}} := \left(\int \left(\int |\langle f, M_\omega T_x g\rangle|^p \, \mathrm{d}x \right)^{q/p} (1+|\omega|)^{sq} \, \mathrm{d}\omega \right)^{1/q}, \tag{16}$$

i.e. for which $V_g f$ belongs to some weighted mixed-norm space over phase space. In the "classical" case the weight depends only on frequency, hence the spaces are isometrically translation invariant. The only important facts about the constraint imposed on $V_g f$ is the membership in a *solid*[9] and *translation invariant* Banach space over $\mathbb{R}^{2d}$. We use the abbreviations $M_s^p := M_s^{p,p}$ and $M^p := M_0^{p,p}$.

The modulation space $M_s^{p,q}(\mathbb{R}^d)$ is a Banach space of tempered distributions, the definition is independent of the analyzing function g, and different g's yield equivalent norms on these spaces. The Gauss function is a good choice. Among the modulation spaces are the following important function spaces:

(a) the space $S_0(\mathbb{R}^d)$ we are after is just $M_0^{1,1}(\mathbb{R}^d) = M^1(\mathbb{R}^d)$;

(b) $L^2(\mathbb{R}^d) = M_0^{2,2}(\mathbb{R}^d)$;

(c) the Bessel potential spaces $\mathscr{H}_s(\mathbb{R}^d)$, defined via the Fourier transform by

$$\mathscr{H}_s(\mathbb{R}^d) = \{f \in \mathscr{S}' : \int |\hat{f}(\omega)|^2 (1+|\omega|)^{2s} \, \mathrm{d}\omega < \infty\} \tag{17}$$

coincide with the modulation spaces $M_s^{2,2}(\mathbb{R}^d)$;

(d) the Shubin classes $Q_s(\mathbb{R}^d)$, which can be characterized by a weighted $L^2(\mathbb{R}^{2d})$-condition with respect to the radial symmetric weight over phase-space of the form $v_s(\lambda) = (1+|\lambda|^2)^{s/2}$ instead of the usual weight $w_s(\omega) := (1+|\omega|)^s$ resp. $(1+|\omega|^2)^{s/2}$.

[9] In a solid space the norm behaves monotonically, i.e. $|F(x)| \le |G(x)|$ for all $x \in \mathbb{R}^d$ implies that the norm of F is smaller than the norm of G.

A lot of details on these spaces can be found in the book of Gröchenig, and in the survey note [46] (written in 1983 and published in 2003).

The original description of the modulation spaces was in terms of generalized Wiener amalgams, on the Fourier transform side:

$$M_s^{pq}(\mathbb{R}^d) = \mathscr{F}^{-1}[W(\mathscr{F}L^p, \ell^q_{w_s})(\mathbb{R}^d)],$$

or equivalently, Banach spaces of distributions obtained using BUPUs (bounded uniform partitions of unity, such as a collection of shifted B-splines) (cf. [43, 78]).

The Banach space $S_0(\mathbb{R}^d)$ and its Various Descriptions

In the following we will establish the basic properties of the Banach space of test functions on which our BGT-approach will be based[10]. $S_0(\mathbb{R}^d)$ can be described in many ways and many equivalent norms can be used to characterize this space. Originally (see [42]) it was introduced as the *Wiener amalgam space* $W(\mathscr{F}L^1, \ell^1)$ (see [43] for generalities of this concept), but the equivalence between discrete and "continuous" norms (using control functions) can be used to show that it coincides with the coorbit space (as developed in full generality in [48]) or with the *modulation space* $(M^1(\mathbb{R}^d), \|\cdot\|_{M^1})$ (see the book [69] for a good introduction to the subject in the context of time-frequency analysis). We will follow the description given there, going back to [45], published in 1989.

According to the description above we can define $S_0(\mathbb{R}^d) := M^1(\mathbb{R}^d)$ by means of the STFT with respect to the Gaussian window $g_0(t) = e^{-\pi|t|^2}$. This choice has the advantage that Fourier invariance of this space is easily verified. It is also not difficult to check that $M^1(\mathbb{R}^d) \subset L^1 \cap C_0(\mathbb{R}^d)$[11]. The following is an alternative definition not making reference to Lebesgue integrals (and thus suitable for applied courses):

Definition 6. $S_0(\mathbb{R}^d) := \{f \in C_0(\mathbb{R}^d) : f$ absolutely Riemann-integrable over $\mathbb{R}^d$, $V_{g_0}f$ absolutely Riemann-integrable over $\mathbb{R}^d \times \widehat{\mathbb{R}}^d\}$, with the norm $\|f\|_{S_0} := \|V_{g_0}f\|_{L^1}$.

An *atomic* characterization[12] also used by H. Reiter (see [101]) is

Theorem 1. *We call a function $f \in \mathscr{F}L^1(\mathbb{R}^d)$ an atom (on the time-side) if* $\operatorname{supp}(f) \subseteq B_1(0)$ *for some $x \in \mathbb{R}^d$. Then $S_0(\mathbb{R}^d)$ consists of all absolutely convergent sums of atoms, i.e. $f \in S_0(\mathbb{R}^d)$ if and only it has a representation as*

$$f = \sum_{n\geq 1} T_{x_n} f_n, \quad \textit{with} \quad \sum_{n\geq 1} \|\widehat{f_n}\|_{L^1} < \infty. \tag{18}$$

[10] It is occasionally referred to as *Feichtinger's algebra* in the literature, see [102].

[11] We write $\left(C_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$ for the space of continuous, complex-valued functions, vanishing at infinity, i.e. with $\lim_{|x|\to\infty} f(x) = 0$, endowed with the sup-norm $\|f\|_\infty := \sup_{x\in\mathbb{R}^d} |f(x)|$.

[12] This atomic characterization should be reminiscent of the atomic characterization of Hardy spaces, given by Coifman and Weiss ([23]).

Endowed with the natural norm, i.e.

$$\|f\|_{S_0} := \inf\left\{\sum \|f_n\|_{\mathscr{F}L^1(\mathbb{R}^d)} : f = \sum T_{x_n} f_n, \sum \|f_n\|_{\mathscr{F}L^1(\mathbb{R}^d)} < \infty\right\}. \tag{19}$$

$S_0(\mathbb{R}^d)$ is the smallest (non-trivial) Banach space $(B, \|\cdot\|_B)$ with the property $\|\pi(\lambda)f\|_B = \|f\|_B$ for all $f \in B, \lambda \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d$, i.e. it is continuously embedded into any other space with this property[13]*. Moreover $S_0(\mathbb{R}^d)$ is invariant under the Fourier transform.*

This is [41], Thm.1. See [94, 95] for further characterizations of $S_0(\mathbb{R}^d)$, and of course the book [69]. It was the clue for many other interesting properties of $S_0(\mathbb{R}^d)$, which are nowadays proved using TF-arguments. Among others one has the following characterization. Since $\hat{g}_0 = g_0$ it sheds some light on the Fourier invariance of $S_0(\mathbb{R}^d)$. Due to the Fourier invariance one can also avoid the $\mathscr{F}L^1$ norm $\|h\|_{\mathscr{F}L^1} := \|f\|_{L^1}$ for $h = \hat{f}$ by doing the decomposition into pieces of equal size on the Fourier transform side. In this way one achieves a description of $f \in S_0(\mathbb{R}^d)$ as a sum of band-pass signals. This is what Hans Reiter really used, in [100, 101].

Lemma 1. *All absolutely convergent series of time-frequency shifts of g_0 are contained in $S_0(\mathbb{R}^d)$, and even make up all of $S_0(\mathbb{R}^d)$, i.e.*

$$S_0(\mathbb{R}^d) = \left\{\sum_{n\in\mathbb{N}} a_n M_{\xi_n} T_{x_n} g_0 : (x_n, \xi_n) \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d, (a_n)_{n\in\mathbb{N}} \in \ell^1(\mathbb{N})\right\}. \tag{20}$$

Since the choice of the window in the definition of modulation spaces gives these definition some smell of arbitrariness, some people prefer the characterization of $S_0(\mathbb{R}^d)$ using the (quadratic) Wigner distribution as a suitable alternatively, despite the fact that from the description below it is a-priori not clear why $S_0(\mathbb{R}^d)$ should be a linear manifold. Let us recall the definition of the cross-Wigner distribution (see [20, 21, 22, 79, 80] for $f, g \in L^2(\mathbb{R}^d)$ first, with $z = (x, \xi)$:

$$W(f,g)(z) = \int_{\mathbb{R}^d} e^{-2\pi i \xi\cdot y} f(x + \tfrac{1}{2}y)\overline{g(x - \tfrac{1}{2}y)}\,\mathrm{d}y. \tag{21}$$

Lemma 2. *$f \in S_0(\mathbb{R}^d)$ if and only if the Wigner function $W(f,f) \in L^1(\mathbb{R}^{2d})$.*

Whereas some basic invariance properties of $S_0(\mathbb{R}^d)$, or properties like the restriction to subgroups or integration along subgroups can be derived quite easily (cf. [40]) the last criterion is the most useful for the derivation of metaplectic invariance (cf. last section).

The dual space $(S_0'(\mathbb{R}^d), \|\cdot\|_{S_0'})$

Together with the space $S_0(\mathbb{R}^d)$ of test functions we will have to consider its dual space, the collection of all bounded linear functionals on $\left(S_0(\mathbb{R}^d), \|\cdot\|_{S_0}\right)$. Since $S_0(\mathbb{R}^d)$

[13] See [102, 41, 99, 101] for background on Segal algebras resp. the Segal algebra viewpoint on $S_0(\mathbb{R}^d)$. It is also the smallest strongly character invariant Segal algebra.

is the smallest isometrically TF-invariant Banach space its dual is essentially the largest space (of distributions) with this property. We will make use of this space, endowed with its standard norm respectively the w^*-topology.

We use the symbol $(S_0'(\mathbb{R}^d), \|\cdot\|_{S_0'})$ for the space of bounded linear functionals on $(S_0(\mathbb{R}^d), \|\cdot\|_{S_0})$, where the norm for $\sigma \in S_0'(\mathbb{R}^d)$ is defined as usual by

$$\|\sigma\|_{S_0'} := \sup\{|\sigma(f)| : f \in S_0(\mathbb{R}^d), \|f\|_{S_0} \leq 1\}. \tag{22}$$

Definition 7. A distribution $\sigma \in S_0'$ is *regular*, if there exists a locally integrable function $\varphi \in {L^1}_{\mathrm{loc}}(\mathbb{R}^d)$ such that

$$\sigma(f) = \int_{\mathbb{R}^d} \varphi(t)\, f(t)\, \mathrm{d}t \quad \text{for all } f \in S_0(\mathbb{R}^d). \tag{23}$$

In this case, we write $\sigma =: \sigma_\varphi$.

Here, we have $\sigma_\varphi = \sigma_\psi$ if and only if $\varphi(t) = \psi(t)$ almost everywhere. Regularity of a distribution does not necessarily imply that the integral in (23) is absolutely convergent for all $f \in S_0$. This holds for φ in an appropriate Wiener amalgam space, though.

Proposition 1. *For $\varphi \in W(L^1, \ell^\infty)$, we have that $\sigma_\varphi \in S_0'$ with $\|\sigma_\varphi\|_{S_0'} \leq \|\varphi\|_{W(L^1,\ell^\infty)}$, and the integral in* (23) *is absolutely convergent for all $f \in S_0$.*

In particular, we see that spaces like S_0, L^p, $W(L^{p_1}, \ell^{p_2})$ are continuously embedded in S_0' for $1 \leq p, p_1, p_2 \leq \infty$, in the sense that for an element φ of one of these spaces, we have $\sigma_\varphi \in S_0'$, and the norm of σ_φ can be estimated from above by the respective norm of φ. For $p_2 = \infty$ this argument implies that we can even consider periodic functions $L^p(\mathbb{T}^d)$ as subspaces of $W(L^p, \ell^\infty) \subset S_0'(\mathbb{R}^d)$. In a similar way every bounded measure $\mu \in M_b(\mathbb{R}^d)$ can be identified with $\sigma_\mu \in S_0'$ via

$$\sigma_\mu(f) = \int_{\mathbb{R}^d} f(t)\, \mathrm{d}\mu(t) \quad \text{for all } f \in S_0\,,$$

with $\|\sigma_\mu\|_{S_0'} \leq C\|\mu\|_{M_b}$. In particular, all finite discrete measures define elements of S_0'. But there are also many other (unbounded) measures within $S_0'(\mathbb{R}^d)$, since the space of *translation-bounded* measures $W(M_b, \ell^\infty)$ is contained S_0'. For example, $\text{⨿}_\Lambda := \sum_{\lambda\in\Lambda} \delta_\lambda \in S_0'(\mathbb{R})$ for any lattice $\Lambda \lhd \mathbb{R}^d$.

The standard methods for Wiener amalgam spaces (cf. [43, 78]) imply that $S_0'(\mathbb{R}^d)$ can be characterized as $W(\mathscr{F}L^\infty, \ell^\infty)(\mathbb{R}^d)$, the space of *translation bounded* quasi-measures, because $\mathscr{F}L^\infty(\mathbb{R}^d) = PM(\mathbb{R}^d) := \{\sigma = \mathscr{F}^{-1}h, \text{ for some } h \in L^\infty(\mathbb{R}^d)\}$, the space of *pseudo-measures*, coincides locally with the space of *quasi-measures* ([92, 39]).

