Journal of Functional Analysis 286 (2024) 110225

Contents lists available at ScienceDirect

Journal of Functional Analysis

journal homepage: www.elsevier.com/locate/jfa

Regular Article

Pseudo-differential calculi and entropy estimates with Orlicz modulation spaces

Anupam Gumber [a], Nimit Rana [b], Joachim Toft [c,*], Rüya Üster [d]

[a] *NuHAG, Faculty of Mathematics, University of Vienna, Vienna, Austria*
[b] *Department of Mathematics, Imperial College London, London, UK*
[c] *Department of Mathematics, Linnæus University, Växjö, Sweden*
[d] *Department of Mathematics, Faculty of Science, İstanbul University, İstanbul, Türkiye*

A R T I C L E   I N F O

Article history:
Received 24 April 2023
Accepted 20 October 2023
Available online 31 October 2023
Communicated by Camil Muscalu

MSC:
primary 35S05, 46E30, 46A16, 42B35
secondary 46F10

Keywords:
Orlicz modulation spaces
Pseudo-differential operator
Essential inverse
Entropy functional

A B S T R A C T

We deduce continuity properties for pseudo-differential operators with symbols in Orlicz modulation spaces when acting on other Orlicz modulation spaces. In particular we extend well-known results in the literature. For example we generalize the classical result by Cordero and Nicola that if

$$\frac{1}{p'}+\frac{1}{q'}\leqslant\frac{1}{p_1}+\frac{1}{p_2},\quad \frac{1}{p'}+\frac{1}{q'}\leqslant\frac{1}{q_1}+\frac{1}{q_2},\quad p_j,q_j\leqslant q',\quad q\leqslant p$$

and $a\in M^{p,q}$, then the pseudo-differential operator $\operatorname{Op}(a)$ is continuous from M^{p_1,q_1} to $M^{p_2',q_2'}$.
We also show that the entropy functional E_ϕ possess suitable continuity properties on a suitable Orlicz modulation space M^Φ satisfying $M^p\subseteq M^\Phi\subseteq M^2$, though E_ϕ is discontinuous on $M^2=L^2$.

* Corresponding author.
E-mail addresses: anupam.gumber@univie.ac.at (A. Gumber), n.rana22@imperial.ac.uk (N. Rana), joachim.toft@lnu.se (J. Toft), ruya.uster@istanbul.edu.tr (R. Üster).

https://doi.org/10.1016/j.jfa.2023.110225

0. Introduction

Pseudo-differential operators are important in several fields of science and technology. In the theory of partial differential equations, they are convenient tools for handling various kinds of problems, e.g. parametrix constructions, micro-local properties and invertibility of (hypo-)elliptic operators. In time-frequency analysis, pseudo-differential operators appear for example when modelling non-stationary filters.

A pseudo-differential operator is a rule in which for every function or distribution a (the symbol), defined on the phase space (or time-frequency shift space) $\mathbf{R}^{2d}$ assigns a linear operator $\operatorname{Op}(a)$ acting on functions or distributions defined on $\mathbf{R}^d$. The assumptions on the symbols, domains and ranks for the pseudo-differential operators, usually resemble on structures where they are applied. Therefore, in the theory of partial differential operators, one usually assumes that the symbols are smooth and that differentiations of the symbols lead to more restrictive growth/decay properties at infinity.

When using pseudo-differential operators for modelling time-dependent filters in time-frequency analysis, any similar assumptions on smoothness are usually not relevant. Here it is more relevant to assume that the involved symbols, inputs and outputs (i.e. the filter constants, ingoing signals and outgoing signals) should fulfill conditions on translation and modulation invariance, as well as certain energy estimates of the time-frequency content. This leads to that the involved functions and distributions should belong to suitable *modulation spaces* a family of functions and distribution spaces, introduced by Feichtinger in [8]. The theory of such spaces was thereafter extended in several ways (see e.g. [7,9–11,13,35,37,40,42] and the references therein).

Recently, some investigations of Orlicz modulation spaces have been performed in [34,45,46]. Such spaces are obtained by imposing an Orlicz norm estimates on the short time Fourier transforms of the involved functions and distributions. By the definition it follows that the family of Orlicz modulation spaces contain all classical modulation spaces $M^{p,q}_{(\omega)}(\mathbf{R}^d)$, introduced by Feichtinger in [8], which essentially follows from the fact that the family of Orlicz spaces contains all Lebesgue spaces. (See [21] and Section 1 for notations.) On the other hand, the Orlicz modulation spaces becomes a subfamily of broader classes of modulation spaces, given in e.g. [9,30,31].

A question which might appear is whether there are relevant situations where it is fruitful to search among Orlicz modulation spaces to deduce sharper estimates compared to classical modulation spaces. For example, consider the entropy functional on short-time Fourier transforms

$$E(f) = E_\phi(f) \equiv - \iint_{\mathbf{R}^{2d}} |V_\phi f(x,\xi)|^2 \log |V_\phi f(x,\xi)|^2 \, dx d\xi + \|V_\phi f\|^2_{L^2} \log \|V_\phi f\|^2_{L^2}. \quad (0.1)$$

Here $\phi \in \mathscr{S}(\mathbf{R}^d) \setminus 0$ is fixed, and as usual we set

$$0 \log 0 \equiv \lim_{t \to 0+} t \log t = 0.$$

We recall the entropy condition

$$E_\phi(f) \geqslant d\left(1+\log(\tfrac{\pi}{2})\right), \quad \text{when} \quad \|f\|_{L^2}\|\phi\|_{L^2}=1, \tag{0.2}$$

which is essential in certain types of estimates of the kinetic energy in quantum systems (see e.g. [23–26] and the references therein).

For the Orlicz modulation space

$$M^\Phi(\mathbf{R}^d), \qquad \Phi(t)=-t^2\log t, \quad 0\leqslant t\leqslant e^{-\frac{2}{3}}, \tag{0.3}$$

we observe that the Young function Φ resembles with the structures of the entropy functional E_ϕ. A question then appear whether the space in (0.3) is better designed as domain for E_ϕ, compared to the strictly larger space $M^2(\mathbf{R}^d)=L^2(\mathbf{R}^d)$, which is usually taken as the domain for E_ϕ (cf. [23,24]).

We also notice that $M^\Phi(\mathbf{R}^d)$ in (0.3) makes sense, while

$$L^\Phi(\mathbf{R}^d), \qquad \Phi(t)=-t^2\log t, \qquad 0\leqslant t<\infty \tag{0.4}$$

does not makes sense as an Orlicz space. (See Theorem 3.1 and Lemma 3.2 in Section 3 for details.)

In the first part of the paper we investigate mapping properties for pseudo-differential operators $\operatorname{Op}(a)$ with symbols a belonging to suitable modulation spaces or Orlicz modulation spaces, when acting on Orlicz modulation spaces. In particular we find suitable conditions on the Young functions Φ_j, Φ, Ψ_j and Ψ, $j=1,2$, in order for the pseudo-differential operators

$$\operatorname{Op}(a): M^{\Phi_1,\Psi_1}(\mathbf{R}^d)\to M^{\Phi_2^*,\Psi_2^*}(\mathbf{R}^d) \tag{0.5}$$

and

$$\operatorname{Op}(a): M^{\Phi^*,\Psi^*}(\mathbf{R}^d)\to W^{\Phi,\Psi}(\mathbf{R}^d) \tag{0.6}$$

are well-defined and continuous.

For example, the following two propositions are consequences of Theorems 2.9 and 2.10 in Section 2. Here and in what follows we let $p'\in[1,\infty]$ be the conjugate Lebesgue exponent of $p\in[1,\infty]$, i.e. p and p' should satisfy $\frac{1}{p}+\frac{1}{p'}=1$, and similarly for other Lebesgue exponents.

Proposition 0.1. *Let $p,q\in[1,\infty]$ be such that $q\leqslant p$ and $p>1$. Also let $\Phi_j,\Psi_j:[0,\infty]\to[0,\infty]$, $j=1,2$, be such that $t\mapsto\Phi_j(t^{\frac{1}{p'}})$ and $t\mapsto\Psi_j(t^{\frac{1}{p'}})$ are Young functions which fulfill the Δ_2-condition, and*

$$\Phi_1(t),\Phi_2(t)\gtrsim t^{q'} \quad \Psi_1(t),\Psi_2(t)\gtrsim t^{q'}, \quad t\geqslant 0,$$

and

$$\Phi_1^{-1}(s)\Phi_2^{-1}(s) \lesssim s^{\frac{1}{p'}+\frac{1}{q'}}, \quad \Psi_1^{-1}(s)\Psi_2^{-1}(s) \lesssim s^{\frac{1}{p'}+\frac{1}{q'}}, \quad s \geqslant 0.$$

If $a \in M^{p,q}(\mathbf{R}^{2d})$*, then* $\operatorname{Op}(a)$ *is continuous from* $M^{\Phi_1,\Psi_1}(\mathbf{R}^d)$ *to* $M^{\Phi_2^*,\Psi_2^*}(\mathbf{R}^d)$*.*

Proposition 0.2. *Let* Φ *and* Ψ *be Young functions which satisfy the* Δ_2*-condition, and let* $a \in W^{\Psi,\Phi}(\mathbf{R}^{2d})$*. Then* $\operatorname{Op}(a)$ *is continuous from* $M^{\Phi^*,\Psi^*}(\mathbf{R}^d)$ *to* $W^{\Psi,\Phi}(\mathbf{R}^d)$*.*

More generally, we deduce weighted versions of such continuity results, as well as relax the assumptions on the Young functions in such way that they only need to fulfill a *local* Δ_2-condition near origin (see Definition 1.7). The essential ingredient for such local condition is the fact that Orlicz modulation spaces are completely determined by the behaviour of the Young functions *near the origin*, and the involved weight functions. (See e.g. [46, Proposition 5.11].) Since Orlicz spaces contain Lebesgue spaces as special cases, it follows that Orlicz modulation spaces contain the classical modulation spaces. Hence, our results also lead to continuity properties for pseudo-differential operators acting on (classical) modulation spaces. More specifically, by choosing the Young functions in Proposition 0.1 and involved weight functions in suitable ways, our main result Theorem 2.9 in Section 2 includes the optimal result [4, Theorem 5.1] by Cordero and Nicola as special case. In the case of unweighted spaces [4, Theorem 5.1] can be expressed as in the following.

Proposition 0.3. *Suppose that* $p, p_j, q, q_j \in [1,\infty]$*,* $j = 1, 2$*, satisfy*

$$\frac{1}{p'} + \frac{1}{q'} \leqslant \frac{1}{p_1} + \frac{1}{p_2}, \quad \frac{1}{p'} + \frac{1}{q'} \leqslant \frac{1}{q_1} + \frac{1}{q_2}, \quad p_1, q_1, p_2, q_2 \leqslant q', \quad q \leqslant p,$$

and let $a \in M^{p,q}(\mathbf{R}^{2d})$*. Then*

$$\operatorname{Op}(a) : M^{p_1,q_1}(\mathbf{R}^d) \to M^{p_2',q_2'}(\mathbf{R}^d), \tag{0.7}$$

is continuous.

Proposition 0.3 is a special case of Theorem 2.9 in Section 2. If in addition $p > 1$, then Proposition 0.3 also follows from Proposition 0.1. We also observe that for weighted (Orlicz) modulation spaces, Theorem 2.9 in Section 2 permits more general weights in the involved spaces, compared to [4, Theorem 5.1].

There are relevant situations where Proposition 0.1 and its extension Theorem 2.9 can be applied, while earlier classical results in e.g. [3,17,18,35–43] seem not to be applicable. For example it follows from Proposition 0.1 that if $p > 2$, $a \in M^{p,p'}(\mathbf{R}^{2d})$, then the map

$$\operatorname{Op}(a) : M^{\Phi}(\mathbf{R}^d) \to M^{\Phi}(\mathbf{R}^d), \quad \Phi(t) = -t^2 \log t,\ t \in [0, e^{-\frac{2}{3}}], \tag{0.8}$$

on Orlicz spaces in (0.3), is continuous. (See Example 2.11 in Section 2 and Remark 3.8 in Section 3.) Any similar continuity property is obviously not reachable from the investigations in [3,17,18,35–43].

In the last part of the paper we investigate continuity of the entropy functional E_ϕ when acting on $M^p(\mathbf{R}^d)$ for $p \in [1,2]$ and $M^\Phi(\mathbf{R}^d)$ in (0.3). More precisely, in Section 3 we show that E_ϕ is continuous on $M^\Phi(\mathbf{R}^d)$ and on $M^p(\mathbf{R}^d)$ for $p \in [1,2)$, but fails to be continuous on $M^2(\mathbf{R}^d)$. This might be surprising due to the dense embeddings

$$M^p(\mathbf{R}^d) \subseteq M^\Phi(\mathbf{R}^d) \subseteq M^2(\mathbf{R}^d), \qquad p < 2,$$

which shows that $M^\Phi(\mathbf{R}^d)$ in some sense is close to $M^2(\mathbf{R}^d)$. See Theorem 3.1 and Lemma 3.2 for details.

Acknowledgment

J. Toft was supported by Vetenskapsrådet, project number 2019-04890. Anupam Gumber and Nimit Rana thanks the Linnaeus university for providing excellent research facilities and kind hospitality during their academic visit. A. Gumber was also supported by the Austrian Science Fund (FWF) project P33217.

1. Preliminaries

In the section we recall some basic facts on Gelfand-Shilov spaces, Orlicz spaces, Orlicz modulation spaces, pseudo-differential operators and Wigner distributions. We also give some examples on Young functions, Orlicz spaces and Orlicz modulation spaces. (See Examples (0.8) and 1.17.) Notice that Young functions are fundamental in the definition of Orlicz spaces and Orlicz modulation spaces).

1.1. Gelfand-Shilov spaces

For a real number $s > 0$, the (standard Fourier invariant) Gelfand-Shilov space $\mathcal{S}_s(\mathbf{R}^d)$ ($\Sigma_s(\mathbf{R}^d)$) of Roumieu type (Beurling type) consists of all $f \in C^\infty(\mathbf{R}^d)$ such that

$$\|f\|_{\mathcal{S}_{s,h}} \equiv \sup_{\alpha,\beta \in \mathbf{N}^d} \left(\frac{\|x^\beta \partial^\alpha f\|_{L^\infty}}{h^{|\alpha+\beta|}(\alpha!\beta!)^s} \right) \tag{1.1}$$

is finite for some $h > 0$ (for every $h > 0$). We equip $\mathcal{S}_s(\mathbf{R}^d)$ ($\Sigma_s(\mathbf{R}^d)$) by the canonical inductive limit topology (projective limit topology) with respect to $h > 0$, induced by the semi-norms in (1.1).

We have

$$\begin{aligned} &\mathcal{S}_s(\mathbf{R}^d) \hookrightarrow \Sigma_t(\mathbf{R}^d) \hookrightarrow \mathcal{S}_t(\mathbf{R}^d) \hookrightarrow \mathscr{S}(\mathbf{R}^d) \\ &\qquad \hookrightarrow \mathscr{S}'(\mathbf{R}^d) \hookrightarrow \mathcal{S}_t'(\mathbf{R}^d) \hookrightarrow \Sigma_t'(\mathbf{R}^d) \hookrightarrow \mathcal{S}_s'(\mathbf{R}^d), \quad \frac{1}{2} \leqslant s < t, \end{aligned} \tag{1.2}$$

with dense embeddings. Here $A \hookrightarrow B$ means that the topological space A is continuously embedded in the topological space B. We also have

$$\mathcal{S}_s(\mathbf{R}^d) = \Sigma_t(\mathbf{R}^d) = \{0\}, \qquad s < \frac{1}{2},\ t \leqslant \frac{1}{2}.$$

The *Gelfand-Shilov distribution spaces* $\mathcal{S}_s'(\mathbf{R}^d)$ and $\Sigma_s'(\mathbf{R}^d)$, of Roumieu and Beurling types respectively, are the (strong) duals of $\mathcal{S}_s(\mathbf{R}^d)$ and $\Sigma_s(\mathbf{R}^d)$, respectively. It follows that if $\mathcal{S}_{s,h}'(\mathbf{R}^d)$ is the L^2-dual of $\mathcal{S}_{s,h}(\mathbf{R}^d)$ and $s \geqslant \frac{1}{2}$ ($s > \frac{1}{2}$), then $\mathcal{S}_s'(\mathbf{R}^d)$ ($\Sigma_s'(\mathbf{R}^d)$) can be identified with the projective limit (inductive limit) of $\mathcal{S}_{s,h}'(\mathbf{R}^d)$ with respect to $h > 0$. It follows that

$$\mathcal{S}_s'(\mathbf{R}^d) = \bigcap_{h>0} \mathcal{S}_{s,h}'(\mathbf{R}^d) \quad \text{and} \quad \Sigma_s'(\mathbf{R}^d) = \bigcup_{h>0} \mathcal{S}_{s,h}'(\mathbf{R}^d) \tag{1.3}$$

for such choices of s and σ, see [14,27,28] for details.

We let the Fourier transform $\mathscr{F}$ be given by

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

when $f \in L^1(\mathbf{R}^d)$. Here $\langle\,\cdot\,,\,\cdot\,\rangle$ denotes the usual scalar product on $\mathbf{R}^d$. The Fourier transform $\mathscr{F}$ extends uniquely to homeomorphisms on $\mathscr{S}'(\mathbf{R}^d)$, $\mathcal{S}_s'(\mathbf{R}^d)$ and on $\Sigma_s'(\mathbf{R}^d)$. Furthermore, $\mathscr{F}$ restricts to homeomorphisms on $\mathscr{S}(\mathbf{R}^d)$, $\mathcal{S}_s(\mathbf{R}^d)$ and on $\Sigma_s(\mathbf{R}^d)$, and to a unitary operator on $L^2(\mathbf{R}^d)$. Similar facts hold true with partial Fourier transforms in place of Fourier transform.

Let $\phi \in \mathscr{S}(\mathbf{R}^d)$ be fixed. Then the *short-time Fourier transform* $V_\phi f$ of $f \in \mathscr{S}'(\mathbf{R}^d)$ with respect to the *window function* ϕ is the tempered distribution on $\mathbf{R}^{2d}$, defined by

$$V_\phi f(x,\xi) = \mathscr{F}(f\,\overline{\phi(\,\cdot\, - x)})(\xi), \quad x,\xi \in \mathbf{R}^d.$$

If $f,\phi \in \mathscr{S}(\mathbf{R}^d)$, then it follows that

$$V_\phi f(x,\xi) = (2\pi)^{-\frac{d}{2}} \int_{\mathbf{R}^d} f(y)\overline{\phi(y-x)} e^{-i\langle y,\xi\rangle}\, dy, \quad x,\xi \in \mathbf{R}^d.$$

By [40, Theorem 2.3] it follows that the definition of the map $(f,\phi) \mapsto V_\phi f$ from $\mathscr{S}(\mathbf{R}^d) \times \mathscr{S}(\mathbf{R}^d)$ to $\mathscr{S}(\mathbf{R}^{2d})$ is uniquely extendable to a continuous map from $\mathcal{S}_s'(\mathbf{R}^d) \times \mathcal{S}_s'(\mathbf{R}^d)$ to $\mathcal{S}_s'(\mathbf{R}^{2d})$, and restricts to a continuous map from $\mathcal{S}_s(\mathbf{R}^d) \times \mathcal{S}_s(\mathbf{R}^d)$ to $\mathcal{S}_s(\mathbf{R}^{2d})$. The same conclusion holds with Σ_s in place of $\mathcal{S}_s$, at each occurrence.

In the following proposition we give characterizations of Gelfand-Shilov spaces and their distribution spaces in terms of estimates of the short-time Fourier transform. We omit the proof since the first part follows from [19, Theorem 2.7] and the second part from [44, Proposition 2.2]. See also [6] for related results. Here and in what follows, $A(\theta) \lesssim B(\theta)$, $\theta \in \Omega$, means that there is a constant $c > 0$ such that $A(\theta) \leqslant cB(\theta)$ holds for all $\theta \in \Omega$. We also set $A(\theta) \asymp B(\theta)$ when $A(\theta) \lesssim B(\theta) \lesssim A(\theta)$.

Proposition 1.1. *Let $s \geqslant \frac{1}{2}$ ($s > \frac{1}{2}$), $\phi \in \mathcal{S}_s(\mathbf{R}^d) \setminus 0$ ($\phi \in \Sigma_s(\mathbf{R}^d) \setminus 0$) and let f be a Gelfand-Shilov distribution on $\mathbf{R}^d$. Then the following is true:*

(1) *$f \in \mathcal{S}_s(\mathbf{R}^d)$ ($f \in \Sigma_s(\mathbf{R}^d)$), if and only if*

$$|V_\phi f(x,\xi)| \lesssim e^{-r(|x|^{\frac{1}{s}}+|\xi|^{\frac{1}{s}})}, \quad x,\xi \in \mathbf{R}^d, \tag{1.4}$$

for some $r > 0$ (for every $r > 0$).

