POLITECNICO DI TORINO

Corso di Laurea Magistrale
in Ingegneria Matematica

Tesi di Laurea Magistrale

Some recent results on the norm of localization operators

Relatori
Prof. Fabio Nicola
firma dei relatori

. . . .

Candidato
Federico Riccardi
firma del candidato

. . . .

Anno Accademico 2022-2023

Contents

Acknowledgements

Summary

Chapter 1

Introduction

Con il blu indico frasi da mettere a posto o dove manca una citazione.
Con il rosso indico frasi da controllare se sono corrette.

Chapter 2

Preliminaries

In the first chapter we briefly recall some basic definitions and results of functional and Fourier analysis. In section 2.1 basic concepts about operators between Banach and Hilbert spaces are presented. In addition, certain classes of operators, namely the trace class and the Hilbert-Schmidt class, are introduced with some of their most important properties. Then, in section 2.2 Fourier transform is defined and some of its essential properties are given.

2.1 Basics of Functional Analysis

In this section we turn our attention to linear operators between Banach spaces. Throughout the section we will refer to a generic Banach space as X (or Y), endowed with the norm $\|\cdot\|_X$. If we are dealing with a Hilbert space, we will denote it by H (or K) and its inner product by $\langle\cdot,\cdot\rangle_H$. Pedex in the norm and scalar product may be dropped in case there is no ambiguity. Moreover, the whole theory is presented under the assumption that we are dealing with infinite dimensional spaces, but, unless otherwise stated, everything can be adapted almost directly for finite dimensional spaces.

A generic linear operator between two Banach spaces X and Y is denoted as $T : X \to Y$. As a standard notation, the image of $x \in X$ through T is indicated as $T(x)$, or equivalently as Tx.

Definition 2.1. *A linear operator $T : X \to Y$ is **bounded** if there exist $C > 0$ such that*

$$\|Tx\|_Y \leq C\|x\|_X \quad \forall x \in X. \tag{2.1}$$

For linear operators, boundedness is strictly related to continuity, as the following theorem states.

Theorem 2.2. *For a linear operator T the following statements are equivalent:*

- *T is continuous;*
- *T is bounded.*

For the sake of completeness we mention that actually, for linear operators, boundedness is equivalent to uniform continuity. We denote the set of linear bounded (continuous) operators from X to Y as $\mathscr{B}(X,Y)$, while if $X = Y$ we will just write $\mathscr{B}(X)$.

The smallest constant for which 2.1 holds is the *norm* of T.

Definition 2.3. *Given a linear bounded operator T we define its **norm** as the following number:*

$$\|T\| := \inf\{C > 0 : \|Tx\|_Y \leq C\|x\|_X \ \forall x \in X\} = \sup\left\{\frac{\|Tx\|_Y}{\|x\|_X} : x \in X \setminus \{0\}\right\}$$

The proof of the equivalence between two definition is straightforward. Sometimes, in order to emphasize the spaces between which T operates, we may write the norm of T as $\|T\|_{X\to Y}$.

In what follows we will mostly deal with X and Y being $L^2(\mathbb{R}^d)$, that is, a Hilbert space. For operators between Hilbert spaces, we can specify the norm of an operator using the dual norm:

$$\|T\|_H = \sup\{|\langle Tx, y\rangle_H| : x, y \in H, \ \ \|x\|_H = \|y\|_H = 1\}. \tag{2.2}$$

Among all operators, a rather important class is the one of *compact* operators.

Definition 2.4. *An operator $T \in \mathscr{B}(X,Y)$ is **compact** if, for every bounded sequence $\{x_n\}_{n\in\mathbb{N}} \subset X$ the sequence of the images $\{Tx_n\}_{n\in\mathbb{N}} \subset Y$ has a converging subsequence.*

Theorem 2.5. *The set of compact operators is a closed subspace of $\mathscr{B}(X,Y)$ (with respect to the operator norm topology).*

A first simple example of compact operators are *finite-rank* operators.

Definition 2.6. *An operator $T \in \mathscr{B}(X,Y)$ is said to be* finite-rank *if $\mathrm{Im}(T)$ is finite-dimensional.*

Although this is a simple example of compact operators, finite rank operators are of some importance in the light of the following immediate corollary of Theorem 2.5.

Corollary 2.7. *Let $\{T_n\}_{n\in\mathbb{N}} \subset \mathscr{B}(X,Y)$ be a sequence of finite-rank operators that converges to $T \in \mathscr{B}(X,Y)$. Then T is compact.*

Another crucial property of compact operators is the following.

Theorem 2.8. *Let X, Y, Z be three Banach spaces and let $T \in \mathscr{B}(X,Y)$, $S \in \mathscr{B}(Y,Z)$. Then, if at least one of T and S is compact, then $ST \in \mathscr{B}(X,Z)$ is compact.*

Now we suppose H and K to be Hilbert spaces. Given $T \in \mathscr{B}(H,K)$ there exist a unique $T^* \in \mathscr{B}(K,H)$ such that:

$$\langle Tx, y\rangle_H = \langle x, T^*y\rangle_K \quad \forall x \in H,\, y \in K$$

T^* is called the **adjoint** operator of T. In the particular case in which $T : H \to H$, if $T = T^*$, we say that T is **self-adjoint**.

From now on we suppose that X is over the field $\mathbb{C}$ and that $T \in \mathscr{B}(X)$.

Definition 2.9. *The set* $\sigma(T) = \{\lambda \in \mathbb{C} : T - \lambda I$ *is not invertible*$\}$ *is called the* ***spectrum*** *of* T.

For operators between finite-dimensional spaces (matrices), the spectrum consists of *eigenvalues*, those $\lambda \in \mathbb{C}$ such that $T - \lambda I$ is not injective. However, this is no longer true for infinte-dimensional spaces. Eigenvalues are in the so called *point spectrum*, which in general is just a part of the whole spectrum.

For compact or self-adjoint operators we can state some properties about the spectrum.

Theorem 2.10 (Fredholm's alternative)**.** *Let* $T \in \mathscr{B}(X)$ *be a compact operator. Then for* $T - I$ *one and only one of the following happens:*

- T *is invertible*
- T *is not injective*

Therefore, for compact operators, all the values in the spectrum, except at most for 0, are eigenvalues.
Another fundamental result arises if we add the condition that T is self-adjoint.

Theorem 2.11. *Let* H *be a separable Hilbert space and* $T \in \mathscr{B}(H)$ *a compact and self-adjoint operator. Then there exist an orthonormal basis* $\{e_n\}_{n\in\mathbb{N}}$ *with eigenvalues* $\{\lambda_n\}_{n\in\mathbb{N}}$ *of* H *composed of eigenvectors of* T. *Moreover* $\lim_{n\to+\infty} \lambda_n = 0$.

Hence self-adjoint compact operators can always be diagonalized in some suitable basis ([2]).

From this Theorem immediately follows the next corollary, which relates the eigenvalues of a compact self-adjoint operator with its norm.

Corollary 2.12. *Let* $T \in \mathscr{B}(H)$ *be a self-adjoint compact operator on a separable Hilbert space* H *and suppose its eigenvalues are ordered in such a way that* $|\lambda_1| \geq |\lambda_2| \geq \cdots$. *Then* $\|T\| = |\lambda_1|$.

In light of Theorem 2.11, it is clear that working with compact self-adjoint operators is of great importance. Thus, if an operator T is compact but not self-adjoint, it would be nice to construct an operator associated with T that is also self-adjoint. This task is easily accomplished considering T^*T.

Corollary 2.13. *Let* $T \in \mathscr{B}(H)$ *be a compact operator. Then there exist orthonormal sets* $\{e_n\}_{n\in\mathbb{N}}$ *and* $\{y_n\}_{n\in\mathbb{N}}$ *and non-negative real numbers* $\{\mu_n\}_{n\in\mathbb{N}}$ *with* $\lim_{n\to+\infty} \mu_n = 0$ *so that*

$$T = \sum_{n=1}^{+\infty} \mu_n (\cdot, e_n) y_n, \tag{2.3}$$

where the series converges in norm. These μ_n *are called* singular values *of* T *and are the square root of the eigenvalues of* T^*T.

2.1.1 Trace class and Hilbert-Schmidt operators

In this section we are going to introduce two important classes of operators: *trace-class* operators and *Hilbert-Schmidt* operators.

The trace of an operator can be defined as it is for matrices. Before proceeding, however, we must define what it means for an operator to be non-negative.

Definition 2.14. *Let H be an Hilbert space. An operator $T \in \mathscr{B}(H)$ is said* ***non-negative*** *if*

$$\langle Tx, x\rangle \geq 0 \quad \forall x \in H. \tag{2.4}$$

Condition (2.4) is sufficient to show that, on Hilbert spaces over the field of complex numbers, non-negative operators are automatically self-adjoint. This result is an immediate corollary of the following form of the polarization identity.

Proposition 2.15. *Let $\mathcal{S} : H \times H \to \mathbb{C}$ be a sesquilinear form over the complex Hilbert space H. Then, for every $x, y \in H$:*

$$\mathcal{S}(x,y) = \frac{1}{4}\sum_{k=0}^{3} i^k \mathcal{S}(x + i^k y, x + i^k y). \tag{2.5}$$

Proof. The proof follows from a direct computation of the right-hand side:

$$\sum_{k=0}^{3} i^k \mathcal{S}(x + i^k y, x + i^k y) = \mathcal{S}(x,x)\sum_{k=0}^{3} i^k + \mathcal{S}(x,y)\sum_{k=0}^{3} i^k i^{-k} + \mathcal{S}(y,x)\sum_{k=0}^{3} i^k i^k +$$
$$+ \mathcal{S}(y,y)\sum_{k=0}^{3} i^k i^k i^{-k} = 4\mathcal{S}(x,y) \implies \mathcal{S}(x,y) = \frac{1}{4}\sum_{k=0}^{3} i^k \mathcal{S}(x + i^k y, x + i^k y).$$

□

Proposition 2.16. *Let $T \in \mathscr{B}(H)$ be a non-negative operator over a complex Hilbert space H. Then T is self-adjoint.*

Proof. First of all we notice that if T is non-negative, the quantity $\langle Tx, x\rangle$ is real, therefore $\langle Tx, x\rangle = \overline{\langle Tx, x\rangle} = \langle x, Tx\rangle$. Letting $\mathcal{T}(\cdot,\cdot) = \langle T\cdot, \cdot\rangle$, and using the polarization identity (2.5):

$$\begin{aligned}
\overline{\mathcal{T}(y,x)} &= \frac{1}{4}\sum_{k=0}^{3} \overline{i^k \mathcal{T}(y + i^k x, y + i^k x)} \overset{T \text{ positive}}{=} \frac{1}{4}\sum_{k=0}^{3} i^{-k} \mathcal{T}(y + i^k x, y + i^k x) = \\
&= \frac{1}{4}\sum_{k=0}^{3} i^{-k} \mathcal{T}\left(i^k(x + i^{-k} y), i^k(x + i^{-k} y)\right) = \\
&= \frac{1}{4}\sum_{k=0}^{3} i^{-k} \mathcal{T}(x + i^{-k}, x + i^{-k} y) = \mathcal{T}(x,y).
\end{aligned} \tag{2.6}$$

□

Non-negative operators are somewhat "special", since their behaviour resembles, in some sense, the one of complex numbers. In particular, one can define the *square root* of non-negative operator. Before we do this, we need a preceding lemma.

Lemma 2.17. *The power series of $\sqrt{1-z}$ about zero converges absolutely for all complex numbers such that $|z| \leq 1$.*

Proof. The power series of $\sqrt{1-z}$ about the origin is given by

$$\sqrt{1-z} = \sum_{n=0}^{+\infty} c_n z^n = 1 + \sum_{n=1}^{+\infty} (-1)^n \binom{1/2}{n} z^n$$

where the binomial $\binom{r}{n}$ is defined by $\frac{r(r-1)\cdots(r-n+1)}{n!}$ for every $r \in \mathbb{R}$ and every $n \in \mathbb{N}$, while if $n = 0$ we have $\binom{r}{0} = 1$. Since $\sqrt{1-z}$ is analytic for $|z| < 1$, the series here converges absolutely, so now we have to consider the case $|z| = 1$. We notice that $c_n < 0$ for every $n \geq 1$ because when n is even $\binom{1/2}{n}$ is negative, while if n is odd $\binom{1/2}{n}$ is positive, therefore

$$1 - \sqrt{1-x} = \sum_{n=1}^{+\infty} (-c_n) x^n$$

is a positive series. Using this fact, given $N \in \mathbb{N}$, we have

$$\sum_{n=0}^{N} |c_n| = 1 + \sum_{n=1}^{N} (-c_n) = 1 + \lim_{x \to 1^-} \sum_{n=1}^{N} (-c_n) x^n \overset{\text{positive series}}{\leq}$$
$$\leq 1 + \lim_{x \to 1^-} \sum_{n=1}^{+\infty} (-c_n) x^n = 1 + \lim_{x \to 1^-} (1 - \sqrt{1-x}) = 2.$$

Since this holds for every N, taking the limit $N \to +\infty$ we obtain $\sum_{n=1}^{+\infty} |c_n| < \infty$, which means exactly that the power series is absolutely convergent. □

Theorem 2.18. *Let $T \in \mathscr{B}(X)$ be a non-negative operator. Then, there exist a unique non-negative operator $S \in \mathscr{B}(X)$ such that $S^2 = T$. Moreover S commutes with all bounded operators commuting with T.*

We call S the ***square root*** *of T and we denote it by $S = \sqrt{T}$.*

Proof. If $T = 0$ we let $\sqrt{T} = 0$, otherwise we define $B = I - \|T\|^{-1} T$, where I is the identity operator. Since T is non-negative, for every $x \in X$ such that $\|x\|_X = 1$, we have

$$\langle Bx, x \rangle = \langle (I - \|T\|^{-1} T)x, x \rangle = \|x\|_X^2 - \|T\|^{-1} \langle Tx, x \rangle \leq \|x\|_X^2 = 1$$

which implies, using polarization identity, that $\|B\| \leq 1$. Thanks to Lemma 2.17, this means that the series $\sum_{n=0}^{+\infty} c_n B^n$ is absolutely convergent, therefore convergent, in $\mathscr{B}(X)$ to an operator we indicate with $B_{1/2}$. We define

$$S = \|T\|^{1/2} B_{1/2} \tag{2.7}$$

and we want to show that S is non-negative and satisfies $S^2 = T$. We start proving that $B_{1/2}$, hence S, is non-negative. Taking $x \in X$, we have

$$\begin{aligned}\langle B_{1/2}x, x\rangle = \langle (I + \sum_{n=1}^{+\infty} c_n B^n), x\rangle = \|x\|_X^2 + \sum_{n=1}^{+\infty} c_n \langle B^n x, x\rangle \geq \|x\|_X^2 + \sum_{n=1}^{+\infty} c_n \|B\|^n \|x\|_X^2 \geq \\ \geq \|x\|_X^2 + \sum_{n=1}^{+\infty} c_n \|x\|_X^2 = 0\end{aligned}$$

where we used the fact that c_n are negative, $\|B\| \leq 1$ and that $1+\sum_{n=1}^{\infty} c_n = \sqrt{1-x}|_{x=1} = 0$.

Now we shall prove that $S^2 = T$, which means $\|T\|(B_{1/2})^2 = T$:

$$(B_{1/2})^2 = \left(\sum_{n=0}^{+\infty} c_n B^n\right)\left(\sum_{m=0}^{+\infty} c_m B^m\right) = \sum_{n=0}^{+\infty}\left(\sum_{m=0}^{n} c_m c_{n-m}\right) B^n = \sum_{n=0}^{+\infty} d_n B^n,$$

where the rearrangement is justified since all series are absolutely converging. In order to compute $d_n = \sum_{m=0}^{n} c_m c_{n-m}$, we notice the following:

$$1 - x = \sqrt{1-x}\sqrt{1-x} = \left(\sum_{n=0}^{+\infty} c_n x^n\right)\left(\sum_{m=0}^{+\infty} c_m x^m\right) = \sum_{n=0}^{+\infty}\left(\sum_{m=0}^{n} c_m c_{n-m}\right) x^n,$$

which implies that $d_0 = 1$, $d_1 = -1$ and $d_n = 0$ for $n \geq 2$, therefore $B_{1/2} = I - B$ and, in the end:

$$S^2 = \|T\|(B_{1/2})^2 = \|T\|(I - B) = \|T\|(I - I + \|T\|^{-1}T) = T.$$

Lastly, since the series that defines $B_{1/2}$, hence S, is absolutely convergent, it commutes with every bounded operator commuting with T.

Up to now we only proved the existence of a square root for T, in particular the one given by the expression (2.7). To prove uniqueness of the square root we start supposing S_0 is another non-negative operator in $\mathscr{B}(X)$ such that $S_0^2 = T$. We notice that S_0 commutes with T, indeed $S_0T = S_0S_0^2 = S_0^2S_0 = TS_0$. Therefore S_0 commutes also with S, thus we have:

$$\begin{aligned}(S - S_0)^2 S + (S - S_0)^2 S_0 &= (S - S_0)[(S - S_0)S + (S - S_0)S_0] = \\ &= (S - S_0)(S^2 - S_0S + SS_0 - S_0^2) = \\ &= (S - S_0)(S^2 - S_0^2) = (S - S_0)(T - T) = 0.\end{aligned}$$

But $(S - S_0)^2 S$ and $(S - S_0)^2 S_0$ are non-negative operator, so they must vanish, in particular also their difference $(S - S_0)^2 S - (S - S_0)_0^S = (S - S_0)^3$ is zero. This implies $(S - S_0)^4 = (S - S_0)(S - S_0)^3 = 0$, but since both S and S_0 are self-adjoint, for every $x \in H$ we have:

$$0 = \langle (S - S_0)^4 x, x\rangle = \langle (S - S_0)^2 u, (S - S_0)^2 u\rangle = \|(S - S_0)^2 x\|^2 \implies (S - S_0)^2 = 0$$

and with the same argument we conclude that also $S - S_0 = 0$. $\square$

Remark. From expression (2.7) it is clear that if T is compact then also $\sqrt{T}$, since it is a limit of compact operators.

Definition 2.19. *Given an operator $T \in \mathscr{B}(X)$ we define its* ***absolute value*** *as*

$$|T| := \sqrt{T^*T}. \tag{2.8}$$

Definition 2.20. *A linear operator $U \in \mathscr{B}(X)$ is a* ***partial isometry*** *if it is an isometry over $(\operatorname{Ker} U)^\perp$, i.e. $\|Ux\|_X = \|x\|_X$ for every $x \in (\operatorname{Ker} U)^\perp$.*

Proposition 2.21. *Let $U \in \mathscr{B}(H)$ be a partial isometry. Then also U^* is a partial isometry.*

Proof. Since U is a partial isometry, from polarization identity it follows that $\langle Ux, Uy\rangle = \langle x, y\rangle$ for every $x, y \in (\operatorname{Ker}(U))^\perp$. Clearly the same equality holds if $x \in (\operatorname{Ker}(U))^\perp$ while $y \in \operatorname{Ker}(U)$, so U^*U is the identity over $(\operatorname{Ker}(U))^\perp$. This implies that U^* is an isometry onto $\overline{\operatorname{Im}(U)} = (\operatorname{Ker}(U^*))^\perp$, hence it is a partial isometry too. □

Theorem 2.22. *Given $T \in \mathscr{B}(X)$ there exist unique a partial isometry U such that*

$$T = U|T|, \tag{2.9}$$

which is uniquely determined by the condition $\operatorname{Ker} U = \operatorname{Ker} T$.

Proof. We start defining $\tilde{U} : \operatorname{Im}|T| \to \operatorname{Im}T$. Every $x \in \operatorname{Im}|T|$ can be written as $x = |T|y$ for some $y \in H$, so we can define $\tilde{U}(x) = \tilde{U}(|T|y) := Ty$. We notice that

$$\|x\|^2 = \||T|y\|^2 = \langle |T|y, |T|y\rangle = \langle |T|^2 y, y\rangle = \langle T^*Ty, y\rangle = \|Ty\|^2 = \|\tilde{U}x\|.$$

This computation ensures us that the definition of $\tilde{U}$ is consistent (if $x = |T|y_1 = |T|y_2$ for some $y_1, y_2 \in H$, then $|T|(y_1 - y_2) = 0$, but $\||T|(y_1 - y_2)\| = \|T(y_1 - y_2)\|$, therefore $Ty_1 = Ty_2$) and it implies that $\tilde{U}$ is an isometry over $\operatorname{Im}(|T|)$, thus it can be uniquely extended to an isometry of $\overline{\operatorname{Im}(|T|)}$ over $\overline{\operatorname{Im}T}$. The definition of the map U is straightforward extending $\tilde{U}$ to all H defining it 0 on $(\operatorname{Im}(|T|))^\perp$. Since $|T|$ is self-adjoint $(\operatorname{Im}(|T|))^\perp = \operatorname{Ker}(|T|)$. Furthermore, since $\||T|y\| = \|Ty\|$, $|T|y = 0$ if and only if $Ty = 0$, therefore $\operatorname{Ker}|T| = \operatorname{Ker}T$ which implies, in the end, that $\operatorname{Ker}U = \operatorname{Ker}T$.

Lastly we shall prove that U is unique. Suppose V is another partial isometry such that $T = V|T|$ and $\operatorname{Ker}V = \operatorname{Ker}T$. First condition implies that $T = U|T| = V|T|$, so U and V coincide over $\operatorname{Im}|T|$ (and by continuity over its closure), while second condition implies that $\operatorname{Ker}T = \operatorname{Ker}U = \operatorname{Ker}T$, so U and V coincide over $\operatorname{Ker}|T| = (\operatorname{Im}|T|)^\perp$, therefore U and V coincide over $H = \overline{\operatorname{Im}|T|} \oplus \operatorname{Ker}|T|$. □

Definition 2.23. *Let X be a separable Hilbert space with orthonormal basis $\{e_n\}_{n\in\mathbb{N}}$. Given $T \in \mathscr{B}(X)$ a non-negative operator we define the* ***trace*** *of T as*

$$\operatorname{tr}(T) = \sum_{n=1}^{\infty} \langle Te_n, e_n\rangle. \tag{2.10}$$

Since T is non-negative, every term of the sum in (2.10) is non-negative, so the series is either convergent or divergent. Nevertheless, in principle, it could depend on the basis $\{e_n\}_{n\in\mathbb{N}}$. The following theorem states that the definition is indeed well-posed, namely that the trace does not depend on the basis.

Proposition 2.24. *The definition of* tr *given by* (2.10) *is independent of the basis.*

Proof. Let $T \in \mathscr{B}(H)$ be a non-negative operator and $\{e_n\}_{n\in\mathbb{N}}$ and $\{f_n\}_{n\in\mathbb{N}}$ be two orthonormal basis of H. We have

$$\begin{aligned}\sum_{n=1}^{+\infty}\langle Te_n, e_n\rangle &= \sum_{n=1}^{+\infty}\|\sqrt{T}e_n\|^2 = \sum_{n=1}^{+\infty}\left(\sum_{m=1}^{+\infty}|\langle\sqrt{T}e_n, f_m\rangle|^2\right) = \\ &= \sum_{m=1}^{+\infty}\left(\sum_{n=1}^{+\infty}|\langle\sqrt{T}f_m, e_n\rangle|^2\right) = \sum_{m=1}^{+\infty}\|\sqrt{T}f_m\|^2 = \\ &= \sum_{m=1}^{+\infty}\langle Tf_m, f_m\rangle,\end{aligned}$$

where we used the fact that $\sqrt{T}$ is self-adjoint, while the exchange of series is allowed because all terms are non-negative. $\square$

Definition 2.25. *An operator $T \in \mathscr{B}(X)$ is called* ***Hilbert-Schmidt*** *if and only if* $\mathrm{tr}(T^*T) < \infty$*. We define the* ***Hilbert-Schmidt norm*** *of an operator as* $\|T\|_{\mathrm{HS}} = \sqrt{\mathrm{tr}(T^*T)}$.

From the definition of the trace we can see that, given an orthonormal basis $\{e_n\}_{n=1}^{+\infty}$,

$$\|T\|_{\mathrm{HS}}^2 = \mathrm{tr}(T^*T) = \sum_{n=1}^{+\infty}\langle T^*Te_n, e_n\rangle = \sum_{n=1}^{+\infty}\langle Te_n, Te_n\rangle = \sum_{n=1}^{+\infty}\|Te_n\|^2, \tag{2.11}$$

so we can say, in an equivalent way, that an operator is Hilbert-Schmidt if and only if $\sum_{n=1}^{+\infty}\|Te_n\|^2 < +\infty$. Thanks to Proposition 2.24 we immediately see that the Hilbert-Schmidt norm is independent on the choice of the basis.

We are now going to show some properties of Hilbert-Schmidt and trace-class operators.

Proposition 2.26. *Let $T \in \mathscr{B}(H)$ be a Hilbert-Schmidt operator. Then also $|T|$ and T^* are Hilbert-Schmidt operators and*

$$\||T|\|_{\mathrm{HS}} = \|T^*\|_{\mathrm{HS}} = \|T\|_{\mathrm{HS}}. \tag{2.12}$$

Proof. We already know that, for every $x \in H$, $\|Tx\|^2 = \||T|x\|^2$, therefore from (2.11) we immediately see that $\||T|\|_{\mathrm{HS}} = \|T\|_{\mathrm{HS}}$.

Consider now an orthonormal basis $\{e_n\}_{n\in\mathbb{N}}$ of H. From (2.11):

$$\begin{aligned}\|T\|_{\mathrm{HS}}^2 &= \sum_{n=1}^{+\infty}\|Te_n\|^2 = \sum_{n=1}^{+\infty}\left(\sum_{m=1}^{+\infty}|\langle Te_n, e_m\rangle|^2\right) = \sum_{m=1}^{+\infty}\left(\sum_{n=1}^{+\infty}|\langle e_n, T^*e_m\rangle|^2\right) = \\ &= \sum_{m=1}^{+\infty}\|T^*e_m\|^2 = \|T^*\|_{\mathrm{HS}}^2,\end{aligned}$$

where exchange of series is allowed since all terms are non-negative. $\square$

Theorem 2.27. *Let $T \in \mathscr{B}(H)$ be a Hilbert-Schmidt operator. Then $\|T\| \leq \|T\|_{\text{HS}}$ and T is compact. Moreover, a compact operator T is Hilbert-Schmidt if and only if $\sum_{n=1}^{+\infty} \mu_n^2$, where $\{\mu_n\}_{n=1}^{+\infty}$ are the singular values of T.*

Proof. Let $\{e_n\}_{n\in\mathbb{N}}$ be an orthonormal basis of H and let $x \in H$. We have:

$$\begin{aligned}\|Tx\|^2 &= \sum_{n=1}^{+\infty} |\langle Tx, e_n\rangle|^2 = \sum_{n=1}^{+\infty} |\langle x, T^*e_n\rangle|^2 \leq \|x\|^2 \sum_{n=1}^{+\infty} \|T^*e_n\|^2 = \\ &= \|x\|^2\|T^*\|_{\text{HS}}^2 \overset{(2.12)}{=} \|x\|^2\|T\|_{\text{HS}}^2.\end{aligned}$$

Taking the supremum over all $x \in H$ gives us that $\|T\| \leq \|T\|_{\text{HS}}$.

To prove that T is compact, consider the following sequence of operators $T_N = \sum_{n=1}^{N} \langle \cdot, e_n\rangle Te_n$. All T_N are finite-rank, so they are compact. Moreover, $T_N \to T$ in $\mathscr{B}(H)$, indeed:

$$\|T - T_N\| \leq \|T - T_N\|_{\text{HS}}^2 = \sum_{n=N+1}^{+\infty} \|Te_n\|^2 \to 0 \quad \text{as } N \to +\infty,$$

therefore, from Corollary 2.7, T is compact.