There are also quite useful convolution relations, such as

$$S_0 * S_0' \subseteq W(\mathscr{F}L^1, \ell^\infty) = \mathscr{M}(S_0)(\mathbb{R}^d), \tag{24}$$

where $\mathscr{M}(S_0)$ are the pointwise multipliers of $S_0(\mathbb{R}^d)$. However, $S_0(\mathbb{R}^d)$ is not dense in $S_0'(\mathbb{R}^d)$ with respect to the norm topology and therefore we have to invoke a second, weaker topology on this dual space.

w^*-convergence in $S_0'(\mathbb{R}^d)$

A sequence $(\sigma_n)_{n\in\mathbb{N}}$ in $S_0'(\mathbb{R}^d)$ is w^*-convergent[14] to $\sigma_0 \in S_0'(\mathbb{R}^d)$, in symbols

$$\sigma_0 = w^* - \lim_n \sigma_n \tag{25}$$

if for every test function $f \in S_0(\mathbb{R}^d)$ one has

$$\lim_n \sigma_n(f) \to \sigma_0(f), \tag{26}$$

i.e. pointwise convergence of the sequence (σ_n) to some limit σ_0. The following equivalent characterization is valid for arbitrary Banach spaces:

Lemma 3. *A (bounded) sequence $(\sigma_n)_{n\in\mathbb{N}}$ in $S_0'(\mathbb{R}^d)$ is w^*-convergent to σ_0 if and only if for every compact $M \subset S_0(\mathbb{R}^d)$ and every $\varepsilon > 0$ one has: There exists some index n_0 such that $n \geq n_0$ implies*

$$|\sigma_n(f) - \sigma_0(f)| \leq \varepsilon, \ \textit{for all} \ \ f \in M. \tag{27}$$

Since the atomic characterization of $\left(S_0(\mathbb{R}^d), \|\cdot\|_{S_0}\right)$ implies that for any non-zero $g \in S_0(\mathbb{R}^d)$ the set of all TF-shifted copies of g, i.e. the family $\{\pi(\lambda)g \,|\, \lambda \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d\}$ is total in $\left(S_0(\mathbb{R}^d), \|\cdot\|_{S_0}\right)$, we arrive at the following characterization of w^*-convergence:

Theorem 2. *A bounded sequence $(\sigma_n)_{n\in\mathbb{N}}$ is w^*-convergent to $\sigma_0 \in S_0'$ if and only if for some (and therefore for any) non-zero $g \in S_0(\mathbb{R}^d)$ one has pointwise, or equivalently uniform convergence over compact sets of the TF-plane of $V_g\sigma_n$ to $V_g\sigma_0$. More explicitely: For every $R > 0$ and $\varepsilon > 0$ there exists some index n_0 such that*

$$|V_g(\sigma_n)(\lambda) - V_g(\sigma_0)(\lambda)| \leq \varepsilon \quad \forall n \geq n_0, \lambda \quad \textit{with} \quad |\lambda| \leq R. \tag{28}$$

A verbal description of this situation is to say that the spectrograms of σ_n look more and more similar to the spectrogram of σ_0 over larger and larger parts of phase space.

There are of course many important examples where w^*-convergence is valid, while in contrast we do not have norm convergence, even for some simple examples as

1. if $x_n \to x_0$, then $\delta_{x_0} = w^* - \lim_n \delta_{x_n}$, while $\|\delta_x - \delta_y\|_{S_0'} = 2$ for $x \neq y$.
2. $\chi_{s_n} \to \chi_{s_0}$ in the w^*-sense if and only if $s_n \to s_0$;
3. $\sqcup\!\sqcup_r = \sum_{k\in\mathbb{Z}^d} \delta_{rk} \to \delta_0 \quad \text{for} \quad r \to \infty$;
4. $(St_\rho g)_{\rho\to 0} \to \delta_0$ in the w^*-topology, for $\rho \to 0$, if $\int_{\mathbb{R}^d} g(x)dx = 1$, where $St_\rho g(x) = \rho^{-d} g(x/\rho)$ is the L^1-normalized, dilated version of g;
5. $h\sqcup\!\sqcup_h \to \mathbf{1} = \sigma_{\mathbf{1}} \quad \text{for} \quad h \to 0$ (Riemannian integrals definition for $f \in S_0(\mathbb{R}^d)$).

[14] The study of this convergence goes back to [17], where *relative completions* have been introduced for the study of multiplier spaces.

For later use let us describe explicitely what it means that a linear mapping T on $S_0'(\mathbb{R}^d)$ is w^*-w^*-continuous using bounded and w^*-convergent sequences:

$$\sigma_n(f) \to \sigma_0(f) \quad \forall f \in S_0(\mathbb{R}^d) \quad \Rightarrow \quad T(\sigma_n)(g) \to T(\sigma_0)(g) \quad \forall g \in S_0(\mathbb{R}^d).$$

Under the boundedness assumption it is enough to test convergence on total subsets of $S_0(\mathbb{R}^d)$ only, e.g. on the set of atoms (or coherent states) $(\pi(\lambda)g)_{\lambda \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d}$.

Later on we will see that the usually vague and heuristic argument, exhibiting the Fourier transform as a limit of Fourier series expansions, can be made precise in such a context. In fact, the Fourier transform $\hat{f}$ of $f \in L^1(\mathbb{R}^d)$ can be viewed as the w^*-limit of the Fourier transforms of the correspondingly periodized version of f (in fact classical Fourier series expansions), with the period length going to infinity.

Practically all the invariance properties of $S_0(\mathbb{R}^d)$, including its invariance under the Fourier transform, can be extended to invariance properties for $S_0'(\mathbb{R}^d)$. One possible explanation for this fact is the w^*-density of $S_0(\mathbb{R}^d)$ in $S_0'(\mathbb{R}^d)$. From the point of view of introducing the extended operators it is more convenient to use adjointness relations, which we will do later on, using Banach Gelfand triples.

The Fourier Transform on (S_0, L^2, S_0')

We now come back to the question we started with, namely to define a convenient setting for the Fourier transform. Using our Banach Gelfand triple (S_0, L^2, S_0'), we find the following, satisfactory answer. It is a perfect demonstration example for the power of *unitary Banach Gelfand triple automorphisms.*

Theorem 3. *The Fourier transform, defined in the usual way via*

$$\hat{f}(s) = \int_{\mathbb{R}^d} f(t) e^{2\pi i s \cdot t} dt \quad \textit{for} \quad f \in S_0(\mathbb{R}^d) \tag{29}$$

extends in a unique way to a (unitary) Banach Gelfand triple automorphism, based on the definition

$$\hat{\sigma}(f) := \sigma(\hat{f}), \quad \textit{for} \quad \sigma \in S_0'(\mathbb{R}^d), f \in S_0(\mathbb{R}^d). \tag{30}$$

It is also characterized by the fact that it is mapping the pure frequencies χ_s are mapped on the corresponding Dirac measures δ_s.

The direct statement is based on the Fourier invariance of $S_0(\mathbb{R}^d)$, while the uniqueness follows from the w^*-density of $S_0(\mathbb{R}^d)$ respectively trigonometric polynomials in $S_0'(\mathbb{R}^d)$.

Gabor characterization of (S_0, L^2, S_0')

The space $\left(S_0(\mathbb{R}^d), \|\cdot\|_{S_0}\right)$ has a number of further equivalent properties, some of them are quite convenient for various purposes. We will use Weyl-Heisenberg families, indexed by *lattices* $\Lambda = A\mathbb{Z}^{2d} \lhd \mathbb{R}^{2d}$, for some non-singular $2d \times 2d$ matrix A:

Definition 8. A family $(\pi(\lambda)g)_{\lambda\in\Lambda}$ is called a *Weyl-Heisenberg family*. It will be convenient to write simply $(g_\lambda)_{\lambda\in\Lambda}$.

A WH-family is also called a *Gabor family*, cf. [65]. If a WH-family is a frame or Riesz basis we will speak of a *Gabor frame* or *Gabor Riesz basis* (for its closed linear span). D. Gabor suggested to use the Gauss-function $g = g_0$, and the ("critical") von Neumann lattice $\Lambda = \mathbb{Z}^{2d}$. Despite the perfect time-frequency localization this family is not a Riesz basis for $\mathscr{H} = L^2(\mathbb{R}^d)$, and the so-called dual γ_{Bast} proposed by M. Bastiaans ([10]) is not in $L^2(\mathbb{R}^d)$. There are two other important results to be mentioned here. For their description we recall the *adjoint lattice* Λ°, which consists of those elements in $\mathbb{R}^{2d}$ which satisfy the commutation property

$$\pi(\lambda^\circ)\pi(\lambda) = \pi(\lambda)\pi(\lambda^\circ) \quad \text{for all } \lambda\in\Lambda. \tag{31}$$

The so-called Wexler-Raz principle (see [118, 33, 82, 53]) says that a WH-family $(g_\lambda)_{\lambda\in\Lambda}$ is a Gabor frame if and only the Gabor frame operator $S : f \mapsto \sum_{\lambda\in\Lambda}\langle f, g_\lambda\rangle g_\lambda$ is invertible, or if and only if there exists a dual WH-family of the form $(\tilde{g}_\lambda)_{\lambda\in\Lambda}$ with a generator $\tilde{g}$, characterized either as the solution of the frame equation $\mathbf{S}\tilde{g} = g$, or equivalently $\tilde{g} = \mathbf{S}^{-1}g$. There are many other possible (non-canonical) dual functions γ, yielding perfect reconstruction, which are characterized according to [118] by the so-called bi-orthogonality relation.

$$V_g\gamma(\lambda^\circ - \mu^\circ) = \langle \pi(\mu^\circ)\gamma, \pi(\lambda^\circ)g \rangle = \delta_{\lambda^\circ,\mu^\circ}, \quad \text{for} \quad \lambda^\circ, \mu^\circ \in \Lambda^\circ. \tag{32}$$

The so-called *Ron-Shen duality* gives more detailed information (cf. [105, 53]): The condition number of the Gabor frame $(g_\lambda)_{\lambda\in\Lambda}$ is the same as condition number of the Gabor Riesz (basic) sequence $(g_{\lambda^\circ})_{\lambda^\circ\in\Lambda^\circ}$, with explicit constants (going back to a symplectic version of Poisson's formula) relating upper and lower frame bounds. This result has a great impact for applications in communication theory. While one tries to use (preferably tight and) *low redundancy* Gabor frames with good localization properties in order to expand signals, avoiding the storage of too many coefficients for the Gabor expansion, one is interested to use Gaborian Riesz bases for the transmission of data, because the well chosen Gabor atoms (obtained using beam-shaping) g ensure that the family $(g_\lambda)_{\lambda\in\Lambda}$ consists of *joint approximate* eigenvectors to all *underspread* resp. *slowly varying* linear systems, i.e. linear operators which have a spreading function supported by a small rectangle in $\mathbb{R}^d \times \widehat{\mathbb{R}}^d$, determined by the maximal time-delay and Doppler shift respectively (see [74]). Ground breaking work in this direction has been done in the PhD thesis [90] of W. Kozek; the link to (S_0, L^2, S_0') has been established in [53].

It is one of the striking recent results due to Gröchenig and Leinert ([75], following the rational case in [50]) to show that $g \in S_0(\mathbb{R}^d)$ implies also that the canonical dual $\tilde{g}$ is in $S_0(\mathbb{R}^d)$, or equivalently (because the frame operator associated with the Gabor system $(\pi(\lambda)\tilde{g})_{\lambda\in\Lambda}$ is just $\mathbf{S}^{-1}$, the inverse of the frame operator for the WH-family $(\pi(\lambda)g)_{\lambda\in\Lambda}$. Expressed in terms of BGT-morphisms their result can be rephrased as follows. The boundedness part of the theorem below is given in detail in [61]:

Theorem 4. *Assume that $(g_\lambda)_{\lambda\in\Lambda}$ be a Gabor frame with $g \in S_0(\mathbb{R}^d)$, and hence $S = S_{g,\Lambda} : f \mapsto \sum_{\lambda\in\Lambda}\langle f, \pi(\lambda)g\rangle g_\lambda$ is a a BGT-morphism on (S_0, L^2, S_0') into itself. If S is invertible at the L^2-level, then it is already BGT-isomorphism.*

In particular, $\tilde{g} = S^{-1}(g)$ is in $S_0(\mathbb{R}^d)$ in this case and

$$f = \sum_{\lambda\in\Lambda} V_{\tilde{g}} f(\lambda) g_\lambda = \sum_{\lambda\in\Lambda} V_g f(\lambda)\tilde{g}_\lambda. \tag{33}$$

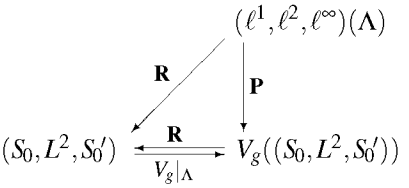
We will call the corresponding families S_0-Gabor families. Another result where the BGT-spirit comes through and the relevance of considering Gabor problems at all three levels is evident can be found in [71] on "Gabor frames without inequalities".

With this background we can give a characterization of elements in each of the levels of (S_0, L^2, S_0') in terms of Gabor coefficients:

Theorem 5. *Let $g \in \mathscr{S}(\mathbb{R}^d)$ be given such that $(g_\lambda)_{\lambda\in\Lambda}$ is a Gabor frame with canonical dual $(\tilde{g}_\lambda)_{\lambda\in\Lambda}$ (also in $\mathscr{S}(\mathbb{R}^d)$). Then one has: A tempered distribution $f \in \mathscr{S}'(\mathbb{R}^d)$ belongs to (S_0, L^2, S_0') if and only if the following (equivalent!) conditions are satisfied:*

1. *f has a representation of the form[15] $f = \sum_{\lambda\in\Lambda} c_\lambda g_\lambda$, with $(c_\lambda)_{\lambda\in\Lambda}$ from $(\ell^1, \ell^2, \ell^\infty)(\Lambda)$;*
2. *The canonical coefficients $(V_{\tilde{g}} f(\lambda))_{\lambda\in\Lambda} \in (\ell^1, \ell^2, \ell^\infty)(\Lambda)$;*
3. *The sampled STFT with window g satisfies: $(V_g f(\lambda))_{\lambda\in\Lambda} \in (\ell^1, \ell^2, \ell^\infty)(\Lambda)$;*

Overall this can be expressed by the fact that the reconstruction mapping $\mathbf{R} : (c_\lambda)_{\lambda\in\Lambda} \mapsto \sum_{\lambda\in\Lambda} c_\lambda \tilde{g}_\lambda$ completes the following diagram:



Poisson's Formula, Sampling and Periodization

Using $S_0(\mathbb{R}^d)$ the classical Poisson's formula can be formulated as follows:

Theorem 6. *For $f \in S_0(\mathbb{R}^d)$ one has*

$$\sum_{k\in\mathbb{Z}^d} f(k) = \sum_{n\in\mathbb{Z}^d} \hat{f}(n), \tag{34}$$

the sum being absolutely convergent on both sides.

