POLITECNICO DI TORINO
Repository ISTITUZIONALE

Filtering of Second Order Generalized Stochastic Processes Corrupted by Additive Noise

Original
Filtering of Second Order Generalized Stochastic Processes Corrupted by Additive Noise / Wahlberg, P.. - In: JOURNAL OF FOURIER ANALYSIS AND APPLICATIONS. - ISSN 1069-5869. - 32:4(2026), pp. 1-41. [10.1007/s00041-026-10282-y]

Availability:
This version is available at: 11583/3015818 since: 2026-09-22T16:22:47Z

Publisher:
Springer

Published
DOI:10.1007/s00041-026-10282-y

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

14 November 2026

Journal of Fourier Analysis and Applications (2026) 32:71
https://doi.org/10.1007/s00041-026-10282-y

Filtering of Second Order Generalized Stochastic Processes Corrupted by Additive Noise

Patrik Wahlberg[1]

Received: 25 April 2025 / Revised: 20 June 2026 / Accepted: 29 June 2026
© The Author(s) 2026

Abstract
We treat the optimal linear filtering problem for a sum of two second order uncorrelated generalized stochastic processes. This is an operator equation involving covariance operators. We study both the wide-sense stationary case and the non-stationary case. In the former case the equation simplifies into a convolution equation. The solution is the Radon–Nikodym derivative between non-negative tempered Radon measures, for signal and signal plus noise respectively, in the frequency domain. In the non-stationary case we work with pseudodifferential operators with symbols in Sjöstrand modulation spaces which admits the use of its spectral invariance properties.

Keywords Second order tempered generalized stochastic processes · Optimal linear filter · Additive uncorrelated noise · Modulation spaces

Mathematics Subject Classification Primary: 60G20 · 60G35 · 47A62 · 94A12 · 47G30 · 47B65 · 42B35 · Secondary: 47A60

1 Introduction

The paper concerns zero mean second order generalized stochastic processes. These are defined as linear continuous maps from a space of test functions defined on $\mathbf{R}^d$ into a Hilbert space of finite variance random variables. Given a sum of two such uncorrelated generalized stochastic processes we study the optimal filtering problem, whose goal is to find the linear operator that recovers one of them with the minimum mean square error.

Communicated by Hans G. Feichtinger.

✉ Patrik Wahlberg
patrik.wahlberg@polito.it

[1] Dipartimento di Scienze Matematiche, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy

Published online: 18 July 2026

If u denotes the useful signal and w the uncorrelated noise, the equation for the optimal linear operator F reads

$$\mathscr{K}_u = F(\mathscr{K}_u + \mathscr{K}_w) \tag{1.1}$$

where $\mathscr{K}_u$ and $\mathscr{K}_w$ are the covariance operators of u and w respectively. These are non-negative continuous linear operators from $C_c^\infty(\mathbf{R}^d)$ to the distributions $\mathscr{D}'(\mathbf{R}^d)$.

If both generalized stochastic processes are wide-sense stationary then the covariance operators are translation invariant which means that they are convolution operators. Their covariance kernels are then Fourier transforms of non-negative tempered Radon measures μ_u and μ_w respectively defined on $\mathbf{R}^d$. It is then natural to impose the operator F in (1.1) to be a convolution operator, that is $F = f*$.

Our first result is the determination of the optimal convolution operator in the wide-sense stationary case. In the frequency domain it turns out to be the Radon–Nikodym derivative

$$\widehat{f} = \frac{\mathrm{d}\mu_u}{\mathrm{d}(\mu_u + \mu_w)}$$

which is a function in $L^\infty(\mu_u + \mu_w)$ that satisfies $0 \leqslant \widehat{f} \leqslant 1$ almost everywhere. This result generalizes and formalizes widely established engineering intuition for the optimal filter for wide-sense stationary processes. The functional framework for equation (1.1) in this case are Hilbert spaces $\mathscr{F}L^2(\mu)$, that is tempered distributions with Fourier transforms that belong to $L^2_{\mathrm{loc}}(\mu)$ and are square integrable with respect to a non-negative tempered Radon measure μ. In fact for wise-sense stationary generalized stochastic processes the covariance operators act continuously on such spaces for certain μ.

Secondly we study equation (1.1) under the assumption that the generalized stochastic processes are non-stationary which is a less well defined problem. We need frameworks and restrictions to be able to formulate solutions.

As an intermediate step from wide-sense stationarity to non-stationarity we first impose the restriction that the covariance operators $\mathscr{K}_u$ and $\mathscr{K}_w$ commute, which holds in the former case, and are continuous and non-negative on a Hilbert space of distributions on $\mathbf{R}^d$. Then we may solve (1.1) using the spectral theorem and its associated functional calculus as

$$F = \int_{\mathbf{C}} f(z)\mathrm{d}\Pi(z) \in \mathscr{L}(\mathscr{H})$$

where Π is a projection-valued measure compactly supported in the first quadrant in $\mathbf{C}$, and where

$$f(z) = \frac{f_u(z)}{f_u(z) + f_w(z)}\,\chi_{\operatorname{supp} f_u}$$

with

$$\mathscr{K}_u = \int_{\mathbf{C}} f_u(z)\mathrm{d}\Pi(z), \quad \mathscr{K}_w = \int_{\mathbf{C}} f_w(z)\mathrm{d}\Pi(z).$$

Finally we relax the commutativity of $\mathscr{K}_u$ and $\mathscr{K}_w$ and study equation (1.1) as an equation for pseudodifferential covariance operators of non-stationary generalized stochastic processes. We work in the functional framework of modulation spaces for the Weyl symbols of the operators and for the spaces on which operators act. Our goal is to extend the analysis for the wide-sense stationary case. We show embeddings of modulation spaces in the spaces $\mathscr{F}L^2(\mu)$. We also show that the Weyl symbols $1 \otimes \mu$, where μ is a non-negative tempered Radon measure, of the covariance operators in the wide-sense stationary case do belong to modulation spaces $M^{\infty,1}_{\omega}(\mathbf{R}^{2d})$ for certain weights ω defined on $\mathbf{R}^{4d}$. This implies that the corresponding operators map between certain modulation spaces, which gives a functional framework for the equation (1.1). The weight ω is however in general decreasing which is too weak for the exploitation of the powerful methods based on Gröchenig's and Sjöstrand's results on the Wiener property of the symbol class $M^{\infty,1}_{1\otimes\omega}(\mathbf{R}^{2d})$ where $\omega(X) = (1+|X|)^r$ for $X \in \mathbf{R}^{2d}$ and $r \geqslant 0$.

Imposing covariance operators to have Weyl symbols in $M^{\infty,1}_{1\otimes\omega}(\mathbf{R}^{2d})$ admits white noise but excludes certain other wide-sense stationary generalized stochastic processes, for example derivatives of white noise. In this framework we show the following result which is a rather immediate consequence of results in [39]. If $\mathscr{K}_u$ and $\mathscr{K}_w$ have Weyl symbols in $M^{\infty,1}_{1\otimes\omega}(\mathbf{R}^{2d})$ for some $r \geqslant 0$ and $\mathscr{K}_u + \mathscr{K}_w$ is invertible as an operator on $L^2(\mathbf{R}^d)$, then the optimal filter F solving (1.1) is again a Weyl pseudodifferential operator with symbol in $M^{\infty,1}_{1\otimes\omega}(\mathbf{R}^{2d})$. If $r > 2d$ then the Gabor coefficients of the Weyl symbols of F and $\mathscr{K}_u$ are related by a multiplication of an infinite matrix with polynomial off-diagonal decay of order smaller than $r/2 - d$.

Finally we discuss a few operator theoretical observations concerning equation (1.1) considered as an equation for bounded linear operators on a Hilbert space on which $\mathscr{K}_u$ and $\mathscr{K}_w$ are non-negative, and $\mathscr{K}_u + \mathscr{K}_w$ is allowed to be non-invertible. Douglas' lemma then says that there is a bounded linear operator F that solves (1.1), with certain uniqueness properties, provided $\operatorname{ran} \mathscr{K}_u \subseteq \operatorname{ran}(\mathscr{K}_u + \mathscr{K}_w)$.

The analysis in this work is second order, which explains the assumption that the processes are uncorrelated rather than statistically independent.

The optimal filtering problem for (generalized) stochastic processes has a long and eclectic history going back to Kolmogorov [25] and Wiener [42], cf. [11, 12, 19, 28, 38, 39, 41]. It has been studied mostly under the assumption of wide-sense stationarity. The time domain has been either $\mathbf{Z}$ or $\mathbf{R}$, and the filter has often been subject to the constraint to be *causal*. This means that the convolutor f has support on $\mathbf{R}_+$ and it is natural in engineering applications. Usually the term Wiener filter refers to this constraint. In this paper we do not restrict the supports of filter kernels. Therefore we do not treat causal filters, neither in the wide-sense stationary nor in the non-stationary case.

The paper is organized as follows. Section 2 is quite long and contains notations, and background material on Radon measures, Weyl pseudodifferential operators, modulation spaces, and Gabor frames. It also specifies the framework of generalized stochastic processes. In Section 3 we deduce the equation (1.1) for the optimal filter operator, and Section 4 treats its solution for wide-sense stationary generalized stochastic processes.

In Section 5 we study equation (1.1) for non-stationary generalized stochastic processes. First we keep the feature of commutating covariance operators which holds in the wide-sense stationary case. Then a solution to (1.1) may be determined using the spectral theorem. Finally we relax the commutativity and study the equation (1.1) as an equation for pseudodifferential operators with symbols in modulation spaces.

2 Preliminaries

2.1 Notations

The symbol B_r denotes the ball in $\mathbf{R}^d$ with center at the origin and radius $r > 0$. The notation $K \Subset \mathbf{R}^d$ means that K is compact. We use $\mathbf{R}_+$ for the non-negative real numbers, and χ_A is the indicator function of a subset $A \subseteq \mathbf{R}^d$. We write $f(x) \lesssim g(x)$ provided there exists $C > 0$ such that $f(x) \leqslant C\, g(x)$ for all x in the domain of f and of g. If $f(x) \lesssim g(x) \lesssim f(x)$ then we write $f \asymp g$. The partial derivative $D_j = -i\partial_j$, $1 \leqslant j \leqslant d$, acts on functions and distributions on $\mathbf{R}^d$, with extension to multi-indices as $D^\alpha = i^{-|\alpha|}\partial^\alpha$ for $\alpha \in \mathbf{N}^d$. We use the bracket $\langle x\rangle = (1 + |x|^2)^{\frac{1}{2}}$ for $x \in \mathbf{R}^d$. The space of bounded linear operators on a Banach space X is denoted $\mathscr{L}(X)$, and inclusions $X \subseteq Y$ of Banach spaces understand embeddings, that is continuity. Peetre's inequality with optimal constant (cf. [30, Lemma 2.1]) is

$$\langle x + y\rangle^s \leqslant \left(\frac{2}{\sqrt{3}}\right)^{|s|} \langle x\rangle^s \langle y\rangle^{|s|} \qquad x, y \in \mathbf{R}^d, \quad s \in \mathbf{R}. \tag{2.1}$$

If $1 \leqslant p \leqslant \infty$ then the conjugate exponent $p' \in [1, \infty]$ satisfies $\frac{1}{p} + \frac{1}{p'} = 1$. The normalization of the Fourier transform is

$$\mathscr{F} f(\xi) = \widehat{f}(\xi) = (2\pi)^{-\frac{d}{2}} \int_{\mathbf{R}^d} f(x) e^{-i\langle x,\xi\rangle}\, \mathrm{d}x, \qquad \xi \in \mathbf{R}^d,$$

for $f \in \mathscr{S}(\mathbf{R}^d)$ (the Schwartz space), where $\langle\, \cdot\,,\, \cdot\,\rangle$ denotes the scalar product on $\mathbf{R}^d$. We have for $f, g \in \mathscr{S}(\mathbf{R}^d)$

$$\widehat{f * g} = (2\pi)^{\frac{d}{2}} \widehat{f}\,\widehat{g} \tag{2.2}$$

and this identity extends to $f \in \mathscr{S}'(\mathbf{R}^d)$ (the tempered distributions) and $g \in \mathscr{S}(\mathbf{R}^d)$, with the Fourier transform defined on $\mathscr{S}'(\mathbf{R}^d)$ as $(\widehat{f}, \widehat{g}) = (f, g)$ for $f \in \mathscr{S}'(\mathbf{R}^d)$ and $g \in \mathscr{S}(\mathbf{R}^d)$.

The conjugate linear (antilinear) action of a distribution u on a test function ϕ is written (u, ϕ), consistent with the L^2 inner product $(\cdot\,, \cdot) = (\cdot\,, \cdot)_{L^2}$ which is

conjugate linear in the second argument. Translation of a function or a distribution f is denoted $T_x f(y) = f(y-x)$ for $x, y \in \mathbf{R}^d$, and modulation as $M_\xi f(x) = e^{i\langle x,\xi\rangle} f(x)$ for $x, \xi \in \mathbf{R}^d$.

2.2 Radon Measures

We use non-negative Radon measures on $\mathbf{R}^d$ [10]. By the Riesz representation theorem [10, Theorem 7.2], [32, Theorem 2.14] we may regard such a measure either as a regular non-negative Borel measure, that is a regular σ-additive function whose domain is the Borel σ-algebra on $\mathbf{R}^d$, denoted $\mathscr{B}(\mathbf{R}^d)$, or equivalently we may regard it as a non-negative linear functional on $C_c(\mathbf{R}^d)$ which denotes the space of compactly supported continuous functions, with certain regularity properties. The latter description implies that the measure μ satisfies an estimate of the form

$$|(\mu, \varphi)| \leqslant C_K \sup_{x \in \mathbf{R}^d} |\varphi(x)|, \quad \varphi \in C_c(K),$$

with $C_K > 0$ for each $K \Subset \mathbf{R}^d$, and the non-negativity means $(\mu, \varphi) \geqslant 0$ when $\varphi \geqslant 0$. We adopt the convention that μ is an antilinear functional, and then using the former description of a non-negative Radon measure μ we may write

$$(\mu, \varphi) = \int_{\mathbf{R}^d} \overline{\varphi(x)} \, \mathrm{d}\mu(x), \quad \varphi \in C_c(K).$$

If the measure μ is finite then it extends uniquely to an antilinear functional on the space $C_0(\mathbf{R}^d)$ which denotes the space of continuous functions that vanish at infinity [10, Theorem 7.17]. The space $C_0(\mathbf{R}^d)$ is the completion of $C_c(\mathbf{R}^d)$ with respect to the uniform topology of $L^\infty(\mathbf{R}^d)$.

If

$$\int_{\mathbf{R}^d} \langle x \rangle^{-s} \, \mathrm{d}\mu(x) < \infty \tag{2.3}$$

for some $s \geqslant 0$ then the measure μ is said to be tempered. The space of non-negative tempered Radon measures is denoted $\mathscr{M}_+(\mathbf{R}^d) \subseteq \mathscr{S}'(\mathbf{R}^d)$. An example is Lebesgue measure which is tempered for any $s > d$. A measure $\mu \in \mathscr{M}_+(\mathbf{R}^d)$ extends uniquely to a continuous antilinear functional on $C_{s,0}(\mathbf{R}^d)$ which denotes the space of continuous functions f that vanish at infinity quicker than $\langle \cdot \rangle^{-s}$:

$$\lim_{|x| \to \infty} |f(x)| \langle x \rangle^s = 0$$

equipped with the weighted supremum norm $\| f \langle \cdot \rangle^s \|_{L^\infty(\mathbf{R}^d)}$.

If $\mu \in \mathscr{M}_+(\mathbf{R}^d)$ then we define the Hilbert space $\mathscr{F}L^2(\mu)$ as the subspace of $f \in \mathscr{S}'(\mathbf{R}^d)$ such that $\widehat{f} \in L^2_{\mathrm{loc}}(\mu)$ and

$$\|f\|_{\mathscr{F}L^2(\mu)} = \left(\int_{\mathbf{R}^d} |\widehat{f}(\xi)|^2 \,\mathrm{d}\mu(\xi)\right)^{\frac{1}{2}} < \infty.$$

It follows from [10, Proposition 7.1] that $C_c^\infty(\mathbf{R}^d) \subseteq L^2(\mu)$ is a dense subspace. From (2.3) we get the estimate for $f \in \mathscr{S}(\mathbf{R}^d)$

$$\|f\|_{L^2(\mu)} = \left(\int_{\mathbf{R}^d} |f(\xi)|^2 \,\mathrm{d}\mu(\xi)\right)^{\frac{1}{2}} \lesssim \sup_{\xi\in\mathbf{R}^d} \langle\xi\rangle^{\frac{s}{2}} |f(\xi)|$$

for some $s \geqslant 0$, which implies that $\mathscr{S}(\mathbf{R}^d) \subseteq L^2(\mu)$ is a continuous inclusion. It follows that $\mathscr{S}(\mathbf{R}^d) \subseteq \mathscr{F}L^2(\mu)$ is a dense and continuous inclusion. Note however that the inclusion $\mathscr{S}(\mathbf{R}^d) \subseteq \mathscr{F}L^2(\mu)$ is not guaranteed to be injective, since $\widehat{f} = 0$ in $\mathscr{F}L^2(\mu)$ means that $\widehat{f} = 0$ μ-a.e. but $\widehat{f} \in \mathscr{S}(\mathbf{R}^d) \setminus \{0\}$ may hold, for instance if $\operatorname{supp}\mu \subseteq \mathbf{R}^d$ is compact.

The inclusion $\mathscr{F}L^2(\mu) \subseteq \mathscr{S}'(\mathbf{R}^d)$ is likewise continuous and dense. In fact the density follows from the density of $\mathscr{S}(\mathbf{R}^d) \subseteq \mathscr{S}'(\mathbf{R}^d)$ [29, Corollary V.3.1]. The inclusion $\mathscr{F}L^2(\mu) \subseteq \mathscr{S}'(\mathbf{R}^d)$ is also injective, as opposed to the possible non-injectivity of $\mathscr{S}(\mathbf{R}^d) \subseteq \mathscr{F}L^2(\mu)$. Nevertheless if $\operatorname{supp}\mu \subseteq \mathbf{R}^d$ has non-empty interior then the possibly non-injective inclusion $\mathscr{S}(\mathbf{R}^d) \subseteq \mathscr{F}L^2(\mu)$ can be modified such that it becomes injective, if the test function space is modified as

$$\mathscr{F}C_c^\infty(\operatorname{supp}\mu) \subseteq \mathscr{F}L^2(\mu).$$

We have $H^\infty_{\frac{s}{2}}(\mu) \subseteq \mathscr{F}L^2(\mu)$ provided (2.3) holds true, where

$$H^\infty_t(\mu) = \{\varphi \in \mathscr{S}'(\mathbf{R}^d) : \quad \widehat{\varphi}\langle\cdot\rangle^t \in L^\infty(\mu)\}, \quad t \in \mathbf{R},$$

is a scale of Sobolev spaces with respect to $\mu \in \mathscr{M}_+(\mathbf{R}^d)$. If μ is Lebesgue measure we write $H^\infty_t(\mu) = H^\infty_t(\mathbf{R}^d)$. When $t = 0$ we have $H^\infty_0(\mathbf{R}^d) = \mathscr{F}L^\infty(\mathbf{R}^d)$ which is the space of pseudo-measures. It can be identified with the topological dual of the Fourier algebra $\mathscr{F}L^1(\mathbf{R}^d)$ [23, 26].

The same conclusion holds for $\mathscr{F}L^\infty(\mu) = H^\infty_0(\mu)$ if $\mu \in \mathscr{M}_+(\mathbf{R}^d)$. In fact by [10, Theorem 6.15] the dual $(L^1(\mu))'$ can be identified isometrically with $L^\infty(\mu)$ via the duality

$$L^1(\mu) \times L^\infty(\mu) \ni (f, g) \mapsto \int_{\mathbf{R}^d} f(x)\overline{g(x)} \,\mathrm{d}\mu(x).$$

We can identify $L^\infty(\mu) \subseteq \mathscr{S}'(\mathbf{R}^d)$ as a subspace as

$$(f, \varphi) = \int_{\mathbf{R}^d} f(x)\overline{\varphi(x)} \,\mathrm{d}\mu(x), \quad f \in L^\infty(\mu), \quad \varphi \in \mathscr{S}(\mathbf{R}^d).$$

The inclusion $L^\infty(\mu) \subseteq \mathscr{S}'(\mathbf{R}^d)$ is continuous and injective. The Fourier transform of $f \in L^\infty(\mu)$, considered as a tempered distribution $f \in \mathscr{S}'(\mathbf{R}^d)$, equals

$$(\widehat{f}, \widehat{\varphi}) = \int_{\mathbf{R}^d} f(x)\overline{\varphi(x)}\,\mathrm{d}\mu(x), \quad \varphi \in \mathscr{S}(\mathbf{R}^d).$$

It follows that $\|f\|_{\mathscr{F}L^\infty(\mu)} = \|\widehat{f}\|_{L^\infty(\mu)}$ and $\mathscr{F}L^\infty(\mu) = \big(\mathscr{F}L^1(\mu)\big)'$, cf. [26, Theorem 4.2.2]. The space $H_0^\infty(\mu) = \mathscr{F}L^\infty(\mu)$ for $\mu \in \mathscr{M}_+(\mathbf{R}^d)$ will play an essential role in this paper.