(2) *$f \in \mathcal{S}_s'(\mathbf{R}^d)$ ($f \in \Sigma_s'(\mathbf{R}^d)$), if and only if*

$$|V_\phi f(x,\xi)| \lesssim e^{r(|x|^{\frac{1}{s}}+|\xi|^{\frac{1}{s}})}, \quad x,\xi \in \mathbf{R}^d, \tag{1.5}$$

for every $r > 0$ (for some $r > 0$).

1.2. Weight functions

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

$$\omega(x+y) \lesssim \omega(x)v(y), \qquad x,y \in \mathbf{R}^d. \tag{1.6}$$

If ω and v are weights on $\mathbf{R}^d$ such that (1.6) holds, then ω is also called *v-moderate*. We note that (1.6) implies that ω fulfills the estimates

$$v(-x)^{-1} \lesssim \omega(x) \lesssim v(x), \quad x \in \mathbf{R}^d. \tag{1.7}$$

We let $\mathscr{P}_E(\mathbf{R}^d)$ be the set of all moderate weights on $\mathbf{R}^d$.

It can be proved that if $\omega \in \mathscr{P}_E(\mathbf{R}^d)$, then ω is v-moderate for some $v(x) = e^{r|x|}$, provided the positive constant r is large enough (cf. [16]). That is, (1.6) implies

$$\omega(x+y) \lesssim \omega(x)e^{r|y|} \tag{1.8}$$

for some $r > 0$. In particular, (1.7) shows that for any $\omega \in \mathscr{P}_E(\mathbf{R}^d)$, there is a constant $r > 0$ such that

$$e^{-r|x|} \lesssim \omega(x) \lesssim e^{r|x|}, \quad x \in \mathbf{R}^d. \tag{1.9}$$

We say that v is *submultiplicative* if v is even and (1.6) holds with $\omega = v$. In the sequel, v and v_j for $j \geqslant 0$, always stand for submultiplicative weights if nothing else is stated.

We let $\mathscr{P}_E^0(\mathbf{R}^d)$ be the set of all $\omega \in \mathscr{P}_E(\mathbf{R}^d)$ such that (1.8) holds for every $r > 0$. We also let $\mathscr{P}(\mathbf{R}^d)$ be the set of all $\omega \in \mathscr{P}_E(\mathbf{R}^d)$ such that

$$\omega(x+y) \lesssim \omega(x)(1+|y|)^r$$

for some $r > 0$. Evidently,

$$\mathscr{P}(\mathbf{R}^d) \subseteq \mathscr{P}_E^0(\mathbf{R}^d) \subseteq \mathscr{P}_E(\mathbf{R}^d).$$

1.3. Orlicz spaces

We recall that a function $\Phi : [0,\infty] \to [0,\infty]$ is called *convex* if

$$\Phi(s_1t_1 + s_2t_2) \leqslant s_1\Phi(t_1) + s_2\Phi(t_2),$$

when $s_j, t_j \in \mathbf{R}$ satisfy $s_j, t_j \geqslant 0$ and $s_1 + s_2 = 1,\ j = 1, 2$.

Definition 1.2. A function Φ_0 from $[0,\infty]$ to $[0,\infty]$ is called a *Young function* if the following is true:

(1) Φ_0 is convex;
(2) $\Phi_0(0) = 0$;
(3) $\lim_{t\to\infty} \Phi_0(t) = \Phi_0(\infty) = \infty$.

A function Φ from $[0,\infty]$ to $[0,\infty]$ is called a *quasi-Young function* (of order $p_0 \in (0,1]$) if there is a Young function Φ_0 such that $\Phi(t) = \Phi_0(t^{p_0})$ when $t \in [0,\infty]$.

We observe that Φ_0 and Φ in Definition 1.2 might not be continuous, because we permit ∞ as function value. For example, if $a > 0$, then

$$\Phi(t) = \begin{cases} 0, & \text{when } t \leqslant a \\ \infty, & \text{when } t > a \end{cases}$$

is convex on $[0,\infty]$ but discontinuous at $t = a$.

It is clear that Φ_0 and Φ in Definition 1.2 are non-decreasing, because if $0 \leqslant t_1 \leqslant t_2$ and $s \in [0,1]$ is chosen such that $t_1 = st_2$ and Φ_0 is the same as in Definition 1.2, then

$$\Phi_0(t_1) = \Phi_0(st_2 + (1-s)0) \leqslant s\Phi_0(t_2) + (1-s)\Phi_0(0) \leqslant \Phi_0(t_2),$$

since $\Phi_0(0) = 0$ and $s \in [0,1]$. Hence every (quasi-)Young function is increasing.

Definition 1.3. Let Φ be a (quasi-)Young function and let $\omega_0 \in \mathscr{P}_E(\mathbf{R}^d)$. Then the Orlicz space $L^\Phi_{(\omega_0)}(\mathbf{R}^d)$ consists of all measurable functions $f : \mathbf{R}^d \to \mathbf{C}$ such that

$$\|f\|_{L^{\Phi}_{(\omega_0)}} \equiv \inf\left\{\lambda > 0\,;\, \int_\Omega \Phi\left(\frac{|f(x)\cdot\omega_0(x)|}{\lambda}\right)\,dx \leqslant 1\right\}$$

is finite. Here f and g in $L^{\Phi}_{(\omega_0)}(\mathbf{R}^d)$ are equivalent if $f = g$ a.e.

We recall that for any (positive) measure μ to a measurable set E, the Orlicz space $L^{\Phi}(\mu)$ is defined in similar ways as in Definition 1.3. (Cf. e.g. [29].)

We will also consider Orlicz spaces parameterized with two (quasi-)Young functions.

Definition 1.4. Let Φ_j be (quasi-)Young functions, $j = 1, 2$ and let $\omega \in \mathscr{P}_E(\mathbf{R}^{2d})$.

(1) The mixed Orlicz space $L^{\Phi_1,\Phi_2}_{(\omega)} = L^{\Phi_1,\Phi_2}_{(\omega)}(\mathbf{R}^{2d})$ consists of all measurable functions $f : \mathbf{R}^{2d} \to \mathbf{C}$ such that

$$\|f\|_{L^{\Phi_1,\Phi_2}_{(\omega)}} \equiv \|f_{1,\omega}\|_{L^{\Phi_2}},$$

is finite, where

$$f_{1,\omega}(x_2) = \|f(\,\cdot\,, x_2)\omega(\,\cdot\,, x_2)\|_{L^{\Phi_1}}.$$

(2) The mixed Orlicz space $L^{\Phi_1,\Phi_2}_{*,(\omega)} = L^{\Phi_1,\Phi_2}_{*,(\omega)}(\mathbf{R}^{2d})$ consists of all measurable functions $f : \mathbf{R}^{2d} \to \mathbf{C}$ such that

$$\|f\|_{L^{\Phi_1,\Phi_2}_{*,(\omega)}} \equiv \|g\|_{L^{\Phi_2,\Phi_1}_{(\omega_0)}},$$

is finite, where

$$g(x,\xi) = f(\xi,x),\ \ \omega_0(x,\xi) = \omega(\xi,x).$$

In most of our situations we assume that Φ and Φ_j above are Young functions. A few properties for Wigner distributions in Section 2 are deduced when Φ and Φ_j are allowed to be quasi-Young functions. The reader who is not interested of such general results may always assume that all quasi-Young functions are Young functions.

It is well-known that if Φ, Φ_1 and Φ_2 in Definitions 1.3 and 1.4 are Young functions, then the spaces $L^{\Phi}_{(\omega_0)}(\mathbf{R}^d)$ and $L^{\Phi_1,\Phi_2}_{(\omega)}(\mathbf{R}^{2d})$ are Banach spaces (see e.g. Theorem 3 of III.3.2 and Theorem 10 of III.3.3 in [29]). If more generally, Φ, Φ_1 and Φ_2 are quasi-Young functions of order $p_0 \in (0, 1]$, then $L^{\Phi}_{(\omega_0)}(\mathbf{R}^d)$ and $L^{\Phi_1,\Phi_2}_{(\omega)}(\mathbf{R}^{2d})$ are quasi-Banach spaces of order p_0. For the reader who is not familiar with quasi-Banach spaces we here give the definition.

Definition 1.5. Let $\mathscr{B}$ be a vector space. Then the functional $\|\cdot\|_{\mathscr{B}}$ on $\mathscr{B}$ is called a quasi-norm of order $p_0 \in (0, 1]$, or an p_0-norm, if the following conditions are fulfilled:

(1) $\|f\|_{\mathscr{B}} \geqslant 0$ with equality only for $f = 0$;
(2) $\|\alpha f\|_{\mathscr{B}} = |\alpha|\|f\|_{\mathscr{B}}$ for every $\alpha \in \mathbf{C}$ and $f \in \mathscr{B}$;
(3) $\|f + g\|_{\mathscr{B}}^{p_0} \leqslant \|f\|_{\mathscr{B}}^{p_0} + \|g\|_{\mathscr{B}}^{p_0}$ for every $f, g \in \mathscr{B}$.

The space $\mathscr{B}$ is called a quasi-Banach space (of order p_0) or a p_0-Banach space, if $\mathscr{B}$ is complete under the topology induced by the quasi-norm $\| \cdot \|_{\mathscr{B}}$.

We refer to [45, Lemma 1.18] for the proof of the following lemma.

Lemma 1.6. *Let Φ, Φ_j be quasi-Young functions, $j = 1, 2$, $\omega_0, v_0 \in \mathscr{P}_E(\mathbf{R}^d)$ and $\omega, v \in \mathscr{P}_E(\mathbf{R}^{2d})$ be such that ω_0 is v_0-moderate and ω is v-moderate. Then $L^{\Phi}_{(\omega_0)}(\mathbf{R}^d)$ and $L^{\Phi_1,\Phi_2}_{(\omega)}(\mathbf{R}^{2d})$ are invariant under translations, and*

$$\|f(\cdot - x)\|_{L^{\Phi}_{(\omega_0)}} \lesssim \|f\|_{L^{\Phi}_{(\omega_0)}} v_0(x), \quad f \in L^{\Phi}_{(\omega_0)}(\mathbf{R}^d),\ x \in \mathbf{R}^d\ ,$$

and

$$\|f(\cdot - (x, \xi))\|_{L^{\Phi_1,\Phi_2}_{(\omega)}} \lesssim \|f\|_{L^{\Phi_1,\Phi_2}_{(\omega)}} v(x, \xi), \quad f \in L^{\Phi_1,\Phi_2}_{(\omega)}(\mathbf{R}^{2d}),\ (x, \xi) \in \mathbf{R}^{2d}.$$

In most situations we assume that the (quasi-)Young functions should satisfy the Δ_2-condition (near origin), whose definition is recalled as follows.

Definition 1.7. Let $\Phi : [0, \infty] \to [0, \infty]$ be a (quasi-)Young function. Then Φ is said to satisfy the Δ_2-*condition* if there exists a constant $C > 0$ such that

$$\Phi(2t) \leqslant C\Phi(t) \tag{1.10}$$

for every $t \in [0, \infty]$. The (quasi-)Young function Φ is said to satisfy *local Δ_2-condition* or *Δ_2-condition near origin*, if there are constants $r > 0$ and $C > 0$ such that (1.10) holds when $t \in [0, r]$.

Remark 1.8. Suppose that $\Phi : [0, \infty] \to [0, \infty]$ is a (quasi-)Young function which satisfies (1.10) when $t \in [0, r]$ for some constants $r > 0$ and $C > 0$. Then it follows by straightforward arguments that there is a quasi-Young function Φ_0 (of the same order) which satisfies the Δ_2-condition (on the whole $[0, \infty)$, and such that $\Phi_0(t) = \Phi(t)$ when $t \in [0, r]$).

Several duality properties for Orlicz spaces can be described in terms of Orlicz spaces with respect to Young conjugates, given in the following definition.

Definition 1.9. Let Φ be a Young function. Then the conjugate Young function Φ^* is given by

$$\Phi^*(t) \equiv \begin{cases} \sup_{s\geqslant 0}(st - \Phi(s)), & \text{when } t \in [0,\infty), \\ \infty, & \text{when } t = \infty. \end{cases} \tag{1.11}$$

Remark 1.10. Let $p \in (0,\infty]$, and set $\Phi_{[p]}(t) = \frac{t^p}{p}$ when $p \in (0,\infty)$, and

$$\Phi_{[\infty]}(t) = \begin{cases} 0, & t \leqslant 1, \\ \infty, & t > 1. \end{cases}$$

Then $L^{\Phi_{[p]}}(\mathbf{R}^d)$ and its (quasi-)norm is equal to the classical Lebesgue space $L^p(\mathbf{R}^d)$ and its (quasi-)norm. We observe that $\Phi_{[p]}$ is a Banach space when $p \geqslant 1$ and a quasi-Banach space of order p when $p \leqslant 1$.

Due to the previous remark we observe that there are Young functions which are not injective and thereby fail to be invertible. In the following definition we define some sort of pseudo-inverses for any quasi-Young function, especially included such functions which are not invertible.

Definition 1.11. Let $\Phi : [0,\infty] \to [0,\infty]$ be a (quasi-)Young function and let

$$t_1 = \sup\{\, t \geqslant 0\,;\, \Phi(t) = 0 \,\},$$
$$t_2 = \sup\{\, t \geqslant 0\,;\, \Phi(t) < \infty \,\}$$

and

$$s_0 = \sup\{\, \Phi(t)\,;\, t < t_2 \,\}.$$

Then t_1 is called the *zero point* and t_2 is called the *infinity point* for Φ, and the *essential inverse* $\Phi^{-\&} : [0,\infty] \to [0,\infty]$ for Φ is given by

$$\Phi^{-\&}(s) = \begin{cases} 0, & s = 0, \\ t, & s = \Phi(t),\ t_1 < t < t_2, \\ t_2, & s \geqslant s_0. \end{cases}$$

Example 1.12. We observe that if $t_1 = 0$ and $t_2 = \infty$ in Definition 1.11, then Φ is invertible and $\Phi^{-\&}$ agrees with the inverse Φ^{-1} of Φ. For example, for $\Phi_{[p]}$ with $p < \infty$ in Remark 1.10 we have

$$\Phi_{[p]}^{-\&}(s) = \Phi_{[p]}^{-1}(s) = \begin{cases} (ps)^{\frac{1}{p}}, & 0 \leqslant s < \infty, \\ \infty, & s = \infty. \end{cases}$$

For $p = \infty$, the inverse to $\Phi_{[\infty]}$ does not exist, while the essential inverse becomes

$$\Phi_{[\infty]}^{-\&}(s) = \begin{cases} 0, & s = 0, \\ 1, & s > 0. \end{cases}$$

Another example of a Young function is

$$\Phi(t) = \begin{cases} \tan t, & 0 \leqslant t < \frac{\pi}{2}, \\ \infty, & t \geqslant \frac{\pi}{2}, \end{cases}$$

which also fails to be invertible. The essential inverse becomes

$$\Phi^{-\&}(s) = \begin{cases} \arctan s, & 0 \leqslant s < \infty, \\ \frac{\pi}{2}, & s = \infty. \end{cases}$$

We also observe that

$$\Phi(t) = \begin{cases} 0, & t = 0, \\ -\frac{t}{\ln t}, & 0 < t < 1, \\ \infty, & t \geqslant 1, \end{cases}$$

is a Young function which satisfies the Δ_2-condition. Its essential inverse is

$$\Phi^{-\&}(s) = \begin{cases} \Phi^{-1}(s), & 0 \leqslant s < \infty, \\ 1, & s = \infty. \end{cases}$$

We notice that the conjugate Young function of Φ is given by

$$\Phi^*(t) = \left(t + \frac{1}{2} - \sqrt{\frac{1}{4} + t} \right) e^{-\frac{1}{t}\left(\frac{1}{2} + \sqrt{\frac{1}{4}+t}\,\right)},$$

when $t \geqslant 0$ is near origin.

We observe that each one of these Young functions gives rise to different Orlicz spaces.

We refer to [20,29,34] for more facts about Orlicz spaces.

1.4. Orlicz modulation spaces

Before considering Orlicz modulation spaces, we recall the definition of classical modulation spaces. (Cf. [8,9].)

Definition 1.13. Let $\phi(x) = \pi^{-\frac{d}{4}} e^{-\frac{|x|^2}{2}}$, $x \in \mathbf{R}^d$, $p, q \in (0, \infty]$, $\omega \in \mathscr{P}_E(\mathbf{R}^{2d})$, and let Φ and Ψ be (quasi-)Young functions.

(1) The *modulation spaces* $M^{p,q}_{(\omega)}(\mathbf{R}^d)$ is set of all $f \in \mathcal{S}'_{1/2}(\mathbf{R}^d)$ such that

$$\|f\|_{M^{p,q}_{(\omega)}} \equiv \|V_\phi f\|_{L^{p,q}_{(\omega)}} \tag{1.12}$$

is finite. The topology of $M^{p,q}_{(\omega)}(\mathbf{R}^d)$ is induced by the norm (1.12).

(2) The *Orlicz modulation spaces* $M^{\Phi}_{(\omega)}(\mathbf{R}^d)$, $M^{\Phi,\Psi}_{(\omega)}(\mathbf{R}^d)$ and $W^{\Phi,\Psi}_{(\omega)}(\mathbf{R}^d)$ are the sets of all $f \in \mathcal{S}'_{1/2}(\mathbf{R}^d)$ such that

$$\|f\|_{M^{\Phi}_{(\omega)}} \equiv \|V_\phi f\|_{L^{\Phi}_{(\omega)}}, \quad \|f\|_{M^{\Phi,\Psi}_{(\omega)}} \equiv \|V_\phi f\|_{L^{\Phi,\Psi}_{(\omega)}} \quad \text{and} \quad \|f\|_{W^{\Phi,\Psi}_{(\omega)}} \equiv \|V_\phi f\|_{L^{\Phi,\Psi}_{*,(\omega)}} \tag{1.13}$$

respectively are finite. The topologies of $M^{\Phi}_{(\omega)}(\mathbf{R}^d)$, $M^{\Phi,\Psi}_{(\omega)}(\mathbf{R}^d)$ and $W^{\Phi,\Psi}_{(\omega)}(\mathbf{R}^d)$ are induced by the respective norms in (1.13).

Let Φ and Ψ be quasi-Young functions, and let $\Phi_{[p]}$ be the same as in Remark 1.10 and $\omega \in \mathscr{P}_E(\mathbf{R}^{2d})$. Then evidently

$$M^{p,q}_{(\omega)}(\mathbf{R}^d) = M^{\Phi_{[p]},\Psi_{[q]}}_{(\omega)}(\mathbf{R}^d). \tag{1.14}$$

We now set

$$M^{p,\Psi}_{(\omega)}(\mathbf{R}^d) = M^{\Phi_{[p]},\Psi}_{(\omega)}(\mathbf{R}^d) \quad \text{and} \quad M^{\Phi,q}_{(\omega)}(\mathbf{R}^d) = M^{\Phi,\Psi_{[q]}}_{(\omega)}(\mathbf{R}^d). \tag{1.15}$$

For conveniency we also set

$$\begin{aligned} M^{p,q} = M^{p,q}_{(\omega)}, \quad M^{p,\Psi} = M^{p,\Psi}_{(\omega)}, \quad M^{\Phi,q} = M^{\Phi,q}_{(\omega)}, \\ M^{\Phi} = M^{\Phi}_{(\omega)}, \quad M^{\Phi,\Psi} = M^{\Phi,\Psi}_{(\omega)} \qquad \text{when} \quad \omega(x,\xi) = 1, \end{aligned}$$

and $M^p = M^{p,p}$ and $M^p_{(\omega)} = M^{p,p}_{(\omega)}$.

Next we explain some basic properties of Orlicz modulation spaces. The following proposition shows that Orlicz modulation spaces are completely determined by the behaviour of the quasi-Young functions near origin. We refer to [46, Proposition 5.11] for the proof.

Proposition 1.14. *Let Φ_j and Ψ_j, $j = 1, 2$, be quasi-Young functions and $\omega \in \mathscr{P}_E(\mathbf{R}^{2d})$. Then the following conditions are equivalent:*

(1) $M^{\Phi_1,\Psi_1}_{(\omega)}(\mathbf{R}^d) \subseteq M^{\Phi_2,\Psi_2}_{(\omega)}(\mathbf{R}^d)$*;*
(2) *for some $t_0 > 0$ it holds $\Phi_2(t) \lesssim \Phi_1(t)$ and $\Psi_2(t) \lesssim \Psi_1(t)$ when $t \in [0, t_0]$.*

The next two proposition show some further convenient properties for Orlicz modulation spaces.