[15] Here $*$ denotes convolution, in contrast to the use of $*$ earlier on, where it was representing matrix multiplication, using MATLAB conventions.

This formula does not hold for arbitrary functions, even if both the left hand side and the right hand side are absolutely convergent, as has been described in the book of Katznelson ([87, 88]). Most of the usual conditions on f which are sufficient for the validity of (34) can be interpreted as sufficient conditions for f to belong to $S_0(\mathbb{R}^d)$ (cf. [84, 68]). The symplectic version of Poisson relation is also highly relevant for Gabor analysis ([8, 112]).

The key properties of $S_0(\mathbb{R}^d)$ needed to verify Thm. 6 are the fact that the restriction of a function $f \in S_0(\mathbb{R}^d)$ is in $\ell^1(\mathbb{Z}^d)$, that the $\mathbb{Z}^d$-periodization of f is uniformly convergent, and the fact that the periodized function f_{per} has as its Fourier coefficients just the samples $(\hat{f}(n))$, which are again in $\ell^1(\mathbb{Z}^d)$, due to the Fourier invariance of $S_0(\mathbb{R}^d)$.

It is an easy exercise to translate the Poisson formula into a statement about the Fourier invariance of the so-called Shah-distribution $\text{ш}_{\mathbb{Z}^d}$ (also called Dirac Comb, etc.):

Theorem 7. *The Shah-distribution $\text{ш}_{\mathbb{Z}^d}$ belongs to $S_0'(\mathbb{R}^d)$, and*

$$\widehat{\text{ш}}_{\mathbb{Z}^d} = \text{ш}_{\mathbb{Z}^d}$$

Using the invariance of $S_0(\mathbb{R}^d)$ under transformation of the argument it is easily extended to other lattices of the form $\Lambda = A * \mathbb{Z}^d \lhd \mathbb{R}^d$, where $det(A) \neq 0$. For the sake of simplicity we will use ordinary dilation, which gives then $\text{ш}_a = \sum_{k\in\mathbb{Z}^d} \delta_{ak}$, which has as its (generalized) Fourier transform $b\text{ш}_b$, with $b = 1/a$.

One of the most important principles in harmonic analysis is the idea that sampling on the "time-side" corresponds to periodization on the frequency side. The most important consequence of this principle is the so-called Shannon sampling theorem, according to which a band-limited signal can be recovered from its regular or *equidistant* samples. Again, we do the detailed discussion only for the normalized case, i.e. for the case that the *Nyquist sampling rate* is the sampling over the integer lattice $\mathbb{Z}^d$, or in other words, that the spectrum (the support of the Fourier transform of $\hat{f}$ under discussion) is contained in the cube $Q := [-1/2, 1/2]^d$. We write $\mathbf{1}_Q$ for the indicator function of Q and define $\text{SINC} = \mathscr{F}^{-1}(\mathbf{1}_Q)$.

Theorem 8. *[Shannon Sampling Theorem]*
For any $f \in L^2(\mathbb{R}^d)$ with $\text{supp}(f) \subseteq Q$ one has

$$f(t) = \sum_{n\in\mathbb{Z}^d} f(n) T_n \text{SINC}(t), \tag{35}$$

with absolute and uniform convergence of the series and norm convergence in $L^2(\mathbb{R}^d)$.

The proof is based on the observation that the family $(T_n \text{SINC})_{n\in\mathbb{Z}^d}$ is an orthonormal basis for the closed subspace $B^Q := \{f \in L^2(\mathbb{R}^d) \mid \text{supp}(\hat{f}) \subseteq Q\}$ of $\left(L^2(\mathbb{R}^d), \|\cdot\|_2\right)$. In fact, one has convergence with respect to the Wiener amalgam norm $W(C_0, \ell^2)(\mathbb{R}^d)$, which implies both uniform and L^2-convergence. The fact that the SINC function is the (inverse) Fourier transform of the indicator function $\mathbf{1}_Q$, which is a Fourier multiplier for

$1 < p < \infty$ also implies, that a similar statement holds true for band-limited functions in $L^p(\mathbb{R}^d)$, for the same range of parameters.

Proof. Given the sampling values $(f(n))_{n\in\mathbb{Z}^d}$ we have all the information in our hands to describe $\sqcup\!\sqcup_{\mathbb{Z}^d} \cdot f = \sum_{n\in\mathbb{Z}^d} f(n)\delta_n$. This is a well defined (unbounded but translation-bounded) discrete measure in $S_0'(\mathbb{R}^d)$ which has a Fourier transform of the form

$$\widehat{\sqcup\!\sqcup_{\mathbb{Z}^d} \cdot f} = \widehat{\sqcup\!\sqcup_{\mathbb{Z}^d}} * \hat{f} = \sum_{k\in\mathbb{Z}^d} \delta_n * \hat{f} = \sum_{k\in\mathbb{Z}^d} T_n \hat{f} \tag{36}$$

which is nothing but the $\mathbb{Z}^d$-periodic version of $\hat{f}$. We can now use the fact that Q is a fundamental domain for the lattice $\mathbb{Z}^d \lhd \mathbb{R}^d$, hence $|n+Q \cap k+Q| = 0$ for $n \neq k$. Multiplying this periodic version by $\mathbf{1}_Q$ gives us exactly the original basic period, which is $\hat{f}$, or back on the time domain

$$f = (\sqcup\!\sqcup_{\mathbb{Z}^d} \cdot f) * \mathscr{F}^{-1}(\mathbf{1}_Q) = \sum_{n\in\mathbb{Z}^d} f(n)\delta_n * \mathrm{SINC} = \sum_{n\in\mathbb{Z}^d} f(n) T_n \mathrm{SINC}. \tag{37}$$

This series is convergent in $L^2(\mathbb{R}^d)$ because on the Fourier transform side we just have the Fourier expansion of $\hat{f}$ (taken as a periodic function on $\mathbb{R}^d$). On the other hand SINC belongs to $L^2(\mathbb{R}^d)$ and even the Wiener amalgam space $W(C_0, \ell^2)(\mathbb{R}^d)$, which implies uniform and pointwise absolute convergence. □

BANACH GELFAND TRIPLES AND OPERATORS

In this section we will indicate the role of BGTs for the description of operators. The same role which is played by the pure frequencies for Fourier analysis (they are perfect building blocks forming an orthonormal basis in the case of finite Abelian groups but fail to belong to the natural Hilbert space) is now taken by other systems of natural objects. From the point of time-frequency analysis of course the collection of $(\pi(\lambda))_{\lambda\in\mathbb{R}^d\times\widehat{\mathbb{R}}^d}$ is a very natural choice, but again they are *not* in the natural Hilbert space, now $\mathscr{HS}$, the space of all Hilbert Schmidt operators on $L^2(\mathbb{R}^d)$, endowed with the $\mathscr{HS}$-scalar product $\langle T, S\rangle_{\mathscr{HS}} := \mathrm{Tr}(TS^*)$.

On the other hand one of the most exciting developments in the field is the realization, that pseudo-differential operators have a very natural description in terms of time-frequency expressions. To give an example: modulation spaces turn out to be the most natural spaces in order to describe *slowly varying channels*, i.e. convolution operators with a *time-variant kernel* (in an engineering terminology), resp. certain classes of pseudo-differential operators. These are the systems which preserve localization in the TF-sense and hence have a matrix representaion which is mostly concentrated along the diagonal. In the extreme case on has Gabor multipliers, i.e. operators which are factorized through a diagonal matrix, acting on the Gabor coefficients.

There is a large number of papers on Gabor multipliers, such as [57, 12, 38, 5] and a self-contained survey (master thesis) by K. Schnass [107] from 2004 or the PhD thesis of P. Balazs ([9]).

AntiWick operators are operators which are defined as STFT-multipliers. They make use of the inversion formula for the STFT (which in turn based on the isometric properties of $f \mapsto V_g(f)$ from $L^2(\mathbb{R}^d)$ into $L^2(\mathbb{R}^{2d})$:

$$\int_{\mathbb{R}^d \times \widehat{\mathbb{R}}^d} (V_{g_1} f)(x,\omega) M_\omega T_x g_2 \, dx \, d\omega = \langle g_2, g_1 \rangle f \tag{38}$$

So one can use any $g_2 \in L^2(\mathbb{R}^d)$ for reconstruction of f from V_{g_1} as long as it is not orthogonal to g_1. Usually the integral has to be understood in the *weak sense* for $g_1, g_2 \in L^2(\mathbb{R}^d)$, but if both of them are in $S_0(\mathbb{R}^d)$ (cf. [117]) then one can even read the above integral as limit of vector-valued Riemannian integrals, which are norm convergent in $\mathscr{H} = L^2(\mathbb{R}^d)$. Due to the good local properties of functions in the range of the STFT one can even work with rough symbols (see [93, 13, 28])

Adjointness Relations

First of all let us mention some principles that allow us to extend bounded linear mappings betwenn S_0-spaces to BGT-morphisms. The following principle is quite useful in order to "automatically extend" a mapping between the "inner spaces" to their dual spaces.

Theorem 9. *Let T be a BGT-homomorphism from $(B_1, \mathscr{H}_1, B_1')$ into $(B_2, \mathscr{H}_2, B_2')$, i.e. a bounded linear mapping which is bounded on all three layers, as well as $w^* - w^*$-continuous. Then there exists a unique adjoint GT-homomorphism, i.e. another BGT-homomorphism (denoted by) T^* from $(B_2, \mathscr{H}_2, B_2')$ into $(B_1, \mathscr{H}_1, B_1')$, such that $T^* : \mathscr{H}_2 \mapsto \mathscr{H}_1$ is the adjoint operator, which extends to a GT-morphism in a unique way. Therefore we have the identity*

$$\langle Tf, g \rangle_{(B_2, \mathscr{H}_2, B_2')} = \langle f, T^* g \rangle_{(B_1, \mathscr{H}_1, B_1')} \tag{39}$$

whenever the pairing makes sense. Moreover $T^{**} = T$, i.e. in this sense any BGT-morphism is the adjoint of another (uniquely determined) adjoint BGT-morphism, denoted by T^*.

The case of unitary operators has been discussed already in [53], Thm.7.3.3 (Extension of Unitary Gelfand Triple Isomorphism), p. 239.

Theorem 10. *A unitary mapping U acting from $L^2(\mathbb{R}^d)$ to $L^2(\mathbb{R}^d)$ extends to an isomorphism between the Gelfand triples $(S_0, L^2, S_0')(\mathbb{R}^d)$ to $(S_0, L^2, S_0')(\mathbb{R}^d)$ if and only if the restriction of U and also of its adjoint U^* are bounded linear operators from $S_0(\mathbb{R}^d)$ to $S_0(\mathbb{R}^d)$.*

Remark 1. There are good reasons why the "central" Hilbert space $\mathscr{H} = L^2(\mathbb{R}^d)$ usually plays the dominant rule, just think of Plancherel's theorem as the central property of the Fourier transform, describing it as a *unitary mapping* on $L^2(\mathbb{R}^d)$. However, from an abstract point of view it is not so important, and in most cases the isomorphism property at the S_0 and S_0'-level (both with the norm and the w^*-topology) implies already that one

has the full continuity claim for the Hilbert space automatically, becausean application of complex interpolation between the dual pair $S_0(G)$ and $S_0'(G)$ yields[16]

$$[S_0, S_0']_{[1/2]} = L^2.$$

Theorem 11. *For any group $(T_i)_{i\in I}$ of unitary operators leaving $S_0(\mathbb{R}^d)$ invariant, which satisfy*

$$\langle f, g\rangle = \langle T_i f, T_i g\rangle \quad \forall f, g \in S_0(\mathbb{R}^d), \tag{40}$$

one has: The action of $(T_i)_{i\in I}$ extends in a unique way to a unitary Banach Gelfand triple automorphism of $(S_0, L^2, S_0')(\mathbb{R}^d)$.

Proof. The assumption (40) implies (just in the same way as Plancherel's theorem is usually proved) that it is well defined and isometric on $S_0(\mathbb{R}^d)$ with respect to the L^2-norm. Due to the density of $S_0(\mathbb{R}^d)$ in $L^2(\mathbb{R}^d)$ it can be uniquely extended to an isometric and in fact unitary automorphism on $L^2(\mathbb{R}^d)$. □

Kernel Theorems and Gelfand Triples

The *nuclear* Frechet space $\mathscr{S}(\mathbb{R}^d)$ and its dual, the space of $\mathscr{S}'(\mathbb{R}^d)$ of tempered distributions are the prototype of function spaces for which on can prove a so-called kernel theorem, a continuous analogue of the existence of a matrix, completely describing the operator. We next prepare a similar principle for our BGT-setting.

Given two functions f^1 and f^2 on $\mathbb{R}^d$ respectively, we set $f^1 \otimes f^2$

$$f^1 \otimes f^2(x_1, x_2) = f^1(x_1) f^2(x_2),\ x_1, x_2 \in \mathbb{R}^d. \tag{41}$$

Given two Banach spaces B_1 and B_2 embedded into $\mathscr{S}'(\mathbb{R}^d)$, $B_1 \hat{\otimes} B_2$ denotes their *projective tensor product*, i.e.

$$\left\{ f \mid f = \sum f_n^1 \otimes f_n^2,\ \sum \|f_n^1\|_{B_1} \|f_n^2\|_{B_2} < \infty \right\}. \tag{42}$$

It is easy to show that this defines a Banach space of tempered distributions on $\mathbb{R}^{2d}$ with respect to the (quotient) norm:

$$\|f\|_{\hat{\otimes}} := \inf \left\{ \sum \|f_n^1\|_{B^1} \|f_n^2\|_{B^2}, \ldots \right\}, \tag{43}$$

where the infimum is taken over all admissible representations.