Lastly, if T is a compact operator also T^*T is compact as well as self-adjoint, hence we can consider as orthonormal basis of H the one given by its eigenvectors $\{e_n\}_{n\in\mathbb{N}}$. Recalling Corollary 2.13, we know that μ_n^2 are exactly the eigenvalues of T^*T. For such basis we have:

$$\|T\|_{\text{HS}}^2 = \sum_{n=1}^{+\infty} \|Te_n\|^2 = \sum_{n=1}^{+\infty} \langle Te_n, Te_n\rangle = \sum_{n=1}^{+\infty} \langle T^*Te_n, e_n\rangle = \sum_{n=1}^{+\infty} \mu_n^2.$$

$\square$

Proposition 2.28. *Let $T \in \mathscr{B}(H)$ be a Hilbert-Schmidt operator and $S \in \mathscr{B}(H)$. Then TS and ST are both Hilbert-Schmidt operators.*

Proof. Given an orthonormal basis $\{e_n\}_{n\in\mathbb{N}}$ of H we have:

$$\|ST\|_{\text{HS}}^2 = \sum_{n=1}^{+\infty} \|STe_n\|^2 \leq \|S\|^2 \sum_{n=1}^{+\infty} \|Te_n\|^2 = \|S\|^2\|T\|_{\text{HS}}^2,$$

so ST is Hilbert-Schmidt. Moreover, from Proposition 2.12 follows:

$$\|TS\|_{\text{HS}} = \|(TS)^*\|_{\text{HS}} = \|S^*T^*\|_{\text{HS}} \leq \|S^*\|\|T^*\|_{\text{HS}} = \|S\|\|T\|_{\text{HS}}.$$

$\square$

We can now turn back to trace-class operator. It is evident that trace-class operators and Hilbert-Schmidt operators are strictly related. Indeed, as seen in the proof of Proposition 2.24, we have

$$\text{tr}|T| = \sum_{n=1}^{+\infty} \langle |T|e_n, e_n\rangle = \sum_{n=1}^{+\infty} \|\sqrt{|T|}e_n\|^2 = \|T\|_{\text{HS}}^2,$$

so, an operator is trace-class if and only if $\sqrt{|T|}$ is Hilbert-Schmidt. By virtue of this link, we can exploit properties of Hilbert-Schmidt operators in order to obtain information about trace-class operators. First of all, we are going to prove that expression (2.10) is well-defined for every trace-class operator and not only for non-negative ones.

Theorem 2.29. *Let $T \in \mathscr{B}(H)$ be a trace-class operator and $\{e_n\}_{n\in\mathbb{N}}$ an orthonormal basis of H. Then $\sum_{n=1}^{+\infty}\langle Te_n, e_n\rangle$ converges absolutely and the limit is independent of the basis.*

Proof. We start proving that the series converges absolutely. Letting $T = U|T|$ be the polar decomposition of T, we want to show that:

$$\sum_{n=1}^{+\infty} |\langle Te_n, e_n\rangle| = \sum_{n=1}^{+\infty} |\langle U\sqrt{|T|}\sqrt{|T|}e_n, e_n\rangle| = \sum_{n=1}^{+\infty} |\langle \sqrt{|T|}e_n, \sqrt{|T|}U^*e_n\rangle| < +\infty. \tag{2.13}$$

From Cauchy-Schwarz' inequality, for every term we have that $|\langle\sqrt{|T|}e_n, \sqrt{|T|}U^*e_n\rangle| \leq \|\sqrt{|T|}e_n\|\,\|\sqrt{|T|}U^*e_n\|$. Since T is trace-class, $\sqrt{|T|}$ is Hilbert-Schmidt and, thanks to Proposition 2.28, also $\sqrt{|T|}U^*$ is a Hilbert-Schmidt operator, therefore both $\{\|\sqrt{|T|}e_n\|\}_{n\in\mathbb{N}}$ and $\{\|\sqrt{|T|}U^*e_n\|\}_{n\in\mathbb{N}}$ are in $\ell^2(\mathbb{N})$. This allows us to use the Cauchy-Schwarz inequality in the last expression of (2.13), thus obtaining:

$$\begin{aligned}\sum_{n=1}^{+\infty} |\langle Te_n, e_n\rangle| &= \sum_{n=1}^{+\infty} |\langle\sqrt{|T|}e_n, \sqrt{|T|}U^*e_n\rangle| \leq \sum_{n=1}^{+\infty} \|\sqrt{|T|}e_n\|\,\|\sqrt{|T|}U^*e_n\| \overset{\text{C−S}}{\leq} \\ &\leq \left(\sum_{n=1}^{+\infty}\|\sqrt{|T|}e_n\|^2\right)^{1/2}\left(\sum_{n=1}^{+\infty}\|\sqrt{|T|}U^*e_n\|^2\right)^{1/2} \leq \|\sqrt{|T|}\|_{\text{HS}}^2.\end{aligned}$$

The proof of the independence of the basis is exactly the same as the one for Proposition 2.24, because now the exchange of series is allowed since the series is convergent. ☐

Another immediate corollary of Proposition 2.28 is the following proposition

Proposition 2.30. *Let $T \in \mathscr{B}(H)$ be a trace-class operator. Then T is also Hilbert-Schmidt.*

Proof. Let $T = U|T|$ be the polar decomposition of T. Since T is trace-class $\sqrt{|T|}$ is an Hilbert-Schmidt operator. Therefore, from Proposition 2.28 follows that $T = U\sqrt{|T|}\sqrt{|T|}$ is an Hilbert-Schmidt operator. ☐

Theorem 2.31. *Let $T \in \mathscr{B}(H)$. Then T is trace-class if and only if it is compact and $\sum_{n=1}^{+\infty}\mu_n$, where $\{\mu_n\}_{n\in\mathbb{N}}$ are the singular values of T. Moreover, if T is trace-class then $\text{tr}|T| = \sum_{n=1}^{+\infty}\mu_n$ and additionally, if it is also self-adjoint, $\text{tr}T = \sum_{n=1}^{+\infty}\lambda_n$ where $\{\lambda_n\}_{n\in\mathbb{N}}$. Finally, we have $\|T\| \leq \text{tr}|T|$.*

Proof. Let $T = U|T|$ be the polar decomposition of T. If T is trace-class $\sqrt{|T|}$ is Hilbert-Schmidt, but from Theorem 2.27 we have that $\sqrt{|T|}$ is compact, therefore also $T = U\sqrt{|T|}\sqrt{|T|}$ and $|T|$ are compact. Not only $|T|$ is compact, but it is also self-adjoint and its eigenvalues are exactly $\{\mu_n\}_{n\in\mathbb{N}}$. Letting $\{e_n\}_{n\in\mathbb{N}}$ be an orthonormal basis made up of eigenvectors of $|T|$, we have

$$\mathrm{tr}|T| = \sum_{n=1}^{+\infty}\langle |T|e_n, e_n\rangle = \sum_{n=1}^{+\infty}\mu_n.$$

Conversely, if T is a compact operator the previous formula still holds and shows that T is also trace-class. With the same reasoning, if T is self-adjoint and trace-class, we can write the trace with respect to the basis made up its eigenvectors thus obtaining $\mathrm{tr}T = \sum_{n=1}^{+\infty}\lambda_n$.

For the last part of the Theorem, we notice that $|T|$ is a compact self-adjoint non-negative operator, therefore, assuming its eigenvalues are decreasingly ordered, from Corollary 2.12 we have that $\||T|\| = \mu_1$, hence:

$$\|T\| = \|U|T|\| \leq \||T|\| = \mu_1 \leq \sum_{n=1}^{+\infty}\mu_n = \mathrm{tr}|T|.$$

□

In the special case $H = L^2(\mathbb{R}^d)$ we are able to give a characterization of Hilbert-Schmidt operators.

Theorem 2.32. *An operator $T \in \mathscr{B}(L^2(\mathbb{R}^d))$ is Hilbert-Schmidt if only if there exist a function $K_T \in L^2(\mathbb{R}^d \times \mathbb{R}^d)$, called integral kernel, such that*

$$(Tf)(x) = \int_{\mathbb{R}^d} K_T(x,y)f(y)dy \quad \forall f \in L^2(\mathbb{R}^d). \tag{2.14}$$

Moreover $\|T\|_{\mathrm{HS}} = \|K_T\|_2$.

Proof. We start proving that, given $K \in L^2(\mathbb{R}^{2d})$, the corresponding integral operator defined by (2.14) is continuos. Denoting with T_K such operator, for every $f \in L^2(\mathbb{R}^d)$, we have:

$$\begin{aligned}\|T_K f\|^2_{L^2(\mathbb{R}^d)} &= \int_{\mathbb{R}^d}|T_K f(x)|^2dx = \int_{\mathbb{R}^d}\Big|\int_{\mathbb{R}^d}K(x,y)f(y)dy\Big|^2dx \leq \\ &\leq \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d}|K(x,y)||f(y)|dy\right)^2dx.\end{aligned}$$

From Fubini's theorem we have that $|K(x,\cdot)| \in L^2(\mathbb{R}^d)$ for almost every $x \in \mathbb{R}^d$. Therefore, we can apply Cauchy-Schwarz's inequality in the inner integral, thus obtaining:

$$\|T_K f\|^2_{L^2(\mathbb{R}^d)} \leq \|f\|^2_{L^2(\mathbb{R}^d)}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}|K(x,y)|^2dydx = \|K\|^2_{L^2(\mathbb{R}^{2d})}\|f\|_{L^2(\mathbb{R}^d)} \tag{2.15}$$

which shows that T_K is a bounded operator. Moreover, T_K is clearly linear, therefore $T_K \in \mathscr{B}(L^2(\mathbb{R}^d))$. Consider now an orthonormal basis $\{e_n\}_{n\in\mathbb{N}}$ of $L^2(\mathbb{R}^d)$. Moreover, always from (2.15), it follows that, for every $f \in L^2(\mathbb{R}^d)$, the operator $K \in L^2(\mathbb{R}^{2d}) \mapsto A_f(K) = T_K f \in L^2(\mathbb{R}^d)$ is linear and continuous.

We can now suppose to have an Hilbert-Schmidt operator $T \in \mathscr{B}(H)$ and an orthonormal basis $\{e_n\}_{n\in\mathbb{N}}$ of $L^2(\mathbb{R}^d)$. We want to show that (2.14) holds with the following kernel:

$$K_T(x,y) = \sum_{(m,n)\in\mathbb{N}^2} \langle Te_n, e_m\rangle_{L^2(\mathbb{R}^d)} e_m(x)\overline{e_n(y)}. \tag{2.16}$$

Since $\{e_m \otimes \overline{e_n}\}_{(m,n)\in\mathbb{N}^2}$ is an orthonormal basis of $L^2(\mathbb{R}^{2d})$, the convergence is unconditional. Therefore, for every $f \in L^2(\mathbb{R}^d)$

$$\begin{aligned}
\int_{\mathbb{R}^d} K_T(x,y)f(y)dy &= A_f(K_T)(x) = A_f\Big(\sum_{(m,n)\in\mathbb{N}^2} \langle Te_n, e_m\rangle e_m \otimes \overline{e_n}\Big)(x) \overset{A.2}{=} \\
&= \sum_{(m,n)\in\mathbb{N}^2} A_f(\langle Te_n, e_m\rangle e_m \otimes \overline{e_n})(x) \overset{A.3}{=} \\
&= \sum_{m=1}^{+\infty}\sum_{n=1}^{+\infty} \langle Te_n, e_m\rangle A_f(e_m \otimes e_n)(x) = \\
&= \sum_{m=1}^{+\infty}\sum_{n=1}^{+\infty} \langle Te_n, e_m\rangle \int_{\mathbb{R}^d} e_m(x)\overline{e_n(y)}f(y)dy = \\
&= \sum_{m=1}^{+\infty}\sum_{n=1}^{+\infty} \langle T(\langle f, e_n\rangle e_n), e_m\rangle e_m(x) = \\
&= \sum_{m=1}^{+\infty} \langle Tf, e_m\rangle e_m(x) = (Tf)(x),
\end{aligned}$$

where the use of A.2 is justified because, $\sum_{(m,n)\in\mathbb{N}^2}\langle Te_n, e_m\rangle e_m \otimes \overline{e_n}$ converges unconditional, while the use of A.3 is justified because A_f is continuous, thus (A.2) implies that $\sum_{(m,n)\in\mathbb{N}^2} A_f(\langle Te_n, e_m\rangle e_m \otimes \overline{e_n})$ converges unconditionally. Finally, we have:

$$\begin{aligned}
\|K_T\|^2_{L^2(\mathbb{R}^{2d})} &= \sum_{(n,m)\in\mathbb{N}^2} |\langle Te_n, e_m\rangle|^2 = \sum_{n=1}^{+\infty}\sum_{m=1}^{+\infty} |\langle Te_n, e_m\rangle|^2 = \\
&= \sum_{n=1}^{+\infty} \|Te_n\|^2_{L^2(\mathbb{R}^d)} = \|T\|^2_{\mathrm{HS}}.
\end{aligned} \tag{2.17}$$

Conversely, if we have $K \in L^2(\mathbb{R}^{2d})$ and we define T_K by (2.14), we have to show that this operator is Hilbert-Schmidt, but this is straightforward since (2.17) still holds. □

In light of this theorem, operators defined by (2.14) are called *Hilbert-Schmidt integral operators.*

Proposition 2.33. *Let T be an Hilbert-Schmidt integral operator over $L^2(\mathbb{R}^d)$ with integral kernel $K \in L^2(\mathbb{R}^{2d})$. Then its adjoin operator is given by*

$$T^*f(x) = \int_{\mathbb{R}^d} \overline{K(y,x)}f(y)dy. \tag{2.18}$$

Therefore T is self-adjoint if and only if $K(x,y) = \overline{K(y,x)}$.

Proof. Let $f, g \in L^2(\mathbb{R}^d)$. We start showing that $K(x,y)f(y)\overline{g(x)} \in L^1(\mathbb{R}^{2d})$:

$$\int_{\mathbb{R}^{2d}} |K(x,y)||f(y)||g(x)|dxdy \overset{\text{Tonelli}}{=} \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d}|K(x,y)||f(y)|dy\right)|g(x)|dx =$$
$$= \int_{\mathbb{R}^d}(T_{|K|}|f|)(y)|g(y)| \overset{\text{C-S}}{\leq} \|T_{|K|}\|_2\|g\|_2 \leq \|K\|_2\|f\|_2\|g\|_2 < +\infty,$$

where $T_{|K|}$ denotes the Hilbert-Schmidt integral operator with kernel $|K|$ and $\|T_{|K|}|f|\|_2 \leq \|K\|_2\|f\|_2$ because $T_{|K|}$ is Hilbert-Schmidt, therefore $\|T_{|K|}\| \leq \|T_{|K|}\|_{\text{HS}} = \|K\|_2$. We are now in the position to use Fubini's theorem:

$$\langle Tf, g\rangle = \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d}K(x,y)f(y)dy\right)\overline{g(x)}dx = \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d}K(x,y)\overline{g(x)}dx\right)f(y)dy =$$
$$= \int_{\mathbb{R}^d}\overline{\left(\int_{\mathbb{R}^d}\overline{K(x,y)}g(x)dx\right)}f(y)dy = \langle f, T^*g\rangle$$

where the expression of T^* is exactly the one in (2.18). □

2.2 Fourier Transform and its properties

In this section we introduce the Fourier transform with its elementary properties, its relation with some fundamental operators in time-frequency analysis and with the convolution product.

Definition 2.34. *Let $f \in L^1(\mathbb{R}^d)$. We define the* ***Fourier transform*** *of f the function*

$$\mathscr{F}f(\omega) = \hat{f}(\omega) \coloneqq \int_{\mathbb{R}^d} e^{-2\pi i\omega\cdot t}f(t)dt \tag{2.19}$$

It is straightforward to see that the definition is well-posed and that $\mathscr{F}f \in L^\infty(\mathbb{R}^d)$ with $\|\mathscr{F}f\|_\infty \leq \|f\|_1$. Therefore $\mathscr{F}$ can be seen as a linear operator between $L^1(\mathbb{R}^d)$ and $L^\infty(\mathbb{R}^d)$ with $\|\mathscr{F}\| \leq 1$. Actually, taking $f \geq 0$ a.e., we have that $\hat{f}(0) = \|f\|_1$, which gives us the equality.

The Fourier transform of an $L^1(\mathbb{R}^d)$ is not only bounded, as stated by the *Riemann-Lebesgue lemma.*

Theorem 2.35 (Riemann-Lebesgue lemma)**.** *Let $f \in L^1(\mathbb{R}^d)$. Then $\hat{f} \in C_0(\mathbb{R}^d) = \{f : \mathbb{R}^d \to \mathbb{C}$ continuous such that $\lim_{|t|\to\infty}|f(t)| = 0\}$.*

Definition 2.36. *Let $f \in L^1(\mathbb{R}^d)$. We define the* ***inverse Fourier transform*** *of the function f*

$$\mathscr{F}^{-1}f(t) = \check{f}(t) \coloneqq \int_{\mathbb{R}^d} e^{2\pi i\omega\cdot t}f(\omega)d\omega \tag{2.20}$$

The inverse Fourier transform is denoted with $\mathscr{F}^{-1}$ because it is actually the inverse operator of the Fourier transform as stated by the *inversion theorem.*

Theorem 2.37 (Inversion theorem)**.** *Let $f \in L^1(\mathbb{R}^d)$ and suppose that also $\hat{f} \in L^1(\mathbb{R}^d)$. Then*

$$f(t) = \mathscr{F}^{-1} \circ \mathscr{F} f(t) = \int_{\mathbb{R}^d} \hat{f}(\omega) e^{2\pi i \omega \cdot t} d\omega$$

If f is in $L^2(\mathbb{R}^d)$, the integral in (2.19) in general will not converge. Nevertheless, we can define the Fourier transform of an L^2 function through a density argument. For example, one can use $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$, which is a dense subspace of $L^2(\mathbb{R}^d)$. On this space one can show that the Fourier transform is an isometry with respect to the L^2 norm and therefore it extends to an isometry on the whole $L^2(\mathbb{R}^d)$. This is stated by the *Plancherel theorem*.

Theorem 2.38 (Plancherel theorem)**.** *If $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ then $\|f\|_2 = \|\hat{f}\|_2$.*

Thanks to the polarization identity this implies that $\mathscr{F}$ preserves the inner product in $L^2(\mathbb{R}^d)$:

$$\langle f, g \rangle_{L^2(\mathbb{R}^d)} = \langle \hat{f}, \hat{g} \rangle_{L^2(\mathbb{R}^d)} \quad \forall f, g \in L^2(\mathbb{R}^d), \tag{2.21}$$

therefore the Fourier transform $\mathscr{F}$ is a unitary operator on $L^2(\mathbb{R}^d)$. Result (2.21) is called *Parseval formula*.

So far we have seen that the Fourier transform is defined on L^1 and L^2. It can be shown, through Riesz-Thorin's interpolation theorem, that it can be extended to all L^p spaces for $1 < p < 2$.

Theorem 2.39 (Hausdorff-Young)**.** *Let $1 \leq p \leq 2$ and let p' such that $\frac{1}{p} + \frac{1}{p'} = 1$. Then $\mathscr{F} : L^p(\mathbb{R}^d) \to L^{p'}(\mathbb{R}^d)$ and $\|\hat{f}\|_{p'} \leq \|f\|_p$.*

In what follows we will need the sharp version of the Hausdorff-Young inequality:

$$\|\hat{f}\|_{p'} \leq \left(\frac{p^{1/p}}{p'^{1/p'}} \right)^{d/2} \|f\|_p = A_p^d \|f\|_p \tag{2.22}$$

where A_p is the so-called Babenko-Bechner constant.

Having considered the spaces over which the Fourier transform is defined, we can focus on some of its properties. In particular, we want to focus on the close relationship that arises between regularity and decay properties. This is explained by the following theorems.

Theorem 2.40. *Let $f \in L^1(\mathbb{R}^d)$. If $|t|^k f \in L^1(\mathbb{R}^d)$ for some $k \in \mathbb{N}$ then $\hat{f} \in C_0^k(\mathbb{R}^d)$ and the following holds for every $\alpha \in \mathbb{N}^d$ with $|\alpha| \leq k$:*

$$\mathscr{F}\left((-2\pi i t)^\alpha f\right)(\omega) = \partial^\alpha \mathscr{F} f(\omega). \tag{2.23}$$

Theorem 2.41. *Let $f \in C^k(\mathbb{R}^d)$ for some $k \in \mathbb{N}$. If $f, \partial^\alpha f \in L^1(\mathbb{R}^d)$ for every $\alpha \in \mathbb{N}^d$ with $|\alpha| \leq k$ then*

$$\mathscr{F}\left(\partial^\alpha f\right)(\omega) = (2\pi i \omega)^\alpha \mathscr{F} f(\omega). \tag{2.24}$$

In particular this implies that $\hat{f}(\omega) = o\left(|\omega|^{-k}\right)$ as $|\omega| \to \infty$.

In summary, previous theorems establish a duality between regularity and decay: if a function is smooth, then its Fourier transform decays rapidly and vice versa.

We now introduce some fundamental operators in Fourier and time-frequency analysis. Given $x, \xi \in \mathbb{R}^d$ and $\lambda > 0$ we define the *time-shift* (or translation) operator T_x

$$T_x f(t) = f(t - x) \quad \forall\, t \in \mathbb{R}^d, \tag{2.25}$$

the *modulation* operator M_ξ

$$M_\xi f(t) = e^{2\pi i \xi \cdot t} f(t) \quad \forall\, t \in \mathbb{R}^d, \tag{2.26}$$

and the *dilation* operator D_λ

$$D_\lambda f(t) = \lambda^d f(\lambda t) \quad \forall\, t \in \mathbb{R}^d. \tag{2.27}$$

Moreover, time-shift and modulation operators can be combined into a *time-frequency shift* operator

$$\pi(x, \xi) f(t) = M_\xi T_x f(t) \quad \forall\, t \in \mathbb{R}^d. \tag{2.28}$$

It is easy to check that all these operators are isometric isomorphisms with respect to the L^1 norm. We show how these operators act under the Fourier transform.

Proposition 2.42. *Let $f \in L^1(\mathbb{R}^d)$. Then the following holds:*

(i) $\mathscr{F}(T_x f)(\omega) = M_{-x} \hat{f}(\omega)$*;*

(ii) $\mathscr{F}(M_\xi f)(\omega) = T_\xi \hat{f}(\omega)$*;*

(iii) $\mathscr{F}(D_\lambda f)(\omega) = \hat{f}\left(\dfrac{\omega}{\lambda}\right)$*.*

Proof.

(i)
$$\begin{aligned}\mathscr{F}(T_x f)(\omega) &= \int_{\mathbb{R}^d} f(t-x) e^{-2\pi i \omega \cdot t} dt \overset{s=t-x}{=} \int_{\mathbb{R}^d} f(s) e^{-2\pi i \omega \cdot (s+x)} ds = \\ &= e^{-2\pi i \omega \cdot x} \int_{\mathbb{R}^d} f(s) e^{-2\pi i \omega \cdot s} ds = M_{-x} \hat{f}(\omega);\end{aligned}$$

(ii)
$$\mathscr{F}(M_\xi f)(\omega) = \int_{\mathbb{R}^d} e^{2\pi i \xi \cdot t} f(t) e^{-2\pi i \omega \cdot t} dt = \int_{\mathbb{R}^d} f(t) e^{-2\pi i (\omega - \xi) \cdot t} dt = T_\xi \hat{f}(\omega);$$

(iii)
$$\mathscr{F}(D_\lambda f)(\omega) = \int_{\mathbb{R}^d} \lambda^d f(\lambda t) e^{-2\pi i \omega \cdot t} dt \overset{s=\lambda t}{=} \int_{\mathbb{R}^d} f(s) e^{-2\pi i \omega \cdot \frac{s}{\lambda}} ds = \hat{f}\left(\frac{\omega}{\lambda}\right).$$

□

We point out that the second property, namely that $\mathscr{F}(M_\xi f) = T_\xi \hat{f}$, is shedding light on the role of modulation operator: while T_x acts as a translation in the time domain, M_ξ is a translation in the frequency domain. Therefore the time-frequency shift operator $\pi(x, \xi)$ is indeed a shift operator because it acts as a translation in the joint time-frequency domain.

Thanks to these properties and Theorems 2.40 and 2.41 we can compute the Fourier transform of Gaussians.

Proposition 2.43. *The Fourier transform maps Gaussian into Gaussians. More precisely, for $\lambda > 0$, we have:*

$$\mathscr{F}(e^{-\lambda\pi|\cdot|^2})(\omega) = \frac{1}{\lambda^{d/2}} e^{-\frac{1}{\lambda}\pi|\omega|^2} \quad \forall\, \omega \in \mathbb{R}^d. \tag{2.29}$$

Proof. We start considering the 1-dimensional case and for the ease of notation we let $\varphi_\lambda(t) = e^{-\lambda\pi t^2}$ for $t \in \mathbb{R}$.

First of all we consider $\varphi = \varphi_1$, for which we have:

$$\frac{d\varphi}{dt}(t) = -2\pi t\varphi(t)$$

If we take the Fourier transform of both members, using (2.24) in the former and (2.23) in the latter we obtain:

$$2\pi i\omega\hat{\varphi}(\omega) = -i\frac{d\hat{\varphi}}{d\omega}(\omega) \implies \frac{d\hat{\varphi}}{d\omega}(\omega) = -2\pi\omega\hat{\varphi}(\omega)$$

which is exactly the same equation satisfied by φ, therefore $\hat{\varphi}(\omega) = Ce^{-\pi\omega^2}$ for some $C \in \mathbb{R}$. Since $\hat{\varphi}(0) = \int_{\mathbb{R}} \varphi(t)dt = \|\varphi\|_1 = 1$ we obtain that $C = 1$.

The general case can be proved using the dilatation operator:

$$\varphi_\lambda(t) = e^{-\lambda\pi t^2} = e^{-\pi\left(\sqrt{\lambda}t\right)^2} = \frac{1}{\sqrt{\lambda}} D_{\sqrt{\lambda}}\varphi(t) \implies$$
$$\implies \hat{\varphi}_\lambda(\omega) = \frac{1}{\sqrt{\lambda}}\mathscr{F}(D_{\sqrt{\lambda}}\varphi)(\omega) = \frac{1}{\sqrt{\lambda}}\hat{\varphi}\left(\frac{\omega}{\sqrt{\lambda}}\right) = \frac{1}{\sqrt{\lambda}}e^{-\frac{1}{\lambda}\pi\omega^2}$$

The passage to the multidimensional case is almost straightforward since $e^{-\lambda|t|^2} = \prod_{j=1}^d e^{-\lambda t_j^2}$, therefore:

$$\mathscr{F}(e^{-\lambda\pi|\cdot|^2})(\omega) = \int_{\mathbb{R}^d} e^{-\lambda\pi|t|^2}e^{-2\pi i\omega\cdot t}dt = \prod_{j=1}^d \int_{\mathbb{R}} e^{-\lambda\pi t_j^2}e^{-2\pi i\omega_j t_j}dt_j =$$
$$= \prod_{j=1}^d \frac{1}{\sqrt{\lambda}}e^{-\frac{1}{\lambda}\pi\omega_j^2} = \frac{1}{\lambda^{d/2}}e^{-\frac{1}{\lambda}\pi|\omega|^2}.$$

□

We conclude this section by considering what a convolution is and how this product relates to the Fourier transform. Given two functions $f, g : \mathbb{R}^d \to \mathbb{R}$, their convolution is given by:

$$(f * g)(x) = \int_{\mathbb{R}^d} f(y)g(x-y)dy.$$

The well-posedness of the convolution is given by Young's theorem.