2.3 Weyl Pseudodifferential Operators

If $a \in \mathscr{S}(\mathbf{R}^{2d})$ is a Weyl symbol then the Weyl pseudodifferential operator [10, 20, 33] is defined as

$$a^w(x, D)f(x) = (2\pi)^{-d} \int_{\mathbf{R}^{2d}} e^{i\langle x-y,\xi\rangle} a\left(\frac{x+y}{2}, \xi\right) f(y)\,\mathrm{d}y\,\mathrm{d}\xi, \quad f \in \mathscr{S}(\mathbf{R}^d), \tag{2.4}$$

and $a^w(x, D) : \mathscr{S}(\mathbf{R}^d) \to \mathscr{S}(\mathbf{R}^d)$ is continuous. By the invariance of $\mathscr{S}'(\mathbf{R}^{2d})$ under linear invertible coordinate transformations and partial Fourier transforms, the Weyl correspondence extends to $a \in \mathscr{S}'(\mathbf{R}^{2d})$ in which case $a^w(x, D) : \mathscr{S}(\mathbf{R}^d) \to \mathscr{S}'(\mathbf{R}^d)$ is continuous.

If $a \in \mathscr{S}'(\mathbf{R}^{2d})$ then

$$(a^w(x, D)f, g) = (2\pi)^{-\frac{d}{2}}(a, W(g, f)), \quad f, g \in \mathscr{S}(\mathbf{R}^d), \tag{2.5}$$

where the cross-Wigner distribution [10, 15] is defined as

$$\begin{aligned} W(g, f)(x, \xi) &= (2\pi)^{-\frac{d}{2}} \int_{\mathbf{R}^d} g(x + y/2)\overline{f(x - y/2)}e^{-i\langle y,\xi\rangle}\mathrm{d}y \\ &= \mathscr{F}_2\left((g \otimes \overline{f}) \circ T\right)(x, \xi), \quad (x, \xi) \in T^*\mathbf{R}^d. \end{aligned} \tag{2.6}$$

Here $\mathscr{F}_2$ denotes the partial Fourier transform with respect to the second $\mathbf{R}^d$ variable in $\mathbf{R}^{2d}$, and T is the matrix

$$T = \begin{pmatrix} I_d & \frac{1}{2}I_d \\ I_d & -\frac{1}{2}I_d \end{pmatrix} \in \mathbf{R}^{2d\times 2d}. \tag{2.7}$$

Conversely, by the Schwartz kernel theorem, for any continuous linear operator $\mathscr{K} : \mathscr{S}(\mathbf{R}^d) \to \mathscr{S}'(\mathbf{R}^d)$ there exists a kernel $k \in \mathscr{S}'(\mathbf{R}^{2d})$ and a Weyl symbol $a \in \mathscr{S}'(\mathbf{R}^{2d})$ such that

$$(\mathscr{K}f, g) = (k, g \otimes \overline{f}) = (2\pi)^{-\frac{d}{2}}(a, W(g, f)), \quad f, g \in \mathscr{S}(\mathbf{R}^d).$$

From this we may extract the relation between the Schwartz kernel and the Weyl symbol of a linear continuous operator $\mathscr{K} : \mathscr{S}(\mathbf{R}^d) \to \mathscr{S}'(\mathbf{R}^d)$:

$$k = (2\pi)^{-\frac{d}{2}}(\mathscr{F}_2^{-1}a) \circ T^{-1} \quad \Longleftrightarrow \quad a = (2\pi)^{\frac{d}{2}}\mathscr{F}_2(k \circ T)\,. \tag{2.8}$$

2.4 The Short-Time Fourier Transform, Modulation Spaces and Gabor Frames

Let $\varphi \in \mathscr{S}(\mathbf{R}^d) \setminus \{0\}$. The short-time Fourier transform (STFT) of a tempered distribution $u \in \mathscr{S}'(\mathbf{R}^d)$ is defined by

$$V_\varphi u(x,\xi) = (2\pi)^{-\frac{d}{2}}(u, M_\xi T_x \varphi) = \mathscr{F}(uT_x\overline{\varphi})(\xi), \quad x,\xi \in \mathbf{R}^d. \tag{2.9}$$

The function $V_\varphi u$ is smooth and polynomially bounded [15, Theorem 11.2.3] as

$$|V_\varphi u(x,\xi)| \lesssim \langle (x,\xi)\rangle^k, \quad (x,\xi) \in T^*\mathbf{R}^d, \tag{2.10}$$

for some $k \geqslant 0$. We have $u \in \mathscr{S}(\mathbf{R}^d)$ if and only if

$$|V_\varphi u(x,\xi)| \lesssim \langle (x,\xi)\rangle^{-k}, \quad (x,\xi) \in T^*\mathbf{R}^d, \quad \forall k \geqslant 0.$$

The inverse transform is given by

$$u = (2\pi)^{-\frac{d}{2}} \iint_{\mathbf{R}^{2d}} V_\varphi u(x,\xi) M_\xi T_x \varphi \, \mathrm{d}x \, \mathrm{d}\xi \tag{2.11}$$

provided $\|\varphi\|_{L^2} = 1$, with action under the integral understood, that is

$$(u, f) = (V_\varphi u, V_\varphi f)_{L^2(\mathbf{R}^{2d})} \tag{2.12}$$

for $u \in \mathscr{S}'(\mathbf{R}^d)$ and $f \in \mathscr{S}(\mathbf{R}^d)$, cf. [15, Theorem 11.2.5].

A *weight* on $\mathbf{R}^d$ is a positive function $\omega \in L^\infty_{\mathrm{loc}}(\mathbf{R}^d)$ such that $1/\omega \in L^\infty_{\mathrm{loc}}(\mathbf{R}^d)$. The weight ω is called v-moderate if there is a positive locally bounded function v such that

$$\omega(x+y) \leqslant C\omega(x)v(y), \quad x,y \in \mathbf{R}^d, \tag{2.13}$$

for some constant $C \geqslant 1$. The set $\mathscr{P}(\mathbf{R}^d)$ consists of weights that are v-moderate for a polynomially bounded weight, that is a weight of the form $v(x) = \langle x\rangle^s$ with $s \geqslant 0$.

Modulation spaces were introduced by Feichtinger 1983 and have been studied thoroughly from many points of view. Gröchenig's book [15] is an excellent source for their basic properties.

Definition 2.1 Let $\omega \in \mathscr{P}(\mathbf{R}^{2d})$, let $p, q \in [1, \infty]$ and let $\varphi \in \mathscr{S}(\mathbf{R}^d) \setminus \{0\}$. The modulation space $M^{p,q}_{\omega}(\mathbf{R}^d)$ is the Banach subspace of $\mathscr{S}'(\mathbf{R}^d)$ defined by the norm

$$\|u\|_{M^{p,q}_{\omega}} = \left(\int_{\mathbf{R}^d} \left(\int_{\mathbf{R}^d} |V_\varphi u(x,\xi)|^p \, \omega(x,\xi)^p \, \mathrm{d}x \right)^{\frac{q}{p}} \mathrm{d}\xi \right)^{\frac{1}{q}} = \|(V_\varphi u)\omega\|_{L^{p,q}} \tag{2.14}$$

when $p, q < \infty$ and the usual modifications otherwise. Here $L^{p,q}(\mathbf{R}^{2d})$ is a mix-normed Lebesgue space [15].

The modulation spaces are independent of $\varphi \in \mathscr{S}(\mathbf{R}^d) \setminus \{0\}$ and increase with the indices as

$$M^{1,1}_{\omega} \subseteq M^{p,q}_{\omega} \subseteq M^{r,s}_{\omega} \subseteq M^{\infty,\infty}_{\omega}, \quad 1 \leqslant p \leqslant r, \quad 1 \leqslant q \leqslant s. \tag{2.15}$$

We write $M^{p,p}_{\omega} = M^p_{\omega}$ and $M^{p,q}_{\omega} = M^{p,q}$ if $\omega = 1$. If $\omega(x,\xi) = \langle x \rangle^t \langle \xi \rangle^s$ for $x, \xi \in \mathbf{R}^d$ and $t, s \in \mathbf{R}$ then we write $M^{p,q}_{\omega}(\mathbf{R}^d) = M^{p,q}_{t,s}(\mathbf{R}^d)$. Note that $\omega \in \mathscr{P}(\mathbf{R}^{2d})$, where the polynomial moderateness is a consequence of (2.1). The spaces $M^2_{\omega}(\mathbf{R}^d)$ are Hilbert spaces and $M^2(\mathbf{R}^d) = L^2(\mathbf{R}^d)$.

The L^2-inner product $(\cdot, \cdot)$ on $\mathscr{S}(\mathbf{R}^d) \times \mathscr{S}(\mathbf{R}^d)$ extends uniquely to a continuous sesquilinear form on $M^{p,q}_{\omega}(\mathbf{R}^d) \times M^{p',q'}_{1/\omega}(\mathbf{R}^d)$, and if $p, q < \infty$ then the dual space of $M^{p,q}_{\omega}(\mathbf{R}^d)$ may be identified with $M^{p',q'}_{1/\omega}(\mathbf{R}^d)$ via the form [15, Theorem 11.3.6].

We will also use modulation spaces with domain $\mathbf{R}^{2d}$ and weight functions defined on $\mathbf{R}^{4d}$ of the form

$$\omega(x_1, x_2, \xi_1, \xi_2) = \langle x_1 \rangle^{\tau_1} \langle x_2 \rangle^{\tau_2} \langle \xi_1 \rangle^{\zeta_1} \langle \xi_2 \rangle^{\zeta_2}, \quad x_1, x_2, \xi_1, \xi_2 \in \mathbf{R}^d,$$

with $\tau_1, \tau_2, \zeta_1, \zeta_2 \in \mathbf{R}$. These spaces are denoted $M^{p,q}_{\tau_1,\tau_2,\zeta_1,\zeta_2}(\mathbf{R}^{2d})$ and will be used as symbols for Weyl pseudodifferential operators acting between modulation spaces $M^{p,q}_{t_1,s_1}(\mathbf{R}^d) \to M^{p,q}_{t_2,s_2}(\mathbf{R}^d)$. In fact according to [18, Proposition 2.9] (cf. [36, Theorem 4.2])

$$a^w(x, D) : M^{p,q}_{\omega_1}(\mathbf{R}^d) \to M^{p,q}_{\omega_2}(\mathbf{R}^d)$$

is continuous for all $p, q \in [1, \infty]$ if $a \in M^{\infty,1}_{\omega}(\mathbf{R}^{2d})$, $\omega_1, \omega_2 \in \mathscr{P}(\mathbf{R}^{2d})$, $\omega \in \mathscr{P}(\mathbf{R}^{4d})$, and

$$\frac{\omega_2(X - Y)}{\omega_1(X + Y)} \lesssim \omega(X, -2\mathcal{J}Y), \quad X, Y \in \mathbf{R}^{2d},$$

where

$$\mathcal{J} = \begin{pmatrix} 0 & I_d \\ -I_d & 0 \end{pmatrix} \in \mathbf{R}^{2d \times 2d}$$

is the matrix which plays a fundamental role in symplectic linear algebra [10, 15].

It follows that

$$a^w(x,D) : M^{p,q}_{t_1,s_1}(\mathbf{R}^d) \to M^{p,q}_{t_2,s_2}(\mathbf{R}^d) \tag{2.16}$$

is continuous for all $p, q \in [1,\infty]$ if $a \in M^{\infty,1}_{\tau_1,\tau_2,\zeta_1,\zeta_2}(\mathbf{R}^{2d})$ and

$$\langle x-y\rangle^{t_2}\langle \xi-\eta\rangle^{s_2}\langle x+y\rangle^{-t_1}\langle \xi+\eta\rangle^{-s_1} \lesssim \langle x\rangle^{\tau_1}\langle \xi\rangle^{\tau_2}\langle \eta\rangle^{\zeta_1}\langle y\rangle^{\zeta_2}, \quad x,\xi,y,\eta \in \mathbf{R}^d. \tag{2.17}$$

A particular case is $a \in M^{\infty,1}_{1\otimes\omega}(\mathbf{R}^{2d})$ with $\omega(X) = \langle X\rangle^r$ for $X \in \mathbf{R}^{2d}$ with $r \geqslant 0$. Indeed $M^{\infty,1}_{1\otimes\omega}(\mathbf{R}^{2d}) \subseteq M^{\infty,1}_{0,0,\frac{r}{2},\frac{r}{2}}(\mathbf{R}^{2d})$ and hence we obtain from (2.1) and (2.17) that

$$a^w(x,D) : M^{p,q}_{t,s}(\mathbf{R}^d) \to M^{p,q}_{t,s}(\mathbf{R}^d) \tag{2.18}$$

is continuous for all $p, q \in [1,\infty]$, and all $t, s \in \mathbf{R}$ such that $\max(|t|,|s|) \leqslant \frac{r}{2}$.

Consequently $a^w(x,D)$ is continuous on $L^2(\mathbf{R}^d) = M^2(\mathbf{R}^d)$. Gröchenig's spectral invariance theorem [17, Theorem 4.6] says that if $a \in M^{\infty,1}_{1\otimes\omega}(\mathbf{R}^{2d})$ and $a^w(x,D)$ is invertible on $L^2(\mathbf{R}^d)$, then $a^w(x,D)^{-1} = b^w(x,D)$ with $b \in M^{\infty,1}_{1\otimes\omega}(\mathbf{R}^{2d})$. This so called Wiener property of $M^{\infty,1}_{1\otimes\omega}(\mathbf{R}^{2d})$ is a refinement of Sjöstrand's original result [35] which concerned the case when $r = 0$.

Thus (2.17) gives good mapping properties for operators with Weyl symbols in the space $M^{\infty,1}_{\tau_1,\tau_2,\zeta_1,\zeta_2}(\mathbf{R}^{2d})$ when $\tau_1, \tau_2, \zeta_1, \zeta_2 \geqslant 0$, and furthermore the Wiener property holds. But we will need also symbols $M^{\infty,1}_{0,\tau_2,\zeta_1,\zeta_2}(\mathbf{R}^{2d})$ with τ_2 and ζ_2 negative. If $\tau_1 = 0$ and $\tau_2, \zeta_2 < 0 \leqslant \zeta_1$ then (2.17) holds if

$$\langle x-y\rangle^{t_2}\langle x+y\rangle^{-t_1}\langle \xi-\eta\rangle^{s_2}\langle \xi+\eta\rangle^{-s_1}\langle \xi\rangle^{-\tau_2}\langle y\rangle^{-\zeta_2} \lesssim \langle \eta\rangle^{\zeta_1}. \tag{2.19}$$

From (2.1) it follows that

$$\langle \xi-\eta\rangle^{s_2}\langle \xi+\eta\rangle^{-s_1}\langle \xi\rangle^{-\tau_2} \lesssim \langle \eta\rangle^{\zeta_1}$$

provided $|s_1| + |s_2| \leqslant \zeta_1$ and $s_2 \leqslant s_1 + \tau_2$, and then (2.19) reduces to

$$\langle x-y\rangle^{t_2}\langle x+y\rangle^{-t_1}\langle y\rangle^{-\zeta_2} \lesssim 1. \tag{2.20}$$

Writing $y = \frac{1}{2}(x+y) - \frac{1}{2}(x-y)$ it follows again from (2.1) that (2.20) is fulfilled provided $t_2 \leqslant \zeta_2$ and $t_1 \geqslant -\zeta_2$. We summarize: If $a \in M^{\infty,1}_{0,\tau_2,\zeta_1,\zeta_2}(\mathbf{R}^{2d})$ with $\tau_2, \zeta_2 < 0 \leqslant \zeta_1$ then the operator (2.16) is continuous for all $p, q \in [1,\infty]$ provided

$$\begin{aligned} &t_1 \geqslant -\zeta_2, \quad t_2 \leqslant \zeta_2, \\ &s_2 \leqslant s_1 + \tau_2, \quad |s_1| + |s_2| \leqslant \zeta_1. \end{aligned} \tag{2.21}$$

Finally we discuss Gabor frames [15, 39] for $L^2(\mathbf{R}^{2d})$ defined by the Gaussian window function

$$\Phi(X) = 2^d \pi^{d/2} \exp(-|X|^2), \quad X \in \mathbf{R}^{2d}. \tag{2.22}$$

Here we denote by

$$\Pi(X, Y) f(Z) = e^{2i\sigma(Y,Z)} f(Z - X), \quad X, Y, Z \in \mathbf{R}^{2d}, \tag{2.23}$$

the composition of the translation operator and the *symplectic* modulation operator $f \mapsto e^{2i\sigma(Y,\cdot)} f$, defined using the symplectic form [10]

$$\sigma((x, \xi), (y, \eta)) = \langle y, \xi \rangle - \langle x, \eta \rangle, \quad (x, \xi), (y, \eta) \in \mathbf{R}^{2d}.$$

If the parameters $a, b > 0$ satisfy $ab < \pi$ then $\{\Pi(an, bk)\Phi\}_{n,k\in\mathbf{Z}^{2d}}$ is a Gabor frame for $L^2(\mathbf{R}^{2d})$ which means that

$$A\|f\|_{L^2}^2 \leqslant \sum_{\mathbf{\Lambda}\in\Theta} |(f, \Pi(\mathbf{\Lambda})\Phi)|^2 \leqslant B\|f\|_{L^2}^2 \quad \forall f \in L^2(\mathbf{R}^{2d}),$$

for some $0 < A \leqslant B < \infty$. Here $\Theta = \{(an, bk)\}_{n,k\in\mathbf{Z}^{2d}} \subseteq \mathbf{R}^{4d}$ is a lattice determined by $a, b > 0$. We denote elements in the lattice as

$$\mathbf{\Lambda} = (\Lambda, \Lambda') \in \Theta, \quad \Lambda = an, \quad \Lambda' = bk, \quad n, k \in \mathbf{Z}^{2d}. \tag{2.24}$$

The Gabor frame operator $S = S(\Phi, \Theta)$, defined by

$$Sf = \sum_{\mathbf{\Lambda}\in\Theta} (f, \Pi(\mathbf{\Lambda})\Phi)\, \Pi(\mathbf{\Lambda})\Phi,$$

is positive and invertible on $L^2(\mathbf{R}^{2d})$. For any $f \in L^2(\mathbf{R}^{2d})$ we have a Gabor expansion

$$f = \sum_{\mathbf{\Lambda}\in\Theta} (f, \Pi(\mathbf{\Lambda})\widetilde{\Phi})\, \Pi(\mathbf{\Lambda})\Phi, \quad f \in L^2(\mathbf{R}^{2d}), \tag{2.25}$$

with unconditional convergence [15]. Here $\widetilde{\Phi} = S^{-1}\Phi \in \mathscr{S}(\mathbf{R}^{2d})$ [15, 22] is the *canonical dual window*.

Gabor theory has been generalized from the Hilbert space L^2 to modulation spaces by Feichtinger, Gröchenig and Leinert [8, 15, 16]. In fact if $\omega \in \mathscr{P}(\mathbf{R}^{4d})$ we have the norm equivalence

$$\begin{aligned} C^{-1}\|f\|_{M^{p,q}_{\omega}(\mathbf{R}^{2d})} &\leqslant \Big(\sum_{k\in\mathbf{Z}^{2d}} \Big(\sum_{n\in\mathbf{Z}^{2d}} |(f, \Pi(an, bk)\widetilde{\Phi})|^p \omega(an, bk)^p \Big)^{\frac{q}{p}} \Big)^{\frac{1}{q}} s \\ &\leqslant C\|f\|_{M^{p,q}_{\omega}(\mathbf{R}^{2d})} \end{aligned} \tag{2.26}$$

where $C > 0$, for the whole scale $1 \leqslant p, q \leqslant \infty$ of modulation spaces. The expansion (2.25) holds with unconditional convergence if $p, q < \infty$, and in the weak* topology of $M^\infty_{1/v}(\mathbf{R}^{2d})$ otherwise for some $v \in \mathscr{P}(\mathbf{R}^{4d})$.