Proposition 1.15. *Let Φ, Φ_j, Ψ, Ψ_j be quasi-Young functions, and let $\omega, \omega_j, v \in \mathscr{P}_E(\mathbf{R}^{2d})$, $j = 1, 2$, be such that ω is v-moderate, and v is submultiplicative and even. Then the following is true:*

(1) $M^{\Phi,\Phi}_{(\omega)}(\mathbf{R}^d) = M^{\Phi}_{(\omega)}(\mathbf{R}^d)$, *with equivalent quasi-norms;*

(2) *if $\phi \in \Sigma_1(\mathbf{R}^d) \setminus 0$, then $f \in M^{\Phi,\Psi}_{(\omega)}(\mathbf{R}^d)$, if and only if $\|V_\phi f\|_{L^{\Phi,\Psi}_{(\omega)}}$ is finite. Moreover, $M^{\Phi,\Psi}_{(\omega)}(\mathbf{R}^d)$ is a quasi-Banach space under the respective quasi-norm in (1.13), and different choices of $\phi \in \Sigma_1(\mathbf{R}^d) \setminus 0$ give rise to equivalent norms. If more restrictive Φ and Ψ are Young functions, then $M^{\Phi,\Psi}_{(\omega)}(\mathbf{R}^d)$ is a Banach space, and similar facts hold true with the condition $\phi \in M^1_{(v)}(\mathbf{R}^d) \setminus 0$ in place of $\phi \in \Sigma_1(\mathbf{R}^d) \setminus 0$ at each occurrence.*

(3) *if $\Phi_2 \lesssim \Phi_1$, $\Psi_2 \lesssim \Psi_1$ and $\omega_2 \lesssim \omega_1$, then*

$$\Sigma_1(\mathbf{R}^d) \subseteq M^{\Phi_1,\Psi_1}_{(\omega_1)}(\mathbf{R}^d) \subseteq M^{\Phi_2,\Psi_2}_{(\omega_2)}(\mathbf{R}^d) \subseteq \Sigma_1'(\mathbf{R}^d).$$

Proposition 1.16. *Let Φ, Ψ be Young functions, and let $\omega \in \mathscr{P}_E(\mathbf{R}^{2d})$. Then the following is true:*

(1) *the sesqui-linear form $(\,\cdot\,,\,\cdot\,)_{L^2}$ on $\Sigma_1(\mathbf{R}^d)$ extends to a continuous map from*

$$M^{\Phi,\Psi}_{(\omega)}(\mathbf{R}^d) \times M^{\Phi^*,\Psi^*}_{(1/\omega)}(\mathbf{R}^d)$$

to $\mathbf{C}$. This extension is unique when Φ and Ψ fulfill a local Δ_2-condition. If $\|f\| = \sup |(f,g)_{L^2}|$, where the supremum is taken over all $b \in M^{\Phi^*,\Psi^*}_{(1/\omega)}(\mathbf{R}^d)$ such that $\|b\|_{M^{\Phi^*,\Psi^*}_{(1/\omega)}} \leqslant 1$, then $\|\cdot\|$ and $\|\cdot\|_{M^{p,q}_{(\omega)}}$ are equivalent norms;

(2) *if Φ and Ψ fulfill a local Δ_2-condition, then $\Sigma_1(\mathbf{R}^d)$ is dense in $M^{\Phi,\Psi}_{(\omega)}(\mathbf{R}^d)$, and the dual space of $M^{\Phi,\Psi}_{(\omega)}(\mathbf{R}^d)$ can be identified with $M^{\Phi^*,\Psi^*}_{(1/\omega)}(\mathbf{R}^d)$, through the form $(\,\cdot\,,\,\cdot\,)_{L^2}$. Moreover, $\Sigma_1(\mathbf{R}^d)$ is weakly dense in $M^{\Phi^*,\Psi^*}_{(\omega)}(\mathbf{R}^d)$.*

Proposition 1.15 follows from Theorem 2.4 in [45], Theorems 3.1 and 5.9 in [46], and Proposition 1.14. The details are left for the reader. (See also Theorem 4.2 and other results in [10].) Proposition 1.16 is well-known in the case of modulation spaces (see e.g. Chapters 11 and 12 in [15]). For general Orlicz modulation spaces, Proposition 1.16 essentially follow from Propositions 4.3 and 4.9 in [10] and the fact that similar results hold for Orlicz spaces.

In order to be self-contained we have included a straight-forward proof of Proposition 1.16 in Appendix A, with arguments adapted to the present situation.

Example 1.17. Let Φ be a Young function given by (0.3). For the entropy functional (0.1) it is announced in the introduction that it might be more suitable to investigate such functional in background of the Orlicz modulation space $M^\Phi(\mathbf{R}^d)$ instead of the classical modulation space or Lebesgue space $M^2(\mathbf{R}^d) = L^2(\mathbf{R}^d)$ when Φ is given by (0.3). (See [25,26] and the references therein.)

In fact, in Section 3 we show that

(1) The functional E is continuous on $M^\Phi(\mathbf{R}^d)$, but fails to be continuous on $M^2(\mathbf{R}^d)$. (Cf. Theorem 3.1.)

(2) The space $M^\Phi(\mathbf{R}^d)$ is close to $M^2(\mathbf{R}^d)$ in the sense of the dense embeddings

$$M^p(\mathbf{R}^d) \subseteq M^\Phi(\mathbf{R}^d) \subseteq M^2(\mathbf{R}^d), \qquad p < 2.$$

(Cf. Lemma 3.2.)

1.5. Pseudo-differential operators

Next we recall some basic facts from pseudo-differential calculus (cf. [21]). Let $s \geqslant 1/2$, $a \in \mathcal{S}_s(\mathbf{R}^{2d})$, and let A belong to $\mathbf{M}(d, \mathbf{R})$, the set of all $d \times d$-matrices with entries in $\mathbf{R}$. Then the pseudo-differential operator $\operatorname{Op}_A(a)$, defined by

$$\operatorname{Op}_A(a)f(x) = (2\pi)^{-d} \iint_{\mathbf{R}^{2d}} a(x - A(x-y), \xi) f(y) e^{i\langle x-y,\xi\rangle}\, dy d\xi, \tag{1.16}$$

is a linear and continuous operator on $\mathcal{S}_s(\mathbf{R}^d)$. For $a \in \mathcal{S}_s'(\mathbf{R}^{2d})$ the pseudo-differential operator $\operatorname{Op}_A(a)$ is defined as the continuous operator from $\mathcal{S}_s(\mathbf{R}^d)$ to $\mathcal{S}_s'(\mathbf{R}^d)$ with distribution kernel given by

$$K_{a,A}(x,y) = (2\pi)^{-\frac{d}{2}} (\mathscr{F}_2^{-1} a)(x - A(x-y), x-y). \tag{1.17}$$

Here $\mathscr{F}_2 F$ is the partial Fourier transform of $F(x,y) \in \mathcal{S}_s'(\mathbf{R}^{2d})$ with respect to the variable $y \in \mathbf{R}^d$. This definition generalizes (1.16) and is well-defined, since the mappings

$$\mathscr{F}_2 \quad \text{and} \quad F(x,y) \mapsto F(x - A(x-y), y-x) \tag{1.18}$$

are homeomorphisms on $\mathcal{S}_s'(\mathbf{R}^{2d})$. The map $a \mapsto K_{a,A}$ is hence a homeomorphism on $\mathcal{S}_s'(\mathbf{R}^{2d})$.

If $A = 0$, then $\operatorname{Op}_A(a)$ is the standard or Kohn-Nirenberg representation $a(x, D)$. If instead $A = \frac{1}{2} I_d$, then $\operatorname{Op}_A(a)$ agrees with the Weyl operator or Weyl quantization $\operatorname{Op}^w(a)$. Here I_d is the $d \times d$ identity matrix.

For any $K \in \mathcal{S}_s'(\mathbf{R}^{d_1+d_2})$, let T_K be the linear and continuous mapping from $\mathcal{S}_s(\mathbf{R}^{d_1})$ to $\mathcal{S}_s'(\mathbf{R}^{d_2})$ defined by

$$(T_K f, g)_{L^2(\mathbf{R}^{d_2})} = (K, g \otimes \overline{f})_{L^2(\mathbf{R}^{d_1+d_2})}, \quad f \in \mathcal{S}_s(\mathbf{R}^{d_1}), \quad g \in \mathcal{S}_s(\mathbf{R}^{d_2}). \tag{1.19}$$

It is a well-known consequence of the Schwartz kernel theorem that if $A \in \mathbf{M}(d, \mathbf{R})$, then $K \mapsto T_K$ and $a \mapsto \operatorname{Op}_A(a)$ are bijective mappings from $\mathscr{S}'(\mathbf{R}^{2d})$ to the space of linear and continuous mappings from $\mathscr{S}(\mathbf{R}^d)$ to $\mathscr{S}'(\mathbf{R}^d)$ (cf. e.g. [21]).

Likewise the maps $K \mapsto T_K$ and $a \mapsto \operatorname{Op}_A(a)$ are uniquely extendable to bijective mappings from $\mathcal{S}_s'(\mathbf{R}^{2d})$ to the set of linear and continuous mappings from $\mathcal{S}_s(\mathbf{R}^d)$ to $\mathcal{S}_s'(\mathbf{R}^d)$. In fact, the asserted bijectivity for the map $K \mapsto T_K$ follows from the kernel theorems for topological vector spaces, using the fact that Gelfand-Shilov spaces are inductive or projective limits of certain Hilbert spaces of Hermite series expansions (see [22,28,33]). This kernel theorem corresponds to the Schwartz kernel theorem in the usual distribution theory. The other assertion follows from the fact that the map $a \mapsto K_{a,A}$ is a homeomorphism on $\mathcal{S}_s'(\mathbf{R}^{2d})$.

In particular, for $a_1 \in \mathcal{S}_s'(\mathbf{R}^{2d})$ and $A_1, A_2 \in \mathbf{M}(d, \mathbf{R})$, there is a unique $a_2 \in \mathcal{S}_s'(\mathbf{R}^{2d})$ such that $\operatorname{Op}_{A_1}(a_1) = \operatorname{Op}_{A_2}(a_2)$. The relationship between a_1 and a_2 is given by

$$\operatorname{Op}_{A_1}(a_1) = \operatorname{Op}_{A_2}(a_2) \quad \Leftrightarrow \quad e^{i\langle A_1 D_\xi, D_x \rangle} a_1(x,\xi) = e^{i\langle A_2 D_\xi, D_x \rangle} a_2(x,\xi). \tag{1.20}$$

(Cf. [21].) Here the expressions on the right in (1.20) makes sense, since the Fourier transform of $e^{i\langle A_j D_\xi, D_x \rangle} a_j(x,\xi)$ is equal to $e^{i\langle (A_j x, \xi \rangle} \widehat{a}_j(\xi, x)$, and that the map, which takes $b(\xi, x)$ into $e^{i\langle Ax, \xi \rangle} b(\xi, x)$, is continuous on $\mathcal{S}_s'(\mathbf{R}^{2d})$ (see e.g. [47]).

The operator $e^{i\langle A D_\xi, D_x \rangle}$ is essential when transferring Wigner distributions to each others. In what follows we have the following continuity result for $e^{i\langle A D_\xi, D_x \rangle}$ when acting on Orlicz modulation spaces.

Proposition 1.18. *Let Φ_1 and Φ_2 be quasi-Young functions, $A \in \mathbf{M}(d, \mathbf{R})$, $s_1 \geqslant \frac{1}{2}$, $s_2 > \frac{1}{2}$, $\omega_0 \in \mathscr{P}_E(\mathbf{R}^{4d})$, where $\mathscr{P}_E(\mathbf{R}^{4d})$ is the set of all moderate functions on $\mathbf{R}^{4d}$, and let*

$$\omega_A(x, \xi, \eta, y) = \omega_0(x + Ay, \xi + A^*\eta, \eta, y).$$

Then the following is true:

(1) *$e^{i\langle A D_\xi, D_x \rangle}$ is continuous from $\mathscr{S}(\mathbf{R}^{2d})$ to $\mathscr{S}'(\mathbf{R}^{2d})$, and restricts to homeomorphisms on*

$$\mathcal{S}_{s_1}(\mathbf{R}^{2d}), \quad \Sigma_{s_2}(\mathbf{R}^{2d}) \quad \text{and} \quad \mathscr{S}(\mathbf{R}^{2d}),$$

and is uniquely extendable to homeomorphisms on

$$\mathcal{S}_{s_1}'(\mathbf{R}^{2d}), \quad \Sigma_{s_2}'(\mathbf{R}^{2d}) \quad \text{and} \quad \mathscr{S}'(\mathbf{R}^{2d});$$

(2) *$e^{i\langle A D_\xi, D_x \rangle}$ from $\Sigma_1'(\mathbf{R}^{2d})$ to $\Sigma_1'(\mathbf{R}^{2d})$ restricts to a homeomorphism from $M^{\Phi_1,\Phi_2}_{(\omega_0)}(\mathbf{R}^{2d})$ to $M^{\Phi_1,\Phi_2}_{(\omega_A)}(\mathbf{R}^{2d})$, and*

$$\|e^{i\langle AD_\xi,D_x\rangle}a\|_{M^{\Phi_1,\Phi_2}_{(\omega_A)}} \asymp \|a\|_{M^{\Phi_1,\Phi_2}_{(\omega_0)}}, \qquad a\in M^{\Phi_1,\Phi_2}_{(\omega_0)}(\mathbf{R}^{2d}). \tag{1.21}$$

Proof. We shall follow the proof of [43, Proposition 2.8].

The assertion (1) and its proof can be found in e.g. [1,47]. Let $T = e^{i\langle AD_\xi,D_x\rangle}$. By (2.12) in [43] we have

$$|(V_{T\phi}(Tf))(x,\xi,\eta,y)| = |(V_\phi f)(x+Ay,\xi+A^*\eta,\eta,y)|,$$

when $\phi\in\Sigma_1(\mathbf{R}^{2d})$. By multiplying with ω_A, applying the L^Φ norm on the x and ξ variables and then taking $x+Ay$ and $\xi+A^*\eta$ as new variables of integration give

$$\|(V_{T\phi}(Tf))(\,\cdot\,,\eta,y)\omega_A(\,\cdot\,,\eta,y)\|_{L^\Phi(\mathbf{R}^{2d})} = \|(V_\phi f)(\,\cdot\,,\eta,y)\omega_0(\,\cdot\,,\eta,y)\|_{L^\Phi(\mathbf{R}^{2d})}.$$

The relation (1.21) now follows by applying the L^Ψ norm with respect to the y and η variables on the latter identity. This gives the result. □

For future references we observe the relationship

$$\begin{gathered}|(V_\phi K_{a,A})(x,y,\xi,-\eta)| = |(V_\psi a)(x-A(x-y),A^*\xi+(I-A^*)\eta,\xi-\eta,y-x)|,\\ \phi(x,y) = (\mathscr{F}_2\psi)(x-A(x-y),x-y)\end{gathered} \tag{1.22}$$

between symbols and kernels for pseudo-differential operators, which follows by straightforward applications of Fourier inversion formula (see also the proof of Proposition 2.5 in [43]).

1.6. Wigner distributions

Next we recall general classes of Wigner distributions parameterized by matrices. Let $A\in\mathbf{M}(d,\mathbf{R})$. Then the A-Wigner distribution (or cross-A-Wigner distribution) of $f_1,f_2\in\mathscr{S}(\mathbf{R}^d)$, is defined by the formula

$$W^A_{f_1,f_2}(x,\xi)\equiv\mathscr{F}\big(f_1(x+A\,\cdot\,)\overline{f_2(x+(A-I)\,\cdot\,)}\big)(\xi), \tag{1.23}$$

which takes the form

$$W^A_{f_1,f_2}(x,\xi) = (2\pi)^{-\frac{d}{2}}\int_{\mathbf{R}^d} f_1(x+Ay)\overline{f_2(x+(A-I)y)}e^{-i\langle y,\xi\rangle}\,dy,$$

when $f_1,f_2\in\mathcal{S}_s(\mathbf{R}^d)$. We set $W_{f_1,f_2} = W^A_{f_1,f_2}$ when $A=\frac12 I_d$ and I_d is the $d\times d$, in which case we get the classical (standard) Wigner distribution.

The definition of Wigner distributions is extendable in various ways, which the following result indicates. For the proof we refer to [43] and its references.

Proposition 1.19. *Let $A \in \mathbf{M}(d,\mathbf{R})$, $s \geqslant \frac{1}{2}$ and T from $\mathscr{S}(\mathbf{R}^d) \times \mathscr{S}(\mathbf{R}^d)$ to $\mathscr{S}'(\mathbf{R}^{2d})$ be the map given by $(f_1,f_2) \mapsto W^A_{f_1,f_2}$. Then the following is true:*

(1) *T restricts to a continuous map from $\mathcal{S}_s(\mathbf{R}^d) \times \mathcal{S}_s(\mathbf{R}^d)$ to $\mathcal{S}_s(\mathbf{R}^{2d})$, and is uniquely extendable to a continuous map from $\mathcal{S}_s'(\mathbf{R}^d) \times \mathcal{S}_s'(\mathbf{R}^d)$ to $\mathcal{S}_s'(\mathbf{R}^{2d})$;*
(2) *T restricts to a continuous map from $\mathcal{S}_s'(\mathbf{R}^d) \times \mathcal{S}_s(\mathbf{R}^d)$ or from $\mathcal{S}_s(\mathbf{R}^d) \times \mathcal{S}_s'(\mathbf{R}^d)$ to $\mathcal{S}_s'(\mathbf{R}^{2d}) \cap C^\infty(\mathbf{R}^{2d})$.*

The same holds true with $\mathscr{S}$ in place of $\mathcal{S}_s$ at each occurrence. If in addition $s > \frac{1}{2}$, then the same holds true with Σ_s in place of $\mathcal{S}_s$ at each occurrence.

The following result shows that Wigner distributions with different matrices can be carried over to each others. We refer to Subsection 1.1 in [43] for the proof (see e.g. (1.10) in [43]).

Lemma 1.20. *Let $A_1, A_2 \in \mathbf{M}(d,\mathbf{R})$ and $f_1, f_2 \in \mathcal{S}_{1/2}'(\mathbf{R}^d)$. Then*

$$e^{i\langle A_1 D_\xi, D_x\rangle} W^{A_1}_{f_1,f_2} = e^{i\langle A_2 D_\xi, D_x\rangle} W^{A_2}_{f_1,f_2}.$$

Finally we recall the links

$$(\operatorname{Op}_A(a)f,g)_{L^2(\mathbf{R}^d)} = (2\pi)^{-\frac{d}{2}}(a, W^A_{g,f})_{L^2(\mathbf{R}^{2d})}, \qquad a \in \mathscr{S}'(\mathbf{R}^{2d}),\ f,g \in \mathscr{S}(\mathbf{R}^d), \tag{1.24}$$

and

$$\operatorname{Op}_A(W^A_{f_1,f_2})f(x) = (2\pi)^{-\frac{d}{2}}(f,f_2)_{L^2} f_1(x), \qquad f_1,f_2 \in \mathscr{S}'(\mathbf{R}^d),\ f \in \mathscr{S}(\mathbf{R}^d), \tag{1.25}$$

between pseudo-differential operators and Wigner distributions, which follows by straight-forward computations. Similar facts hold true with $\mathcal{S}_s$ or Σ_s in place of $\mathscr{S}$ at each occurrence.

Remark 1.21. We observe that the definition of Wigner distributions can be extended in various ways. For example, metaplectic Wigner distributions are given in [2].

2. Continuity for pseudo-differential operators when acting on Orlicz modulation spaces

In this section we deduce continuity properties for Wigner distributions when acting on Orlicz modulation spaces. Thereafter we apply such results to obtain continuity properties for pseudo-differential operators with symbols in Orlicz modulation spaces when acting on other Orlicz modulation spaces.

We need the following result on Hölder's inequality for Orlicz spaces, and refer to [29, III.3.3] for the proof (cf. Theorem 7 in [29, III.3.3]). Here we let $S(\mu)$ be the set of all simple and (μ-)measurable functions on the measurable space (E,μ).

Proposition 2.1 *(Hölder inequality). Let (E, μ) be a measurable space, and Φ_j, $j = 0, 1, 2$ be Young's functions such that*

$$\Phi_0(t_1 t_2) \leqslant \Phi_1(t_1) + \Phi_2(t_2) \quad \text{or} \quad \Phi_0^{-1}(s) \geqslant \Phi_1^{-1}(s) \cdot \Phi_2^{-1}(s), \qquad s, t_1, t_2 \geqslant 0.$$

Then the map $(f_1, f_2) \mapsto f_1 \cdot f_2$ from $S(\mu) \times S(\mu)$ to $S(\mu)$ extends uniquely to a continuous map from $L^{\Phi_1}(\mu) \times L^{\Phi_2}(\mu)$ to $L^{\Phi_0}(\mu)$, and

$$\|f_1 \cdot f_2\|_{L^{\Phi_0}} \leqslant 2\|f_1\|_{L^{\Phi_1}} \|f_2\|_{L^{\Phi_2}}, \qquad f_j \in L^{\Phi_j}(\mu),\ j = 1, 2. \tag{2.1}$$

2.1. Continuity for Wigner distributions and short-time Fourier transforms

In some situations, we need some more restrictions on our (quasi-)Young functions.

Definition 2.2. Let $\Phi : [0, \infty] \to [0, \infty]$ and $p \in (0, \infty)$. Then Φ is called *p-steered* if one of the following conditions are fulfilled:

(1) $\limsup_{t \to 0+} \frac{\Phi(t)}{t^p} = \infty$;
(2) $t \mapsto \Phi(t^{\frac{1}{p}})$ is equal to a Young function near origin.