Theorem 2.44 (Young)**.** *Given $f \in L^p(\mathbb{R}^d)$ and $g \in L^q(\mathbb{R}^d)$, suppose that $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$ with $r \geq 1$. Then $f * g \in L^r(\mathbb{R}^d)$ and $\|f * g\|_r \leq \|f\|_p\|g\|_q$.*

We notice that, if $p = q = 1$, then $r = 1$ so the convolution of two L^1 functions is still in L^1. Thanks to this, using Fubini's theorem it is immediate to see that:

$$\mathscr{F}(f * g) = \mathscr{F} f \cdot \mathscr{F} g,$$

which explains the connection between convolution and Fourier transform. Just like as Young's inequality, in what follows we will need the sharp version of Hausdorff-Young's inequality, namely:

$$\|f * g\|_r \leq (A_p A_q A_{r'})^d \|f\|_p \|g\|_q. \tag{2.30}$$

Chapter 3

Short-Time Fourier Transform

The Fourier transform is a widely used tool in both theoretical and applied settings. In particular, from an applied point of view, the importance of the Fourier transform lies primarily in the possibility of examining a signal in the frequency domain, and thus obtaining information that would otherwise be difficult to derive from the time domain. However, the distinction between these domains is rigid, whereas it is of great importance to have a tool with which time and frequency characteristics can be examined simultaneously. There are many representations that can accomplish this task, such as the Wigner distribution, the ambiguity function and so on. The basic time-frequency representation of a signal is probably the *short-time Fourier transform* or *STFT*. In this chapter we will define the STFT and describe some of its properties.

3.1 STFT

The *short-time Fourier transform* or *STFT* is a powerful tool used in signal processing and time-frequency analysis to study the properties of a signal locally in both time and frequency. The main idea behind STFT is the following: if we want some information about the spectrum of a signal f at a certain time, say T, we could choose an interval $(T-\Delta T, T+\Delta T)$ and take the Fourier transform of $f\chi_{(T-\Delta T,T+\Delta T)}$. Normally, multiplication by a characteristic function will not yield a regular function (not even continuous) and, in light of the duality between regularity and decay, the Fourier transform of $f\chi_{(T-\Delta T,T+\Delta T)}$ will not decay quickly. The problem with this not quick decay is that in this case the energy in the frequency domain will be spread all over the domain. Therefore, a sharp cutoff in the time domain leads to a "poor" localization in the frequency domain. In order to avoid this kind of problem, we could think to multiply the signal f by a smooth function.

Definition 3.1. *Fix a function $y \neq 0$, called* window function. *The* ***short-time Fourier transform*** *of a function f with window ϕ is defined as*

$$\mathcal{V}_\phi f(x,\omega) = \int_{\mathbb{R}^d} f(t)\overline{\phi(t-x)}e^{-2\pi i\omega\cdot t}dt, \quad (x,\omega)\in\mathbb{R}^{2d} \tag{3.1}$$

In the above definition we did not specify where f and ϕ are chosen. We notice that (3.1) can be seen in an alternative way:

$$\mathcal{V}_\phi f(x,\omega) = \mathscr{F}(fT_x\overline{\phi})(\omega), \tag{3.2}$$

therefore STFT is well-defined whenever the Fourier transform of this function is. For example, if both f and ϕ are in $L^2(\mathbb{R}^d)$ then $fT_x\overline{\phi}$ is in $L^1(\mathbb{R}^d)$ for every $x \in \mathbb{R}^d$ and so the integral in (3.1) is defined. In this special case the STFT can be written as a scalar product in $L^2(\mathbb{R}^d)$:

$$\mathcal{V}_\phi f(x,\omega) = \langle f, M_\omega T_x \phi\rangle = \langle f, \pi(x,\omega)\phi\rangle.$$

In general, the STFT of f with respect to ϕ will be defined whenever $\langle f, M_\omega T_x\phi\rangle$ is an expression of some sort of duality. For example, if $f \in \mathcal{S}'(\mathbb{R}^d)$ and $\phi \in \mathcal{S}(\mathbb{R}^d)$, then $M_\omega T_x\phi \in \mathcal{S}(\mathbb{R}^d)$, therefore $\langle f, M_\omega T_x\phi\rangle$ can be seen as the usual duality between tempered distributions and functions in the Schwartz space. Despite this remark, we will mainly focus on the case in which both the window and the signal are in $L^2(\mathbb{R}^d)$.

3.1.1 Properties of STFT

In this section we will introduce and prove some basic properties of STFT. In particular, as we did for the Fourier transform, we want to know to which spaces $\mathcal{V}_\phi f$ belongs.

Theorem 3.2. *Let $f_1, f_2, \phi_1, \phi_2 \in L^2(\mathbb{R}^d)$. Then $\mathcal{V}_{\phi_i} f_i \in L^2(\mathbb{R}^{2d})$ and the following holds:*

$$\langle \mathcal{V}_{\phi_1} f_1, \mathcal{V}_{\phi_2} f_2\rangle = \langle f_1, f_2\rangle\overline{\langle \phi_1, \phi_2\rangle}. \tag{3.3}$$

Proof. We start proving that, if $f \in L^2(\mathbb{R}^d)$ and $\phi \in \mathcal{S}(\mathbb{R}^d)$, the STFT of f with window ϕ is in $L^2(\mathbb{R}^{2d})$. Since we supposed $\phi \in \mathcal{S}(\mathbb{R}^d)$, the function $fT_x\overline{\phi}$ is in $L^2(\mathbb{R}^d)$ for every $x \in \mathbb{R}^d$, hence:

$$\begin{aligned}\|\mathcal{V}_\phi f\|_2^2 &= \int_{\mathbb{R}^{2d}} |\mathcal{V}_\phi f(x,\omega)|^2 dx d\omega \overset{\text{Tonelli}}{=} \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} |\mathcal{V}_\phi f(x,\omega)|^2\right) d\omega \overset{(3.2)}{=} \\ &= \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} |\mathscr{F}(fT_x\overline{\phi})|^2(\omega) d\omega\right) dx \overset{\text{Plancherel}}{=} \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} |f(t)\overline{\phi(t-x)}|^2 dt\right) dx \overset{\text{Tonelli}}{=} \\ &= \int_{\mathbb{R}^d} |f(t)|^2 \left(\int_{\mathbb{R}^d} |\phi(t-x)|^2 dx\right) dt = \|f\|_2^2\|\phi\|_2^2.\end{aligned}$$

Now, consider $f_1, f_2 \in L^2(\mathbb{R}^d)$ and $\phi_1, \phi_2 \in \mathcal{S}(\mathbb{R}^d)$. Both $\mathcal{V}_{\phi_1} f_1$ and $\mathcal{V}_{\phi_2} f_2$ are in $L^2(\mathbb{R}^{2d})$, so their product is in $L^1(\mathbb{R}^{2d})$, hence we can use Fubini's theorem in the following:

$$\begin{aligned}\langle \mathcal{V}\phi_1 f_1, \mathcal{V}_{\phi_2} f_2\rangle &= \int_{\mathbb{R}^{2d}} \mathcal{V}_{\phi_1} f_1(x,\omega)\overline{\mathcal{V}_{\phi_2} f_2(x,\omega)} dx d\omega \overset{\text{Fubini}}{=} \\ &= \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} \mathscr{F}(f_1T_x\overline{\phi_1})(\omega)\overline{\mathscr{F}(f_2T_x\overline{\phi_2})(\omega)} d\omega\right) dx \overset{\text{Parseval}}{=} \\ &= \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} f_1(t)\overline{\phi_1(t-x)}\,\overline{f_2(t)}\phi_2(t-x) dt\right) dx \overset{\text{Fubini}}{=} \\ &= \int_{\mathbb{R}^d} f_1(t)\overline{f_2(t)}\left(\int_{\mathbb{R}^d} \phi_2(t-x)\overline{\phi_1(t-x)} dx\right) dt = \langle f_1, f_2\rangle\langle \phi_2, \phi_1\rangle.\end{aligned}$$

The transition from $\mathcal{S}(\mathbb{R}^d)$ to whole $L^2(\mathbb{R}^d)$ is done trough a density argument. Indeed, for $f_1, f_2 \in L^2(\mathbb{R}^d)$ and $\phi_1 \in \mathcal{S}(\mathbb{R}^d)$ fixed, the mapping $\phi_2 \in L^2(\mathbb{R}^d) \mapsto \langle \mathcal{V}_{\phi_1} f_1, \mathcal{V}_{\phi_2} f_2 \rangle$ is a linear functional and we just showed that it is bounded over $\mathcal{S}(\mathbb{R}^d)$, where it coincides with $\langle f_1, f_2 \rangle \langle \phi_2, \phi_1 \rangle$. Since Schwartz's class is a dense subspace of $L^2(\mathbb{R}^d)$ it extends to a bounded linear operator for every $\phi_2 \in L^2(\mathbb{R}^d)$. Similarly, for fixed $f_1, f_2 \in L^2(\mathbb{R}^d)$ and $\phi_2 \in L^2(\mathbb{R}^d)$ the mapping $\phi_1 \in L^2(\mathbb{R}^d) \mapsto \langle \mathcal{V}_{\phi_1} f_1, \mathcal{V}_{\phi_2} f_2 \rangle$ is an antilinear functional that coincides with $\langle f_1, f_2 \rangle \langle \phi_2, \phi_1 \rangle$ over $\mathcal{S}(\mathbb{R}^d)$, therefore it extends to a bounded linear functional over whole $L^2(\mathbb{R}^d)$. □

Corollary 3.3. *If $f, \phi \in L^2(\mathbb{R}^d)$ then*

$$\|\mathcal{V}_\phi f\|_2 = \|f\|_2 \|\phi\|_2. \tag{3.4}$$

In particular, if $\|\phi\|_2 = 1$, $\mathcal{V}_\phi$ is an isometry from $L^2(\mathbb{R}^d)$ into $L^2(\mathbb{R}^{2d})$.

Proof. It is sufficient to consider (3.3) with $\phi_1 = \phi_2 = \phi$ and $f_1 = f_2 = f$. □

From the Cauchy-Schwarz inequality we immediately see that $\mathcal{V}_\phi f$ is in $L^\infty(\mathbb{R}^{2d})$:

$$|\mathcal{V}_\phi f(x,\omega)| = |\langle f, M_\omega T_x \phi \rangle| \overset{\text{C−S}}{\leq} \|f\|_2 \|M_\omega T_x \phi\|_2 = \|f\|_2 \|\phi\|_2. \tag{3.5}$$

Combing this with (3.4) and using a simple interpolation argument we see that $\mathcal{V}_\phi f \in L^p(\mathbb{R}^{2d})$ for every $p \in [2, +\infty]$ and that

$$\|\mathcal{V}_\phi f\|_p \leq \|f\|_2 \|\phi\|_2. \tag{3.6}$$

This result is improved by the following theorem due to Lieb [19].

Theorem 3.4. *If $f, \phi \in L^2(\mathbb{R}^d)$ and $2 \leq p < \infty$, then:*

$$\|\mathcal{V}_\phi\|_p^p = \int_{\mathbb{R}^{2d}} |\mathcal{V}_\phi(x,\omega)|^p dx d\omega \leq \left(\frac{2}{p}\right)^d \|f\|_2^p \cdot \|g\|_2^p. \tag{3.7}$$

Proof. Using the Cauchy-Schwarz inequality it is immediate to see that $fT_x\overline{\phi} \in L^1(\mathbb{R}^d)$ for every $x \in \mathbb{R}^d$. In addition to that, since $\mathcal{V}_\phi f = \mathscr{F}(fT_x\overline{\phi}) \in L^2(\mathbb{R}^{2d})$, from Fubini's theorem we can say that $\mathscr{F}(fT_x\overline{\phi}) \in L^2(\mathbb{R}^d)$ for almost every $x \in R^d$ and therefore also $fT_x\overline{\phi} \in L^2(\mathbb{R}^d)$ for a.e. $x \in \mathbb{R}^d$. Through an interpolation argument we obtain that $fT_x\overline{\phi} \in L^q(\mathbb{R}^d)$ for every $q \in$ [1,2].

We start considering the L^p norm of $\mathcal{V}_\phi f$:

$$\begin{aligned}
\|\mathcal{V}_\phi f\|_p &= \left(\int_{\mathbb{R}^{2d}} |\mathcal{V}_\phi f(x,\omega)|^p dx d\omega\right)^{1/p} \overset{\text{Tonelli}}{=} \left[\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |\mathcal{V}_\phi f(x,\omega)|^p d\omega\right) dx\right]^{1/p} = \\
&= \left[\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |\mathscr{F}(fT_x\overline{\phi})(\omega)|^p d\omega\right) dx\right]^{1/p} \overset{(2.30)}{\leq} \\
&\leq A_{p'}^d \left[\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |f(t)\overline{\phi(t-x)}|^{p'} dt\right)^{p/p'} dx\right]^{1/p},
\end{aligned} \tag{3.8}$$

where the use of Young's inequality (2.30) is justified since we noticed that $f T_x \overline{\phi}$ is in $L^q(\mathbb{R}^d)$ for every $q \in [1,2]$, so in particular it is in $L^{p'}(\mathbb{R}^d)$. Letting $\phi^*(t) = \overline{\phi(-t)}$ and considering the inner integral we have

$$\int_{\mathbb{R}^d} |f(t)\overline{\phi(t-x)}|^{p'} dt = \int_{\mathbb{R}^d} |f(t)|^{p'} |\phi^*(x-t)|^{p'} dt = \left(|f|^{p'} * |\phi^*|^{p'}\right)(x),$$

so the expression in (3.8) is the $L^{p/p'}(\mathbb{R}^d)$ norm of $|f|^{p'} * |\phi^*|^{p'}$. Since both f and ϕ are in $L^2(\mathbb{R}^d)$ and $p' \leq 2$ we have that $|f|^{p'}, |\phi^*|^{p'} \in L^{2/p'}(\mathbb{R}^d)$. Thanks to Young's theorem 2.44 $|f|^{p'} * |\phi^*|^{p'}$ belongs to $L^r(\mathbb{R}^d)$, where r is given by:

$$\frac{1}{(2/p')} + \frac{1}{(2/p')} = 1 + \frac{1}{r} \implies r = \frac{1}{p'-1} = \frac{1}{\frac{p}{p-1} - 1} = p - 1 = \frac{p}{p'},$$

therefore, using the sharp version of Young's inequality (2.30) in (3.8) we obtain:

$$\|\mathcal{V}_\phi f\|_p \leq A_{p'}^d \left(A_{2/p'}^d A_{2/p'}^d A_{(p/p')'}^d \| \, |f|^{p'} \|_{2/p'} \| \, |\phi^*|^{p'} \|_{2/p'}\right)^{1/p'}.$$

However $\| \, |f|^{p'} \|_{2/p'} = (\int_{\mathbb{R}^d} (|f(x)|^{p'})^{2/p'} dx)^{p'/2} = \|f\|_2^{p'}$ and from a direct calculation (which can be found in B.1) one can see that $A_{p'}^d A_{2/p'}^{2d/p'} A_{(p/p')'}^{d/p'} = (2/p)^{d/p}$, which corresponds to the desired result. □

From a direct computation one can see that the adjoint operator of the STFT operator $\mathcal{V}_\phi$ is given by the following expression:

$$\mathcal{V}_\phi^* g(t) = \int_{\mathbb{R}^{2d}} g(x,\omega)\phi(t-x)e^{2\pi i \omega \cdot t} dx d\omega = \int_{\mathbb{R}^{2d}} g(x,\omega) M_\omega T_x \phi(t) dx d\omega \quad \forall g \in L^2(\mathbb{R}^{2d}). \tag{3.9}$$

This adjoint operator appears in the following nice property, named *inversion formula for the STFT*.

Theorem 3.5. *Let $f \in L^2(\mathbb{R}^d)$ and $\phi, \gamma \in L^2(\mathbb{R}^{2d})$ such that $\langle \phi, \gamma \rangle \neq 0$. Then:*

$$f(t) = \frac{1}{\langle \phi, \gamma \rangle} \mathcal{V}_\gamma^* \mathcal{V}_\phi f(t) = \frac{1}{\langle \phi, \gamma \rangle} \int_{\mathbb{R}^{2d}} \mathcal{V}_\phi f(x,\omega) M_\omega T_x \gamma(t) dx d\omega \quad \forall t \in \mathbb{R}^d. \tag{3.10}$$

Proof. Given $f, g \in L^2(\mathbb{R}^d)$ from (3.3) we have:

$$\langle \mathcal{V}_\phi f, \mathcal{V}_\gamma g \rangle = \langle f, g \rangle \langle \gamma, \phi \rangle.$$

On the other hand:

$$\langle \mathcal{V}_\phi f, \mathcal{V}_\gamma g \rangle = \langle \mathcal{V}_\gamma^* \mathcal{V}_\phi f, g \rangle,$$

therefore, letting I be the identity operator over $L^2(\mathbb{R}^d)$, we have:

$$\langle \mathcal{V}_\gamma^* \mathcal{V}_\phi f, g \rangle = \langle f, g \rangle \langle \gamma, \phi \rangle \implies \langle (\mathcal{V}_\gamma^* \mathcal{V}_\phi - \langle \gamma, \phi \rangle I) f, g \rangle = 0.$$

Since this holds for every $g \in L^2(\mathbb{R}^d)$ necessarily:

$$(\mathcal{V}_\gamma^* \mathcal{V}_\phi - \langle \gamma, \phi \rangle I) f = 0 \implies \frac{1}{\langle \gamma, \phi \rangle} \mathcal{V}_\gamma^* \mathcal{V}_\phi f = f.$$

□

Therefore, the adjoint operator $\mathcal{V}_\gamma^*$ acts, in some sense, as an inverse operator. This will be of paramount importance afterwards.

3.2 Bargmann Transform and Fock Space

Throughout this section we will consider the STFT with Gaussian window. We choose

$$\varphi(x) = 2^{d/4} e^{-\pi|x|^2} \tag{3.11}$$

where $|x|^2 = \sum_{k=1}^d x_k^2$ is the Euclidean norm of x in $\mathbb{R}^d$. The factor $2^{d/4}$ is chosen so that $\|\varphi\|_2 = 1$. The STFT with Gaussian window becomes

$$\mathcal{V}_\varphi f(x,\omega) = 2^{d/4} \int_{\mathbb{R}^d} f(t) e^{-\pi|t-x|^2} e^{-2\pi i\omega\cdot t} dt \tag{3.12}$$

Our aim now is to rearrange the terms in the above expression in order to make $z = x + i\omega \in \mathbb{C}^d$ appear. We want to highlight the fact that when talking about complex quantities $|z|^2 = z\overline{z} = |x|^2 + |\omega|^2$. Aggiungere remark sulla notazione usata.

$$\begin{aligned}
\mathcal{V}_\varphi f(x,\omega) &= 2^{d/4} \int_{\mathbb{R}^d} f(t) e^{-\pi|t|^2 + 2\pi x\cdot t - \pi|\omega|^2} e^{-2\pi i\omega\cdot t} dt \\
&= 2^{d/4} \int_{\mathbb{R}^d} f(t) e^{-\pi|t|^2} e^{2\pi(x-i\omega)\cdot t} e^{-\frac{\pi}{2}\left(|x|^2 - 2ix\cdot\omega - |\omega|^2\right)} e^{-\frac{\pi}{2}\left(|x|^2 + |\omega|^2 + 2ix\cdot\omega\right)} dt \\
&= 2^{d/4} e^{-\pi ix\cdot\omega} e^{-\frac{\pi}{2}\left(|x|^2+|\omega|^2\right)} \int_{\mathbb{R}^d} f(t) e^{-\pi|t|^2} e^{2\pi(x-i\omega)\cdot t} e^{-\frac{\pi}{2}(x-i\omega)^2} dt.
\end{aligned}$$

The rearrangement may seem arbitrary, but actually it is done in such a way that inside the integral x and ω enter only via $\overline{z}$.

Definition 3.6. *The* ***Bargmann transform*** *of a function f on $\mathbb{R}^d$ is the function $\mathcal{B}f$ on $\mathbb{C}^d$ given by*

$$\mathcal{B}f(z) = 2^{d/4} \int_{\mathbb{R}^d} f(t) e^{2\pi t\cdot z - \pi|t|^2 - \frac{\pi}{2}z^2} dt. \tag{3.13}$$

We recall that a function defined over $\mathbb{C}^d$ is *entire* if it is holomorphic over all $\mathbb{C}^d$.

Definition 3.7. *The* ***Fock space*** *$\mathcal{F}^2(\mathbb{C}^d)$ is the Hilbert space of all entire functions F on $\mathbb{C}^d$ for which the norm*

$$\|F\|_{\mathcal{F}^2}^2 = \int_{\mathbb{C}^d} |F(z)|^2 e^{-\pi|z|^2} dz \tag{3.14}$$

is finite.

Clearly the norm of the Fock space is induced by the following scalar product

$$\langle F, G\rangle_{\mathcal{F}^2} \int_{\mathbb{C}^d} F(z)\overline{G(z)} e^{-\pi|z|^2} dz \tag{3.15}$$

Proposition 3.8. *If f is a function on $\mathbb{R}^d$ with polynomial growth then its Bargmann transform $\mathcal{B}f$ is an entire function on $\mathbb{C}^d$. Moreover, letting $z = x + i\omega$, the Bargmann transform of f is related to its STFT through the following*

$$\mathcal{V}_\varphi f(x,-\omega) = e^{\pi ix\cdot\omega} \mathcal{B}f(z) e^{-\pi|z|^2/2} \tag{3.16}$$

Proof. contenuto... □

Proposition 3.9. *If $f \in L^2(\mathbb{R}^d)$ then*

$$\|f\|_2 = \left(\int_{\mathbb{C}^d} |\mathcal{B}f(z)|^2 e^{-\pi|z|^2} dz \right)^{1/2} = \|\mathcal{B}f\|_{\mathcal{F}^2}. \tag{3.17}$$

Thus $\mathcal{B}$ is an isometry from $L^2(\mathbb{R}^d)$ into $\mathcal{F}^2(\mathbb{C}^d)$.

Chapter 4

Localization Operators

In the previous chapter we defined the STFT, which, roughly speaking, gives us a 'time-frequency picture' of a signal. Once we have our joint representation, we might be interested in highlighting some of its features, for example, to understand where most of the energy is in phase-space.

In this chapter we will look at two ways in which we can solve the problem of creating operators able to localize a signal.

4.1 Localization with projections

Our first attempt to localize a signal is arguably the most straightforward one, namely using a sharp cutoff. If we suppose to have a signal $f \in L^2(\mathbb{R}^d)$ and we want to localize it in a measurable subset $T \subseteq \mathbb{R}^d$ of the time domain we can consider the natural projection operator:

$$P_T : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d) \quad P_T f(t) = \chi_T(t) f(t). \tag{4.1}$$

This is clearly a projection operator, which means that $P_T^2 = P_T = P_T^*$.

In the same fashion we can define an operator able to localize on a measurable subset $\Omega \subseteq \mathbb{R}^d$ in the frequency domain. Its definition it is not as direct as the one for time projections but it is still easy to understand:

$$Q_\Omega : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d) \quad Q_\Omega f(t) = \mathscr{F}^{-1}\left(\chi_\Omega \mathscr{F} f\right)(t) = \int_\Omega \hat{f}(\omega) e^{2\pi i \omega \cdot t} d\omega. \tag{4.2}$$

It is also quite simple to show that this is a projection operator:

$$\begin{aligned} Q_\Omega^2 &= \mathscr{F}^{-1}\chi_\Omega \mathscr{F}\mathscr{F}^{-1}\chi_\Omega\mathscr{F} = \mathscr{F}^{-1}\chi_\Omega\chi_\Omega\mathscr{F} = \mathscr{F}^{-1}\chi_\Omega\mathscr{F} = Q_\Omega; \\ Q_\Omega^* &= \left(\mathscr{F}^{-1}\chi_\Omega\mathscr{F}\right)^* = \mathscr{F}^*\chi_\Omega^*\left(\mathscr{F}^{-1}\right)^* = \mathscr{F}^{-1}\chi_\Omega\mathscr{F} = Q_\Omega, \end{aligned}$$

where we used the fact that the Fourier transform is a unitary operator on $L^2(\mathbb{R}^d)$, namely that $\mathscr{F}^* = \mathscr{F}^{-1}$.

Since both operators are projections, their norm is less or equal than 1, independently of T and Ω, indeed:

$$\|P_T f\|_{L^2(\mathbb{R}^d)} = \|f\|_{L^2(T)} \leq \|f\|_{L^2(\mathbb{R}^d)};$$
$$\|Q_\Omega f\|_{L^2(\mathbb{R}^d)} = \|\mathscr{F}^{-1}\chi_\Omega \mathscr{F} f\|_{L^2(\mathbb{R}^d)} \overset{\text{Plancherel}}{=} \|\mathscr{F} f\|_{L^2(\Omega)} \leq \|\mathscr{F} f\|_{L^2(\mathbb{R}^d)} \overset{\text{Plancherel}}{=} \|f\|_{L^2(\mathbb{R}^d)}.$$

Clearly, those operators do not answer our original question about localization in time-frequency, since P_T and Q_Ω act only in time and frequency, respectively. However, we may think to combine these projections into a single operator:

$$Q_\Omega P_T,\ P_T Q_\Omega : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d),$$

which hopefully are able to localize a signal both in time and frequency "near" to the set $T \times \Omega$.

It is clear that these operators are linear and bounded, in particular their norms are less or equal than 1. Moreover, they are one the adjoint of the other, indeed:

$$(Q_\Omega P_T)^* = P_T^* Q_\Omega^* = Q_\Omega P_T. \tag{4.3}$$

Up to now the only (essential) hypothesis on T and Ω is that they are measurable. Clearly, by adding some requirements on T and Ω we expect $Q_\Omega P_T$ and $P_T Q_\Omega$ to gain some properties.

Proposition 4.1. *Let $T, \Omega \subset \mathbb{R}^d$ with finite measure. Then $Q_\Omega P_T$ and $P_T Q_\Omega$ are Hilbert-Schmidt integral operators of the form*

$$Q_\Omega P_T f(x) = \int_{\mathbb{R}^d} K(x,t) f(t) dt; \tag{4.4}$$
$$P_T Q_\Omega f(x) = \int_{\mathbb{R}^d} \overline{K(t,x)} f(t) dt, \tag{4.5}$$

where

$$K(x,t) = \chi_T(t) \int_\Omega e^{2\pi i \omega \cdot (x-t)} d\omega, \tag{4.6}$$

which has $\|K\|_{L^2(\mathbb{R}^{2d})} = \sqrt{|T||\Omega|}$.