2.5 Stochastic Processes and Generalized Stochastic Processes

Let Ω be a sample space equipped with a σ-algebra $\mathscr{B}$ of subsets of Ω and let $\mathbf{P}$ be a probability measure defined on $\mathscr{B}$. The space of $\mathbf{C}$-valued random variables is the Hilbert space $L^2(\Omega)$ equipped with the inner product

$$L^2(\Omega) \times L^2(\Omega) \ni (X, Y) \mapsto \mathbf{E}(X\overline{Y}) = (X, Y)_{L^2(\Omega)}$$

where

$$\mathbf{E}X = \int_\Omega X(\omega)\mathbf{P}(\mathrm{d}\omega)$$

is the expectation functional (integral). We write $X \perp Y$ if $\mathbf{E}(X\overline{Y}) = 0$. The Hilbert subspace of $L^2(\Omega)$ of zero mean random variables is denoted $L^2_0(\Omega)$, and thus $\mathbf{E}X = 0$ and $\mathbf{E}|X|^2 < \infty$ for each element $X \in L^2_0(\Omega)$.

A second order zero mean stochastic process is a locally Bochner integrable map $f : \mathbf{R}^d \to L^2_0(\Omega)$. This space of stochastic processes is denoted $L^1_{\mathrm{loc}}(\mathbf{R}^d, L^2_0(\Omega))$, and $C(\mathbf{R}^d, L^2_0(\Omega)) \subseteq L^1_{\mathrm{loc}}(\mathbf{R}^d, L^2_0(\Omega))$ denotes the subspace of continuous stochastic processes. The cross-covariance function of $f, g \in L^1_{\mathrm{loc}}(\mathbf{R}^d, L^2_0(\Omega))$ is

$$k_{fg}(x, y) = \mathbf{E}(f(x)\overline{g(y)}), \quad x, y \in \mathbf{R}^d.$$

The function $k_f = k_{ff}$ is the (auto-)covariance function of f. By the Cauchy–Schwarz inequality in $L^2(\Omega)$ we have

$$|k_{fg}(x, y)|^2 \leqslant k_f(x, x)\, k_g(y, y), \quad x, y \in \mathbf{R}^d,$$

which implies that k_{fg} extends to an element in $\mathscr{D}'(\mathbf{R}^{2d})$ [14] when $f, g \in L^1_{\mathrm{loc}}(\mathbf{R}^d, L^2_0(\Omega))$. Defining the cross-covariance operator

$$(\mathscr{K}_{fg}\varphi, \psi) = (k_{fg}, \psi \otimes \overline{\varphi})_{L^2(\mathbf{R}^{2d})}, \quad \varphi, \psi \in C^\infty_c(\mathbf{R}^d),$$

yields a continuous linear cross-covariance operator $\mathscr{K}_{fg} : C^\infty_c(\mathbf{R}^d) \to \mathscr{D}'(\mathbf{R}^d)$.

Since Fubini's theorem gives

$$(k_f, \varphi \otimes \overline{\varphi})_{L^2(\mathbf{R}^{2d})} = \mathbf{E}\left|\int_{\mathbf{R}^d} f(x)\overline{\varphi(x)}\,\mathrm{d}x\right|^2 \geqslant 0 \quad \forall \varphi \in C^\infty_c(\mathbf{R}^d),$$

it follows that k_f is the kernel of a non-negative linear continuous covariance operator $\mathscr{K}_f = \mathscr{K}_{ff} : C^\infty_c(\mathbf{R}^d) \to \mathscr{D}'(\mathbf{R}^d)$.

We use Gelfand and Vilenkin's concept of generalized stochastic process defined as follows [14], cf. [9, 21, 24].

Definition 2.2 A second order zero mean generalized stochastic process (GSP) $u \in \mathscr{L}(C_c^\infty(\mathbf{R}^d), L_0^2(\Omega))$ is a conjugate linear continuous operator $u : C_c^\infty(\mathbf{R}^d) \to L_0^2(\Omega)$, written as $(u, \varphi) \in L_0^2(\Omega)$ for $\varphi \in C_c^\infty(\mathbf{R}^d)$.

Thus for each $K \Subset \mathbf{R}^d$ there exist $C > 0$ and $k \in \mathbf{N}$ such that

$$\|(u, \varphi)\|_{L_0^2(\Omega)} \leqslant C \sum_{|\alpha| \leqslant k} \sup_{x \in \mathbf{R}^d} |\partial^\alpha \varphi(x)|, \quad \varphi \in C_c^\infty(K).$$

As in ordinary distribution theory [13, 20] a GSP is always differentiable as

$$(D^\alpha u, \varphi) = (u, D^\alpha \varphi), \quad \varphi \in C_c^\infty(\mathbf{R}^d), \quad \alpha \in \mathbf{N}^d.$$

For $u \in \mathscr{L}(C_c^\infty(\mathbf{R}^d), L_0^2(\Omega))$ we denote by $L_u^2(\Omega) \subseteq L_0^2(\Omega)$ the *time domain* [24] of u, i.e. the linear subspace of the closure of its image:

$$L_u^2(\Omega) = \text{closure}\,\{(u, \varphi),\ \varphi \in C_c^\infty(\mathbf{R}^d)\} \subseteq L_0^2(\Omega) \tag{2.27}$$

Let $u, v \in \mathscr{L}(C_c^\infty(\mathbf{R}^d), L_0^2(\Omega))$. The cross-covariance distribution is defined by

$$(k_{uv}, \varphi \otimes \overline{\psi}) = \mathbf{E}((u, \varphi)\overline{(v, \psi)}), \quad \varphi, \psi \in C_c^\infty(\mathbf{R}^d). \tag{2.28}$$

For each pair $K_1, K_2 \Subset \mathbf{R}^d$ there is $C > 0$ and $k_1, k_2 \in \mathbf{N}$ such that

$$|(k_{uv}, \varphi \otimes \overline{\psi})| \leqslant C \sum_{|\alpha| \leqslant k_1} \sup_{x \in \mathbf{R}^d} |\partial^\alpha \varphi(x)| \sum_{|\beta| \leqslant k_2} \sup_{x \in \mathbf{R}^d} |\partial^\beta \psi(x)|,$$
$$\varphi \in C_c^\infty(K_1), \quad \psi \in C_c^\infty(K_2).$$

Thus k_{uv} is a sesquilinear (conjugate linear in the first argument) continuous form on $C_c^\infty(\mathbf{R}^d) \times C_c^\infty(\mathbf{R}^d)$. If we equip $\mathscr{D}'(\mathbf{R}^d)$ with its weak* topology then

$$(\mathscr{K}_{uv}\psi, \varphi) = (k_{uv}, \varphi \otimes \overline{\psi}), \quad \varphi, \psi \in C_c^\infty(\mathbf{R}^d),$$

defines a linear continuous operator $\mathscr{K}_{uv} : C_c^\infty(\mathbf{R}^d) \to \mathscr{D}'(\mathbf{R}^d)$, called the cross-covariance operator. Note that $(\mathscr{K}_{uv}\psi, \varphi) = \overline{(\mathscr{K}_{vu}\varphi, \psi)}$. From the Schwartz kernel theorem [20, Theorem 5.2.1] it follows that $k_{uv} \in \mathscr{D}'(\mathbf{R}^{2d})$. If $k_{uv} \equiv 0$ then $\mathscr{K}_{uv} = 0$ and we say that u and v are uncorrelated. This means that

$$L_u^2(\Omega) \perp L_v^2(\Omega).$$

If $v = u$ then we call $k_u = k_{uu} \in \mathscr{D}'(\mathbf{R}^{2d})$ the (auto-)covariance distribution and $\mathscr{K}_u = \mathscr{K}_{uu} \in \mathscr{L}(C_c^\infty(\mathbf{R}^d), \mathscr{D}'(\mathbf{R}^d))$ the (auto-)covariance operator of u. Then

$(k_u, \varphi \otimes \overline{\varphi}) \geqslant 0$ for all $\varphi \in C_c^\infty(\mathbf{R}^d)$ so $\mathscr{K}_u \geqslant 0$ is non-negative in the sense of

$$(\mathscr{K}_u \varphi, \varphi) \geqslant 0 \quad \forall \varphi \in C_c^\infty(\mathbf{R}^d).$$

Remark 2.3 In [14, Chapter 3 §2.3] and [40, proof of Proposition 8.1] it is shown that any sesquilinear continuous form k on $C_c^\infty(\mathbf{R}^d) \times C_c^\infty(\mathbf{R}^d)$ which gives rise to a non-negative continuous operator $C_c^\infty(\mathbf{R}^d) \to \mathscr{D}'(\mathbf{R}^d)$ is the auto-covariance distribution of a Gaussian GSP $u \in \mathscr{L}(C_c^\infty(\mathbf{R}^d), L_0^2(\Omega))$ such that

$$\begin{aligned} \mathbf{E}\left((u, \varphi)\overline{(u, \psi)}\right) &= (k, \varphi \otimes \overline{\psi}), \\ \mathbf{E}\left((u, \varphi)(u, \psi)\right) &\equiv 0, \quad \varphi, \psi \in C_c^\infty(\mathbf{R}^d). \end{aligned}$$

Example 2.4 If $p > 0$ there exists a GSP $u \in \mathscr{L}(C_c^\infty(\mathbf{R}^d), L_0^2(\Omega))$ such that $(k_u, \varphi \otimes \overline{\psi}) = p(\psi, \varphi)_{L^2}$. This is a consequence of [14, Chapter 3 §2.3]. Then the covariance operator equals a positive multiple of the identity: $\mathscr{K}_u = pI$, and $k_u(x, y) = p\delta_0(x - y)$. This GSP is called white noise with power p, and is an example of a GSP which is not a stochastic process.

Example 2.5 Let $p > 0$, let u be a white noise GSP with power p, and let $\alpha \in \mathbf{N}^d$. Then $D^\alpha u$ is a GSP with covariance distribution

$$\begin{aligned} (k_{D^\alpha u}, \varphi \otimes \overline{\psi}) &= \mathbf{E}((u, D^\alpha \varphi)\overline{(u, D^\alpha \psi)}) = (k_u, D^\alpha \varphi \otimes \overline{D^\alpha \psi}) \\ &= p(D^\alpha \psi, D^\alpha \varphi)_{L^2} = p(D^{2\alpha}\psi, \varphi)_{L^2}, \quad \varphi, \psi \in C_c^\infty(\mathbf{R}^d), \end{aligned}$$

where we integrate by parts. Thus $\mathscr{K}_{D^\alpha u} = pD^{2\alpha}$, and the covariance distribution is

$$k_{D^\alpha u} = p\left(D^\alpha \otimes (-D)^\alpha\right)\left((1 \otimes \delta_0) \circ T^{-1}\right)$$

with the matrix (cf. (2.7))

$$T^{-1} = \begin{pmatrix} \frac{1}{2} I_d & \frac{1}{2} I_d \\ I_d & -I_d \end{pmatrix} \in \mathbf{R}^{2d \times 2d}.$$

In fact $k_u = p(1 \otimes \delta_0) \circ T^{-1}$ according to Example 2.4.

A stochastic process $f \in L^1_{\mathrm{loc}}(\mathbf{R}^d, L_0^2(\Omega))$ may be considered a generalized stochastic process in $\mathscr{L}(C_c^\infty(\mathbf{R}^d), L_0^2(\Omega))$ by means of

$$C_c^\infty(\mathbf{R}^d) \ni \varphi \mapsto (f, \varphi) = \int_{\mathbf{R}^d} f(x)\,\overline{\varphi(x)}\mathrm{d}x,$$

and then there is consistency between the covariance function and the covariance distribution as

$$(k_f, \psi \otimes \overline{\varphi})_{L^2(\mathbf{R}^{2d})} = \iint_{\mathbf{R}^{2d}} \mathbf{E}\left(f(x)\overline{f(y)}\right)\overline{\psi(x)}\varphi(y)\mathrm{d}x\,\mathrm{d}y = \mathbf{E}\left((f, \psi)\overline{(f, \varphi)}\right).$$

Remark 2.6 As in ordinary distribution theory [13, 20] a generalized stochastic process may extend to the domain $\mathscr{S}(\mathbf{R}^d) \supseteq C_c^\infty(\mathbf{R}^d)$ and is then called *tempered*. If u is a tempered GSP then $u \in \mathscr{L}(\mathscr{S}(\mathbf{R}^d), L_0^2(\Omega)) \subseteq \mathscr{L}(C_c^\infty(\mathbf{R}^d), L_0^2(\Omega))$. The Fourier transform of u can then be defined as $(\widehat{u}, \widehat{\varphi}) = (u, \varphi)$ for $\varphi \in \mathscr{S}(\mathbf{R}^d)$.

If $u \in \mathscr{L}(\mathscr{S}(\mathbf{R}^d), L_0^2(\Omega))$ is a tempered GSP then the covariance operator is continuous $\mathscr{K}_u : \mathscr{S}(\mathbf{R}^d) \to \mathscr{S}'(\mathbf{R}^d)$ and non-negative. Its Schwartz kernel $k_u \in \mathscr{S}'(\mathbf{R}^{2d})$ is a tempered distribution, and its Weyl symbol is denoted $a_u \in \mathscr{S}'(\mathbf{R}^{2d})$. The distributions $k_u, a_u \in \mathscr{S}'(\mathbf{R}^{2d})$ are connected by (2.8).

3 Filtering and the Equation for the Optimal Filter

Suppose a stochastic process g is a noisy observation of a message stochastic process f. A common assumption is the additive model

$$g = f + w,$$

where w is a noise stochastic process which is uncorrelated with f, i.e. $\mathbf{E}(f(x)\overline{w(y)}) = 0$ for all $x, y \in \mathbf{R}^d$. To recover f approximately from g we may try to filter g linearly using a linear operator F with kernel function k to get an optimal approximation of f, denoted f_o, [11, 42] as

$$f_o(x) = (Fg)(x) = \int_{\mathbf{R}^d} k(x, y)g(y)\mathrm{d}y. \tag{3.1}$$

This integral makes sense e.g. if we assume $k \in L^2(\mathbf{R}^{2d})$ and $g \in L^2(\mathbf{R}^d, L_0^2(\Omega))$.

The optimality of the approximation f_o refers to the postulate to pick a filter kernel k that minimizes the mean square error

$$\mathbf{E}|f(x) - f_o(x)|^2 = \|f(x) - f_o(x)\|_{L^2(\Omega)}^2$$

for all $x \in \mathbf{R}^d$.

The idea of filtering can be extended from stochastic processes to generalized stochastic processes as follows. Let $u, v \in \mathscr{L}(C_c^\infty(\mathbf{R}^d), L_0^2(\Omega))$ have auto-covariance distributions $k_u, k_v \in \mathscr{D}'(\mathbf{R}^{2d})$, respectively, and cross-covariance distribution $k_{uv} \in \mathscr{D}'(\mathbf{R}^{2d})$. These are sesquilinear continuous forms on $C_c^\infty(\mathbf{R}^d) \times C_c^\infty(\mathbf{R}^d)$.

Each of these kernels gives rise to continuous linear (auto-, cross-)covariance operators denoted $\mathscr{K}_u, \mathscr{K}_v, \mathscr{K}_{uv} : C_c^\infty(\mathbf{R}^d) \to \mathscr{D}'(\mathbf{R}^d)$ respectively. Then $\mathscr{K}_u \geqslant 0$ and $\mathscr{K}_v \geqslant 0$ on $C_c^\infty(\mathbf{R}^d)$.

We will assume that

$$F : C_c^\infty(\mathbf{R}^d) \to C_c^\infty(\mathbf{R}^d) \tag{3.2}$$

is a linear continuous operator, called *filter*. The adjoint F^* defined by

$$(F\varphi, \psi) = (\varphi, F^*\psi), \quad \varphi, \psi \in C_c^\infty(\mathbf{R}^d), \tag{3.3}$$

is then a continuous linear operator $C_c^\infty(\mathbf{R}^d) \to \mathscr{D}'(\mathbf{R}^d)$. We assume continuity of the adjoint:

$$F^* : C_c^\infty(\mathbf{R}^d) \to C_c^\infty(\mathbf{R}^d). \tag{3.4}$$

Replacing $\varphi \in C_c^\infty(\mathbf{R}^d)$ with $\varphi \in \mathscr{D}'(\mathbf{R}^d)$, the formula (3.3) extends F uniquely to a linear continuous operator on $\mathscr{D}'(\mathbf{R}^d)$, equipped with the weak* topology.

Given a linear operator F that satisfy (3.2), (3.4) and $v \in \mathscr{L}(C_c^\infty(\mathbf{R}^d), L_0^2(\Omega))$ we define the filtered generalized stochastic process $u_o \in \mathscr{L}(C_c^\infty(\mathbf{R}^d), L_0^2(\Omega))$ as $u_o = Fv$, that is

$$(u_o, \varphi) = (Fv, \varphi) = (v, F^*\varphi), \quad \varphi \in C_c^\infty(\mathbf{R}^d), \tag{3.5}$$

which extends (3.1) from stochastic processes $g \in L^1_{\mathrm{loc}}(\mathbf{R}^d, L_0^2(\Omega))$ to $v \in \mathscr{L}(C_c^\infty(\mathbf{R}^d), L_0^2(\Omega))$. To wit, if F has integral kernel k and $v \in L^1_{\mathrm{loc}}(\mathbf{R}^d, L_0^2(\Omega))$ then (3.5) reduces to

$$\iint_{\mathbf{R}^{2d}} k(x, y) v(y) \overline{\varphi(x)} \, \mathrm{d}x \, \mathrm{d}y.$$

Let $u, v \in \mathscr{L}(C_c^\infty(\mathbf{R}^d), L_0^2(\Omega))$. We assume that we have access to v but not to u. The GSP v is assumed to be a noise corrupted version of u, and we want to recover the latter with optimally small error using a filter F as in (3.5).

For a fixed arbitrary $\varphi \in C_c^\infty(\mathbf{R}^d) \setminus \{0\}$ we may derive an equation for the filter operator F that is optimal in the sense of minimizing the mean square error

$$\mathbf{E}|(u, \varphi) - (u_o, \varphi)|^2.$$

By (3.5) we have $(u_o, \varphi) \in L_v^2(\Omega)$ for any filter linear operator F that satisfies (3.2) and (3.4). We may formulate optimality as follows.

Proposition 3.1 *Let $u, v \in \mathscr{L}(C_c^\infty(\mathbf{R}^d), L_0^2(\Omega))$, let $\varphi \in C_c^\infty(\mathbf{R}^d) \setminus \{0\}$ be fixed, and suppose F is a linear operator that satisfy (3.2) and (3.4). The filter F in (3.5) is optimal for φ if and only if*

$$(u - u_o, \varphi) \perp L_v^2(\Omega). \tag{3.6}$$

Proof We may uniquely decompose $(u, \varphi) = X_0 + X_1$ where $X_0 \in L_v^2(\Omega)$ and $X_1 \in L_v^2(\Omega)^\perp$. By elementary Hilbert space theory X_0 is the unique optimal vector in $L_v^2(\Omega)$ that satisfies

$$X_0 = \arg\inf_{X \in L_v^2(\Omega)} \|(u, \varphi) - X\|_{L^2(\Omega)}, \tag{3.7}$$

and $\|X_1\|_{L^2(\Omega)} = \inf_{X \in L_v^2(\Omega)} \|(u, \varphi) - X\|_{L^2(\Omega)}$.

Suppose $(u - u_o, \varphi) \perp L^2_v(\Omega)$. Then $(u, \varphi) = (u_o, \varphi) + Y_1$ where $Y_1 \in L^2_v(\Omega)^\perp$. By (3.5) we have $(u_o, \varphi) \in L^2_v(\Omega)$, and it follows from the uniqueness of the decomposition $(u, \varphi) = X_0 + X_1$ that $Y_1 = X_1$ and $(u_o, \varphi) = X_0$. By (3.7) it thus follows that (u_o, φ) is optimal.

On the other hand, suppose that $(u - u_o, \varphi) \notin L^2_v(\Omega)^\perp$, i.e. $(u, \varphi) = (u_o, \varphi) + Y_0 + Y_1$ where $Y_0 \in L^2_v(\Omega)$, $Y_1 \in L^2_v(\Omega)^\perp$ and $Y_0 \neq 0$. Again by the uniqueness of the decomposition $(u, \varphi) = X_0 + X_1$ we have $(u_o, \varphi) = X_0 - Y_0$. There exists $\psi \in C^\infty_c(\mathbf{R}^d) \setminus 0$ such that $\|Y_0 - (v, \psi)\|^2_{L^2(\Omega)} < \|Y_0\|^2_{L^2(\Omega)}$.