One of the most important properties of $S_0(\mathbb{R}^d)$ (leading to a characterization given by V. Losert, [94]) is the tensor-product factorization:

Lemma 4.

$$S_0(\mathbb{R}^k) \hat{\otimes} S_0(\mathbb{R}^n) \cong S_0(\mathbb{R}^{k+n}). \tag{44}$$

---

[16] One way to understand/accept this fact is to invoke the fact that the Wilson bases establish, at least for elementary locally compact Abelian groups, a BGT-isomorphism between (S_0, L^2, S_0') and $(\ell^1, \ell^2, \ell^\infty)$.

The easiest way to realize this relationship is to make use of the atomic decomposition of $S_0(\mathbb{R}^m)$, observing that both time- and frequency shifts, but also the multi-dimensional Gaussian function factorize into lower dimensional partial ingredients. This tensor product property of S_0 is on the other hand the basis for the realization of the so called kernel theorem (see [53], Chap.7.4).

The setting of BGTs is also well suited for the description of linear operators. The so-called *kernel theorem* shows how essentially every reasonable[17] operator T can be interpreted as a kind of *integral operator*, also called *kernel* from a suitable class of generalized functions. One may expect similarity to the finite discrete case (= matrix multiplication), but now with continuous variables:

$$Tf(x) = \int_{\mathbb{R}^d} K(x,y)f(y)dy \tag{45}$$

in analogy to the description of $u = Tf \simeq u = A * z \in \mathbb{C}^m$ via coordinates

$$(Tz)_s = \sum_{k=1}^{n} a_{s,k}z_k, \quad \text{for} \quad s = 1,\dots n. \tag{46}$$

The usual way of finding the appropriate $m \times n$ matrix A for a linear mapping from $\mathbb{C}^n$ to $\mathbb{C}^m$ is easy: recall that one obtains coordinate number s of the vector u via scalar product with e_s, that the k-th column of A has to correspond to $T(e_k)$ if (46) is supposed to be valid, hence the individual entry must be

$$a_{s,k} = A[s,k] = \langle T(e_k, e_s)\rangle. \tag{47}$$

Viewing $A = (A(n,k))$ as a function over the product index set one can say that has to take the scalar product of A (in the sense of the Euclidean space $\mathbb{C}^{m\dot{n}}$) with the unit matrix $e_s * \otimes e_k$, which can also be expressed as via a trace formula of operators:

$$A[s,k] = \langle A, e_s \otimes e_k^* \rangle = \mathrm{Tr}(A * (e_k * e_s^*)).^{18} \tag{48}$$

There is a lot of literature about Dirac's formalism. On the one hand it is very intuitive, on the other hand it has created a lot of discussion concerning the strict mathematical interpretation of these formal symbols. Even engineers are by now aware of the fact δ_0 is no just another function, which is zero everywhere except at 0, but "so strongly infinite" that the integral equals 1. Nowadays it is well known that the Dirac *measure* $\delta_x : f \mapsto f(x)$ is a good way to formalize this procedure, but this still does not explain what the connection between Kronecker's δ, usually written as $\delta_{i,j}$ and Dirac's symbol which is nowadays just a distribution (of *one* variable, so to say) applied on a test function, while in many early interpretations of what Dirac might have had in mind with his symbol the idea of using the symbols he had introduced often comes with the recommendation of only using it within an integral, and not as an individual object. In

[17] E.g. T is a bounded linear operator from some L^q-space to another L^p-space.
[18] Where $*$ corresponds to matrix multiplication in a MATLAB setting.

fact, we will argue that one should consider δ_{x-y} as a distributional kernel representing the identity mapping, so simply the continuous analogue of Kronecker's symbol.

Moreover it appears as a way to express a kind of orthogonality relations between "pure frequencies" (and because they are not square-integrable the scalar product between $\langle \chi_s, \chi_s \rangle$ has to be $+\infty$) which allows to do derivations of very useful formulas[19].

Sequences of regularizing operators

Once we have continuous variables one comes into a world where finite dimensional arguments break down, where one may have unbounded operators, and even point evaluations are not always possible, i.e. the use of e.g. Dirac distributions is required. Nevertheless one has a number of different products, which often can be written as integrals (convolution, twisted convolution) or using point values (defining the "ordinary pointwise product"), and sometimes such products immediately make sense, in some other case on has first approximate the involved ingredients before applying the operation, using a regularized version of one or both partners involved, and then let the regularization parameter tend to 0 or ∞, as appropriate.

Such a principle is not really new, as many special cases can be located in the literature. The definition of the Fourier transform is one case, where one has to "push" general $L^2(\mathbb{R}^d)$-functions into $L^1(\mathbb{R}^d)$ (in case they are not already within $L^1 \cap L^2(\mathbb{R}^d)$, which is fortunately a dense subspace of $\big(L^2(\mathbb{R}^d), \|\cdot\|_2\big)$, e.g. via pointwise multiplication with the indicator function $\mathbf{1}_{[-N,N]}$, for $N \to \infty$. For the inversion of the Fourier transform a similar strategy can be applied, now by doing a pointwise multiplication with some suitable summability kernel. Although it would again be enough to use any localizing function, it has been realized that a sharp frequency cut-off is not a good way, since $\mathscr{F}^{-1}\mathbf{1}_{[-N,N]} \notin L^1(\mathbb{R}^d)$. Choosing a summability kernel from $S_0(\mathbb{R}^d)$ will help and ensure that its inverse Fourier transform is in $L^1(\mathbb{R}^d)$ as well. Since stretching in Fourier space is the same as L^1-norm preserving dilation the resulting sequence of Dirac-like convolution kernels is an approximate identity for the Banach Gelfand triple (S_0, L^2, S_0'), while the SINC-function is not having this good property.

Wiener amalgam convolution and pointwise multiplier results ([78]) imply that

$$S_0(\mathbb{R}^d) \cdot (S_0'(\mathbb{R}^d) * S_0(\mathbb{R}^d)) \subseteq S_0(\mathbb{R}^d), \quad S_0(\mathbb{R}^d) * (S_0'(\mathbb{R}^d) \cdot S_0(\mathbb{R}^d)) \subseteq S_0(\mathbb{R}^d) \tag{49}$$

Proof. The key arguments for both of these regularization procedures, be it convolution followed by pointwise multiplication (a so-called *product-convolution operator*, for short PC-operator), or corresponding CP-operators, are based on the pointwise and convolutive behaviour of generalized Wiener amalgam spaces, such as the relation $S_0(\mathbb{R}^d) * S_0'(\mathbb{R}^d) = W(\mathscr{F}L^1, \ell^1) * W(\mathscr{F}L^\infty, \ell^\infty) \subseteq W(\mathscr{F}L^1, \ell^\infty)$. □

[19] We suggest to view the well known identity $e^{2\pi i} = 1$ in a similar way, as an extremely useful formula which makes use of the complex numbers, the irrational number π, which is never explicitly and constructively realized, let alone the power series expression of the exponential function.

Let now $h \in \mathscr{F}L^1(\mathbb{R}^d)$ be given with $h(0) = 1$. Then the dilated version $h_n(t) = h(t/n)$ are a uniformly bounded family of multipliers on (S_0, L^2, S_0'), tending to the identity operator in a suitable way. Similarly, the usual Dirac sequences, obtained by compressing a function $g \in L^1(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d} g(x)dx = 1$ are showing a similar behavior: $g_n(t) = n \cdot g(nt)$

Following the above rules the combination of the two procedures, i.e. product-convolution or convolution-product operators of the form provide suitable regularizers: $A_n f = g_n * (h_n \cdot f)$ or $B_n f = h_n \cdot (g_n * f)$.

Following Theorem 5 we know that elements in $f \in (S_0, L^2, S_0')$ can be characterized (among others) by their minimal norm coefficients, given in the form $(V_{\tilde{g}} f(\lambda))_{\lambda \in \Lambda}$. It is therefore clear that the partial sum operators for this canonical Gabor expansions, such as

$$A_N f := \sum_{max(|\lambda_1|,|\lambda_2|) \leq N} V_{\tilde{g}} f(\lambda) g_\lambda \tag{50}$$

is mapping $S_0'(\mathbb{R}^d)$ into $S_0(\mathbb{R}^d)$, while on the other hand one has obviously that $A_N f \to f$ as $N \to \infty$ for any f in $S_0(\mathbb{R}^d)$ or $L^2(\mathbb{R}^d)$ respectively, in the corresponding norm, while the convergence occurs in the w^*-sense, for all $f \in S_0'(\mathbb{R}^d)$.

Similar statements can be made for rectangular or any other kind of'exhausting partial sums, also with respect to Wilson basis. The better the building blocks are (in terms of time-frequency localization, typically expressed using membership in the modulation spaces $M^1_{v_s}(\mathbb{R}^d)$) the more can be said about the rate of approximation, given the quality of the signal, i.e. speed of approximation of f in some Shubin class $Q_s(\mathbb{R}^d)$, measured in the L^2-norm.

Various types of regularizations are also used in the discussion about the most general definition of convolution between distributions, see the work of M. Oberguggenberger ([1, 18]). In fact, one can say, that the basic idea is to assume the the limit of $A_N\sigma_1 * A_N\sigma_2$ exists (for a sufficiently rich class of regularization operators, implying that this limit is then independent of the particular choice of the sequence (A_N)).

Kernel Theorem for $S_0(\mathbb{R}^d)$

There are many different ways to show that the space of test functions $S_0(\mathbb{R}^d)$ is w^*-dense in $S_0'(\mathbb{R}^d)$. One very important and natural way (also valid in a similar way for the space of Schwartz test functions from the space $\mathscr{S}(\mathbb{R}^d)$ of rapidly decreasing functions and its corresponding dual space, $\mathscr{S}'(\mathbb{R}^d)$ of tempered distributions)

Theorem 12. *If K is a bounded operator from $S_0(\mathbb{R}^d)$ to $S_0'(\mathbb{R}^d)$, then there exists a unique kernel $k \in S_0'(\mathbb{R}^{2d})$ such that $\langle Kf, g\rangle = \langle k, g \otimes f\rangle$ for $f, g \in S_0(\mathbb{R}^d)$, where $g \otimes f(x,y) = g(x)f(y)$.*

Formally sometimes one writes by "abuse of language"

$$Kf(x) = \int_{\mathbb{R}^d} k(x,y)f(y)dy \tag{51}$$

with the understanding that one can define the action of the functional $Kf \in S_0'(\mathbb{R}^d)$ as

$$Kf(g) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} k(x,y)f(y)dyg(x)dx = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} k(x,y)g(x)f(y)dxdy. \tag{52}$$

This result is the "outer shell of the Gelfand triple isomorphism, which corresponds to the well-known result that Hilbert Schmidt operators on $L^2(\mathbb{R}^d)$ are just those compact operators which arise as integral operators with $L^2(\mathbb{R}^{2d})$-kernels. The complete picture can again be expressed by a unitary Gelfand triple isomorphism. Let us start with the classical setting: The Hilbert space $\mathscr{HS}$ of Hilbert Schmidt operators on $L^2(\mathbb{R}^d)$ is important, because the classical *kernel theorem* establishes a unitary mapping between operators $T \in \mathscr{HS}$ and their kernels K. The scalar product of $\mathscr{HS}-$operators is given by $\langle T,S\rangle_{\mathscr{HS}} = \mathrm{Tr}(T * S')$ and turns $\mathscr{HS}$ into a Hilbert space.

Theorem 13 (Kernel Theorem for S_0)**.** *Let $T \in \mathscr{HS}$ be given, with kernel $K \in L^2(\mathbb{R}^{2d})$. Such an operator has a kernel in $S_0(\mathbb{R}^{2d})$ if and only if it maps bounded, w^*-convergent sequences in $S_0'(\mathbb{R}^d)$ into norm convergent in $S_0(\mathbb{R}^{2d})$. The most general operators from $\mathscr{L}(S_0,S_0')$ are in a one-to-one correspondence with $S_0'(\mathbb{R}^{2d})$.*

Overall the kernel theorem allows us to establish a unitary BGT isomophism between the BGT $(\mathscr{L}(S_0',S_0),\mathscr{HS},\mathscr{L}(S_0,S_0'))$ of operator spaces and the corresponding kernels in $(S_0,L^2,S_0')(\mathbb{R}^{2d})$.

Remark 2. Note that for *regularizing* kernels in $S_0(\mathbb{R}^{2d})$ the usual identification (recall that the entry of a matrix $a_{n,k}$ is the coordinate number n of the image of the $k-$th unit vector under that action of the matrix $A = (a_{n,k})$) holds:

$$k(x,y) = K(\delta_y)(x) = \delta_x(K(\delta_y)). \tag{53}$$

Since $\delta_y \in S_0'(\mathbb{R}^d)$ and thus $K(\delta_y) \in S_0(\mathbb{R}^d)$ the pointwise evaluation makes sense.