The first main result of the section is the following theorem which concerns continuity property for Wigner distributions acting on Orlicz modulation spaces. Here the involved weight functions should satisfy

$$\omega(x, \xi, \eta, y) \lesssim \omega_1(x - Ay, \xi + (I - A^*)\eta)\omega_2(x + (I - A)y, \xi - A^*\eta). \tag{2.2}$$

Theorem 2.3. *Let $A \in \mathbf{M}(d, \mathbf{R})$, $p, q \in [1, \infty]$ be such that $p \leqslant q$, and let $\Phi_j, \Psi_j : [0, \infty] \to [0, \infty]$, $j = 1, 2$, be such that the following is true:*

- *if $p = \infty$, then Φ_j and Ψ_j are Young functions;*
- *if $p < \infty$, then Φ_j and Ψ_j are p-steered Young functions which fulfill a local Δ_2-condition, and for some $r > 0$, it holds*

$$\Phi_1(t), \Phi_2(t) \gtrsim t^q, \quad \Psi_1(t), \Psi_2(t) \gtrsim t^q, \quad t \in [0, r], \tag{2.3}$$

and

$$\Phi_1^{-\&}(s)\Phi_2^{-\&}(s) \lesssim s^{\frac{1}{p}+\frac{1}{q}}, \quad \Psi_1^{-\&}(t)\Psi_2^{-\&}(s) \lesssim s^{\frac{1}{p}+\frac{1}{q}}, \quad s \in [0, r]. \tag{2.4}$$

Also let $\omega \in \mathscr{P}_E(\mathbf{R}^{4d})$ and $\omega_1, \omega_2 \in \mathscr{P}_E(\mathbf{R}^{2d})$ be such that (2.2) holds. Then the map $(f_1, f_2) \mapsto W^A_{f_1, f_2}$ from $\Sigma_1'(\mathbf{R}^d) \times \Sigma_1'(\mathbf{R}^d)$ to $\Sigma_1'(\mathbf{R}^{2d})$ restricts to a continuous map from $M^{\Phi_1, \Psi_1}_{(\omega_1)}(\mathbf{R}^d) \times M^{\Phi_2, \Psi_2}_{(\omega_2)}(\mathbf{R}^d)$ to $M^{p,q}_{(\omega)}(\mathbf{R}^{2d})$, and

$$\|W^A_{f_1, f_2}\|_{M^{p,q}_{(\omega)}} \lesssim \|f_1\|_{M^{\Phi_1, \Psi_1}_{(\omega_1)}} \|f_2\|_{M^{\Phi_2, \Psi_2}_{(\omega_2)}}, \quad f_j \in M^{\Phi_j, \Psi_j}_{(\omega_j)}(\mathbf{R}^d),\ j = 1, 2. \tag{2.5}$$

Remark 2.4. Suppose that $p, q \in (0, \infty]$ satisfy $p \leqslant q$, $\Phi : [0, \infty] \to [0, \infty]$ and let $\Phi_{[q]}(t) = \Phi(t^{\frac{1}{q}})$. (Observe that such expressions appear behind the conditions in Theorem 2.3.) Then the following is true:

(1) if $q < \infty$ and $\Phi_{[q]}$ is a Young function, then $\Phi(t) \lesssim t^q$ near origin and $\Phi_{[p]}$ is a Young function;
(2) if $\Phi_{[q]}$ is a Young function which satisfies the Δ_2-condition, then $\Phi_{[p]}$ is a Young function which satisfies the Δ_2-condition;
(3) Suppose that $p < \infty$, and $\Phi_j, \Psi_j : [0, \infty] \to [0, \infty]$ are such that $t \mapsto \Phi_j(t^{\frac{1}{p}})$ and $t \mapsto \Psi_j(t^{\frac{1}{p}})$ are Young functions, $j = 1, 2$, and that (2.4) holds. Then (2.3) holds. In particular, for some $r_1, r_2 > 0$ it holds

$$t^q \lesssim \Phi_j(t), \Psi_j(t) \lesssim t^p, \quad t \in [0, r_1], \tag{2.6}$$

or equivalently,

$$s^{\frac{1}{p}} \lesssim \Phi_j^{-\&}(s), \Psi_j^{-\&}(s) \lesssim s^{\frac{1}{q}}, \quad s \in [0, r_2]. \tag{2.7}$$

For $q = \infty$ we observe the following consequence of Theorem 2.3, for extreme choices of Φ_j or Ψ_j, for some $j = 1, 2$.

Corollary 2.5. *Let $A \in \mathbf{M}(d, \mathbf{R})$, $p \in [1, \infty)$ be such that $p \leqslant q$, and let Φ_j and Ψ_j be such that $t \mapsto \Phi_j(t^{\frac{1}{p}})$ and $t \mapsto \Psi_j(t^{\frac{1}{p}})$ are Young functions which fulfill the Δ_2-condition, $j = 1, 2$, and such that*

$$\Phi_1^{-1}(s)\Phi_2^{-1}(s) \leqslant s^{\frac{1}{p}} \quad \text{and} \quad \Psi_1^{-1}(s)\Psi_2^{-1}(s) \leqslant s^{\frac{1}{p}}. \tag{2.8}$$

Also let $\omega \in \mathscr{P}_E(\mathbf{R}^{4d})$ and $\omega_1, \omega_2 \in \mathscr{P}_E(\mathbf{R}^{2d})$ be such that (2.2) holds. Then the map $(f_1, f_2) \mapsto W^A_{f_1,f_2}$ from $\Sigma_1'(\mathbf{R}^d) \times \Sigma_1'(\mathbf{R}^d)$ to $\Sigma_1'(\mathbf{R}^{2d})$ restricts to a continuous map from

$$\begin{aligned} &M^{p,\Psi_1}_{(\omega_1)}(\mathbf{R}^d) \times M^{\infty,\Psi_2}_{(\omega_2)}(\mathbf{R}^d), \qquad && M^{\infty,\Psi_1}_{(\omega_1)}(\mathbf{R}^d) \times M^{p,\Psi_2}_{(\omega_2)}(\mathbf{R}^d),\\ &M^{\Phi_1,p}_{(\omega_1)}(\mathbf{R}^d) \times M^{\Phi_2,\infty}_{(\omega_2)}(\mathbf{R}^d) \quad \text{or} && M^{\Phi_1,\infty}_{(\omega_1)}(\mathbf{R}^d) \times M^{\Phi_2,p}_{(\omega_2)}(\mathbf{R}^d), \end{aligned}$$

to $M^{p,\infty}_{(\omega)}(\mathbf{R}^{2d})$.

We need the following Young type results for Orlicz spaces for the proof of Theorem 2.3 (see Theorem 9 in [29, III.3.3]).

Lemma 2.6. *Let $\Phi_j, j = 0, 1, 2$, be Young functions which fulfill the Δ_2-condition and such that*

$$\Phi_1^{-1}(s) \cdot \Phi_2^{-1}(s) \leqslant s\Phi_0^{-1}(s), \qquad s \geqslant 0. \tag{2.9}$$

Then the convolution map $(f_1, f_2) \mapsto f_1 * f_2$ *from* $L^{\Phi_1}(\mathbf{R}^d) \times L^{\Phi_2}(\mathbf{R}^d)$ *to* $L^{\Phi_0}(\mathbf{R}^d)$ *is continuous and*

$$\|f_1 * f_2\|_{L^{\Phi_0}} \leqslant 2\|f_1\|_{L^{\Phi_1}}\|f_2\|_{L^{\Phi_2}}, \quad f_j \in L^{\Phi_j}(\mathbf{R}^d),\ j = 1, 2. \tag{2.10}$$

Proof of Theorem 2.3. First suppose $p = \infty$. Then it follows from e.g. [43, Proposition 2.4] that $(f_1, f_2) \mapsto W^A_{f_1,f_2}$ is continuous from $M^\infty_{(\omega_1)}(\mathbf{R}^d) \times M^\infty_{(\omega_1)}(\mathbf{R}^d)$ to $M^\infty_{(\omega)}(\mathbf{R}^{2d})$. The result now follows from the facts that $q \geqslant p = \infty$ and $M^{\Phi_j,\Psi_j}_{(\omega_j)}(\mathbf{R}^d)$ are continuously embedded in $M^\infty_{(\omega_j)}(\mathbf{R}^d)$, $j = 1, 2$, in view of Proposition 1.14.

It remains to consider the case when $p < \infty$. Since $M^{\Phi_j,\Psi_j}_{(\omega_j)}(\mathbf{R}^d)$ only depends on Φ_j and Ψ_j near origin, in view of [46, Proposition 5.9], we may replace these Young functions with new ones such that (2.4) holds for all $s \in [0, \infty]$ (see Definition 1.11). Furthermore, by (1.10) and Lemma 2.6 in [43], it follows that we may reduce ourselves to the case when $A = \frac{1}{2}I$, giving the standard (cross-)Wigner distribution.

First we consider the case when $t \mapsto \Phi_j(t^{\frac{1}{p}})$ and $t \mapsto \Psi_j(t^{\frac{1}{p}})$ are Young functions. As a first step on this we also assume that $q < \infty$, giving that $p < \infty$.

Let

$$F = W_{f_1,f_2} \quad \text{and} \quad \psi = W_{\phi_1,\phi_2}.$$

Then [43, Lemma 2.6] gives

$$|V_\psi F(x, \xi, \eta, y)| = |V_{\phi_1} f_1(x - \tfrac{1}{2}y, \xi + \tfrac{1}{2}\eta)| \cdot |V_{\phi_2} f_2(x + \tfrac{1}{2}y, \xi - \tfrac{1}{2}\eta)|.$$

Hence, if

$$G(x, \xi, \eta, y) = |V_\psi F(x, \xi, \eta, y)\omega(x, \xi, \eta, y)|,$$
$$G_1(x, \xi) = |V_{\phi_1} f_1(-x, \xi)\omega_1(-x, \xi)|$$

and

$$G_2(x, \xi) = |V_{\phi_2} f_2(x, -\xi)\omega_2(x, -\xi)|,$$

then it follows from the assumptions that

$$0 \leqslant G(x, \xi, \eta, y) \lesssim G_1(\tfrac{1}{2}y - x, \tfrac{1}{2}\eta + \xi) \cdot G_2(\tfrac{1}{2}y + x, \tfrac{1}{2}\eta - \xi). \tag{2.11}$$

By first applying the L^p-norm on the x and ξ variables, and then the L^q norm on the y variable we obtain

$$\|H_0(\eta, \cdot\,)\|_{L^q(\mathbf{R}^d)} \lesssim R(\eta), \qquad H_0(\eta, y) = \|G(\,\cdot\,, \eta, y)\|_{L^p(\mathbf{R}^{2d})}, \tag{2.12}$$

where

$$R(\eta) \equiv \left(\int \left(\iint G_1(\tfrac{1}{2}y - x, \tfrac{1}{2}\eta + \xi)^p G_2(\tfrac{1}{2}y + x, \tfrac{1}{2}\eta - \xi)^p \, dx d\xi \right)^{\frac{q}{p}} \, dy \right)^{\frac{1}{q}}.$$

We need to estimate $R(\eta)$ in suitable ways.

By taking $x + \frac{1}{2}y$, $\xi - \frac{1}{2}\eta$ and y as new variables of integrations, and using Minkowski's inequality, we obtain

$$\begin{aligned} R(\eta) &= \left(\int \left(\iint G_1(y - x, \eta - \xi)^p G_2(x, \xi)^p \, dx d\xi \right)^{\frac{q}{p}} \, dy \right)^{\frac{1}{q}} \\ &\leqslant \left(\int \left(\int \left(\int G_1(y - x, \eta - \xi)^p G_2(x, \xi)^p \, dx \right)^{\frac{q}{p}} \, dy \right)^{\frac{p}{q}} \, d\xi \right)^{\frac{1}{p}} \\ &= \left(\int \left(\| G_1(\,\cdot\,, \eta - \xi)^p * G_2(\,\cdot\,, \xi)^p \|_{L^{\frac{q}{p}}} \right) \, d\xi \right)^{\frac{1}{p}}. \end{aligned} \tag{2.13}$$

Now recall that if

$$\widetilde{\Phi}_j(t) = \Phi_j(t^{\frac{1}{p}}) \quad \text{and} \quad \widetilde{\Psi}_j(t) = \Psi_j(t^{\frac{1}{p}}), \quad j = 1, 2,$$

then, since Φ_j and Ψ_j for $j = 1, 2$ are p-steered, $\widetilde{\Phi}_j(t)$ and $\widetilde{\Psi}_j(t)$ are Young functions such that

$$\widetilde{\Phi}_1^{-1}(s)\widetilde{\Phi}_2^{-1}(s) \leqslant s^{\frac{p}{q}+1} \quad \text{and} \quad \widetilde{\Psi}_1^{-1}(s)\widetilde{\Psi}_2^{-1}(s) \leqslant s^{\frac{p}{q}+1}.$$

Hence Lemma 2.6 gives

$$\| G_1(\,\cdot\,, \eta - \xi)^p * G_2(\,\cdot\,, \xi)^p \|_{L^{\frac{q}{p}}} \lesssim H_1(\eta - \xi)^p H_2(\xi)^p, \tag{2.14}$$

where

$$H_j(\xi) \equiv \| G_j(\,\cdot\,, \xi)^p \|_{L^{\tilde{\Phi}_1}}^{\frac{1}{p}}, \quad j = 1, 2. \tag{2.15}$$

By combining (2.13) with (2.14) we obtain

$$R(\eta) \lesssim ((H_1^p * H_2^p)(\eta))^{\frac{1}{p}}.$$

By applying the L^q norm, and using that $\|F\|_{M^{p,q}_{(\omega)}} \lesssim \|R\|_{L^q}$, due to (2.12) and Lemma 2.6 we get

$$\begin{aligned} \|F\|_{M^{p,q}_{(\omega)}} &\lesssim \|R\|_{L^q} \lesssim \|H_1^p * H_2^p\|_{L^{q/p}}^{\frac{1}{p}} \lesssim \left(\|H_1^p\|_{L^{\widetilde{\Psi}_1}} \|H_2^p\|_{L^{\widetilde{\Psi}_2}} \right)^{\frac{1}{p}} \\ &= \left(\|G_1^p\|_{L^{\tilde{\Phi}_1, \widetilde{\Psi}_1}} \|G_2^p\|_{L^{\tilde{\Phi}_2, \widetilde{\Psi}_2}} \right)^{\frac{1}{p}} = \|G_1\|_{L^{\Phi_1, \Psi_1}} \|G_2\|_{L^{\Phi_2, \Psi_2}} \\ &\asymp \|f_1\|_{M^{\Phi_1, \Psi_1}_{(\omega_1)}} \|f_2\|_{M^{\Phi_2, \Psi_2}_{(\omega_2)}}, \end{aligned}$$

giving the result in the case $q < \infty$.

Next suppose that $q = \infty$. By first applying the L^p-norm on the x and ξ variables, and then the L^∞ norm on the y variable in (2.11) we obtain

$$\sup_{y \in \mathbf{R}^d} \left(\|G(\,\cdot\,, \eta, y)\|_{L^p(\mathbf{R}^{2d})} \right) \lesssim R(\eta), \tag{2.12$'$}$$

where $R(\eta)$ is now redefined as

$$R(\eta) \equiv \sup_{y \in \mathbf{R}^d} \left(\iint G_1(\tfrac{1}{2}y - x, \xi + \tfrac{1}{2}\eta)^p G_2(x + \tfrac{1}{2}y, \tfrac{1}{2}\eta - \xi)^p \, dx d\xi \right)^{\frac{1}{p}}.$$

An application of (2.13) gives

$$R(\eta) \leqslant \left(\int \|G_1(\,\cdot\,, \eta - \xi)^p * G_2(\,\cdot\,, \xi)^p\|_{L^\infty} \, d\xi \right)^{\frac{1}{p}}. \tag{2.13$'$}$$

We have

$$\widetilde{\Psi}_1^{-1}(s) \widetilde{\Psi}_2^{-1}(s) \leqslant s,$$

and by Hölder's inequality we obtain

$$\|G_1(\,\cdot\,, \eta - \xi)^p * G_2(\,\cdot\,, \xi)^p\|_{L^\infty} \lesssim H_1(\eta - \xi)^p H_2(\xi)^p, \tag{2.14$'$}$$

where H_j are the same as in (2.15).

A combination of (2.13)$'$ and (2.14)$'$ gives

$$R(\eta) \lesssim (H_1^p * H_2^p)(\eta)^{\frac{1}{p}}.$$

By applying the L^∞ norm, and using that $\|F\|_{M^{p,\infty}_{(\omega)}} \lesssim \|R\|_{L^\infty}$, due to (2.12)$'$, we obtain

$$\begin{aligned} \|F\|_{M^{p,\infty}_{(\omega)}} &\lesssim \|R\|_{L^\infty} \lesssim \|H_1^p * H_2^p\|_{L^\infty} \\ &\lesssim \left(\|H_1^p\|_{L^{\widetilde{\Psi}_1}} \|H_2^p\|_{L^{\widetilde{\Psi}_2}} \right)^{\frac{1}{p}} = \left(\|G_1^p\|_{L^{\tilde{\Phi}_1, \widetilde{\Psi}_1}} \|G_2^p\|_{L^{\tilde{\Phi}_2, \widetilde{\Psi}_2}} \right)^{\frac{1}{p}} \\ &= \|G_1\|_{L^{\Phi_1, \Psi_1}} \|G_2\|_{L^{\Phi_2, \Psi_2}} \asymp \|f_1\|_{M^{\Phi_1, \Psi_1}_{(\omega_1)}} \|f_2\|_{M^{\Phi_2, \Psi_2}_{(\omega_2)}}, \end{aligned}$$

giving the result in the case when $t \mapsto \Phi_j(t^{\frac{1}{p}})$ and $t \mapsto \Psi_j(t^{\frac{1}{p}})$ are Young functions.

It remains to consider the case when $t \mapsto \Phi_j(t^{\frac{1}{p}})$ or $t \mapsto \Psi_k(t^{\frac{1}{p}})$ are not Young functions for some $j = 1, 2$ and some $k = 1, 2$. We shall here mainly use similar arguments as in the proof of Theorem 1.1 in [3]. Then

$$\Phi_j^{-\&}(s) \lesssim s^{\frac{1}{p}} \quad \text{or} \quad \Psi_k^{-\&}(s) \lesssim s^{\frac{1}{p}}$$

near origin, for some $j = 1, 2$ and some $k = 1, 2$. First suppose that $\Phi_1^{-\&}(s) \lesssim s^{\frac{1}{p}}$, and that $\Psi_j(t^{\frac{1}{p}})$ are Young functions, $j = 1, 2$. Then

$$\Phi_1^{-\&}(s)s^{\frac{1}{q}} \lesssim s^{\frac{1}{p}+\frac{1}{q}},$$

and

$$M^{\Phi_1,\Psi_1}_{(\omega_1)}(\mathbf{R}^d) \subseteq M^{p,\Psi_1}_{(\omega_1)}(\mathbf{R}^d) \quad \text{and} \quad M^{\Phi_2,\Psi_2}_{(\omega_2)}(\mathbf{R}^d) \subseteq M^{q,\Psi_2}_{(\omega_2)}(\mathbf{R}^d), \tag{2.16}$$

where the last embedding follows from (2.3). By the previous part of the proof we have that $(f_1, f_2) \mapsto W^A_{f_1,f_2}$ is continuous from $M^{p,\Psi_1}_{(\omega_1)}(\mathbf{R}^d) \times M^{q,\Psi_2}_{(\omega_2)}(\mathbf{R}^d)$ to $M^{p,q}_{(\omega)}(\mathbf{R}^{2d})$. The result now follows in this case by combining the latter continuity property with the embeddings in (2.16).

By similar arguments, the same conclusion holds true if instead

$$\Phi_2^{-\&}(s) \lesssim s^{\frac{1}{p}}, \quad \Psi_1^{-\&}(s) \lesssim s^{\frac{1}{p}} \quad \text{or} \quad \Psi_2^{-\&}(s) \lesssim s^{\frac{1}{p}}.$$

The details are left for the reader.

Finally suppose that

$$\Phi_j^{-\&}(s) \lesssim s^{\frac{1}{p}}, \qquad \text{and} \quad \Psi_k^{-\&}(s) \lesssim s^{\frac{1}{p}},$$

for some $j = 1, 2$ and some $k = 1, 2$. Then the previous arguments lead to

$$M^{\Phi_j,\Psi_j}_{(\omega_j)}(\mathbf{R}^d) \subseteq M^{p_j,q_j}_{(\omega_j)}(\mathbf{R}^d), \tag{2.17}$$

for some $p_j, q_j \in \{p, q\}$, $j = 1, 2$, and such that $p_1 \neq p_2$ and $q_1 = q_2$. Again we have that $(f_1, f_2) \mapsto W^A_{f_1,f_2}$ is continuous from $M^{p_1,q_1}_{(\omega_1)}(\mathbf{R}^d) \times M^{p_2,q_2}_{(\omega_2)}(\mathbf{R}^d)$ to $M^{p,q}_{(\omega)}(\mathbf{R}^{2d})$. The asserted continuity now follows from (2.17), and the result follows. □

Beside the estimates for Wigner distributions on Orlicz modulation spaces in Theorem 2.3, we also have the following result on estimates for the short-time Fourier transform. The result generalizes [5, Proposition 3.3] which involves non-weighted modulation spaces as well as [39, Proposition 2.2] which involves weighted modulation spaces. (See Definition 1.4 for broader spectrum of Orlicz spaces.)