If F_0^* is the linear continuous operator on $C^\infty_c(\mathbf{R}^d)$

$$F_0^* g = \|\varphi\|_{L^2}^{-2} (g, \varphi)\, \psi, \quad g \in C^\infty_c(\mathbf{R}^d),$$

then

$$F_0 g = \|\varphi\|_{L^2}^{-2} (g, \psi)\varphi, \quad g \in C^\infty_c(\mathbf{R}^d).$$

Thus F_0 and F_0^* are both continuous on $C^\infty_c(\mathbf{R}^d)$, and $F_0^*\varphi = \psi$. We have

$$\begin{aligned} \|(u, \varphi) - (v, (F + F_0)^*\varphi\|^2_{L^2(\Omega)} &= \|X_1 + (Y_0 - (v, \psi)\|^2_{L^2(\Omega)} \\ &= \|X_1\|^2_{L^2(\Omega)} + \|Y_0 - (v, \psi)\|^2_{L^2(\Omega)} \\ &< \|X_1\|^2_{L^2(\Omega)} + \|Y_0\|^2_{L^2(\Omega)} = \|(u - u_o, \varphi)\|^2_{L^2(\Omega)}, \end{aligned}$$

which means that the filter F in (u_o, φ) is not optimal. □

Condition (3.6) generalized to all $\varphi \in C^\infty_c(\mathbf{R}^d)$ can be expressed with (2.28) as

$$(k_{uv}, \varphi \otimes \overline{\psi}) = (k_v, F^*\varphi \otimes \overline{\psi}), \quad \forall \varphi, \psi \in C^\infty_c(\mathbf{R}^d).$$

Thus

$$(\mathscr{K}_{uv}\psi, \varphi) = (\mathscr{K}_v\psi, F^*\varphi) = (F\mathscr{K}_v\psi, \varphi) \quad \forall \varphi, \psi \in C^\infty_c(\mathbf{R}^d),$$

which means that

$$\mathscr{K}_{uv} = F\mathscr{K}_v \tag{3.8}$$

as operators $C^\infty_c(\mathbf{R}^d) \to \mathscr{D}'(\mathbf{R}^d)$. Given the operators $\mathscr{K}_{uv}, \mathscr{K}_v \in \mathscr{L}(C^\infty_c(\mathbf{R}^d), \mathscr{D}'(\mathbf{R}^d))$, this is an operator equation for the optimal linear filter F, which as noted above may be considered a continuous linear operator $F : \mathscr{D}'(\mathbf{R}^d) \to \mathscr{D}'(\mathbf{R}^d)$.

Remark 3.2 In [39] the operator equation (3.8) is deduced in a different functional framework. In fact we use $\mathscr{L}(M^1(\mathbf{R}^d), L^2_0(\Omega))$ as the class of GSPs, and the filters are pseudodifferential operators with symbols in modulation spaces. The

space $\mathscr{L}(M^1(\mathbf{R}^d), L^2_0(\Omega))$ of GSPs is studied more carefully in [9, 21, 24]. The test function space is Feichtinger's algebra $M^1(\mathbf{R}^d)$ which is a Fourier invariant Banach space of continuous integrable functions with integrable Fourier transform, and $\mathscr{S}(\mathbf{R}^d) \subseteq M^1(\mathbf{R}^d)$ is an embedding. This gives a smaller space of GSPs than the space of tempered GSPs: $\mathscr{L}(M^1(\mathbf{R}^d), L^2_0(\Omega)) \subseteq \mathscr{L}(\mathscr{S}(\mathbf{R}^d), L^2_0(\Omega))$. If $u \in \mathscr{L}(M^1(\mathbf{R}^d), L^2_0(\Omega))$ then the covariance operator is continuous $\mathscr{K}_u : M^1(\mathbf{R}^d) \to M^\infty(\mathbf{R}^d)$ with $M^\infty(\mathbf{R}^d)$ equipped with its weak* topology.

By the Pythagorean theorem in $L^2(\Omega)$ the optimal (minimal) mean square error for $\varphi \in C_c^\infty(\mathbf{R}^d)$ is

$$\begin{aligned} J(\varphi) &= \mathbf{E}|(u - u_o, \varphi)|^2 = \mathbf{E}|(u, \varphi)|^2 - \mathbf{E}|(u_o, \varphi)|^2 = \mathbf{E}|(u, \varphi)|^2 - \mathbf{E}|(v, F^*\varphi)|^2 \\ &= (k_u, \varphi \otimes \overline{\varphi}) - (k_v, F^*\varphi \otimes \overline{F^*\varphi}) \\ &= ((\mathscr{K}_u - F\mathscr{K}_v F^*)\varphi, \varphi) = ((\mathscr{K}_u - \mathscr{K}_{uv}F^*)\varphi, \varphi). \end{aligned} \tag{3.9}$$

Suppose now the more specific model, common in engineering applications,

$$v = u + w \tag{3.10}$$

where $u, w \in \mathscr{L}(C_c^\infty(\mathbf{R}^d), L^2_0(\Omega))$, u is considered a signal and w is considered as noise uncorrelated to u, that is $k_{uw} \equiv 0$. Then $\mathscr{K}_{uv} = \mathscr{K}_u$ and $\mathscr{K}_v = \mathscr{K}_u + \mathscr{K}_w$, since $k_v = k_u + k_w$. The operator equation (3.8) is then

$$\mathscr{K}_u = F(\mathscr{K}_u + \mathscr{K}_w) \tag{3.11}$$

in the cone of non-negative continuous linear operators $C_c^\infty(\mathbf{R}^d) \to \mathscr{D}'(\mathbf{R}^d)$. The minimal mean square error for $\varphi \in C_c^\infty(\mathbf{R}^d)$ is obtained from (3.9) and (3.11) as

$$\begin{aligned} J(\varphi) &= \overline{J(\varphi)} = \overline{(\mathscr{K}_u(I - F^*)\varphi, \varphi)} = (\mathscr{K}_u\varphi, (I - F^*)\varphi) \\ &= ((I - F)\mathscr{K}_u\varphi, \varphi) = (F\mathscr{K}_w\varphi, \varphi). \end{aligned} \tag{3.12}$$

As a tool to solve (3.11) we will use

Lemma 3.3 *If $T : \mathscr{S}(\mathbf{R}^d) \to \mathscr{S}'(\mathbf{R}^d)$ is a linear continuous operator then $T = 0$ if and only if $(Tf, f) = 0$ for all $f \in \mathscr{S}(\mathbf{R}^d)$.*

Proof The claim is an immediate consequence of the polarization identity

$$\begin{aligned} &(T(f + g), f + g) - (T(f - g), f - g) \\ &+i\Big((T(f + ig), f + ig) - (T(f - ig), f - ig)\Big) = 4(Tf, g). \end{aligned}$$

□

4 The Optimal Filter for Wide-Sense Stationary Generalized Stochastic Processes

4.1 Wide-Sense Stationary GSPs

A zero mean stochastic process f is said to be wide-sense stationary (WSS) if its covariance function satisfies $k_f(x, y) = \kappa_f(x - y)$ for a function $\kappa_f : \mathbf{R}^d \to \mathbf{C}$. This means that the stochastic process is second order translation invariant. The function κ_f is non-negative definite in the following sense:

$$\sum_{j,k=1}^{n} \kappa_f(x_j - x_k) z_j \bar{z}_k \geqslant 0 \quad \forall \{x_j\}_{j=1}^n \subseteq \mathbf{R}^d, \ \{z_j\}_{j=1}^n \subseteq \mathbf{C}, \quad n \in \mathbf{N} \setminus 0.$$

Bochner's theorem [23, 31] says that a function $\kappa : \mathbf{R}^d \to \mathbf{C}$ is continuous and non-negative definite if and only if $\mu = \mathscr{F}\kappa$ is a non-negative bounded Radon measure on $\mathbf{R}^d$.

Let $u \in \mathscr{L}(C_c^\infty(\mathbf{R}^d), L_0^2(\Omega))$ be a GSP. Then $k_u \in \mathscr{D}'(\mathbf{R}^{2d})$ as explained in Section 2.5. We call u WSS if its covariance distribution k_u is translation invariant:

$$(k_u, T_x\varphi \otimes T_x\overline{\psi}) = (k_u, \varphi \otimes \overline{\psi}) \quad \forall \varphi, \psi \in C_c^\infty(\mathbf{R}^d) \quad \forall x \in \mathbf{R}^d,$$

cf. [14, Chapter 3, §3]. This means that the covariance operator is translation invariant as $T_x \mathscr{K}_u = \mathscr{K}_u T_x$ for all $x \in \mathbf{R}^d$. Under the assumption WSS there exists $\kappa_u \in \mathscr{D}'(\mathbf{R}^d)$ such that

$$(k_u, \varphi \otimes \overline{\psi}) = (\kappa_u, \varphi * \psi^*) \quad \forall \varphi, \psi \in C_c^\infty(\mathbf{R}^d),$$

where $\psi^*(x) = \overline{\psi(-x)}$, see [14, Chapter 2, §3.5] and [20, Theorem 3.1.4′]. By the Bochner–Schwartz theorem [14, Chapter 2, §3.3, Theorem 3] there exists a spectral non-negative tempered Radon measure $\mu_u = (2\pi)^{\frac{d}{2}} \widehat{\kappa}_u \in \mathscr{M}_+(\mathbf{R}^d)$ that satisfies (2.3) for some $s \geqslant 0$. Hence (cf. (2.2))

$$(\kappa_u, \varphi * \psi^*) = \int_{\mathbf{R}^d} \widehat{\psi}(\xi)\, \overline{\widehat{\varphi}(\xi)}\, \mathrm{d}\mu_u(\xi).$$

This gives for $\varphi, \psi \in C_c^\infty(\mathbf{R}^d)$

$$\mathbf{E}\left((u, \varphi)\overline{(u, \psi)}\right) = (k_u, \varphi \otimes \overline{\psi}) = (\mathscr{K}_u \psi, \varphi) = (\kappa_u, \varphi * \psi^*) = (\psi, \varphi)_{\mathscr{F}L^2(\mu_u)} \tag{4.1}$$

and in particular

$$\|(u, \varphi)\|_{L^2(\Omega)}^2 = \int_{\mathbf{R}^d} |\widehat{\varphi}(\xi)|^2 \, \mathrm{d}\mu_u(\xi). \tag{4.2}$$

Let $u \in \mathscr{L}(C_c^\infty(\mathbf{R}^d), L_0^2(\Omega))$ be WSS and denote by $\mu_u \in \mathscr{M}_+(\mathbf{R}^d)$ the corresponding spectral non-negative tempered Radon measure. The identity (4.1) implies that the linear map $C_c^\infty(\mathbf{R}^d) \ni \varphi \mapsto \overline{(u, \varphi)}$ extends uniquely to a unitary operator between Hilbert spaces $\mathscr{F}L^2(\mu_u) \to L_u^2(\Omega)$ (cf. (2.27)), and $\mathscr{K}_u$ extends uniquely to the identity operator on $\mathscr{F}L^2(\mu_u)$. Alternatively we may regard $\mathscr{K}_u$ as the convolution operator

$$\mathscr{K}_u f = \kappa_u * f \tag{4.3}$$

and we may extend the domain to $f \in \mathscr{S}(\mathbf{R}^d)$. (Note that $\kappa_u \in \mathscr{S}'(\mathbf{R}^d)$.) This yields a continuous operator

$$\mathscr{K}_u : \mathscr{S}(\mathbf{R}^d) \to \left(C^\infty \cap \mathscr{S}'\right)(\mathbf{R}^d). \tag{4.4}$$

Remark 4.1 For any given $\mu \in \mathscr{M}_+(\mathbf{R}^d)$ there exists a WSS GSP u such that $\mu_u = \mu$. This is a consequence of Remark 2.3.

Remark 4.2 If u is a WSS GSP then from Remark 2.6 and $\mathscr{S}(\mathbf{R}^d) \subseteq \mathscr{F}L^2(\mu_u)$ it follows that u is tempered, possess a Fourier transform $\widehat{u} : L^2(\mu_u) \to L_u^2(\Omega)$ such that $\varphi \mapsto \overline{(\widehat{u}, \varphi)}$ is unitary, and $\operatorname{supp} \widehat{u} = \operatorname{supp} \mu_u$.

Remark 4.3 If u is a WSS GSP and $\alpha \in \mathbf{N}^d$ then (4.1) gives

$$\left(k_{D^\alpha u}, \varphi \otimes \overline{\psi}\right) = (k_u, D^\alpha \varphi \otimes \overline{D^\alpha \psi}) = \int_{\mathbf{R}^d} \widehat{\psi}(\xi)\, \overline{\widehat{\varphi}(\xi)} \xi^{2\alpha}\, \mathrm{d}\mu_u(\xi).$$

Thus $D^\alpha u$ is WSS and its spectral measure is $\mathrm{d}\mu_{D^\alpha u} = \xi^{2\alpha}\, \mathrm{d}\mu_u$.

Remark 4.4 Consider the framework $\mathscr{L}(M^1(\mathbf{R}^d), L_0^2(\Omega))$ with test functions in Feichtinger's algebra, cf. Remark 3.2. If $u \in \mathscr{L}(M^1(\mathbf{R}^d), L_0^2(\Omega))$ is WSS then its spectral non-negative measure μ_u is translation bounded [9, Corollary 6] which means that

$$\sup_{x \in \mathbf{R}^d} |(\mu_u, T_x \varphi)| < \infty \quad \forall \varphi \in C_c(\mathbf{R}^d).$$

Remark 4.5 If $s = 0$ in (2.3) then the measure $\mu_u \in \mathscr{M}_+(\mathbf{R}^d)$ is bounded. This implies that

$$\kappa_u(x) = (2\pi)^{-d} \int_{\mathbf{R}^d} e^{i\langle x, \xi\rangle} \mathrm{d}\mu_u(\xi) = (2\pi)^{-\frac{d}{2}} \mathscr{F}^{-1} \mu_u(x) \tag{4.5}$$

is a function in $(C \cap L^\infty)(\mathbf{R}^d)$. In turn this means that the initial assumption $u \in \mathscr{L}(\mathscr{S}(\mathbf{R}^d), L_0^2(\Omega))$ can be strengthened to $u \in C(\mathbf{R}^d, L_0^2(\Omega))$, that is u is a continuous stochastic process.

Remark 4.6 If the measure $\mu_u \in \mathscr{M}_+(\mathbf{R}^d)$ is absolutely continuous with respect to Lebesgue measure then we may write $\mathrm{d}\mu_u(\xi) = f_u(\xi)\mathrm{d}\xi$ where $f_u \in L^1_{\mathrm{loc}}(\mathbf{R}^d)$ [10]. If further $f_u \in L^1(\mathbf{R}^d)$ then $\mu_u \in \mathscr{M}_+(\mathbf{R}^d)$ is finite. By (4.5) $\kappa_u = (2\pi)^{-\frac{d}{2}}\mathscr{F}^{-1}f_u \in C_0(\mathbf{R}^d)$.

Suppose u is a WSS GSP with spectral measure $\mu_u \in \mathscr{M}_+(\mathbf{R}^d)$, and suppose $\mu \in \mathscr{M}_+(\mathbf{R}^d)$ satisfies $\mu \geqslant \mu_u$. From (4.1) and the Cauchy–Schwarz inequality in $L^2(\mu)$ we get

$$|(k_u, \varphi \otimes \overline{\psi})| = |(\mathscr{K}_u\psi, \varphi)| = |(\psi, \varphi)_{\mathscr{F}L^2(\mu_u)}| \leqslant \|\psi\|_{\mathscr{F}L^2(\mu)}\,\|\varphi\|_{\mathscr{F}L^2(\mu)}. \quad (4.6)$$

Thus k_u extends to a sesquilinear form on $\mathscr{F}L^2(\mu) \times \mathscr{F}L^2(\mu)$. From [29, Corollary II.2] we get the following conclusion. There exists a unique $\mathscr{T}_{u,\mu} \in \mathscr{L}(\mathscr{F}L^2(\mu))$ such that $\|\mathscr{T}_{u,\mu}\|_{\mathscr{L}(\mathscr{F}L^2(\mu))} \leqslant 1$ and

$$(\mathscr{K}_u\psi, \varphi) = (\mathscr{T}_{u,\mu}\psi, \varphi)_{\mathscr{F}L^2(\mu)}, \quad \psi, \varphi \in \mathscr{F}L^2(\mu). \quad (4.7)$$

If $\mu = \mu_u$ then $\mathscr{T}_{u,\mu} = \mathscr{K}_u = \mathrm{id}_{\mathscr{F}L^2(\mu_u)}$ as already observed.

Example 4.7 For the white noise GSP u in Example 2.4 we have $(k_u, \varphi \otimes \overline{\psi}) = p(\psi, \varphi)_{L^2} = p(\widehat{\psi}, \widehat{\varphi})_{L^2}$ by Plancherel's theorem. The corresponding spectral tempered measure is therefore a positive multiple of Lebesgue measure $\mathrm{d}\mu_u = p\,\mathrm{d}\xi$ on $\mathbf{R}^d$.

Example 4.8 Let u be white noise with power $p > 0$ as in Example 2.4 and let $\alpha \in \mathbf{N}^d$. Remark 4.3 combined with Example 4.7 gives $\mu_{D^\alpha u} = p\,\xi^{2\alpha}\mathrm{d}\xi$, and the covariance operator is $\mathscr{K}_{D^\alpha u} = pD^{2\alpha}$. This example reveals that although the covariance operator $\mathscr{K}_u$ of a WSS GSP u always extends to the identity operator on $\mathscr{F}L^2(\mu_u)$, it may happen that $\mathscr{K}_u$ fails to be continuous on $L^2(\mathbf{R}^d)$. In fact in this example $\mathscr{K}_u$ is unbounded considered as an operator in $L^2(\mathbf{R}^d)$, equipped with domain $\mathscr{S}(\mathbf{R}^d)$.

Remark 4.9 In general it is not possible to consider $\mathscr{K}_u$ as an unbounded operator in $L^2(\mathbf{R}^d)$. In fact suppose $\mu_u = \delta_0 \in \mathscr{M}_+(\mathbf{R}^d)$. Then $\kappa_u = (2\pi)^{-\frac{d}{2}}\mathscr{F}^{-1}\mu_u = (2\pi)^{-d}$ which implies that $\mathscr{K}_u f(x) = \kappa_u * f(x) = (2\pi)^{-d}\int_{\mathbf{R}^d} f(y)\mathrm{d}y$. Thus $\mathscr{K}_u f \in L^2(\mathbf{R}^d)$ can happen only for functions f with integral zero, and then $\mathscr{K}_u f = 0$. There is no interesting function space (domain) for f for which $\mathscr{K}_u f \in L^2(\mathbf{R}^d) \setminus \{0\}$. For the stochastic process $u \in C(\mathbf{R}^d, L^2_0(\Omega))$ that corresponds to $\mu_u = \delta_0$ we have $\mathbf{E}|u(x) - u(0)|^2 = 0$ so $u(x) = u_0 \in L^2_0(\Omega)$ is constant for all $x \in \mathbf{R}^d$.

Example 4.10 We return to Example 4.8 and investigate the corresponding space $\mathscr{F}L^2(\mu_u)$. If $p = d = 1$ and $n \in \mathbf{N}$ then $\mu = \mu_{D^n u} = |\xi|^{2n}\mathrm{d}\xi$, and hence

$$\|f\|^2_{\mathscr{F}L^2(\mu)} = \int_{\mathbf{R}} |\widehat{f}(\xi)|^2|\xi|^{2n}\mathrm{d}\xi.$$

This is the square norm of the homogeneous Hilbert Sobolev space $\dot{H}^2_n(\mathbf{R})$ [1]. If $s \in \mathbf{R}$ then by Remark 4.1 there exists a WSS tempered GSP u on $\mathbf{R}^d$ such that

$\mu_u = \langle\xi\rangle^{2s}\mathrm{d}\xi$. In this case $\mathscr{F}L^2(\mu) = H_s^2(\mathbf{R}^d)$ which denotes the usual Hilbert Sobolev space [1].