Remark 3. It is of course interesting to ask how the w^*-topology can be transferred to the operator level. Here again a characterization of general linear operators using Gabor expansions comes into the picture:

Definition 9. Assume that $(g_\lambda)_{\lambda\in\Lambda}$ and $(\tilde{g}_\lambda)_{\lambda\in\Lambda}$ is a dual pair of Gabor frames, with $g,\tilde{g} \in S_0(\mathbb{R}^d)$, and assume $T \in \mathscr{L}(S_0,S_0')$, i.e. that T is a bounded linear operator from $S_0(\mathbb{R}^d)$ into $S_0'(\mathbb{R}^d)$. Then the *matrix elements* of T with respect to the Gabor frame are

$$a_{\lambda,\lambda'} := \langle Tg_\lambda,\tilde{g}_{\lambda'}\rangle, \quad \lambda,\lambda' \in \Lambda. \tag{54}$$

Using these matrix coefficients (one can use either g or $\tilde{g}$), both in the first or the second place, on obtains

Lemma 5. *Let T_n be a sequence of operators from $S_0(\mathbb{R}^d)$ into $S_0'(\mathbb{R}^d)$, such that the corresponding kernels K^n form a bounded sequence in $S_0'(\mathbb{R}^{2d})$, convergent to K^0 in the w^*-sense. Then T_nf is w^*-convergent for every $f \in S_0(\mathbb{R}^d)$ to some limiting operator $T^0(f) = \lim_n T_nf$ and conversely. In particular, $K^0 = w^* - \lim_n K_n$ if and only if all the matrix coefficients converge pointwise, i.e. for each pair $(\lambda,\lambda') \in \Lambda\times\Lambda$ one has*

$$a^n_{\lambda,\lambda'} \to a^0_{\lambda,\lambda'} \quad \text{for } n \to \infty.$$

Kohn-Nirenberg Symbol and Spreading Function

The *Kohn-Nirenberg symbol* of an operator (respectively its *symplectic* Fourier transform, the so-called *spreading symbol*) can be obtained from the kernel by applying suitable coordinate transforms (automorphisms) and partial Fourier transforms [20]. Hence they define functions or distributions over $\mathbb{R}^d \times \widehat{\mathbb{R}}^d$. Since all these ingredients are unitary BGT isomorphisms of (S_0, L^2, S_0') the known correspondences at the level of $\mathscr{HS}$-operators can be extended to BGT isomorphisms.

Theorem 14. *The correspondence between an operator T with kernel K from the Banach Gelfand triple $(\mathscr{L}(S_0', S_0), \mathscr{HS}, \mathscr{L}(S_0, S_0'))$ and the corresponding spreading distribution $\eta(T)$ in $S_0'(\mathbb{R}^{2d})$ is the uniquely defined Gelfand triple isomorphism between $(\mathscr{L}(S_0', S_0), \mathscr{HS}, \mathscr{L}(S_0, S_0'))$ and $(S_0, L^2, S_0')(\mathbb{R}^d \times \widehat{\mathbb{R}}^d)$ which maps the time-frequency shift operators $M_y \circ T_x$ onto the Dirac measure $\delta_{(x,y)}$.*

The w^*-continuity of this mapping allows among others to calculate (in the sense of approximate) $\eta(T)$ by first dealing with regularizing operators from $\mathscr{L}(S_0', S_0)$ with kernels and symbols in S_0. For this "core" space one can apply transformations and partial Fourier transform in a direct way, while more general case is realized either by taking w^*-limits of using an adjointness argument.

The Kohn-Nirenberg description of operators is particularly interesting in the discussion of *Gabor multipliers*, i.e. of operators of the form

$$Tf = \sum_{\lambda \in \Lambda} m_\lambda \langle f, \pi(\lambda) g \rangle g_\lambda = \sum_{\lambda \in \Lambda} m_\lambda P_\lambda(f), \tag{55}$$

where P_λ is the projection of f onto the one-dimensional space generated by g_λ. Equivalently, $P_\lambda = \pi(\lambda) \circ P_0 \pi(\lambda)'$. The mapping $H \mapsto \pi(\lambda) \circ P_0 \pi(\lambda)'$ is a unitary group representation of the additive group $\mathbb{R}^{2d}$ on the Hilbert space $\mathscr{HS}$, and one crucial facts is the relation

$$\kappa[\pi(\lambda) \circ P_0 \pi(\lambda)'] = T_\lambda \kappa(H), \quad \lambda \in \Lambda. \tag{56}$$

Composition of Operators

Given the kernel representation (or whatever other form of "symbol", from Weyl- to Kohn-Nirenberg or spreading representation) it is clear that the composition of operators corresponds to some kind of composition rule at the level of symbols. For the case of matrices we *know* that we have to perform matrix multiplication, i.e. the *matrix-product* $C := A * B$ is given (coefficientwise) by the rule

$$c_{k,l} = \sum_{s=1}^{n} a_{k,s} b_{s,l} \tag{57}$$

[20] i.e. Fourier transforms with respect to one variable only with $\mathbb{R}^d \times \mathbb{R}^d$.

whenever the matrix product is possible, resp. whenever the composition of operators is possible (the range of B has to be equal to the domain of A, in our example $\mathbb{C}^m$).

Of course the situation is - from a purely technical point - much more delicate in the case of (perhaps even unbounded) linear mappings between (infinite) dimensional vector spaces (typically Hilbert spaces or Banach spaces), and even if we are next discussing the composition of BGT-morphism it is not absolutely clear how to interpret their composition (which is kind of obvious from the point of view of operators).

Let us therefore consider first the composition of two simple integral operators, with the corresponding kernels $K_2(x,s)$ and $K_1(s,y)$ in $S_0(\mathbb{R}^{2d})$. It is not difficult to verify that one has in such a case, in complete analogy to case of matrix multiplication:

Lemma 6. *The composition of two operators $T_2 \circ T_1$, both of which have a kernel representation with $S_0(\mathbb{R}^{2d})$-kernels $K_2(x,s)$ and $K_1(s,y)$ respectively, has a kernel in $S_0(\mathbb{R}^{2d})$ of the form*

$$K(x,y) = \int_{\mathbb{R}^d} K_2(x,s)K_1(s,y)\,ds \tag{58}$$

This formula is also valid if one of the kernels belongs to $L^\infty(\mathbb{R}^{2d}) \subset S_0'(\mathbb{R}^{2d})$[21].

Proof. The kernels in K_1 and K_2 define a bounded linear operator from $S_0'(\mathbb{R}^d)$ space into $S_0(\mathbb{R}^d)$, converting w^*-convergent sequences into norm convergent sequences. Hence one can compose the operators, but also verify without difficulties (under the L^∞-condition) the existence of the corresponding integrals in (58) [22]. □

For more general cases, e.g. for the composition of general bounded linear operators on $\mathscr{H} = L^2(\mathbb{R}^d)$ it turns out that a composition rule like the simple integral composition of (58) may become questionable. Among others, because it is known to be hard to characterize the L^2-boundedness of the operator T in terms of the kernel $K(x,s)$. Therefore one has to use the approximation of operators by "good" ones before calculating the "product-kernel", i.e. the (distributional) kernel of the composite linear mapping. In order to realize this in a systematic way (admitting that there are many other ways of doing it) we formulate an auxiliary result. It is based on the use of sequences of regularzing operators with kernels in $S_0(\mathbb{R}^{2d})$, i.e. a bounded sequence A_n of BGT-morphism with kernels $K^n(s,u) \in S_0(\mathbb{R}^{2d})$ such that the sequence (as well as its adjoint) acts as an approximation to the identity operator on $S_0'(\mathbb{R}^d)$ (hence on the larger spaces), i.e. satisfies $\|A_n f - f\|_{S_0} \to 0$ for $n \to \infty$, for each $f \in S_0(\mathbb{R}^d)$.

Lemma 7. *For each regularizing sequence $A_n, n \geq 1$ and linear mappings T_1 and T_2 one finds that $A_n \circ T_1$ resp. $T_2 \circ A_n$ are regularizing operators in $\mathscr{L}(S_0', S_0)(\mathbb{R}^d)$. Hence their kernel of composite mappings such as $T_2 \circ A_n \circ A_n \circ T_1$ can be composed according to formula (58), and the product kernels K^n obtained in this way are w^*-convergent to the*

[21] Various properties of the kernel of the composite mapping can be derived from the properties of the resulting product operators. The composition itself need not be "well-defined" in the sense of Lebesgue-integrability almost everywhere. This problem can be overcome using regularization techniques described below.

[22] Here is a warning in place: even if the kernels are given as bounded and continuous functions we do not claim in the most general case that the integration has to make sense in the Lebesgue sense!

kernel of $T_2 \circ T_1$. At the same time the corresponding operators are convergent (in the sense of pointwise convergence) to the action of $T_2 \circ T_1$, because

$$\|T_2(T_1 f) - T_2(A_n(T_1 f))\|_{S_0'} \to 0 \quad \textit{for all} \quad f \in S_0(\mathbb{R}^d). \tag{59}$$

There are of course many variations of this principle, and the concrete form of the regularizing operator can vary from case to case. Again the Fourier transform is a perfect example. We know that the Fourier transform, viewed as an *integral transform* on $\left(L^1(\mathbb{R}^d), \|\cdot\|_1\right)$, maps into $\left(\mathscr{F}L^1(\mathbb{R}^d), \|\cdot\|_{\mathscr{F}L^1}\right) \hookrightarrow \left(C_0(\mathbb{R}^d), \|\cdot\|_\infty\right)$, according to the Riemann-Lebesgue Lemma, as a proper but dense subspace. The problem with Fourier inversion on $\mathscr{F}L^1(\mathbb{R}^d)$ is not the roughness of those functions, but their lack of decay, since they need not be integrable. Since $L^1(\mathbb{R}^d) * S_0(\mathbb{R}^d) \subseteq S_0(\mathbb{R}^d)$ we have the pointwise relationship $\mathscr{F}L^1(\mathbb{R}^d) \cdot S_0(\mathbb{R}^d) \subseteq S_0(\mathbb{R}^d)$, or $\mathscr{F}L^1(\mathbb{R}^d) \hookrightarrow \mathscr{M}(S_0)(\mathbb{R}^d)$ (the pointwise multipliers of $S_0(\mathbb{R}^d)$). Hence it is enough that regularization takes place in the form of pointwise multiplication with any function $h \in S_0(\mathbb{R}^d)$, typically $h_n(t) = h(t/n)$ for $n \to \infty$, with $h(0) = 1$. That indeed all known classical summability kernels are in fact elements of $S_0(\mathbb{R}^d)$ has been investigated in some detail in joint work with F. Weisz ([58, 59, 60]). Of course choices such as the Gauss-Weierstrass kernel $g(t) = e^{-\pi|t|^2}$, the inverse exponential $h(t) = exp(-|t|)$ or $h(t) = 1/(1+t^2)$ on $\mathbb{R}$ come to mind.

FURTHER APPLICATIONS, COMMENTS, OUTLOOK

So far we have only outlined some general principles where the setting of Banach Gelfand triples, and specifically the S_0-BGT come very handy and natural. In the rest of this paper let us just give some indications about further areas where such a setting appears to be quite natural.

Generalized stochastic processes

Already the PhD thesis of A.J.E.M. Janssen [81] indicates that generalized stochastic processes can be modeled appropriately using distribution theoretic methods. His space of test functions did not allow for compactly supported elements, hence he could not define the support of linear functionals in his setting. In this respect the setting of the BGT (S_0, L^2, S_0') is more suitable for a treatment of generalized stochastic processes. We can give only a quick indication of how this works (up to the topic of "spectral representations" of stationary stochastic processes), see [52].

First of all we view a *generalized stochastic process* as a generalization of an ordinary stochastic process, in the sense that an ordinary stochastic process assigns to each $x \in \mathbb{R}^d$ some random variable, abstractly speaking some element $\rho(x)$ in some Hilbert space[23] [of L^2-functions on some measure space, usually with expected value $E(X) = 0$]. As in

[23] Usually it is the set of all square-integrable functions random variables with zero expectation over some probability space, but this does not matter in our description of basic concepts.

the case of regular distributions one can integrate against a test function, i.e. extend the mapping $x \mapsto \rho(x) \in \mathscr{H}$ to a linear mapping

$$k \mapsto \rho(k) := \int_{\mathbb{R}^d} k(x)\rho(x)dx,$$

which is well defined at least for $C_c(\mathbb{R}^d)$, the space of compactly supported, continuous and complex-valued functions on $\mathbb{R}^d$. For us $S_0(\mathbb{R}^d)$ is more attractive as a (Banach) space of test functions and therefore we give the following definition:

Definition 10. We call a bounded linear mapping $\rho : f \mapsto \rho(f)$ from $S_0(\mathbb{R}^d)$ into some Hilbert space $\mathscr{H}$ a *generalized stochastic process*, for short a GSP.

In the standard approach to stochastic processe it is quite cumbersome, at least from the technical point of view, to check the existence of an autocorrelation function (resp. distribution) or to provide the spectral representation of a GSP, using vector-valued measures, such things become quite smooth and natural in our setting:

Definition 11. For any GSP ρ one defines its Fourier transform $\hat{\rho}$ via

$$\hat{\rho}(f) := \rho(\hat{f}), \quad \forall f \in S_0(\mathbb{R}^d). \tag{60}$$

Obviously the inverse Fourier transform of a GSP is defined in an analogous manner, and thus every GSP has a spectral representation in this sense. An important object for GSPs is the autocorrelation of such a process, which is given as follows:

Definition 12. Let ρ be a GSP. The **autocovariance** is characterized via

$$\langle \sigma_\rho, f \otimes g \rangle := \langle \rho(f) | \rho(\bar{g}) \rangle \; \forall f,g \in S_0(\mathbb{R}^d). \tag{61}$$

Theorem 15. *For a GSP ρ the following properties are equivalent:*
a) ρ stationary $\Longleftrightarrow$ σ_ρ diagonally invariant, i.e. $L_{(x,x)}\sigma_\rho = \sigma_\rho \; \forall x \in \mathbb{R}^d$;
b) ρ bounded $\Longleftrightarrow$ σ_ρ extends in a unique way to a bimeasure on $\mathbb{R}^d \times \mathbb{R}^d$;
c) ρ orthogonally scattered
$\Longleftrightarrow$ σ_ρ has support on the diagonal, i.e. $supp(\sigma_\rho) \subseteq \Delta_{\mathbb{R}^d} := \{(x,x) \mid x \in \mathbb{R}^d\}$;
$\Longleftrightarrow$ there exists a positive and translation bounded measure τ_ρ with:

$$\langle \sigma_\rho, f \otimes g \rangle = \langle \tau_\rho, fg \rangle \; \forall f,g \in S_0(\mathbb{R}^d).$$

Corollary 1. *A GSP ρ is bounded and orthogonally scattered if and only if there exists a bounded measure μ_ρ on $\mathbb{R}^d$ such that*

$$\langle \sigma_\rho, f \otimes g \rangle = \langle \mu_\rho, fg \rangle = \int_{\mathbb{R}^d} f(x)g(x)d\mu_\rho(x) \quad \forall f,g \in S_0(\mathbb{R}^d) \tag{62}$$

These statements should only indicate that the BGT (S_0, L^2, S_0') is also very helpful in this context, and therefore likely to be useful in the context of stochastic signal processing, where most often differentiation does not play any role (which in turn would justify using the Schwartz space $\mathscr{S}(\mathbb{R}^d)$ instead), cf. also the PhD thesis of B. Keville ([89]). There is more recent work using such tools by P. Wahlberg ([114, 115]).

