The Journal of Geometric Analysis (2024) 34:159
https://doi.org/10.1007/s12220-024-01601-0

Mean-Dispersion Principles and the Wigner Transform

Chiara Boiti[1] · David Jornet[2] · Alessandro Oliaro[3]

Received: 15 November 2023 / Accepted: 23 February 2024 / Published online: 6 April 2024

Abstract

Given a function $f \in L^2(\mathbb{R})$, we consider means and variances associated to f and its Fourier transform $\hat{f}$, and explore their relations with the Wigner transform $W(f)$, obtaining, as particular cases, a simple new proof of Shapiro's mean-dispersion principle, as well as a stronger result due to Jaming and Powell. Uncertainty principles for orthonormal sequences in $L^2(\mathbb{R})$ involving linear partial differential operators with polynomial coefficients and the Wigner distribution, or different Cohen class representations, are obtained, and an extension to the case of Riesz bases is studied.

Keywords Mean-dispersion principle · Wigner transform · Uncertainty principle · Orthonormal systems · Hermite functions

Mathematics Subject Classification Primary 42B10 · 42C05; Secondary 33C45 · 33C50

1 Introduction

This paper treats uncertainty principles for families of orthonormal functions in $L^2(\mathbb{R})$ in connection with time-frequency analysis. When talking about uncertainty principles, in harmonic analysis, one refers to a class of theorems giving limitations on how much

✉ Alessandro Oliaro
alessandro.oliaro@unito.it

Chiara Boiti
chiara.boiti@unife.it

David Jornet
djornet@mat.upv.es

1 Dipartimento di Matematica e Informatica, Università di Ferrara, Via Machiavelli n. 30, 44121 Ferrara, Italy

2 Instituto Universitario de Matemática Pura y Aplicada IUMPA, Universitat Politècnica de València, Camino de Vera, s/n, 46071 Valencia, Spain

3 Dipartimento di Matematica, Università di Torino, Via Carlo Alberto n. 10, 10123 Turin, Italy

a function and its Fourier transform can be both localized at the same time. Different meanings of the word "localized" give rise to different uncertainty principles. For instance, referring to the most classical results (see [12] for a survey), in the Heisenberg uncertainty principle the localization of f and its Fourier transform $\hat{f}$ has to do with their associated variances, in Benedicks [1] it has to do with the measure of their supports, in Donoho-Stark [10] with the concept of ε-concentration, in Hardy [15] with (exponential) decay at infinity, and so on. Over the years many other authors studied uncertainty principles involving various time-frequency distributions (see, for instance, [2, 5, 9, 11, 13, 14]). There are, moreover, uncertainty principles giving not only limitations on the localization of a single function and its Fourier transform, but on how such limitations behave, becoming stronger and stronger, when adding more and more elements of an orthonormal system in L^2. In this paper we focus in particular on results of this type involving means and variances. For $f \in L^2(\mathbb{R})$ we define the *associated mean*

$$\mu(f) := \frac{1}{\|f\|^2} \int_{\mathbb{R}} t|f(t)|^2 dt \tag{1.1}$$

and the *associated variance*

$$\Delta^2(f) := \frac{1}{\|f\|^2} \int_{\mathbb{R}} |t - \mu(f)|^2 |f(t)|^2 dt; \tag{1.2}$$

observe that, for $\|f\|_2 = 1$, such quantities are the mean and the variance of $|f|^2$. The *dispersion* associated with f is $\Delta(f) := \sqrt{\Delta^2(f)}$.

Our aim is to study uncertainty principles of mean-dispersion type involving quadratic time-frequency representations applied to the elements of an orthonormal system in $L^2(\mathbb{R})$. In order to state our main results we need some basic definitions. The classical cross-Wigner distribution is defined as

$$W(f, g)(x, \xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f\left(x + \frac{t}{2}\right) \overline{g\left(x - \frac{t}{2}\right)} e^{-it\xi} dt, \quad f, g \in L^2(\mathbb{R}), \tag{1.3}$$

and we set for convenience $W(f) := W(f, f)$. Let moreover $\hat{L}$ be the linear partial differential operator in $\mathbb{R}^2$ defined as

$$\hat{L} := \left(\frac{1}{2} D_\xi + x\right)^2 + \left(\frac{1}{2} D_x - \xi\right)^2. \tag{1.4}$$

The following result (that we prove in Theorem 4.4 and Corollary 4.5 below) constitutes a Mean-Dispersion uncertainty principle associated to the Wigner transform.

Theorem 1.1 *Let* $\{f_k\}_{k\in\mathbb{N}_0}$ *be an orthonormal sequence in* $L^2(\mathbb{R})$*. Then for every* $n \geq 0$

$$\sum_{k=0}^{n} \langle \hat{L}W(f_k), W(f_k)\rangle \geq (n+1)^2, \tag{1.5}$$

where as usual $\langle\cdot,\cdot\rangle$ *indicates the inner product in* L^2 *(see Sect. 3 for a discussion on the domain of* $\hat{L}$ *and the corresponding meaning of* $\langle \hat{L}W(f_k), W(f_k)\rangle$*). Equality in* (1.5) *holds for every* $0 \leq n \leq n_0$*,* $n_0 \in \mathbb{N}_0$*, if and only if there exist* $c_k \in \mathbb{C}$ *with* $|c_k| = 1$ *such that* $f_k = c_k h_k$*,* $k = 0, \dots, n_0$*, where* h_k *are the Hermite functions* (1.8).

This result comes motivated by an uncertainty principle for orthonormal sequences due to Shapiro. We shall use throughout the paper the notation $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$, and adopt the following normalization of the Fourier transform:

$$\hat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(t) e^{-it\xi}\, dt, \qquad \xi \in \mathbb{R}. \tag{1.6}$$

Theorem 1.2 (Shapiro's Mean-Dispersion Principle) *There does not exist an infinite orthonormal sequence* $\{f_k\}_{k\in\mathbb{N}_0}$ *in* $L^2(\mathbb{R})$ *such that all* $\mu(f_k)$*,* $\mu(\hat{f}_k)$*,* $\Delta(f_k)$*,* $\Delta(\hat{f}_k)$ *are uniformly bounded.*

This theorem appeared in an unpublished manuscript of Shapiro from 1991; in [19] a stronger result has been proved, namely, there does not exist an orthonormal basis $\{f_k\}_{k\in\mathbb{N}_0}$ of $L^2(\mathbb{R})$ such that

$$\Delta(f_k),\ \Delta(\hat{f}_k),\ \mu(f_k)$$

are uniformly bounded, while there exists an orthonormal basis $\{f_k\}_{k\in\mathbb{N}_0}$ of $L^2(\mathbb{R})$ such that

$$\mu(f_k),\ \mu(\hat{f}_k),\ \Delta(f_k)$$

are uniformly bounded (see also [18]). Moreover, the following quantitative version of Shapiro's Mean-Dispersion Principle is proved in [16].

Theorem 1.3 [16, Theorem 2.3] *Let* $\{f_k\}_{k\in\mathbb{N}_0}$ *be an orthonormal sequence in* $L^2(\mathbb{R})$*. Then for every* $n \geq 0$

$$\sum_{k=0}^{n} \left(\Delta^2(f_k) + \Delta^2(\hat{f}_k) + |\mu(f_k)|^2 + |\mu(\hat{f}_k)|^2 \right) \geq (n+1)^2. \tag{1.7}$$

Equality holds for every $0 \leq n \leq n_0$*,* $n_0 \in \mathbb{N}_0$*, if and only if there exist* $c_k \in \mathbb{C}$ *with* $|c_k| = 1$ *such that* $f_k = c_k h_k$ *for* $k = 0, \dots, n_0$*, where* h_k *are the Hermite functions* *on $\mathbb{R}$ defined as follows:*

$$h_k(t) = \frac{1}{(2^k k! \sqrt{\pi})^{1/2}} e^{-t^2/2} H_k(t), \qquad t \in \mathbb{R}, \tag{1.8}$$

where H_k is the Hermite polynomial of degree k given by

$$H_k(t) = (-1)^k e^{t^2} \frac{d^k}{dt^k} e^{-t^2}, \qquad t \in \mathbb{R}.$$

Observe that (1.7) differs for a constant from the result in [16], due to a different normalization of the Fourier transform. Theorem 1.2 is an easy consequence of Theorem 1.3; moreover, Theorem 1.3 also says that the limitation on the concentration of f_k and $\hat{f}_k$ become stronger and stronger by adding more and more elements from the orthonormal system, as the lower bound $(n+1)^2$ increases faster than the number of involved functions.

We show that Theorem 1.1 implies Theorem 1.3 (and then also Theorem 1.2), and in this sense it can be interpreted as a Mean-Dispersion principle associated to the Wigner transform. Theorems 1.1 and 1.3 are in fact intimately connected, and this is linked to a deep relation between (1.4) and the Hermite operator; such a relation has many consequences and applications in different fields, see for instance [8] where dispersive estimates of the wave flow for the twisted Laplacian (and the Hermite operator) are investigated, or [3, 6, 22] where it has been used to study global regularity properties of partial differential equations. The advantage of the formulation of Theorem 1.1 on the side of the Wigner transform is twofold. First, the proof is simpler than the one of Theorem 1.3 in [16]. In particular, it does not need the Rayleigh-Ritz technique used there. Moreover, $\hat{L}$ is not the only operator that can be used in (1.5) in order to have Mean-Dispersion principles of the kind of Theorem 1.1. In Sects. 4 and 5 we give more details on this fact. Here, we just point out that we can use instead of $\hat{L}$ the multiplication operator by $x^2+\xi^2$, obtaining that (see Theorem 5.1 below) if $\{f_k\}_{k\in\mathbb{N}_0}$ is an orthonormal sequence in $L^2(\mathbb{R})$, then for every $n \geq 0$

$$\sum_{k=0}^{n} \int_{\mathbb{R}^2} (x^2+\xi^2)|W(f_k)(x,\xi)|^2 dx d\xi \geq \frac{(n+1)^2}{2}, \tag{1.9}$$

and equality is characterized as in Theorem 1.1. We show that if f_k satisfies $\mu(f_k) = \mu(\hat{f}_k) = 0$, then the quantity

$$\int_{\mathbb{R}^2} (x^2+\xi^2)|W(f_k)(x,\xi)|^2 dx d\xi$$

is the trace of the covariance matrix of $|W(f_k)(x,\xi)|^2$; then, comparing (1.9) with (1.7) (in the case $\mu(f_k) = \mu(\hat{f}_k) = 0$) we observe that we have replaced the two variances associated with f_k and $\hat{f}_k$ in (1.7), with (a constant times) the trace of the covariance matrix associated with $W(f_k)$, which reflects the fact that $W(f_k)$ includes at the same time both information on f_k and on $\hat{f}_k$.

Other extensions of Theorem 1.1 are also studied. Since there are many different time-frequency representations besides the classical Wigner, we consider the so-called *Cohen class*, given by all the representations $Q(f,g)$ of the form

$$Q(f,g) = \frac{1}{\sqrt{2\pi}} \sigma * W(f,g), \qquad \sigma \in \mathcal{S}'(\mathbb{R}^2),\ f,g \in \mathcal{S}(\mathbb{R}); \tag{1.10}$$

such class contains all the most used time-frequency representations. A natural question is if in Theorem 1.1 one can substitute $W(f_k)$ with $Q(f_k) := Q(f_k, f_k)$, and which operators can be considered instead of $\hat{L}$. We prove in Sect. 6 that for a suitable class of *kernels* σ in (1.10) a result of the kind of Theorem 1.1 can be formulated for representations Q in the Cohen class. Finally, the Mean-Dispersion principle for the Wigner transform can be extended to Riesz bases instead of orthonormal bases.

The paper is organized as follows. In Sects. 2 and 3 we give basic results on the Wigner transform and on the action of the Wigner transform on Hermite functions. In Sect. 4 we prove Theorem 1.1. Section 5 is devoted to the study of the case of the covariance matrix associated with $W(f_k)$ and to the proof of (1.9). In Sects. 6 and 7 we extend the results to the Cohen class and Riesz bases.