4.2 Filtering of WSS GSPs

Let $v \in \mathscr{L}(C_c^\infty(\mathbf{R}^d), L_0^2(\Omega))$ be WSS so that $C_c^\infty(\mathbf{R}^d) \ni \varphi \mapsto \overline{(v,\varphi)}$ extends to a unitary map $\mathscr{F}L^2(\mu_v) \to L_v^2(\Omega)$ where $\mu_v \in \mathscr{M}_+(\mathbf{R}^d)$ is the spectral Radon measure corresponding to v. Let $f \in \mathscr{F}L^\infty(\mu_v)$. If $\varphi \in \mathscr{F}L^2(\mu_v)$ then $\widehat{f}\,\widehat{\varphi} \in L^2(\mu_v)$, that is $\mathscr{F}^{-1}(\widehat{f}\,\widehat{\varphi}) \in \mathscr{F}L^2(\mu_v)$. Defining a linear filter operator F as the convolution

$$F\varphi = (2\pi)^{-\frac{d}{2}} f * \varphi = \mathscr{F}^{-1}(\widehat{f}\,\widehat{\varphi}), \quad \varphi \in \mathscr{F}L^2(\mu_v), \tag{4.8}$$

cf. (2.2), gives a translation invariant continuous operator on $\mathscr{F}L^2(\mu_v)$. In fact

$$(2\pi)^{-\frac{d}{2}}\|f * \varphi\|_{\mathscr{F}L^2(\mu_v)} \leqslant \|f\|_{\mathscr{F}L^\infty(\mu_v)}\|\varphi\|_{\mathscr{F}L^2(\mu_v)}. \tag{4.9}$$

If $\varphi, \psi \in \mathscr{F}L^2(\mu_v)$ then we have, using $\widehat{f^*} = \overline{\widehat{f}}$,

$$\begin{aligned}(F\varphi, \psi)_{\mathscr{F}L^2(\mu_v)} &= \int_{\mathbf{R}^d} \widehat{f}(\xi)\,\widehat{\varphi}(\xi)\,\overline{\widehat{\psi}(\xi)}\,\mathrm{d}\mu_v(\xi)\\ &= \int_{\mathbf{R}^d} \widehat{\varphi}(\xi)\,\overline{\widehat{f^*}(\xi)\,\widehat{\psi}(\xi)}\,\mathrm{d}\mu_v(\xi)\\ &= (2\pi)^{-\frac{d}{2}}(\varphi, f^* * \psi)_{\mathscr{F}L^2(\mu_v)}\end{aligned}$$

which means that the adjoint of F with respect to $(\cdot,\cdot)_{\mathscr{F}L^2(\mu_v)}$ is

$$F^*\varphi = (2\pi)^{-\frac{d}{2}} f^* * \varphi. \tag{4.10}$$

We may hence define a filtered GSP $u_o \in \mathscr{L}(C_c^\infty(\mathbf{R}^d), L_0^2(\Omega))$, cf. (3.5), as

$$(u_o, \varphi) = (2\pi)^{-\frac{d}{2}}(v, f^* * \varphi), \quad \varphi \in \mathscr{F}L^2(\mu_v). \tag{4.11}$$

The filtered GSP u_o is WSS. Indeed since convolution is translation invariant as $g * T_x\varphi = T_x(g * \varphi)$, and v is assumed to be WSS, we have for any $x \in \mathbf{R}^d$ and $\varphi, \psi \in C_c^\infty(\mathbf{R}^d)$

$$\begin{aligned}(k_{u_o}, T_x\varphi \otimes T_x\overline{\psi}) &= \mathbf{E}\left((u_o, T_x\varphi)\overline{(u_o, T_x\psi)}\right)\\ &= (2\pi)^{-d}\,\mathbf{E}\left((v, f^* * T_x\varphi)\overline{(v, f^* * T_x\psi)}\right)\\ &= (2\pi)^{-d}\,\mathbf{E}\left((v, f^* * \varphi)\overline{(v, f^* * \psi)}\right) = (k_{u_o}, \varphi \otimes \overline{\psi}).\end{aligned}$$

4.3 The Optimal Filter for WSS GSPs

The next lemma will be needed in the proof of Theorem 4.12.

Lemma 4.11 *Let $\mu \in \mathscr{M}_+(\mathbf{R}^d)$. If $f \in L^1_{\mathrm{loc}}(\mu)$ is real-valued then*

$$\int_A f(x)\,\mathrm{d}\mu \geqslant 0 \quad \forall A \in \mathscr{B}(\mathbf{R}^d) \quad \Longrightarrow \quad f(x) \geqslant 0 \text{ for } \mu\text{-a.e. } x \in \mathbf{R}^d.$$

Proof Set $N = \{x \in \mathbf{R}^d : f(x) \leqslant 0\} \in \mathscr{B}(\mathbf{R}^d)$. Then

$$0 \leqslant \int_N f(x)\,\mathrm{d}\mu \leqslant 0$$

that is $\int_N f(x)\,\mathrm{d}\mu = 0$. Now [32, Theorem 1.39 (a)] implies $f(x) = 0$ for μ-a.e. $x \in N$. □

The following result is the main statement of this section. It reveals that the optimal filter for the uncorrelated additive noise problem for WSS GSPs is a Radon–Nikodym derivative in the frequency domain. It formalizes the well established convention in engineering to think of the Wiener filter in the frequency domain as a fraction of spectral densities [11, 38, 41, 42].

Theorem 4.12 *Suppose that $u, w \in \mathscr{L}(C_c^\infty(\mathbf{R}^d), L_0^2(\Omega))$ are two uncorrelated WSS generalized stochastic processes, let the corresponding spectral tempered non-negative Radon measures be denoted by $\mu_u, \mu_w \in \mathscr{M}_+(\mathbf{R}^d)$ respectively, and let $v = u + w$.*

The unique convolution filter $f \in \mathscr{F}L^\infty(\mu_u + \mu_w)$ that applied to v as in (4.11) which is mean square error optimal to recover u is $f = \mathscr{F}^{-1}\widehat{f} \in \mathscr{S}'(\mathbf{R}^d)$ where $\widehat{f}$ is the Radon–Nikodym derivative given by

$$\mu_u = \widehat{f}\,(\mu_u + \mu_w). \tag{4.12}$$

We have $0 \leqslant \widehat{f}(\xi) \leqslant 1$ for $(\mu_u + \mu_w)$-a.e. $\xi \in \mathbf{R}^d$,

$$\operatorname{ess\,supp} \widehat{f} = \operatorname{supp} \mu_u, \tag{4.13}$$

and the minimal error for $\varphi \in \mathscr{F}L^2(\mu_u + \mu_w)$ is

$$J(\varphi) = \mathbf{E}\,|(u - u_o, \varphi)|^2 = \int_{\mathbf{R}^d} \widehat{f}(\xi)\,|\widehat{\varphi}(\xi)|^2\,\mathrm{d}\mu_w = \int_{\mathbf{R}^d} \big(1 - \widehat{f}(\xi)\big)\,|\widehat{\varphi}(\xi)|^2\,\mathrm{d}\mu_u. \tag{4.14}$$

Proof We denote the auto-covariance operators for u and w by $\mathscr{K}_u$ and $\mathscr{K}_w$ respectively, and we set $\mu = \mu_u + \mu_w$ with domain $\mathscr{B}(\mathbf{R}^d)$. The measures μ_u, μ_w and μ satisfy (2.3) for some common $s \geqslant 0$. All three measures being non-negative entails

$$\mathscr{F}L^2(\mu) \subseteq \mathscr{F}L^2(\mu_u) \bigcap \mathscr{F}L^2(\mu_w).$$

From (4.2) it follows that for $\varphi \in C_c^\infty(\mathbf{R}^d)$ we have

$$\|(u,\varphi)\|_{L^2(\Omega)} \leqslant \|\varphi\|_{\mathscr{F}L^2(\mu)}, \quad \|(w,\varphi)\|_{L^2(\Omega)} \leqslant \|\varphi\|_{\mathscr{F}L^2(\mu)},$$

which implies that u and w restricts to continuous linear maps $u : \mathscr{F}L^2(\mu) \to L^2_u(\Omega)$ and $w : \mathscr{F}L^2(\mu) \to L^2_w(\Omega)$, respectively. From (4.6) we obtain

$$\begin{aligned} |(\mathscr{K}_u\psi,\varphi)| &\leqslant \|\psi\|_{\mathscr{F}L^2(\mu)}\,\|\varphi\|_{\mathscr{F}L^2(\mu)}, \\ |(\mathscr{K}_w\psi,\varphi)| &\leqslant \|\psi\|_{\mathscr{F}L^2(\mu)}\,\|\varphi\|_{\mathscr{F}L^2(\mu)}, \quad \varphi,\psi \in \mathscr{F}L^2(\mu), \end{aligned}$$

and by (4.7) and the preceding discussion we may regard $\mathscr{K}_u = \mathscr{T}_{u,\mu}$ and $\mathscr{K}_w = \mathscr{T}_{w,\mu}$ as bounded operators on $\mathscr{F}L^2(\mu)$ with operator norm upper bounded by one. Defining F by (4.8) for $f \in \mathscr{F}L^\infty(\mu)$ we know by (4.9), (4.10) that $\|F\|_{\mathscr{L}(\mathscr{F}L^2(\mu))} \leqslant \|f\|_{\mathscr{F}L^\infty(\mu)}$ and $\|F^*\|_{\mathscr{L}(\mathscr{F}L^2(\mu))} \leqslant \|f\|_{\mathscr{F}L^\infty(\mu)}$.

These considerations lead to the conclusion that we may consider the operator equation (3.11), that is $\mathscr{K}_u = F(\mathscr{K}_u + \mathscr{K}_w)$, for the optimal filter F as an equation for bounded operators on $\mathscr{F}L^2(\mu)$. By Lemma 3.3 and (4.10) this equation is equivalent to the equality

$$\begin{aligned} \int_{\mathbf{R}^d} |\widehat{\varphi}(\xi)|^2 \,\mathrm{d}\mu_u(\xi) &= (\mathscr{K}_u\varphi,\varphi) = (F(\mathscr{K}_u+\mathscr{K}_w)\varphi,\varphi) = (2\pi)^{-\frac{d}{2}}((\mathscr{K}_u+\mathscr{K}_w)\varphi, f^**\varphi) \\ &= (2\pi)^{-\frac{d}{2}}\left((\mathscr{K}_u\varphi, f^**\varphi)_{\mathscr{F}L^2(\mu_u)} + (\mathscr{K}_w\varphi, f^**\varphi)_{\mathscr{F}L^2(\mu_w)}\right) \\ &= (2\pi)^{-\frac{d}{2}}\int_{\mathbf{R}^d} \widehat{\varphi}(\xi)\,\overline{\widehat{f^**\varphi}(\xi)}\,\mathrm{d}\mu(\xi) \\ &= \int_{\mathbf{R}^d} |\widehat{\varphi}(\xi)|^2\,\widehat{f}(\xi)\,\mathrm{d}\mu(\xi) \end{aligned} \tag{4.15}$$

for all $\varphi \in \mathscr{S}(\mathbf{R}^d)$. Due to the density of $\mathscr{F}C_c^\infty(\mathbf{R}^d)$ in $\mathscr{S}(\mathbf{R}^d)$, the equation (4.15) for all $\varphi \in \mathscr{S}(\mathbf{R}^d)$ is equivalent to the same equation for all $\varphi \in \mathscr{F}C_c^\infty(\mathbf{R}^d)$.

Again the equation (4.15) for all $\widehat{\varphi} \in C_c^\infty(\mathbf{R}^d)$ is equivalent to the same equation with $|\widehat{\varphi}(\xi)|^2$ replaced by any non-negative function $g \in C_c(\mathbf{R}^d)$. In fact let $\psi \in C_c^\infty(\mathbf{R}^d)$ fulfill $\operatorname{supp}\psi \subseteq \mathrm{B}_1, \psi \geqslant 0, \int \psi\,\mathrm{d}x = 1$, set $\psi_n(x) = n^d\psi(xn)$ for $n \in \mathbf{N}\backslash 0$, let $\chi \in C_c^\infty(\mathbf{R}^d)$ satisfy $0 \leqslant \chi \leqslant 1$, $\chi|_{\operatorname{supp} g} \equiv 1$, and set

$$g_n = \chi\sqrt{g * \psi_n + \frac{1}{n}} \in C_c^\infty(\mathbf{R}^d), \quad n \in \mathbf{N}\setminus 0.$$

Then

$$\sup_{\xi\in\mathbf{R}^d} \left||g_n(\xi)|^2 - g(\xi)\right| \to 0, \quad n \to \infty,$$

which implies that (4.15) holds for all $\varphi \in \mathscr{F}C_c^\infty(\mathbf{R}^d)$ if and only if

$$\int_{\mathbf{R}^d} g(\xi)\,\mathrm{d}\mu_u(\xi) = \int_{\mathbf{R}^d} g(\xi)\,\widehat{f}(\xi)\,\mathrm{d}\mu(\xi)$$

holds for all non-negative functions $g \in C_c(\mathbf{R}^d)$. This means that the operator equation (3.11) reduces to the equation for tempered measures

$$\mu_u = \widehat{f}\,\mu = \widehat{f}\,(\mu_u + \mu_w)\,. \tag{4.16}$$

Since μ_u, μ_w are non-negative measures, it follows that μ_u is absolutely continuous with respect to μ. By the Radon–Nikodym theorem [10, Theorem 3.8] there exists a μ-measurable function $\widehat{f} \geqslant 0$ μ-a.e., with $\widehat{f} \in L^1_{\mathrm{loc}}(\mu)$, uniquely determined μ-a.e., such that (4.16) is satisfied in the sense of

$$\mu_u(A) = \int_A \widehat{f}(\xi)\,\mathrm{d}\mu(\xi), \quad A \in \mathscr{B}(\mathbf{R}^d). \tag{4.17}$$

We have now shown (4.12).

If $A \in \mathscr{B}(\mathbf{R}^d)$ then

$$\int_A \mathrm{d}\mu(\xi) \geqslant \int_A \mathrm{d}\mu_u(\xi) = \mu_u(A) = \int_A \widehat{f}(\xi)\,\mathrm{d}\mu(\xi)$$

that is

$$\int_A (1 - \widehat{f}(\xi))\mathrm{d}\mu(\xi) \geqslant 0, \quad A \in \mathscr{B}(\mathbf{R}^d).$$

An application of Lemma 4.11 yields $\widehat{f}(\xi) \leqslant 1$ μ-a.e. It follows that $\widehat{f} \in L^\infty(\mu)$, that is $f \in \mathscr{F}L^\infty(\mu)$. Thus the Radon–Nikodym derivative $\widehat{f}$ is indeed the unique solution in $\mathscr{F}L^\infty(\mu)$ to (4.16).

In order to show (4.13) let $\Omega \subseteq \mathbf{R}^d \setminus \operatorname{ess\,supp} \widehat{f}$ be open. Then $\widehat{f}\big|_\Omega = 0$ μ-a.e. [27] which implies $\mu_u(\Omega) = 0$ by (4.17), and thus $\Omega \subseteq \mathbf{R}^d \setminus \operatorname{supp} \mu_u$. This shows

$$\operatorname{ess\,supp} \widehat{f} \supseteq \operatorname{supp} \mu_u. \tag{4.18}$$

If instead $\Omega \subseteq \mathbf{R}^d \setminus \operatorname{supp} \mu_u$ is open then $\mu_u(\Omega) = 0$. Again from (4.17) combined with [10, Proposition 2.16] it follows that $\widehat{f}\big|_\Omega = 0$ μ-a.e. Hence $\Omega \subseteq \mathbf{R}^d \setminus \operatorname{ess\,supp} \widehat{f}$ and we have shown

$$\operatorname{ess\,supp} \widehat{f} \subseteq \operatorname{supp} \mu_u. \tag{4.19}$$

Now (4.13) is a consequence of (4.18) and (4.19).

Finally we obtain from (3.12) for $\varphi \in \mathscr{F}L^2(\mu)$, using (3.11), (4.1), (4.10) and (4.16)

$$\begin{aligned} J(\varphi) &= ((I - F)\mathscr{K}_u\varphi, \varphi) = (\mathscr{K}_u\varphi, (I - F^*)\varphi) = (\varphi, (I - F^*)\varphi)_{\mathscr{F}L^2(\mu_u)} \\ &= \int_{\mathbf{R}^d} \big(1 - \widehat{f}(\xi)\big)\,|\widehat{\varphi}(\xi)|^2\,\mathrm{d}\mu_u \end{aligned}$$

$$= \int_{\mathbf{R}^d} \widehat{f}(\xi)|\widehat{\varphi}(\xi)|^2 \, \mathrm{d}\mu_w$$

which proves (4.14). □

Remark 4.13 Note that $J(T_x\varphi) = J(\varphi)$, that is J is translation invariant. This feature is an intuitively natural consequence of the WSS assumption. The frequency support of $\varphi \in \mathscr{F}C_c(\mathbf{R}^d) \subseteq \mathscr{F}L^2(\mu)$ affects the minimal error as follows:

$$\begin{aligned} \operatorname{supp}\widehat{\varphi} \cap \operatorname{supp}\mu_w = \emptyset \quad &\Longrightarrow \quad J(\varphi) = 0, \\ \operatorname{supp}\widehat{\varphi} \cap \operatorname{supp}\mu_u = \emptyset \quad &\Longrightarrow \quad J(\varphi) = 0. \end{aligned}$$

The following consequence of Theorem 4.12 states a sufficient condition for perfect reconstruction (zero error) of the GSP u.

Corollary 4.14 *Under the assumptions of Theorem 4.12, and*

$$\operatorname{supp}\mu_u \cap \operatorname{supp}\mu_w = \emptyset$$

we have $\widehat{f} = \chi_{\operatorname{supp}\mu_u}$ *and* $J(\varphi) = 0$ *for all* $\varphi \in \mathscr{F}L^2(\mu)$.

Proof By (4.13) we have $\operatorname{ess\,supp}\widehat{f} = \operatorname{supp}\mu_u$ and (4.14) then yields $J(\varphi) = 0$ for all $\varphi \in \mathscr{F}L^2(\mu)$. The claim $\widehat{f} = \chi_{\operatorname{supp}\mu_u}$ follows from (4.12) and $\operatorname{ess\,supp}\widehat{f} \cap \operatorname{supp}\mu_w = \emptyset$. □

Example 4.15 Suppose as in Remark 4.6 that the measure μ_u is absolutely continuous with respect to Lebesgue measure on $\mathbf{R}^d$, that is $\mathrm{d}\mu_u = f_u \mathrm{d}\xi$ for some $f_u \in L^1_{\mathrm{loc}}(\mathbf{R}^d)$. If the GSP w is white noise with power $p > 0$ then by Theorem 4.12 and Example 4.7 the optimal filter is

$$\widehat{f}(\xi) = \frac{f_u(\xi)}{f_u(\xi) + p}, \quad \xi \in \mathbf{R}^d.$$

If the GSP w is white noise with power $p > 0$ differentiated to order $\alpha \in \mathbf{N}^d$ as in Example 4.8 then

$$\widehat{f}(\xi) = \frac{f_u(\xi)}{f_u(\xi) + p\xi^{2\alpha}}.$$

So if f_u has slower growth that $\xi^{2\alpha}$ then the optimal filter will attenuate high frequencies ("low-pass filter" [38]).

4.4 The Optimal Error for WSS Stochastic Processes

When the WSS GSPs u and w in Theorem 4.12 can be identified with stochastic processes we get an error that does not depend on a test function, under certain restrictions. As we will see the error is in fact constant with respect to $x \in \mathbf{R}^d$.

Let $u \in C(\mathbf{R}^d, L^2_0(\Omega))$ be a continuous WSS stochastic process with covariance function $k_u(x, y) = \mathbf{E}\left(u(x)\overline{u(y)}\right) = \kappa_u(x - y)$, $x, y \in \mathbf{R}^d$. By Bochner's theorem $\mu_u = (2\pi)^{\frac{d}{2}}\mathscr{F}\kappa_u$ is a non-negative bounded Radon measure. Thus

$$\kappa_u(x) = (2\pi)^{-d}\int_{\mathbf{R}^d} e^{i\langle x,\xi\rangle}\,\mathrm{d}\mu_u(\xi), \quad x \in \mathbf{R}^d, \tag{4.20}$$

and $\kappa_u \in (C \cap L^\infty)(\mathbf{R}^d)$.