Modulation spaces and Coorbit Theory

The Banach Gelfand triple (S_0, L^2, S_0') is just a prototype of the much more general family of modulation spaces, introduced in the early 80's (see ([44, 46])). The by now classical modulation spaces $M_s^{p,q}(\mathbb{R}^d)$ have been modeled in similarity to the family of (inhomogenous) Besov spaces and have similar properties. Moreover, the classical L^2-Sobolev spaces belong to both families (by choosing $p = 2 = q$). The parameter $s \in \mathbb{R}$ is the most important one, describing the smoothness. The family of modulation spaces is closed under duality (at least for finite parameters) or complex interpolation. A summary of the state of the art is given in the survey article ([47], in the special issue of STSIP on modulation spaces).

In the last few years these spaces have found a lot of interest both as a family of Banach spaces of (tempered) distributions of its own right, but above all as a natural tool to describe pseudo-differential operators ([109, 104, 110, 72, 91, 73, 70, 111] and many others, or Chap. 14 of ([69]).)

At the beginning there was the impression that the defining property of a modulation space is the fact that it is a *Wiener amalgam space* (see ([43, 78]) on the Fourier transform side, meaning that it is characterized by *uniform* decomposition of $\hat{f}$, for $f \in \mathscr{S}'(\mathbb{R}^d)$ (as opposed to standard dyadic decompositions used for Besov spaces, see ([97, 113])), or perhaps because a mixed norm-space was used over the TF-plane $\mathbb{R}^d \times \widehat{\mathbb{R}}^d$, with a specific order of integration (first along the time axis, with respect to the L^p-norm, and then in the frequency direction, using an L^q-norm with polynomial weight m), be it a partially discrete or continuous norm of the form

$$\|f\|_{M_m^{p,q}} = \|V_g f\|_{L_m^{p,q}} = \left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |V_g f(x,\omega)|^p m(x,\omega)^p \, dx \right)^{q/p} d\omega \right)^{1/q}, \qquad (63)$$

However, soon the time-frequency point of view suggested to make also use of radial symmetric weights in phase space, not only of weights depending only on the frequency parameter, but rather on polynomial weights of the form $v_s(x,\omega) := (1+|x|^2+|\omega|^2)^{s/2}$. The advantage of the corresponding space $M_{v_s}^p(\mathbb{R}^d)$, defined via a weighted L^p-condition with weight $v_s, s \in \mathbb{R}$, on the short-time Fourier transform $V_{g_0} f$ (with respect to any non-zero window g_0, say the Gauss function) is the fact, that they are invariant, not only under the Fourier transform, but even under fractional Fourier transform (and even the whole metaplectic group, cf. ([69]), Chap.9.4).

The realization that modulation spaces and the classical family of Besov-Triebel-Lizorkin spaces have a lot in common, namely the fact that they can be described using so-called representation coefficients of (square-) integrable and irreducible group representations[24] had a great impact on the view on modulation spaces. They appear now as a special example of a more general principle, which is described through the theory of so-called coorbit spaces $\mathscr{C}o(Y)$ (see ([48])). From this point of view modulation spaces are those spaces which are described by the (global) behaviour of the STFT of its

[24] For the affine $ax+b$-group one obtains the continuous wavelet transform, while one has the STFT in the case of the Heisenberg group $\mathbb{R}^d \times \widehat{\mathbb{R}}^d \times \mathbb{T}$, using the Schrödinger representation on $\mathscr{H} = L^2(\mathbb{R}^d)$.

elements, expressed by some solid and translation invariant Banach space of functions over phase space $\mathbb{R}^d \times \widehat{\mathbb{R}}^d$.

METAPLECTIC OPERATORS AND SCHRÖDINGER EQUATION

In this last section we try to indicate that the invariance properties of the BGT (S_0, L^2, S_0') can be used to describe in the case of quadratic Hamiltonians the properties of solutions of the Schrödinger equation in the BGT setting.

Metaplectic and Heisenberg–Weyl invariance properties

Recall that the metaplectic group $\mathrm{Mp}(2d,\mathbb{R})$ is the unitary representation of the connected double covering of the symplectic group $\mathrm{Sp}(2d,\mathbb{R})$ (see e.g. [62]). The metaplectic group is generated by the following elementary unitary operators:

- The Fourier transform $\widehat{J} = i^{-n/2}F$, that is

$$\widehat{J}\psi(x) = \int_{\mathbb{R}^d} e^{2\pi i x\cdot x'}\psi(x')dx' \tag{64}$$

  whose projection on $\mathrm{Sp}(2d,\mathbb{R})$ is the standard symplectic matrix $J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$;

- The "chirps" $\widehat{V_{-P}}$ defined, for $P = P^T$, by

$$\widehat{V_{-P}}\psi(x) = e^{2\pi i Px\cdot x}\psi(x) \tag{65}$$

  and whose projection on $\mathrm{Sp}(2,\mathbb{R})$ are the symplectic shears $\begin{pmatrix} I & 0 \\ P & I \end{pmatrix}$;

- The unitary changes of variables, defined for invertible L, by

$$\widehat{M_{L,m}}\psi(x) = i^m\sqrt{|\det L|}\psi(Lx) \tag{66}$$

  where the integer m corresponds to a choice of $\arg\det L$; the projection of $\widehat{M_{L,m}}$ on $\mathrm{Sp}(2d,\mathbb{R})$ is $\begin{pmatrix} L^{-1} & 0 \\ 0 & L^T \end{pmatrix}$.

Proposition 2. *The Segal algebra $S_0(\mathbb{R}^d)$ is invariant under the action of* $\mathrm{Mp}(2d,\mathbb{R})$*; in particular $\psi \in S_0(\mathbb{R}^d)$ if and only if $F\psi \in S_0(\mathbb{R}^d)$.*

Proof. This is an immediate property of the metaplectic covariance property

$$W(\widehat{S}\psi)(z) = W\psi(S^{-1}z) \tag{67}$$

of the Wigner distribution ($S \in \mathrm{Sp}(2d,\mathbb{R})$ the projection of $\widehat{S} \in \mathrm{Mp}(2d,\mathbb{R})$) and of the characterization given in Lemma (2). □

Further references are [40, 43, 44, 62, 34]. Finally let us mention that $S_0(\mathbb{R}^d)$ can be charactized via Wilson bases (see [53]) and local Fourier bases (see [76]) . One can show that $f \in L^2(\mathbb{R}^d)$ is in $S_0(\mathbb{R}^d)$ if and only if it has Wilson coefficients in $\ell^1(I)$, where I is essentially a half-space in $\mathbb{Z}^d \times \mathbb{Z}^d$. In this sense Wilson bases over $\mathbb{R}^d$ are like the Fourier basis (defining $A(\mathbb{T})$ and $PM(\mathbb{T})$) for the torus group.

The construction of Wilson bases was published by Daubechies/Jaffard/Journe in [32]. This author learned about Wilson bases from I. Daubechies already in 1989 it was possible to publish the follow-up result (connecting it with modulation spaces) already one year later in [51]. Wilson bases in the discrete domain are given in [14, 15, 16]. They have also been used to prove the kernel theorem in [53].

The Schrödinger equation for quadratic Hamiltonians

The metaplectic group $\mathrm{Mp}(2d,\mathbb{R})$ plays a crucial role in quantum mechanics because of the following property. Consider a Hamiltonian function H which is quadratic in the x_j, p_k variables:

$$H(x,p) = \frac{1}{2}(x,p)M(x,p)^T \tag{68}$$

(M a real symmetric $2d \times 2d$ matrix). Such Hamiltonians generalize the "harmonic oscillator"

$$H(x,p) = \frac{1}{2m}(|p|^2 + m^2\omega^2|x|^2) \tag{69}$$

familiar from elementary physics. The solution of the Hamilton equations

$$\frac{dx}{dt} = \frac{\partial H}{\partial p} \ , \ \frac{dp}{dt} = -\frac{\partial H}{\partial x} \tag{70}$$

is explicitly given by

$$\begin{gathered}(x(t),p(t)) = S_t(x(0),p(0)) \\ \text{with } S_t = \exp(tJM);\end{gathered}$$

since JM is in the Lie algebra of $\mathrm{Sp}(2d,\mathbb{R})$ we have $S_t \in \mathrm{Sp}(2d,\mathbb{R})$ for every $t \in \mathbb{R}$. Now, when t varies the symplectic matrices S_t describe a differentiable curve in $\mathrm{Sp}(2d,\mathbb{R})$ passing through the identity at time $T = 0$ (in fact, (S_t) is a one-parameter subgroup). It follows from a classical result from the theory of covering spaces (the "unique path lifting property") that there exists a unique path $t \longmapsto \widehat{S}_t$ in $\mathrm{Mp}(2d,\mathbb{R})$ whose projection is precisely the path $t \longmapsto S_t$; in particular $\widehat{S}_0$ is the identity in $\mathrm{Mp}(2d,\mathbb{R})$. The interest of these considerations comes from the following well-known result, whose second part trivially follows from Proposition 2 above:

Proposition 3. *(i) Consider the Schrödinger equation*

$$i\hbar\frac{\partial \psi}{\partial t} = H(x,-i\hbar\partial_x)\psi \tag{71}$$

where $H(x,-i\hbar\partial_x)$ is the partial differential operator obtained by Weyl quantization from the quadratic Hamiltonian H. Its solution is given by the formula

$$\psi(x,t) = \widehat{S}_t\psi(x,0). \tag{72}$$

(ii) Thus, if $\psi(\cdot,0) \in S_0(\mathbb{R}^d)$ then $\psi \in S_0(\mathbb{R}^d)$ and the solution depends continuously in the S_0-norm on the time parameter t.

Part (i) has been known for a very long time, it has been implicit in the early work of Hermann Weyl ([119]) , and proofs can be found in [62].

We said above that there is an alternative description of the metaplectic group $\mathrm{Mp}(2d,\mathbb{R})$ in terms of generators. We set

$$W(x,x') = \frac{1}{2}Px\cdot x - Lx\cdot x' + \frac{1}{2}Qx'\cdot x' \tag{73}$$

where P (resp. Q) and L are as above, and consider the Fourier integral operator $\widehat{S}_{W,m}$ defined by

$$\begin{aligned}\widehat{S}_{W,m}\psi(x) &= \left(\frac{1}{2\pi i\hbar}\right)^{n/2}\Delta(W)\int_{\mathbb{R}^d} e^{\frac{i}{\hbar}W(x,x')}\psi(x')dx'\\ \Delta(W) &= i^m\sqrt{|\det L|}.\end{aligned}$$

One verifies, by simple inspection, that $\widehat{S}_{W,m}$ is easily expressed in terms of the elementary generators of $\mathrm{Mp}(2d,\mathbb{R})$, in fact:

$$\widehat{S}_{W,m} = \widehat{V_{-P}}\,\widehat{M_{L,m}}\,\widehat{JV_{-Q}}. \tag{74}$$

It follows that $\widehat{S}_{W,m} \in \mathrm{Mp}(2d,\mathbb{R})$; one proves that the Fourier integral operators $\widehat{S}_{W,m}$ generate the metaplectic group, more precisely: every $\widehat{S} \in \mathrm{Mp}(2d,\mathbb{R})$ can be written (non-uniquely) as a product of exactly two such operators: $\widehat{S} = \widehat{S}_{W,m}\widehat{S}_{W',m'}$. In the case of the Schrödinger equation, it turns out that if the Hessian matrix M of the Hamiltonian is non-singular, the operators $\widehat{S}_t$ are, except for a set of exceptional values of t, of the type $\widehat{S}_{W,m}$.

Note that more concrete realizations of this principle allow E. Cordero and coauthors to derive Strichartz-type estimates for the solutions of the Schrödinger equation (see [26, 25, 27]).

A Fresh Look on Dirac's Functional Calculus

In the case of matrices unitary matrices U are the most useful ones. A complex-valued $n\times n$-matrix U is *unitary* if and only one has

$$U*U' = Id_n = U'*U. \tag{75}$$

Sometimes it is also interesting to consider rectangular matrices U of size $m \times n$, satisfying one or the other of these two properties. If $n \leq m$ it is still possible that $U' * U = Id_n$, or equivalently, that the columns of U form an *orthonormal system*. In particular, they are a linear independent system of vectors. Alternatively for $n \geq m$ one can have a "perfect set of generators", or a so called *tight frame*, satisfying $U * U' = Id_m$, or $x = \sum_{k=1}^{n} \langle x, u_k \rangle u_k$ for all $x \in \mathbb{C}^{m}$[25]. It is obvious that only for the case $n = m$ (finite dimension!) these two properties are *equivalent*, while in one can have one without the other, also for the case $\mathscr{H} = \ell^2$.