2 The Wigner Distribution

Besides the classical cross-Wigner distribution $W(f,g)$ for $f,g \in L^2(\mathbb{R})$ defined in (1.3) we also consider the following Wigner-like transform introduced in [6]

$$\mathrm{Wig}[u](x,\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} u\left(x + \frac{t}{2}, x - \frac{t}{2}\right) e^{-it\xi}\, dt, \qquad u \in L^2(\mathbb{R}^2),$$

with standard extensions to $f,g \in \mathcal{S}'(\mathbb{R})$ and $u \in \mathcal{S}'(\mathbb{R}^2)$. Such operators are strictly related since

$$W(f,g) = \mathrm{Wig}[f \otimes \bar{g}].$$

However, the second one has the advantage, with respect to the classical Wigner transform, that

$$\mathrm{Wig} : \mathcal{S}(\mathbb{R}^2) \longrightarrow \mathcal{S}(\mathbb{R}^2)$$
$$\mathrm{Wig} : \mathcal{S}'(\mathbb{R}^2) \longrightarrow \mathcal{S}'(\mathbb{R}^2)$$

is a linear invertible operator, being composition of a linear invertible change of variables and a partial Fourier transform. Indeed, denoting by $\mathcal{F}(f)(\xi) = \hat{f}(\xi)$ the classical Fourier transform (1.6), by

$$\mathcal{F}_2(u)(x,\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} u(x,t) e^{-it\xi}\, dt, \qquad (x,\xi) \in \mathbb{R}^2,$$

the partial Fourier transform with respect to the second variable, and by

$$\tau_s u(x,t) = u\left(x + \frac{t}{2}, x - \frac{t}{2}\right),$$

we have that

$$\mathrm{Wig}[u] = \mathcal{F}_2 \tau_s u.$$

The inverses of the operators above are

$$\mathcal{F}^{-1}(F)(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} F(\xi) e^{ix\xi} d\xi$$

and

$$\tau_s^{-1} F(x,t) = F\left(\frac{x+t}{2}, x - t\right).$$

Moreover, denoting by

$$\begin{aligned} M_1 u(x,y) &= xu(x,y), \quad M_2 u(x,y) = yu(x,y), \\ D_1 u(x,y) &= D_x u(x,y), \quad D_2 u(x,y) = D_y u(x,y), \end{aligned}$$

for $D_x = -i\partial_x$ and $D_y = -i\partial_y$, a straightforward computation (see also [6]) shows that

$$D_1 \mathrm{Wig}[u] = \mathrm{Wig}[(D_1 + D_2)u] \tag{2.1}$$

$$D_2 \mathrm{Wig}[u] = \mathrm{Wig}[(M_2 - M_1)u] \tag{2.2}$$

$$M_1 \mathrm{Wig}[u] = \mathrm{Wig}\left[\frac{1}{2}(M_1 + M_2)u\right] \tag{2.3}$$

$$M_2 \mathrm{Wig}[u] = \mathrm{Wig}\left[\frac{1}{2}(D_1 - D_2)u\right] \tag{2.4}$$

for all $u \in \mathcal{S}(\mathbb{R}^2)$.

We write M and D for the multiplication and differentiation operators when just one variable is involved, so for $u \in \mathcal{S}(\mathbb{R})$

$$Mu(t) = tu(t), \quad Du(t) = -iu'(t).$$

Moreover, we also adopt, for convenience, the following notations. First, we write $\langle \cdot, \cdot \rangle$ to indicate both the inner product in L^2, the duality $\mathcal{S}'$-$\mathcal{S}$ (we consider here distributions as conjugate-linear functionals), and in general the integral

$$\langle g, h \rangle = \int_{\mathbb{R}} g(t) \overline{h(t)}\, dt$$

each time such integral is finite, even though g, h are not L^2 functions (in particular, we use $\langle \cdot, \cdot \rangle$ for the duality $L^2_v - L^2_{1/v}$ between weighted L^2-spaces with weight v). Second, we write

$$\langle D^n f, D^m g \rangle \tag{2.5}$$

for the integral

$$\int_{\mathbb{R}} \xi^{n+m} \hat{f}(\xi)\overline{\hat{g}(\xi)}\, d\xi \tag{2.6}$$

when the last one makes sense and is finite. It coincides with

$$\int_{\mathbb{R}} D^n f(t)\overline{D^m g(t)}\, dt$$

if $D^n f, D^m g \in L^2$ by *Parseval's formula*

$$\langle f, g \rangle = \langle \hat{f}, \hat{g} \rangle, \qquad \forall f, g \in L^2(\mathbb{R}). \tag{2.7}$$

We use the symbol $\langle \cdot, \cdot \rangle$ with analogous meaning in dimension greater than 1.

With this notation, formulas (2.1)-(2.4) hold also for $u \in \mathcal{S}'(\mathbb{R}^2)$. Let us prove, for instance, (2.1). Since it's valid in $\mathcal{S}(\mathbb{R}^2)$, then for all $u, \varphi \in \mathcal{S}(\mathbb{R}^2)$:

$$\begin{aligned}\langle D_1 \mathrm{Wig}[u], \varphi \rangle &= \langle \mathrm{Wig}[(D_1 + D_2)u], \varphi \rangle = \langle \mathcal{F}_2 \tau_s (D_1 + D_2)u, \varphi \rangle \\ &= \langle \tau_s (D_1 + D_2)u, \mathcal{F}_2^{-1}\varphi \rangle = \langle (D_1 + D_2)u, \tau_s^{-1}\mathcal{F}_2^{-1}\varphi \rangle \\ &= \langle u, (D_1 + D_2)(\tau_s^{-1}\mathcal{F}_2^{-1}\varphi) \rangle \end{aligned} \tag{2.8}$$

by Parseval's formula and

$$\langle \tau_s u, \tau_s v \rangle = \langle u, v \rangle, \qquad \forall u, v \in L^2(\mathbb{R}^2). \tag{2.9}$$

On the other hand, for all $u, \varphi \in \mathcal{S}(\mathbb{R}^2)$,

$$\langle D_1 \mathrm{Wig}[u], \varphi \rangle = \langle \mathrm{Wig}[u], D_1\varphi \rangle = \langle \mathcal{F}_2 \tau_s u, D_1\varphi \rangle = \langle u, \tau_s^{-1}\mathcal{F}_2^{-1}(D_1\varphi) \rangle,$$

which yields, together with (2.8),

$$\tau_s^{-1}\mathcal{F}_2^{-1}(D_1\varphi) = (D_1 + D_2)(\tau_s^{-1}\mathcal{F}_2^{-1}\varphi).$$

Therefore, if $u \in \mathcal{S}'(\mathbb{R}^2)$ and $\varphi \in \mathcal{S}(\mathbb{R}^2)$:

$$\begin{aligned}\langle D_1 \mathrm{Wig}[u], \varphi \rangle &= \langle \mathrm{Wig}[u], D_1\varphi \rangle = \langle \mathcal{F}_2 \tau_s u, D_1\varphi \rangle = \langle u, \tau_s^{-1}\mathcal{F}_2^{-1}(D_1\varphi) \rangle \\ &= \langle u, (D_1 + D_2)(\tau_s^{-1}\mathcal{F}_2^{-1}\varphi) \rangle = \langle (D_1 + D_2)u, \tau_s^{-1}\mathcal{F}_2^{-1}\varphi \rangle \\ &= \langle \mathrm{Wig}[(D_1 + D_2)u], \varphi \rangle,\end{aligned}$$

so that (2.1) is valid also for $u \in \mathcal{S}'(\mathbb{R}^2)$.

Similarly also (2.2)-(2.4) hold for $u \in \mathcal{S}'(\mathbb{R}^2)$.

More generally, we have the following result (proved in [3] for $u \in \mathcal{S}(\mathbb{R}^2)$):

Proposition 2.1 *Let $P(x, y, D_x, D_y)$ be a linear partial differential operator with polynomial coefficients. Then for all $u \in \mathcal{S}'(\mathbb{R}^2)$:*

$$\begin{aligned} &P(M_1, M_2, D_1, D_2)\mathrm{Wig}[u] \\ &\quad = \mathrm{Wig}\left[P\left(\frac{1}{2}(M_1 + M_2), \frac{1}{2}(D_1 - D_2), D_1 + D_2, M_2 - M_1\right)u\right], \quad (2.10) \\ &\mathrm{Wig}[P(M_1, M_2, D_1, D_2)u] \\ &\quad = P\left(M_1 - \frac{1}{2}D_2, M_1 + \frac{1}{2}D_2, \frac{1}{2}D_1 + M_2, \frac{1}{2}D_1 - M_2\right)\mathrm{Wig}[u]. \quad (2.11) \end{aligned}$$

The above proposition will be useful to relate the classical Wigner distribution $W(f)$ to the mean (1.1) and the variance (1.2) associated with a function $f \in L^2(\mathbb{R})$ and its Fourier transform $\hat{f} \in L^2(\mathbb{R})$.

Proposition 2.2 *Given $f \in L^2(\mathbb{R})$ with $\|f\| = 1$ and finite associated means and variances of f and $\hat{f}$, the following properties hold:*

(a) $\langle M^2 f, f\rangle = \mu^2(f) + \Delta^2(f)$
(b) $\langle D^2 f, f\rangle = \mu^2(\hat{f}) + \Delta^2(\hat{f})$
(c) $\langle M_1 W(f), W(f)\rangle = \mu(f)$
(d) $\langle M_2 W(f), W(f)\rangle = \mu(\hat{f})$
(e) $\langle D_1 W(f), W(f)\rangle = 0$
(f) $\langle D_2 W(f), W(f)\rangle = 0$
(g) $\langle D_1^2 W(f), W(f)\rangle = 2\Delta^2(\hat{f})$
(h) $\langle D_2^2 W(f), W(f)\rangle = 2\Delta^2(f)$
(i) $\langle M_1 D_1 W(f), W(f)\rangle = \frac{i}{2}$
$\langle D_1 M_1 W(f), W(f)\rangle = -\frac{i}{2}$
(j) $\langle M_2 D_2 W(f), W(f)\rangle = \frac{i}{2}$
$\langle D_2 M_2 W(f), W(f)\rangle = -\frac{i}{2}$
(k) $\langle M_1^2 W(f), W(f)\rangle = \mu^2(f) + \frac{1}{2}\Delta^2(f)$
(l) $\langle M_2^2 W(f), W(f)\rangle = \mu^2(\hat{f}) + \frac{1}{2}\Delta^2(\hat{f})$.

Proof Let us first recall that (2.7) and (2.9) imply the following *Moyal's formula* for the cross-Wigner distribution (cf. [13, p. 66])

$$\langle W(f_1, g_1), W(f_2, g_2)\rangle = \langle f_1, f_2\rangle\overline{\langle g_1, g_2\rangle}, \qquad \forall f_1, f_2, g_1, g_2 \in L^2(\mathbb{R}). \quad (2.12)$$

Note that the assumption that f has finite associated mean and variance implies that $Mf \in L^2(\mathbb{R})$. Indeed,

$$\begin{aligned}\langle Mf, Mf\rangle &= \int_{\mathbb{R}} y^2|f(y)|^2 dy = \int_{\mathbb{R}} (y-\mu(f)+\mu(f))^2|f(y)|^2 dy \\ &= \int_{\mathbb{R}} |y-\mu(f)|^2|f(y)|^2 dy \\ &\quad +2\mu(f)\int_{\mathbb{R}} (y-\mu(f))|f(y)|^2 dy + \mu^2(f)\|f\|^2 \\ &= \|f\|^2\Delta^2(f) + 2\mu^2(f)\|f\|^2 - 2\mu^2(f)\|f\|^2 + \mu^2(f)\|f\|^2 \\ &= \Delta^2(f) + \mu^2(f), \end{aligned} \tag{2.13}$$

which is the well-known relation between the variance of a random variable and the expectations of the randon variable and its square.

In the same way, the fact that $\hat{f}$ has finite associated mean and variance implies that $Df \in L^2(\mathbb{R})$. This means that Moyal's formula (2.12) can be applied when, in its left-hand side, Mf or Df appear in the arguments of the Wigner transform.

Now we analyze the case when in the left-hand side of (2.12) the expression $W(f, M^2g)$ appears, for $f, g \in L^2(\mathbb{R})$ with finite associated means and variances of $f, g, \hat{f}, \hat{g}$.

From (2.3) we have that

$$2M_1W(f,g)(x,\xi) = W(f, Mg)(x,\xi) + W(Mf, g)(x,\xi) \in L^2(\mathbb{R}^2),$$

since $Mf, Mg \in L^2(\mathbb{R})$. Therefore $W(f,g) \in L^2_v(\mathbb{R}^2)$ for $v(x,\xi) = 1+|x|$. On the other hand, for $f, g \in \mathcal{S}(\mathbb{R})$,

$$\begin{aligned} W(f, M^2g)(x,\xi) &= \int_{\mathbb{R}} f\left(x+\frac{t}{2}\right)\overline{\left(x-\frac{t}{2}\right)^2 g\left(x-\frac{t}{2}\right)} e^{-it\xi}\, dt \\ &= \int_{\mathbb{R}} \left[2x - \left(x+\frac{t}{2}\right)\right] f\left(x+\frac{t}{2}\right)\overline{\left(x-\frac{t}{2}\right) g\left(x-\frac{t}{2}\right)} e^{-it\xi}\, dt \\ &= 2xW(f, Mg)(x,\xi) - W(Mf, Mg)(x,\xi). \end{aligned}$$

Such an equality holds in fact for $f, g \in \mathcal{S}'(\mathbb{R})$ and for tempered distributions it reads

$$W(f, M^2g) = 2M_1W(f, Mg) - W(Mf, Mg). \tag{2.14}$$

Being $Mf, Mg \in L^2(\mathbb{R})$, it follows from (2.14) that $W(f, M^2g) \in L^2_{1/v}(\mathbb{R}^2)$. Then, by the duality between $L^2_{1/v}(\mathbb{R}^2)$ and $L^2_v(\mathbb{R}^2)$, we have that

$$\langle W(f, M^2g), W(f,g)\rangle = \int_{\mathbb{R}^2} W(f, M^2g)(x,\xi)\overline{W(f,g)(x,\xi)}\, dx\, d\xi$$

is convergent.