Let $f \in \mathscr{F}C^\infty_c(\mathbf{R}^d)$ and define the convolution

$$u_o(x) = (2\pi)^{-\frac{d}{2}}\int_{\mathbf{R}^d} f(x - y)u(y)\,\mathrm{d}y, \quad x \in \mathbf{R}^d. \tag{4.21}$$

Then

$$\begin{aligned}\|u_o(x)\|^2_{L^2(\Omega)} &= (2\pi)^{-d}\iint_{\mathbf{R}^{2d}} f(x - y)\overline{f(x - z)}\,\kappa_u(y - z)\,\mathrm{d}y\,\mathrm{d}z\\ &= (2\pi)^{-d}(f * \kappa_u, f) = (2\pi)^{-d}\int_{\mathbf{R}^d}|\widehat{f}(\xi)|^2\mathrm{d}\mu_u\end{aligned} \tag{4.22}$$

for any $x \in \mathbf{R}^d$. Thus the convolution (4.21) extends to a well defined stochastic process $u_o \in L^\infty(\mathbf{R}^d, L^2_0(\Omega))$ when $f \in \mathscr{F}L^\infty(\mu_u)$, and $\|u_o(x)\|_{L^2(\Omega)}$ does not depend on $x \in \mathbf{R}^d$. The filtered stochastic process u_o is thus well defined for $f \in \mathscr{F}L^\infty(\mu_u)$. Note that $\mathscr{F}L^\infty(\mu_u) \subseteq \mathscr{F}L^2(\mu_u)$ since the measure μ_u is bounded. We have

$$\begin{aligned}&\|u_o(x + y) - u_o(x)\|^2_{L^2(\Omega)}\\ &= (2\pi)^{-d}\iint_{\mathbf{R}^{2d}} f(z)\overline{f(w)}\,\left(u(x + y - z) - u(x - z), u(x + y - w) - u(x - w)\right)_{L^2(\Omega)}\,\mathrm{d}z\,\mathrm{d}w\\ &= (2\pi)^{-d}\iint_{\mathbf{R}^{2d}} f(z)\overline{f(w)}\left(2\kappa_u(w - z) - \kappa_u(w - z + y) - \kappa_u(w - z - y)\right)\mathrm{d}z\,\mathrm{d}w\\ &= (2\pi)^{-d}\int_{\mathbf{R}^d}\left(2f * \kappa_u(w) - f * \kappa_u(w + y) - f * \kappa_u(w - y)\right)\overline{f(w)}\,\mathrm{d}w\\ &= (2\pi)^{-d}\left(2\widehat{f * \kappa_u} - M_y\widehat{f * \kappa_u} - M_{-y}\widehat{f * \kappa_u}, \widehat{f}\right)_{L^2}\\ &= (2\pi)^{-d}\int_{\mathbf{R}^d}|\widehat{f}(\xi)|^2\left(2 - e^{i\langle y,\xi\rangle} - e^{-i\langle y,\xi\rangle}\right)\mathrm{d}\mu_u\\ &= 2(2\pi)^{-d}\int_{\mathbf{R}^d}|\widehat{f}(\xi)|^2\,(1 - \cos(\langle y, \xi\rangle))\,\mathrm{d}\mu_u\end{aligned}$$

so it follows from dominated convergence that $u_o \in C(\mathbf{R}^d, L^2_0(\Omega))$ if $f \in \mathscr{F}L^\infty(\mu_u)$.

Let $v = u + w$ where u, w are uncorrelated continuous stochastic processes. We pick $f \in \mathscr{F}L^\infty(\mu_u + \mu_w)$ in (4.21) in order to minimize the mean square error $\mathbf{E}|u_o(x) - u(x)|^2$ where u_o is defined by (4.21) with u replaced by v. This is a particular case of Theorem 4.12, and hence $\widehat{f} \in L^\infty(\mu_u + \mu_w)$ is to be chosen as the

Radon–Nikodym derivative, that by definition satisfies

$$\mu_u = \widehat{f}\,(\mu_u + \mu_w). \tag{4.23}$$

Using (4.20), (4.21), (4.22) with μ_u replaced by $\mu_v = \mu_u + \mu_w$, (4.23), $\kappa_u^* = \kappa_u$ and $\widehat{f} \geqslant 0$ we obtain the minimal mean square error as

$$\begin{aligned}
\|u_o(x) - u(x)\|^2_{L^2(\Omega)} &= \|u_o(x)\|^2_{L^2(\Omega)} + \|u(x)\|^2_{L^2(\Omega)} - 2\,\mathrm{Re}\,(u_o(x), u(x))_{L^2(\Omega)} \\
&= (2\pi)^{-d} \int_{\mathbf{R}^d} |\widehat{f}(\xi)|^2 (\mathrm{d}\mu_u + \mathrm{d}\mu_w) + \kappa_u(0) - 2(2\pi)^{-\frac{d}{2}}\,\mathrm{Re}\,(f, \kappa_u)_{L^2} \\
&= (2\pi)^{-d} \int_{\mathbf{R}^d} \big(\widehat{f}(\xi) + 1 - 2\widehat{f}(\xi)\big)\,\mathrm{d}\mu_u \\
&= (2\pi)^{-d} \int_{\mathbf{R}^d} \big(1 - \widehat{f}(\xi)\big)\,\mathrm{d}\mu_u = (2\pi)^{-d} \int_{\mathbf{R}^d} \widehat{f}(\xi)\,\mathrm{d}\mu_w
\end{aligned}$$

independently of $x \in \mathbf{R}^d$, as claimed.

Remark 4.16 Writing

$$J = 2(2\pi)^d \|u_o(x) - u(x)\|^2_{L^2(\Omega)} = \int_{\mathbf{R}^d} \big(1 - \widehat{f}(\xi)\big)\,\mathrm{d}\mu_u + \int_{\mathbf{R}^d} \widehat{f}(\xi)\,\mathrm{d}\mu_w = J_1 + J_2$$

we split the total error as $J = J_1 + J_2$ where the bargain of the optimal filter $\widehat{f}$ becomes visible. In fact, recalling $0 \leqslant \widehat{f}(\xi) \leqslant 1$, $\widehat{f}$ should be close to 1 at the frequencies where the signal in μ_u is strong in order to minimize J_1. On the other hand $\widehat{f}$ should be close to zero at the frequencies where the noise in μ_w is strong in order to attenuate J_2. The compromise may give a small error if $\operatorname{supp}\mu_u \cap \operatorname{supp}\mu_w$ is small.

Remark 4.17 For a WSS generalized stochastic process the minimal error depends on the test function $\varphi \in \mathscr{F}L^2(\mu)$ according to (4.14). Also in this case we may split the minimal error as

$$2J(\varphi) = \int_{\mathbf{R}^d} \big(1 - \widehat{f}(\xi)\big)\,|\widehat{\varphi}(\xi)|^2\,\mathrm{d}\mu_u + \int_{\mathbf{R}^d} \widehat{f}(\xi)\,|\widehat{\varphi}(\xi)|^2\,\mathrm{d}\mu_w.$$

This admits a similar interpretation as for a stochastic process (Remark 4.16), with the modification of the test function frequency weight $|\widehat{\varphi}(\xi)|^2$ in the integrals.

5 Optimal Filtering of Non-stationary Generalized Stochastic Processes

Let $u, w \in \mathscr{L}(C_c^\infty(\mathbf{R}^d), L^2_0(\Omega))$ be uncorrelated generalized stochastic processes. In this section we do not assume that u and w are WSS. We return to the model (3.10) and the corresponding non-stationary operator equation (3.11)

$$\mathscr{K}_u = F(\mathscr{K}_u + \mathscr{K}_w) \tag{5.1}$$

in the space of continuous linear operators $C_c^\infty(\mathbf{R}^d) \to \mathscr{D}'(\mathbf{R}^d)$. Recall that $\mathscr{K}_u \geqslant 0$, $\mathscr{K}_w \geqslant 0$, and that F is assumed to be a continuous linear operator on $C_c^\infty(\mathbf{R}^d)$ which extends uniquely to a continuous linear operator on $\mathscr{D}'(\mathbf{R}^d)$.

The equation (5.1) when u or w or both are not WSS is more cumbersome to handle than the case when both are WSS. In the latter case we could in Section 4 for a given WSS $u \in \mathscr{L}(\mathscr{S}(\mathbf{R}^d), L_0^2(\Omega))$ choose a Hilbert subspace $\mathscr{F}L^2(\mu_u) \subseteq \mathscr{S}'(\mathbf{R}^d)$ where $\mu_u \in \mathscr{M}_+(\mathbf{R}^d)$ is the spectral measure corresponding to u, such that the covariance operator $\mathscr{K}_u$ extends to the identity operator on $\mathscr{F}L^2(\mu_u)$. Note that the space $\mathscr{F}L^2(\mu_u)$ is adapted to the specific $u \in \mathscr{L}(\mathscr{S}(\mathbf{R}^d), L_0^2(\Omega))$. In the non-stationary case, without further restrictions, there is no natural space on which a covariance operator is continuous.

The covariance operators $\mathscr{K}$ for WSS GSPs are translation invariant as $\mathscr{K}_u T_x = T_x \mathscr{K}_u$ for all $x \in \mathbf{R}^d$, that is they are convolution operators. Moreover the space $\mathscr{F}L^2(\mu_u)$ is also translation invariant (more precisely translation isometric).

5.1 Commuting Covariance Operators

In this subsection we start from the observation that the operators $\mathscr{K}_u$ and $\mathscr{K}_w$ commute when u and w are WSS GSPs, since $\mathscr{K}_u$ and $\mathscr{K}_w$ are convolution operators then.

We let $u, w \in \mathscr{L}(\mathscr{S}(\mathbf{R}^d), L_0^2(\Omega))$, assume they are uncorrelated as before, allow them to be non-stationary, but we keep the restriction that $\mathscr{K}_u$ and $\mathscr{K}_w$ commute. We also assume that $\mathscr{K}_u, \mathscr{K}_w \in \mathscr{L}(\mathscr{H})$, $\mathscr{K}_u \geqslant 0$ and $\mathscr{K}_w \geqslant 0$ where $\mathscr{H}$ is a Hilbert space of distributions defined on $\mathbf{R}^d$. This framework is realized e.g. with modulation spaces if $\mathscr{H} = M_{t,s}^2(\mathbf{R}^d)$ and the covariance operators are Weyl pseudodifferential operators with symbols in $M_{0,0,|s|,|t|}^{\infty,1}(\mathbf{R}^{2d})$ for any $t, s \in \mathbf{R}$ as we discussed in Section 2.4, cf. (2.18).

Then we may use the powerful methods that are based on the spectral theorem. The spectral theorem (cf. [3, Theorem IX §2.2.2] and [37, Theorem 9.14]) concerns a normal operator $N \in \mathscr{L}(\mathscr{H})$ which by definition satisfies $N^*N = NN^*$. For such an operator there exists a measure Π with domain $\mathscr{B}(\sigma(N))$ where $\sigma(N) \Subset \mathbf{C}$ denotes the spectrum of N, and values in the projections in $\mathscr{L}(\mathscr{H})$, such that N can be spectrally decomposed as

$$N = \int_{\sigma(N)} z \mathrm{d}\Pi(z).$$

By [3, Theorem IX §2.2.3] or [37, Theorem 9.8] there exists a bounded functional calculus defined by

$$f(N) = \int_{\sigma(N)} f(z) \mathrm{d}\Pi(z) \in \mathscr{L}(\mathscr{H})$$

for $f \in L^\infty(\sigma(N), \Pi)$ (cf. [37, Proposition 9.11]) which is a *-homomorphism and an isomorphism. If N is selfadjoint then $\operatorname{supp} \Pi \subseteq \mathbf{R}$ and if $N \geqslant 0$ then $\operatorname{supp} \Pi \subseteq \mathbf{R}_+$ (cf. [37, Corollary 9.18]).

We use von Neumann's result [37, Theorem 9.24] which depends on the spectral theorem for normal operators on a Hilbert space. This theorem concerns in fact the more general assumption of two commuting self-adjoint operators. Adapted to our setup of commuting non-negative operators, the theorem and its proof says the following. There exists a measure Π with domain $\mathscr{B}(\mathbf{C})$, compactly supported in the first quadrant, i.e. $\operatorname{supp}\Pi \Subset \{x+iy \in \mathbf{C} : x \geqslant 0,\ y \geqslant 0\} := \mathbf{C}_+$, which is projection-valued in $\mathscr{L}(\mathscr{H})$, and there exists a normal operator $N \in \mathscr{L}(\mathscr{H})$ with spectral decomposition

$$N = \int_{\mathbf{C}_+} z \mathrm{d}\Pi(z).$$

Moreover there exist two functions $f_u, f_w \in C_c(\mathbf{C}_+)$ such that $f_u, f_w \geqslant 0$, and $\mathscr{K}_u$ and $\mathscr{K}_w$ can be simultaneously spectrally decomposed with respect to Π as

$$\mathscr{K}_u = \int_{\mathbf{C}_+} f_u(z) \mathrm{d}\Pi(z), \tag{5.2}$$

$$\mathscr{K}_w = \int_{\mathbf{C}_+} f_w(z) \mathrm{d}\Pi(z). \tag{5.3}$$

If we set

$$f(z) = \frac{f_u(z)}{f_u(z) + f_w(z)} \chi_{\operatorname{supp} f_u} \tag{5.4}$$

then $f : \mathbf{C}_+ \to \mathbf{R}_+$ is a well defined bounded measureable function. The functional calculus of the operator N hence yields that

$$F = \int_{\mathbf{C}_+} f(z) \mathrm{d}\Pi(z) \in \mathscr{L}(\mathscr{H}) \tag{5.5}$$

is an operator that solves (5.1). The solution F is unique in the bounded functional calculus of N. The fact that $f \geqslant 0$ implies by [37, Theorem 9.8] that F is non-negative.

We have shown the following result.

Theorem 5.1 *Let $u, w \in \mathscr{L}(\mathscr{S}(\mathbf{R}^d), L^2_0(\Omega))$ be uncorrelated. Suppose that the covariance operators satisfy $\mathscr{K}_u, \mathscr{K}_w \in \mathscr{L}(\mathscr{H})$ where $\mathscr{H}$ is a Hilbert space of distributions defined on $\mathbf{R}^d$, $\mathscr{K}_u \geqslant 0$, $\mathscr{K}_w \geqslant 0$ and $\mathscr{K}_u \mathscr{K}_w = \mathscr{K}_w \mathscr{K}_u$.*

Then there exists a Borel measure Π, compactly supported on $\mathbf{C}_+$, with values in the projections in $\mathscr{L}(\mathscr{H})$, such that $\mathscr{K}_u$ and $\mathscr{K}_w$ can be expressed as (5.2) and (5.3) respectively, with non-negative $f_u, f_w \in C_c(\mathbf{C}_+)$, and the solution to (5.1) is given by (5.4) and (5.5). The solution satisfies $F \geqslant 0$ and is unique in the bounded functional calculus of Π.

5.2 Pseudodifferential Operators

In this subsection we relax both the conditions that u, w are WSS and that $\mathscr{K}_u$ and $\mathscr{K}_w$ commute. The covariance operators are allowed to be pseudodifferential operators

rather than convolution operators as in Section 4. We use the functional framework of Section 2.4, that is operators with symbols in modulation spaces, acting on modulation spaces. In an effort to extend the analysis in Section 4 we try to choose a framework that contains the WSS case as far as possible.

First we deduce an inclusion of modulation spaces into the spaces on which WSS covariance operators act in Section 4. This gives a relation between spaces used in Section 4 and spaces used in this section.

Proposition 5.2 *Let $1 \leqslant p, q < \infty$ and let $\mu \in \mathscr{M}_+(\mathbf{R}^d)$ with $s \geqslant 0$ chosen such that (2.3) is satisfied. If $\varepsilon > 0$ and*

$$\omega(x,\xi) = \langle x \rangle^{\frac{1}{p'}(d+\varepsilon)} \langle \xi \rangle^{\frac{s}{2}+\frac{1}{q'}(d+\varepsilon)}$$

then

$$M^{p,q}_{\omega}(\mathbf{R}^d) \subseteq \mathscr{F}L^2(\mu). \tag{5.6}$$

Proof Let $f \in \mathscr{S}(\mathbf{R}^d)$ and let $\varphi \in \mathscr{S}(\mathbf{R}^d)$ satisfy $\|\varphi\|_{L^2} = 1$. We use (2.11) which gives

$$\widehat{f} = (2\pi)^{-\frac{d}{2}} \iint_{\mathbf{R}^{2d}} V_\varphi f(x,\xi) T_\xi M_{-x} \widehat{\varphi} \, \mathrm{d}x \, \mathrm{d}\xi.$$

Using (2.1) and Hölder's inequality we obtain for all $\eta \in \mathbf{R}^d$

$$\begin{aligned}
\langle \eta \rangle^{\frac{s}{2}} |\widehat{f}(\eta)| &\lesssim \iint_{\mathbf{R}^{2d}} |V_\varphi f(x,\xi)| \, |\widehat{\varphi}(\eta-\xi)| \langle \eta \rangle^{\frac{s}{2}} \, \mathrm{d}x \, \mathrm{d}\xi \lesssim \iint_{\mathbf{R}^{2d}} |V_\varphi f(x,\xi)| \langle \xi \rangle^{\frac{s}{2}} \, \mathrm{d}x \, \mathrm{d}\xi \\
&= \iint_{\mathbf{R}^{2d}} |V_\varphi f(x,\xi)| \, \omega(x,\xi) \, \langle x \rangle^{-\frac{1}{p'}(d+\varepsilon)} \langle \xi \rangle^{-\frac{1}{q'}(d+\varepsilon)} \mathrm{d}x \, \mathrm{d}\xi \\
&\lesssim \int_{\mathbf{R}^d} \|V_\varphi f(\cdot,\xi) \, \omega(\cdot,\xi)\|_{L^p(\mathbf{R}^d)} \langle \xi \rangle^{-\frac{1}{q'}(d+\varepsilon)} \mathrm{d}\xi \\
&\lesssim \|(V_\varphi f)\omega\|_{L^{p,q}(\mathbf{R}^{2d})} = \|f\|_{M^{p,q}_{\omega}(\mathbf{R}^d)}.
\end{aligned}$$

From (2.3) it hence follows that

$$\|f\|_{\mathscr{F}L^2(\mu)} = \left(\int_{\mathbf{R}^d} |\widehat{f}(\eta)|^2 \mathrm{d}\mu(\eta) \right)^{\frac{1}{2}} \lesssim \sup_{\eta \in \mathbf{R}^d} \langle \eta \rangle^{\frac{s}{2}} |\widehat{f}(\eta)| \lesssim \|f\|_{M^{p,q}_{\omega}(\mathbf{R}^d)}. \tag{5.7}$$

The inclusion (5.6) now follows from the density of $\mathscr{S}(\mathbf{R}^d)$ in $M^{p,q}_{\omega}(\mathbf{R}^d)$ [15, Proposition 11.3.4]. □

Remark 5.3 The space $\mathscr{F}L^2(\mu)$ is translation invariant, and in fact also translation isometric. The modulation spaces $M^{p,q}_{\omega}$ are translation invariant but in general not translation isometric [15].

Remark 5.4 From (2.10) it follows

$$\mathscr{S}'(\mathbf{R}^d) = \bigcup_{t\in\mathbf{R}} M^{p,q}_{\omega_t}(\mathbf{R}^d) \tag{5.8}$$

for any $p, q \in [1,\infty]$ with $\omega_t(X) = \langle X\rangle^t$ and $X \in \mathbf{R}^{2d}$. From Section 2.2 we know that the inclusions

$$\mathscr{S}(\mathbf{R}^d) \subseteq \mathscr{F}L^2(\mu) \subseteq \mathscr{S}'(\mathbf{R}^d) \tag{5.9}$$

are continuous and dense but the left inclusion is in general not injective if $\mu \in \mathscr{M}_+(\mathbf{R}^d)$. Combining (5.8) and (5.9) we obtain the injective inclusion

$$\mathscr{F}L^2(\mu) \subseteq \bigcup_{t\in\mathbf{R}} M^{p,q}_{\omega_t}(\mathbf{R}^d)$$

which may be seen as an upper inclusion opposite to (5.6).

However we cannot get continuous inclusions in individual modulation spaces of the form $\mathscr{F}L^2(\mu) \subseteq M^{p,q}_{\omega}(\mathbf{R}^d)$ in general, that is an estimate of the form

$$\|f\|_{M^{p,q}_{\omega}} \lesssim \|f\|_{\mathscr{F}L^2(\mu)}.$$

Indeed suppose $\operatorname{supp}\mu \Subset \mathbf{R}^d$, take $f \in \mathscr{S}(\mathbf{R}^d)\setminus\{0\}$ such that $\widehat{f} = 0$ in $L^2(\mu)$. Then $f \neq 0$ in $M^{p,q}_{\omega}(\mathbf{R}^d)$ for any $p, q \in [1,\infty]$ and any $\omega \in \mathscr{P}(\mathbf{R}^d)$.