Let us now go for the analogue of these two identities in the case of the Fourier transform. Recall that $\mathscr{F}$ is a unitary Banach Gelfand morphism, which however is typically used in the spirit of an *analysis mapping* $\mathscr{F} : f \mapsto (\langle f, \chi_s \rangle)$ or $(\langle f, \chi_s \rangle)$, while the inverse Fourier transform (*synthesis* of a function or distribution from "pure frequencies") is more in the spirit of the adjoint mapping. Since in both cases the kernel for the corresponding mapping is continuous and bounded (actually smooth) , namely $K(x,s) = e^{-2\pi i s \cdot x}$ for $\mathscr{F}$ and $K(s,y) = e^{2\pi i s \cdot x}$ for $\mathscr{F}^{-1}$ the fact that they are corresponding to two mappings which are inverse to each other, i.e. they satisfy

$$\mathscr{F} \circ \mathscr{F}^{-1} = Id_{\mathscr{H}} = \mathscr{F}^{-1} \circ \mathscr{F}, \tag{76}$$

implies obviously that the composition of their symbols according to (58) has to result in the kernel of the identity operator. This brings us to a short discussion of the connection between the *Kronecker's Delta* and *Dirac's Delta*. Modern distribution theory (and in fact the kernel theorem) tell us that the identity mapping (or equivalently multiplication by the constant 1) is given by a kind of δ-distribution concentrated along the diagonal, namely the functional (we just use an "arbitrary symbol" reminding of this idea) δ_Δ, given as an element of $S_0'(\mathbb{R}^{2d})$ via the action $f \mapsto \int_{\mathbb{R}^d} f(t,t)dt$ for $f \in S_0(\mathbb{R}^{2d})$. In the matrix setting we can view the unit matrix Id_n as a collection of unit vectors, which is clearly described in an equivalent way by the Kronecker δ-function Δ_{kron}. Viewed as a matrix kernel we have of course $\Delta_{kron} * x = x$. The continuous analogue of such a situation is a kernel $K(x,y)$ such that

$$\int_{\mathbb{R}^d} K(x,y) f(y) dx = f(x),$$

so somehow one should have for any fixed x that $K(x,y)$ represents δ_x (in the sense of the point measure at x). However, starting from a general distributional kernel $x \in S_0'(\mathbb{R}^{2d})$, even if we write it symbolically in the form the "restriction" to x, i.e. the distribution $K(x,\cdot)$ does not make sense a priori. So we should probably really interpret the Dirac symbol as a continuous analogue of the Kronecker symbol.

In books and papers on *quantum mechanics*, (using slightly different symbols) one often finds relationships such as the following formulas:

$$\langle \chi_s, \chi_t \rangle = \delta(s-t), s,t \in \mathbb{R}^d, \tag{77}$$

[25] One can show that these systems are nothing else than orthogonal 1 projections of orthonormal bases in higher dimensions.

as a replacement for the orthogonality relation, mostly in the form

$$\langle \chi_s, \chi_t \rangle = \delta_t(s), \quad s,t \in \mathbb{R}^d, \tag{78}$$

which is view-point similar to the interpretation of the $n \times n$ identity matrix as a collection of unit vectors (cf. (75)). On the other hand one finds expressions such as

$$Id_{\mathscr{H}} = \int_{\mathbb{R}^d} |\chi_s\rangle\langle\chi_s| ds \tag{79}$$

claiming that the identity operator can viewed as a superposition of rank-one operators (cf. [98], Chap.1), evidently expressing *completeness* of the system of characters $\widehat{\mathbb{R}^d}$.

We can say: both formulas can be given *their proper meanings* in different ways. First of all we view the Fourier transform and its inverse (or equivalently its conjugate kernel) as Banach Gelfand triple morphisms. In some cases it is the "how", i.e. the way how the transformation is first defined, at least on the space of test function $S_0(\mathbb{R}^d)$, which catches our intention. One is lead to believe that the Fourier transform is primarily an integral transform, which has the Lebesgue space $L^1(\mathbb{R}^d)$ as natural domain. On the other hand (when we talk about Fourier synthesis) the w^*-convergence is helpful, because it does not make sense to interpret the Fourier inversion formula (we write f as a superposition of pure frequencies) in any other natural topology.

Despite the fact that $\chi_s \in C_b(\mathbb{R}^d) \subset S_0'(\mathbb{R}^d)$ is only applicable (via integration) on test functions from $S_0(\mathbb{R}^d)$, concrete (hard analysis) arguments allow to show that they determine even a *unitary Banach Gelfand Gelfand triple automorphism*. Obviously one has $\mathscr{F} \circ \mathscr{F}^{-1} = Id$ as well as $\mathscr{F}^{-1} \circ \mathscr{F} = Id$. Both identities can be useful, e.g. in order to show that the Fourier transform is injective, or that a given function is the Fourier transform of another function (or distribution) of the same kind. If we try to describe these two mappings through their kernels and try to compose the kernels using the standard composition formula for kernels we end up with exactly the relations (**??**) and (79) respectively. Of course, one can combine these kernels with regularizing operators, in order to have kernels from $S_0(\mathbb{R}^{2d})$, and in this case the composition can be carried out in the usual way, using Riemannian integrals. Their products are then well defined (according to (58)), and then the claim is: in the w^*-sense the limit of these kernels is the (kernel of the) identity operator, or $\delta(t-s)$, which is in standard terminology the tensor product of δ_0 (the usual Dirac measures at zero) with the function constant one, rotated by 45 degrees. From this point of view Dirac's intention might not have been too far away from simply going from the well-known Kronecker symbol with discrete entries to a continuous version. The fact that the Fourier transform is using building blocks from "outside the Hilbert space" gives troubles to anybody who tries to stay within the world of Hilbert spaces, while the viewpoint of Banach Gelfand triples (in our view only a convenient realization of the idea underlying the concept of rigged Hilbert spaces) opens up a new view and a technically sound perspective. As mathematician we suggest therefore to provide in any concrete application the details of the involved BGT-morphism instead of relying on the symbolic calculus per se. Most likely one can overcome the purely technical problems using the idea of approximation by test functions using regularization ideas, while obviously at critical points (where the

symbolic manipulations lead to misleading conclusions) such justification will fail for good reasons.

ACKNOWLEDGMENTS

The author gratefully acknowledges support in the preparation of this manuscript by a number of present and earlier coworkers, especially Werner Kozek, Maurice De Gosson, Harald Stockinger, and Franz Luef. He was partially supported by the Marie Curie Excellence Grant project EUCETIFA (FP6-517154). Thanks go also to professor Abul Hasan Siddiqi for inviting the author to the ISIAM Conference in Agra, January 2009. The notes as it was held (with a number of illustrations) can be viewed at
`http://univie.ac.at/nuhag-php/dateien/talks/1055_agra09SIAM.pdf`

REFERENCES

1. Extrapolating a band-limited function from its samples taken in a finite. *IEEE Trans. Inf. Theory*, 32:464–470, 1986.
2. S. Albeverio, R. Bozhok, and V. Koshmanenko. The rigged Hilbert spaces approach in singular perturbation Theory. *Rep. Math. Phys.*, 58(2):227–246, 2006.
3. S. T. Ali, J.-P. Antoine, and J. P. Gazeau. Continuous frames in Hilbert space. *Ann. Phys.*, 222(1):1–37, 1993.
4. S. T. Ali, J.-P. Antoine, and J.-P. Gazeau. *Coherent States, Wavelets and their Generalizations*. Graduate Texts in Contemporary Physics. Springer, New York, 2000.
5. A. Amann, H. G. Feichtinger, A. Klotz, G. Kracher, T. Werther, H. Gilly, and M. Baubin. CPR Artefact Removal in ECG Signals Using Gabor Multipliers. *IEEE Trans. Biomedical Engineering*, 2007.
6. J.-P. Antoine. Quantum Mechanics Beyond Hilbert Space. In A. Bohm, H.-D. Doebner, and P. Kileanowski, editors, *Irreversibility and Causality, Semigroups and Rigged Hilbert Spaces. A Selection of Articles Presented at the 21st International Colloquium on Group Theoretical Methods in Physics (ICGTMP) at Goslar, Germany, 16-21 July, 1996.*, volume 504 of *Lecture Notes in Physics*, Berlin, 1998. Springer Verlag.
7. J.-P. Antoine and C. Trapani. Partial Inner Product Spaces - Theory and Applications. 2009.
8. L. Auslander and Y. Meyer. A generalized Poisson summation formula. *Appl. Comput. Harmon. Anal.*, 3(4):372–376, 1996.
9. P. Balazs. *Regular and Irregular Gabor Multiplier With Application To Psychoacoustic Masking*. PhD thesis, University of Vienna, 2005.
10. M. J. Bastiaans. Gabor's expansion of a signal into Gaussian elementary signals. *Proc. IEEE*, 68(4):538– 539, 1980.
11. J. J. Benedetto. *Spectral Synthesis*. Academic Press, Inc, Francisco, 1975.
12. J. J. Benedetto and G. E. Pfander. Frame expansions for Gabor multipliers. *Appl. Comput. Harm. Anal.*, 20(1):26–40, 2006.
13. P. Boggiatto, E. Cordero, and K. Gröchenig. Generalized anti-Wick operators with symbols in distributional Sobolev spaces. *Integral Equations Operator Theory*, 48(4):427–442, 2004.
14. H. Bölcskei, H. G. Feichtinger, K. Gröchenig, and F. Hlawatsch. Discrete-time Wilson expansions. In *Proc. IEEE-SP 1996 Int. Sympos. Time-Frequency Time-Scale Analysis*, pages 525–528, jun 1996.
15. H. Bölcskei, K. Gröchenig, F. Hlawatsch, and H. G. Feichtinger. Oversampled Wilson expansions. *IEEE Signal Proc. Letters*, 4(4):106–108, 1997.
16. H. Bölcskei and F. Hlawatsch. Oversampled Wilson-type cosine modulated filter banks with linear phase. In *Asilomar Conf. on Signals, Systems, and Computers*, pages 998–1002, nov 1996.

17. W. Braun and H. G. Feichtinger. Banach spaces of distributions having two module structures. *J. Funct. Anal.*, 51:174–212, 1983.
18. P. L. Butzer, S. Jansche, and R. L. Stens. Functional analytic methods in the solution of the fundamental theorems on best-weighted algebraic approximation. In *Approximation theory, Proc. 6th Southeast. Approximation Theory Conf., Memphis/TN (USA) 1991*, volume 138 of *Lect. Notes Pure Appl. Math.*, pages 151–205, 1992.
19. O. Christensen. *An Introduction to Frames and Riesz Bases*. Applied and Numerical Harmonic Analysis. Birkhäuser, Boston, 2003.
20. T. A. C. M. Claasen and W. F. G. Mecklenbraeuker. The Wigner distribution - a tool for time-frequency signal analysis. I: Continuous-time signals. *Philips J. Res.*, 35:217–250, 1980.
21. T. A. C. M. Claasen and W. F. G. Mecklenbraeuker. The Wigner distribution - a tool for time-frequency signal analysis. II: Discrete-time signals. *Philips J. Res.*, 35:276–300, 1980.
22. T. A. C. M. Claasen and W. F. G. Mecklenbraeuker. The Wigner distribution - a tool for time-frequency signal analysis. III: Relations with other time-frequency signal transformations. *Philips J. Res.*, 35:372–389, 1980.
23. R. R. Coifman and G. Weiss. *Analyse harmonique non-commutative sur certains espaces homog'enes. Etude de certaines int'egrales singuli'eres. (Non-commutative harmonic analysis on certain homogeneous spaces. Study of certain singular integrals.).*, volume 242 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin-Heidelberg-New York, 1971.
24. E. Cordero, H. G. Feichtinger, and F. Luef. Banach Gelfand triples for Gabor analysis. In *'Pseudo-Differential Operators, Quantization and Signals', C.I.M.E. 2006*, Lecture Notes in Mathematics. Springer, 2007.
25. E. Cordero and F. Nicola. Some new Strichartz estimates for the Schrödinger equation. *J. Differ. Equations*, 245(7):1945–1974, 2008.
26. E. Cordero and F. Nicola. Strichartz estimates in Wiener amalgam spaces for the Schrödinger equation. *Math. Nachr.*, 281(1):25–41, 2008.
27. E. Cordero, F. Nicola, and L. Rodino. Sparsity of Gabor representation of Schrödinger propagators. *Applied and Computational Harmonic Analysis,*, In Press, Corrected Proof,(,):– „ 2008,.
28. E. Cordero and L. Rodino. Wick calculus: a time-frequency approach. *Osaka J. Math.*, 42(1):43–63, 2005.
29. S. Dahlke, M. Fornasier, H. Rauhut, G. Steidl, and G. Teschke. Generalized coorbit theory, Banach frames, and the relation to *alpha*-modulation spaces. *Proc. Lond. Math. Soc. (3)*, 96(2):464–506, 2008.
30. S. Dahlke, G. Steidl, and G. Teschke. Coorbit spaces and Banach frames on homogeneous spaces with applications to the sphere. *Adv. Comput. Math.*, 21(1-2):147–180, 2004.
31. S. Dahlke, G. Steidl, and G. Teschke. Weighted coorbit spaces and Banach frames on homogeneous spaces. *J. Fourier Anal. Appl.*, 10(5):507–539, 2004.
32. I. Daubechies, S. Jaffard, and J. L. Journ'e. A simple Wilson orthonormal basis with exponential decay. *SIAM J.Math. Anal.*, 22:554–573, 1991.
33. I. Daubechies, H. J. Landau, and Z. Landau. Gabor time-frequency lattices and the Wexler-Raz identity. *J. Fourier Anal. Appl.*, 1(4):437–478, 1995.
34. M. A. de Gosson. *Symplectic Geometry and Quantum Mechanics*, volume 166 of *Operator Theory: Advances and Applications. Advances in Partial Differential Equations*. Birkhäuser, Basel, 2006.
35. R. de la Madrid. Rigged Hilbert space approach to the Schrödinger equation. *J. Phys. A, Math. Gen.*, 35(2):319–342, 2002.
36. R. de la Madrid. The role of the rigged Hilbert space in quantum mechanics. *Eur. J. Phys.*, 26(2):277–312, 2005.
37. R. de la Madrid, A. Bohm, and M. Gadella. Rigged Hilbert space treatment of continuous spectrum. *Fortschr. Phys.*, 50(2):185–216, 2002.
38. M. Dörfler and B. Torresani. On the time-frequency representation of operators and generalized Gabor multiplier approximations. *preprint*, 2007.
39. H. G. Feichtinger. Un espace de Banach de distributions temp'er'ees sur les groupes localement compacts ab'eliens. *Compt.Rend.Acad.Sci.Paris, Ser.A*, 290(17):791–794, 1980.
40. H. G. Feichtinger. A characterization of minimal homogeneous Banach spaces. *Proc. Amer. Math. Soc.*, 81:55–61, 1981.
41. H. G. Feichtinger. On a new Segal algebra. *Monatsh. Math.*, 92:269–289, 1981.