Using again (2.14) and (2.3), we can then write

$$
\begin{aligned}
\langle W(f, M^2g), W(f,g)\rangle &= \langle 2M_1W(f,Mg), W(f,g)\rangle - \langle W(Mf,Mg), W(f,g)\rangle \\
&= \langle W(f,Mg), 2M_1W(f,g)\rangle - \langle W(Mf,Mg), W(f,g)\rangle \\
&= \langle W(f,Mg), W(Mf,g)\rangle + \langle W(f,Mg), W(f,Mg)\rangle \\
&\quad -\langle W(Mf,Mg), W(f,g)\rangle.
\end{aligned}
$$

Now, all functions are in $L^2(\mathbb{R})$ and applying Moyal's formula (2.12) we finally get

$$
\begin{aligned}
\langle W(f, M^2g), W(f,g)\rangle &= \langle f, Mf\rangle\overline{\langle Mg, g\rangle} + \langle f, f\rangle\overline{\langle Mg, Mg\rangle} - \langle Mf, f\rangle\overline{\langle Mg, g\rangle} \\
&= \langle f, f\rangle\overline{\langle M^2g, g\rangle}. \qquad (2.15)
\end{aligned}
$$

Recall now that for every $u_1, u_2 \in \mathcal{S}'(\mathbb{R})$ the following formula holds

$$
W(\hat{u}_1, \hat{u}_2)(x,\xi) = W(u_1,u_2)(-\xi, x); \qquad (2.16)
$$

then, since $\hat{f}$ and $\hat{g}$ have finite associated means and variances, the same procedure can be applied when we have $W(f, D^2g)$ instead of $W(f, M^2g)$ obtaining that, with the notation (2.5)-(2.6),

$$
\langle W(f, D^2g), W(f,g)\rangle = \langle W(\hat{f}, M^2\hat{g}), W(\hat{f},\hat{g})\rangle = \langle f, f\rangle\overline{\langle D^2g, g\rangle}. \quad (2.17)
$$

Similar considerations can be done for MDf, obtaining that $W(MDf, g) \in L^2_{1/v}(\mathbb{R}^2)$ because

$$
\begin{aligned}
W(MDf, g) &= \int \left(x+\frac{t}{2}\right) Df\left(x+\frac{t}{2}\right)\overline{g\left(x-\frac{t}{2}\right)}e^{-it\xi}dt \\
&= \int \left[2x - \left(x-\frac{t}{2}\right)\right] Df\left(x+\frac{t}{2}\right)\overline{g\left(x-\frac{t}{2}\right)}e^{-it\xi}dt \\
&= 2xW(Df, g) - W(Df, Mg),
\end{aligned}
$$

with $Df, Mg \in L^2(\mathbb{R})$ under the assumptions of finite associated means and variances. Arguing as for M^2f we then have

$$
\langle W(MDf, g), W(f,g)\rangle = \langle MDf, f\rangle\overline{\langle g, g\rangle} = \langle Df, Mf\rangle\overline{\langle g, g\rangle}. \qquad (2.18)
$$

All the above considerations will be implicit from now on.

Let us now prove point (a): it follows from (2.13) since $\langle M^2f, f\rangle = \langle Mf, Mf\rangle$.

(b): Let us first remark that by Parseval's formula (2.7) and the assumption $\|f\| = 1$ we also have that $\|\hat{f}\| = 1$. With the notations (2.5)-(2.6), by point (a) applied to $\hat{f}$ we thus obtain:

$$
\langle D^2f, f\rangle = \langle \xi^2\hat{f}, \hat{f}\rangle = \mu^2(\hat{f}) + \Delta^2(\hat{f}).
$$

(c): From (2.3) and Moyal's formula (2.12):

$$\begin{aligned}\langle M_1 W(f), W(f)\rangle &= \langle M_1 \mathrm{Wig}[f \otimes \bar{f}], W(f)\rangle \\ &= \langle \mathrm{Wig}[\frac{1}{2}(M_1 + M_2)(f \otimes \bar{f})], W(f)\rangle \\ &= \frac{1}{2}(\langle W(Mf, f), W(f, f)\rangle + \langle W(f, Mf), W(f, f)\rangle) \\ &= \frac{1}{2}(\langle Mf, f\rangle\overline{\langle f, f\rangle} + \langle f, f\rangle\overline{\langle Mf, f\rangle}) = \mu(f),\end{aligned}$$

since $\mu(f) \in \mathbb{R}$.

(d): From (2.4), Moyal's and Parseval's formulas (2.12) and (2.7):

$$\begin{aligned}\langle M_2 W(f), W(f)\rangle &= \langle \mathrm{Wig}[\frac{1}{2}(D_1 - D_2) f \otimes \bar{f}], W(f)\rangle \\ &= \frac{1}{2}(\langle W(Df, f), W(f, f)\rangle + \langle W(f, Df), W(f, f)\rangle) \\ &= \frac{1}{2}(\langle Df, f\rangle\overline{\langle f, f\rangle} + \langle f, f\rangle\overline{\langle Df, f\rangle}) \\ &= \frac{1}{2}(\langle \xi\hat{f}, \hat{f}\rangle + \overline{\langle \xi\hat{f}, \hat{f}\rangle}) = \mu(\hat{f}),\end{aligned}$$

since $\mu(\hat{f}) \in \mathbb{R}$.

(e): From (2.1), (2.12) and (2.7):

$$\begin{aligned}\langle D_1 W(f), W(f)\rangle &= \langle \mathrm{Wig}[(D_1 + D_2) f \otimes \bar{f}], W(f)\rangle \\ &= \langle W(Df, f) - W(f, Df), W(f, f)\rangle \\ &= \langle Df, f\rangle\overline{\langle f, f\rangle} - \langle f, f\rangle\overline{\langle Df, f\rangle} \\ &= \langle \xi\hat{f}, \hat{f}\rangle - \overline{\langle \xi\hat{f}, \hat{f}\rangle} = 0.\end{aligned}$$

(f): From (2.2) and (2.12):

$$\begin{aligned}\langle D_2 W(f), W(f)\rangle &= \langle \mathrm{Wig}[(M_2 - M_1) f \otimes \bar{f}], W(f)\rangle \\ &= \langle W(f, Mf) - W(Mf, f), W(f, f)\rangle \\ &= \langle f, f\rangle\overline{\langle Mf, f\rangle} - \langle Mf, f\rangle\overline{\langle f, f\rangle} = 0.\end{aligned}$$

(g): From (2.1), (2.12), (2.17), (2.7) and point (a):

$$\begin{aligned}\langle D_1^2 W(f), W(f)\rangle &= \langle \mathrm{Wig}[(D_1 + D_2)^2 f \otimes \bar{f}], W(f)\rangle \\ &= \langle W(D^2 f, f) - 2W(Df, Df) + W(f, D^2 f), W(f, f)\rangle \\ &= \langle D^2 f, f\rangle\overline{\langle f, f\rangle} - 2\langle Df, f\rangle\overline{\langle Df, f\rangle} + \langle f, f\rangle\overline{\langle D^2 f, f\rangle} \\ &= \langle \xi^2\hat{f}, \hat{f}\rangle - 2|\langle \xi\hat{f}, \hat{f}\rangle|^2 + \overline{\langle \xi^2\hat{f}, \hat{f}\rangle} \\ &= 2(\mu^2(\hat{f}) + \Delta^2(\hat{f})) - 2\mu^2(\hat{f}) = 2\Delta^2(\hat{f}).\end{aligned}$$

(h): From (2.2), (2.12), (2.15) and point (a):

$$\begin{aligned}\langle D_2^2 W(f), W(f)\rangle &= \langle \mathrm{Wig}[(M_2 - M_1)^2 f \otimes \bar{f}], W(f)\rangle \\ &= \langle W(f, M^2 f) - 2W(Mf, Mf) + W(M^2 f, f), W(f, f)\rangle \\ &= \langle f, f\rangle \overline{\langle M^2 f, f\rangle} - 2\langle Mf, f\rangle \overline{\langle Mf, f\rangle} + \langle M^2 f, f\rangle \overline{\langle f, f\rangle} \\ &= 2(\mu^2(f) + \Delta^2(f)) - 2\mu^2(f) = 2\Delta^2(f).\end{aligned}$$

(i): From (2.1), (2.3), (2.12), (2.18) and (2.7):

$$\begin{aligned}&\langle M_1 D_1 W(f), W(f)\rangle \\ &\quad = \left\langle \mathrm{Wig}\left[\frac{1}{2}(M_2 + M_1)(D_1 + D_2) f \otimes \bar{f}\right], W(f)\right\rangle \\ &\quad = \frac{1}{2}\langle \mathrm{Wig}[(M_2 D_1 + M_1 D_1 + M_2 D_2 + M_1 D_2) f \otimes \bar{f}], W(f)\rangle \\ &\quad = \frac{1}{2}\langle W(Df, Mf) + W(MDf, f) - W(f, MDf) - W(Mf, Df), W(f, f)\rangle \\ &\quad = \frac{1}{2}(\langle Df, f\rangle \overline{\langle Mf, f\rangle} + \langle Df, Mf\rangle \overline{\langle f, f\rangle} - \langle f, f\rangle \overline{\langle Df, Mf\rangle} - \langle Mf, f\rangle \overline{\langle Df, f\rangle}) \\ &\quad = \frac{1}{2}(\langle \xi \hat{f}, \hat{f}\rangle \overline{\mu(f)} + \langle Df, Mf\rangle - \overline{\langle Df, Mf\rangle} - \mu(f)\overline{\langle \xi \hat{f}, \hat{f}\rangle}).\end{aligned}$$

Since $\mu(f) \in \mathbb{R}$, $\langle \xi \hat{f}, \hat{f}\rangle = \mu(\hat{f}) \in \mathbb{R}$ and

$$\begin{aligned}\langle Df, Mf\rangle &= \langle f, DMf\rangle = i\langle f, f\rangle + \langle f, MDf\rangle \\ &= i + \langle Mf, Df\rangle = i + \overline{\langle Df, Mf\rangle},\end{aligned} \tag{2.19}$$

we finally have that

$$\langle M_1 D_1 W(f), W(f)\rangle = \frac{i}{2}.$$

Therefore

$$\begin{aligned}\langle D_1 M_1 W(f), W(f)\rangle &= \langle M_1 W(f), D_1 W(f)\rangle = \langle W(f), M_1 D_1 W(f)\rangle \\ &= \overline{\langle M_1 D_1 W(f), W(f)\rangle} = -\frac{i}{2}.\end{aligned}$$

(j): From (2.2), (2.4), (2.12), (2.18), (2.7) and (2.19):

$$
\begin{aligned}
&\langle M_2D_2W(f), W(f)\rangle \\
&= \left\langle \mathrm{Wig}\left[\frac{1}{2}(D_1 - D_2)(M_2 - M_1)f \otimes \bar{f}\right], W(f)\right\rangle \\
&= \frac{1}{2}\langle \mathrm{Wig}[(D_1M_2 - D_2M_2 - D_1M_1 + D_2M_1)f \otimes \bar{f}], W(f)\rangle \\
&= \frac{1}{2}\langle W(Df, Mf) - \frac{1}{i}W(f, f) + W(f, MDf), W(f, f)\rangle \\
&\quad -\frac{1}{2}\langle \frac{1}{i}W(f, f) + W(MDf, f) + W(Mf, Df), W(f, f)\rangle \\
&= \frac{1}{2}(\langle Df, f\rangle\overline{\langle Mf, f\rangle} - \frac{1}{i}\langle f, f\rangle\overline{\langle f, f\rangle} + \langle f, f\rangle\overline{\langle Df, Mf\rangle}) \\
&\quad -\frac{1}{2}\left(\frac{1}{i} + \langle Df, Mf\rangle\overline{\langle f, f\rangle} + \langle Mf, f\rangle\overline{\langle Df, f\rangle}\right) \\
&= \frac{1}{2}\langle \xi\hat{f}, \hat{f}\rangle\mu(f) + i + \frac{1}{2}(\overline{\langle Df, Mf\rangle} - \langle Df, Mf\rangle) - \frac{1}{2}\mu(f)\overline{\langle \xi\hat{f}, \hat{f}\rangle} \\
&= i - \frac{i}{2} = \frac{i}{2}.
\end{aligned}
$$

It follows that

$$\langle D_2M_2W(f), W(f)\rangle = \frac{1}{i}\langle W(f, f), W(f, f)\rangle + \langle M_2D_2W(f), W(f)\rangle = -\frac{i}{2}.$$

(k): From (2.3), (2.12) and point (a):

$$
\begin{aligned}
&\langle M_1^2W(f), W(f)\rangle \\
&= \langle M_1\mathrm{Wig}[f \otimes \bar{f}], M_1\mathrm{Wig}[f \otimes \bar{f}]\rangle \\
&= \left\langle \mathrm{Wig}\left[\frac{1}{2}(M_2 + M_1)f \otimes \bar{f}\right], \mathrm{Wig}\left[\frac{1}{2}(M_2 + M_1)f \otimes \bar{f}\right]\right\rangle \\
&= \frac{1}{4}\langle W(f, Mf) + W(Mf, f), W(f, Mf) + W(Mf, f)\rangle \\
&= \frac{1}{4}(\langle f, f\rangle\overline{\langle Mf, Mf\rangle} + \langle Mf, f\rangle\overline{\langle f, Mf\rangle} + \langle f, Mf\rangle\overline{\langle Mf, f\rangle} + \langle Mf, Mf\rangle\overline{\langle f, f\rangle}) \\
&= \frac{1}{4}(\overline{\langle M^2f, f\rangle} + 2\mu^2(f) + \langle M^2f, f\rangle) \\
&= \frac{1}{4}\cdot 2(\mu^2(f) + \Delta^2(f)) + \frac{1}{2}\mu^2(f) = \frac{1}{2}\Delta^2(f) + \mu^2(f).
\end{aligned}
$$