The covariance operator $\mathscr{K}_u$ of a WSS GSP u can be seen as a convolution operator as in (4.3) and (4.4), where $\kappa_u = (2\pi)^{-\frac{d}{2}}\mathscr{F}^{-1}\mu_u \in \mathscr{S}'(\mathbf{R}^d)$ and $\mu_u \in \mathscr{M}_+(\mathbf{R}^d)$. From (2.4) we may identify the Weyl symbol of $\mathscr{K}_u = a^w_u(x,D)$ as $a_u = 1\otimes\mu_u \in \mathscr{M}_+(\mathbf{R}^{2d})$. Since we want to include the WSS case we need to work with pseudodifferential operators with symbols that are not smooth, as opposed to the generic frameworks for pseudodifferential operators when they are applied to PDEs [10, 20, 33]. In fact we need spaces of non-regular symbols, including elements of the form $1\otimes\mu \in \mathscr{M}_+(\mathbf{R}^{2d})$ with $\mu \in \mathscr{M}_+(\mathbf{R}^d)$. The next result shows that $1\otimes\mu$ belongs to a modulation space when $\mu \in \mathscr{M}_+(\mathbf{R}^d)$.

Proposition 5.5 *Suppose $\mu \in \mathscr{M}_+(\mathbf{R}^d)$ and let $s \geqslant 0$ be chosen so that (2.3) is fulfilled. Then for any $n \in \mathbf{N}$ and any $\varepsilon > 0$ we have $1\otimes\mu \in M^{\infty,1}_{\omega}(\mathbf{R}^{2d})$ with*

$$\omega(x_1,x_2,\xi_1,\xi_2) = \langle x_2\rangle^{-s}\langle\xi_1\rangle^n\langle\xi_2\rangle^{-d-\varepsilon}, \quad x_1,x_2,\xi_1,\xi_2 \in \mathbf{R}^d. \tag{5.10}$$

Proof Let $\varphi \in \mathscr{S}(\mathbf{R}^d)\setminus 0$. We have

$$V_{\varphi\otimes\varphi}(1\otimes\mu)(x_1,x_2,\xi_1,\xi_2) = V_\varphi 1(x_1,\xi_1)V_\varphi\mu(x_2,\xi_2)$$

and $V_\varphi 1(x_1,\xi_1) = e^{-i\langle x_1,\xi_1\rangle}\overline{\widehat{\varphi}(-\xi_1)}$. Using (2.1) and (2.3) this gives for any $n \in \mathbf{N}$ and any $\varepsilon > 0$

$$\begin{aligned}\big|V_{\varphi\otimes\varphi}(1\otimes\mu)(x_1,x_2,\xi_1,\xi_2)\big| &= (2\pi)^{-\frac{d}{2}}|\widehat{\varphi}(-\xi_1)|\left|\int_{\mathbf{R}^d} e^{-i\langle\xi_2,y_2\rangle}\overline{\varphi(y_2-x_2)}\mathrm{d}\mu(y_2)\right| \\ &\lesssim \langle\xi_1\rangle^{-n-d-\varepsilon}\int_{\mathbf{R}^d}\langle y_2-x_2\rangle^{-s}\mathrm{d}\mu(y_2) \\ &\lesssim \langle\xi_1\rangle^{-n-d-\varepsilon}\langle x_2\rangle^{s}.\end{aligned}$$

It follows that $\big|\big(V_{\varphi\otimes\varphi}(1\otimes\mu)\omega\big)(x_1,x_2,\xi_1,\xi_2)\big| \lesssim \langle\xi_1\rangle^{-d-\varepsilon}\langle\xi_2\rangle^{-d-\varepsilon}$ which implies

$$\|1\otimes\mu\|_{M^{\infty,1}_\omega(\mathbf{R}^{2d})} = \int_{\mathbf{R}^{2d}}\sup_{X\in\mathbf{R}^{2d}}\Big(\big|\big(V_{\varphi\otimes\varphi}(1\otimes\mu)\omega\big)(X,\Xi)\big|\Big)\mathrm{d}\Xi < \infty.$$

□

With the notation of Section 2.4, Proposition 5.5 says that $1\otimes\mu \in M^{\infty,1}_{0,-s,n,-d-\varepsilon}(\mathbf{R}^{2d})$ for any $n \in \mathbf{N}$, any $\varepsilon > 0$, and some $s \geqslant 0$. From (2.21) and the preceding discussion we get the following consequence.

Corollary 5.6 *Suppose $\mu \in \mathscr{M}_+(\mathbf{R}^d)$, let $s \geqslant 0$ be chosen so that (2.3) is fulfilled, set $a = 1\otimes\mu$, let $p,q \in [1,\infty]$, and let $t_j, s_j \in \mathbf{R}$ for $j = 1,2$ satisfy for some $\varepsilon > 0$*

$$t_1 \geqslant d+\varepsilon, \quad t_2 \leqslant -d-\varepsilon, \quad \text{and} \quad s_2 \leqslant s_1 - s. \tag{5.11}$$

Then $a^w(x,D) : M^{p,q}_{t_1,s_1}(\mathbf{R}^d) \to M^{p,q}_{t_2,s_2}(\mathbf{R}^d)$ is continuous.

Remark 5.7 Since $t_2 < t_1$ and $s_2 \leqslant s_1$ in the assumptions (5.11) we have $M^{p,q}_{t_1,s_1}(\mathbf{R}^d) \subsetneq M^{p,q}_{t_2,s_2}(\mathbf{R}^d)$. Corollary 5.6 does therefore not give a common space for domain and range of the operator.

Proposition 5.5 gives evidence that $M^{\infty,1}_\omega(\mathbf{R}^{2d})$ for certain weights $\omega \in \mathscr{P}(\mathbf{R}^{4d})$ may be an interesting class for covariance operators of non-stationary GSPs that contains the covariance operators for WSS GSPs. Corollary 5.6 says that the corresponding WSS covariance operators act continuously between certain modulation spaces, and Proposition 5.2 relates modulation spaces to the spaces $\mathscr{F}L^2(\mu)$ used in Section 4.

Nevertheless we need to introduce some restrictions in order to proceed. In fact we want to use the spectral invariance results for modulation spaces discussed in Section 2.4. The space $M^{\infty,1}_\omega(\mathbf{R}^{2d})$ with the weight (5.10) is not quite good enough as a symbol class. We need instead $M^{\infty,1}_{1\otimes\omega}(\mathbf{R}^{2d})$ with $\omega(X) = \langle X\rangle^r$ for $r \geqslant 0$.

For the Weyl symbols of covariance operators of WSS GSPs this leads to the following criterion.

Lemma 5.8 *Let $f \in \mathscr{S}'(\mathbf{R}^d)$, $r \geqslant 0$, $\omega_1(\xi) = \langle\xi\rangle^r$ for $\xi \in \mathbf{R}^d$, and $\omega_2(X) = \langle X\rangle^r$ for $X \in \mathbf{R}^{2d}$. Then $1\otimes f \in M^{\infty,1}_{1\otimes\omega_2}(\mathbf{R}^{2d})$ if and only if $f \in M^{\infty,1}_{1\otimes\omega_1}(\mathbf{R}^d)$.*

Proof Let $\varphi \in \mathscr{S}(\mathbf{R}^d) \setminus 0$. As in the proof of Proposition 5.5 we have

$$|V_{\varphi\otimes\varphi}(1\otimes f)(x_1,x_2,\xi_1,\xi_2)| = |\widehat{\varphi}(-\xi_1)|\,|V_\varphi f(x_2,\xi_2)|.$$

This gives, using (2.1),

$$\begin{aligned}|\widehat{\varphi}(-\xi_1)|\,|V_\varphi f(x_2,\xi_2)|\langle\xi_2\rangle^r &\leqslant |V_{\varphi\otimes\varphi}(1\otimes f)(x_1,x_2,\xi_1,\xi_2)|\langle(\xi_1,\xi_2)\rangle^r\\ &\lesssim |\widehat{\varphi}(-\xi_1)|\langle\xi_1\rangle^r|V_\varphi f(x_2,\xi_2)|\langle\xi_2\rangle^r\end{aligned}$$

and it follows that

$$\begin{aligned}\|\widehat{\varphi}\|_{L^1}\|f\|_{M^{\infty,1}_{1\otimes\omega_1}(\mathbf{R}^d)} &= \int_{\mathbf{R}^{2d}} |\widehat{\varphi}(-\xi_1)|\Big(\sup_{x_2\in\mathbf{R}^d}|V_\varphi f(x_2,\xi_2)|\Big)\langle\xi_2\rangle^r\mathrm{d}\xi_1\mathrm{d}\xi_2 \leqslant \|1\otimes f\|_{M^{\infty,1}_{1\otimes\omega_2}(\mathbf{R}^{2d})}\\ &\lesssim \int_{\mathbf{R}^d}|\widehat{\varphi}(-\xi_1)|\langle\xi_1\rangle^r\mathrm{d}\xi_1\int_{\mathbf{R}^d}\Big(\sup_{x_2\in\mathbf{R}^d}|V_\varphi f(x_2,\xi_2)|\Big)\langle\xi_2\rangle^r\mathrm{d}\xi_2\\ &\lesssim \|f\|_{M^{\infty,1}_{1\otimes\omega_1}(\mathbf{R}^d)}.\end{aligned}$$

□

The requirement $\mu = f \in M^{\infty,1}_{1\otimes\langle\cdot\rangle^r}(\mathbf{R}^d)$ for some $r \geqslant 0$ on the spectral measure $\mu \in \mathscr{M}_+(\mathbf{R}^d)$ for a WSS GSP is hence equivalent to its covariance operator having a Weyl symbol in $M^{\infty,1}_{1\otimes\omega}(\mathbf{R}^{2d})$ with $\omega(X) = \langle X\rangle^r$.

The condition $\mu \in M^{\infty,1}_{1\otimes\langle\cdot\rangle^r}(\mathbf{R}^d)$ is satisfied for white noise for any $r \geqslant 0$. Indeed then we have $\mu = p\,\mathrm{d}\xi$ for a constant $p > 0$, cf. Example 4.7, and it follows from the proof of Lemma 5.8 that $1 \in M^{\infty,1}_{1\otimes\langle\cdot\rangle^r}(\mathbf{R}^d)$ for any $r \geqslant 0$. But if $u \in \mathscr{L}(\mathscr{S}(\mathbf{R}^d), L^2_0(\Omega))$ is white noise of power $p = 1$ then $D^\alpha u$ has spectral measure $\mu_{D^\alpha u} = \xi^{2\alpha}\mathrm{d}\xi$ if $\alpha \in \mathbf{N}^d$, cf. Example 4.8. The Weyl symbol of the covariance operator $\mathscr{K}_{D^\alpha u}$ is therefore $1\otimes\xi^{2\alpha}$. This Weyl symbol does not belong to $M^{\infty,1}_{1\otimes\omega}(\mathbf{R}^{2d})$ with $\omega(X) = \langle X\rangle^r$ for $X \in \mathbf{R}^{2d}$, for any $r \geqslant 0$ if $\alpha \neq 0$. In fact $M^{\infty,1}_{1\otimes\omega}(\mathbf{R}^{2d}) \subseteq M^{\infty,1}(\mathbf{R}^{2d}) \subseteq (C\cap L^\infty)(\mathbf{R}^{2d})$ [34]. Thus white noise is included but its derivatives are excluded when we work with the symbol classes $M^{\infty,1}_{1\otimes\langle\cdot\rangle^r}(\mathbf{R}^{2d})$ for some $r \geqslant 0$. Furthermore $\mu \in M^{\infty,1}_{1\otimes\langle\cdot\rangle^r}(\mathbf{R}^d) \subseteq (C\cap L^\infty)(\mathbf{R}^d)$ implies that μ is absolutely continuous with respect to Lebesgue measure.

Remark 5.9 Let $u \in \mathscr{L}(\mathscr{S}(\mathbf{R}^d), L^2_0(\Omega))$ and assume that its covariance operator satisfies $\mathscr{K}_u = a_u^w(x,D)$ with $a_u \in M^{\infty,1}_{1\otimes\omega}(\mathbf{R}^{2d})$ with $\omega(X) = \langle X\rangle^r$ for some $r \geqslant 0$. By (2.18) and the surrounding discussion it follows that $\mathscr{K}_u \in \mathscr{L}(M^1(\mathbf{R}^d))$. Using $(M^1)' = M^\infty$ and (2.15) this gives for $\psi,\varphi \in M^1(\mathbf{R}^d)$

$$|(\mathscr{K}_u\psi,\varphi)| \leqslant \|\mathscr{K}_u\psi\|_{M^1(\mathbf{R}^d)}\|\varphi\|_{M^\infty(\mathbf{R}^d)} \lesssim \|\psi\|_{M^1(\mathbf{R}^d)}\|\varphi\|_{M^1(\mathbf{R}^d)}$$

and it follows that $u \in \mathscr{L}(M^1(\mathbf{R}^d), L^2_0(\Omega))$, cf. Remark 3.2. The assumption $a_u \in M^{\infty,1}_{1\otimes\omega}(\mathbf{R}^{2d})$ hence implies that the corresponding class of GSPs is a subspace of the class studied in [9, 21, 24, 39] with test function space Feichtinger's algebra.

As preparation for the next result on non-stationary filters we will discuss operators $\mathscr{K} \in \mathscr{L}(L^2(\mathbf{R}^d))$ that satisfy an assumption of the form

$$(\mathscr{K} f, f) \geqslant \varepsilon \|f\|^2_{L^2(\mathbf{R}^d)}, \quad f \in L^2(\mathbf{R}^d), \tag{5.12}$$

for some $\varepsilon > 0$. This assumption is a strengthening of $\mathscr{K} \geqslant 0$ on $L^2(\mathbf{R}^d)$.

Lemma 5.10 *If $\mathscr{K} \in \mathscr{L}(L^2(\mathbf{R}^d))$ satisfies $\mathscr{K} \geqslant 0$ then $\mathscr{K}$ is invertible on $L^2(\mathbf{R}^d)$ if and only if (5.12) is satisfied for some $\varepsilon > 0$.*

Proof Suppose (5.12) is satisfied for some $\varepsilon > 0$. Since $(\mathscr{K} f, f) = \|\mathscr{K}^{\frac{1}{2}} f\|^2_{L^2}$ it follows that $\ker \mathscr{K}^{\frac{1}{2}} = \{0\}$, $\operatorname{ran} \mathscr{K}^{\frac{1}{2}}$ is closed in L^2, and hence $\operatorname{ran} \mathscr{K}^{\frac{1}{2}} = \left(\ker \mathscr{K}^{\frac{1}{2}}\right)^{\perp} = L^2(\mathbf{R}^d)$. By the open mapping theorem $\mathscr{K}^{\frac{1}{2}}$ is invertible on L^2 and hence $\mathscr{K}^{-1} = \left(\mathscr{K}^{-\frac{1}{2}}\right)^2 \in \mathscr{L}(L^2)$.

Suppose on the other hand that (5.12) is not satisfied for any $\varepsilon > 0$. Then there exists a sequence $\{f_n\}_{n \geqslant 1} \subseteq L^2(\mathbf{R}^d)$ such that $\|f_n\|_{L^2} = 1$ for all $n \geqslant 1$ and $\lim_{n\to\infty}(\mathscr{K} f_n, f_n) = 0$. It follows that

$$\lim_{n\to\infty} \|\mathscr{K}^{\frac{1}{2}} f_n\|_{L^2} = 0. \tag{5.13}$$

If we assume that $\mathscr{K}$ is invertible on L^2 then $\mathscr{K}^{-1} \geqslant 0$ and hence $\mathscr{K}^{-\frac{1}{2}} \in \mathscr{L}(L^2)$ is a well defined operator. Combined with (5.13) this gives $f_n = \mathscr{K}^{-\frac{1}{2}} \mathscr{K}^{\frac{1}{2}} f_n \to 0$ in L^2 which contradicts the assumption $\|f_n\|_{L^2} = 1$ for all $n \geqslant 1$. Hence $\mathscr{K}$ cannot be invertible on L^2. □

The following theorem is the main result on the operator equation (5.1) for non-stationary GSPs with non-commuting covariance operators. It is formulated in terms of Weyl symbols of covariance operators, using the relation (2.8) between Schwartz kernels and Weyl symbols for operators (cf. Remark 2.6). The result has a short proof since it is a consequence of results in [39], with clarified assumptions using (5.12) and Lemma 5.10.

For Gabor expansions of $a \in M^{\infty,1}_{1\otimes\omega}(\mathbf{R}^{2d})$ we use the notation

$$a = \sum_{\mathbf{\Lambda}\in\Theta} g_a(\mathbf{\Lambda}) \Pi(\mathbf{\Lambda}) \Phi, \quad g_a(\mathbf{\Lambda}) = (a, \Pi(\mathbf{\Lambda})\widetilde{\Phi}) \tag{5.14}$$

where $\{g_a(\mathbf{\Lambda})\}_{\mathbf{\Lambda}\in\Theta}$ are the Gabor coefficients for a, Φ is the Gaussian (2.22), $\widetilde{\Phi} \in \mathscr{S}(\mathbf{R}^{2d})$ is its canonical dual window, and $\Theta = \{(an, bk)\}_{n,k\in\mathbf{Z}^{2d}} \subseteq \mathbf{R}^{4d}$ is a lattice determined by $a, b > 0$ that satisfy $ab < \pi$, cf. (2.25).

Theorem 5.11 *Suppose $u, w \in \mathscr{L}(M^1(\mathbf{R}^d), L^2_0(\Omega))$ are uncorrelated tempered* GSP*s with covariance operators $\mathscr{K}_u, \mathscr{K}_w : M^1(\mathbf{R}^d) \to M^\infty(\mathbf{R}^d)$. Suppose the Weyl symbols of $\mathscr{K}_u$ and $\mathscr{K}_w$ satisfy $a_u, a_w \in M^{\infty,1}_{1\otimes\omega}(\mathbf{R}^{2d})$ with $\omega(X) = \langle X\rangle^r$ for some $r \geqslant 0$,*

and suppose

$$((\mathscr{K}_u + \mathscr{K}_w)f, f) \geqslant \varepsilon \|f\|^2_{L^2(\mathbf{R}^d)}, \quad f \in L^2(\mathbf{R}^d), \tag{5.15}$$

holds for some $\varepsilon > 0$. *Then the following holds:*

(i) *The operator equation (5.1) is solved by*

$$F = \mathscr{K}_u(\mathscr{K}_u + \mathscr{K}_w)^{-1} = a^w(x, D) \tag{5.16}$$

where $a \in M^{\infty,1}_{1\otimes\omega}(\mathbf{R}^{2d})$, *and* $a^w(x, D) : M^{p,q}_{t,s}(\mathbf{R}^d) \to M^{p,q}_{t,s}(\mathbf{R}^d)$ *is continuous for all* $1 \leqslant p, q \leqslant \infty$ *and all* $t, s \in \mathbf{R}$ *such that* $\max(|t|, |s|) \leqslant \frac{1}{2}r$.

(ii) *The Gabor coefficients* $\{g_a(\mathbf{\Lambda})\}_{\mathbf{\Lambda}\in\Theta}$ *of* a *can be expressed in those of* a_u *as*

$$g_a = M(g_b) \cdot g_{a_u} \tag{5.17}$$

where $M(g_b)$ *is an infinite matrix defined on* $\Theta \times \Theta$, *depending linearly on the Gabor coefficients* $\{g_b(\mathbf{\Lambda})\}_{\mathbf{\Lambda}\in\Theta}$ *of* $b \in M^{\infty,1}_{1\otimes\omega}(\mathbf{R}^{2d})$ *where* $b^w(x, D) = (\mathscr{K}_u + \mathscr{K}_w)^{-1}$.

(iii) *If* $r > 2d$ *then the matrix* $M(g_b)$ *has off-diagonal decay*

$$|M(g_b)(\mathbf{\Lambda}, \mathbf{\Omega})| \lesssim \langle \mathbf{\Lambda} - \mathbf{\Omega}\rangle^{-t}, \quad \mathbf{\Lambda}, \mathbf{\Omega} \in \Theta, \tag{5.18}$$

for any $t < r/2 - d$.