42. H. G. Feichtinger. A new family of functional spaces on the Euclidean n-space. In *Proc.Conf. on Theory of Approximation of Functions*, Teor. Priblizh., 1983.
43. 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.
44. H. G. Feichtinger. Modulation spaces on locally compact Abelian groups. Technical report, January 1983.
45. H. G. Feichtinger. An elementary approach to the generalized Fourier transform. In T. Rassias, editor, *Topics in Mathematical Analysis*, pages 246–272. World Sci.Pub., 1989.
46. H. G. Feichtinger. Modulation spaces of locally compact Abelian groups. In R. Radha, M. Krishna, and S. Thangavelu, editors, *Proc. Internat. Conf. on Wavelets and Applications*, pages 1–56, Chennai, January 2002, 2003. New Delhi Allied Publishers.
47. H. G. Feichtinger. Modulation Spaces: Looking Back and Ahead. *Sampl. Theory Signal Image Process.*, 5(2):109–140, 2006.
48. H. G. Feichtinger and K. Gröchenig. Banach spaces related to integrable group representations and their atomic decompositions, I. *J. Funct. Anal.*, 86:307–340, 1989.
49. H. G. Feichtinger and K. Gröchenig. Banach spaces related to integrable group representations and their atomic decompositions, II. *Monatsh. Math.*, 108:129–148, 1989.
50. H. G. Feichtinger and K. Gröchenig. Gabor frames and time-frequency analysis of distributions. *J. Funct. Anal.*, 146(2):464–495, 1997.
51. H. G. Feichtinger, K. Gröchenig, and D. F. Walnut. Wilson bases and modulation spaces. *Math. Nachr.*, 155:7–17, 1992.
52. H. G. Feichtinger and W. Hörmann. Harmonic analysis of generalized stochastic processes on locally compact abelian groups, 1990.
53. H. G. Feichtinger and W. Kozek. Quantization of TF lattice-invariant operators on elementary LCA groups. In H. Feichtinger and T. Strohmer, editors, *Gabor Analysis and Algorithms. Theory and Applications.*, Applied and Numerical Harmonic Analysis, pages 233–266, 452–488, Boston, MA, 1998. Birkhäuser Boston.
54. H. G. Feichtinger, W. Kozek, and F. Luef. Gabor Analysis over finite Abelian groups. *Appl. Comput. Harmon. Anal.*, 26(2):230–248, 2009.
55. H. G. Feichtinger and F. Luef. Wiener Amalgam Spaces for the Fundamental Identity of Gabor Analysis. *Collect. Math.*, 57(Extra Volume (2006)):233–253, 2006.
56. H. G. Feichtinger, F. Luef, and T. Werther. A Guided Tour from Linear Algebra to the Foundations of Gabor Analysis. In *Gabor and Wavelet Frames*, volume 10 of *IMS Lecture Notes Series*, pages 1–49. 2007.
57. H. G. Feichtinger and K. Nowak. A first survey of Gabor multipliers. In H. Feichtinger and T. Strohmer, editors, *Advances in Gabor Analysis*, Appl. Numer. Harmon. Anal., pages 99–128. Birkhäuser, 2003.
58. H. G. Feichtinger and F. Weisz. The Segal algebra $S_0(R^d)$ and norm summability of Fourier series and Fourier transforms. *Monatsh. Math.*, 148:333–349, 2006.
59. H. G. Feichtinger and F. Weisz. Wiener amalgams and pointwise summability of Fourier transforms and Fourier series. *Math. Proc. Cambridge Philos. Soc.*, 140:509–536, 2006.
60. H. G. Feichtinger and F. Weisz. Herz spaces and summability of Fourier transforms. *Math. Nachr.*, 281(3):1–16, 2008.
61. H. G. Feichtinger and G. Zimmermann. A Banach space of test functions for Gabor analysis. In H. Feichtinger and T. Strohmer, editors, *Gabor Analysis and Algorithms: Theory and Applications*, Applied and Numerical Harmonic Analysis, pages 123–170, Boston, MA, 1998. Birkhäuser Boston.
62. G. B. Folland. *Harmonic Analysis in Phase Space*. Princeton University Press, Princeton, N.J., 1989.
63. M. Fornasier and H. Rauhut. Continuous Frames, Function Spaces, and the Discretization Problem. *J. Fourier Anal. Appl.*, 11(3):245–287, 2005.
64. J.-P. Gabardo and D. Han. Frames associated with measurable spaces. *Adv. Comput. Math.*, 18(2-4):127–147, 2003.
65. D. Gabor. Theory of communication. *J. IEE*, 93(26):429–457, 1946.

66. M. Gadella and F. Gomez. A unified mathematical formalism for the Dirac formulation of quantum mechanics. *Found. Phys.*, 32(6):815–869, 2002.
67. K. Gröchenig. Describing functions: atomic decompositions versus frames. *Monatsh. Math.*, 112(3):1–41, 1991.
68. K. Gröchenig. An uncertainty principle related to the Poisson summation formula. *Stud. Math.*, 121(1), 1996.
69. K. Gröchenig. *Foundations of Time-Frequency Analysis.* Appl. Numer. Harmon. Anal. Birkhäuser Boston, Boston, MA, 2001.
70. K. Gröchenig. Composition and spectral invariance of pseudodifferential operators on modulation spaces. *J. Anal. Math.*, 98:65–82, 2006.
71. K. Gröchenig. Gabor frames without inequalities. *International Mathematics Research Notices 2007*, (23), 2007.
72. K. Gröchenig and C. Heil. Modulation spaces and pseudodifferential operators. *Integral Equations Operator Theory*, 34(4):439–457, 1999.
73. K. Gröchenig and C. Heil. Modulation spaces as symbol classes for pseudodifferential operators. In R. R. M. Krishna, editor, *Wavelets and Their Applications*, pages 151–170. Allied Publishers, Chennai, 2003.
74. K. Gröchenig, F. Hlawatsch, A. Klotz, G. Matz, and A. Skupch. Advanced Mathematical Models for the Design and Optimization of Low-Interference Wireless Multicarrier Systems. 2006.
75. K. Gröchenig and M. Leinert. Wiener's lemma for twisted convolution and Gabor frames. *J. Amer. Math. Soc.*, 17:1–18, 2004.
76. K. Gröchenig and S. Samarah. Non-linear approximation with local Fourier bases. *Constr. Approx.*, 16(3):317–331, 2000.
77. D. Han, K. Kornelson, D. Larson, and E. Weber. *Frames for Undergraduates.*, volume 40 of *Student Mathematical Library*. American Mathematical Society (AMS), Providence, RI, 2007.
78. 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.
79. F. Hlawatsch and P. Flandrin. The interference structure of the Wigner distribution and related time-frequency signal representations. In F. Hlawatsch and W. Mecklenbräuker, editors, *The Wigner Distribution - Theory and Applications in Signal Processing*, pages 59–133. Elsevier Science, 1997.
80. F. Hlawatsch and W. Kozek. The Wigner distribution of a linear signal space. *IEEE Trans. Signal Processing*, 41:1248–1258, 1993.
81. A. Janssen. *Application of the Wigner Distribution to Harmonic Analysis of Generalized Stochastic Processes. Dissertation.* 1979.
82. A. J. E. M. Janssen. Duality and biorthogonality for Weyl-Heisenberg frames. *J. Fourier Anal. Appl.*, 1(4):403–436, 1995.
83. J. P. Kahane. *Some Random Series of Functions.* 1968.
84. J. P. Kahane and P. G. Lemari'e Rieusset. Remarques sur la formule sommatoire de Poisson. *Studia Math.*, 109:303–316, 1994.
85. G. Kaiser. *A Friendly Guide to Wavelets.* Birkhäuser, Boston, 1994.
86. P. Kasperkovitz. Quasiclassical descriptions of quantum systems based on coherent states: Product formulae. *J. Phys. A, Math. Gen.*, 23(23):5493–5512, 1990.
87. Y. Katznelson. Une remarque concernant la formule de Poisson. *Studia Math.*, 19:107–108, 1967.
88. Y. Katznelson. *An Introduction to Harmonic Analysis. 2nd corr. ed.* Dover Publications Inc, New York, 1976.
89. B. Keville. *Multidimensional Second Order Generalised Stochastic Processes on Locally Compact Abelian Groups.* PhD thesis, Trinity College Dublin, 2003.
90. W. Kozek. *Matched Weyl-Heisenberg Expansions of Nonstationary Environments.* PhD thesis, University of Technology Vienna, 1997.
91. D. Labate. Pseudodifferential operators on modulation spaces. *J. Math. Anal. Appl.*, 262(1):242–255, 2001.
92. R. Larsen. *An Introduction to the Theory of Multipliers.* Springer, 1971.
93. N. Lerner. The Wick calculus of pseudo-differential operators and energy estimates. In J.-M. Bony and et al., editors, *New trends in microlocal analysis.*, pages 23–37. Springer, Tokyo, 1997.

94. V. Losert. A characterization of the minimal strongly character invariant Segal algebra. *Ann. Inst. Fourier*, 30:129–139, 1980.
95. V. Losert. A characterization of groups with the one-sided Wiener property. *J. Reine Angew. Math.*, 331:47–57, 1982.
96. S. MacLane. *Categories for the working mathematician. 4th corrected printing.* Graduate Texts in Mathematics, 5. New York etc.: Springer-Verlag. ix, 1988.
97. J. Peetre. *New Thoughts on Besov Spaces.* Mathematics Department, Duke University, Durham, 1976.
98. A. Perelomov. *Generalized Coherent States and their Applications.* Springer, Berlin, 1986.
99. H. Reiter. *L^1-algebras and Segal Algebras.* Springer, Berlin, Heidelberg, New York, 1971.
100. H. Reiter. Theta functions and symplectic groups. *Monatsh. Math.*, 97:219–232, 1984.
101. H. Reiter. *Metaplectic Groups and Segal Algebras.* Lect. Notes in Mathematics. Springer, Berlin, 1989.
102. H. Reiter and J. D. Stegeman. *Classical Harmonic Analysis and Locally Compact Groups. 2nd ed.* Clarendon Press, Oxford, 2000.
103. J. E. Roberts. Rigged Hilbert spaces in quantum mechanics. *Commun. Math. Phys.*, 3:98–119, 1966.
104. R. Rochberg and K. Tachizawa. Pseudodifferential operators, Gabor frames, and local trigonometric bases. In H. Feichtinger and T. Strohmer, editors, *Gabor Analysis and Algorithms: Theory and Applications*, pages 171–192, 453–488. Birkhäuser, Boston, 1998.
105. A. Ron and Z. Shen. Weyl-Heisenberg frames and Riesz bases in $L_2(\mathbb{R}^d)$. *Duke Math. J.*, 89(2):237–282, 1997.
106. W. Rudin. *Functional Analysis.* McGraw-Hill Book Company, New York, 1973.
107. K. Schnass. Gabor Multipliers. A self-contained survey. Master's thesis, University of Vienna, 2004.
108. G. Strang. *Introduction to Linear Algebra. 3rd ed.* Wellesley-Cambridge Press, Wellesley, MA, 2003.
109. K. Tachizawa. The boundedness of pseudodifferential operators on modulation spaces. *Math. Nachr.*, 168:263–277, 1994.
110. K. Tachizawa. A generalization of Calder'on-Vaillancourt's theorem. *RIMS Kokyuroku*, 1102:64–75, 1999.
111. J. Toft. Continuity properties for modulation spaces, with applications to pseudo-differential calculus. I. *J. Funct. Anal.*, 207(2):399–429, 2004.
112. R. Tolimieri and R. Orr. Poisson summation, the ambiguity function, and the theory of Weyl–Heisenberg frames. *J.Four. Anal. Appl.*, 1:233–247, 1995.
113. H. Triebel. *Theory of Function Spaces.*, volume 78 of *Monographs in Mathematics.* Birkhäuser, Basel, 1983.
114. P. Wahlberg. Vector-valued modulation spaces and localization operators with operator-valued symbols. *Integral Equations Oper. Theory*, 59(1):99–128, 2007.
115. P. Wahlberg and P. Schreier. Gabor discretization of the Weyl product for modulation spaces and filtering of nonstationary stochastic processes. *Appl. Comput. Harmon. Anal.*, 26:97–120, January 2009.
116. A. Weil. *L'integration dans les Groupes Topologiques et ses Applications.* Hermann and Cie, Paris, 1940.
117. F. Weisz. Inversion of the short-time Fourier transform using Riemannian sums. *J. Fourier Anal. Appl.*, 13(3):357–368, 2007.
118. J. Wexler and S. Raz. Discrete Gabor expansions. *Signal Proc.*, 21:207–220, 1990.
119. H. Weyl. *Gruppentheorie und Quantenmechanik.* S. Hirzel, Leipzig, 1928.
120. S. Wickramasekara and A. Bohm. Representation of semigroups in rigged Hilbert spaces: subsemigroups of the Weyl-Heisenberg group. *J. Math. Phys.*, 44(2):930–942, 2003.