(l): From (2.4), (2.12), (2.7) and point (b):

$$\langle M_2^2W(f), W(f)\rangle$$

$$\begin{aligned}
&= \langle M_2 \mathrm{Wig}[f \otimes \bar{f}], M_2 \mathrm{Wig}[f \otimes \bar{f}] \rangle \\
&= \left\langle \mathrm{Wig}\left[\frac{1}{2}(D_1 - D_2) f \otimes \bar{f}\right], \mathrm{Wig}\left[\frac{1}{2}(D_1 - D_2) f \otimes \bar{f}\right] \right\rangle \\
&= \frac{1}{4}\langle W(Df, f) + W(f, Df), W(Df, f) + W(f, Df) \rangle \\
&= \frac{1}{4}(\langle Df, Df\rangle \overline{\langle f, f\rangle} + \langle f, Df\rangle \overline{\langle Df, f\rangle} + \langle Df, f\rangle \overline{\langle f, Df\rangle} + \langle f, f\rangle \overline{\langle Df, Df\rangle}) \\
&= \frac{1}{4}(\langle D^2 f, f\rangle + \langle \hat{f}, \xi \hat{f}\rangle \overline{\langle \xi \hat{f}, \hat{f}\rangle} + \langle \xi \hat{f}, \hat{f}\rangle \overline{\langle \hat{f}, \xi \hat{f}\rangle} + \overline{\langle D^2 f, f\rangle}) \\
&= \frac{1}{2}\Delta^2(\hat{f}) + \mu^2(\hat{f}).
\end{aligned}$$

The proof is complete. □

Let us remark that the identities of Proposition 2.2 can be rewritten for any $f \in L^2(\mathbb{R})$, without the restriction that $\|f\| = 1$, by considering $\frac{f}{\|f\|}$ instead of f. For instance, points (a), (b) of Proposition 2.2 would be replaced by

$$\langle M^2 f, f\rangle = \|f\|^2(\mu^2(f) + \Delta^2(f)), \tag{2.20}$$

$$\langle D^2 f, f\rangle = \|f\|^2(\mu^2(\hat{f}) + \Delta^2(\hat{f})). \tag{2.21}$$

3 The Hermite Basis

For $k \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$, let h_k be the Hermite functions on $\mathbb{R}$ defined by (1.8). It is well known that h_k are eigenfunctions of the Fourier transform and form an orthonormal basis in $L^2(\mathbb{R})$. Moreover, they are an absolute basis in $\mathcal{S}(\mathbb{R})$ (see [17]).

Denoting by

$$h_{j,k} := \mathcal{F}^{-1} W(h_j, h_k),$$

by [22, Thms. 3.2 and 3.4] we have that the functions $\{h_{j,k}\}_{j,k\in\mathbb{N}_0}$ form an orthonormal basis in $L^2(\mathbb{R}^2)$ and are eigenfunctions of the twisted Laplacian:

$$L h_{j,k}(y, t) = (2k + 1) h_{j,k}(y, t), \qquad j, k \in \mathbb{N}_0,$$

for

$$L := \left(D_y - \frac{1}{2}t\right)^2 + \left(D_t + \frac{1}{2}y\right)^2.$$

By Fourier transform (see [4, Ex. 3.20])

$$\hat{h}_{j,k}(x, \xi) = W(h_j, h_k)(x, \xi) \tag{3.1}$$

are eigenfunctions of the operator $\hat{L}$ defined in (1.4), with the same eigenvalues as before, in the sense that

$$\hat{L}\hat{h}_{j,k} = (2k+1)\hat{h}_{j,k}. \tag{3.2}$$

Note that also $\{\hat{h}_{j,k}\}_{j,k\in\mathbb{N}_0}$ are in $\mathcal{S}(\mathbb{R}^2)$ and form an orthonormal basis in $L^2(\mathbb{R}^2)$.

More in general, following the same ideas as in [21, Thm. 21.2], we can prove:

Theorem 3.1 *If $\{f_k\}_{k\in\mathbb{N}_0}$ is an orthonormal basis in $L^2(\mathbb{R})$, then $\{W(f_j, f_k)\}_{j,k\in\mathbb{N}_0}$ is an orthonormal basis in $L^2(\mathbb{R}^2)$.*

Proof Let us first remark that if $\{f_k\}_k$ is an othonormal sequence in $L^2(\mathbb{R})$, then $\{W(f_j, f_k)\}_{j,k}$ is an orthonormal sequence in $L^2(\mathbb{R}^2)$ since, by (2.12),

$$\begin{aligned}\langle W(f_j, f_k), W(f_i, f_h)\rangle &= \langle f_j, f_i\rangle\overline{\langle f_k, f_h\rangle}\\ &= \delta_{j,i}\cdot\delta_{k,h} = \begin{cases}1 & \text{if } (j,k)=(i,h)\\ 0 & \text{if } (j,k)\neq(i,h).\end{cases}\end{aligned}$$

In order to prove that $\{W(f_j, f_k)\}_{j,k\in\mathbb{N}_0}$ is a basis for $L^2(\mathbb{R}^2)$, by [7, Thm. 3.4.2], it is enough to prove that if $F\in L^2(\mathbb{R}^2)$ is such that

$$\int_{\mathbb{R}^2} F(x,\xi)W(f_j, f_k)(x,\xi)dxd\xi = 0, \qquad \forall j,k\in\mathbb{N}_0, \tag{3.3}$$

then $F = 0$ a.e. in $\mathbb{R}^2$.

By [21, Thms. 4.4 and 7.5] the operator

$$\begin{aligned}&L^2(\mathbb{R}^2) \longrightarrow \mathcal{L}(L^2(\mathbb{R}), L^2(\mathbb{R}))\\ &F \longmapsto W_F\end{aligned}$$

defined by

$$\langle W_F\varphi, \psi\rangle = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}^2} F(x,\xi)W(\varphi,\psi)(x,\xi)dxd\xi$$

is a bounded linear operator satisfying

$$\|W_F\|_{\mathcal{L}(L^2,L^2)} \le \frac{1}{\sqrt{2\pi}}\|F\|_{L^2(\mathbb{R}^2)} = \|W_F\|_{HS}, \tag{3.4}$$

where $\|\cdot\|_{HS}$ is the Hilbert-Schmidt norm defined by (see [21, formula (7.1)]):

$$\|W_F\|_{HS}^2 := \sum_{j=0}^{+\infty}\|W_F f_j\|_{L^2(\mathbb{R})}^2 \tag{3.5}$$

for an orthonormal basis $\{f_j\}_{j\in\mathbb{N}_0}$ of $L^2(\mathbb{R})$. The operator W_F is in fact the classical Weyl operator with symbol F. Then

$$\langle W_F f_j, f_k\rangle = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}^2} F(x,\xi)W(f_j,f_k)(x,\xi)dxd\xi = 0, \qquad \forall j,k\in\mathbb{N}_0,$$

by assumption, which implies that

$$W_F f_j = 0, \qquad \forall j\in\mathbb{N}_0, \tag{3.6}$$

since $\{f_j\}_{j\in\mathbb{N}_0}$ is an orthonormal basis in $L^2(\mathbb{R})$.

From (3.4) and (3.5) we finally have that $F=0$ a.e. in $\mathbb{R}^2$. □

The operator $\hat{L}$ defined in (1.4) is unbounded on $L^2(\mathbb{R}^2)$ (see Remark 6.6 below) and defined (at least) in $\mathcal{S}(\mathbb{R}^2)\subset L^2(\mathbb{R}^2)$. Now, since the functions (3.1) are an orthonormal basis for $L^2(\mathbb{R}^2)$, every element $F\in L^2(\mathbb{R}^2)$ can be written as

$$F=\sum_{j,k=0}^{+\infty} c_{j,k}\hat{h}_{j,k}$$

where $c_{j,k}=\langle F,\hat{h}_{j,k}\rangle$. Then, writing

$$F_N=\sum_{j,k=0}^{N} c_{j,k}\hat{h}_{j,k}\in\mathcal{S}(\mathbb{R}^2)$$

we have from (3.2)

$$\hat{L}F_N(x,\xi)=\sum_{j,k=0}^{N} c_{j,k}(2k+1)\hat{h}_{j,k}(x,\xi).$$

The operator $\hat{L}$ is then the unbounded and densely defined operator with domain

$$D(\hat{L})=\left\{F\in L^2(\mathbb{R}^2):\sum_{j,k=0}^{+\infty} c_{j,k}(2k+1)\hat{h}_{j,k}\text{ converges in }L^2(\mathbb{R}^2)\right\}$$

for $c_{j,k}=\langle F,\hat{h}_{j,k}\rangle$, acting on $F\in D(\hat{L})$ as

$$\hat{L}F=\sum_{j,k=0}^{+\infty} c_{j,k}(2k+1)\hat{h}_{j,k}\in L^2(\mathbb{R}^2).$$

In this case

$$\langle \hat{L}F, F\rangle = \sum_{j,k=0}^{+\infty} |c_{j,k}|^2(2k+1). \tag{3.7}$$

In general, we shall write (3.7) for all $F \in L^2(\mathbb{R}^2)$, meaning that $\langle \hat{L}F, F\rangle = +\infty$ if the series diverges. Note that, being $\{\hat{h}_{j,k}\}_{j,k\in\mathbb{N}_0}$ an orthonormal basis for $L^2(\mathbb{R}^2)$, we have that $F \in D(\hat{L})$ if and only if $\{c_{j,k}(2k+1)\}_{j,k\in\mathbb{N}_0} \in \ell^2$. This implies that the series (3.7) converges (but not vice versa).

4 Mean-Dispersion Principle

From the results of the previous sections we obtain now an alternative formulation and a simple proof of the Shapiro's Mean-Dispersion Principle (see [16] and the references therein). To this aim let us first prove some preliminary results.

Lemma 4.1 *Let $\{h_k\}_{k\in\mathbb{N}_0}$ be the Hermite functions defined in (1.8) and $\hat{L}$ as in (1.4). Then for every $j \in \mathbb{N}_0$ we have*

$$\sum_{k=0}^{n}\langle \hat{L}W(h_j, h_k), W(h_j, h_k)\rangle = (n+1)^2, \qquad \forall n \in \mathbb{N}_0.$$

Proof From (3.2) for all $j, k \in \mathbb{N}_0$ we have

$$\langle \hat{L}W(h_j, h_k), W(h_j, h_k)\rangle = \langle \hat{L}\hat{h}_{j,k}, \hat{h}_{j,k}\rangle = \langle (2k+1)\hat{h}_{j,k}, \hat{h}_{j,k}\rangle = 2k+1,$$

since $\{\hat{h}_{j,k}\}_{j,k}$ is an orthonormal basis in $L^2(\mathbb{R}^2)$.

It follows that

$$\sum_{k=0}^{n}\langle \hat{L}W(h_j, h_k), W(h_j, h_k)\rangle = \sum_{k=0}^{n}(2k+1) = (n+1)^2,$$

where the last equality is the formula for the sum of all odd numbers from 1 to $2n+1$. □

Lemma 4.2 *Let $\hat{L}$ be the operator in (1.4). Then for all $f, g \in L^2(\mathbb{R})$ with finite associated means and variances of $f, g, \hat{f}, \hat{g}$:*

(i) $\hat{L}W(f, g) = W(f, (M^2 + D^2)g)$,
(ii) $\langle \hat{L}W(f, g), W(f, g)\rangle = \|f\|^2\|g\|^2(\Delta^2(g) + \Delta^2(\hat{g}) + \mu^2(g) + \mu^2(\hat{g}))$.

In particular, $\langle \hat{L}W(f, g), W(f, g)\rangle \in \mathbb{R}$ and if $\|f\| = \|g\| = 1$, then

$$\langle \hat{L}W(f, g), W(f, g)\rangle = \Delta^2(g) + \Delta^2(\hat{g}) + \mu^2(g) + \mu^2(\hat{g}).$$

Proof (i): From (2.10) we have

$$\begin{aligned}\hat{L}W(f,g) &= \left[\left(\frac{1}{2}D_2 + M_1\right)^2 + \left(\frac{1}{2}D_1 - M_2\right)^2\right]\mathrm{Wig}[f\otimes\bar{g}]\\ &= \mathrm{Wig}\left[\left(\left(\frac{1}{2}(M_2 - M_1) + \frac{1}{2}(M_2 + M_1)\right)^2\right.\right.\\ &\qquad \left.\left. + \left(\frac{1}{2}(D_1 + D_2) - \frac{1}{2}(D_1 - D_2)\right)^2\right) f\otimes\bar{g}\right]\\ &= \mathrm{Wig}[(M_2^2 + D_2^2)f\otimes\bar{g}] = W(f,(M^2 + D^2)g).\end{aligned}$$

(ii): From (i), (2.15), (2.17), (2.20) and (2.21):

$$\begin{aligned}\langle\hat{L}W(f,g), W(f,g)\rangle &= \langle W(f,(M^2 + D^2)g), W(f,g)\rangle\\ &= \langle W(f,M^2 g), W(f,g)\rangle + \langle W(f,D^2 g), W(f,g)\rangle\\ &= \langle f,f\rangle\overline{\langle M^2 g,g\rangle} + \langle f,f\rangle\overline{\langle D^2 g,g\rangle}\\ &= \|f\|^2\|g\|^2(\Delta^2(g) + \mu^2(g)) + \|f\|^2\|g\|^2(\Delta^2(\hat{g}) + \mu^2(\hat{g}))\\ &= \|f\|^2\|g\|^2(\Delta^2(g) + \Delta^2(\hat{g}) + \mu^2(g) + \mu^2(\hat{g})).\end{aligned}$$

□

Lemma 4.3 *Let $0 \le \alpha_\ell \le 1$ satisfy $\sum_{\ell=0}^{+\infty}\alpha_\ell = n+1$ for some $n \in \mathbb{N}_0$. Then*

$$\sum_{\ell=0}^{+\infty}\alpha_\ell(2\ell+1) \ge (n+1)^2.$$

Moreover, the inequality is an identity if and only if $\alpha_0 = \cdots = \alpha_n = 1$.