Theorem 2.7. *Let $f_1, f_2 \in \mathcal{S}'_{1/2}(\mathbf{R}^d)$, Φ_j and Ψ_j be quasi-Young functions $j = 1, 2$, $\omega_0 \in \mathscr{P}_E(\mathbf{R}^{4d})$ and $\omega_1, \omega_2 \in \mathscr{P}_E(\mathbf{R}^{2d})$. Also let $\phi_1, \phi_2 \in \mathcal{S}_{1/2}(\mathbf{R}^d)$, and let $\phi = V_{\phi_1}\phi_2$. Then the following is true:*

(1) *if*

$$\omega_0(x, \xi, \eta, -y) \leqslant C\omega_1(y - x, \eta)\omega_2(y, \xi + \eta), \qquad x, y, \xi, \eta \in \mathbf{R}^d, \tag{2.18}$$

for some constant $C > 0$*, then*

$$\|V_\phi(V_{f_1}f_2)\|_{L^{\Phi_1,\Phi_2,\Psi_1,\Psi_2}_{(\omega_0)}} \leqslant C\|V_{\phi_1}f_1\|_{L^{\Phi_1,\Psi_1}_{(\omega_1)}}\|V_{\phi_2}f_2\|_{L^{\Psi_2,\Phi_2}_{*,(\omega_2)}}; \tag{2.19}$$

(2) *if*

$$\omega_1(y-x,\eta)\omega_2(y,\xi+\eta) \leqslant C\omega_0(x,\xi,\eta,-y), \qquad x,y,\xi,\eta \in \mathbf{R}^d, \tag{2.20}$$

for some constant C*, then*

$$\|V_{\phi_1}f_1\|_{L^{\Phi_1,\Psi_1}_{(\omega_1)}}\|V_{\phi_2}f_2\|_{L^{\Psi_2,\Phi_2}_{*,(\omega_2)}} \leqslant C\|V_\phi(V_{f_1}f_2)\|_{L^{\Phi_1,\Phi_2,\Psi_1,\Psi_2}_{(\omega_0)}}; \tag{2.21}$$

(3) *if* (2.18) *and* (2.20) *hold for some constant* C*, then* $f_1 \in M^{\Phi_1,\Psi_1}_{(\omega_1)}(\mathbf{R}^d)$ *and* $f_2 \in W^{\Psi_2,\Phi_2}_{(\omega_2)}(\mathbf{R}^d)$*, if and only if* $V_{f_1}f_2 \in M^{\Phi_1,\Phi_2,\Psi_1,\Psi_2}_{(\omega_0)}(\mathbf{R}^{2d})$*, and*

$$\|V_{f_1}f_2\|_{M^{\Phi_1,\Phi_2,\Psi_1,\Psi_2}_{(\omega_0)}} \asymp \|f_1\|_{M^{\Phi_1,\Psi_1}_{(\omega_1)}}\|f_2\|_{W^{\Psi_2,\Phi_2}_{(\omega_2)}}. \tag{2.22}$$

Proof. We shall mainly follow the proofs of Proposition 3.3 in [5] and Proposition 2.2 in [39].

It suffices to prove (1) and (2), and then we only prove (1), since (2) follows by similar arguments.

By Fourier's inversion formula we have

$$|V_{\phi_1}f_1(-x-y,\eta)V_{\phi_2}f_2(-y,\xi+\eta)| = |V_\phi(V_{f_1}f_2)(x,\xi,\eta,y)|$$

(cf. e.g. [12,15,37,43]). Hence, if

$$F_1(x,\xi) = |V_{\phi_1}f_1(x,\xi)|\omega_1(x,\xi) \quad \text{and} \quad F_2(x,\xi) = V_{\phi_2}f_2(x,\xi)\omega_2(x,\xi),$$

then

$$\begin{aligned}\|V_\phi(V_{f_1}f_2)(\,\cdot\,,\xi,\eta,y)\omega_0(\,\cdot\,,\xi,\eta,y)\|_{L^{\Phi_1}(\mathbf{R}^d)} &\leqslant C\|F_1(-y-\,\cdot\,,\eta)F_2(-y,\xi+\eta)\|_{L^{\Phi_1}(\mathbf{R}^d)}\\ &= C\|F_1(\,\cdot\,,\eta)\|_{L^{\Phi_1}(\mathbf{R}^d)}F_2(-y,\xi+\eta).\end{aligned}$$

By applying the L^{Φ_2} quasi-norm with respect to the ξ-variables we obtain

$$\begin{aligned}&\|V_\phi(V_{f_1}f_2)(\,\cdot\,,\eta,y)\omega_0(\,\cdot\,,\eta,y)\|_{L^{\Phi_1,\Phi_2}(\mathbf{R}^{2d})}\\ &\leqslant C\|F_1(\,\cdot\,,\eta)\|_{L^{\Phi_1}(\mathbf{R}^d)}\|F_2(-y,\,\cdot+\eta)\|_{L^{\Phi_2}(\mathbf{R}^d)} = C\|F_1(\,\cdot\,,\eta)\|_{L^{\Phi_1}(\mathbf{R}^d)}\|F_2(-y,\,\cdot\,)\|_{L^{\Phi_2}(\mathbf{R}^d)}.\end{aligned}$$

The result now follows by first applying the L^{Ψ_1} quasi-norm on the η-variables, and then the L^{Ψ_2} quasi-norm on the y-variables. □

Corollary 2.8. *Let $f_1, f_2 \in \mathcal{S}_{1/2}'(\mathbf{R}^d)$, Φ and Ψ be quasi-Young functions and let $\omega_0 \in \mathscr{P}_E(\mathbf{R}^{4d})$ and $\omega_1, \omega_2 \in \mathscr{P}_E(\mathbf{R}^{2d})$ be such that*

$$\omega_0(x, \xi, \eta, -y) \asymp \omega_1(y - x, \eta)\omega_2(y, \xi + \eta).$$

Then $f_1 \in M^{\Phi,\Psi}_{(\omega_1)}(\mathbf{R}^d)$ and $f_2 \in W^{\Psi,\Phi}_{(\omega_2)}(\mathbf{R}^d)$, if and only if $V_{f_1} f_2 \in M^{\Phi,\Psi}_{(\omega_0)}(\mathbf{R}^{2d})$, and

$$\|V_{f_1} f_2\|_{M^{\Phi,\Psi}_{(\omega_0)}} \asymp \|f_1\|_{M^{\Phi,\Psi}_{(\omega_1)}} \|f_2\|_{W^{\Psi,\Phi}_{(\omega_2)}}.$$

2.2. Continuity for pseudo-differential operators when acting on Orlicz modulation spaces

Next we apply the previous results to deduce continuity for pseudo-differential operators with symbols in modulation spaces which act on Orlicz modulation spaces. The involved weight functions should satisfy

$$\frac{\omega_2(x, \xi)}{\omega_1(y, \eta)} \lesssim \omega_0(x - A(x - y), A^*\xi + (I - A^*)\eta, \xi - \eta, y - x). \tag{2.23}$$

The following result extend [4, Theorem 5.1].

Theorem 2.9. *Let $A \in \mathbf{M}(d, \mathbf{R})$, $p, q \in [1, \infty]$ be such that $q \leqslant p$, and let $\Phi_j, \Psi_j : [0, \infty] \to [0, \infty]$, $j = 1, 2$, be such that the following is true:*

- *if $p = 1$, then Φ_j and Ψ_j are Young functions;*
- *if $p > 1$, then Φ_j and Ψ_j are p'-steered Young functions which fulfill a local Δ_2-condition, and for some $r > 0$, it holds*

$$\Phi_1(t), \Phi_2(t) \gtrsim t^{q'} \quad \Psi_1(t), \Psi_2(t) \gtrsim t^{q'}, \quad t \in [0, r], \tag{2.24}$$

and

$$\Phi_1^{-\&}(s)\Phi_2^{-\&}(s) \lesssim s^{\frac{1}{p'} + \frac{1}{q'}}, \quad \Psi_1^{-\&}(s)\Psi_2^{-\&}(s) \lesssim s^{\frac{1}{p'} + \frac{1}{q'}}, \quad s \in [0, r]. \tag{2.25}$$

Also let $\omega_0 \in \mathscr{P}_E(\mathbf{R}^{2d} \oplus \mathbf{R}^{2d})$ and $\omega_1, \omega_2 \in \mathscr{P}_E(\mathbf{R}^{2d})$ satisfy (2.23). If $a \in M^{p,q}_{(\omega_0)}(\mathbf{R}^{2d})$, then $\operatorname{Op}_A(a)$ from $\mathcal{S}_{1/2}(\mathbf{R}^d)$ to $\mathcal{S}_{1/2}'(\mathbf{R}^d)$ extends uniquely to a continuous map from $M^{\Phi_1,\Psi_1}_{(\omega_1)}(\mathbf{R}^d)$ to $M^{\Phi_2^*,\Psi_2^*}_{(\omega_2)}(\mathbf{R}^d)$, and

$$\|\operatorname{Op}_A(a)\|_{M^{\Phi_1,\Psi_1}_{(\omega_1)} \to M^{\Phi_2^*,\Psi_2^*}_{(\omega_2)}} \lesssim \|a\|_{M^{p,q}_{(\omega_0)}}. \tag{2.26}$$

Moreover, if in addition a belongs to the closure of $\mathcal{S}_{1/2}$ under the $M^{p,q}_{(\omega_0)}$ norm, then $\operatorname{Op}_A(a) : M^{\Phi_1,\Psi_1}_{(\omega_1)}(\mathbf{R}^d) \to M^{\Phi_2^*,\Psi_2^*}_{(\omega_2)}(\mathbf{R}^d)$ is compact.

Proof. First suppose that $p < \infty$. Then $q < \infty$, and it follows that $\mathcal{S}_{1/2}(\mathbf{R}^{2d})$ is dense in $M^{p,q}_{(\omega_0)}(\mathbf{R}^{2d})$. Let $f, g \in \mathcal{S}'_{1/2}(\mathbf{R}^d)$ and $a \in \mathcal{S}_{1/2}(\mathbf{R}^{2d})$. Then (1.24) and Theorem 2.3 gives

$$
\begin{aligned}
&|(\operatorname{Op}_A(a)f, g)| \asymp |(a, W^A_{g,f})| \\
&\qquad \lesssim \|a\|_{M^{p,q}_{(\omega_0)}} \|W^A_{g,f}\|_{M^{p',q'}_{(1/\omega_0)}} \lesssim \|a\|_{M^{p,q}_{(\omega_0)}} \|f\|_{M^{\Phi_1,\Psi_1}_{(\omega_1)}} \|g\|_{M^{\Phi_2,\Psi_2}_{(1/\omega_2)}},
\end{aligned}
\tag{2.27}
$$

and by duality it follows that $\operatorname{Op}_A(a)$ from $\mathcal{S}'_{1/2}(\mathbf{R}^d)$ to $\mathcal{S}_{1/2}(\mathbf{R}^d)$ restricts to a continuous map from $M^{\Phi_1,\Psi_1}_{(\omega_1)}(\mathbf{R}^d)$ to $M^{\Phi_2^*,\Psi_2^*}_{(\omega_2)}(\mathbf{R}^d)$, and that (2.27) gives (2.26). The result now follows in this case by (2.27) and the fact that $\mathcal{S}_{1/2}(\mathbf{R}^{2d})$ is dense in $M^{p,q}_{(\omega_0)}(\mathbf{R}^{2d})$.

Next suppose that $p = \infty$ and $q < \infty$. If $a \in \mathcal{S}_{1/2}(\mathbf{R}^{2d})$, then (2.27) implies that (2.26) holds in this case as well. By Hahn-Banach's theorem it follows that the definition of $\operatorname{Op}_A(a)$ is extendable to any $a \in M^{p,q}_{(\omega_0)}(\mathbf{R}^{2d})$, and that (2.26) still holds. The uniqueness of the extension now follows from the fact that $\mathcal{S}_{1/2}(\mathbf{R}^{2d})$ is dense in $M^{p,q}_{(\omega_0)}(\mathbf{R}^{2d})$ with respect to the narrow convergence, when $q < \infty$ (see [40]).

Finally, if $p = q = \infty$, then (2.4) implies that

$$
\Phi_j(t) \asymp t \quad \text{and} \quad \Psi_j(t) \asymp t, \quad j = 1, 2,
$$

giving that $M^{\Phi_1,\Psi_1}_{(\omega_1)}(\mathbf{R}^d) = M^{1,1}_{(\omega_1)}(\mathbf{R}^d)$ and $M^{\Phi_2^*,\Psi_2^*}_{(\omega_2)}(\mathbf{R}^d) = M^{\infty,\infty}_{(\omega_2)}(\mathbf{R}^d)$. The result now follows by choosing $p = q = \infty$ in [43, Theorem 2.2]. □

As a special case we obtain the following extension of Proposition 0.1 in the introduction. The details are left for the reader.

Proposition 0.1′. *Let $A \in \mathbf{M}(d, \mathbf{R})$, $p, q \in [1, \infty]$ be such that $q \leqslant p$ and $p > 1$, $\omega_0 \in \mathscr{P}_E(\mathbf{R}^{2d} \oplus \mathbf{R}^{2d})$ and $\omega_1, \omega_2 \in \mathscr{P}_E(\mathbf{R}^{2d})$ satisfy (2.23). Also let $\Phi_j, \Psi_j : [0, \infty] \to [0, \infty]$, $j = 1, 2$, be such that $t \mapsto \Phi_j(t^{\frac{1}{p'}})$ and $t \mapsto \Psi_j(t^{\frac{1}{p'}})$ are Young functions which fulfill the Δ_2-condition, and*

$$
\Phi_1(t), \Phi_2(t) \gtrsim t^{q'} \quad \Psi_1(t), \Psi_2(t) \gtrsim t^{q'}, \quad t \geqslant 0,
$$

and

$$
\Phi_1^{-1}(s)\Phi_2^{-1}(s) \lesssim s^{\frac{1}{p'}+\frac{1}{q'}}, \quad \Psi_1^{-1}(s)\Psi_2^{-1}(s) \lesssim s^{\frac{1}{p'}+\frac{1}{q'}}, \quad s \geqslant 0.
$$

If $a \in M^{p,q}_{(\omega_0)}(\mathbf{R}^{2d})$, then $\operatorname{Op}_A(a)$ is continuous from $M^{\Phi_1,\Psi_1}_{(\omega_1)}(\mathbf{R}^d)$ to $M^{\Phi_2^*,\Psi_2^*}_{(\omega_2)}(\mathbf{R}^d)$.

By similar type of duality arguments, using Theorem 2.7 instead of Theorem 2.3, we obtain the following extension of [43, Theorem 2.1]. Here we observe that (2.23) takes the form

$$
\frac{\omega_2(x, \xi)}{\omega_1(y, \eta)} \lesssim \omega_0(x, \eta, \xi - \eta, y - x) \tag{2.23$'$}
$$

when $A = 0$.

Theorem 2.10. *Let Φ and Ψ be Young functions which satisfy local Δ_2-condition, and let $\omega_0 \in \mathscr{P}_E(\mathbf{R}^{4d})$ and $\omega_1, \omega_2 \in \mathscr{P}_E(\mathbf{R}^{2d})$ be such that (2.23)$'$ holds. Also let $a \in W^{\Psi,\Phi}_{(\omega_0)}(\mathbf{R}^{2d})$. Then the definition of $\operatorname{Op}_0(a)$ from $\mathcal{S}_{1/2}(\mathbf{R}^d)$ to $\mathcal{S}'(\mathbf{R}^d)$ extends uniquely to a continuous map from $M^{\Phi^*,\Psi^*}_{(\omega_1)}(\mathbf{R}^d)$ to $W^{\Psi,\Phi}_{(\omega_2)}(\mathbf{R}^d)$, and*

$$\|\operatorname{Op}_0(a)f\|_{W^{\Psi,\Phi}_{(\omega_2)}} \lesssim \|a\|_{W^{\Psi,\Phi}_{(\omega_0)}} \|f\|_{M^{\Phi^*,\Psi^*}_{(\omega_1)}}, \qquad a \in W^{\Psi,\Phi}_{(\omega_0)}(\mathbf{R}^{2d}),\ f \in M^{\Phi^*,\Psi^*}_{(\omega_1)}(\mathbf{R}^d). \tag{2.28}$$

Proof. We shall follow the proof of Theorem 2.1 in [43]. We may assume that equality holds in (2.23)$'$. We start to prove the result in the case $\Phi^*(t) > 0$ and $\Psi^*(t) > 0$ when $t > 0$. Then we may replace Φ and Ψ such that the Orlicz modulation spaces are the same and Φ and Ψ satisfy (global) Δ_2-conditions.

Let

$$\omega(x,\xi,\eta,y) = \omega_0(-y,\eta,\xi,-x)^{-1},$$

$a \in W^{\Psi,\Phi}_{(\omega_0)}(\mathbf{R}^{2d})$ and $f, g \in \mathcal{S}_{1/2}(\mathbf{R}^d)$. Then $\operatorname{Op}_0(a)f$ makes sense as an element in $\mathcal{S}'_{1/2}(\mathbf{R}^d)$.

By Theorem 2.7 we get

$$\|V_f g\|_{M^{\Phi^*,\Psi^*}_{(\omega)}} \lesssim \|f\|_{M^{\Phi^*,\Psi^*}_{(\omega_1)}} \|g\|_{W^{\Psi^*,\Phi^*}_{(1/\omega_2)}}. \tag{2.29}$$

Furthermore, if T is the torsion operator defined by $TF(x,\xi) = F(\xi,-x)$ when $F \in \mathcal{S}'_{1/2}(\mathbf{R}^{2d})$, then it follows by Fourier's inversion formula that

$$(V_\phi(T\widehat{a}))(x,\xi,\eta,y) = e^{-i(\langle x,\eta\rangle + \langle y,\xi\rangle)}(V_{T\widehat{\phi}}a)(-y,\eta,\xi,-x).$$

This gives

$$|(V_\phi(T\widehat{a}))(x,\xi,\eta,y)\omega(x,\xi,\eta,y)^{-1}| = |(V_{\phi_1}a)(-y,\eta,\xi,-x)\omega_0(-y,\eta,\xi,-x)|,$$

when $\phi_1 = T\widehat{\phi}$. Hence, by applying the $L^{\Phi,\Psi}$ norm we obtain

$$\|T\widehat{a}\|_{M^{\Phi,\Psi}_{(1/\omega)}} = \|a\|_{W^{\Psi,\Phi}_{(\omega_0)}}.$$

It now follows from (2.29) that

$$\begin{aligned} |(\operatorname{Op}_0(a)f, g)| &= (2\pi)^{-d/2}|(T\widehat{a}, V_f g)| \\ &\lesssim \|T\widehat{a}\|_{M^{\Phi,\Psi}_{(1/\omega)}} \|V_f g\|_{M^{\Phi^*,\Psi^*}_{(\omega)}} \lesssim \|a\|_{W^{\Psi,\Phi}_{(\omega_0)}} \|f\|_{M^{\Phi^*,\Psi^*}_{(\omega_1)}} \|g\|_{W^{\Psi^*,\Phi^*}_{(1/\omega_2)}}. \end{aligned} \tag{2.30}$$

The result now follows by the facts that $\mathcal{S}_{1/2}(\mathbf{R}^d)$ is dense in $M^{\Phi^*,\Psi^*}_{(\omega_1)}(\mathbf{R}^d)$, and that the dual of $W^{\Psi^*,\Phi^*}_{(1/\omega_2)}$ is $W^{\Phi,\Psi}_{(\omega_2)}$ when Φ^* and Ψ^* satisfies the Δ_2-condition.

If instead $\Phi^*(t) = 0$ and $\Psi(t) > 0$, or $\Phi(t) > 0$ and $\Psi^*(t) = 0$, when $t > 0$ is close to origin, then let $f \in M^{\Phi^*,\Psi^*}_{(\omega_1)}$ and $a \in \mathcal{S}_{1/2}(\mathbf{R}^{2d})$. Then $\operatorname{Op}_0(a)f$ makes sense as an element in $\mathcal{S}_{1/2}(\mathbf{R}^d)$, and from the first part of the proof it follows that (2.30) still holds. The result now follows by duality and the fact that $\mathscr{S}(\mathbf{R}^{2d})$ is dense in $W^{\Psi,\Phi}_{(\omega_0)}(\mathbf{R}^{2d})$, since it follows from the assumptions that Φ and Ψ fulfill the Δ_2-condition.