Proof. Given $f \in L^2(\mathbb{R}^d)$ we have:

$$Q_\Omega P_T f(x) = \int_\Omega e^{2\pi i \omega \cdot x} \left(\int_T e^{-2\pi i \omega \cdot t} f(t) dt \right) d\omega \overset{\text{Fubini}}{=} \int_{\mathbb{R}^d} \chi_T(t) \left(\int_\Omega e^{2\pi i \omega \cdot (x-t)} d\omega \right) f(t) dt,$$

where the use of Fubini's theorem is allowed since Ω and T have finite measure. This gives us the expression of $Q_\Omega P_T$. In order to obtain also the expression for $P_T Q_\Omega$ it suffices to recall from (4.3) that $(Q_\Omega P_T)^* = P_T Q_\Omega$. Therefore, the integral kernel of $P_T Q_\Omega$ is given by Proposition 2.33. Lastly, from Theorem 2.32, we have that $\|Q_\Omega P_T\|_{\text{HS}} = \|P_T Q_\Omega\|_{\text{HS}} =$

$\|K\|_{L^2(\mathbb{R}^{2d})}$, and:

$$\begin{aligned}
\|K\|_{L^2(\mathbb{R}^{2d})} &= \left(\int_{\mathbb{R}^{2d}} |K(x,t)|^2 dxdt\right)^{1/2} = \\
&= \left(\int_{\mathbb{R}^{2d}} \chi_T(t) \left| \int_{\mathbb{R}^d} e^{2\pi i\omega\cdot(x-t)} \chi_\Omega(\omega) d\omega \right|^2 dxdt\right)^{1/2} = \\
&= \left(\int_{\mathbb{R}^{2d}} \chi_T(t) |\mathscr{F}^{-1}(\chi_\Omega)(x-t)|^2 dxdt\right)^{1/2} \overset{\text{Tonelli}}{=} \\
&= \left[\int_{\mathbb{R}^d} \chi_T(t) \left(\int_{\mathbb{R}^d} |\mathscr{F}^{-1}(\chi_\Omega)(x-t)|^2 dx\right) dt\right]^{1/2} = \\
&= \left(\int_{\mathbb{R}^d} \|T_t \mathscr{F}^{-1}(\chi_\Omega)\|_2^2 \chi_T(t) dt\right)^{1/2} \overset{T_t \text{ isometry+Plancherel}}{=} \\
&= \|\chi_\Omega\|_2 \|\chi_T\|_2 = \sqrt{|T||\Omega|}.
\end{aligned}$$

□

If we compare the integral kernels of $Q_\Omega P_T$ and $P_T Q_\Omega$ we see that $K(x,t) \neq \overline{K(t,x)}$, hence, by Proposition 2.33, we immediately conclude that both operator are not self-adjoint. We already know that, if possible, it is better to deal with self-adjoint operators, so we should consider the following operators:

$$(Q_\Omega P_T)^* Q_\Omega P_T = P_T^* Q_\Omega^* Q_\Omega P_T = P_T Q_\Omega P_T; \tag{4.7}$$

$$(P_T Q_\Omega)^* P_T Q_\Omega = Q_\Omega^* P_T^* P_T Q_\Omega = Q_\Omega P_T Q_\Omega. \tag{4.8}$$

By construction, these are self-adjoint operators, and since both $Q_\Omega P_T$ and $P_T Q_\Omega$ are compact, thanks to Proposition 4.1, Theorem 2.27 and Theorem 2.8 they are also compact. Hence, by Theorem 2.11, they can be diagonalized. In the particular but relevant case where T and Ω are intervals (disks in the multi-dimensional case) the eigenfunctions of these operators are the *prolate spheroidal wave functions* and have been studied by Slepian, Pollak and Landau in a series of papers [30, 17, 18, 29].

4.2 Daubechies' localization operators

The projection operators considered in the previous section fulfil the task of localizing a signal in both time and frequency. However, those are still treated separately. Indeed, if we consider, for example, $Q_\Omega P_T$, we see that at the first moment we perform localization in time and only then in frequency. Since our task is to localize in both domains at the same time, it would be more natural to have an operator that treats time and frequency in a joint way. This is exactly what Ingrid Daubechies did in her remarkable 1988-paper [5].

In Chapter 3 we defined a time-frequency representation of a signal, namely the STFT. Therefore, to reach our goal, it seems more natural to use STFT instead of the Fourier transform. Moreover, from Theorem 3.5 we know that the adjoint operator of $\mathcal{V}_\phi$ acts, in some sense, as an inverse operator. If we choose a window $\phi \in L^2(\mathbb{R}^{2d})$ normalized, 3.10 becomes

$$f(t) = \mathcal{V}_\phi^* \mathcal{V}_\phi f(t).$$

The key idea is to multiply $\mathcal{V}_\phi f$ by a *weight function* $F(x, \omega)$, which logically should highlight some features of $\mathcal{V}_\phi f$, before applying the adjoint operator. This leads to the definition of *time-frequency localization operators*:

$$L_{F,\phi} : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d), \quad L_{F,\phi} f(t) = \mathcal{V}_\phi^* F \mathcal{V}_\phi f(t). \tag{4.9}$$

Related to this localization operator is the sesquilinear form $\mathscr{L}_{F,\phi} : L^2(\mathbb{R}^d) \times L^2(\mathbb{R}^d) \to \mathbb{R}$ defined by the expression:

$$\mathscr{L}_{F,\phi}(f, g) = \int_{\mathbb{R}^{2d}} F(x, \omega) \mathcal{V}_\phi f(x, \omega) \overline{\mathcal{V}_\phi g(x, \omega)} dx d\omega. \tag{4.10}$$

Indeed, assuming $\mathscr{L}_{F,\phi}$ is bounded, we could define $L_{F,\phi} f$ through Riesz' representation theorem as the only element of $L^2(\mathbb{R}^d)$ such that:

$$\mathscr{L}_{F,\phi}(f, g) = \langle L_{F,\phi} f, g \rangle = \int_{\mathbb{R}^d} L_{F,\phi} f(t) \overline{g(t)} \quad \forall g \in L^2(\mathbb{R}^d). \tag{4.11}$$

and therefore $L_{F,\phi}$ as the function which maps f into its representation.

Now that we have defined time-frequency localization operators our goal is to study their properties, starting from boundedness, compactness and the appartenenza to trace class or Hilbert-Schmidt class. We recall that the window function $\phi \in L^2(\mathbb{R}^d)$ is fixed, so it is clear that the properties of $L_{F,\phi}$ will depend upon F.

Proposition 4.2. *Let $F \in L^p(\mathbb{R}^{2d})$ for $p \in [1, +\infty]$. Then $L_{F,\phi}$ is bounded and $\|L_{F,\phi}\| \leq \|F\|_p$.*

Proof. Letting $f, g \in L^2(\mathbb{R}^d)$, we have:

$$|\mathscr{L}_{F,\phi}(f, g)| \leq \int_{\mathbb{R}^{2d}} |F(x, \omega)| |\mathcal{V}_\phi f(x, \omega)| |\mathcal{V}_\phi g(x, \omega)| dx d\omega.$$

From (3.6) we know that, given $f \in L^2(\mathbb{R}^d)$, $\mathcal{V}_\phi f \in L^p(\mathbb{R}^d)$ for every $p \in [2, +\infty]$ and $\|\mathcal{V}_\phi f\|_p \leq \|f\|_2$. We want to find an exponent $q \geq 2$ in order to apply (generalized) Hölder's inequality:

$$\frac{1}{p} + \frac{1}{q} + \frac{1}{q} = 1 \implies q = \frac{2p}{p-1},$$

which is greater or equal than 2, regardless of p. Applying Hölder's inequality with exponents p, q and q we have:

$$|\mathscr{L}_{F,\phi}(f, g)| \leq \|F\|_p \|\mathcal{V}_\phi f\|_q \|\mathcal{V}_\phi g\|_q \leq \|F\|_p \|f\|_2 \|g\|_2.$$

Taking the supremum above all normalized $f, g \in L^2(\mathbb{R}^d)$ gives us the boundedness of $\mathscr{L}_{F,\phi}$ and, in the end, of $L_{F,\phi}$. □

In previous section we managed to prove that projection operators $Q_\Omega P_T$ and $P_T Q_\Omega$ are Hilbert-Schmidt operators, provided both T and Ω have finite measure, which is equivalent to asking that χ_T and χ_Ω are in $L^1(\mathbb{R}^d)$. An analogous result holds for Daubechies' localization operators.

Theorem 4.3. *Let $F \in L^1(\mathbb{R}^{2d})$. Then $L_{F,\phi}$ is an Hilbert-Schmidt integral operator with kernel*

$$K_F(s,t) = \int_{\mathbb{R}^{2d}} F(x,\omega) M_\omega T_x \phi(s) \overline{M_\omega T_x \phi(t)} dx d\omega. \tag{4.12}$$

Moreover, $\|K_F\|_2 \leq \|F\|_1$.

Proof. Let $f, g \in L^2(\mathbb{R}^d)$. We begin showing that $F(x,\omega) f(t) \overline{M_\omega T_x \phi(t)}\, \overline{g(s)} M_\omega T_x \phi(s)$ belongs to $L^1(\mathbb{R}^{2d} \times \mathbb{R}^d \times \mathbb{R}^d)$:

$$\begin{aligned}
&\int_{\mathbb{R}^{4d}} |F(x,\omega) f(t) \overline{M_\omega T_x \phi(t)}\, \overline{g(s)} M_\omega T_x \phi(s)| dx d\omega dt ds \overset{\text{Tonelli}}{=} \\
&= \int_{\mathbb{R}^{2d}} |F(x,\omega)| \left(\int_{\mathbb{R}^d} |f(t)||M_\omega T_x \phi(t)| dt \right) \left(\int_{\mathbb{R}^d} |g(s)||M_\omega T_x g(s)| ds \right) dx d\omega \overset{\text{C−S}}{\leq} \\
&\leq \|f\|_2 \|\phi\|_2 \|g\|_2 \|\phi\|_2 \int_{\mathbb{R}^{2d}} |F(x,\omega)| dx d\omega = \|F\|_1 \|f\|_2 \|g\|_2.
\end{aligned}$$

Now we can apply Fubini's theorem in the expression of $\langle L_{F,\phi} f, g \rangle$:

$$\begin{aligned}
\langle L_{F,\phi} f, g \rangle &= \int_{\mathbb{R}^{2d}} F(x,\omega) \left(\int_{\mathbb{R}^d} f(t) \overline{M_\omega T_x \phi(t)} dt \right) \overline{\left(\int_{\mathbb{R}^d} g(s) \overline{M_\omega T_x \phi(s)} ds \right)} dx d\omega \overset{\text{Fubini}}{=} \\
&= \int_{\mathbb{R}^d} \left[\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^{2d}} F(x,\omega) M_\omega T_x \phi(s) \overline{M_\omega T_x \phi(t)} dx d\omega \right) f(t) dt \right] \overline{g(s)} ds = \\
&= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} K_F(s,t) f(t) dt \right) \overline{g(s)} ds.
\end{aligned}$$

Since this holds for every f and g we can conclude that $L_{F,\phi} f = \int_{\mathbb{R}^d} K_F(\cdot, t) f(t) dt$. Thanks to Proposition 2.14, we know that such integral operator is an Hilbert-Schmidt operator if and only if $K_F \in L^2(\mathbb{R}^{2d})$, so all we have to do is to compute its norm.

$$\begin{aligned}
\|K_F\|_2^2 = |\langle K_F, K_F \rangle| &\leq \int_{\mathbb{R}^{2d}} \left(\int_{\mathbb{R}^{2d}} |F(x,\omega)||M_\omega T_x \phi(s)||M_\omega T_x \phi(t)| dx d\omega \right) \\
&\qquad \cdot \left(\int_{\mathbb{R}^{2d}} |F(y,\xi)||M_\xi T_y \phi(s)||M_\xi T_y \phi(t)| dy d\xi \right) ds dt \overset{\text{Fubini}}{=} \\
&= \int_{\mathbb{R}^{4d}} |F(x,\omega)||F(y,\xi)| \left(\int_{\mathbb{R}} |M_\omega T_x \phi(t)||M_\xi T_y \phi(t)| dt \right) \\
&\qquad \cdot \left(\int_{\mathbb{R}} |M_\omega T_x \phi(s)||M_\xi T_y \phi(s)| ds \right) dx d\omega dy d\xi \overset{\text{C−S}}{\leq} \\
&\leq \|\phi\|_2^4 \int_{\mathbb{R}^{2d}} |F(x,\omega)| dx d\omega \int_{\mathbb{R}^{2d}} |F(y,\xi)| dy d\xi = \|F\|_1^2.
\end{aligned}$$

$\square$

Reminding that Hilbert-Schmidt operators are compact (Theorem 2.27) we observe that, if F is integrable, the corresponding localization operator $L_{F,\phi}$ is compact. Moreover, since we have the explicit expression of the integral kernel, from Proposition 2.33 follows immediately the next sufficient condition on F in order to make $L_{F,\phi}$ self-adjoint.

Proposition 4.4. *If $F \in L^1(\mathbb{R}^{2d})$ is a real-valued function then $L_{F,\phi}$ is self-adjoint.*

We will now prove that localization operators with integrable weight function are trace-class operators. We want to emphasize that, in light of Proposition 2.30, this is a stronger condition than just being an Hilbert-Schmidt operator.

Theorem 4.5. *Let $F \in L^1(\mathbb{R}^{2d})$. Then $L_{F,\phi}$ is a trace-class operator. Moreover, given an orthonormal basis $\{e_n\}_{n\in\mathbb{N}}$ of $L^2(\mathbb{R}^d)$, the following holds:*

$$\sum_{n=1}^{+\infty} |\langle L_{F,\phi} e_n, e_n\rangle| \leq \|F\|_1 \quad \mathrm{tr} L_{F,\phi} = \int_{\mathbb{R}^{2d}} F(x,\omega)\,dx d\omega. \tag{4.13}$$

Proof. We start proving that $L_{F,\phi}$ is a trace-class operator. Given an orthonormal basis $\{e_n\}_{n\in\mathbb{N}}$ of $L^2(\mathbb{R}^d)$:

$$\begin{aligned}
\sum_{n=1}^{+\infty} |\langle L_{F,\phi} e_n, e_n\rangle| &= \sum_{n=1}^{+\infty} \Big| \int_{\mathbb{R}^{2d}} F(x,\omega) \mathcal{V}_\phi e_n(x,\omega) \overline{\mathcal{V}_\phi e_n(x,\omega)} dx d\omega \Big| \leq \\
&\leq \sum_{n=1}^{+\infty} \int_{\mathbb{R}^{2d}} |F(x,\omega)| |\mathcal{V}_\phi e_n(x,\omega)|^2 dx d\omega = \\
&= \int_{\mathbb{R}^{2d}} |F(x,\omega)| \sum_{n=1}^{+\infty} |\langle e_n, M_\omega T_x \phi\rangle|^2 dx d\omega \overset{\text{Parseval}}{=} \\
&= \int_{\mathbb{R}^{2d}} |F(x,\omega)| \|M_\omega T_x \phi\|_2^2 dx d\omega = \|\phi\|_2^2 \|F\|_1 = \|F\|_1,
\end{aligned}$$

where the exchange between series and integral is due to the monotone convergence theorem. Now that we know that $L_{F,\phi}$ is trace-class we can compute its trace:

$$\begin{aligned}
\mathrm{tr} L_{F,\phi} &= \sum_{n=1}^{+\infty} \langle L_{F,\phi} e_n e_n\rangle = \sum_{n=1}^{+\infty} \int_{\mathbb{R}^{2d}} F(x,\omega) |\mathcal{V}_\phi e_n(x,\omega)|^2 dx d\omega = \\
&= \lim_{N\to+\infty} \sum_{n=1}^{N} \int_{\mathbb{R}^{2d}} F(x,\omega) |\mathcal{V}_\phi e_n(x,\omega)|^2 dx d\omega.
\end{aligned}$$

Since

$$|F(x,\omega)| \sum_{n=1}^{N} |\mathcal{V}_\phi e_n(x,\omega)|^2 \leq |F(x,\omega)| \sum_{n=1}^{+\infty} |\mathcal{V}_\phi e_n(x,\omega)|^2 = |F(x,\omega)| \in L^1(\mathbb{R}^{2d}),$$

we can apply Lebesgue's dominated convergence theorem to conclude that:

$$\mathrm{tr} L_{F,\phi} = \lim_{N\to+\infty} \sum_{n=1}^{N} \int_{\mathbb{R}^{2d}} F(x,\omega) |\mathcal{V}_\phi e_n(x,\omega)|^2 dx d\omega = \int_{\mathbb{R}^{2d}} F(x,\omega)\,dx d\omega.$$

□

So far, except for 4.2, we considered only the case $F \in L^1(\mathbb{R}^{2d})$. As the last result of the section we will deal with the more generic case $F \in L^p(\mathbb{R}^{2d})$ for $p < +\infty$.

Proposition 4.6. *Let $F \in L^p(\mathbb{R}^{2d})$ with $1 \leq p < \infty$. Then the corresponding localization operator $L_{F,\phi}$ is compact.*

Proof. Given $F \in L^p(\mathbb{R}^{2d})$, thanks to Theorem 2.5 it is sufficient to consider a sequence F_n of functions in $L^1(\mathbb{R}^{2d})$ such that $F_n \to F$ in $L^p(\mathbb{R}^{2d})$. For example, we can suppose that F_n are in Schwartz's class $\mathcal{S}(\mathbb{R}^{2d})$, which is a well-known dense subspace of $L^p(\mathbb{R}^{2d})$ for $p < +\infty$. Indeed, from Proposition 4.2, we have that $\|L_{F_n,\phi} - L_{F,\phi}\| \leq \|F_n - F\|_p$, so $L_{F_n,\phi} \to L_{F,\phi}$ in $\mathscr{B}(L^2(\mathbb{R}^d))$. Since $\mathcal{S}(\mathbb{R}^{2d}) \subset L^1(\mathbb{R}^{2d})$, $L_{F_n,\phi}$ are compact, thus also $L_{F,\phi}$ is. □

4.2.1 Spherically Symmetric Weights

In previous section we managed to prove that, if the weight function F is in $L^p(\mathbb{R}^{2d})$ for some $p < +\infty$ and it is real-valued then the corresponding localization operator $L_{F,\phi}$ is compact and self-adjoint. Thus, it is a natural to ask which are its eigenfuctions with corresponding eigenvalues. However, this in general is not feasible. Hence, we shall consider some specific class of weight and window functions. In particular, in this section we will consider the special case in which the window for the STFT is a Gaussian (3.11) and the weight F is spherically symmetric. Letting $r_j^2 = x_j^2 + \omega_j^2$ for $j = 1, \ldots, d$ and $r^2 = (r_1^2 \ldots, r_d^2) \in \mathbb{R}^d$, the hypothesis about F can be rephrased in the following way

$$F(x, \omega) = \mathcal{F}(r^2). \tag{4.14}$$

In order to highlight the dependence of F through $\mathcal{F}$ the corresponding localization operator will be denoted as $L_{\mathcal{F},\varphi}$. For this operators a complete characterization of the spectrum and eigenspaces is given in the already cited paper of Daubechies [5].

Before stating we need to introduce some special function, namely *Hermite functions*. In dimension $d = 1$, Hermite functions are given by

$$H_k(t) = \frac{2^{1/4}}{\sqrt{k!}} \left(-\frac{1}{2\sqrt{\pi}} \right)^k e^{\pi t^2} \frac{d^k}{dt^k} \left(e^{-2\pi t^2} \right), \tag{4.15}$$

where $k \in \mathbb{N} \cup \{0\} \coloneqq \mathbb{N}_0$. Hermite functions have lots of interesting and useful properties. A standard reference is [6, Section 1.7]. We cite some of them which will be useful in the following.

(i) $\{H_k\}_{k \in \mathbb{N}_0}$ is an orthonormal basis of $L^2(\mathbb{R})$;

(ii) $H_0(t) = \varphi(t)$, where φ is the normalized Gaussian given by (3.11);

(iii) Setting $H_{-1} = 0$, the following recursive relation holds

$$2\sqrt{\pi} t H_k = \sqrt{k+1} H_{k+1} + \sqrt{k} H_{k-1} \quad \text{for } k = 0,1,\ldots; \tag{4.16}$$

(iv) Hermite functions are eigenfunctions of $\mathscr{F}$, specifically

$$\mathscr{F} H_k = (-i)^k H_k. \tag{4.17}$$

Hermite functions in generic dimension are just the tensor product of 1-dimensional Hermite functions. Explicitly, given a multi-index $k = (k_1, \dots, k_d) \in \mathbb{N}_0^d$ the corresponding Hermite function is given by:

$$H_k(t) = \prod_{j=1}^{d} H_{k_j}(t_j). \tag{4.18}$$

It is still true that d-dimensional Hermite functions (now ranging between all possible multi-indices) form an orthonormal basis of $L^2(\mathbb{R}^d)$. Moreover, using (4.17), it is easy to see that d-dimensional Hermite functions are still eigenfunction of the Fourier transform and that the following holds:

$$\mathscr{F} H_k = (-i)^{|k|} H_k, \tag{4.19}$$

where $|k| = k_1 + \cdots k_d$ is the length of the multi-index.

The introduction of Hermite functions is necessary, since they are exactly the eigenfunctions of $L_{\mathcal{F},\varphi}$.

Theorem 4.7. *Eigenfuctions of $L_{\mathcal{F},\varphi}$ are the d-dimensional Hermite functions H_k, with corresponding eigenvalues:*

$$\lambda_k = \frac{1}{k!} \int_0^{+\infty} \cdots \int_0^{+\infty} \mathcal{F}\left(\frac{s_1}{\pi}, \dots, \frac{s_d}{\pi}\right) \left(\prod_{j=1}^{d} s_j^{k_j}\right) e^{-(s_1 + \cdots + s_d)} ds_1 \cdots ds_d, \tag{4.20}$$

where $k \in \mathbb{N}_0^d$ and $k! = k_1! \cdots k_d!$.

Before proving the theorem we need the following lemma.

Lemma 4.8. *Given $z \in \mathbb{C}$, it holds:*

$$\int_{\mathbb{R}} e^{-2\pi(t^2 + zt)} dt = \frac{1}{2^{1/2}} e^{\pi z^2/2}. \tag{4.21}$$

Proof. Letting $z = \mathrm{Re} z + i \mathrm{Im} z = u + iv$ we have:

$$\begin{aligned}
\int_{\mathbb{R}} e^{-2\pi(t^2+zt)} dt &= \int_{\mathbb{R}} e^{-2\pi(t^2+ut)} e^{-2\pi i v t} dt = e^{\pi u^2/2} e^{-2\pi(t^2+ut+u^2/4)} e^{-2\pi i v t} dt \\
&= e^{\pi u^2/2} \int_{\mathbb{R}} e^{-2\pi(t+u/2)^2} e^{-2\pi i v t} dt = e^{\pi u^2/2} \mathscr{F}(T_{-u/2} e^{-2\pi(\cdot)^2})(v) \overset{2.42(i)+2.43}{=} \\
&= e^{\pi u^2/2} e^{2\pi i (u/2) v} \frac{1}{2^{1/2}} e^{-\pi v^2/2} = \frac{1}{2^{1/2}} e^{\pi(u^2+2iuv-v^2)/2} = \frac{1}{2^{1/2}} e^{\pi z^2/2}
\end{aligned}$$

□

Proof of Theorem 4.7. Since Hermite functions are an orthonormal basis of $L^2(\mathbb{R}^n)$ it is sufficient to prove that $\langle L_{\mathcal{F},\varphi}H_k, H_l\rangle = \langle \mathcal{F}\mathcal{V}_\varphi H_k, \mathcal{V}_\varphi H_l\rangle = \lambda_k \prod_{j=1}^d \delta_{k_j,l_j}$, which means that the scalar product is different from zero if and only if $k = l$. We start by computing the STFT of an Hermite function:

$$\begin{aligned}
\mathcal{V}_\varphi H_k(x,\omega) &= \int_{\mathbb{R}^d} H_k(t)e^{-2\pi i\omega\cdot t}\, 2^{d/4}e^{-\pi|t-x|^2}dt = \int_{\mathbb{R}^d}\prod_{j=1}^d H_{k_j}(t_j)e^{-2\pi i\omega\cdot t}\, 2^{d/4}e^{-\pi|t-x|^2}dt = \\
&= \prod_{j=1}^d 2^{1/4}\int_{\mathbb{R}} H_{k_j}(t_j)e^{-2\pi i\omega_j t_j}e^{-\pi(t_j-x_j)^2}dt_j = \\
&= \prod_{j=1}^d 2^{1/4}\int_{\mathbb{R}} \frac{2^{1/4}}{\sqrt{k_j!}}\left(-\frac{1}{2\sqrt{\pi}}\right)^{k_j} e^{\pi t_j^2}\frac{d^{k_j}}{dt^{k_j}}\left(e^{-2\pi t_j^2}\right)e^{-2\pi i\omega_j t_j}e^{-\pi(t_j-x_j)^2}dt_j = \\
&= \prod_{j=1}^d \frac{2^{1/2}}{\sqrt{k_j!}}\left(-\frac{1}{2\sqrt{\pi}}\right)^{k_j}\int_{\mathbb{R}}\frac{d^{k_j}}{dt^{k_j}}\left(e^{-2\pi t_j^2}\right)e^{\pi(t_j^2-2i\omega_j t_j-t_j^2+2t_jx_j-x_j^2)}dt_j = \\
&= \prod_{j=1}^d \frac{2^{1/2}}{\sqrt{k_j!}}\left(-\frac{1}{2\sqrt{\pi}}\right)^{k_j}e^{-\pi x_j^2}\int_{\mathbb{R}}\frac{d^{k_j}}{dt^{k_j}}\left(e^{-2\pi t_j^2}\right)e^{2\pi(x_j-i\omega_j)t_j}dt_j \overset{\text{integration by parts}}{=} \\
&= \prod_{j=1}^d \frac{2^{1/2}}{\sqrt{k_j!}}\left(-\frac{1}{2\sqrt{\pi}}\right)^{k_j}e^{-\pi x_j^2}[2\pi(x_j-i\omega_j)]^{k_j}(-1)^{k_j}\int_{\mathbb{R}}e^{-2\pi[t_j^2-(x_j-i\omega_j)t_j]}dt_j = \\
&= \prod_{j=1}^d \sqrt{\frac{\pi^{k_j}}{k_j!}}2^{1/2}(x_j-i\omega_j)^{k_j}e^{-\pi x_j^2}\frac{1}{2^{1/2}}e^{\pi(x_j-i\omega_j)^2/2} = \\
&= \prod_{j=1}^d \sqrt{\frac{\pi^{k_j}}{k_j!}}(x_j-i\omega_j)^{k_j}e^{-\pi i\omega_j x_j}e^{-\pi(x_j^2+\omega_j^2)/2} = \\
&= \left(\prod_{j=1}^d \sqrt{\frac{\pi^{k_j}}{k_j!}}(x_j-i\omega_j)^{k_j}\right)e^{-\pi i\omega\cdot x}e^{-\pi(r_1^2+\dots+r_d^2)/2}. \qquad (4.22)
\end{aligned}$$

Before computing the scalar product between $L_{\mathcal{F},\varphi}H_k$ and H_l, we introduce the angular coordinate θ_j, such that $x_j + i\omega_j = r_j e^{i\theta_j}$. Therefore we have:

$$\begin{aligned}
\langle L_{\mathcal{F},\varphi}H_k, H_l\rangle &= \int_{\mathbb{R}^{2d}} F(x,\omega)\left(\prod_{j=1}^d\sqrt{\frac{\pi^{k_j}}{k_j!}}r_j^{k_j}e^{-ik_j\theta_j}\right)e^{-\pi i\omega\cdot x}e^{-\pi(r_1^2+\dots+r_d^2)/2}\cdot \\
&\qquad \overline{\left(\prod_{m=1}^d\sqrt{\frac{\pi^{l_m}}{l_m!}}r_m^{l_m}e^{-il_m\theta_m}\right)e^{-\pi i\omega\cdot x}e^{-\pi(r_1^2+\dots+r_d^2)/2}}dxd\omega = \\
&= \int_{\mathbb{R}^{2d}} F(x,\omega)\left(\prod_{j=1}^d\sqrt{\frac{\pi^{k_j+l_j}}{k_j!l_j!}}r_j^{k_j+l_j}e^{i(l_j-k_j)\theta_j}\right)e^{-\pi(r_1^2+\dots+r_d^2)}dxd\omega. \qquad (4.23)
\end{aligned}$$

For every pair of coordinates (x_j,ω_j) we can switch to polar coordinates (r_j,θ_j). Since $F(x,\omega) = \mathcal{F}(r^2)$ is independent of angular coordinates, only functions depending on those are $e^{i(l_j-k_j)\theta_j}$, for which:

$$\int_0^{2\pi} e^{i(l_j-k_j)\theta_j}d\theta_j = \begin{cases} \left.\dfrac{e^{i(l_j-k_j)\theta_j}}{i(l_j-k_j)}\right|_0^{2\pi} = 0 & \text{if } k_j \neq l_j \\ \displaystyle\int_0^{2\pi} d\theta_j = 2\pi & \text{if } k_j = l_j \end{cases}.$$

Therefore, if $k \neq l$, the whole integral is 0, otherwise, letting $k_j = l_j$ in (4.23):

$$\langle L_{\mathcal{F},\varphi}H_k, H_k\rangle = (2\pi)^d \frac{\pi^{|k|}}{k!}\int_0^{+\infty}\cdots\int_0^{+\infty} \mathcal{F}(r_1^2,\dots,r_d^2)e^{-\pi(r_1^2+\dots+r_d^2)}\left(\prod_{j=1}^d r_j^{2k_j+1}\right)dr_1\cdots dr_d =$$
$$= \frac{1}{k!}\int_0^{+\infty}\cdots\int_0^{+\infty} \mathcal{F}(r_1^2,\dots,r_d^2)e^{-\pi(r_1^2+\dots+r_d^2)}\left(\prod_{j=1}^d (\pi r_j^2)^{k_j}\right)\pi r_1 dr_1\cdots \pi r_d dr_d$$

With the change of variable $s_j = \pi r_j^2$, we finally obtain:

$$\langle L_{\mathcal{F},\varphi}H_k, H_k\rangle = \frac{1}{k!}\int_0^{+\infty}\cdots\int_0^{+\infty} \mathcal{F}\left(\frac{s_1}{\pi},\dots,\frac{s_d}{\pi}\right)\left(\prod_{j=1}^d s_j^{k_j}\right)e^{-(s_1+\dots+s_d)}ds_1\cdots ds_d$$

which is exactly the expression (4.20). □

In conclusion, we will consider two meaningful examples, namely when the weight F is the characteristic function of a disk centred around the origin and when it is a Gaussian. In order to make computations easier we confine ourselves in the case $d = 1$.