Proof By assumption we can write

$$n+1 = \sum_{\ell=0}^{+\infty}\alpha_\ell = \alpha_0 + \cdots + \alpha_n + R_n$$

for a reminder

$$R_n = \sum_{\ell=n+1}^{+\infty}\alpha_\ell. \tag{4.1}$$

Note that $\alpha_0 = \cdots = \alpha_n = 1$ if $R_n = 0$.

For all $0 \le k \le n$ we set

$$c_k = \begin{cases} 0, & \text{if } R_n = 0\\ \frac{1-\alpha_k}{R_n}, & \text{if } R_n > 0.\end{cases} \tag{4.2}$$

Then

$$\alpha_k + c_k R_n = 1 \qquad \forall 0 \le k \le n \tag{4.3}$$

and $(c_0 + \cdots + c_n)R_n = R_n$, so that

$$c_0 + \cdots + c_n = \begin{cases} 1 & \text{if } R_n > 0 \\ 0 & \text{if } R_n = 0 \end{cases}$$

and we can write

$$(c_0 + \cdots + c_n) \sum_{\ell=n+1}^{+\infty} \alpha_\ell (2\ell + 1) = \sum_{\ell=n+1}^{+\infty} \alpha_\ell (2\ell + 1), \tag{4.4}$$

being $R_n = 0$ iff $\alpha_\ell = 0$ for all $\ell \ge n+1$.

By (4.4) and (4.3)

$$\begin{aligned}
\sum_{\ell=0}^{+\infty} \alpha_\ell (2\ell + 1) &= \sum_{\ell=0}^{n} \alpha_\ell (2\ell + 1) + (c_0 + \cdots + c_n) \sum_{\ell=n+1}^{+\infty} \alpha_\ell (2\ell + 1) \\
&= \sum_{\ell=0}^{n} \alpha_\ell (2\ell + 1) + c_0 \sum_{\ell=n+1}^{+\infty} \alpha_\ell \underbrace{(2\ell + 1)}_{>1} + c_1 \sum_{\ell=n+1}^{+\infty} \alpha_\ell \underbrace{(2\ell + 1)}_{>3} \\
&\quad \cdots + c_{n-1} \sum_{\ell=n+1}^{+\infty} \alpha_\ell \underbrace{(2\ell + 1)}_{>2n-1} + c_n \sum_{\ell=n+1}^{+\infty} \alpha_\ell \underbrace{(2\ell + 1)}_{>2n+1} \\
&\ge \left(\alpha_0 + c_0 \sum_{\ell=n+1}^{+\infty} \alpha_\ell \right) + \left(\alpha_1 \cdot 3 + c_1 \sum_{\ell=n+1}^{+\infty} \alpha_\ell \cdot 3 \right) \\
&\quad \cdots + \left(\alpha_{n-1} \cdot (2n-1) + c_{n-1} \sum_{\ell=n+1}^{+\infty} \alpha_\ell \cdot (2n-1) \right) \\
&\quad + \left(\alpha_n \cdot (2n+1) + c_n \sum_{\ell=n+1}^{+\infty} \alpha_\ell \cdot (2n+1) \right) \\
&= \sum_{k=0}^{n} \underbrace{(\alpha_k + c_k R_n)}_{=1} (2k+1) = \sum_{k=0}^{n} (2k+1) = (n+1)^2.
\end{aligned} \tag{4.5}$$

Finally note that the inequality in (4.5) is strict unless $c_0 = \cdots = c_n = 0$, that is $\alpha_0 = \cdots = \alpha_n = 1$. □

Theorem 4.4 *Let $\{f_k\}_{k\in\mathbb{N}_0}$ be such that $\|f_k\| = 1$ for every $k \in \mathbb{N}_0$, and let $\{g_k\}_{k\in\mathbb{N}_0}$ be an orthonormal sequence in $L^2(\mathbb{R})$. Then*

$$\sum_{k=0}^{n} \langle \hat{L} W(f_k, g_k), W(f_k, g_k)\rangle \geq (n+1)^2, \qquad \forall n \in \mathbb{N}_0. \tag{4.6}$$

Moreover, the estimate is optimal, in the sense that if g_k are the Hermite functions, then equality holds in (4.6) and, conversely, given $n_0 \in \mathbb{N}$, if equality holds in (4.6) for all $n \leq n_0$, then there exist $c_k \in \mathbb{C}$ with $|c_k| = 1$ such that $g_k = c_k h_k$ for all $0 \leq k \leq n_0$.

Proof Since $W(f_k, g_k) \in L^2(\mathbb{R}^2)$ and the sequence $\{\hat{h}_{j,\ell}\} = \{W(h_j, h_\ell)\}$ defined in (3.1) is an orthonormal basis in $L^2(\mathbb{R}^2)$, we can write

$$W(f_k, g_k) = \sum_{j,\ell=0}^{+\infty} c_{j,\ell}^{(k)} W(h_j, h_\ell)$$

with

$$c_{j,\ell}^{(k)} = \langle W(f_k, g_k), W(h_j, h_\ell)\rangle = \langle f_k, h_j\rangle \overline{\langle g_k, h_\ell\rangle}, \tag{4.7}$$

by (2.12). As in (3.7) we have

$$\sum_{k=0}^{n} \langle \hat{L} W(f_k, g_k), W(f_k, g_k)\rangle = \sum_{k=0}^{n} \sum_{j,\ell=0}^{+\infty} |c_{j,\ell}^{(k)}|^2 (2\ell+1), \tag{4.8}$$

and we can assume that for every $0 \leq k \leq n$ the series in (4.8) converges, otherwise (4.6) would be trivial, being the left-hand side equal to $+\infty$.

By (4.7) and (4.8), we get

$$\begin{aligned}\sum_{k=0}^{n} \langle \hat{L} W(f_k, g_k), W(f_k, g_k)\rangle &= \sum_{k=0}^{n} \sum_{j,\ell=0}^{+\infty} |\langle f_k, h_j\rangle|^2 |\langle g_k, h_\ell\rangle|^2 (2\ell+1) \\ &= \sum_{k=0}^{n} \left(\sum_{j=0}^{+\infty} |\langle f_k, h_j\rangle|^2 \right) \sum_{\ell=0}^{+\infty} |\langle g_k, h_\ell\rangle|^2 (2\ell+1) = \sum_{\ell=0}^{+\infty} \left(\sum_{k=0}^{n} |\langle g_k, h_\ell\rangle|^2 \right) (2\ell+1),\end{aligned} \tag{4.9}$$

since $\|f_k\|^2 = 1$. Setting

$$\alpha_\ell := \sum_{k=0}^{n} |\langle g_k, h_\ell\rangle|^2,$$

we remark that

$$\sum_{\ell=0}^{+\infty} \alpha_\ell = \sum_{\ell=0}^{+\infty} \sum_{k=0}^{n} |\langle g_k, h_\ell \rangle|^2 = \sum_{k=0}^{n} \sum_{\ell=0}^{+\infty} |\langle g_k, h_\ell \rangle|^2 = \sum_{k=0}^{n} \|g_k\|^2 = n + 1. \quad (4.10)$$

But for each $\ell \in \mathbb{N}_0$

$$\alpha_\ell \leq \sum_{k=0}^{+\infty} |\langle g_k, h_\ell \rangle|^2 \leq \|h_\ell\|^2 = 1,$$

so that we can apply Lemma 4.3 to obtain from (4.9) that

$$\sum_{k=0}^{n} \langle \hat{L} W(f_k, g_k), W(f_k, g_k) \rangle = \sum_{\ell=0}^{+\infty} \alpha_\ell (2\ell + 1) \geq (n+1)^2,$$

and equality holds if and only if

$$\alpha_\ell = \sum_{k=0}^{n} |\langle g_k, h_\ell \rangle|^2 = 1, \qquad \forall \ell = 0, \ldots, n.$$

If $g_k = h_k$ for every $k \in \mathbb{N}_0$, then $\alpha_\ell = 1$ for all $\ell \in \mathbb{N}_0$ and equality holds in (4.6) for all $n \in \mathbb{N}_0$.

Conversely, if equality holds in (4.6) for all $n \leq n_0$, then

$$\sum_{k=0}^{n} |\langle g_k, h_\ell \rangle|^2 = 1, \qquad \forall \ell = 0, \ldots, n, \ \forall n \leq n_0.$$

This implies $h_\ell \in \operatorname{span}(g_0, \ldots, g_\ell)$ for all $\ell \leq n_0$. Therefore $g_0 = c_0 h_0$ for some $c_0 \in \mathbb{C}$ with $|c_0| = 1$ and then, recursively, $g_k = c_k h_k$ for some $c_k \in \mathbb{C}$ with $|c_k| = 1$ and for $k = 1, \ldots, n_0$. □

Let us remark that we can choose in Theorem 4.4 a constant sequence $f_k = f \in L^2(\mathbb{R})$ with $\|f\| = 1$, obtaining that if $\{g_k\}_{k\in\mathbb{N}_0}$ is an orthonormal sequence in $L^2(\mathbb{R})$, then

$$\sum_{k=0}^{n} \langle \hat{L} W(f, g_k), W(f, g_k) \rangle \geq (n+1)^2, \qquad \forall n \in \mathbb{N}_0, \ \forall f \in L^2(\mathbb{R}), \ \|f\| = 1.$$

Choosing $g_k = f_k$ in Theorem 4.4, we immediately get:

Corollary 4.5 *If $\{f_k\}_{k\in\mathbb{N}_0}$ is an orthonormal sequence in $L^2(\mathbb{R})$, then*

$$\sum_{k=0}^{n} \langle \hat{L} W(f_k), W(f_k) \rangle \geq (n+1)^2, \qquad \forall n \in \mathbb{N}_0, \quad (4.11)$$

and the estimate is optimal, in the sense that if f_k are the Hermite functions, then equality holds in (4.11) *and, conversely, given $n_0 \in \mathbb{N}$, if equality holds in* (4.11) *for all $n \leq n_0$, then there exist $c_k \in \mathbb{C}$ with $|c_k| = 1$ such that $f_k = c_k h_k$ for all $0 \leq k \leq n_0$.*

From Corollary 4.5 we have that if

$$\langle \hat{L} W(f_j), W(f_j) \rangle \leq A, \qquad \forall j \in J,$$

then J must be finite.

Moreover, since

$$\langle \hat{L} W(f_k), W(f_k) \rangle = \mu^2(f_k) + \mu^2(\hat{f}_k) + \Delta^2(f_k) + \Delta^2(\hat{f}_k) \tag{4.12}$$

by Lemma 4.2, we have obtained a simple proof of Theorem 1.3 (the sharp Mean-Dispersion Principle [16, Thm. 2.3]), and then also of Theorem 1.2 (the original Shapiro's Mean-Dispersion Principle).

Formula (4.12) says that Corollary 4.5 is exactly a reformulation of Theorem 1.3, and in this sense Theorem 4.4 and Corollary 4.5 can be seen as Mean-Dispersion principles related with the Wigner transform. On the other hand we observe that working with the Wigner transform gives several advantages. First of all we have more generality since in Theorem 4.4 we can consider different arguments f_k, g_k in the cross-Wigner distribution; moreover the proofs with the Wigner transform are simpler and more self-contained with respect to [16]. Another advantage is that we have information on the Wigner transform of an orthonormal sequence $\{f_k\}_{k \in \mathbb{N}_0}$ rather than on f_k and $\hat{f}_k$ themselves, and this gives more possibilities on how such information can be treated and written. In Sect. 5 we give a Mean-Dispersion principle on the trace of the covariance matrix associated to the Wigner transform; here we start by noting that, from Proposition 2.2, the quantity $\mu^2(f_k) + \mu^2(\hat{f}_k) + \Delta^2(f_k) + \Delta^2(\hat{f}_k)$ in (4.12) can be written not only as $\langle \hat{L} W(f_k), W(f_k) \rangle$, but also through many other operators, as we can see in the following examples.