It remains to consider the case when $\Phi(t) = \Psi^*(t) = 0$ and the case when $\Phi^*(t) = \Psi(t) = 0$ when $t > 0$ is near origin. In this case, we have

$$W^{\Psi,\Phi} = W^{1,\infty} \quad \text{and} \quad M^{\Phi^*,\Psi^*} = M^{1,\infty},$$

or

$$W^{\Psi,\Phi} = W^{\infty,1} \quad \text{and} \quad M^{\Phi^*,\Psi^*} = M^{\infty,1}.$$

The result then follows by letting $p = q' = \infty$ or $p = q' = 1$ in [36, Theorem 3.9] or in the proof of [43, Theorem 2.1]. $\square$

Example 2.11. Let $A \in \mathbf{M}(d,\mathbf{R})$, $p > 2$, $a \in M^{p,p'}(\mathbf{R}^{2d})$ and Φ be a Young function which fullfils (0.8). That is, we let $q = p'$ in our results. Then it follows that Φ fullfils a local Δ_2-condition,

$$\Phi(t) \gtrsim t^p = t^{q'} \quad \text{and} \quad \Phi^{-1}(s)^2 \lesssim s = s^{\frac{1}{p'}+\frac{1}{q'}}.$$

Hence the hypothesis in Propositions 0.1 and 0.1′ (as well as in Theorem 2.9) are fulfilled with

$$\Phi_1 = \Phi_2 = \Psi_1 = \Psi_2 = \Phi.$$

It now follows from any of these results that

$$\operatorname{Op}_A(a) : M^{\Phi}(\mathbf{R}^d) \to M^{\Phi}(\mathbf{R}^d) \tag{2.31}$$

is continuous.

We also observe that if instead a belongs to $M^2(\mathbf{R}^{2d})$, which is near $M^{p,p'}(\mathbf{R}^{2d})$ when $p > 2$ is closed to 2, then the map (2.31) may be discontinuous (cf. Remark 3.8 in the end of the next section).

3. Continuity for entropy functionals in background of Orlicz modulation spaces

In this section we show that the entropy functional in (0.3) is continuous on the modulation spaces $M^p(\mathbf{R}^d)$, $1 \leqslant p < 2$, and the Orlicz modulation space $M^{\Phi}(\mathbf{R}^d)$ with

$\Phi(t) = -t^2 \log t$ near origin. For completeness we also give a proof of that the same functional is discontinuous on $M^2(\mathbf{R}^d) = L^2(\mathbf{R}^d)$. (Cf. Theorem 3.1.) In order to reach such properties we need to prove some preparing results which might be of independent interest. For example we deduce estimates for entropy functionals when changing window functions (see Lemma 3.6).

We observe that the entropy functional (0.3) can be written as

$$\begin{aligned} E(f) &= E_\phi(f) \\ &\equiv -\iint_{\mathbf{R}^{2d}} |V_\phi f(x,\xi)|^2 \log |V_\phi f(x,\xi)|^2 \, dx d\xi + \|\phi\|_{L^2}^2 \|f\|_{L^2}^2 \log(\|\phi\|_{L^2}^2 \|f\|_{L^2}^2), \end{aligned} \tag{0.1$'$}$$

by using Moyal's identity

$$\|V_\phi f\|_{L^2} = \|f\|_{L^2} \|\phi\|_{L^2} \tag{3.1}$$

(see e.g. [15]). In particular, if $\|f\|_{L^2} = \|\phi\|_{L^2} = 1$ which is a common condition in the applications, the entropy of f becomes

$$E(f) = E_\phi(f) = -\iint_{\mathbf{R}^{2d}} |V_\phi f(x,\xi)|^2 \log |V_\phi f(x,\xi)|^2 \, dx d\xi, \quad \|f\|_{L^2} = \|\phi\|_{L^2} = 1 \tag{0.1$''$}$$

(see e.g. [23,24]). For general $f, \phi \in L^2(\mathbf{R}^d)$ we observe that the entropy possess homogeneity properties of the form

$$E_{\lambda\phi}(f) = E_\phi(\lambda f) = |\lambda|^2 E_\phi(f), \qquad f, \phi \in L^2(\mathbf{R}^d),\ \lambda \in \mathbf{C}. \tag{3.2}$$

In fact, Moyal's identity gives

$$\begin{aligned} E_{\lambda\phi}(f) &= E_\phi(\lambda f) \\ &= -\iint_{\mathbf{R}^{2d}} |\lambda|^2 |V_\phi f(x,\xi)|^2 \log(|\lambda||V_\phi f(x,\xi)|)^2 \, dx d\xi \\ &+ |\lambda|^2 \|\phi\|_{L^2}^2 \|f\|_{L^2}^2 \log(|\lambda|^2 \|\phi\|_{L^2}^2 \|f\|_{L^2}^2) \\ &= |\lambda|^2 \left(-\iint_{\mathbf{R}^{2d}} |V_\phi f(x,\xi)|^2 \log |V_\phi f(x,\xi)|^2 \, dx d\xi + \|\phi\|_{L^2}^2 \|f\|_{L^2}^2 \log(\|\phi\|_{L^2}^2 \|f\|_{L^2}^2) \right) \\ &+ (\log |\lambda|^2)(\|\phi\|_{L^2}^2 \|f\|_{L^2}^2 - \|V_\phi f\|_{L^2}^2) \\ &= |\lambda|^2 E_\phi(f). \end{aligned}$$

In order to discuss continuity for the entropy functional, we restrict ourself and assume that the window functions belong to the subspace $M^1(\mathbf{R}^d)$ of $L^2(\mathbf{R}^d)$. The main result of the section is the following.

Theorem 3.1. *Let Φ be a Young function which satisfies (0.3), $\phi \in M^1(\mathbf{R}^d) \setminus 0$ and let E_ϕ be as in (0.1). Then the following is true:*

(1) *E_ϕ is continuous on $M^p(\mathbf{R}^d)$ and on $M^\Phi(\mathbf{R}^d)$, $0 < p < 2$;*
(2) *E_ϕ is discontinuous on $M^p(\mathbf{R}^d)$, $2 \leqslant p \leqslant \infty$.*

We need some preparations for the proof of Theorem 3.1. First we observe that $M^\Phi(\mathbf{R}^d)$ is in some sense close to $M^2(\mathbf{R}^d)$.

Lemma 3.2. *Let Φ be a Young function which satisfies (0.3). Then*

$$M^p(\mathbf{R}^d) \subseteq M^\Phi(\mathbf{R}^d) \subseteq M^2(\mathbf{R}^d), \qquad p < 2, \tag{3.3}$$

with continuous and dense inclusions, and

$$\lim_{p \to 2-} \|f\|_{M^p} = \|f\|_{M^2}, \quad \text{when} \quad f \in M^{p_0}(\mathbf{R}^d), \text{ for some } p_0 < 2. \tag{3.4}$$

For the limit in (3.4) it is understood that the same window function is used in the modulation space norms.

Proof. By Proposition 1.14 it follows that $M^\Phi(\mathbf{R}^d)$ is independent of the choice of Φ outside the interval $[0, e^{-\frac{2}{3}}]$. It is therefore no restriction to assume that Φ is given by

$$\Phi(t) = \begin{cases} -t^2 \log t, & t \in [0, e^{-\frac{2}{3}}], \\ \frac{1}{3} e^{-\frac{2}{3}}(t + e^{-\frac{2}{3}}), & t \in (e^{-\frac{2}{3}}, \infty), \\ \infty, & t = \infty, \end{cases} \tag{0.1$'$}$$

which is obviously a Young function.

By Remark 1.10 and the limits

$$\lim_{t \to 0+} \frac{t^2}{\Phi(t)} = \lim_{t \to 0+} -\frac{t^2}{t^2 \log t} = 0 \quad \text{and} \quad \lim_{t \to 0+} \frac{t^2}{\Phi(t)} = \lim_{t \to 0+} -\frac{t^p}{t^2 \log t} = \infty,$$

when $p < 2$, it follows from Proposition 1.14 that the inclusions in (3.3) holds and are continuous. Since $M^p(\mathbf{R}^d)$ is dense in $M^2(\mathbf{R}^d)$, it also follows that $M^\Phi(\mathbf{R}^d)$ is dense in $M^2(\mathbf{R}^d)$.

The limit in (3.4) follows by straight-forward computations in measure theory (cf. e.g. the exercise part of Chapter 3 in [32]). □

Remark 3.3. Let Φ be a Young function which satisfies (0.3). A consequence of Theorem 3.1, Lemma 3.2 and the open mapping theorem is that $M^\Phi(\mathbf{R}^d) \subsetneq M^2(\mathbf{R}^d)$.

Next we show that E_ϕ is well-defined and finite on $M^\Phi(\mathbf{R}^d)$.

Lemma 3.4. *Let Φ be a Young function which satisfies (0.3), $f,\phi \in M^2(\mathbf{R}^d)$. Then the following is true:*

(1) $|V_\phi f(x,\xi)|^2 \log_+ |V_\phi f(x,\xi)| \in L^1(\mathbf{R}^{2d})$ *and*

$$\iint\limits_{\mathbf{R}^{2d}} |V_\phi f(x,\xi)|^2 \log |V_\phi f(x,\xi)|\, dx d\xi \in [-\infty,\infty);$$

(2) *if in addition $f \in M^\Phi(\mathbf{R}^d)$ and $\phi \in M^1(\mathbf{R}^d)$, then*

$$\iint\limits_{\mathbf{R}^{2d}} |V_\phi f(x,\xi)|^2 \big| \log |V_\phi f(x,\xi)| \big|\, dx d\xi < \infty.$$

Proof. The assertion (1) follows from the fact that $V_\phi f \in L^2(\mathbf{R}^{2d}) \cap L^\infty(\mathbf{R}^{2d})$, in view of Moyal's identity and the embedding $M^2(\mathbf{R}^d) \subseteq M^\infty(\mathbf{R}^d)$, ensured by Proposition 1.15 (3).

Since (2) is obviously true when f or ϕ are identically equal to zero, we may assume that $f \in M^\Phi(\mathbf{R}^d) \setminus 0$ and $\phi \in M^1(\mathbf{R}^d) \setminus 0$. Let ϕ be chosen as the window function in the modulation space norms, $C > 1$ be a fixed constant and for every $f \in M^\Phi(\mathbf{R}^d)$, choose the number $\lambda = \lambda_f$ such that

$$\|f\|_{M^\Phi} < \lambda < C\|f\|_{M^\Phi}.$$

For conveniency we also let $F = V_\phi f$,

$$\Omega_1 = \{\, (x,\xi) \in \mathbf{R}^{2d}\, ;\, |F(x,\xi)| \leqslant \lambda e^{-\frac{2}{3}}\, \},$$
$$\Omega_2 = \{\, (x,\xi) \in \mathbf{R}^{2d}\, ;\, \lambda e^{-\frac{2}{3}} \leqslant |F(x,\xi)| \leqslant \lambda\, \}$$

and

$$\Omega_3 = \{\, (x,\xi) \in \mathbf{R}^{2d}\, ;\, |F(x,\xi)| \geqslant \lambda\, \}.$$

Then

$$\left| - \iint\limits_{\mathbf{R}^{2d}} |F(x,\xi)|^2 \log |F(x,\xi)|\, dx d\xi \right| \leqslant \sum_{k=1}^{4} J_k(f),$$

where

$$J_k(f) = \lambda^2 \left| \iint\limits_{\Omega_k} \left(\frac{|F(x,\xi)|}{\lambda} \right)^2 \log \left(\frac{|F(x,\xi)|}{\lambda} \right) dx d\xi \right|, \quad k = 1, 2, 3,$$

and

$$J_4(f) = |\log \lambda| \cdot \|f\|_{M^2}^2,$$

and the result follows if we prove

$$J_k(f) < \infty, \qquad k = 1, 2, 3, 4. \tag{3.5}$$

By the definition of Φ and the Orlicz modulation space norm, we have

$$J_1(f) \leqslant \lambda^2 \leqslant C^2 \|f\|_{M^\Phi}^2 < \infty,$$

which shows that (3.5) holds for $k = 1$.

In order to prove (3.5) for $k = 2$ and $k = 3$ we recall that $\|f\|_{M^\infty} \lesssim \|f\|_{M^2} \lesssim \|f\|_{M^\Phi}$, which implies that $|F(x,\xi)| \lesssim \|f\|_{M^\Phi}$. On the other hand, $|F(x,\xi)| \gtrsim \|f\|_{M^\Phi}$ when $(x,\xi) \in \complement\Omega_1$. A combination of these relations yields $|F(x,\xi)| \asymp \|f\|_{M^\Phi}$ when $(x,\xi) \in \complement\Omega_1$, which implies that the logarithm in the integral expression of $J_k(f)$ is bounded when $k = 2$ or $k = 3$. This gives

$$0 \leqslant J_k(f) \lesssim \iint\limits_{\complement\Omega_1} |F(x,\xi)|^2 \, dx d\xi \leqslant \|f\|_{M^2}^2 \lesssim \|f\|_{M^\Phi}^2, \quad k = 2, 3,$$

and (3.5) follows in the cases $k = 2$ and $k = 3$.

Finally, for $J_4(f)$ we have

$$0 \leqslant J_3(f) \leqslant |\log \lambda| \|f\|_{M^2}^2 \lesssim |\log \lambda| \|f\|_{M^\Phi}^2 < \infty,$$

and the result follows. □

The next lemma gives an essential step when deducing the asserted continuity in Theorem 3.1.

Lemma 3.5. *Let Φ and ϕ be the same as in Lemma 3.4. Then*

$$M^\Phi(\mathbf{R}^d) \ni f \mapsto \iint\limits_{\mathbf{R}^{2d}} |V_\phi f(x,\xi)|^2 \big| \log |V_\phi f(x,\xi)| \big| \, dx d\xi \tag{3.6}$$

is continuous near origin.

Proof. The result follows if we prove

$$-\iint_{\mathbf{R}^{2d}} |V_\phi f(x,\xi)|^2 \big| \log |V_\phi f(x,\xi)| \big| \, dx d\xi \to 0 \quad \text{as} \quad \|f\|_{M^\Phi} \to 0, \; f \in M^\Phi(\mathbf{R}^d). \tag{3.7}$$

Let ϕ, C, λ and $J_k(f)$ be the same as in the proof of Lemma 3.4. Then

$$\left| \iint_{\mathbf{R}^{2d}} |V_\phi f(x,\xi)|^2 \left| \log |V_\phi f(x,\xi)| \right| \, dx d\xi \right| \leqslant \sum_{k=1}^{4} J_k(f),$$

and (3.7) follows if we prove

$$J_k(f) \to 0 \quad \text{as} \quad \|f\|_{M^\Phi} \to 0, \; f \in M^\Phi(\mathbf{R}^d), \qquad k = 1,2,3,4. \tag{3.8}$$

By the definition of Φ and the Orlicz modulation space norm, we have

$$0 \leqslant J_1(f) \leqslant \lambda^2 \leqslant C^2 \|f\|^2_{M^\Phi} \to 0 \quad \text{as} \quad \|f\|_{M^\Phi} \to 0,$$

which shows that (3.8) holds for $k = 1$.

In order to prove (3.8) for $k = 2$ and $k = 3$ we recall from the proof of Lemma 3.4 that the logarithm in (3.8) is bounded when $k = 2$ or $k = 3$. This gives

$$0 \leqslant J_k(f) \lesssim \iint_{\Omega_k} |V_\phi f(x,\xi)|^2 \, dx d\xi \leqslant \|f\|^2_{M^2} \lesssim \|f\|^2_{M^\Phi} \to 0$$

as $\|f\|_{M^\Phi} \to 0$, and (3.8) follows in the cases $k = 2$ and $k = 3$.

For $J_4(f)$ with $\|f\|_{M^\Phi} \leqslant 1$ we have

$$0 \leqslant J_4(f) \leqslant \|f\|^2_{M^2} |\log \|f\|_{M^\Phi}| \lesssim \|f\|^2_{M^\Phi} |\log \|f\|_{M^\Phi}| \to 0$$

as $\|f\|_{M^\Phi} \to 0$. This gives (3.8) in the case $k = 4$, and (3.8) follows, and we have proved that the map (3.6) is continuous at origin. □

The next lemma concerns estimates for E_ϕ in transitions between different window functions ϕ. The result is needed in the proof of discontinuity of E_ϕ on $M^2(\mathbf{R}^d)$.

Lemma 3.6. *Let Φ be a Young function and $\phi, \psi \in M^1(\mathbf{R}^d) \setminus 0$. Then there is a constant C which only depends on ϕ and ψ such that*

$$E_\phi(f) \leqslant C(E_\psi(f) + \|f\|^2_{L^2}), \qquad f \in M^2(\mathbf{R}^d). \tag{3.9}$$

Proof. Let $f \in M^2(\mathbf{R}^d)$,

$$F_1 = |V_\phi f|, \quad F_2 = |V_\psi f| \quad \text{and} \quad H = |V_\phi \psi|.$$

We recall that $\|H\|_{L^1} \asymp \|\phi\|_{M^1}\|\psi\|_{M^1} < \infty$ in view of [15, Proposition 12.1.2]. Since

$$\mathscr{S}(\mathbf{R}^d) \ni \phi \mapsto E_\phi(f_0), \quad M^2(\mathbf{R}^d) \ni f \mapsto E_{\phi_0}(f) \quad \text{and} \quad L^2(\mathbf{R}^d) \ni f \mapsto \|f\|_{L^2}^2 \tag{3.10}$$

are positively homogeneous of order 2 for every fixed $\phi_0 \in \mathscr{S}(\mathbf{R}^d)\backslash 0$ and $f_0 \in M^2(\mathbf{R}^d)\backslash 0$, we reduce ourselves to the case when $\|H\|_{L^1} = 1$.

Since $M^2(\mathbf{R}^d)$ is continuously embedded in $M^\infty(\mathbf{R}^d)$, there is a constant $C_1 > 0$ such that $\|V_\psi f\|_{L^\infty} \leqslant C_1\|V_\psi f\|_{L^2}$ for every $f \in M^2(\mathbf{R}^d)$. First assume that ϕ, ψ and f are chosen such that $\|H\|_{L^1} = 1$ and

$$\|F_2\|_{L^2} = \|V_\psi f\|_{L^2} = e^{-\frac{2}{3}}/C_1. \tag{3.11}$$

Then $0 \leqslant F_2(x,\xi) \leqslant \|V_\psi f\|_{L^\infty} \leqslant e^{-\frac{2}{3}}$. By [15, Lemma 11.3.3] we obtain

$$\begin{aligned} 0 \leqslant F_1(x,\xi) \leqslant (F_2 * H)(x,\xi) &= \iint\limits_{\mathbf{R}^{2d}} F_2(x-y,\xi-\eta)\, d\mu(y,\eta) \\ &\leqslant \|F_2\|_{L^\infty}\|H\|_{L^1} \leqslant e^{-\frac{2}{3}}. \end{aligned} \tag{3.12}$$

Here μ is the positive measure given by $d\mu(y,\eta) = H(y,\eta)\, dyd\eta$, giving that

$$\int\limits_{\mathbf{R}^{2d}} d\mu = \|H\|_{L^1} = 1.$$

Since $t \mapsto \varphi(t) = -t^2 \log t$ is increasing and convex on $[0, e^{-\frac{2}{3}}]$, it follows from (3.12) and Jensen's inequality that

$$\begin{aligned} E_{0,\phi}(f) \equiv -\iint\limits_{\mathbf{R}^{2d}} F_1(x,\xi)^2 \log F_1(x,\xi)\, dxd\xi &= \iint\limits_{\mathbf{R}^{2d}} \varphi(F_1(x,\xi))\, dxd\xi \\ &\leqslant \iint\limits_{\mathbf{R}^{2d}} \varphi\left(\iint\limits_{\mathbf{R}^{2d}} F_2(x-y,\xi-\eta)\, d\mu(y,\eta)\right) dxd\xi \\ &\leqslant \iint\limits_{\mathbf{R}^{2d}} \left(\iint\limits_{\mathbf{R}^{2d}} \varphi\left(F_2(x-y,\xi-\eta)\right)\, d\mu(y,\eta)\right) dxd\xi \\ &= \|H\|_{L^1} \iint\limits_{\mathbf{R}^{2d}} \varphi\left(F_2(x,\xi)\right)\, dxd\xi = E_{0,\psi}(f). \end{aligned} \tag{3.13}$$

Now choose $C_0 \geqslant \max(e, e^{\frac{5}{3}} C_1)$ such that

$$\|V_\phi f\|_{L^2} \leqslant C_0 \|V_\psi f\|_{L^2} \quad \text{and} \quad \|V_\psi f\|_{L^2} \leqslant C_0 \|f\|_{L^2},$$

for every $f \in M^2(\mathbf{R}^d)$, which is possible because

$$f \mapsto \|V_\phi f\|_{L^2} \quad \text{and} \quad f \mapsto \|V_\psi f\|_{L^2}$$

are two equivalent norms for $M^2(\mathbf{R}^d) = L^2(\mathbf{R}^d)$. Then $\log C_0 \geqslant 1$. A combination of (3.11) and (3.13) gives $\log(C_0 \|F_2\|_{L^2}) \geqslant 1$ and

$$\begin{aligned}
E_\phi(f) &= 2E_{0,\phi}(f) + 2\|F_1\|_{L^2}^2 \log \|F_1\|_{L^2} \\
&\leqslant 2E_{0,\psi}(f) + 2(C_0\|F_2\|_{L^2})^2 \log(C_0\|F_2\|_{L^2}) \\
&= 2E_{0,\psi}(f) + 2C_0^2 \|F_2\|_{L^2}^2 \log \|F_2\|_{L^2} + 2(C_0^2 \log C_0)\|F_2\|_{L^2}^2 \\
&\leqslant C(E_{0,\psi}(f) + \|F_2\|_{L^2}^2) \log \|F_2\|_{L^2} + \|f\|_{L^2}^2),
\end{aligned}$$

when $C = 2C_0^4 \log C_0$. Hence (3.9) follows under the additional condition (3.11). The estimate (3.9) now follows for general $f \in M^2(\mathbf{R}^d)$ by the homogeneity of the mappings in (3.10), and the result follows. □

Proof of Theorem 3.1. We choose Φ as in (0.1)$'$. First we prove the continuity for E_ϕ on $M^\Phi(\mathbf{R}^d)$ at origin.