Example 4.9 (Localization on a disk)**.** We study arguably the most simple case, namely when F is the characteristic function of the disk $\mathcal{B}_R = \{(x,\omega)\in\mathbb{R}^2 : x^2+\omega^2 \leq R^2\}$:

$$F(x,\omega) = \mathcal{F}(r^2) = \begin{cases} 1 & \text{if } x^2+\omega^2 = r^2 \leq R^2 \\ 0 & \text{otherwise} \end{cases}.$$

In order to highlight that F is the characteristic function of $\mathcal{B}_R$ we let $L_{\mathcal{F},\varphi} = L_{\mathcal{B}_R,\varphi}$. Noticing that $\mathcal{F}(\frac{s}{\pi}) = \chi_{[0,\pi R^2]}(s)$, expression (4.20) brings to:

$$\lambda_k(R) = \frac{1}{k!}\int_0^{+\infty}\chi_{[0,\pi R^2]}(s)s^k e^{-s}ds = \frac{1}{k!}\int_0^{\pi R^2} s^k e^{-s}ds = \gamma(k+1,\pi R^2)$$

where γ is the lower incomplete gamma function. An easy integration by parts, when $k \geq 1$, leads to:

$$\int_0^{\pi R^2} s^k e^{-s}ds = -(\pi R^2)^k e^{-\pi R^2} + k\int_0^{\pi R^2} s^{k-1}e^{-s}ds.$$

Iterating this process gives us the following formula for the k-th eigenvalue:

$$\lambda_k = 1 - e^{-\pi R^2} \sum_{j=0}^{k} \frac{(\pi R^2)^j}{j!}, \quad k = 0,1,\dots$$

Since $(\pi R^2)^j/j!$ is strictly positive, it follows immediately that the sequence of eigenvalues is strictly decreasing. Moreover, since F is real-valued, from Proposition 4.4 we have that $L_{\mathcal{B}_R,\varphi}$ is self-adjoint, as well as compact. Therefore, from Corollay 2.12 we conclude that:

$$\|L_{\mathcal{B}_R,\varphi}\| = |\lambda_0| = 1 - e^{-\pi R^2}.$$

Recalling the definition of the norm for operators between Hilbert spaces 2.2, we obtain that, for every normalized $f \in L^2(\mathbb{R}^d)$:

$$\lambda_0 = \|L_{\mathcal{B}_R,\varphi}\| \geq |\langle L_{\mathcal{B}_R,\varphi} f, f\rangle| = \int_{\mathcal{B}_R} |\mathcal{V}_\varphi f(x,\omega)|^2 dx d\omega \implies$$
$$\implies \int_{\mathcal{B}_R} |\mathcal{V}_\varphi f(x,\omega)|^2 dx d\omega \leq 1 - e^{-\pi R^2}. \tag{4.24}$$

We point out that the left-hand side of the last expression represents the energy of $\mathcal{V}_\varphi f$ concentrated on the disk $\mathcal{B}_R$.

Example 4.10 (Localization with Gaussian weight)**.** Another natural choice for the weight function F is a Gaussian:

$$F(x,\omega) = e^{-\alpha\pi(x^2+\omega^2)},$$

where $\alpha > 0$ is a dilation parameter. In this case $\mathcal{F}(\frac{s}{\pi}) = e^{-\alpha s}$, so from (4.20) and integrating by parts $k+1$ times we obtain the eigenvalues of $L_{\mathcal{F},\varphi}$:

$$\lambda_k = \frac{1}{k!}\int_0^{+\infty} s^k e^{-(1+\alpha)s} ds = (1+\alpha)^{-(k+1)}, \quad k = 0,1,\dots$$

Like the previous case, eigenvalues are already ordered in decreasing order and F is still real-valued, therefore

$$\|L_{\mathcal{F},\varphi}\| = 1 + \alpha.$$

Chapter 5

Uncertainty principles

Up to know we put some effort in constructing some tools that have the ability to concentrate a signal in the time-frequency domain. In the introduction of Section 3.1 we also pointed out that a characteristic function is not a "good" window for the STFT because, in light of the duality between regularity and decay, the Fourier transform of a not regular functions has a slow decay. Indeed, in Section 4.2.1, we considered the STFT with a Gaussian window function, which has nice regularity and decay properties. However, a natural question arises: how good can we concentrate a signal? Is it possible to have a signal arbitrarily concentrated both in time and frequency? The answer to this questions is definitely no and it is given by *uncertainty principles*, which are ubiquitous results in Fourier and time-frequency analysis. Uncertainty principles arise in different versions but the main underlying idea is the following:
a function cannot be too concentrated both in time and frequency.

Even if not explicit, we already had a first glimpse of this phenomenon when we computed the Fourier transform of dilated Gaussian (Proposition 2.43). Indeed, given $\lambda > 0$ we recall that:

$$\mathscr{F}(e^{-\lambda\pi|\cdot|^2})(\omega) = \frac{1}{\lambda^{d/2}} e^{-\frac{1}{\lambda}\pi|\omega|^2} \quad \forall\, \omega \in \mathbb{R}^d.$$

If we choose λ to be very large, the Gaussian in the time domain $e^{-\lambda\pi|t|^2}$ will be strongly concentrated around the origin. However, in the corresponding Gaussian in the frequency domain the dilation parameter appears in the denominator of the exponent, so this will be poorly concentrated.

In this chapter we will present some uncertainty principles, both for the Fourier transform and the STFT and we will give a quantitative description of how good we can localize a signal.

5.1 Heisenberg's uncertainty principle

Arguably, the most famous uncertainty principle is the one named after Heisenberg ([12]). Despite being a fascinating topic, we will not discuss all the implications that this uncertainty principles has in quantum mechanics. Therefore, our attention is driven to the mathematical formulation of the principle.

In literature there are several proofs of Heisenberg's uncertainty principle. The one that we will present was given (in the 1-dimensional case) by de Bruijn in [3] and involves, once again, Hermite functions. Before stating and proving Heisenberg's uncertainty principle we are going to need some lemmas.

Lemma 5.1. *Let $f \in L^2(\mathbb{R})$. Then*

$$\int_{\mathbb{R}} t^2|f(t)|^2 dt + \int_{\mathbb{R}} \omega^2 |\hat{f}(\omega)|^2 d\omega = \frac{1}{2\pi} \sum_{k=0}^{+\infty} (2k+1)|\langle f, H_k\rangle|^2, \tag{5.1}$$

where H_k is the k-th Hermite function. In particular we have

$$\int_{\mathbb{R}} t^2|f(t)|^2 dt + \int_{\mathbb{R}} \omega^2 |\hat{f}(\omega)|^2 d\omega \geq \frac{\|f\|_2^2}{2\pi}, \tag{5.2}$$

with equality if and only if f is a multiple of H_0.

Proof. Our aim is to exploit the fact that Hermite functions are an orthonormal basis of $L^2(\mathbb{R})$ by computing $\langle tf, H_k\rangle$. Before going on we remark that the previous expression is not an L^2 scalar product but is a duality between a tempered distribution and a function in the Schwartz class. However, from the theory of tempered distributions we have that $\langle tf, H_k\rangle = \langle f, tH_f\rangle$ and the latter expression is indeed a scalar product in $L^2(\mathbb{R}^d)$. Moreover, from this observation, we see that tH_k appears. Therefore, using (4.16) we have:

$$\langle tf, H_k\rangle = \langle f, tH_k\rangle = \frac{1}{2\sqrt{\pi}} \left(\sqrt{k+1}\langle f, H_{k+1}\rangle + \sqrt{k}\langle f, H_{k-1}\rangle \right).$$

To compute the similar quantity for $\hat{f}$ we also need to recall that Hermite functions are eigenfunction of the Fourier transform (4.17):

$$\begin{aligned}
\langle \omega\hat{f}, H_k\rangle = \langle \hat{f}, \omega H_k\rangle &= \frac{1}{2\sqrt{\pi}} \left(\sqrt{k+1}\langle \hat{f}, H_{k+1}\rangle + \sqrt{k}\langle \hat{f}, H_{k-1}\rangle \right) \overset{(4.17)+\mathscr{F}\text{ unitary}}{=} \\
&= \frac{1}{2\sqrt{\pi}} \left(\sqrt{k+1}(-i)^{k+1}\langle f, H_{k+1}\rangle + \sqrt{k}(-i)^{k-1}\langle f, H_{k-1}\rangle \right) = \\
&= \frac{1}{2\sqrt{\pi}} (-i)^{k-1} \left(\sqrt{k}\langle f, H_{k-1}\rangle - \sqrt{k+1}\langle f, H_{k+1}\rangle \right).
\end{aligned}$$

Now, since $\{H_k\}_{k\in\mathbb{N}_0^d}$ is an orthonormal basis of $L^2(\mathbb{R})$, we can use Parseval's identity:

$$\begin{aligned}
\int_{\mathbb{R}} t^2|f(t)|^2 dt + \int_{\mathbb{R}} \omega^2 |\hat{f}(\omega)|^2 d\omega &= \sum_{k=0}^{+\infty} \left(|\langle tf, H_k\rangle|^2 + |\langle \omega\hat{f}, H_k\rangle|^2 \right) = \\
&= \frac{1}{2\pi} \sum_{k=0}^{+\infty} \left[(k+1)|\langle f, H_{k+1}\rangle|^2 + k|\langle f, H_{k-1}\rangle| \right] = \\
&= \frac{1}{2\pi} \sum_{k=0}^{+\infty} (2k+1)|\langle f, H_k\rangle|^2.
\end{aligned}$$

Since $(2k+1) \geq 1$, it is immediate to see that inequality (5.2) holds and that equality is achieved if and only if $\langle f, H_k \rangle = 0$ for every $k > 1$, which means exactly that f is a multiple of H_0. □

In order to extend previous lemma to the multi-dimensional case we need the following result related to the Fourier transform of restrictions. Before stating the lemma we introduce the following notation:

$$t' = (t_2, \ldots, t_d) \in \mathbb{R}^{d-1}, \quad \omega' = (\omega_2, \ldots, \omega_d) \in \mathbb{R}^{d-1},$$
$$(\mathscr{F}_1 f)(\omega_1, t') = \mathscr{F}(f(\cdot, t'))(\omega_1), \quad (\mathscr{F}' f)(t_1, \omega') = \mathscr{F}(f(t_1, \cdot))(\omega').$$

Lemma 5.2. *Let $f \in L^2(\mathbb{R}^d)$. Then $\mathscr{F}_1 f(\omega_1, \cdot) \in L^2(\mathbb{R}^{d-1})$ for almost every $\omega_1 \in \mathbb{R}$ and $\mathscr{F}'(\mathscr{F}_1 f(\omega_1, \cdot))(\omega') = \mathscr{F} f(\omega)$.*

Proof. Before starting, we point out that, since $f \in L^2(\mathbb{R}^d)$, $f(\cdot, t') \in L^2(\mathbb{R})$ for almost every $t' \in \mathbb{R}^{d-1}$, therefore $\mathscr{F}_1 f(\cdot, t')$ is well-defined for almost every $t' \in \mathbb{R}^{d-1}$. Then we have:

$$\int_{\mathbb{R}^d} |\mathscr{F}_1 f(\omega_1, t')|^2 d\omega_1 dt' \overset{\text{Tonelli}}{=} \int_{\mathbb{R}^{d-1}} \int_{\mathbb{R}} |\mathscr{F}(f(\cdot, t'))(\omega_1)|^2 d\omega_1 dt' \overset{\text{Plancherel}}{=}$$
$$= \int_{\mathbb{R}^{d-1}} \int_{\mathbb{R}} |f(t_1, t')|^2 dt_1 dt' = \|f\|_2.$$

which proves that $\mathscr{F}_1 f(\omega_1, \cdot)$ is in $L^2(\mathbb{R}^{d-1})$ for almost every $\omega_1 \in \mathbb{R}$.

Now suppose that $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$. Since f is in $L^1(\mathbb{R}^d)$ we can use (2.19) to make $\mathscr{F}_1 f$ explicit:

$$\mathscr{F}_1 f(\omega_1, t') = \int_{\mathbb{R}} f(t_1, t') e^{-2\pi i \omega_1 t_1} dt_1 \implies$$
$$\mathscr{F}'(\mathscr{F}_1 f(\omega_1, \cdot))(\omega') = \int_{\mathbb{R}^{d-1}} \left(\int_{\mathbb{R}} f(t_1, t') e^{-2\pi i \omega_1 t_1} dt_1 \right) e^{-2\pi i \omega' \cdot t'} dt' \overset{\text{Fubini}}{=}$$
$$= \int_{\mathbb{R}^d} f(t_1, t') e^{-2\pi i \omega_1 t_1} e^{-2\pi i \omega' \cdot t'} dt_1 dt' = \mathscr{F} f(\omega).$$

Through the density of $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ in $L^2(\mathbb{R}^d)$ the last part of the statement follows. □

Lemma 5.3. *Let $f \in L^2(\mathbb{R}^d)$. Then, for every $j = 1, \ldots, d$:*

$$\int_{\mathbb{R}^d} t_j^2 |f(t)|^2 dt + \int_{\mathbb{R}^d} \omega_j^2 |\hat{f}(\omega)|^2 d\omega = \frac{1}{2\pi} \sum_{k \in \mathbb{N}_0^d} (2k_j + 1) |\langle f, H_k \rangle|^2. \tag{5.3}$$

In particular we have:

$$\int_{\mathbb{R}^d} t_j^2 |f(t)|^2 dt + \int_{\mathbb{R}^d} \omega_j^2 |\hat{f}(\omega)|^2 d\omega \geq \frac{\|f\|_2^2}{2\pi}, \tag{5.4}$$

with equality if and only if f is a multiple of H_0.

Proof. Without loss of generality we consider the case $j = 1$. Moreover, we introduce the following notation:

$$\langle f, g\rangle_1 = \langle f(\cdot, t'), g(\cdot, t')\rangle \quad \langle f, g\rangle' = \langle f(t_1, \cdot), g(t_1, \cdot)\rangle$$

From Lemma 5.1, since $f(\cdot, t') \in L^2(\mathbb{R})$ for a.e. $t' \in \mathbb{R}^{d-1}$, we have:

$$\int_{\mathbb{R}} t_1^2 |f(t_1, t')|^2 dt_1 + \int_{\mathbb{R}} \omega_1^2 |\mathscr{F}_1 f(\omega_1, t')|^2 d\omega_1 = \frac{1}{2\pi} \sum_{k_1=0}^{+\infty} (2k_1 + 1) |\langle f, H_{k_1}\rangle_1|^2$$

which holds for almost every $t' \in \mathbb{R}^{d-1}$. We can now integrate with respect to t' every member:

- $\displaystyle\int_{\mathbb{R}^{d-1}} \int_{\mathbb{R}} t_1^2 |f(t_1, t')|^2 dt_1 dt' \overset{\text{Tonelli}}{=} \int_{\mathbb{R}^d} t_1^2 |f(t)|^2 dt;$
- $\displaystyle\int_{\mathbb{R}^{d-1}} \int_{\mathbb{R}} \omega_1^2 |\mathscr{F}_1 f(\omega_1, t')|^2 d\omega_1 dt' \overset{\text{Tonelli}}{=} \int_{\mathbb{R}} \omega_1^2 \int_{\mathbb{R}^{d-1}} |\mathscr{F}_1 f(\omega_1, t')|^2 dt' d\omega_1.$

  From Lemma 5.2 we know that $\mathscr{F}_1 f(\omega_1) \in L^2(\mathbb{R}^{d-1})$ for a.e. $\omega_1 \in \mathbb{R}$. Therefore, we can use Plancherel's theorem in the inner integral for a.e. $\omega_1 \in \mathbb{R}$ and obtain:

$$\begin{aligned} \int_{\mathbb{R}^{d-1}} \int_{\mathbb{R}} \omega_1^2 |\mathscr{F}_1 f(\omega_1, t')|^2 d\omega_1 dt' &= \int_{\mathbb{R}} \omega_1^2 \int_{\mathbb{R}^{d-1}} |\mathscr{F}'(\mathscr{F}_1 f(\omega_1, \cdot))(\omega')|^2 d\omega' d\omega_1 \overset{5.2}{=} \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}^{d-1}} \omega_1^2 |\mathscr{F} f(\omega_1, \omega')|^2 d\omega' d\omega_1 = \int_{\mathbb{R}^d} \omega_1^2 |\mathscr{F}(\omega)|^2 d\omega. \end{aligned}$$

- For the last term we start pointing out that integral and series can be exchanged because every term is non-negative. Then, from Parseval's identity we have:

$$\begin{aligned} \int_{\mathbb{R}^{d-1}} |\langle f, H_{k_1}\rangle_{t_1}|^2 dt' &= \sum_{k' \in \mathbb{N}_0^{d-1}} |\langle\langle f, H_{k_1}\rangle_{t_1}, H_{k'}\rangle_{t'}|^2 = \\ &= \sum_{k' \in \mathbb{N}_0^{d-1}} \left| \int_{\mathbb{R}^{d-1}} \left(\int_{\mathbb{R}} f(t_1, t') \overline{H_{k_1}(t_1)} dt_1 \right) \overline{H_{k'}(t')} dt' \right|^2 \overset{\text{Fubini}}{=} \\ &= \sum_{k' \in \mathbb{N}_0^{d-1}} \left| \int_{\mathbb{R}^d} f(t) \overline{H_{(k,k')}(t)} dt \right|^2 = \sum_{k' \in \mathbb{N}_0^{d-1}} |\langle f, H_{(k,k')}\rangle|^2. \end{aligned}$$

  where we used the fact that multi-dimensional Hermite functions are just the tensor product of 1-dimensional ones. Plugging this result in the series leads to:

$$\begin{aligned} \int_{\mathbb{R}^{d-1}} \sum_{k_1=0}^{+\infty} (2k_1 + 1) |\langle f, H_{k_1}\rangle|^2 &= \sum_{k_1=0}^{+\infty} \sum_{k' \in \mathbb{N}_0^{d-1}} (2k_1 + 1) |\langle f, H_{(k,k')}\rangle|^2 = \\ &= \sum_{k \in \mathbb{N}_0^d} (2k_1 + 1) |\langle f, H_k\rangle|^2, \end{aligned}$$

and we notice that the rearrangement of the series is allowed since convergence is unconditional. Putting all these results together leads to (5.3).

The proof of the last part of the statement is exactly the same as the one in Lemma 5.1.

□

We are now in the position to prove Heisenberg's uncertainty principle.

Theorem 5.4. *Let $f \in L^2(\mathbb{R}^d)$ and $a, b \in \mathbb{R}^d$. Then:*

$$\left(\int_{\mathbb{R}^d} |t-a|^2 |f(t)|^2 dt\right)^{1/2} \left(\int_{\mathbb{R}^d} |\omega - b|^2 |\hat{f}(\omega)|^2 d\omega\right)^{1/2} \geq \frac{d\|f\|_2^2}{4\pi}, \tag{5.5}$$

where $|t-a| = \sum_{j=1}^d (t_j - a_j)^2$ is the Euclidean distance. Moreover, equality is achieved if and only if $f(t) = c\varphi(\lambda t)$, where φ is the normalized Gaussian given by (3.11), $c \in \mathbb{C}$ and $\lambda > 0$.

Proof. Firsly, we notice that it is sufficient to prove the inequality when $a = b = 0$, since the generic case can be recovered from this special one by means of a phase-space translation. Indeed, given $f \in L^2(\mathbb{R}^d)$ we can consider $g = M_{-b}T_{-a}f$ for which we have:

$$|f(t)|^2 = |(T_a M_b g)(t)|^2 = |e^{2\pi i b\cdot(t-a)} g(t-a)|^2 = |g(t-a)|^2,$$
$$|\hat{f}(\omega)|^2 = |\mathscr{F}(T_a M_b g)(\omega)|^2 \overset{2.42}{=} |M_{-a} T_b \hat{g}(\omega)|^2 = |e^{-2\pi i a\cdot\omega} \hat{g}(\omega - b)|^2 = |\hat{g}(\omega - b)|^2,$$

and $\|f\|_2 = \|g\|_2$. In light of this, we will consider $a = b = 0$.

We start proving that, for every component, the following holds:

$$\left(\int_{\mathbb{R}^d} t_j^2 |f(t)|^2 dt\right)^{1/2} \left(\int_{\mathbb{R}^d} \omega_j^2 |\hat{f}(\omega)|^2 d\omega\right)^{1/2} \geq \frac{\|f\|_2^2}{4\pi}. \tag{5.6}$$

We observe that in the left-hand side of (5.6), apart from a square root, we have the product of two integrals, while in (5.4) we have an estimate for the sum of these. The transition from the latter to the former estimate can be done through a dilation argument. Precisely, given $f \in L^2(\mathbb{R}^d)$, we consider the following dilation:

$$g(t) = \lambda^{-d/2} f(t/\lambda) \implies \hat{g}(\omega) = \lambda^{d/2} \mathscr{F}(D_{1/\lambda} g) \overset{2.42(iii)}{=} \lambda^{d/2} \hat{f}(\lambda\omega).$$

We remark that this dilation is different from the one considered in Section 2.2, expression (2.27). Indeed, now we chose the dilation so that $\|g\|_2 = \|f\|_2$, while previously the dilation was chosen in order to preserve the L^1 norm. Putting g in (5.4) provides us:

$$\begin{aligned}
\frac{\|f\|_2^2}{2\pi} &\leq \int_{\mathbb{R}^d} t_j^2 |g(t)|^2 dt + \int_{\mathbb{R}^d} \omega_j^2 |\hat{g}(\omega)|^2 d\omega = \qquad (5.7)\\
&= \frac{1}{\lambda^d} \int_{\mathbb{R}^d} t_j^2 |f(t/\lambda)|^2 dt + \lambda^d \int_{\mathbb{R}^d} \omega^2 |\hat{f}(\lambda\omega)|^2 d\omega = \\
&= \lambda^2 \int_{\mathbb{R}^d} t_j^2 |f(t)|^2 dt + \frac{1}{\lambda^2} \int_{\mathbb{R}^d} \omega_j^2 |\hat{f}(\omega)|^2 d\omega.
\end{aligned}$$

We can choose λ in order to minimize the last expression. Thus, deriving with respect to λ^2 and putting the derivative to 0 we obtain that the minimum is achieved when

$$\lambda^2 = \left(\int_{\mathbb{R}^d} \omega_j^2 |\hat{f}(\omega)|^2 d\omega\right)^{1/2} \left(\int_{\mathbb{R}^d} t_j^2 |f(t)|^2 dt\right)^{-1/2}. \tag{5.8}$$

If we substitute this λ into the last expression we obtain exactly (5.6). Moreover, equality is achieved if and only if it is achieved in (5.7), but Lemma 5.3 then implies that $g(t) = cH_0$ for some $c \in \mathbb{C}$. Recalling the definition of g we obtain $f(t) = c\lambda^{d/2} H_0(\lambda t)$. If we put the explicit expression of f into (5.8) we will see that λ can be chosen arbitrarily, indeed:

$$\begin{aligned}
\lambda^2 &= \left(\int_{\mathbb{R}^d} \omega_j^2 |c|^2 \lambda^{-d} |\mathscr{F}(D_\lambda H_0)(\omega)|^2 d\omega\right)^{1/2} \left(\int_{\mathbb{R}^d} t_j^2 |c|^2 \lambda^d |H_0(\lambda t)|^2 dt\right)^{-1/2} \overset{(2.42)(iii)+(4.17)}{=} \\
&= \left(\int_{\mathbb{R}^d} \omega_j^2 \lambda^{-d} |H_0(\omega/\lambda)|^2 d\omega\right)^{1/2} \left(\int_{\mathbb{R}^d} t_j^2 \lambda^d |H_0(\lambda t)|^2 dt\right)^{-1/2} \underset{s=\lambda t}{\overset{\xi=\omega/\lambda}{=}} \\
&= \lambda^2 \left(\int_{\mathbb{R}^d} \xi_j^2 |H_0(\xi)|^2 d\xi\right)^{1/2} \left(\int_{\mathbb{R}^d} s_j^2 |H(s)|^2 ds\right)^{-1/2} = \lambda^2.
\end{aligned}$$

We notice that this result is independent of j. So, to sum up, equality in (5.6) is achieved for every $j = 1, \ldots, d$ if and only if $f(t) = cH_0(\lambda t)$ for some $c \in \mathbb{C}$ and $\lambda > 0$.