Example 4.6 For all $f \in L^2(\mathbb{R})$ with $\|f\| = 1$ and finite associated mean and variance of f and $\hat{f}$

$$\mu^2(f) + \mu^2(\hat{f}) + \Delta^2(f) + \Delta^2(\hat{f}) = \langle M^2 f, f \rangle + \langle D^2 f, f \rangle$$

by Proposition 2.2(a), (b). Therefore formula (1.7) for an orthonormal sequence $\{f_k\}_{k \in \mathbb{N}_0}$ in $L^2(\mathbb{R})$ can be rewritten as

$$\sum_{k=0}^{n} \langle (M^2 + D^2) f_k, f_k \rangle \geq (n+1)^2, \qquad \forall n \in \mathbb{N}_0.$$

Example 4.7 For all $f \in L^2(\mathbb{R})$ with $\|f\| = 1$ and finite associated mean and variance of f and $\hat{f}$ we have from Proposition 2.2(g), (h), (k), (l):

$$\left\langle \left[\frac{1}{4}(D_1^2 + D_2^2) + (M_1^2 + M_2^2) \right] W(f), W(f) \right\rangle$$
$$= \Delta^2(f) + \Delta^2(\hat{f}) + \mu^2(f) + \mu^2(\hat{f})$$

and hence for an orthonormal sequence $\{f_k\}_{k \in \mathbb{N}_0} \subset L^2(\mathbb{R})$

$$\sum_{k=0}^{n} \langle PW(f_k), W(f_k) \rangle \geq (n+1)^2, \qquad \forall n \in \mathbb{N}_0,$$

for $P = \frac{1}{4}(D_1^2 + D_2^2) + (M_1^2 + M_2^2)$, by Theorem 1.3.

We can also combine, for example, the operators of Examples 4.6 and 4.7, or add combinations of D_1, D_2, $M_1 D_1 - M_2 D_2$, by Proposition 2.2(e), (f), (i), (j).

5 Covariance

In this section we give an uncertainty principle involving the trace of the covariance matrix of the square of the Wigner distribution $|W(f)(x, \xi)|^2$, and explore its relations with Theorem 1.3.

To this aim, let us first recall some notions about mean and covariance for a function of two variables $\rho(x, y) \in L^1(\mathbb{R}^2)$. We set

$$\rho_X(x) := \int_{\mathbb{R}} \rho(x, y) dy, \qquad \rho_Y(y) := \int_{\mathbb{R}} \rho(x, y) dx, \tag{5.1}$$

and then consider the *means*

$$M(X) := \int_{\mathbb{R}} x \rho_X(x) dx, \qquad M(Y) := \int_{\mathbb{R}} y \rho_Y(y) dy, \tag{5.2}$$

and the *covariances*

$$C(X, Y) := \int_{\mathbb{R}^2} (x - M(X))(y - M(Y)) \rho(x, y) dx dy = C(Y, X)$$
$$C(X, X) = \int_{\mathbb{R}^2} (x - M(X))^2 \rho(x, y) dx dy$$
$$C(Y, Y) = \int_{\mathbb{R}^2} (y - M(Y))^2 \rho(x, y) dx dy.$$

The *covariance matrix*

$$\begin{pmatrix} C(X, X) & C(X, Y) \\ C(Y, X) & C(Y, Y) \end{pmatrix}$$

is symetric and its *trace* is given by

$$
\begin{aligned}
C(X,X)+C(Y,Y) &= \int_{\mathbb{R}^2}\left((x-M(X))^2+(y-M(Y))^2\right)\rho(x,y)dxdy \\
&= \int_{\mathbb{R}^2}(x^2+y^2)\rho(x,y)dxdy \\
&\quad -2M(X)\int_{\mathbb{R}^2}x\rho(x,y)dxdy-2M(Y)\int_{\mathbb{R}^2}y\rho(x,y)dxdy \\
&\quad +(M^2(X)+M^2(Y))\int_{\mathbb{R}^2}\rho(x,y)dxdy. \qquad (5.3)
\end{aligned}
$$

If $\rho(x,y)$ has null means $M(X)=M(Y)=0$, then (5.3) represents the trace of the covariance matrix of $\rho(x,y)$.

For $f\in L^2(\mathbb{R})$ we can consider $\rho(x,\xi)=|W(f)(x,\xi)|^2\in L^1(\mathbb{R}^2)$ since $W(f)\in L^2(\mathbb{R}^2)$. It is then interesting to consider the quantity in (5.3)

$$
\int_{\mathbb{R}^2}(x^2+\xi^2)|W(f)(x,\xi)|^2dxd\xi,
$$

which is related to means and variances of f and $\hat{f}$; indeed, if $f\in L^2(\mathbb{R})$ with $\|f\|=1$, by Proposition 2.2(*k*), (*l*) we have

$$
\begin{aligned}
&\int_{\mathbb{R}^2}(x^2+\xi^2)|W(f)(x,\xi)|^2dxd\xi \\
&= \langle (M_1^2+M_2^2)W(f),W(f)\rangle \qquad (5.4)\\
&= \mu^2(f)+\frac{1}{2}\Delta^2(f)+\mu^2(\hat{f})+\frac{1}{2}\Delta^2(\hat{f}) \qquad (5.5)\\
&\geq \frac{1}{2}(\mu^2(f)+\mu^2(\hat{f})+\Delta^2(f)+\Delta^2(\hat{f})) \qquad (5.6)
\end{aligned}
$$

and the equality in (5.6) holds if and only if $\mu(f)=\mu(\hat{f})=0$. In particular, since the Hermite functions satisfy $\mu(h_k)=\mu(\hat{h}_k)=0$ by [16, Ex. 2.4], from Theorem 1.3 we have the following:

Theorem 5.1 *If $\{f_k\}_{k\in\mathbb{N}_0}$ is an orthonormal sequence in $L^2(\mathbb{R})$, then*

$$
\sum_{k=0}^{n}\int_{\mathbb{R}^2}(x^2+\xi^2)|W(f_k)(x,\xi)|^2dxd\xi\geq\frac{(n+1)^2}{2},\qquad\forall n\in\mathbb{N}_0. \qquad (5.7)
$$

Moreover, given $n_0\in\mathbb{N}$, the equality holds for all $n\leq n_0$ if and only if there exist $c_k\in\mathbb{C}$ with $|c_k|=1$ such that $f_k=c_kh_k$ for all $0\leq k\leq n_0$.

Proof The inequality (5.7) immediately follows from (5.6) and Theorem 1.3. If f_k are multiples of the Hermite functions c_kh_k with $|c_k|=1$, then the equality holds because of (5.5), the fact that $\mu(h_k)=\mu(\hat{h}_k)=0$, and Theorem 1.3.

In the other direction, if the equality holds in (5.7) for all $n \leq n_0$, then from (5.5) we have, for $n \leq n_0$,

$$\sum_{k=0}^{n}(\mu^2(f_k) + \mu^2(\hat{f}_k) + \frac{1}{2}\Delta^2(f_k) + \frac{1}{2}\Delta^2(\hat{f}_k)) = \frac{(n+1)^2}{2}$$

and hence, from Theorem 1.3:

$$\begin{cases} \mu(f_k) = \mu(\hat{f}_k) = 0 & \forall\, 0 \leq k \leq n \\ \sum_{k=0}^{n}(\Delta^2(f_k) + \Delta^2(\hat{f}_k)) = (n+1)^2. & \end{cases}$$

Then we conclude from Theorem 1.3. □

Let us remark that from Theorem 5.1 we immediately get the following uncertainty principle for the covariance matrix:

Corollary 5.2 *If $\{f_j\}_{j\in J}$ is an orthonormal sequence in $L^2(\mathbb{R})$ with zero means $\mu(f_j) = \mu(\hat{f}_j) = 0$, and if the trace of the covariance matrix of $|W(f_j)(x,\xi)|^2$ is uniformly bounded in j, then J is finite.*

Proof From Proposition 2.2(c), (d) we have that

$$M(X) = \int_{\mathbb{R}^2} x|W(f_k)(x,\xi)|^2 dx d\xi = \langle M_1 W(f_k), W(f_k)\rangle = \mu(f_k) = 0$$
$$M(Y) = \langle M_2 W(f_k), W(f_k)\rangle = \mu(\hat{f}_k) = 0$$

by assumption, and hence from (5.3):

$$C(X,X) + C(Y,Y) = \int_{\mathbb{R}^2}(x^2+\xi^2)|W(f_k)(x,\xi)|^2 dx d\xi.$$

The thesis thus immediately follows from Theorem 5.1. □

Note that Corollary 5.2 can be stated also in terms of the variances of $|W(f_j)(x,\xi)|^2$ since, in general, the *variances*

$$V(X) = \int_{\mathbb{R}}(x - M(X))^2 \rho_X(x) dx,$$
$$V(Y) = \int_{\mathbb{R}}(y - M(Y))^2 \rho_Y(y) dy,$$

for ρ_X, ρ_Y, $M(X)$, $M(Y)$ defined as in (5.1)-(5.2), satisfy:

$$C(X,X) = V(X), \qquad C(Y,Y) = V(Y),$$

if $\rho(x,y) \in L^1(\mathbb{R}^2)$.

6 Cohen Classes

Infinitely many operators playing the same role as in the previous sections may be constructed by means of the *Cohen class*

$$Q(f, g) = \frac{1}{\sqrt{2\pi}} \sigma * W(f, g), \qquad f, g \in \mathcal{S}(\mathbb{R}),$$

for some tempered distribution $\sigma \in \mathcal{S}'(\mathbb{R}^2)$. For $f, g \in \mathcal{S}(\mathbb{R})$ we have $W(f, g) \in \mathcal{S}(\mathbb{R}^2)$, and then $Q(f, g)$ is well-defined for every $\sigma \in \mathcal{S}'(\mathbb{R}^2)$. As for the Wigner we define

$$Q[w] = \frac{1}{\sqrt{2\pi}} \sigma * \mathrm{Wig}[w], \qquad w \in \mathcal{S}(\mathbb{R}^2).$$

If $\sigma = \mathcal{F}^{-1}(e^{-iP(\xi,\eta)}) \in \mathcal{S}'(\mathbb{R}^2)$ for some polynomial $P \in \mathbb{R}[\xi, \eta]$, we have the following result (see [3, Thms. 3.1 and 3.2]):

Theorem 6.1 *Let $B(x, y, D_x, D_y)$ be a linear partial differential operator with polynomial coefficients and let $\sigma = \mathcal{F}^{-1}(e^{-iP(\xi,\eta)}) \in \mathcal{S}'(\mathbb{R}^2)$ for some $P \in \mathbb{R}[\xi, \eta]$. Then for every $w \in \mathcal{S}(\mathbb{R}^2)$:*

(i)

$$\begin{aligned} &Q[B(M_1, M_2, D_1, D_2)w] \\ &\quad = B\left(M_1 - \frac{1}{2}D_2 - P_1, M_1 + \frac{1}{2}D_2 - P_1, \frac{1}{2}D_1 + M_2 - P_2, \frac{1}{2}D_1 - M_2 + P_2\right) Q[w] \end{aligned}$$

for

$$P_1 = (iD_1P)(D_1, D_2), \quad P_2 = (iD_2P)(D_1, D_2). \tag{6.1}$$

(ii)

$$\begin{aligned} &B(M_1, M_2, D_1, D_2)Q[w] \\ &\quad = Q\left[B\left(\frac{M_2 + M_1}{2} + P_1^*, \frac{D_1 - D_2}{2} + P_2^*, D_1 + D_2, M_2 - M_1\right) w\right] \end{aligned}$$

for

$$P_1^* = (iD_1P)(D_1 + D_2, M_2 - M_1), \quad P_2^* = (iD_2P)(D_1 + D_2, M_2 - M_1).$$

Let us remark that if $\sigma = \mathcal{F}^{-1}(e^{-iP(\xi,\eta)})$, then $|\hat{\sigma}| = 1$ and hence, for all $f_1, f_2, g_1, g_2 \in \mathcal{S}(\mathbb{R})$, from (2.7) and (2.12):

$$\begin{aligned}
\langle Q(f_1, g_1), Q(f_2, g_2)\rangle &= \frac{1}{2\pi}\langle \sigma * W(f_1, g_1), \sigma * W(f_2, g_2)\rangle \\
&= \frac{1}{2\pi}\langle \mathcal{F}^{-1}(\sqrt{2\pi}\,\hat{\sigma}\cdot \widehat{W(f_1, g_1)}), \mathcal{F}^{-1}(\sqrt{2\pi}\,\hat{\sigma}\cdot \widehat{W(f_2, g_2)})\rangle \\
&= \langle \hat{\sigma}\cdot \widehat{W(f_1, g_1)}, \hat{\sigma}\cdot \widehat{W(f_2, g_2)}\rangle \\
&= \langle |\hat{\sigma}|^2 \widehat{W(f_1, g_1)}, \widehat{W(f_2, g_2)}\rangle = \langle \widehat{W(f_1, g_1)}, \widehat{W(f_2, g_2)}\rangle \\
&= \langle W(f_1, g_1), W(f_2, g_2)\rangle = \langle f_1, f_2\rangle\overline{\langle g_1, g_2\rangle}, \qquad (6.2)
\end{aligned}$$

since $\widehat{f * g} = \sqrt{2\pi}\,\hat{f}\cdot\hat{g}$.