By (3.3) it follows that $\|f\|_{M^2} \leqslant C\|f\|_{M^\Phi}$, for some constant $C \geqslant 1$ which is independent of $f \in M^\Phi(\mathbf{R}^d)$. Hence, for $f \in M^\Phi(\mathbf{R}^d)$ with $\|f\|_{M^\Phi}$ being small enough we obtain

$$\begin{aligned}
\left| \|f\|_{M^2}^2 \log \|f\|_{M^2} \right| &\leqslant C^2 \left| \|f\|_{M^\Phi}^2 \log(C\|f\|_{M^\Phi}) \right| \\
&\leqslant C^2 \left(\left| \|f\|_{M^\Phi}^2 \log(\|f\|_{M^\Phi}) \right| + (\log C)\|f\|_{M^\Phi}^2 \right) \to 0
\end{aligned}$$

as $\|f\|_{M^\Phi} \to 0$. A combination of the latter continuity and (3.7) now gives

$$E_\phi(f) = -\iint_{\mathbf{R}^{2d}} |V_\phi f(x,\xi)|^2 \log |V_\phi f(x,\xi)^2| \, dx d\xi + \|f\|_{L^2}^2 \log \|f\|_{L^2}^2 \to 0$$
$$\text{as} \quad \|f\|_{M^\Phi} \to 0, \ f \in M^\Phi(\mathbf{R}^d),$$

and the asserted continuity for E_ϕ near origin follows.

Next we prove that E_ϕ is continuous at a general $f \in M^\Phi(\mathbf{R}^d)$. Due to the first part it suffices to prove that E_ϕ is continuous outside origin. Therefore assume that $f \in M^\Phi(\mathbf{R}^d) \setminus 0$. By using the homogeneity $E_\phi(\lambda f) = |\lambda|^2 E_\phi(f)$ when $f \in M^\Phi(\mathbf{R}^d)$ in combination with (3.3), it follows that it suffices to prove the result under the additional condition

$$\|f\|_{M^\Phi} + \|f\|_{M^2} + \|f\|_{M^\infty} < 1.$$

For conveniency we set $F = V_\phi f$, $G = V_\phi g$, $H = F + G$ and

$$J(f,g) \equiv \left| \iint\limits_{\mathbf{R}^{2d}} \big(|H(x,\xi)|^2 \log |H(x,\xi)| - |F(x,\xi)|^2 \log |F(x,\xi)| \big) \, dx d\xi \right|$$

when $g \in M^\Phi(\mathbf{R}^d)$. We have

$$|E_\phi(f+g) - E_\phi(f)| \leqslant 2 \left(J(f,g) + \left| \|H\|_{L^2}^2 \log \|H\|_{L^2} - \|F\|_{L^2}^2 \log \|F\|_{L^2} \right| \right). \tag{3.14}$$

If $\|g\|_{M^\Phi} \to 0$, then $\|g\|_{M^2} \to 0$, which implies that $\|H\|_{L^2} \to \|F\|_{L^2}$ as $\|g\|_{M^2} \to 0$. Hence, by the continuity of $t^2 \log t$ on $[0, \infty)$, it follows that last modulus in (3.14) tends to zero as $\|g\|_{M^\Phi} \to 0$. This implies that the asserted continuity follows if we prove

$$J(f,g) \to 0 \quad \text{as} \quad \|g\|_{M^\Phi} \to 0. \tag{3.15}$$

Let $R > 1$ be fixed and let

$$\Omega = \{\, (x,\xi) \in \mathbf{R}^{2d}\, ;\, |F(x,\xi)| > R|G(x,\xi)| \,\}.$$

Then

$$0 \leqslant J(f,g) \leqslant \sum_{k=1}^{3} J_k(f,g), \tag{3.16}$$

where

$$J_1(f,g) = \left| \iint\limits_{\Omega} (|H(x,\xi)|^2 - |F(x,\xi)|^2) \log |F(x,\xi)| \, dx d\xi \right|,$$

$$J_2(f,g) = \left| \iint\limits_{\Omega} |H(x,\xi)|^2 \log \left| \frac{H(x,\xi)}{F(x,\xi)} \right| \, dx d\xi \right|,$$

and

$$J_3(f,g) = \left| \iint\limits_{\complement\Omega} \big(|H(x,\xi)|^2 \log |H(x,\xi)| - |F(x,\xi)|^2 \log |F(x,\xi)| \big) \, dx d\xi \right|.$$

We shall estimate $J_k(f,g)$ in suitable ways, $k = 1, 2, 3$.

For the integrand in $J_1(f,g)$, taken into account that

$$R|G(x,\xi)| < |F(x,\xi)| < 1,$$

we have

$$\begin{aligned}0 &\leqslant \big|(|H(x,\xi)|^2 - |F(x,\xi)|^2)\log|F(x,\xi)|\big| \\ &= -\big||F(x,\xi)+G(x,\xi)|^2 - |F(x,\xi)|^2\big|\log|F(x,\xi)| \\ &\leqslant -(2|F(x,\xi)|\,|G(x,\xi)| + |G(x,\xi)|^2)\log|F(x,\xi)| \\ &\leqslant -\left(\frac{2}{R}+\frac{1}{R^2}\right)|F(x,\xi)|^2\log|F(x,\xi)|,\end{aligned}$$

which gives

$$J_1(f,g) \leqslant -\left(\frac{2}{R}+\frac{1}{R^2}\right)\iint\limits_{\mathbf{R}^{2d}} |F(x,\xi)|^2\log|F(x,\xi)|\,dxd\xi. \tag{3.17}$$

For the logarithm in $J_2(f,g)$ we have

$$\begin{aligned}\left|\log\left|\frac{H(x,\xi)}{F(x,\xi)}\right|\right| &= \left|\log\left|1+\frac{G(x,\xi)}{F(x,\xi)}\right|\right| \leqslant -\log\left(1-\frac{|G(x,\xi)|}{|F(x,\xi)|}\right) \\ &\leqslant -\log\left(1-\frac{1}{R}\right) = \sum_{j=1}^{\infty}\frac{R^{-j}}{j} \leqslant \sum_{j=1}^{\infty} R^{-j} = \frac{1}{R-1}.\end{aligned}$$

In the second inequality we have used the fact that $R > 1$ and that $|F(x,\xi)| > R|G(x,\xi)|$ when $(x,\xi) \in \Omega$.

This gives

$$\begin{aligned}J_2(f,g) &\leqslant \frac{1}{R-1}\iint\limits_{\Omega} |H(x,\xi)|^2\,dxd\xi \\ &\leqslant \frac{2}{R-1}\iint\limits_{\Omega} (|F(x,\xi)|^2 + |G(x,\xi)|^2)\,dxd\xi \\ &< \frac{2}{R-1}\iint\limits_{\Omega} (|F(x,\xi)|^2 + \frac{1}{R^2}|F(x,\xi)|^2)\,dxd\xi,\end{aligned}$$

which in turn gives

$$J_2(f,g) < \frac{2}{R-1}\left(1+\frac{1}{R^2}\right)\|F\|_{L^2}^2. \tag{3.18}$$

Next we estimate $J_3(f,g)$. By (3.3) there is a $\delta_0 > 0$ such that

$$|G(x,\xi)| \leqslant \frac{e^{-\frac{1}{2}}}{(R+1)}, \quad \text{when} \quad \|g\|_{M^\Phi} < \delta_0. \tag{3.19}$$

Since $|t^2 \log t| = -t^2 \log t$ is increasing on $[0, e^{-\frac{1}{2}}]$,

$$|H(x,\xi)| \leqslant |F(x,\xi)| + |G(x,\xi)| \leqslant (R+1)|G(x,\xi)| \leqslant e^{-\frac{1}{2}}$$

and

$$|F(x,\xi)| \leqslant R|G(x,\xi)| \leqslant e^{-\frac{1}{2}}$$

when $(x,\xi) \in \complement\Omega$ by (3.19), we obtain

$$\begin{aligned}
J_3(f,g) &\leqslant \left| \iint\limits_{\complement\Omega} |H(x,\xi)|^2 \log |H(x,\xi)|\, dxd\xi \right| + \left| \iint\limits_{\complement\Omega} |F(x,\xi)|^2 \log |F(x,\xi)|\, dxd\xi \right| \\
&\leqslant \left| \iint\limits_{\complement\Omega} |(R+1)G(x,\xi)|^2 \log |(R+1)G(x,\xi)|\, dxd\xi \right| \\
&\quad + \left| \iint\limits_{\complement\Omega} |RG(x,\xi)|^2 \log |RG(x,\xi)|\, dxd\xi \right| \\
&\leqslant \big((R+1)^2 + R^2\big) \left| \iint\limits_{\complement\Omega} |G(x,\xi)|^2 \log |G(x,\xi)|\, dxd\xi \right| \\
&\quad + \big((R+1)^2 \log(R+1) + R^2 \log R\big) \iint\limits_{\complement\Omega} |G(x,\xi)|^2\, dxd\xi
\end{aligned}$$

when $\|g\|_{M^\Phi} < \delta_0$. A combination of these estimates and the fact that $\log|G(x,\xi)| < 0$ in view of (3.19) gives

$$\begin{aligned}
J_3(f,g) \leqslant &- \big((R+1)^2 + R^2\big) \iint\limits_{\mathbf{R}^{2d}} |G(x,\xi)|^2 \log |G(x,\xi)|\, dxd\xi \\
&+ \big((R+1)^2 \log(R+1) + R^2 \log R\big) \|G\|_{L^2}^2, \quad \|g\|_{M^\Phi} < \delta_0.
\end{aligned} \tag{3.20}$$

Now let $\varepsilon > 0$ be arbitrary. By Lemma 3.4, (3.3), (3.17) and (3.18) it follows that $J_1(f,g) < \frac{\varepsilon}{3}$ and $J_2(f,g) < \frac{\varepsilon}{3}$, provided R is chosen large enough. A combination of Lemma 3.5, (3.3) and (3.20) shows that there is a positive number $\delta < \delta_0$ such that $J_3(f,g) < \frac{\varepsilon}{3}$ when $\|g\|_{M^\Phi} < \delta$.

By combining these estimates with (3.16) now gives

$$0 \leqslant J(f,g) < \varepsilon \quad \text{when} \quad g \in M^{\Phi}(\mathbf{R}^d),\ \|g\|_{M^{\Phi}} < \delta.$$

This shows that (3.15) holds true, and the continuity for E_ϕ on $M^{\Phi}(\mathbf{R}^d)$ follows.

The continuity for E_ϕ on $M^p(\mathbf{R}^d)$, $0 < p < 2$ now follows from the fact that $M^p(\mathbf{R}^d)$ is continuously embedded in $M^{\Phi}(\mathbf{R}^d)$. (See Proposition 1.14.)

It remains to prove the discontinuity for E_ϕ on $M^p(\mathbf{R}^d)$, $p \geqslant 2$, and then it follows from Lemma 3.6 that we may assume that $\phi(x) = \pi^{-\frac{d}{4}} e^{-\frac{1}{2}|x|^2}$. Since $M^2(\mathbf{R}^d)$ is continuously embedded in $M^p(\mathbf{R}^d)$ when $p \geqslant 2$, it suffices to prove the asserted discontinuity for $p = 2$.

We shall investigate $E_\phi(f)$ with

$$f(x) = f_\lambda(x) = \pi^{-\frac{d}{4}} \lambda^{\frac{d}{4}} e^{-\frac{\lambda}{2}|x|^2}, \qquad \lambda > 1.$$

Then $\|\phi\|_{L^2} = \|f_\lambda\|_{L^2} = 1$, and by straight-forward computations it follows that

$$V_\phi f_\lambda(x,\xi) = \left(\frac{\lambda^{\frac{1}{2}}}{\pi(\lambda+1)} \right)^{\frac{d}{2}} e^{-\frac{i}{\lambda+1}\langle x,\xi\rangle} e^{-\frac{1}{2(\lambda+1)}(\lambda|x|^2+|\xi|^2)},$$

and since f_λ is L^2-normalized we get

$$\begin{aligned} E_\phi(f_\lambda) &= -\iint_{\mathbf{R}^{2d}} |V_\phi f_\lambda(x,\xi)|^2 \log |V_\phi f_\lambda(x,\xi)|^2 \, dx d\xi \\ &= \left(\frac{\lambda^{\frac{1}{2}}}{\pi(\lambda+1)} \right)^{d} \iint_{\mathbf{R}^{2d}} h_\lambda\Big(\frac{1}{\lambda+1}(\lambda|x|^2 + |\xi|^2) \Big) \, dx d\xi, \end{aligned}$$

where

$$h_\lambda(t) = e^{-t} \left(\frac{t}{2} + d \log \left(\pi(\lambda^{\frac{1}{2}} + \lambda^{-\frac{1}{2}}) \right) \right).$$

By taking $(\frac{\lambda}{\lambda+1})^{\frac{1}{2}} x$ and $(\frac{1}{\lambda+1})^{\frac{1}{2}} \xi$ as new variables of integrations we obtain

$$\begin{aligned} E_\phi(f_\lambda) &= \pi^{-d} \iint_{\mathbf{R}^{2d}} h_\lambda(|x|^2 + |\xi|^2) \, dx d\xi \\ &= \pi^{-d} \iint_{\mathbf{R}^{2d}} e^{-(|x|^2+|\xi|^2)} \left(\frac{1}{2}(|x|^2 + |\xi|^2) + d \log \left(\pi(\lambda^{\frac{1}{2}} + \lambda^{-\frac{1}{2}}) \right) \right) dx d\xi \\ &= d \left(\frac{1}{4} + \log \left(\pi(\lambda^{\frac{1}{2}} + \lambda^{-\frac{1}{2}}) \right) \right). \end{aligned}$$

This implies

$$\lim_{\lambda\to 0+} E_\phi(f_\lambda) = \lim_{\lambda\to\infty} E_\phi(f_\lambda) = \infty \quad \text{but} \quad \|f_\lambda\|_{L^2} = 1,$$

which shows that E_ϕ is discontinuous on $L^2(\mathbf{R}^d) = M^2(\mathbf{R}^d)$, and the result follows. □

By Theorem 3.1 and its proof it follows that Lemma 3.5 can be improved into the following.

Lemma 3.5′. *Let Φ and ϕ be the same as in Lemma 3.4. Then*

$$M^\Phi(\mathbf{R}^d) \ni f \mapsto \iint_{\mathbf{R}^{2d}} |V_\phi f(x,\xi)|^2 \big| \log |V_\phi f(x,\xi)| \big| \, dx d\xi$$

is locally uniformly continuous.

Remark 3.7. In view of Theorem 3.1 and its proof it follows that (2) in that theorem can be extended into the following:

(2)′ *E_ϕ in (0.1) is locally uniformly continuous on $M^p(\mathbf{R}^d)$ and on $M^\Phi(\mathbf{R}^d)$, $0 < p < 2$, and discontinuous on $M^p(\mathbf{R}^d)$ for $2 \leqslant p \leqslant \infty$.*

Remark 3.8. Let $A \in \mathbf{M}(d, \mathbf{R})$ and Φ be a Young function which fullfils (0.8). We claim that there is a symbol a in $M^2(\mathbf{R}^{2d})$ (which is close to $M^{p,p'}(\mathbf{R}^{2d})$ when $p > 2$ is close to 2) such that the map (2.31) is discontinuous. (Cf. Example 2.11.)

In fact, by Remark 3.3, there are $f_1 \in M^2(\mathbf{R}^d) \setminus M^\Phi(\mathbf{R}^d)$ and $f_2 \in \mathscr{S}(\mathbf{R}^d) \setminus 0$. Then $a = W^A_{f_1,f_2} \in M^2(\mathbf{R}^{2d})$. By (1.25) it follows that

$$\operatorname{Op}_A(a) f(x) = (2\pi)^{-\frac{d}{2}} (f, f_2)_{L^2} f_1(x) \in M^2(\mathbf{R}^d) \setminus M^\Phi(\mathbf{R}^d)$$

for every $f \in \mathscr{S}(\mathbf{R}^d) \subseteq M^\Phi(\mathbf{R}^d)$ which is not orthogonal to f_2, and the asserted discontinuity follows.

Data availability

No data was used for the research described in the article.

Appendix A. STFT projections on Orlicz modulation spaces

In this appendix we first recall some facts on projections on Orlicz modulation spaces which appear after compositions between short-time Fourier transforms and their adjoints.

Thereafter we apply our results to give a proof of Proposition 1.16.

A.1. STFT projections and twisted convolutions

Let $s \geqslant \frac{1}{2}$. If $\phi \in \mathcal{S}_s(\mathbf{R}^d) \setminus 0$, then it follows from Fourier's inversion formula that

$$\mathrm{Id} = \mathrm{Id}_{\mathcal{S}_s'} = \left(\|\phi\|_{L^2}^{-2}\right) \cdot V_\phi^* \circ V_\phi, \tag{A.1}$$

is the identity operator on $\mathcal{S}_s'(\mathbf{R}^d)$. The same and following results hold true with Σ_s and $\mathscr{S}$ in place of $\mathcal{S}_s$ at each occurrence. The identity (A.1) is equivalent to Moyal's identity (3.1). If we swap the order of this composition we get certain types of projections. More precisely, let

$$P_\phi \equiv \|\phi\|_{L^2}^{-2} \cdot V_\phi \circ V_\phi^*. \tag{A.2}$$

We observe that P_ϕ is continuous on $\mathcal{S}_s(\mathbf{R}^{2d})$, $L^2(\mathbf{R}^{2d})$ and on $\mathcal{S}_s'(\mathbf{R}^{2d})$ due to the mapping properties for V_ϕ and V_ϕ^*.

It is clear that $P_\phi^* = P_\phi$, i.e. P_ϕ is self-adjoint. Furthermore,

$$P_\phi^2 = \|\phi\|_{L^2}^{-2} \cdot V_\phi \circ \Big(\underbrace{\|\phi\|_{L^2}^{-2} \cdot V_\phi^* \circ V_\phi}_{\text{The identity operator}} \Big) \circ V_\phi^* = \|\phi\|_{L^2}^{-2} \cdot V_\phi \circ V_\phi^* = P_\phi,$$

giving that P_ϕ is an orthonormal projection, that is,

$$P_\phi^* = P_\phi \quad \text{and} \quad P_\phi^2 = P_\phi. \tag{A.3}$$

The ranks of P_ϕ are given by

$$P_\phi(\mathcal{S}_s(\mathbf{R}^{2d})) = V_\phi(\mathcal{S}_s(\mathbf{R}^d)) \quad \text{and} \quad P_\phi(\mathcal{S}_s'(\mathbf{R}^{2d})) = V_\phi(\mathcal{S}_s'(\mathbf{R}^d)). \tag{A.4}$$

In fact, if $F \in \mathcal{S}_s'(\mathbf{R}^{2d})$, then

$$P_\phi F = V_\phi f, \tag{A.5}$$

where $f = \|\phi\|_{L^2}^{-2} V_\phi^* F \in \mathcal{S}_s'(\mathbf{R}^d)$. This shows that $P_\phi(\mathcal{S}_s'(\mathbf{R}^{2d})) \subseteq V_\phi(\mathcal{S}_s'(\mathbf{R}^d))$. On the other hand, if $f \in \mathcal{S}_s'(\mathbf{R}^d)$ and $F = V_\phi f$, then

$$P_\phi F = \Big(V_\phi \circ \Big(\|\phi\|_{L^2}^{-2} \cdot V_\phi^* \circ V_\phi \Big) \Big) f = V_\phi f,$$

which shows that any element in $V_\phi(\mathcal{S}_s'(\mathbf{R}^d))$ is equal to an element in $P_\phi(\mathcal{S}_s'(\mathbf{R}^{2d}))$, i.e. $P_\phi(\mathscr{S}'(\mathbf{R}^{2d})) = V_\phi(\mathscr{S}'(\mathbf{R}^d))$. The same holds true with $\mathcal{S}_s$ in place of $\mathcal{S}_s'$ at each occurrence, and (A.4) follows.