Now that we have (5.6) we can prove (5.5), starting from:

$$\begin{aligned}
&\left(\int_{\mathbb{R}^d} |t|^2 |f(t)|^2 dt\right)^{1/2} \left(\int_{\mathbb{R}^d} |\omega|^2 |\hat{f}(\omega)|^2 d\omega\right)^{1/2} = \\
&= \left(\sum_{j=1}^d \int_{\mathbb{R}^d} t_j^2 |f(t)|^2 dt\right)^{1/2} \left(\sum_{j=1}^d \int_{\mathbb{R}^d} \omega_j^2 |\hat{f}(\omega)|^2 d\omega\right)^{1/2}.
\end{aligned}$$

We notice that the last expression is the product of the Euclidean norm of vectors $(\|t_j f\|_2)_{j=1}^d$ and $(\|\omega_j \hat{f}\|_2)_{j=1}^d$. Thus, from Cauchy-Schwarz inequality in $\mathbb{R}^d$, we obtain:

$$\begin{aligned}
&\left(\int_{\mathbb{R}^d} |t|^2 |f(t)|^2 dt\right)^{1/2} \left(\int_{\mathbb{R}^d} |\omega|^2 |\hat{f}(\omega)|^2 d\omega\right)^{1/2} \geq \\
&\geq \sum_{j=1}^d \left(\int_{\mathbb{R}^d} t_j^2 |f(t)|^2 dt\right)^{1/2} \left(\int_{\mathbb{R}^d} \omega_j^2 |\hat{f}(\omega)|^2 d\omega\right)^{1/2} \overset{(5.6)}{\geq} \frac{d\|f\|_2^2}{4\pi}.
\end{aligned}$$

Finally, equality is achieved if and only if both inequalities in the last expression become equalities. From the first part of the proof we know that equality in the latter inequality is achieved if and only if $f(t) = cH_0(\lambda t)$ for some $c \in \mathbb{C}$ and $\lambda > 0$. Moreover, always from previous computations we saw that in this case f satisfies (5.8) for every $j = 1, \ldots, d$. This means exactly that vectors $(\|t_j f\|_2)_{j=1}^d$ and $(\|\omega_j \hat{f}\|_2)_{j=1}^d$ are parallel, therefore equality is achieved also when using Cauchy-Schwarz's inequality. $\square$

We shall comment a mathematical interpretation of Heisenberg's uncertainty principle. This can be written in the following form:

$$\left(\int_{\mathbb{R}^d} |t-a|^2 \frac{|f(t)|^2}{\|f\|_2^2} dt\right)^{1/2} \left(\int_{\mathbb{R}^d} |\omega-b|^2 \frac{|\hat{f}(\omega)|^2}{\|\hat{f}\|_2^2} d\omega\right)^{1/2} \geq \frac{d}{4\pi},$$

so we may directly assume that f is normalized. In such a case, $|f|^2$ can be seen as a probability distribution. If these integrals are finite for some a and b, through the same argument of time-frequency shift we already used, it is easy that they are always finite. Then, from a formal point of view, we can take the derivative with respect a and b, thus obtaining that their minimum is achieved when

$$a = \bar{t} = \int_{\mathbb{R}^d} t|f(t)|^2 dt, \quad b = \bar{\omega} = \int_{\mathbb{R}^d} \omega|\hat{f}(\omega)|^2 d\omega,$$

which are the mean of $|f|^2$ and $|\hat{f}|^2$, respectively. In this case, previous integrals represent the standard deviation of $|f|^2$ and $|\hat{f}|^2$, which we indicate with $\Delta_x f$ and $\Delta_\omega f$. From an heuristic perspective, it is fair to believe that a function $|f|^2$ is mostly concentrated around its mean and that its standard deviation is a measure of how spread it is. In light of these arguments, Heisenberg's uncertainty principle can written as:

$$\Delta_x f \cdot \Delta_\omega f \geq \frac{d}{4\pi},$$

which is a quantification of the main point of an uncertainty principle, namely that a function and its Fourier cannot be simultaneously concentrated.

5.2 Donoho-Stark's uncertainty principle

As we saw, Heisenberg's uncertainty principle measures the concentration of a function in terms of the variance. However, this is not the only way concentration can be stated. In this section we present an uncertainty principle about the so-called *essential support* of a function which, roughly speaking, is the set where a function has most of its energy.

Definition 5.5. *A function $f \in L^2(\mathbb{R}^d)$ is ε-**concetrated** on a measurable set $T \subseteq \mathbb{R}^d$ for some $\varepsilon \in [0,1]$ if*

$$\left(\int_{T^c} |f(t)|^2 dt\right)^{1/2} \leq \varepsilon \|f\|_2,$$

where $T^c = \mathbb{R}^d \setminus T$ denotes the complement set of T.

If $\varepsilon \leq \frac{1}{2}$, this tells us that most of the energy of f is inside T. Therefore, in such case we may call T the *essential support* of f.

Theorem 5.6 (Donoho-Stark's uncertainty principle). *Let $f \in L^2(\mathbb{R}^d) \setminus \{0\}$, suppose that f is ε_T-concentrated on $T \subseteq \mathbb{R}^d$ while $\hat{f}$ is ε_Ω-concentrated on $\Omega \subseteq \mathbb{R}^d$. Then*

$$|T|\,|\Omega| \geq (1-\varepsilon_T-\varepsilon_\Omega)^2 \tag{5.9}$$

Proof. The result is trivial if T or Ω have infinite measure. Hence we will suppose that they both have finite measure.
Concentration can be stated in an equivalent way through projection operators introduced in Section 4.1, indeed:

$$\left(\int_{T^c} |f(t)|^2 dt\right)^{1/2} = \|f - \chi_T f\|_2 = \|f - P_T f\|_2 \leq \varepsilon_T \|f\|_2,$$
$$\left(\int_{\Omega^c} |\hat{f}(\omega)|^2 d\omega\right)^{1/2} = \|\hat{f} - \chi_\Omega \hat{f}\|_2 = \|f - \mathscr{F}^{-1}(\chi_\Omega \hat{f})\|_2 = \|f - Q_\Omega f\|_2 \leq \varepsilon_\Omega \|f\|_2.$$

In Section 4.1 we also noticed that $\|Q_\Omega\| \leq 1$, hence

$$\|f - Q_\Omega P_T f\|_2 = \|f - Q_\Omega f + Q_\Omega f - Q_\Omega P_T f\|_2 \leq \|f - Q_\Omega f\|_2 + \|Q_\Omega(f - P_T f)\|_2 \leq$$
$$\leq \|f - Q_\Omega f\|_2 + \|f - P_T f\|_2 \leq (\varepsilon_\Omega + \varepsilon_T)\|f\|_2.$$

and consequently

$$\|f\|_2 = \|f - Q_\Omega P_T f + Q_\Omega P_T f\|_2 \leq \|f - Q_\Omega P_T f\|_2 + \|Q_\Omega P_T f\|_2 \implies$$
$$\implies \|Q_\Omega P_T f\|_2 \geq \|f\|_2 - \|f - Q_\Omega P_T f\|_2 \geq (1 - \varepsilon_\Omega - \varepsilon_T)\|f\|_2$$

Thanks to Proposition 4.1 we know that $\|Q_\Omega P_T\|_{\mathrm{HS}} = \sqrt{|T|\,|\Omega|}$ and from Theorem 2.27 we know that $\|Q_\Omega P_T\| \leq \|Q_\Omega P_T\|_{\mathrm{HS}}$, therefore

$$(1 - \varepsilon_\Omega - \varepsilon_T)\|f\|_2 \leq \|Q_\Omega P_T f\|_2 \leq \sqrt{|T|\,|\Omega|}\|f\|_2.$$

□

Taking $\varepsilon = 0$ in (5.9) gives us the following corollary.

Corollary 5.7. *Let $f \in L^2(\mathbb{R}^d) \setminus \{0\}$, $\mathrm{supp} f \subseteq T$, $\mathrm{supp} \hat{f} \subseteq \Omega$. Then $|T||\Omega| \geq 1$.*

Loosely speaking, this result is telling us that f and $\hat{f}$ cannot concentrate too much energy in a small subset of the phase space. However, a much stronger result holds.

Theorem 5.8. *Suppose $f \in L^1(\mathbb{R}^d)$, $\mathrm{supp} f \subseteq T$ and $\mathrm{supp} \hat{f} \subseteq \Omega$. If $|T||\Omega| < +\infty$ then $f = 0$.*

As pointed out in [10], this result forces us not to talk about "support of a function" in the usual way, since in this case the only information we have is that either $f = 0$ or $|\mathrm{supp} f||\mathrm{supp} \hat{f}| = +\infty$. Therefore, a more suitable notion is the one of essential support, since for this we can give a quantitative result, namely 5.6.

5.3 Lieb's uncertainty principle

Up to now we presented two uncertainty principles related to the Fourier transform. However uncertainty principles can be stated for every time of time-frequency analysis. In this and in the following section we present some uncertainty principles for the STFT.

We start considering a weak form and then we will show how to turn Lieb's inequality (3.4) into an uncertainty principle. Like for the Donoho-Stark's uncertainty principle, the notion of concentration is measured in terms of essential support.

Proposition 5.9. *Let $f, \phi \in L^2(\mathbb{R}^d)$ normalized, $U \subseteq \mathbb{R}^{2d}$ and $\varepsilon \in [0,1]$. Suppose that*

$$\int_U |\mathcal{V}_\phi f(x,\omega)|^2 dx d\omega \geq 1 - \varepsilon.$$

Then $|U| \geq 1 - \varepsilon$.

Proof. From (3.5) we see that $|\mathcal{V}_\phi f(x,\omega)| \leq 1$ for all $(x,\omega) \in \mathbb{R}^{2d}$, therefore

$$1 - \varepsilon \leq \int_U |\mathcal{V}_\phi f(x,\omega)|^2 dx d\omega \leq \|\mathcal{V}_\phi f\|_\infty^2 |U| \leq |U|. \tag{5.10}$$

□

Theorem 5.10 (Lieb's uncertainty principle)**.** *Suppose that $\|f\|_2 = \|\phi\|_2 = 1$. If $U \subseteq \mathbb{R}^{2d}$ and $\varepsilon \in [0,1]$ are such that*

$$\int_U |\mathcal{V}_\phi f(x,\omega)|^2 dx d\omega \geq 1 - \varepsilon.$$

Then

$$|U| \geq \sup_{p>2} (1-\varepsilon)^{\frac{p}{p-2}} \left(\frac{p}{2}\right)^{\frac{2d}{p-2}}.$$

Proof. If $|U| = \infty$ the result is trivial hence we can suppose that U has finite measure. It is sufficient to use Hölder's inequality with exponents $p/2$ and $(p/2)' = p/(p-2)$:

$$\begin{aligned}
1 - \varepsilon &\leq \int_U |\mathcal{V}_\phi f(x,\omega)|^2 dx d\omega = \int_{\mathbb{R}^{2d}} |\mathcal{V}_\phi f(x,\omega)|^2 \chi_U(x,\omega) dx d\omega \overset{\text{Hölder}}{\leq} \\
&\leq \left(\int_{\mathbb{R}^{2d}} |\mathcal{V}_\phi f(x,\omega)|^{2\frac{p}{2}} dx d\omega \right)^{\frac{2}{p}} \left(\int_{\mathbb{R}^{2d}} \chi_U(x,\omega)^{\frac{p}{p-2}} dx d\omega \right)^{\frac{p-2}{p}} \overset{(3.7)}{\leq} \\
&\leq \left(\frac{2}{p}\right)^{\frac{2d}{p}} \|f\|_2^2 \, \|g\|_2^2 \, |U|^{\frac{p-2}{p}} = \left(\frac{2}{p}\right)^{\frac{2d}{p}} |U|^{\frac{p-2}{p}}.
\end{aligned}$$

We point out that the use of Hölder's inequality is justified because U has finite measure and, since $\mathcal{V}_\phi f \in L^q(\mathbb{R}^{2d})$ for every $q \geq 2$, $|\mathcal{V}_\phi f|^2 \in L^q(\mathbb{R}^{2d})$ for every $q \geq 1$. Because this result holds for every $p > 2$, we can take the supremum over all possible p, which leads to (5.10). □

5.4 Faber-Krahn Inequality for the STFT

Lieb's uncertainty principle and Lieb's inequality are general results for the STFT because they hold for every possible window $\phi \in L^2(\mathbb{R}^d)$. One may think that for specific choices of the window it is possible to obtain improved results. In this last section we present a recent result, due to Nicola and Tilli and presented in [23], about the STFT with Gaussian window. In this work, they considered the following variational problem:

$$\max_{f \in L^2(\mathbb{R}^d)\setminus\{0\}} \frac{\int_\Omega |\mathcal{V}_\varphi f(x,\omega)|^2 dx d\omega}{\|f\|_2^2} \tag{5.11}$$

where $\Omega \subset \mathbb{R}^{2d}$ is a measurable set with prescribed measure $s > 0$. Therefore, we are asking for the maximal energy of the STFT that can be trapped into a set of a prescribed measure s and, possibly, which functions achieve the maximum. The problem is completely solved and the solution is presented in the following theorem.

Theorem 5.11 (Theorem 4.1 [24]). *For every $f \in L^2(\mathbb{R}^d)$ such that $\|f\|_{L^2} = 1$ and every measurable subset $\Omega \subset \mathbb{R}^{2d}$ with finite measure we have*

$$\int_\Omega |\mathcal{V}_\varphi f(x,\omega)|^2 dx d\omega \le G(|\Omega|), \tag{5.12}$$

where $G(s)$ is given by

$$G(s) := \int_0^s e^{(-d!\tau)^{1/d}} d\tau. \tag{5.13}$$

Moreover, equality occurs if and only if f is a Gaussian of the kind

$$f(x) = ce^{2\pi i x\cdot\omega_0}\varphi(x - x_0) = c\,\pi(x_0,\omega_0)\varphi(x), \quad x \in \mathbb{R}^d \tag{5.14}$$

for some unimodular $c \in \mathbb{C}$ and some $(x_0,\omega_0) \in \mathbb{R}^{2d}$ and Ω is equivalent, in measure, to a ball of centre (x_0,ω_0).

The proof of this theorem is non trivial since it requires some tools from geometric measure theory, such as the coarea formula and the isoperimetric inequality. However, it is worth mentioning that the very first step of the proof is rephrasing the problem in the Fock space introduced in 3.2. Indeed, recalling the relation between the STFT with Gaussian window and the Bargmann transform (3.16) and that the latter is an isometry from $L^2(\mathbb{R}^d)$ into $\mathcal{F}^2(\mathbb{C}^d)$ we have:

$$\frac{\int_\Omega |\mathcal{V}_\varphi f(x,\omega)|^2 dx d\omega}{\|f\|_2^2} = \frac{\int_{\Omega'} |\mathcal{B}f(z)|^2 e^{-\pi|z|^2} dz}{\|\mathcal{B}f\|_{\mathcal{F}^2}^2},$$

where $\Omega' = \{(x,\omega) : (x,-\omega) \in \Omega\}$. Since the Bargmann transform is an unitary operator, variational problem (5.11) can be rephrased in the following way:

$$\max_{F\in\mathcal{F}^2(\mathbb{C}^d)\setminus\{0\}} \frac{\int_\Omega |F(z)|^2 e^{-\pi|z|^2} dz}{\|F\|_{\mathcal{F}^2}^2}.$$

While, at first sight, this might just seem a rewriting of the problem, actually the presence of the Bargmann transform is crucial. It is clear that, for $F \in \mathcal{F}^2(\mathbb{C}^d)$, the quantity $\int_\Omega |F(z)|^2 e^{-\pi|z|^2} dz$ is maximized when Ω is the a super-level set of $|F(z)|^2 e^{-\pi|z|^2}$. Thus, it is natural to study the integral of $|F(z)|^2 e^{-\pi|z|^2}$ over its super-level sets and this is where regularity of functions in the Fock space comes into play.

For the sake of completeness, we mention that the Theorem in [23] is presented in a slightly different way, namely:

$$\int_\Omega |\mathcal{V}f(x,\omega)|^2 dx d\omega \le \frac{\gamma\left(d, \pi(|\Omega|/\omega_{2d})^{1/d}\right)}{(d-1)!},$$

where ω_{2d} is the volume of the unit ball in $\mathbb{R}^{2d}$ and γ is the lower incomplete gamma function. Recalling the definition of γ:

$$\gamma\left(d, \pi(|\Omega|/\omega_{2d})^{1/d}\right) = \int_0^{\pi(|\Omega|/\omega_{2d})^{1/d}} t^{d-1}e^{-t}dt$$

and since $\omega_{2d} = \pi^d/d!$, through the change of variable $t^d = d!\tau$ one obtains (5.12).

Remark. We notice that the numerator of (5.11) can be written also in the following way:

$$\int_\Omega |\mathcal{V}_\varphi f(x,\omega)|^2 dxd\omega = \langle \chi_\Omega \mathcal{V}_\varphi f, \mathcal{V}_\varphi f\rangle = \langle \mathcal{V}_\varphi^* \chi_\Omega \mathcal{V}_\varphi f, f\rangle = \langle L_{\Omega,\varphi} f, f\rangle, \tag{5.15}$$

where $L_{\Omega,\varphi}$ is the localization operator with weight χ_Ω. Therefore, taking the maximum for all possible $f \in L^2(\mathbb{R}^d)\setminus\{0\}$ such that $\|f\|_2 = 1$ leads to a bound for the norm of $L_{\Omega,\varphi}$ (to obtain the norm we have to take the maximum of $\langle L_{\Omega,\varphi} f, g\rangle$ for all possible f and g normalized). However, we know *a posteriori* that the maximum is attained when Ω is a ball and f is a Gaussian, both with the same centre $(x_0, \omega_0) \in \mathbb{R}^{2d}$. We mention that results we obtain in Section 4.2.1 for localization operators with spherically symmetric weights can be obtained also under the action of a time-frequency shift. Indeed, this case is considered in [5] and, as expected, the eigenfunctions of these shifted localization operators are time-frequency shifted Hermite functions. So, if f is a Gaussian it is an eigenfunction of $L_{\Omega,\varphi}$ and, in the end, the maximum of (5.15) is not only a bound for the norm of $L_{\Omega,\varphi}$ but it is the actual norm.

Once Theorem 5.11 has been established, arguing like previous section we immediately obtain an uncertainty principle, which is sharp.

Corollary 5.12. *Let $f \in L^2(\mathbb{R}^d)$ with $\|f\|_2 = 1$, $\Omega \subset \mathbb{R}^{2d}$ measurable, $\varepsilon \in [0,1)$ and suppose that*

$$\int_\Omega |\mathcal{V}_\varphi(x,\omega)|^2 dxd\omega \geq 1 - \varepsilon.$$

Then

$$|\Omega| \geq G^{-1}(1-\varepsilon). \tag{5.16}$$

We point out that G is invertible since it is monotonically increasing. Moreover, its image is $[0,1)$, therefore its inverse $G^{-1} : [0,1) \to [0, +\infty)$ is itself monotonically increasing. This implies that, letting $\varepsilon \to 0$ in (5.16), which means that Ω contains more and more energy, we have $|\Omega| \to +\infty$. If we compare this with Lieb's uncertainty principle we immediately realize how strong this result is, since letting $\varepsilon = 0$ in (5.10) yields to:

$$|\Omega| \geq \sup_{p>2} \left(\frac{p}{2}\right)^{\frac{2d}{p-2}},$$

which is a finite number.

In conclusion, we consider the special case $d = 1$, when G^{-1} can be actually computed. Indeed, in this case we have:

$$G(s) = 1 - e^{-s} \implies G^{-1}(s) = \log\left(\frac{1}{1-G(s)}\right),$$

therefore (5.16) becomes

$$|\Omega| \geq \log\left(\frac{1}{\varepsilon}\right).$$

Even if we did not remark it, we already obtain this bound in the case Ω is a ball. Indeed, in 4.9 we obtained the following expression (4.24):

$$\int_{\mathcal{B}_R} |\mathcal{V}_\phi f(x,\omega)|^2 dx d\omega \leq 1 - e^{-\pi R^2}.$$

If we suppose $\int_{\mathcal{B}_R} |\mathcal{V}_\varphi f(x,\omega)|^2 dx d\omega \geq 1-\varepsilon$ for some $\varepsilon \in [0,1)$, we obtain that:

$$1-\varepsilon \leq 1 - e^{-\pi R^2} \implies \pi R^2 \geq \log\left(\frac{1}{\varepsilon}\right).$$

where πR^2 is exactly the measure of $\mathcal{B}_R$.

Chapter 6

Recent results

Theorem 5.11 is not only remarkable by itself, but it turns out to be a powerful tool in the study of the maximal norm of localization operator when the window of the STFT is a normalized Gaussian. We already mentioned that 5.11 can be rephrased into a result for the norm of localization operators of the kind $L_{\Omega,\varphi}$, where $\Omega \subset \mathbb{R}^{2d}$ is a measurable set of prescribed measure. The latter condition can be seen as a constraint for the $L^1(\mathbb{R}^{2d})$ norm of χ_Ω. In light of this observation, we may think to consider an analogous problem where χ_Ω is replaced by a generic weight function F that satisfies an integrability, and possibly boundedness, condition. This last chapter is devoted to the study of this problem. We start presenting a result from Nicola and Tilli [24] where F is chosen under an L^p and L^∞ constraint. Then, we consider a more generic case where the L^∞ constraint is replaced by a L^q one.

6.1 Norm of localization operators: results from Nicola-Tilli

In this section we show the results in [24]. The aforementioned problem can be precisely stated as follows: find the optimal constant $C > 0$ such that:

$$\|L_{F,\varphi}\|_{L^2(\mathbb{R}^d)\to L^2(\mathbb{R}^d)} \leq C, \tag{6.1}$$

where F satisfies the following constraints:

$$\|F\|_\infty \leq A \quad \text{and} \quad \|F\|_p \leq B. \tag{6.2}$$

Clearly the constant C will depend on p, A and B. In [24] this problem is completely solved: the constant C is computed (explicitly in some cases), weight functions F which achieve this bound are explicitly found and also function f and g such that $|\langle L_{F,\varphi} f, g\rangle| = \|L_{F,\varphi}\| = C$ are found. Before reporting the main Theorem of [24], we define the following number which will appear many times:

$$\kappa_p := \frac{p-1}{p}. \tag{6.3}$$

Moreover, for the sake of brevity, we denote the variable $(x, \omega) \in \mathbb{R}^{2d}$ as z and therefore $dxd\omega$ as dz.

Theorem 6.1. *Assume $p \in [1, \infty)$, $A \in (0, \infty]$ and $B \in (0, \infty)$ with the additional condition that $A < \infty$ when $p = 1$. Let F satisfy the constraints in (6.2).*

(i) If $p = 1$, then

$$\|L_{F,\varphi}\| \leq A\, G(B/A), \tag{6.4}$$

and equality occurs if and only if, for some $\theta \in \mathbb{R}$ and some $z_0 \in \mathbb{R}^{2d}$

$$F(z) = Ae^{i\theta}\chi_{\mathcal{B}}(z - z_0) \quad \forall z \in \mathbb{R}^{2d} \tag{6.5}$$

where $\mathcal{B} \subset \mathbb{R}^{2d}$ is the ball of measure B/A centred at the origin.

(ii) If $p > 1$ and $B/A \leq \kappa_p^{d/p}$, then

$$\|L_{F,\varphi}\| \leq \kappa_p^{d\kappa_p} B, \tag{6.6}$$

with equality if and only if, for some $\theta \in \mathbb{R}$ and some $z_0 \in \mathbb{R}^{2d}$,

$$F(z) = e^{i\theta}\lambda e^{\frac{\pi}{p-1}|z-z_0|^2} \quad \forall z \in \mathbb{R}^{2d}, \tag{6.7}$$

where $\lambda = \kappa_p^{-d/p} B$.

(iii) If $p > 1$ and $B/A > \kappa_p^{d/p}$, then

$$\|L_{F,\varphi}\| \leq \int_0^A G(u_\lambda(t))dt, \tag{6.8}$$

where $u_\lambda(t) = \left[-\log\left((t/\lambda)^{p-1}\right)\right]^d$ and $\lambda > A$ is uniquely determined by the condition $p\int_0^A t^{p-1}u_\lambda(t)dt = B^p$. Equality in (6.8) is achieved if and only if, for some $\theta \in \mathbb{R}$ and some $z_0 \in \mathbb{R}^{2d}$,

$$F(z) = e^{i\theta}\min\{\lambda e^{-\frac{\pi}{p-1}|z-z_0|^2}, A\} \tag{6.9}$$

Finally, in all the cases, condition $|\langle L_{F,\varphi}f, g\rangle| = \|L_{F,\varphi}\|$ holds for some, $f, g \in L^2(\mathbb{R}^d)$ such that $\|f\|_2 = \|g\|_2 = 1$, if and only if both f and g are of the kind (5.14), possibly with different c's, but with the same $(x_0, \omega_0) \in \mathbb{R}^{2d}$ which coincides with the centre of F.

We will not give the proof of these results since some of its parts are similar to the one we will see in the following section. Moreover, we point out that the case $A = \infty$ means we are dropping the L^∞ constraint.

6.2 Generic case

In this section we will deal with a generalized version of the problem considered in [24]. Indeed, we want to find the optimal constant C such that

$$\|L_{F,\varphi}\|_{L^2(\mathbb{R}^d)\to L^2(\mathbb{R}^d)} \leq C$$

under the following constraints on F:

$$\|F\|_p \leq A \quad \text{and} \quad \|F\|_q \leq B, \tag{6.10}$$

where $p, q \in (1,\infty)$ and $A, B \in (0,\infty)$. In this setting it is no more possible to find an explicit expression for C and F, although they can be computed numerically.

Theorem 6.1(ii) includes the case when F satisfies just an L^p constraint by taking A (L^∞ constraint) equal to ∞. In the current setting we have an L^p and an L^q bound, hence, thanks to 6.6, it is straightforward to see that

$$\|L_{F,\varphi}\|_{L_2\to L_2} \leq \min\{\kappa_p^{d\kappa_p}A,\, \kappa_q^{d\kappa_q}B\}.$$

Suppose that the first term is smaller than the second, which means:

$$\kappa_p^{d\kappa_p}A \leq \kappa_q^{d\kappa_q}B \iff \frac{B}{A} \geq \left(\frac{\kappa_p^{\kappa_p}}{\kappa_q^{\kappa_q}}\right)^d. \tag{6.11}$$

Clearly, for B sufficiently large we expect that the solution of current problem is the same as the one with just an L^p constraint, namely the one given by (6.7). Therefore, we want to compare its L^q norm with the bound given by B:

$$\begin{aligned}\|F\|_q^q &= \int_{\mathbb{R}^{2d}} |F(z)|^q dz = \lambda^q \int_{\mathbb{R}^{2d}} e^{-\frac{q\pi}{p-1}|z-z_0|^2} dz \overset{z'=\left(\frac{q\pi}{p-1}\right)^{1/2}(z-z_0)}{=} \\ &= \lambda^q \left(\frac{p-1}{q\pi}\right)^d \int_{\mathbb{R}^{2d}} e^{-|z'|^2} dz' = \lambda^q \left(\frac{p-1}{q\pi}\right)^d \pi^d = \lambda^q \left(\frac{p-1}{q}\right)^d.\end{aligned}$$

Since we want F to satisfy the L^q constraint we should have

$$\lambda \left(\frac{p-1}{q}\right)^{d/q} \leq B \overset{\lambda=\kappa_p^{-d/p}A}{\Longrightarrow} \left(\frac{p}{p-1}\right)^{d/p} \left(\frac{p-1}{q}\right)^{d/q} A \leq B,$$

which is equivalent to

$$\frac{B}{A} \geq \kappa_p^{d\left(\frac{1}{q}-\frac{1}{p}\right)} \left(\frac{p}{q}\right)^{\frac{d}{q}}.$$

If this condition were less restrictive than the one given by (6.11) we would have solved the problem. Unfortunately, this is not the case. Indeed it is always true, regardless of p and q, that

$$\kappa_p^{d\left(\frac{1}{q}-\frac{1}{p}\right)} \left(\frac{p}{q}\right)^{\frac{d}{q}} \geq \left(\frac{\kappa_p^{\kappa_p}}{\kappa_q^{\kappa_q}}\right)^d. \tag{6.12}$$

The proof of this inequality can be found in B.2.

Despite this fact, at least we can say that, if $B/A \geq \kappa_p^{d\left(\frac{1}{q}-\frac{1}{p}\right)}\left(\frac{p}{q}\right)^{\frac{d}{q}}$ or $B/A \leq \kappa_q^{d\left(\frac{1}{q}-\frac{1}{p}\right)}\left(\frac{p}{q}\right)^{\frac{d}{p}}$, the problem is already solved and the solution is given by Theorem 6.1. Therefore, from now on, we will suppose to be in the intermediate case, that is:

$$\kappa_q^{d\left(\frac{1}{q}-\frac{1}{p}\right)}\left(\frac{p}{q}\right)^{\frac{d}{p}} < \frac{B}{A} < \kappa_p^{d\left(\frac{1}{q}-\frac{1}{p}\right)}\left(\frac{p}{q}\right)^{\frac{d}{q}}. \tag{6.13}$$

We notice that the condition is well-posed, since it is actually true that

$$\kappa_q^{d\left(\frac{1}{q}-\frac{1}{p}\right)}\left(\frac{p}{q}\right)^{\frac{d}{p}} < \kappa_p^{d\left(\frac{1}{q}-\frac{1}{p}\right)}\left(\frac{p}{q}\right)^{\frac{d}{q}} \tag{6.14}$$

whenever $p \neq q$ (proof is in B.3).