Moreover:

Theorem 6.2 *Let $\{f_k\}_{k\in\mathbb{N}_0} \subset \mathcal{S}(\mathbb{R})$ be an orthonormal basis in $L^2(\mathbb{R})$ and Q, σ as in Theorem 6.1. Then $\{Q(f_j, f_k)\}_{j,k\in\mathbb{N}_0}$ is an orthonormal basis in $L^2(\mathbb{R}^2)$.*

Proof Let us first remark that $Q(f_j, f_k) \in \mathcal{S}(\mathbb{R}^2) \subset L^2(\mathbb{R}^2)$. Moreover $\{Q(f_j, f_k)\}_{j,k\in\mathbb{N}_0}$ is an orthonormal sequence by (6.2).

We only have to prove that if $F \in L^2(\mathbb{R}^2)$ satisfies

$$\langle F, Q(f_j, f_k)\rangle = 0, \qquad \forall j, k \in \mathbb{N}_0,$$

then $F = 0$ a.e. in $\mathbb{R}^2$ (see [7, Thm. 3.4.2]). Let $G = \mathcal{F}^{-1}(\hat{F}/\hat{\sigma}) \in L^2(\mathbb{R}^2)$, so that from (2.7)

$$\begin{aligned}
0 &= \langle F, Q(f_j, f_k)\rangle = \langle \hat{F}, \widehat{Q(f_j, f_k)}\rangle = \langle \hat{G}\cdot\hat{\sigma}, \hat{\sigma}\cdot\widehat{W(f_j, f_k)}\rangle \\
&= \langle |\hat{\sigma}|^2\hat{G}, \widehat{W(f_j, f_k)}\rangle = \langle \hat{G}, \widehat{W(f_j, f_k)}\rangle = \langle G, W(f_j, f_k)\rangle, \qquad \forall j, k \in \mathbb{N}_0,
\end{aligned}$$

which implies $G = 0$ a.e. in $\mathbb{R}^2$ since $\{W(f_j, f_k)\}_{j,k\in\mathbb{N}_0}$ is a basis in $L^2(\mathbb{R}^2)$ by Theorem 3.1.

Then $\hat{F} = \hat{G}\cdot\hat{\sigma} = 0$, i.e. $F = 0$ a.e. in $\mathbb{R}^2$. □

Let us remark that if $f, g \in \mathcal{S}(\mathbb{R}) \subset L^2(\mathbb{R})$ with $\|f\| = \|g\| = 1$ then, by Lemma 4.2, (6.2) and Theorem 6.1, we have

$$\begin{aligned}
&\mu^2(g) + \mu^2(\hat{g}) + \Delta^2(g) + \Delta^2(\hat{g}) \\
&= \langle \hat{L}W(f, g), W(f, g)\rangle \\
&= \langle W(f, (M^2 + D^2)g), W(f, g)\rangle = \langle Q(f, (M^2 + D^2)g), Q(f, g)\rangle \\
&= \left\langle \left[\left(M_1 + \frac{1}{2}D_2 - P_1\right)^2 + \left(\frac{1}{2}D_1 - M_2 + P_2\right)^2\right] Q(f, g), Q(f, g)\right\rangle \qquad (6.3)
\end{aligned}$$

for P_1, P_2 as in (6.1).

Then Theorem 4.4 can be rephrased as follows, for any choice of $P \in \mathbb{R}[\xi, \eta]$:

Theorem 6.3 *Let* $\{f_k\}_{k\in\mathbb{N}_0}, \{g_k\}_{k\in\mathbb{N}_0} \subset \mathcal{S}(\mathbb{R})$ *be two orthonormal sequences in* $L^2(\mathbb{R})$. *Then*

$$\sum_{k=0}^{n}\langle \tilde{L}Q(f_k, g_k), Q(f_k, g_k)\rangle \geq (n+1)^2, \qquad \forall n \in \mathbb{N}_0, \tag{6.4}$$

for any linear partial differential operator $\tilde{L}$ *of the form*

$$\tilde{L}(M_1, M_2, D_1, D_2) = \left(M_1 + \frac{1}{2}D_2 - P_1\right)^2 + \left(\frac{1}{2}D_1 - M_2 + P_2\right)^2$$

with

$$\begin{aligned} &P_1 = (iD_1P)(D_1, D_2), \qquad P_2 = (iD_2P)(D_1, D_2),\\ &P \in \mathbb{R}[\xi, \eta], \qquad \sigma = \mathcal{F}^{-1}(e^{-iP(\xi,\eta)}),\\ &Q(f_k, g_k) = \frac{1}{\sqrt{2\pi}}\sigma * W(f_k, g_k). \end{aligned}$$

Example 6.4 Let $P(D_1, D_2) = \frac{1}{2}D_1D_2$. Then

$$\begin{aligned} P_1 &= iD_1P(\xi_1, \xi_2)\big|_{(\xi_1,\xi_2)=(D_1,D_2)} = \frac{1}{2}D_2\\ P_2 &= iD_2P(\xi_1, \xi_2)\big|_{(\xi_1,\xi_2)=(D_1,D_2)} = \frac{1}{2}D_1 \end{aligned}$$

and hence

$$\begin{aligned} \tilde{L} &= \left(M_1 + \frac{1}{2}D_2 - \frac{1}{2}D_2\right)^2 + \left(\frac{1}{2}D_1 - M_2 + \frac{1}{2}D_1\right)^2\\ &= M_1^2 + (D_1 - M_2)^2. \end{aligned}$$

Therefore, by Theorem 6.3, we obtain

$$\sum_{k=1}^{n}\langle (M_1^2 + (D_1 - M_2)^2)Q(f_k, f_k), Q(f_k, f_k)\rangle \geq (n+1)^2, \qquad \forall n \in \mathbb{N}_0.$$

Example 6.5 Similar results can be obtained considering the operator $P(M_1, M_2) = M_1^2 + M_2^2$ in (5.4) instead of $\hat{L}$ and then Theorem 5.1 instead of Corollary 4.5. Indeed, for $f \in \mathcal{S}(\mathbb{R})$ with $\|f\| = 1$ we can write, by Proposition 2.1, (6.2) and Theorem 6.1:

$$
\begin{aligned}
&\langle (M_1^2 + M_2^2) W(f), W(f) \rangle \\
&= \left\langle \mathrm{Wig}\left[\left(\frac{1}{4}(M_1 + M_2)^2 + \frac{1}{4}(D_1 - D_2)^2 \right) f \otimes \bar{f} \right], W(f) \right\rangle \\
&= \frac{1}{4} \langle Q[((M_1 + M_2)^2 + (D_1 - D_2)^2) f \otimes \bar{f}], Q(f) \rangle \\
&= \frac{1}{4} \left\langle \left(M_1 - \frac{1}{2} D_2 - P_1 + M_1 + \frac{1}{2} D_2 - P_1 \right)^2 Q(f), Q(f) \right\rangle \\
&\quad + \frac{1}{4} \left\langle \left(\frac{1}{2} D_1 + M_2 - P_2 - \frac{1}{2} D_1 + M_2 - P_2 \right)^2 Q(f), Q(f) \right\rangle \\
&= \langle ((M_1 - P_1)^2 + (M_2 - P_2)^2) Q(f), Q(f) \rangle
\end{aligned}
$$

for any P_1, P_2 as in (6.1).

It follows that if $\{f_k\}_{k \in \mathbb{N}_0} \subset \mathcal{S}(\mathbb{R})$ is an orthonormal sequence in $L^2(\mathbb{R})$, then, from Theorem 5.1,

$$
\sum_{k=0}^{n} \langle L^* Q(f_k), Q(f_k) \rangle \geq \frac{(n+1)^2}{2}, \qquad \forall n \in \mathbb{N}_0, \tag{6.5}
$$

for any linear partial differential operator L^* of the form

$$
L^*(M_1, M_2, D_1, D_2) = (M_1 - P_1)^2 + (M_2 - P_2)^2
$$

with

$$
\begin{aligned}
&P_1 = (i D_1 P)(D_1, D_2), \qquad P_2 = (i D_2 P)(D_1, D_2), \\
&P \in \mathbb{R}[\xi, \eta], \qquad \sigma = \mathcal{F}^{-1}(e^{-i P(\xi, \eta)}), \\
&Q(f_k) = \frac{1}{\sqrt{2\pi}} \sigma * W(f_k).
\end{aligned}
$$

Remark 6.6 Any linear operator $T : L^2(\mathbb{R}^2) \to L^2(\mathbb{R}^2)$ (not necessarily everywhere defined) satisfying, for some orthonormal sequence $\{f_k\}_{k \in \mathbb{N}_0} \subset L^2(\mathbb{R})$,

$$
\sum_{k=0}^{n} \langle T W(f_k, f_k), W(f_k, f_k) \rangle \geq (n+1)^2, \qquad \forall n \in \mathbb{N}_0, \tag{6.6}
$$

cannot be a bounded operator on $L^2(\mathbb{R}^2)$. Indeed, assuming by contradiction that T is bounded, by Theorem 3.1 we would have, for all $n \in \mathbb{N}_0$:

$$\begin{aligned}(n+1)^2 &\leq \sum_{k=0}^{n} \langle TW(f_k, f_k), W(f_k, f_k)\rangle \\ &\leq \sum_{k=0}^{n} \|T\|_{\mathcal{L}(L^2,L^2)} \|W(f_k, f_k)\|_{L^2}^2 = (n+1)\|T\|_{\mathcal{L}(L^2,L^2)}\end{aligned}$$

which gives a contradiction for large n. The above considerations can be applied to the partial differential operators with polynomial coefficients appearing in the various results where we have proved estimates of the kind of (6.6). This is not surprising since all non-constant differential operators with polynomial coefficients are in fact unbounded in $L^2(\mathbb{R}^n)$.

7 Riesz Bases

In this section we consider a general Riesz basis of $L^2(\mathbb{R})$ instead of an orthonormal basis. We recall that a *Riesz basis* in a Hilbert space H is the image of an orthonormal basis for H under an invertible linear bounded operator. In particular, if $\{u_k\}_{k\in\mathbb{N}_0}$ is a Riesz basis for $L^2(\mathbb{R})$, we can find an invertible linear bounded operator $U_1 : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ such that

$$U_1(u_k) = h_k, \qquad \forall k \in \mathbb{N}_0,$$

for the Hermite functions $\{h_k\}_{k\in\mathbb{N}_0}$. Moreover (see [7, Lemma 3.6.2])

$$0 < C_1 := \inf_{k\in\mathbb{N}_0} \|u_k\|^2 \leq \sup_{k\in\mathbb{N}_0} \|u_k\|^2 =: C_2 < +\infty. \tag{7.1}$$

We can thus generalize Theorem 4.4 to Riesz bases:

Theorem 7.1 *If $\{u_k\}_{k\in\mathbb{N}_0}$ and $\{v_k\}_{k\in\mathbb{N}_0}$ are Riesz bases for $L^2(\mathbb{R})$ and $\hat{L}$ is the operator in (1.4), then for all $n \in \mathbb{N}_0$*

$$\sum_{k=0}^{n} \langle \hat{L}W(u_k, v_k), W(u_k, v_k)\rangle \geq \frac{\|U_2^{-1}\|_{\mathcal{L}(L^2,L^2)}^2}{\|U_1\|_{\mathcal{L}(L^2,L^2)}^2} \left[\frac{n+1}{\|U_2^{-1}\|_{\mathcal{L}(L^2,L^2)}^2 \|U_2\|_{\mathcal{L}(L^2,L^2)}^2}\right]^2, \tag{7.2}$$

where $U_j : L^2(\mathbb{R}) \to L^2(\mathbb{R})$, $j = 1, 2$, are such that $U_1(u_k) = h_k$ and $U_2(v_k) = h_k$, for the Hermite functions h_k defined in (1.8), and $[x]$ denotes the integer part of x.

Proof As in (4.8)-(4.9) we obtain that

$$\sum_{k=0}^{n}\langle \hat{L}W(u_k,v_k),W(u_k,v_k)\rangle = \sum_{k=0}^{n}\left(\sum_{j=0}^{+\infty}|\langle u_k,h_j\rangle|^2\right)\sum_{\ell=0}^{+\infty}|\langle v_k,h_\ell\rangle|^2(2\ell+1)$$
$$= \sum_{\ell=0}^{+\infty}\sum_{k=0}^{n}\|u_k\|^2|\langle v_k,h_\ell\rangle|^2(2\ell+1),$$

and we can suppose that the series in the right-hand side is convergent, otherwise (7.2) would be trivial. We thus obtain, for the constant C_1 defined in (7.1):

$$\sum_{k=0}^{n}\langle \hat{L}W(u_k,v_k),W(u_k,v_k)\rangle \geq C_1\sum_{\ell=0}^{+\infty}\sum_{k=0}^{n}|\langle v_k,h_\ell\rangle|^2(2\ell+1) = C_1\sum_{\ell=0}^{+\infty}\alpha_\ell(2\ell+1) \tag{7.3}$$

for

$$\alpha_\ell := \sum_{k=0}^{n}|\langle v_k,h_\ell\rangle|^2 \leq \sum_{k=0}^{+\infty}|\langle v_k,h_\ell\rangle|^2 \leq B\|h_\ell\|^2 = B$$

for $B = \|U_2^{-1}\|^2_{\mathcal{L}(L^2,L^2)}$ because of [7, Prop. 3.6.4].