Remark A.1. Let $F \in \mathcal{S}_s'(\mathbf{R}^{2d})$. Then (A.4) shows that $F = V_\phi f$ for some $f \in \mathcal{S}_s'(\mathbf{R}^d)$, if and only if

$$F = P_\phi F. \tag{A.6}$$

Furthermore, if (A.6) holds, then $F = V_\phi f$ with

$$f = (\|\phi\|_{L^2}^{-2}) \cdot V_\phi^* F. \tag{A.7}$$

Let $F \in \mathcal{S}_s(\mathbf{R}^{2d})$ and $\phi \in \mathcal{S}_s(\mathbf{R}^d) \setminus 0$. Then by expanding the integrals for V_ϕ and V_ϕ^* in (A.2) one obtains

$$P_\phi F = \|\phi\|_{L^2}^{-2} \cdot V_\phi \phi *_V F, \qquad F \in \mathscr{S}'(\mathbf{R}^{2d}), \tag{A.8}$$

where the *twisted convolution* $*_V$ is defined by

$$(F *_V G)(x,\xi) = (2\pi)^{-\frac{d}{2}} \iint\limits_{\mathbf{R}^{2d}} F(x-y,\xi-\eta) G(y,\eta) e^{-i\langle y,\xi-\eta\rangle}\, dy d\eta, \tag{A.9}$$

when $F, G \in \mathcal{S}_s(\mathbf{R}^{2d})$. We observe that the definition of $*_V$ extends in different ways. For example, Young's inequality for ordinary convolution also holds for $*_V$. Moreover, the map $(F,G) \mapsto F *_V G$ extends uniquely to continuous mappings from $\mathcal{S}_s(\mathbf{R}^{2d}) \times \mathcal{S}_s'(\mathbf{R}^{2d})$ or $\mathcal{S}_s'(\mathbf{R}^{2d}) \times \mathcal{S}_s(\mathbf{R}^{2d})$ to $\mathcal{S}_s'(\mathbf{R}^{2d})$. By straight-forward computations it follows that

$$(F *_V G) *_V H = F *_V (G *_V H), \tag{A.10}$$

when $F, H \in \mathcal{S}_s(\mathbf{R}^{2d})$ and $G \in \mathcal{S}_s'(\mathbf{R}^{2d})$, or $F, H \in \mathcal{S}_s'(\mathbf{R}^{2d})$ and $G \in \mathcal{S}_s(\mathbf{R}^{2d})$.

Let $f \in \mathcal{S}_s'(\mathbf{R}^d)$ and $\phi_j \in \mathcal{S}(\mathbf{R}^d)$, $j = 1, 2, 3$. By straight-forward applications of Parseval's formula it follows that

$$\big((V_{\phi_2}\phi_3) *_V (V_{\phi_1} f)\big)(x,\xi) = (\phi_3, \phi_1)_{L^2} \cdot (V_{\phi_2} f)(x,\xi), \tag{A.11}$$

which is some sort of reproducing kernel of short-time Fourier transforms in the background of $*_V$. (See also Chapter 11 in [15].)

A.2. Applications to Orlicz modulation spaces

We have now the following which essentially follows from Proposition 4.3 and its proof in [10].

Lemma A.2. *Let Φ and Ψ be Young functions, $\phi \in \Sigma_1(\mathbf{R}^d)$ be such that $\|\phi\|_{L^2} = 1$ and let $\omega \in \mathscr{P}_E(\mathbf{R}^{2d})$. Then the following is true:*

(1) *P_ϕ from $\Sigma_1'(\mathbf{R}^{2d})$ to $V_\phi(\Sigma_1'(\mathbf{R}^d))$ restricts to a continuous projection from $L^{\Phi,\Psi}_{(\omega)}(\mathbf{R}^{2d})$ to $V_\phi(M^{\Phi,\Psi}_{(\omega)}(\mathbf{R}^d))$;*

(2) *if* $F \in L^{\Phi,\Psi}_{(\omega)}(\mathbf{R}^{2d})$ *and* $f = V_\phi^* F$, *then* $V_\phi f = P_\phi F$ *and*

$$\|f\|_{M^{\Phi,\Psi}_{(\omega)}} \asymp \|P_\phi F\|_{L^{\Phi,\Psi}_{(\omega)}} \lesssim \|F\|_{L^{\Phi,\Psi}_{(\omega)}}, \qquad f = V_\phi^* F. \tag{A.12}$$

Proof. By (A.5) and Remark A.1, the result follows if we prove (A.12).

Let $v \in \mathscr{P}_E(\mathbf{R}^{2d})$ be submultiplicative such that ω is v-moderate. By (A.8) we have

$$|F *_V G| \leqslant |F| * |G|.$$

Hence (A.5) and (A.9) give

$$\begin{aligned}\|f\|_{M^{\Phi,\Psi}_{(\omega)}} &\asymp \|V_\phi f\|_{L^{\Phi,\Psi}_{(\omega)}} = \|P_\phi F\|_{L^{\Phi,\Psi}_{(\omega)}} \\ &\leqslant \||F| * |V_\phi \phi|\|_{L^{\Phi,\Psi}_{(\omega)}} \lesssim \|F\|_{L^{\Phi,\Psi}_{(\omega)}} \|V_\phi \phi\|_{L^1_{(v)}}.\end{aligned}$$

The asserted continuity now follows from the fact that for some $r > 0$ we have

$$v(x,\xi) \lesssim e^{r(|x|+|\xi|)} \quad \text{and} \quad |V_\phi \phi(x,\xi)| \lesssim e^{-2r(|x|+|\xi|)},$$

in view of Proposition 1.1 and (1.9). □

Proof of Proposition 1.16. We have

$$|(F,G)_{L^2(\mathbf{R}^{2d})}| \lesssim \|F\|_{L^{\Phi,\Psi}_{(\omega)}} \|G\|_{L^{\Phi^*,\Psi^*}_{(\omega)}}$$

when $F, G \in \Sigma_1(\mathbf{R}^{2d})$, by Hölder's inequality for Orlicz spaces (cf. e.g. [20,29]). By Hahn-Banach's theorem it follows that the map $(F,G) \to (F,G)_{L^2(\mathbf{R}^{2d})}$ from $\Sigma_1(\mathbf{R}^{2d}) \times \Sigma_1(\mathbf{R}^{2d})$ to $\mathbf{C}$ extends to a continuous map from $L^{\Phi,\Psi}_{(\omega)}(\mathbf{R}^{2d}) \times L^{\Phi^*,\Psi^*}_{(\omega)}(\mathbf{R}^{2d})$ to $\mathbf{C}$.

If $\phi \in \Sigma_1(\mathbf{R}^d) \setminus 0$ satisfies $\|\phi\|_{L^2} = 1$, $f \in M^{\Phi,\Psi}_{(\omega)}(\mathbf{R}^d)$ and $g \in M^{\Phi^*,\Psi^*}_{(\omega)}(\mathbf{R}^d)$, we now use Moyal's identity to define $(f,g)_{L^2(\mathbf{R}^d)} = (V_\phi f, V_\phi g)_{L^2(\mathbf{R}^{2d})}$, which satisfies the requested properties, because

$$\begin{aligned}|(f,g)_{L^2(\mathbf{R}^d)}| &= |(V_\phi f, V_\phi g)_{L^2(\mathbf{R}^{2d})}| \\ &\lesssim \|V_\phi f\|_{L^{\Phi,\Psi}_{(\omega)}} \|V_\phi g\|_{L^{\Phi^*,\Psi^*}_{(\omega)}} \asymp \|f\|_{M^{\Phi,\Psi}_{(\omega)}} \|g\|_{M^{\Phi^*,\Psi^*}_{(\omega)}},\end{aligned} \tag{A.13}$$

and the continuity extension in (1) follows. Suppose from now on that Φ and Ψ in addition satisfy the Δ_2-condition. Then $\Sigma_1(\mathbf{R}^d)$ is dense in $M^{\Phi,\Psi}_{(\omega)}(\mathbf{R}^d)$ which implies that the latter continuity extension is unique.

Next suppose that T is a continuous linear form on $M^{\Phi,\Psi}_{(\omega)}(\mathbf{R}^d)$. Then

$$T_1(V_\phi f) \equiv T(f)$$

satisfies

$$|T_1(V_\phi f)| \lesssim |T(f)| \lesssim \|f\|_{M^{\Phi,\Psi}_{(\omega)}} \asymp \|V_\phi f\|_{L^{\Phi,\Psi}_{(\omega)}}.$$

Hence T_1 is a continuous linear form on $V_\phi(M^{\Phi,\Psi}_{(\omega)}(\mathbf{R}^d))$. Since the injection from $V_\phi(M^{\Phi,\Psi}_{(\omega)}(\mathbf{R}^d))$ to $L^{\Phi,\Psi}_{(\omega)}(\mathbf{R}^{2d})$ is norm preserving, it follows by Hahn-Banach's theorem that T_1 extends to a linear form on $L^{\Phi,\Psi}_{(\omega)}(\mathbf{R}^{2d}))$ with the same norm. By [29] it follows that the dual of the latter space is equal to $L^{\Phi^*,\Psi^*}_{(1/\omega)}(\mathbf{R}^{2d})$ through the $(\,\cdot\,,\,\cdot\,)_{L^2(\mathbf{R}^{2d})}$ form. Hence

$$T_1(F) = (F,G)_{L^2(\mathbf{R}^{2d})} = \iint\limits_{\mathbf{R}^{2d}} F(x,\xi)\overline{G(x,\xi)}\,dxd\xi, \qquad F \in L^{\Phi,\Psi}_{(\omega)}(\mathbf{R}^{2d}),$$

for some fixed $G \in L^{\Phi^*,\Psi^*}_{(1/\omega)}(\mathbf{R}^{2d}))$ which satisfies

$$\|G\|_{L^{\Phi^*,\Psi^*}_{(1/\omega)}} \asymp \|T_1\| = \|T\|. \tag{A.14}$$

By Lemma A.2 we also have $P_\phi G = V_\phi g$ for some $g \in M^{\Phi,\Psi}_{(\omega)}(\mathbf{R}^d)$. A combination of these identities and Moyal's identity gives that for any $f \in M^{\Phi,\Psi}_{(\omega)}(\mathbf{R}^d)$ we have

$$\begin{aligned} T(f) = (V_\phi f, G)_{L^2(\mathbf{R}^{2d})} &= (P_\Phi(V_\phi f), G)_{L^2(\mathbf{R}^{2d})} \\ &= (V_\phi f, P_\phi G)_{L^2(\mathbf{R}^{2d})} = (V_\phi f, V_\phi g)_{L^2(\mathbf{R}^{2d})} \\ &= (f,g)_{L^2((\mathbf{R}^d))}, \end{aligned} \tag{A.15}$$

which gives (2).

Finally, by (A.13) it follows that $\|f\| \lesssim \|f\|_{M^{\Phi,\Psi}_{(\omega)}}$ when $f \in M^{\Phi,\Psi}_{(\omega)}(\mathbf{R}^d)$.

On the other hand, let $f_0 \in M^{\Phi,\Psi}_{(\omega)}(\mathbf{R}^d)$ be fixed and let T be the linear form on $\{\,\lambda f_0\,;\,\lambda \in \mathbf{C}\,\} \subseteq M^{\Phi,\Psi}_{(\omega)}(\mathbf{R}^d)$ given by

$$T(\lambda f_0) = \lambda \|f_0\|_{M^{\Phi,\Psi}_{(\omega)}}.$$

Then $\|T\| = 1$. By Hahn-Banach's theorem, there is a $G \in L^{\Phi^*,\Psi^*}_{(1/\omega)}(\mathbf{R}^{2d}))$ such that T extends to a form on $M^{\Phi,\Psi}_{(\omega)}(\mathbf{R}^d)$ and such that (A.14) and (A.15) hold. Since $\|g\|_{M^{\Phi^*,\Psi^*}_{(1/\omega)}} \lesssim \|G\|_{L^{\Phi^*,\Psi^*}_{(1/\omega)}}$ in view of Lemma A.2 we get by choosing $f = f_0$ that

$$\|f_0\|_{M^{\Phi,\Psi}_{(\omega)}} = T(f_0) = (f_0,g)_{L^2} \lesssim \sup |(f_0,g)_{L^2}| = \|f_0\|,$$

where the hidden constants are independent of $f_0 \in M^{\Phi,\Psi}_{(\omega)}(\mathbf{R}^d)$. Here the supremum is taken over all $g \in M^{\Phi^*,\Psi^*}_{(1/\omega)}(\mathbf{R}^d)$ such that $\|g\|_{M^{\Phi^*,\Psi^*}_{(1/\omega)}} \leqslant 1$. Consequently we have $\|f\|_{M^{\Phi,\Psi}_{(\omega)}} \asymp \|f_0\|$, giving that (1), and thereby the result follow. $\square$

References

[1] M. Cappiello, J. Toft, Pseudo-differential operators in a Gelfand–Shilov setting, Math. Nachr. 290 (2017) 738–755.

[2] E. Cordero, G. Giacchi, Symplectic analysis of time-frequency spaces, J. Math. Pures Appl. (2023) 154–177.

[3] E. Cordero, F. Nicola, Pseudodifferential operators on L^p, Wiener amalgam and modulation spaces, Int. Math. Res. Not. 8 (2010) 1860–1893.

[4] E. Cordero, F. Nicola, Sharp integral bounds for Wigner distributions, Int. Math. Res. Not. 6 (2018) 1779–1807.

[5] E. Cordero, K. Okoudjou, Multilinear localization operators, J. Math. Anal. Appl. 325 (2007) 1103–1116.

[6] E. Cordero, S. Pilipović, L. Rodino, N. Teofanov, Quasianalytic Gelfand-Shilov spaces with applications to localization operators, Rocky Mt. J. Math. 40 (2010) 1123–1147.

[7] E. Cordero, L. Rodino, Time-Frequency Analysis of Operators, De Gruyter Studies in Mathematics, vol. 75, De Gruyter, Berlin, 2020.

[8] H.G. Feichtinger, Modulation spaces on locally compact abelian groups, Technical report, University of Vienna, Vienna, 1983; also in: M. Krishna, R. Radha, S. Thangavelu (Eds.), Wavelets and their applications, Allied Publishers Private Limited, NewDelhi Mumbai Kolkata Chennai Nagpur Ahmedabad Bangalore Hyderabad Lucknow, 2003, pp. 99–140.

[9] H.G. Feichtinger, Modulation spaces: looking back and ahead, Sampl. Theory Signal Image Process. 5 (2006) 109–140.

[10] H.G. Feichtinger, K. Gröchenig, Banach spaces related to integrable group representations and their atomic decompositions, I, J. Funct. Anal. 86 (1989) 307–340.

[11] H.G. Feichtinger, K. Gröchenig, Banach spaces related to integrable group representations and their atomic decompositions, II, Monatshefte Math. 108 (1989) 129–148.

[12] G.B. Folland, Harmonic Analysis in Phase Space, Princeton U. P., Princeton, 1989.

[13] Y.V. Galperin, S. Samarah, Time-frequency analysis on modulation spaces $M_m^{p,q}$, $0 < p, q \leqslant \infty$, Appl. Comput. Harmon. Anal. 16 (2004) 1–18.

[14] I.M. Gelfand, G.E. Shilov, Generalized Functions, II-III, Academic Press, NewYork London, 1968.

[15] K. Gröchenig, Foundations of Time-Frequency Analysis, Birkhäuser, Boston, 2001.

[16] K. Gröchenig, Composition and spectral invariance of pseudodifferential operators on modulation spaces, J. Anal. Math. 98 (2006) 65–82.

[17] K. Gröchenig, C. Heil, Modulation spaces and pseudo-differential operators, Integral Equ. Oper. Theory (4) 34 (1999) 439–457.

[18] K. Gröchenig, C. Heil, Modulation spaces as symbol classes for pseudodifferential operators, in: M. Krishna, R. Radha, S. Thangavelu (Eds.), Wavelets and their applications, Allied Publishers Private Limited, NewDelhi Mumbai Kolkata Chennai Nagpur Ahmedabad Bangalore Hyderabad Lucknow, 2003, pp. 151–170.

[19] K. Gröchenig, G. Zimmermann, Spaces of test functions via the STFT, J. Funct. Spaces Appl. 2 (2004) 25–53.

[20] P. Harjulehto, P. Hästö, Orlicz Spaces and Generalized Orlicz Spaces, Springer, Cham, 2019.

[21] L. Hörmander, The Analysis of Linear Partial Differential Operators, vol I–III, Springer-Verlag Berlin Heidelberg NewYork Tokyo, 1983, 1985.

[22] G. Köthe, Topological Vector Spaces I, Die Grundlehren der mathematischen Wissenschaften, vol. 159, Springer-Verlag, NewYork-Berlin, 1969.

[23] E.H. Lieb, Integral bounds for radar ambiguity functions and Wigner distributions, J. Math. Phys. 31 (1990) 594–599.

[24] E.H. Lieb, J.P. Solovej, Quantum coherent operators: a generalization of coherent states, Lett. Math. Phys. 22 (1991) 145–154.

[25] W.A. Majewski, L.E. Labuschagne, On applications of Orlicz spaces to statistical physics, Ann. Henri Poincaré 15 (2014) 1197–1221.

[26] W.A. Majewski, L.E. Labuschagne, On entropy for general quantum systems, Adv. Theor. Math. Phys. 24 (2020) 491–526.

[27] S. Pilipović, Generalization of Zemanian spaces of generalized functions which have orthonormal series expansions, SIAM J. Math. Anal. 17 (1986) 477–484.

[28] S. Pilipović, Tempered ultradistributions, Boll. UMI 7 (1988) 235–251.

[29] M.M. Rao, Z.D. Ren, Theory of Orlicz Spaces, Marcel Dekker, New York, 1991.

[30] H. Rauhut, Wiener amalgam spaces with respect to quasi-Banach spaces, Colloq. Math. 109 (2007) 345–362.
[31] H. Rauhut, Coorbit space theory for quasi-Banach spaces, Stud. Math. 180 (2007) 237–253.
[32] W. Rudin, Real and Complex Analysis, McGraw-Hill Book Co., New York, 1987.
[33] H.H. Schaefer, M.P. Wolff, Topological Vector Spaces, 2nd. ed., Graduate Texts in Mathematics, vol. 3, Springer-Verlag, New York, 1999.
[34] C. Schnackers, H. Führ, Orlicz Modulation Spaces, in: Proceedings of the 10th International Conference on Sampling Theory and Applications, 2013.
[35] N. Teofanov, Ultradistributions and time-frequency analysis, in: P. Boggiatto, L. Rodino, J. Toft, M.W. Wong (Eds.), Pseudo-Differential Operators and Related Topics, in: Operator Theory: Advances and Applications, vol. 164, Birkhäuser, Basel, 2006, pp. 173–192.
[36] N. Teofanov, J. Toft, Pseudo-differential calculus in a Bargmann setting, Ann. Acad. Sci. Fenn., Math. 45 (2020) 227–257.
[37] J. Toft, Continuity properties for modulation spaces, with applications to pseudo-differential operators, I, J. Funct. Anal. 207 (2004) 399–429.
[38] J. Toft, Continuity properties for modulation spaces, with applications to pseudo-differential operators, II, Ann. Glob. Anal. Geom. 26 (2004) 73–106.
[39] J. Toft, Pseudo-differential operators with symbols in modulation spaces, in: B.-W. Schulze, M.W. Wong (Eds.), Pseudo-Differential Operators: Complex Analysis and Partial Differential Equations, in: Operator Theory Advances and Applications, vol. 205, Birkhäuser Verlag, Basel, 2010, pp. 223–234.
[40] J. Toft, The Bargmann transform on modulation and Gelfand-Shilov spaces, with applications to Toeplitz and pseudo-differential operators, J. Pseudo-Differ. Oper. Appl. 3 (2012) 145–227.
[41] J. Toft, Gabor analysis for a broad class of quasi-Banach modulation spaces, in: S. Pilipović, J. Toft (Eds.), Pseudo-Differential Operators, Generalized Functions, in: Operator Theory: Advances and Applications, vol. 245, Birkhäuser, 2015, pp. 249–278.
[42] J. Toft, Continuity and compactness for pseudo-differential operator with symbols in quasi-Banach spaces or Hörmander classes, Anal. Appl. 15 (2017) 353–389.
[43] J. Toft, Matrix parameterized pseudo-differential calculi on modulation spaces, in: M. Oberguggenberger, J. Toft, J. Vindas, P. Wahlberg (Eds.), Generalized Functions and Fourier Analysis, in: Operator Theory: Advances and Applications, vol. 260, Birkhäuser, 2017, pp. 215–235.
[44] J. Toft, Images of function and distribution spaces under the Bargmann transform, J. Pseudo-Differ. Oper. Appl. 8 (2017) 83–139.
[45] J. Toft, R. Üster, Pseudo-differential operators on Orlicz modulation spaces, J. Pseudo-Differ. Oper. Appl. 14 (2023) 6.
[46] J. Toft, R. Üster, E. Nabizadeh, S. Öztop, Continuity and Bargmann mapping properties of quasi-Banach Orlicz modulation spaces, Forum Math. 34 (2022) 1205–1232.
[47] G. Tranquilli, Global normal forms and global properties in function spaces for second order Shubin type operators, PhD Thesis, 2013.