Before tackling the problem in this intermediate regime, we prove a Theorem from [24] which gives a bound for $\|L_{F,\varphi}\|$ in terms of the distribution function of $|F|$.

Theorem 6.2. *Assume $F \in L^p(\mathbb{R}^{2d})$ for some $p \in [1, +\infty)$ and let $\mu(t) = |\{|F| > t\}|$ be the distribution function of $|F|$. Then*

$$\|L_{F,\varphi}\| \leq \int_0^{\infty} G(\mu(t))dt. \tag{6.15}$$

Equality occurs if and only if $F(z) = e^{i\theta}\rho(|z - z_0|)$ for some $\theta \in \mathbb{R}$, $z_0 \in \mathbb{R}^{2d}$ and some nonincreasing function $\rho : [0, +\infty) \to [0, +\infty)$

Proof. Let $f, g \in L^2(\mathbb{R}^d)$ such that $\|f\|_2 = \|g\|_2 = 1$. Since we are in a Hilbert space $\|L_{F,\varphi}\|$ can be computed as the supremum of $|\langle L_{F,\varphi}f, g\rangle|$ over all normalized f and g. Therefore we are interested in estimating the previous scalar product:

$$\begin{aligned} |\langle L_{F,\varphi}f, g\rangle| = |\mathscr{L}_{F,\varphi}(f,g)| &\leq \int_{\mathbb{R}^{2d}} |F(z)| \cdot |\mathcal{V}_\varphi f(z)| \cdot |\mathcal{V}_\varphi g(z)| dz \overset{\text{C-S}}{\leq} \\ &\leq \left(\int_{\mathbb{R}^{2d}} |F(z)| \cdot |\mathcal{V}_\varphi f(z)|^2 dz\right)^{1/2} \left(\int_{\mathbb{R}^{2d}} |F(z)| \cdot |\mathcal{V}_\varphi g(z)|^2 dz\right)^{1/2}. \end{aligned} \tag{6.16}$$

Since the result is symmetric in f and g we can study just one of the terms. Letting $m = \operatorname{ess\,sup} |F(z)|$ and assuming $m > 0$ (otherwise every result is trivial) we can use the "layer cake" representation [20, Theorem 1.13]

$$|F(z)| = \int_0^m \chi_{\{|F|>t\}}(z)dt$$

in order to find

$$\begin{aligned} \int_{\mathbb{R}^{2d}} |F(z)| \cdot |\mathcal{V}_\varphi f(z)|^2 dz &= \int_{\mathbb{R}^{2d}} \left(\int_0^m \chi_{\{|F|>t\}}(z)dt\right) |\mathcal{V}_\varphi f(z)|^2 dz \overset{\text{Tonelli}}{=} \\ &= \int_0^m \left(\int_{\mathbb{R}^{2d}} \chi_{\{|F|>t\}}(z)|\mathcal{V}_\varphi f(z)|^2 dz\right) dt = \\ &= \int_0^m \left(\int_{\{|F|>t\}} |\mathcal{V}_\varphi f(z)|^2 dz\right) dt. \end{aligned}$$

We notice that the quantity in the inner integral is exactly the one in Theorem 5.11, hence

$$\int_{\mathbb{R}^{2d}} |F(z)| \cdot |\mathcal{V}_\varphi f(z)|^2 dz \leq \int_0^m G\left(|\{|F| > t\}|\right) dt = \int_0^m G(\mu(t)) dt \tag{6.17}$$

We point out that since $\mu(t) = 0$ for $t > m$ and that $G(0) = 0$, the previous expression is equivalent to (6.15).

Because $p < \infty$, from Proposition 4.6 we know that L_F is a compact operator, therefore there exist normalized f and g which achieve equality in the supremum of the norm, namely $\langle L_{F,\varphi} f, g\rangle = \|L_{F,\varphi}\|$. Therefore, equality in (6.15) occurs if and only if all the previous inequalities become equalities. Equality in (6.17) occurs if and only if

$$\int_{\{|F|>t\}} |\mathcal{V}_\varphi f(z)|^2 dz = G(\mu(t)) \tag{6.18}$$

for a.e. $t \in (0, m)$. Now, fix $t_0 \in (0, m)$ such that equality holds. From Theorem 5.11 we can infer that $\{|F| > t_0\}$ is (equivalent to) a ball centred in $z_0 = (x_0, \omega_0)$ and that f is a Gaussian of the kind (5.14) with the same centre z_0. Now that the centre of f is fixed, still from Theorem 5.11, we obtain that equality in (6.18) a.e implies that also the other levels sets $\{|F| > t\}$ are equivalent to balls centred at the same z_0. Finally, we can extend the result to every $t \in (0, m)$ because $\{|F| > t\} = \bigcup_{s>t}\{|F| > s\}$. Since Theorem 5.11 is a “if and only if”, these conditions on F and f are also sufficient to guarantee equality in (6.17). Clearly, the same result holds for g which has to be a Gaussian, possibly with different coefficient c but the same centre z_0.

In the end, since all the super-level sets of $|F|$ are balls we conclude that it is spherically symmetric and radially decreasing, as claimed in the theorem's statement.

Conditions for f and g imply that $\mathcal{V}_\varphi g = e^{i\alpha}\mathcal{V}_\varphi f$ for some $\alpha \in \mathbb{R}$. This provides equality in (6.16) when using Cauchy-Schwarz inequality. Lastly we shall prove that also the first inequality in (6.16), that is

$$\left|\int_{\mathbb{R}^{2d}} F(z)\mathcal{V}_\varphi f(z)\overline{\mathcal{V}_\varphi g(z)} dz\right| \leq \int_{\mathbb{R}^{2d}} |F(z)| \cdot |\mathcal{V}_\varphi f(z)| \cdot |\mathcal{V}_\varphi g(z)| dz$$

becomes an equality. With the additional information that $\mathcal{V}_\varphi g = e^{i\alpha}\mathcal{V}_\varphi f$ this is equivalent to prove that:

$$\left|\int_{\mathbb{R}^{2d}} F(z)|\mathcal{V}_\varphi f(z)|^2 dz\right| = \int_{\mathbb{R}^{2d}} |F(z)| \cdot |\mathcal{V}_\varphi f(z)|^2 dz. \tag{6.19}$$

The integral on the left-hand side is just a complex number, therefore we have:

$$\int_{\mathbb{R}^{2d}} F(z)|\mathcal{V}_\varphi f(z)|^2 dz = e^{i\theta}\left|\int_{\mathbb{R}^{2d}} F(z)|\mathcal{V}_\varphi f(z)|^2 dz\right|,$$

so (6.19) becomes

$$\int_{\mathbb{R}^{2d}} e^{-i\theta} F(z)|\mathcal{V}_\varphi f(z)|^2 dz = \int_{\mathbb{R}^{2d}} |F(z)| \cdot |\mathcal{V}_\varphi f(z)|^2 dz.$$

The proof would be complete if $|\mathcal{V}_\varphi f(z)|^2$ was always positive, since this would imply $F(z) = e^{i\theta}|F(z)|$. In the proof of Theorem 4.7 we computed the STFT of Hermite functions. Since $\varphi = H_0$, letting $k = 0 \in \mathbb{N}_0^d$ into (4.22) leads to:

$$\mathcal{V}_\varphi \varphi(x,\omega) = e^{-\pi i \omega\cdot x} e^{-\pi(|x|^2+|\omega|^2)/2}.$$

In order to compute the STFT of $f = c\pi(x_0,\omega_0)\varphi$, all we have to do is to understand how $\mathcal{V}_\varphi$ interacts with a time-frequency shift:

$$\mathcal{V}_\varphi f(x,\omega) = \langle c\pi(x_0,\omega_0)\varphi, \pi(x,\omega)\varphi\rangle = c\langle \varphi, T_{-x_0}M_{\omega-\omega_0}T_x\varphi\rangle.$$

Then:

$$T_{-x_0}M_{\omega-\omega_0}T_x\varphi(t) = e^{2\pi i(\omega-\omega_0)\cdot(t+x_0)} f(t+x_0-x) = e^{2\pi i(\omega-\omega_0)\cdot x_0} M_{\omega-\omega_0}T_{x-x_0}\varphi(t),$$

so that we obtain

$$\mathcal{V}_\varphi f(x,\omega) = ce^{2\pi i(\omega-\omega_0)\cdot x_0}\langle \varphi, M_{\omega-\omega_0}T_{x-x_0}\varphi\rangle = ce^{2\pi i(\omega-\omega_0)\cdot x_0}\mathcal{V}_\varphi\varphi(x-x_0,\omega-\omega_0).$$

In the end, taking the modulus of both side we conclude that $|\mathcal{V}_\varphi f(x,\omega)|^2$ is always strictly positive, which concludes the proof. □

In light of the previous Theorem, it is natural to seek for a sharp upper bound for the right-hand side of (6.15). Since this involves the distribution function $|F|$, we shall search this bound between all the possible distribution functions. In order to do so, we need to rephrase constraints (6.10) in terms of μ. This can be easily done thanks to a more general version of the "layer cake" representation (see [20, Theorem 1.13] or [9, Proposition 1.1.4]):

$$\|F\|_p^p = p\int_0^\infty t^{p-1}|\{|F|>t\}|dt.$$

Hence, constraints (6.10) become

$$p\int_0^\infty t^{p-1}u(t)dt \le A^p \quad \text{and} \quad q\int_0^\infty t^{q-1}u(t)dt \le B^q \tag{6.20}$$

and we can define the proper space of possible distribution functions

$$\mathcal{C} = \{u : (0,+\infty) \to [0,+\infty) \text{ such that } u \text{ is decreasing and satisfies (6.20)}\}. \tag{6.21}$$

We have reached the point where our original question is rephrased in the following variational problem:

$$\sup_{v\in\mathcal{C}} I(v) \quad \text{where} \quad I(v) := \int_0^{+\infty} G(v(t))dt \tag{6.22}$$

Firstly, we shall prove existence of maximizers.

Proposition 6.3. *The supremum in (6.22) is finite and it is attained by at least one function $u \in \mathcal{C}$. Moreover, every extremal function u achieves equality in at least one of the constraints (6.20).*

Proof. Considering, for example, the first constraint in (6.20), we see that

$$t^p u(t) = p \int_0^t \tau^{p-1} u(t) d\tau \overset{u \text{ decreasing}}{\leq} p \int_0^t \tau^{p-1} u(\tau) d\tau \leq A^p,$$

hence functions in $\mathcal{C}$ are pointwise bounded by A^p/t^p. It is straightforward to verify that G in (5.13) is increasing, that $G(s) \leq s$ and that $G(s) \leq 1$. Using these properties we have:

$$I(u) = \int_0^{+\infty} G(u(t))dt = \int_0^1 G(u(t))dt + \int_1^{+\infty} G(u(t))dt \overset{G(s)\leq 1}{\leq} 1 + \int_1^{+\infty} G(u(t))dt \overset{G \text{ increasing}}{\leq}$$
$$\leq 1 + \int_1^{+\infty} G(A^p/t^p)dt \overset{G(s)\leq s}{\leq} 1 + \int_1^{+\infty} \frac{A^p}{t^p} dt < \infty,$$

therefore the supremum in (6.22) is finite.

Let $\{u_n\}_{n\in\mathbb{N}} \subset \mathcal{C}$ be a maximizing sequence. Since every u_n is pointwise bounded by A^p/t^p, thanks to Helly's selection theorem C.1 we can say that, up to a subsequence, u_n converges pointwise to a decreasing function u. Moreover, u is still in $\mathcal{C}$, indeed:

$$\int_0^{+\infty} t^{p-1} u(t) = \int_0^{+\infty} \lim_{n\to\infty} t^{p-1} u_n(t) dt \overset{\text{Fatou's lemma}}{\leq} \liminf_{n\to\infty} \int_0^{+\infty} t^{p-1} u_n(t) dt \leq \frac{A^p}{p},$$

and clearly the same holds for q instead of p.

Now we have to prove that u is actually achieving the supremum. We already saw that the following holds:

$$|G(u_n(t))| \leq \chi_{(0,1)}(t) + \frac{A^p}{t^p} \chi_{(1,+\infty)}(t)$$

and that the left-hand side is a function in $L^1(0, +\infty)$. This allows us to use dominated convergence theorem to conclude that

$$I(u) = \int_0^{+\infty} G(u(t)) = \lim_{n\to\infty} \int_0^{\infty} G(u_n(t))dt = \lim_{n\to\infty} I(u_n) = \sup_{v\in\mathcal{C}} I(v).$$

Lastly, we need to show that u achieves equality at least in one of the constraints (6.20). Suppose that this is not true. If we let $u_\varepsilon(t) = (1+\varepsilon)u(t)$, then for $\varepsilon > 0$ sufficiently small constraints are still satisfied and since G is strictly increasing $I(u_\varepsilon) > I(u)$, which contradicts the hypothesis that u is a maximizer. □

In order to do some "meaningful" calculus of variations over I we need to enlarge $\mathcal{C}$, because the monotonicity assumption is quite strict. We will show that removing this hypothesis leaves the supremum unchanged and that maximizers are indeed monotonic.

Proposition 6.4. *Let $\mathcal{C}' = \{u : (0, +\infty) \to [0, +\infty)$ such that u is measurable and satisfies (6.20)$\}$. Then*

$$\sup_{v\in\mathcal{C}} I(v) = \sup_{v\in\mathcal{C}'} I(v). \tag{6.23}$$

In particular, any function $u \in \mathcal{C}$ achieving the supremum on the left-hand side also achieves it on the right-hand side.

Proof. Let $u \in \mathcal{C}'$. We define its *decreasing rearrangement* as:

$$u^*(s) = \sup\{t \geq 0 : |\{u > t\}| > s\}, \tag{6.24}$$

with the convention that $\sup \emptyset = 0$. It is clear from the definition that u^* is a non-increasing function. Moreover, one can see ([11, Section 10.12], [9, Proposition 1.4.5]) that u^* is right-continuous and that u and u^* are *equi-measurable*, which means that they have the same distribution function. Moreover, we already pointed out that constraints (6.20) imply that u is pointwise bounded by A^p/t^p, therefore u^* takes only finite values. Our aim is to show that $u^* \in \mathcal{C}$. Letting ν be the Radon measure with density t^{p-1}, we start proving that $\nu(\{u > s\}) \geq \nu(\{u^* > s\})$, indeed:

$$\begin{aligned}\nu(\{u > s\}) = \int_{\{u>s\}} t^{p-1} dt &\overset{t^{p-1}\text{ increasing}}{\geq} \int_0^{|\{u>s\}|} t^{p-1} dt \overset{\text{equi-measurability}}{=} \\ &= \int_0^{|\{u^*>s\}|} t^{p-1} dt \underset{\text{right-continuous}}{\overset{u^*\text{ decreasing}}{=}} \int_{\{u^*>s\}} t^{p-1} dt = \nu(\{u^* > s\}).\end{aligned}$$

Then, using one more time the "layer cake" representation:

$$\begin{aligned}\int_0^{+\infty} t^{p-1} u(t) dt = \int_0^{+\infty} u(t) d\nu(t) &= \int_0^{+\infty} \nu(\{u > s\}) ds \geq \\ = \int_0^{+\infty} \nu(\{u^* > s\}) ds &= \int_0^{+\infty} u^*(t) d\nu(t) = \int_0^{+\infty} t^{p-1} u^*(t) dt.\end{aligned}$$

If we swap p with q we conclude that $u^* \in \mathcal{C}$. Moreover, always from equi-measurability, we have:

$$\begin{aligned}I(u) = \int_0^{+\infty} G(u(t)) dt = \int_0^{+\infty} \int_0^{u(t)} e^{-(d!\tau)^{1/d}} d\tau dt = \int_0^{+\infty} \int_0^{+\infty} \chi_{\{u>\tau\}}(t) e^{-(d!\tau)^{1/d}} d\tau dt \overset{\text{Tonelli}}{=} \\ = \int_0^{+\infty} |\{u > \tau\}| e^{-(d!\tau)^{1/d}} d\tau = \int_0^{+\infty} |\{u^* > \tau\}| e^{-(d!\tau)^{1/d}} d\tau = I(u^*).\end{aligned}$$

Taking the supremum over all possible $u \in \mathcal{C}'$ we have:

$$\sup_{v \in \mathcal{C}'} I(v) = \sup_{v \in \mathcal{C}'} I(v^*) \leq \sup_{v \in \mathcal{C}} I(v).$$

Inequality $\sup_{v \in \mathcal{C}'} I(v) \geq \sup_{v \in \mathcal{C}} I(v)$ is trivial since $\mathcal{C}' \supset \mathcal{C}$.

□

Before stating and proving the theorem that gives the only maximal function of (6.22), we introduce the following notation:

$$\mathrm{Log}_-(x) = \max\{-\log(x),\ 0\}, \quad x > 0.$$

Theorem 6.5. *There exist a unique function $u \in \mathcal{C}$ achieving the supremum in (6.22) that is:*

$$u(t) = \frac{1}{d!} \left[\mathrm{Log}_-\left(\lambda_1 t^{p-1} + \lambda_2 t^{q-1}\right)\right]^d, \quad t > 0 \tag{6.25}$$

where λ_1, λ_2 are both positive and uniquely determined by

$$p\int_0^{+\infty} t^{p-1}u(t)dt = A^p, \quad q\int_0^{+\infty} t^{q-1}u(t)dt = B^q.$$

Proof. We will split the proof in several parts. Firstly we will show that maximizers are given by (6.25). Then we will show that multipliers λ_1 and λ_2 are both strictly positive and unique.

• **Expression of maximizers**

Let $M = \sup\{t \in (0, +\infty) \ : u(t) > 0\}$. From Proposition 6.3 we know that u has to achieve at least one of the constraints, therefore $M > 0$. Consider now a closed interval $[a, b] \subset (0, M)$ and a function $\eta \in L^\infty(0, M)$ supported in $[a, b]$. Without loss of generality we can suppose that η is orthogonal, in the L^2 sense, to t^{p-1} and t^{q-1}, explicitly

$$\int_a^b t^{p-1}\eta(t)dt = 0, \quad \int_a^b t^{q-1}\eta(t)dt = 0. \tag{6.26}$$

On $[a, b]$ we have that $u(t) \geq u(b) > 0$, hence, for $|\varepsilon|$ sufficiently small, $u + \varepsilon\eta$ is still a nonnegative function which satisfies (6.20), therefore $u + \varepsilon\eta \in \mathcal{C}'$. Since we are supposing that u is a maximizer, the function $\varepsilon \mapsto I(u + \varepsilon\eta)$ has a maximum for $\varepsilon = 0$. Since η is supported in a compact interval we can differentiate under the integral sign and obtain

$$0 = \frac{d}{d\varepsilon} I(u + \varepsilon\eta)|_{\varepsilon=0} = \int_a^b G'(u(t))\eta(t)dt.$$

We would like to extend this result to every η in $L^2(a, b)$ satisfying (6.26). Since $L^\infty(a, b)$ is dense in $L^2(a, b)$, there exist a sequence $\{\eta_k\}_{k\in\mathbb{N}} \subset L^\infty(a, b)$ such that $\eta_k \to \eta$ in $L^2(a, b)$. We can consider the projection operator P such that, given $\psi \in L^2(a, b)$, $P\psi$ is the orthogonal projection of ψ onto $X = \mathrm{span}\{t^{p-1}, t^{q-1}\}^\perp \subset L^2(a, b)$. Since P is continuous we have that $P\eta_k \to P\eta = \eta$, hence

$$0 = \int_a^b G'(u(t))P\eta_k(t)dt = \langle G'(u), P\eta_n\rangle_{L^2(a,b)} \to \langle G'(u), \eta\rangle_{L^2(a,b)} = \int_a^b G'(u(t))\eta(t)dt \implies$$
$$\implies \int_a^b G'(u(t))\eta(t)dt = 0. \tag{6.27}$$

Since (6.27) holds for every $\eta \in X$ it must be that

$$G'(u) \in X^\perp = \left(\mathrm{span}\{t^{p-1}, t^{q-1}\}^\perp\right)^\perp = \mathrm{span}\{t^{p-1}, t^{q-1}\} \quad \text{in } (a, b).$$

Letting $a \to 0^+$ and $b \to M^-$ we then obtain

$$G'(u(t)) = \lambda_1 t^{p-1} + \lambda_2 t^{q-1} \quad \text{for a.e. } t \in (0, M) \tag{6.28}$$

for some multipliers $\lambda_1, \lambda_2 \in \mathbb{R}$. Since u is decreasing actually (6.28) holds for every $t \in (0, M)$. Finally, recalling the expression of (5.13) we see that $G'(s) = e^{-(d!s)^{1/d}}$. Since

u is monotonically decreasing we can invert (6.28) thus obtaining the explicit expression of maximizers:

$$u(t) = \begin{cases} \frac{1}{d!}\left[-\log\left(\lambda_1 t^{p-1} + \lambda_2 t^{q-1}\right)\right]^d & t \in (0, M) \\ 0 & t \in (M, +\infty) \end{cases} \tag{6.29}$$

We remark that a priori it was possible that $M = +\infty$, but from the explicit expression of maximizers we see that this is not possible since u has to be nonnegative.

- **Maximizers achieve equality in both constraints and multipliers are non-zero**

The argument we used to determine the expression of maximizers enables us to say that these have to achieve equality in both constraints in (6.20). Indeed, if, for example, we had that $q\int_0^\infty t^{q-1}u(t)dt < B^q$, the second condition of orthogonality in (6.26) could be removed, because for sufficiently small ε a variation non-orthogonal to t^{q-1} would be admissible. This would provide us the solution of the same variational problem but without the L^q constraint. Since we already know that actually this solution does not satisfy the L^q constraint, we conclude that u has to achieve equality in both constraints. With the very same reasoning we can say that neither λ_1 nor λ_2 can be 0.

- **Multipliers are positive**

Suppose that one of the multipliers, for example λ_2, is negative. Consider an interval $[a, b] \subset (0, M)$ and a variation $\eta \in L^\infty(0, M)$ supported in $[a, b]$. Thanks to the Gram-Schmidt process we can construct a variation orthogonal to t^{p-1}. Since η is arbitrary we can also suppose that it is not orthogonal to t^{q-1}, in particular we can ask that $\int_a^b t^{q-1}\eta(t)dt < 0$. Therefore, the directional derivative of G along η is:

$$\int_a^b G'(u(t))\eta(t)dt = \int_a^b (\lambda_1 t^{p-1} + \lambda_2 t^{q-1})\eta(t)dt = \lambda_2 \int_a^b t^{q-1}\eta(t)dt > 0,$$

which contradicts the fact that u is a maximizer.

- **u is continuous**

Now that we now that both multipliers are positive we can prove that u is continuous, which is equivalent to say that $M = T$, where T is the unique positive number such that $\lambda_1 T^{p-1} + \lambda_2 T^{q-1} = 1$ (uniqueness of T follows from the positivity of multipliers).

We start supposing that $M < T$, which means that $\lim_{t\to M^-} u(t) > 0$. Consider the following variation

$$\eta(t) = \begin{cases} -1 + \alpha\dfrac{t}{M} + \beta, & t \in (M - M\delta, M) \\ 1, & t \in (M, M + M\delta) \\ 0, & \text{otherwise} \end{cases}$$

where $\delta > 0$ is small enough so that $M - M\delta > 0$ and $M + M\delta < T$, while α and β are constants, depending on δ, to be determined. Since we want this to be an admissible variation, we need to impose that η is orthogonal to t^{p-1} and t^{q-1}. For example, the first condition is:

$$0 = \int_{M-M\delta}^{M+M\delta} t^{p-1}\eta(t)dt = -\int_{M-M\delta}^{M} t^{p-1}dt + \int_{M-M\delta}^{M} t^{p-1}\left(\alpha\frac{t}{M}+\beta\right)dt + \int_{M}^{M+M\delta} t^{p-1}dt \overset{\tau=t/M}{=}$$
$$= M^p\int_{1-\delta}^{1}\tau^{p-1}(\alpha\tau+\beta)d\tau - M^p\int_{1-\delta}^{1}\tau^{p-1}d\tau + M^p\int_{1}^{1+\delta}\tau^{p-1}d\tau \overset{1/\delta}{\Longrightarrow}$$
$$\Longrightarrow \fint_{1-\delta}^{1}\tau^{p-1}(\alpha\tau+\beta)d\tau = \alpha\fint_{1-\delta}^{1}\tau^{p}d\tau + \beta\fint_{1-\delta}^{1}\tau^{p-1}d\tau = \fint_{1-\delta}^{1}\tau^{p-1}d\tau - \fint_{1}^{1+\delta}\tau^{p-1}d\tau.$$

The equation stemming from the orthogonality with t^{q-1} is analogous. Therefore, we obtained a nonhomogeneous linear system for α and β:

$$\begin{pmatrix} \fint_{1-\delta}^{1}\tau^{p}d\tau & \fint_{1-\delta}^{1}\tau^{p-1}d\tau \\ \fint_{1-\delta}^{1}\tau^{q}d\tau & \fint_{1-\delta}^{1}\tau^{q-1}d\tau \end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \begin{pmatrix} \fint_{1-\delta}^{1}\tau^{p-1}d\tau - \fint_{1}^{1+\delta}\tau^{p-1}d\tau \\ \fint_{1-\delta}^{1}\tau^{q-1}d\tau - \fint_{1}^{1+\delta}\tau^{q-1}d\tau \end{pmatrix}. \tag{6.30}$$

This system has a unique solution if and only if the determinant of the matrix is not 0. We can show this directly:

$$\fint_{1-\delta}^{1}\tau^{p}d\tau\fint_{1-\delta}^{1}\tau^{q-1}d\tau - \fint_{1-\delta}^{1}\tau^{q}d\tau\fint_{1-\delta}^{1}\tau^{p-1}d\tau =$$
$$= \frac{1}{\delta^2}\int_{(1-\delta,1)^2}\left(\tau^p\sigma^{q-1} - \tau^{p-1}\sigma^{q}\right)d\tau d\sigma = \frac{1}{\delta^2}\int_{(1-\delta,1)^2}\tau^{p-1}\sigma^{q-1}\left(\tau-\sigma\right)d\tau d\sigma =$$
$$= \frac{1}{\delta^2}\left(\int_{Q_1}\tau^{p-1}\sigma^{q-1}\left(\tau-\sigma\right)d\tau d\sigma + \int_{Q_2}\tau^{p-1}\sigma^{q-1}\left(\tau-\sigma\right)d\tau d\sigma\right),$$

where $Q_1 = (1-\delta,1)^2 \cap \{\tau > \sigma\}$ and $Q_2 = (1-\delta,1)^2 \cap \{\tau < \sigma\}$. In the second integral we can consider the change of variable that swaps τ and σ. In this case, the new domain is Q_1, hence:

$$\fint_{1-\delta}^{1}\tau^{p}d\tau\fint_{1-\delta}^{1}\tau^{q-1}d\tau - \fint_{1-\delta}^{1}\tau^{q}d\tau\fint_{1-\delta}^{1}\tau^{p-1}d\tau =$$
$$= \frac{1}{\delta^2}\int_{Q_1}\left(\tau^{p-1}\sigma^{q-1} - \tau^{q-1}\sigma^{p-1}\right)\left(\tau-\sigma\right)d\tau d\sigma.$$

In Q_1 we have that $\tau - \sigma > 0$ and the sign of $\tau^{p-1}\sigma^{q-1} - \tau^{q-1}\sigma^{p-1}$ is constant, indeed:

$$\tau^{p-1}\sigma^{q-1} - \tau^{q-1}\sigma^{p-1} > 0 \iff \left(\frac{\tau}{\sigma}\right)^{p-q} > 1 \overset{\tau>\sigma}{\iff} p > q.$$

Therefore the determinant of the matrix is always not 0.