We have

$$\sum_{\ell=0}^{+\infty}\alpha_\ell = \sum_{\ell=0}^{+\infty}\sum_{k=0}^{n}|\langle v_k,h_\ell\rangle|^2 = \sum_{k=0}^{n}\sum_{\ell=0}^{+\infty}|\langle v_k,h_\ell\rangle|^2 = \sum_{k=0}^{n}\|v_k\|^2 \geq \tilde{C}_1(n+1),$$

for $\tilde{C}_1 := \inf_{k\in\mathbb{N}_0}\|v_k\|^2$. Note that $0 < \tilde{C}_1 \leq \sup_{k\in\mathbb{N}_0}\|U_2^{-1}\|^2_{\mathcal{L}(L^2,L^2)}\|h_k\|^2 = B$.

Let us now assume $n \geq \frac{B}{\tilde{C}_1} - 1$ so that

$$m := \left[\frac{n+1}{B}\tilde{C}_1\right] - 1 \in \mathbb{N}_0$$

and write

$$\tilde{C}_1(n+1) \leq \sum_{\ell=0}^{+\infty}\alpha_\ell = \alpha_0 + \cdots + \alpha_m + R_m$$

with

$$R_m := \sum_{\ell\geq m+1}\alpha_\ell.$$

If $R_m = 0$, then $\frac{n+1}{B}\tilde{C}_1 \in \mathbb{N}$ and $\alpha_0 = \cdots = \alpha_m = B$ because otherwise or $m+1 < \frac{n+1}{B}\tilde{C}_1$ or $\alpha_k < B$ for some $k = 0, \ldots, m$ and

$$\tilde{C}_1(n+1) \leq \alpha_0 + \cdots + \alpha_m < B \cdot \frac{n+1}{B}\tilde{C}_1 = (n+1)\tilde{C}_1$$

would give a contradiction. It follows that setting

$$c_k := \begin{cases} 0 & \text{if } R_m = 0 \\ \frac{B-\alpha_k}{R_m} & \text{if } R_m > 0, \end{cases}$$

we have $c_k \geq 0$ and $\alpha_k + c_k R_m = B$ for all $0 \leq k \leq m$.

Moreover, if $R_m > 0$

$$\begin{aligned}(c_0 + \cdots + c_m)R_m = & \; B\left[\frac{n+1}{B}\tilde{C}_1\right] - (\alpha_0 + \cdots + \alpha_m) \\ \leq & \; \tilde{C}_1(n+1) - (\alpha_0 + \cdots + \alpha_m) \leq R_m.\end{aligned}$$

It follows that $c_0 + \cdots + c_m \leq 1$ and hence, for all $n \geq \frac{B}{\tilde{C}_1} - 1$,

$$\begin{aligned}\sum_{\ell=0}^{+\infty}\alpha_\ell(2\ell+1) &= \sum_{\ell=0}^{m}\alpha_\ell(2\ell+1) + \sum_{\ell\geq m+1}\alpha_\ell(2\ell+1) \\ &\geq \sum_{\ell=0}^{m}\alpha_\ell(2\ell+1) + (c_0+\cdots+c_m)\sum_{\ell\geq m+1}\alpha_\ell(2\ell+1) \\ &= \sum_{\ell=0}^{m}\alpha_\ell(2\ell+1) + c_0\sum_{\ell\geq m+1}\alpha_\ell\underbrace{(2\ell+1)}_{>1} + \cdots + c_m\sum_{\ell\geq m+1}\alpha_\ell\underbrace{(2\ell+1)}_{>2m+1} \\ &\geq \underbrace{(\alpha_0 + c_0R_m)}_{=B}\cdot 1 + \cdots + \underbrace{(\alpha_m + c_mR_m)}_{=B}\cdot(2m+1) \\ &= B\sum_{k=0}^{m}(2k+1) = B(m+1)^2 = B\left[\frac{n+1}{B}\tilde{C}_1\right]^2.\end{aligned}$$

Note that the above inequality is trivial if $\frac{n+1}{B}\tilde{C}_1 < 1$ so that from (7.3) we have, for all $n \in \mathbb{N}_0$,

$$\sum_{k=0}^{n}\langle \hat{L}W(u_k, v_k), W(u_k, v_k)\rangle \geq C_1B\left[\frac{n+1}{B}\tilde{C}_1\right]^2. \tag{7.4}$$

Let us now remark that, from the continuity of $U_j : L^2(\mathbb{R}) \to L^2(\mathbb{R})$, $j = 1, 2$,

$$1 = \|h_k\|_{L^2} \leq \|U_1\|_{\mathcal{L}(L^2,L^2)} \cdot \|u_k\|_{L^2}, \quad 1 = \|h_k\|_{L^2} \leq \|U_2\|_{\mathcal{L}(L^2,L^2)} \cdot \|v_k\|_{L^2}$$

for every $k \in \mathbb{N}_0$, and therefore

$$C_1 = \inf_{k\in\mathbb{N}_0} \|u_k\|_{L^2}^2 \geq \frac{1}{\|U_1\|^2_{\mathcal{L}(L^2,L^2)}}, \quad \tilde{C}_1 = \inf_{k\in\mathbb{N}_0} \|v_k\|_{L^2}^2 \geq \frac{1}{\|U_2\|^2_{\mathcal{L}(L^2,L^2)}}.$$

Since $B = \|U_2^{-1}\|^2_{\mathcal{L}(L^2,L^2)}$ we finally have from (7.4) that

$$\sum_{k=0}^{n} \langle \hat{L}W(u_k, v_k), W(u_k, v_k)\rangle \geq \frac{\|U_2^{-1}\|^2_{\mathcal{L}(L^2,L^2)}}{\|U_1\|^2_{\mathcal{L}(L^2,L^2)}} \left[\frac{n+1}{\|U_2^{-1}\|^2_{\mathcal{L}(L^2,L^2)}\|U_2\|^2_{\mathcal{L}(L^2,L^2)}}\right]^2$$

for all $n \in \mathbb{N}_0$.

From Theorem 7.1 and Lemma 4.2 we have the mean-dispersion principle for Riesz bases:

Corollary 7.2 *Let $\{u_k\}_{k\in\mathbb{N}_0}$ be a Riesz basis in $L^2(\mathbb{R})$, with $U(u_k) = h_k$, for the Hermite functions $\{h_k\}_{k\in\mathbb{N}_0}$ defined in (1.8). Then for all $n \in \mathbb{N}_0$*

$$\sum_{k=0}^{n} (\Delta^2(u_k) + \Delta^2(\hat{u}_k) + \mu^2(u_k) + \mu^2(\hat{u}_k)) \geq \frac{1}{\|U^{-1}\|^2\|U\|^2}\left[\frac{n+1}{\|U^{-1}\|^2\|U\|^2}\right]^2,$$

where $\|\cdot\| = \|\cdot\|_{\mathcal{L}(L^2,L^2)}$.

Proof From Lemma 4.2 we have that

$$\langle \hat{L}W(u_k), W(u_k)\rangle = \|u_k\|^4(\Delta^2(u_k) + \Delta^2(\hat{u}_k) + \mu^2(u_k) + \mu^2(\hat{u}_k)).$$

Since $\|u_k\| \leq \|U^{-1}\| \cdot \|h_k\| = \|U^{-1}\|$, by Theorem 7.1 we obtain:

$$\begin{aligned}
&\sum_{k=0}^{n} (\Delta^2(u_k) + \Delta^2(\hat{u}_k) + \mu^2(u_k) + \mu^2(\hat{u}_k)) \\
&\quad \geq \frac{1}{\|U^{-1}\|^4} \sum_{k=0}^{n} \|u_k\|^4 (\Delta^2(u_k) + \Delta^2(\hat{u}_k) + \mu^2(u_k) + \mu^2(\hat{u}_k)) \\
&\quad \geq \frac{1}{\|U^{-1}\|^2\|U\|^2}\left[\frac{n+1}{\|U^{-1}\|^2\|U\|^2}\right]^2.
\end{aligned}$$

□

Note that if the Riesz basis $\{u_k\}_{k\in\mathbb{N}_0}$ is orthonormal, then $\|U\| = 1$ since

$$U(f) = \sum_{k=0}^{+\infty} \langle f, u_k\rangle U(u_k) = \sum_{k=0}^{+\infty} \langle f, u_k\rangle h_k, \qquad f \in L^2(\mathbb{R}),$$

and hence $\|U(f)\|_{L^2} = \|f\|_{L^2}$ for all $f \in L^2(\mathbb{R})$.

From Corollary 7.2 in the case of orthonormal Riesz bases we thus find again (1.7), i.e. Shapiro's Mean Dispersion principle. This improves [16, Cor. 2.8] for $\|U\| = 1$, where a weaker estimate is obtained with respect to Shapiro's Mean Dispersion principle [16, Thm. 2.3].

Acknowledgements The authors are very grateful to the referees for their helpful suggestions that really improved the paper. The authors are also sincerely grateful to Prof. Paolo Boggiatto for helpful discussions about the subject of this work. Boiti and Oliaro were partially supported by the INdAM-GNAMPA Project 2020 "Analisi microlocale e applicazioni: PDEs stocastiche e di evoluzione, analisi tempo-frequenza, varietà", and by the Project FAR 2019 (University of Ferrara). Boiti was also supported by the Projects FAR 2020, FAR 2021, FIRD 2022 (University of Ferrara) and Oliaro was also supported by the Projects "Ricerca locale 2020 - linea A" and "Ricerca locale 2021 - linea A" (University of Torino). Jornet was partially supported by the projects PID2020-119457GB-100 funded by MCIN/AEI/10.13039/501100011033 and by "ERDF A way of making Europe", and by the project GV AICO/2021/170.

Funding Open access funding provided by Università degli Studi di Torino within the CRUI-CARE Agreement.

References

1. Benedicks, M.: On Fourier transforms of functions supported on sets of finite Lebesgue measure. J. Math. Anal. Appl. **106**(1), 180–183 (1985)
2. Boggiatto, P., Fernández, C., Galbis, A.: Supports of representations in the Cohen class. J. Fourier Anal. Appl. **17**(6), 1180–1197 (2011)
3. Boiti, C., Jornet, D., Oliaro, A.: Regularity of partial differential operators in ultradifferentiable spaces and Wigner type transforms. J. Math. Anal. Appl. **446**(1), 920–944 (2017)
4. Boiti, C., Jornet, D., Oliaro, A.: Real Paley-Wiener theorems in spaces of ultradifferentiable functions. J. Funct. Anal. **278**(4), 108348 (2020)
5. Bonami, A., Demange, B., Jaming, P.: Hermite functions and uncertainty principles for the Fourier and the windowed Fourier transforms. Rev. Mat. Iberoam. Eur. Math. Soc. **19**, 23–55 (2003)
6. Buzano, E., Oliaro, A.: Global regularity of second order twisted differential operators. J. Differ. Equ. **268**, 7364–7416 (2020)
7. Christensen, O.: An Introduction to Frames and Riesz Bases. Birkhäuser, Boston (2003)
8. D'ancona, P., Pierfelice, V., Ricci, F.: On the wave equation associated to the Hermite and the twisted Laplacian. J. Fourier Anal. Appl. **16**(2), 294–310 (2010)
9. de Gosson, M.A., Luef, F.: Quantum states and Hardy's formulation of the uncertainty principle: a symplectic approach. Lett. Math. Phys. **80**, 69–82 (2007)
10. Donoho, D.L., Stark, P.B.: Uncertainty principles and signal recovery. SIAM J. Appl. Math. **49**(3), 906–931 (1989)
11. Fernández, C., Galbis, A.: Annihilating sets for the short time Fourier transform. Adv. Math. **224**(5), 1904–1926 (2010)
12. Folland, G., Sitaram, A.: The uncertainty principle: a mathematical survey. J. Fourier Anal. Appl. **3**(3), 207–238 (1997)
13. Gröchenig, K.: Foundations of Time-Frequency Analysis. Birkhäuser, Boston (2001)

14. Gröchenig, K., Zimmermann, G.: Hardy's theorem and the short-time Fourier transform of Schwartz functions. J. Lond. Math. Soc. **63**(2), 205–214 (2001)
15. Hardy, G.H.: A theorem concerning Fourier transforms. J. Lond. Math. Soc. **8**(3), 227–231 (1933)
16. Jaming, P., Powell, A.M.: Uncertainty principles for orthonormal sequences. J. Funct. Anal. **243**, 611–630 (2007)
17. Langenbruch, M.: Hermite functions and weighted spaces of generalized functions. Manuscripta Math. **119**(3), 269–285 (2006)
18. Malinnikova, E.J.: Orthonormal sequences in $L^2(\mathbb{R}^d)$ and time frequency localization. J. Fourier Anal. Appl. **16**(6), 983–1006 (2010)
19. Powell, A.M.: Time-frequency mean and variance sequences of orthonormal bases. J. Fourier Anal. Appl. **11**, 375–387 (2005)
20. Slepian, D., Pollak, H.O.: Prolate spheroidal wave functions, Fourier analysis and uncertainty I. Bell Syst. Tech. J. **40**, 43–63 (1961)
21. Wong, M.-W.: Weyl Transforms. Springer, New York (1998)
22. Wong, M.-W.: Weyl transforms, the heat kernel and Green functions of a degenerate elliptic operator. Ann. Glob. Anal. Geom. **28**(3), 271–283 (2005)