Proof The assumptions, (2.18) and the surrounding discussion imply that

$$\mathscr{K}_u, \mathscr{K}_w : M^{p,q}_{t,s}(\mathbf{R}^d) \to M^{p,q}_{t,s}(\mathbf{R}^d)$$

are continuous for all $p, q \in [1, \infty]$, and all $t, s \in \mathbf{R}$ such that $\max(|t|, |s|) \leqslant \frac{1}{2}r$. In particular $\mathscr{K}_u$ and $\mathscr{K}_w$ are continuous on $L^2(\mathbf{R}^d) = M^2(\mathbf{R}^d)$, and hence $\mathscr{K}_u + \mathscr{K}_w$ is continuous on L^2. By the assumption (5.15) and Lemma 5.10 the operator $\mathscr{K}_u + \mathscr{K}_w$ is invertible on L^2.

It follows from Gröchenig's spectral invariance theorem [17, Theorem 4.6] that $(\mathscr{K}_u+\mathscr{K}_w)^{-1} = b^w(x, D)$ where $b \in M^{\infty,1}_{1\otimes\omega}(\mathbf{R}^{2d})$. From the algebraic result [17, Theorem 4.3] we get the conclusion that $F = \mathscr{K}_u(\mathscr{K}_u + \mathscr{K}_w)^{-1} = a^w_u(x, D)b^w(x, D) = a^w(x, D)$ with $a \in M^{\infty,1}_{1\otimes\omega}(\mathbf{R}^{2d})$. It again follows from (2.18) and the surrounding discussion that $a^w(x, D) : M^{p,q}_{t,s}(\mathbf{R}^d) \to M^{p,q}_{t,s}(\mathbf{R}^d)$ is continuous for all $1 \leqslant p, q \leqslant \infty$ and all $t, s \in \mathbf{R}$ such that $\max(|t|, |s|) \leqslant \frac{1}{2}r$. We have shown (5.16) and statement (i).

Statement (ii) follows from [39, Section 4] (cf. also [2]). Indeed let

$$\{g_a(\mathbf{\Lambda})\}_{\mathbf{\Lambda}\in\Theta}, \quad \{g_b(\mathbf{\Lambda})\}_{\mathbf{\Lambda}\in\Theta}, \quad \{g_{a_u}(\mathbf{\Lambda})\}_{\mathbf{\Lambda}\in\Theta}$$

denote the Gabor coefficients defined as in (5.14) for for a, b, a_u, respectively. By [39, Proposition 7] these Gabor coefficients are related as in (5.17), that is

$$g_a(\mathbf{\Lambda}) = \sum_{\mathbf{\Omega}\in\Theta} M(g_b)(\mathbf{\Lambda}, \mathbf{\Omega})\, g_{a_u}(\mathbf{\Omega}), \quad \mathbf{\Lambda} \in \Theta,$$

where

$$M(g_b)(\mathbf{\Lambda}, \mathbf{\Omega}) = \sum_{\mathbf{\Gamma}\in\Theta} \mathcal{M}(\mathbf{\Omega}, \mathbf{\Gamma}, \mathbf{\Lambda}) g_b(\mathbf{\Gamma}), \quad \mathbf{\Lambda}, \mathbf{\Omega} \in \Theta,$$

is a $\mathbf{C}$-valued infinite matrix indexed by $\Theta \times \Theta \subseteq \mathbf{R}^{8d}$ depending linearly on the Gabor coefficients $\{g_b(\mathbf{\Gamma})\}_{\mathbf{\Gamma}\in\Theta}$ of b, and defined in terms of

$$\begin{aligned}
&\mathcal{M}(\mathbf{\Omega}, \mathbf{\Gamma}, \mathbf{\Lambda}) \\
&= (2\pi)^d \pi^{\frac{d}{2}} \exp\left(i\left(\sigma(\Omega + \Omega' + \Gamma - \Gamma', \Lambda') + \sigma(\Omega' + \Gamma', \Omega + \Gamma)\right)\right) \\
&\quad \times \exp\left(-\frac{1}{4}\left|\Omega - \Omega' - \Gamma - \Gamma'\right|^2\right) \times \mathcal{V}_{\widetilde{\Phi}}\Phi \\
&\left(\Lambda - \frac{\Omega + \Omega' + \Gamma - \Gamma'}{2}, \Lambda' - \frac{\Omega + \Omega' - \Gamma + \Gamma'}{2}\right), \\
&\mathbf{\Omega}, \mathbf{\Gamma}, \mathbf{\Lambda} \in \Theta,
\end{aligned}$$

with notation as in (2.24), and where

$$\mathcal{V}_{\widetilde{\Phi}}\Phi(X, Y) = (2\pi)^{-d}(\Phi, \Pi(X, Y)\widetilde{\Phi}), \quad X, Y \in \mathbf{R}^{2d},$$

denotes a symplectic version of the STFT (2.9), cf. (2.23), with $\widetilde{\Phi}$ the canonical dual window of Φ.

It remains to show statement (iii). The assumption $r > 2d$ admits the use of [39, Corollary 11] which says that the matrix $M(g_b)$ has off-diagonal decay as in (5.18) for any $t < r/2 - d$. □

The conclusion (5.17) with the matrix $M(g_b)$ having off-diagonal decay as in (5.18) may be regarded as conceptually reminiscent of the conclusions in Theorem 4.12, and in (5.4) and (5.5) respectively. In fact (5.17) can be seen almost as a pointwise multiplication when the matrix has rapid off-diagonal decay, of the Gabor coefficients of the Weyl symbol of $\mathscr{K}_u$ with a matrix that depends linearly on the Gabor coefficients of the Weyl symbol of $b^w(x, D) = (\mathscr{K}_u + \mathscr{K}_w)^{-1}$.

An engineering point of view for optimal filters for non-stationary signals is treated in [19].

Remark 5.12 The assumptions of Theorem 5.11 and Parseval's theorem imply that

$$((\mathscr{K}_u + \mathscr{K}_w)f, f) \asymp \|f\|_{L^2}^2 = \|\widehat{f}\|_{L^2}^2, \quad f \in L^2(\mathbf{R}^d). \tag{5.19}$$

If u and w are WSS then by (4.1) we have

$$((\mathscr{K}_u + \mathscr{K}_w)f, f) = \|\widehat{f}\|^2_{L^2(\mu_u+\mu_w)}, \quad f \in L^2(\mathbf{R}^d). \tag{5.20}$$

It follows from (5.19) and (5.20) that $L^2(\mu_u + \mu_w) = L^2(\mathbf{R}^d)$, and by the Radon–Nikodym theorem we have $\mu_u + \mu_w = f\mathrm{d}x$ where $f \in L^\infty(\mathbf{R}^d)$ and $f^{-1} \in L^\infty(\mathbf{R}^d)$. The measure $\mu_u + \mu_w$ satisfies

$$C^{-1}\mathrm{d}x(A) \leqslant (\mu_u + \mu_w)(A) \leqslant C\mathrm{d}x(A), \quad A \in \mathscr{B}(\mathbf{R}^d),$$

where $C > 0$ and $\mathrm{d}x$ denotes Lebesgue measure. The measure $\mu_u + \mu_w$ is thus upper and lower bounded by Lebesgue measure. We may conclude that $u + w$ resembles white noise in that all frequencies are supported in $\mu_u + \mu_w$, albeit with intensities that are uniformly upper and lower bounded rather than constant.

Remark 5.13 Finally we extract an observation concerning the Gröchenig–Sjöstrand space $M^{\infty,1}_{1\otimes\langle\cdot\rangle^r}(\mathbf{R}^d)$ for $r \geqslant 0$. If $f \in M^{\infty,1}_{1\otimes\langle\cdot\rangle^r}(\mathbf{R}^d)$ is real-valued and $f(x) \geqslant \varepsilon > 0$ for all $x \in \mathbf{R}^d$, then $1/f \in M^{\infty,1}_{1\otimes\langle\cdot\rangle^r}(\mathbf{R}^d)$. Indeed by Lemma 5.8 we have $a = 1 \otimes f \in M^{\infty,1}_{1\otimes\langle\cdot\rangle^r}(\mathbf{R}^{2d})$, and

$$(a^w(x, D)g, g)_{L^2} = (f\widehat{g}, \widehat{g})_{L^2} \geqslant \varepsilon\|g\|^2_{L^2}, \quad g \in L^2(\mathbf{R}^d),$$

so it follows from Lemma 5.10 that $a^w(x, D)$ is invertible on L^2. The inverse has Weyl symbol $a^w(x, D)^{-1} = b^w(x, D)$ with $b = 1 \otimes 1/f$. By Gröchenig's spectral invariance theorem [17, Theorem 4.6] we obtain $b \in M^{\infty,1}_{1\otimes\langle\cdot\rangle^r}(\mathbf{R}^{2d})$, and finally Lemma 5.8 yields the claim $1/f \in M^{\infty,1}_{1\otimes\langle\cdot\rangle^r}(\mathbf{R}^d)$.

5.3 Further Remarks on the Operator Equation

In this final subsection we make a few operator theoretical remarks on the operator equation

$$\mathscr{K}_u = F(\mathscr{K}_u + \mathscr{K}_w). \tag{5.21}$$

We assume that $\mathscr{K}_u, \mathscr{K}_w \in \mathscr{L}(\mathscr{H})$ for a Hilbert space $\mathscr{H}$, and $\mathscr{K}_u \geqslant 0$ and $\mathscr{K}_w \geqslant 0$ on $\mathscr{H}$ (cf. Section 5.1), but we do not assume that $\mathscr{K}_u + \mathscr{K}_w$ is invertible on $\mathscr{H}$. We seek a solution $F \in \mathscr{L}(\mathscr{H})$ to (5.21), or equivalently, since $\mathscr{K}_u^* = \mathscr{K}_u$ and $\mathscr{K}_w^* = \mathscr{K}_w$, to the equation

$$\mathscr{K}_u = (\mathscr{K}_u + \mathscr{K}_w)F^*. \tag{5.22}$$

We have

$$\ker(\mathscr{K}_u + \mathscr{K}_w) \subseteq \ker \mathscr{K}_u \cap \ker \mathscr{K}_w. \tag{5.23}$$

In fact if $(\mathscr{K}_u + \mathscr{K}_w)f = 0$ for $f \in \mathscr{H}$ then

$$0 = ((\mathscr{K}_u + \mathscr{K}_w)f, f) = (\mathscr{K}_u f, f) + (\mathscr{K}_w f, f)$$

which implies

$$0 = (\mathscr{K}_u f, f) = \|\mathscr{K}_u^{\frac{1}{2}} f\|_{\mathscr{H}}^2 \quad \text{and} \quad 0 = (\mathscr{K}_w f, f) = \|\mathscr{K}_w^{\frac{1}{2}} f\|_{\mathscr{H}}^2.$$

Thus $\mathscr{K}_u f = \mathscr{K}_u^{\frac{1}{2}} \mathscr{K}_u^{\frac{1}{2}} f = 0$ so $f \in \ker \mathscr{K}_u$, and likewise $f \in \ker \mathscr{K}_w$, which proves (5.23).

From (5.23) and $(\ker \mathscr{K}_u)^{\perp} = \overline{\operatorname{ran} \mathscr{K}_u}$ (where $\overline{X}$ denotes the closure of a subspace $X \subseteq \mathscr{H}$) we obtain

$$\operatorname{ran} \mathscr{K}_u \subseteq \overline{\operatorname{ran} \mathscr{K}_u} \subseteq \overline{\operatorname{ran}(\mathscr{K}_u + \mathscr{K}_w)}. \tag{5.24}$$

If we assume the strengthened inclusion

$$\operatorname{ran} \mathscr{K}_u \subseteq \operatorname{ran}(\mathscr{K}_u + \mathscr{K}_w) \tag{5.25}$$

then according to Douglas' lemma [5–7], the inclusion (5.25) is equivalent to the existence of $F \in \mathscr{L}(\mathscr{H})$ that solves (5.22). The filter linear operator F is the unique solution to (5.22) that satisfies

$$\begin{aligned} \operatorname{ran} F^* &\subseteq \overline{\operatorname{ran}(\mathscr{K}_u + \mathscr{K}_w)}, \\ \ker F^* &= \ker \mathscr{K}_u, \\ \|F\| &= \sup_{\|f\|=1} \frac{\|\mathscr{K}_u f\|}{\|(\mathscr{K}_u + \mathscr{K}_w)f\|}. \end{aligned} \tag{5.26}$$

We may summarize this as follows.

Proposition 5.14 *Let $\mathscr{H}$ be a Hilbert space, suppose $\mathscr{K}_u, \mathscr{K}_w \in \mathscr{L}(\mathscr{H})$, $\mathscr{K}_u, \mathscr{K}_w \geqslant 0$ on $\mathscr{H}$, and (5.25). Then there exists $F \in \mathscr{L}(\mathscr{H})$ that solves (5.21). The operator F is the unique solution that satisfies (5.26).*

Note that we do not have to assume that $\mathscr{K}_u + \mathscr{K}_w$ is invertible, as we did in Theorem 5.11.

Under the assumptions of Proposition 5.14 it follows from [6, Corollary 1] and (5.23) that F is invertible as an operator in $\mathscr{L}(\mathscr{H})$ if and only if (5.25) is strengthened into

$$\operatorname{ran} \mathscr{K}_u = \operatorname{ran}(\mathscr{K}_u + \mathscr{K}_w)$$

and $\ker \mathscr{K}_u \subseteq \ker \mathscr{K}_w$.

Remark 5.15 If we assume that $\operatorname{ran}(\mathscr{K}_u + \mathscr{K}_w)$ is closed then (5.25) is a consequence of (5.24). Then the Moore–Penrose pseudo-inverse $(\mathscr{K}_u + \mathscr{K}_w)^+ \in \mathscr{L}(\mathscr{H})$ for $\mathscr{K}_u + \mathscr{K}_w$ is well defined [4]. A solution to (5.21) is then $F = \mathscr{K}_u(\mathscr{K}_u + \mathscr{K}_w)^+$.

Acknowledgements The author is grateful to Fabio Nicola for helpful comments. The author is a member of Gruppo Nazionale per l'Analisi Matematica, la Probabilità e le loro Applicazioni (GNAMPA) – Istituto Nazionale di Alta Matematica (INdAM).

Funding Open access funding provided by Politecnico di Torino within the CRUI-CARE Agreement.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Bergh, J., Löfström, J.: Interpolation Spaces. Springer-Verlag, Berlin Heidelberg New York (1976)
2. Chen, Y., Toft, J., Wahlberg, P.: The Weyl product on quasi-Banach modulation spaces. Bull. Math. Sci. **9**(2), 1950018 (2019)
3. Conway, J.B.: A Course in Functional Analysis. Graduate Texts in Mathematics, vol. 96. Springer-Verlag, Berlin Heidelberg New York (1990)
4. Desoer, C.A., Whalen, B.H.: A note on pseudoinverses. J. Soc. Indust. Appl. Math. **11**(2), 442–447 (1963)
5. Douglas, R.G.: On majorization, factorization, and range inclusion of operators on Hilbert space. Proc. AMS **17**(2), 413–415 (1966)
6. Fillmore, P.A., Williams, J.P.: On operator ranges. Adv. Math. **7**, 254–281 (1971)
7. Forough, M.: Majorization, range inclusion, and factorization for unbounded operators on Banach spaces. Lin. Alg. Appl. **449**, 60–67 (2014)
8. Feichtinger, H.G., Gröchenig, K.: Gabor frames and time-frequency analysis of distributions. J. Funct. Anal. **146**(2), 464–495 (1997)
9. Feichtinger, H.G., Hörmann, W.: A distributional approach to generalized stochastic processes on locally compact Abelian groups In: New Perspectives on Approximation and Sampling Theory. Festschrift in Honor of Paul Butzer's 85th birthday, pp. 423–446. Birkhäuser/Springer, Cham (2014)
10. Folland, G.B.: Real Analysis. Modern Techniques and Their Applications. John Wiley (1999)
11. Fomin, V.: Optimal Filtering. Volume 1: Filtering of Stochastic Processes. Kluwer Academic Publishers (1999)
12. Fomin, V., Ruzhansky, M.: Abstract optimal linear filtering. SIAM J. Control. Optim. **38**(5), 1334–1352 (2000)
13. Gel'fand, I.M., Shilov, G.E.: Generalized Functions, vol. I. Academic Press, New York (1968)
14. Gel'fand, I.M., Vilenkin, N.Y.: Generalized Functions, vol. IV. Academic Press, New York (1968)
15. Gröchenig, K.: Foundations of Time-Frequency Analysis. Birkhäuser (2001)
16. Gröchenig, K., Leinert, M.: Wiener's lemma for twisted convolution and Gabor frames. J. Amer. Math. Soc. **17**(1), 1–18 (2004)
17. Gröchenig, K.: Time-frequency analysis of Sjöstrand's class. Rev. Mat. Iberoam. **22**(2), 703–724 (2006)
18. Gröchenig, K., Toft, J.: Isomorphism properties of Toeplitz operators and pseudo-differential operators between modulation spaces. J. Anal. Math. **114**, 255–283 (2011)
19. Hlawatsch, F., Matz, G., Kirchauer, H., Kozek, W.: Time-frequency formulation, design, and implementation of time-varying optimal filters for signal estimation. IEEE Trans. Signal Process. **48**(5), 1417–1432 (2000)
20. Hörmander, L.: The Analysis of Linear Partial Differential Operators, vol. I. III, IV. Springer, Berlin (1990)
21. Hörmann, W.: Generalized Stochastic Processes and Wigner Distribution, PhD thesis. Universität Wien (1989)

22. Janssen, A.J.E.M.: Duality and biorthogonality for Weyl-Heisenberg frames. J. Fourier Anal. Appl. **1**(4), 403–436 (1995)
23. Katznelson, Y.: An Introduction to Harmonic Analysis. John Wiley, New York (1968)
24. Keville, B.: Multidimensional Second Order Generalised Stochastic Processes on Locally Compact Abelian Groups, PhD thesis. University of Dublin (2003)
25. Kolmogorov, A.N.: Stationary sequences in Hilbert's space (Russian). Bolletin Moskovskogo Gosudarstvenogo Universiteta. Matematika **2**, (1941)
26. Larsen, R.: An Introduction to the Theory of Multipliers. Springer-Verlag, Berlin Heidelberg (1971)
27. Lieb, E. H., Loss, M.: Analysis. Graduate Studies in Mathematics, vol. 14. American Mathematical Society, Providence (1997)
28. Papoulis, A.: Signal Analysis. McGraw-Hill, Auckland (1984)
29. Reed, M., Simon, B.: Methods of Modern Mathematical Physics, vol. 1: Functional Analysis. Academic Press, New York (1980)
30. Rodino, L., Wahlberg, P.: Microlocal analysis for Gelfand-Shilov spaces. Ann. Mat. Pura Appl. **202**(5), 2379–2420 (2023)
31. Rudin, W.: Fourier Analysis on Groups. John Wiley (1962)
32. Rudin, W.: Real and Complex Analysis, McGraw-Hill (1987)
33. Shubin, M.A.: Pseudodifferential Operators and Spectral Theory. Springer (2001)
34. Sjöstrand, J.: An algebra of pseudodifferential operators. Math. Res. Lett. **1**(2), 185–192 (1994)
35. Sjöstrand, J.: Wiener Type Algebras of Pseudodifferential Operators. Séminaire sur les Équations aux Dérivées Partielles, 1994–1995, Exp. No. IV. pp. 21. École Polytech., Palaiseau (1995)
36. Toft, J.: Continuity and Schatten properties for pseudo-differential operators on modulation spaces. In: Toft, J., Wong, M.W., Zhu, H. (eds.) Modern Trends in Pseudo-Differential Operators. Operator Theory: Advances and Applications, vol. 172, pp. 173–206. Birkhäuser, Basel (2007)
37. van Neerven, J.: Functional Analysis. Cambridge Studies in Advanced Mathematics, vol. 201. Cambridge University Press, Cambridge (2022)
38. Van Trees, H.L.: Detection, Estimation, and Modulation Theory. John Wiley (1971)
39. Wahlberg, P., Schreier, P.J.: Gabor discretization of the Weyl product for modulation spaces and filtering of nonstationary stochastic processes. Appl. Comput. Harm. Anal. **26**(1), 97–120 (2009)
40. Wahlberg, P.: The Wigner distribution of Gaussian tempered generalized stochastic processes. Appl. Comput. Harmon. Anal. **79**, 101799 (2025)
41. Weinstein, L.A., Zubakov, V.D.: Extraction of Signals from Noise, Prentice-Hall (1962)
42. Wiener, N.: Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications. MIT Press (1949)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.