Now that we have an admissible variation, we can compute the directional derivative of G along η. Since u is supposed to be a maximizer, this derivative has to be nonpositive, therefore:

$$
\begin{aligned}
0 \geq \int_{M-M\delta}^{M+M\delta} G'(u(t))\eta(t)dt &= -\int_{M-M\delta}^{M} \left(\lambda_1 t^{p-1} + \lambda_2 t^{q-1}\right) dt + \\
&+ \int_{M-M\delta}^{M} \left(\lambda_1 t^{p-1} + \lambda_2 t^{q-1}\right) \left(\alpha \frac{t}{M} + \beta\right) dt + \int_{M}^{M+M\delta} dt = \\
&= -\int_{M-M\delta}^{M} \left(\lambda_1 t^{p-1} + \lambda_2 t^{q-1}\right) dt + \lambda_1 M^p \int_{1-\delta}^{1} t^{p-1}(\alpha t + \beta)dt + \\
&+ \lambda_2 M^q \int_{1-\delta}^{1} t^{q-1}(\alpha t + \beta)dt + M\delta.
\end{aligned}
$$

Dividing by $M\delta$ and rearranging we obtain:

$$
\begin{aligned}
\fint_{M-M\delta}^{M} \left(\lambda_1 t^{p-1} + \lambda_2 t^{q-1}\right) dt \geq 1 &+ \lambda_1 M^{p-1} \fint_{1-\delta}^{1} t^{p-1}(\alpha t + \beta)dt \\
&+ \lambda_2 M^{q-1} \fint_{1-\delta}^{1} t^{q-1}(\alpha t + \beta)dt.
\end{aligned}
\tag{6.31}
$$

We notice that the last two terms are exactly the ones that appear in the orthogonality condition, therefore, to understand their behavior as δ approaches 0, we need to study the right-hand side of the system (6.30). If we expand the first term in its Taylor series with respect to δ we have:

$$
\left(1 - \frac{p-1}{2}\delta + o(\delta)\right) - \left(1 + \frac{p-1}{2}\delta + o(\delta)\right) = -(p-1)\delta + o(\delta)
$$

and the same is for the other term. Since they both are of order δ, if we let $\delta \to 0^+$ in (6.31) we obtain

$$
\lambda_1 M^{p-1} + \lambda_2 M^{q-1} \geq 1
$$

The function $\lambda_1 t^{p-1} + \lambda_2 t^{q-1}$ is strictly increasing because λ_1 and λ_2 are both positive, therefore this implies that $M \geq T$, which is absurd because we supposed that $M < T$. This allows us to write u as in (6.25).

- **Uniqueness of multipliers**

  Lastly we shall prove that multipliers λ_1, λ_2, and hence maximizer, are unique. For this proof it is convenient to express u in a slightly different way:

$$
u(t) = \frac{1}{d!}\left[\mathrm{Log}_-\left((c_1 t)^{p-1} + (c_2 t)^{q-1}\right)\right]^d.
$$

To emphasize that u is parametrized by c_1, c_2 we may write $u(t; c_1, c_2)$. Now we define

$$
f(c_1, c_2) = p \int_0^T t^{p-1} u(t; c1, c2)dt, \quad g(c_1, c_2) = q \int_0^T t^{q-1} u(t; c1, c2)dt.
$$

We want to highlight that, even if it is not explicit, also T depends on c_1 and c_2. Nevertheless, these functions are differentiable since both T and u are differentiable with respect to (c_1, c_2) and $t^{p-1}u$, $t^{q-1}u$ and their derivatives are bounded in $(0, T)$. Our maximizer u satisfies the constraints only if $f(c_1, c_2) = A^p$, $g(c_1, c_2) = B^q$. Therefore, to prove uniqueness of the maximizer we need to show that level sets $\{f = A^p\}$ and $\{g = B^q\}$ intersect only in one point.

First of all we need to study endpoints, namely when one of c_1 or c_2 is 0. For example, if $c_2 = 0$:

$$\begin{aligned} f(c_1,0) &= p \int_0^{1/c_1} t^{p-1} \frac{1}{d!} \left[-\log(c_1 t)^{p-1}\right]^d dt \overset{\tau = c_1 t}{=} \\ &= \frac{p(p-1)^d}{c_1^p d!} \int_0^1 \tau^{p-1} \left[-\log(\tau)\right]^d d\tau = \frac{\kappa_p^d}{c_1^p} = A^p \implies c_{1,f} = \frac{\kappa_p^{d/p}}{A}. \end{aligned}$$

The same can be done for g and setting $c_1 = 0$ instead of $c_2 = 0$. Thus, we obtain four points:

$$c_{1,f} = \frac{\kappa_p^{d/p}}{A}, \; c_{1,g} = \left(\frac{p-1}{q}\right)^{d/q} \frac{1}{B}, \; c_{2,f} = \left(\frac{q-1}{p}\right)^{d/p} \frac{1}{A}, \; c_{2,g} = \frac{\kappa_q^{d/q}}{B}.$$

In the regime we are considering one has that $c_{1,f} < c_{1,g}$ and $c_{2,f} > c_{2,g}$, indeed:

$$\begin{aligned} c_{1,f} < c_{1,g} &\iff \frac{\kappa_p^{d/p}}{A} < \left(\frac{p-1}{q}\right)^{d/q} \frac{1}{B} \iff \frac{B}{A} < \kappa_p^{d\left(\frac{1}{q} - \frac{1}{p}\right)} \left(\frac{p}{q}\right)^{d/q}, \\ c_{2,f} > c_{2,g} &\iff \left(\frac{q-1}{p}\right)^{d/p} \frac{1}{A} > \frac{\kappa_q^{d/q}}{B} \iff \frac{B}{A} > \kappa_q^{d\left(\frac{1}{q} - \frac{1}{p}\right)} \left(\frac{p}{q}\right)^{d/p}, \end{aligned}$$

which are exactly conditions in (6.13). Since there is this arrangement of these points we expect there is an intersection between level sets. Firstly we notice that, for every $c_1 \in (0, c_{1,f})$, there exist a unique value of c_2 for which $f(c_1, c_2) = A^p$. Indeed, from previous computations we notice that $f(c_1,0)$ is a decreasing function hence $f(c_1,0) > A^p$, while $\lim_{c_2 \to +\infty} f(c_1, c_2) = 0$, therefore from the intermediate value theorem it follows that $f(c_1, c_2) = A^p$ for some c_2. The uniqueness of this value follows from strict monotonicity of $f(c_1, \cdot)$, indeed:

$$\frac{\partial f}{\partial c_1}(c_1, c_2) = -\frac{p(p-1)}{(d-1)!} c_1^{p-2} \int_0^T \frac{t^{2(p-1)}}{(c_1 t)^{p-1} + (c_2 t)^{q-1}} \left[-\log\left((c_1 t)^{p-1} + (c_2 t)^{q-1}\right)\right]^{d-1} dt. \tag{6.32}$$

is always strictly negative. We point out that the term $\frac{\partial T}{\partial c_1}(c_1, c_2) u(T; c_1, c_2)$, that should appear since T depends on c_1, is 0 because u is 0 in T. The same is true for g, therefore on the interval $(0, c_{1,f})$ the level sets of f and g can be seen as the graph of two functions φ, γ. Since f and g are both differentiable, from the implicit function theorem we have that φ and γ are differentiable with respect to c_1.

After defining φ and γ we want to prove that $(\varphi - \gamma)' < 0$. Still from the implicit function theorem we have

$$\begin{aligned} \frac{d}{dc_1}(\varphi - \gamma)(c_1) = -\frac{\frac{\partial f}{\partial c_1}(c_1, \varphi(c_1))}{\frac{\partial f}{\partial c_2}(c_1, \varphi(c_1))} + \frac{\frac{\partial g}{\partial c_1}(c_1, \gamma(c_1))}{\frac{\partial g}{\partial c_2}(c_1, \gamma(c_1))} < 0 \iff \\ \mathcal{I}(c_1) = \frac{\partial f}{\partial c_1}(c_1, \varphi(c_1)) \frac{\partial g}{\partial c_2}(c_1, \gamma(c_1)) - \frac{\partial f}{\partial c_2}(c_1, \varphi(c_1)) \frac{\partial g}{\partial c_1}(c_1, \gamma(c_1)) > 0 \end{aligned}$$

As for (6.32) the other derivatives are computed. To simplify the notation we define $h(t;c_1,c_2) = \frac{1}{(d-1)!}\frac{1}{(c_1t)^{p-1}+(c_2t)^{q-1}}\left[-\log\left((c_1t)^{p-1}+(c_2t)^{q-1}\right)\right]^{d-1}$. From Fubini's theorem, we can write the product of the integrals as a double integral:

$$\begin{aligned}\mathcal{I}(c_1) &= p(p-1)q(q-1)c_1^{p-2}\gamma(c_1)^{q-2}\int_{[0,T]^2} h(t;c_1,\varphi(c_2))h(s;c_1,\gamma(c_2))t^{2(p-1)}s^{2(q-1)}dtds+\\ &\quad - p(q-1)q(p-1)c_1^{p-2}\varphi(c_1)^{q-2}\int_{[0,T]^2} h(t;c_1,\varphi(c_2))h(s;c_1,\gamma(c_2))t^{p+q-2}s^{p+q-2}dtds.\end{aligned}$$

When level sets intersect we have $\varphi(c_1) = \gamma(c_1)$. In this situation we can factorize the terms outside the integral and notice that the sign of $\mathcal{I}$ depends only on the sign of:

$$\begin{aligned}&\int_{[0,T]^2} h(t;c_1,\varphi(c_1))h(s;c_1,\gamma(c_1))\left(t^{2(p-1)}s^{2(q-1)} - t^{p+q-2}s^{p+q-2}\right)dtds =\\ &= \int_{[0,T]^2} h(t;c_1,\varphi(c_1))h(s;c_1,\gamma(c_1))t^{p-2}s^{q-2}\left(t^ps^q - t^qs^p\right)dtds.\end{aligned}$$

In order to simplify the notation once again, we set $H(t,s;c_1) = h(t;c_1,\varphi(c_1))h(s;c_1,\gamma(c_1))$. Let $T_1 = [0,T]^2 \cap \{t > s\}$ and $T_2 = [0,T]^2 \cap \{t < s\}$. We can split the above integral in two parts:

$$\int_{T_1} H(t,s;c_1)t^{p-2}s^{q-2}\left(t^ps^q - t^qs^p\right)dtds + \int_{T_2} H(t,s;c_1)t^{p-2}s^{q-2}\left(t^ps^q - t^qs^p\right)dtds.$$

Then, considering the change of variables that swaps t and s, the domain of integration becomes T_1 and since H is symmetric in t and s, we have that the previous quantity is equal to

$$\begin{aligned}&\int_{T_1} H(t,s;c_1)\left(t^{p-2}s^{q-2} - t^{q-2}s^{p-2}\right)\left(t^ps^q - t^qs^p\right)dtds =\\ &= \int_{T_1} H(t,s;c_1)\frac{1}{t^2s^2}\left(t^ps^q - t^qs^p\right)^2 dtds,\end{aligned}$$

which is strictly positive.We are now in the position to prove the uniqueness of multipliers.

First of all, since $(\varphi-\gamma)' < 0$ whenever $\varphi(c_1) = \gamma(c_1)$, we point out that for every point of intersection there exist $\delta > 0$ such that $\varphi(t) > \gamma(t)$ for $t \in (c_1-\delta, c_1)$ and $\varphi(t) < \gamma(t)$ for $t \in (c_1, c_1+\delta)$.

Define $c_1^* := \sup\{c_1 \in [0, c_{1,f}] : \forall t \in [0, c_1]\ \varphi(t) \geq \gamma(t)\}$. This is an intersection point between φ and γ (if $\varphi(c_1^*) > \gamma(c_1^*)$ due to continuity there would be $\varepsilon > 0$ such that $\varphi(c_1^*+\varepsilon) > \gamma(c_1^*+\varepsilon)$ which contradicts the definition of c_1^*) and it is the first one, because we saw that after every intersection point there is an interval where $\varphi < \gamma$. Lastly, since $\varphi(0) > \gamma(0)$ and $\varphi(c_{1,f}) = 0 < \gamma(c_1, f)$, we have that $0 < c_1^* < c_{1,f}$.

Suppose now that there is a second point of intersection $\tilde{c}_1$ after the first one. Since immediately after c_1^* we have that φ becomes smaller than γ, this second point of intersection is given by $\tilde{c}_1 = \sup\{c_1 \in [c_1^*, c_{1,f}] : \forall t \in [c_1^*, c_1]\ \varphi(t) \leq \gamma(t)\}$. Considering that this is an intersection point, there exist an interval before $\tilde{c}_1$ where φ is strictly greater than γ

which is absurd, hence c_1^* is the only intersection point between φ and γ.
Therefore, the pair $(c_1^*, \varphi(c_1^*) = c_2^*)$ is the unique pair of multipliers for which

$$p \int_0^T t^{p-1} u(t; c_1^*, c_2^*) dt = A^p, \quad q \int_0^T t^{q-1} u(t; c_1^*, c_2^*) dt = B^q$$

and, in the end, $u(t; c_1^*, c_2^*)$ is the unique maximizer for (6.22). $\square$

Appendix A

Unconditional convergence

In some cases we had to deal with series over multiple indices. Formally, the possibility of manipulating these series is related to the property of *unconditional convergence.*

Definition A.1. *Let $\{x_j\}_{j\in\mathcal{J}}$ be a countable subset of a Banach space X. The series $\sum_{j\in\mathcal{J}} x_j$ is said to **converge unconditionally** to some $x \in X$ if, for every $\varepsilon > 0$, there exist a finite subset J_0 of $\mathcal{J}$ such that*

$$\|x - \sum_{j\in J} x_j\| \leq \varepsilon$$

for every finite set $J \supseteq J_0$.

The notion of unconditional convergence is of crucial importance in cases where an exchange between an operator and a series or a certain order of summation is required. Here we present two results that are useful in such situations.

Proposition A.2. *Let $A \in \mathscr{B}(X, Y)$. If $\sum_{j\in\mathcal{J}} x_j$ converges unconditionally to x in X, then $\sum_{j\in\mathcal{J}} Ax_j$ converges unconditionally to Ax in Y.*

Proof. Let $\varepsilon > 0$. By definition, there exist $J_0 \subseteq \mathcal{J}$ finite such that $\|x - \sum_{j\in J} x_j\|_X \leq \varepsilon$ for every finite set $J \supseteq J_0$. Then:

$$\|Ax - \sum_{j\in J} Ax_j\|_Y = \|A(x - \sum_{j\in J} x_j)\|_Y \leq \|A\|\|x - \sum_{j\in J} x_j\|_X < \|A\|\varepsilon,$$

thus the series $\sum_{j\in\mathcal{J}} Ax_j$ converges unconditionally to Ax. $\square$

Proposition A.3. *Suppose that $\sum_{(n,m)\in\mathbb{N}^2} x_{n,m}$ converges unconditionally to $x \in X$. Then the inner partial sum $s_{n,M} = \sum_{m=1}^{M} x_{n,m}$ converges to some $y_n \in X$ for every $n \in \mathbb{N}$ and $x = \sum_{n\in\mathbb{N}} y_n$ with unconditional convergence. Similarly, $\sum_{n=1}^{N} x_{n,m}$ converges to some $z_m \in X$ for every $m \in \mathbb{N}$ and $x = \sum_{m\in\mathbb{N}} z_m$.*

Proof. Since $\sum_{(n,m)\in\mathbb{N}^2} x_{n,m}$ is unconditionally convergent to x, given $\varepsilon > 0$, by definition there exist $J_0 \subset \mathbb{N}^2$ such that $\|x - \sum_{(n,m)\in J} x_{n,m}\| < \varepsilon$ for every finite set J containing

J_0. Without loss of generality, we can suppose that J_0 is of the form $J_0 = \{(n,m) \in \mathbb{N}^2 : n \leq N_0,\ m \leq M_0\}$ for some $N_0, M_0 \in \mathbb{N}$. In such a way we have:

$$\|x - \sum_{n=1}^{N}\sum_{m=1}^{M} x_{n,m}\| < \varepsilon \tag{A.1}$$

for every $N \geq N_0$ and $M \geq M_0$. Consider now $I \subset \mathbb{N}$ finite and $M_1, M_2 \in \mathbb{N}$ such that $M_0 \leq M_1 < M_2$. Then:

$$\begin{aligned}\|\sum_{n\in I} s_{n,M_2} - \sum_{n\in I} s_{n,M_1}\| &= \|\sum_{n\in I}\sum_{m=1}^{M_2} x_{n,m} - \sum_{n\in I}\sum_{m=1}^{M_1} x_{n,m}\| = \|\sum_{n\in I}\sum_{m=M_1+1}^{M_2} x_{n,m}\| = \\ &= \|\sum_{n\in I}\sum_{m=M_1+1}^{M_2} x_{n,m} + \sum_{(n,m)\in J_0} x_{n,m} - \sum_{(n,m)\in J_0} x_{n,m}\|\end{aligned}$$

Letting $J = J_0 \cup (I \times \{M_1+1,\dots,M_2\}) \subset \mathbb{N}^2$ and using triangular inequality we obtain:

$$\|\sum_{n\in I}(s_{n,M_2} - s_{n,M_1})\| \leq \|x - \sum_{(n,m)\in J} x_{n,m}\| + \|x - \sum_{(n,m)\in J_0} x_{n,m}\| < 2\varepsilon \tag{A.2}$$

because $J \supseteq J_0$. This proves that, for every $n \in \mathbb{N}$ and for every $I \subset \mathbb{N}$ finite, the sequence $\{\sum_{n\in I} s_{n,M}\}_{M\in\mathbb{N}}$ is a Cauchy sequence in X which is a Banach space, therefore it is convergent. In particular, taking $I = \{n\}$, we obtain that the sequence $\{s_{n,M}\}_{M\in\mathbb{N}}$ converges to some $y_n \in X$ for every $n \in \mathbb{N}$. Moreover, since I is finite, we have that $\sum_{n\in I} s_{n,M} \overset{M\to+\infty}{\longrightarrow} \sum_{n\in I} y_n$.

Now we have to show that $\sum_{n\in\mathbb{N}} y_n$ converges unconditionally to x. Consider $I \subset \mathbb{N}$ finite such that $\{1,\dots,N_0\} \subseteq I$. First of all we notice that:

$$\lim_{M_2\to+\infty}\left(\sum_{n\in I} s_{n,M_1} - \sum_{n\in I} s_{n,M_2}\right) = \sum_{n\in I} s_{n,M_1} - \sum_{n\in I} y_n.$$

Therefore, taking the $\limsup_{M_2\geq M_1}$ in (A.3) we obtain

$$\|\sum_{n\in I} s_{n,M_1} - \sum_{n\in I} y_n\| = \limsup_{M_2\geq M_1}\|\sum_{n\in I} s_{n,M_1} - \sum_{n\in I} s_{n,M_2}\| \leq 2\varepsilon. \tag{A.3}$$

Then we have:

$$\|x - \sum_{n\in I} y_n\| \leq \|x - \sum_{n\in I} s_{n,M_1}\| + \|\sum_{n\in I} s_{n,M_1} - \sum_{n\in I} y_n\| \overset{(A.1)+(A.3)}{<} 3\varepsilon.$$

The last part of the statement easily follows swapping n and m in previous computations. ☐

Appendix B

Calculations

B.1 Constant in Lieb's inequality

In the last part of the proof of Lieb's inequality 3.4 we used the following equality $A_{p'}^{d} A_{2/p'}^{2d/p'} A_{(p/p')'}^{d/p'} = (2/p)^{d/p}$ without proving it. We recall that the Babenko-Bechner constant A_p is given by:

$$A_p = \left(\frac{p^{1/p}}{p'^{1/p'}} \right)^{1/2}.$$

Since there is an exponent 1/2 in the Babenko-Bechner constant it is better to compute the square of $A_{p'}^{d} A_{2/p'}^{2d/p'} A_{(p/p')'}^{d/p'}$. For the sake of clarity we are going to compute every single term and then we are going to multiply them.

- $A_{p'}^2 = \dfrac{p'^{1/p'}}{p^{1/p}}$;

- In order to compute $A_{2/p'}^{2\cdot 2/p'}$ we start computing $(2/p')'$:

$$\left(\frac{2}{p'}\right)' = \frac{2/p'}{2/p' - 1} = \frac{2}{2 - p'},$$

  therefore

$$A_{2/p'}^{2\cdot 2/p'} = \left[\left(\frac{2}{p'}\right)^{p'/2} \left(\frac{2-p'}{2}\right)^{(2-p')/2} \right]^{2/p'} = \frac{2}{p'} \left(\frac{2-p'}{2}\right)^{(2-p')/(p')} =$$
$$= \frac{2^{2(1-1/p')}}{p'} (2-p')^{2/p'-1} = \frac{2^{2/p}}{p'} (2-p')^{1/p'-1/p};$$

- Like the previous case, we start computing $(p/p')'$:

$$\left(\frac{p}{p'}\right)' = \frac{p/p'}{p/p' - 1} = \frac{p}{p - p'},$$

hence

$$A_{(p/p')'}^{2/p'} = \left[\left(\frac{p}{p-p'} \right)^{(p-p')/p} \left(\frac{p'}{p} \right)^{p'/p} \right]^{1/p'} = \left(\frac{p}{p-p'} \right)^{1/p'-1/p} \left(\frac{p'}{p} \right)^{1/p} =$$
$$= \left(1 - \frac{p'}{p} \right)^{1/p-1/p'} \left(\frac{p'}{p} \right)^{1/p} = (2-p')^{1/p-1/p'} \left(\frac{p'}{p} \right)^{1/p}.$$

We are now ready to calculate the product of the three constant:

$$A_{p'}^2 A_{2/p'}^{2\cdot 2/p'} A_{(p/p')'}^{2/p'} = \frac{p'^{1/p'}}{p^{1/p}} \frac{2^{2/p}}{p'} (2-p')^{1/p'-1/p} (2-p')^{1/p-1/p'} \left(\frac{p'}{p} \right)^{1/p} =$$
$$= 2^{2/p} p^{-2/p} p'^{1/p'+1/p-1} = \left(\frac{2}{p} \right)^{2/p}$$

which is the desired result.

B.2 Curious inequality between conjugate exponent

We consider inequality (6.12), which we rewrite for the sake of clarity:

$$\kappa_p^{d\left(\frac{1}{q}-\frac{1}{p}\right)} \left(\frac{p}{q} \right)^{\frac{d}{q}} \geq \left(\frac{\kappa_p^{\kappa_p}}{\kappa_q^{\kappa_q}} \right)^d$$

Firstly we want to restate in a more concise way. Recalling that $\kappa_p = \frac{1}{p'}$, where p' is the conjugate exponent of p, we have:

$$\left(\frac{1}{p'} \right)^{\frac{1}{q}-\frac{1}{p}} \left(\frac{p}{q} \right)^{\frac{1}{q}} \geq \left(\frac{1}{p'} \right)^{\frac{1}{p'}} \left(\frac{1}{q'} \right)^{-\frac{1}{q'}} \iff \left(\frac{1}{p'} \right)^{\frac{1}{q}-\frac{1}{p}-\frac{1}{p'}} \left(\frac{1}{q'} \right)^{\frac{1}{q'}} \left(\frac{p}{q} \right)^{\frac{1}{q}} \geq 1$$

but, since $\frac{1}{p} + \frac{1}{p'} = 1$ and $\frac{1}{q} - 1 = \frac{1}{q'}$, in conclusion we have:

$$\left(\frac{p'}{q'} \right)^{\frac{1}{q'}} \left(\frac{p}{q} \right)^{\frac{1}{q}} \geq 1$$

In order to prove that this inequality holds for every pair of $p, q > 1$ we consider the left-hand side as function of $x = \frac{1}{p}$ and $y = \frac{1}{q}$ (therefore $\frac{1}{p'} = 1 - x$ and $\frac{1}{q'} = 1 - y$). If we take the logarithm of this quantity we want to show that:

$$f(x, y) = (1-y)\left[\log(1-y) - \log(1-x)\right] + y\left[\log(y) - \log(x)\right] \geq 0.$$

The partial derivative of f with respect to x is:

$$\frac{\partial f}{\partial x}(x, y) = \frac{1-y}{1-x} - \frac{y}{x} = \frac{x-y}{x(1-x)}.$$

Since $x \in (0,1)$, $\frac{\partial f}{\partial x}(x, y)$ is negative for $x < y$ and positive for $x > y$, so f has a minimum for $x = y$ where $f(x, x) = 0$.

B.3 Another inequality between conjugate exponents

We want to prove that inequality (6.14) holds, namely that:

$$\kappa_q^{d\left(\frac{1}{q}-\frac{1}{p}\right)}\left(\frac{p}{q}\right)^{\frac{d}{p}} < \kappa_p^{d\left(\frac{1}{q}-\frac{1}{p}\right)}\left(\frac{p}{q}\right)^{\frac{d}{q}}$$

whenever $p \neq q$. As for the previous section, since $\kappa_p = \frac{1}{p'}$ and $\kappa_q = \frac{1}{q'}$ we can write the inequality in a more concise way:

$$\left(\frac{1}{q'}\right)^{\frac{1}{q}-\frac{1}{p}}\left(\frac{p}{q}\right)^{\frac{1}{p}} < \left(\frac{1}{p'}\right)^{\frac{1}{q}-\frac{1}{p}}\left(\frac{p}{q}\right)^{\frac{1}{q}} \iff \left(\frac{q\ p'}{q'\,p}\right)^{\frac{1}{q}-\frac{1}{p}} < 1 \iff \left(\frac{q-1}{p-1}\right)^{\frac{1}{q}-\frac{1}{p}} < 1.$$

If $q > p$ the base is greater than 1 while the exponent is less than 1, whereas if $q < p$ the converse happens, which proves that inequality holds whenever $p \neq q$.

Appendix C

Helly's selection theorem

In this chapter we prove Helly's selection theorem, which was used in the proof of Proposition 6.3.

Theorem C.1. *Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of monotonically increasing functions on $\mathbb{R}$ and suppose that the sequence is uniformly bounded, namely that $f_n(x) \in [a,b]$ for every $n \in \mathbb{N}$, every $x \in \mathbb{R}$ and some $a < b$ in $\mathbb{R}$. Then, there exist a monotonically increasing function f and a subsequence $\{f_{n_k}\}_{k\in\mathbb{N}}$ such that f_{n_k} is pointwise convergent to f.*

Proof. We begin by constructing f over rational numbers. Let $\{q_1, q_2, \ldots\}$ be an enumeration of $\mathbb{Q}$.

Since $\{f_n(q_1)\}_{n\in\mathbb{N}}$ is a bounded sequence in $\mathbb{R}$, it follows from Bolzano-Weiestrass' theorem that there exist $S_1 \subseteq \mathbb{N}$ countable such that $\{f_n(q_1)\}_{n\in S_1}$ is convergent to some $y_1 \in [a,b]$. Then, if we consider the sequence $\{f_n(q_2)\}_{n\in S_1}$, we see that this is a bounded sequence in $\mathbb{R}$ and again, from Bolzano-Weiestrass' theorem it follows that there exist $S_2 \subseteq S_1$ countable such that $\{f_n(q_2)\}_{n\in S_2}$ converges to some $y_2 \in [a,b]$. Repeating this argument lead to a family of countable subsets of $\mathbb{N}$ such that $S_1 \supseteq S_2 \supseteq \cdots$ such that $\{f_n(q_i)\}_{n\in S_i}$ converges to y_i.
Consider now the set $S \subseteq \mathbb{N}$ built in the following way: the first element of S is the first element of S_1, the second element of S is the second element of S_2 and so on. Then, if we consider $q_i \in \mathbb{Q}$, we have that, up to the first $i-1$ terms, $\{f_n(q_i)\}_{n\in S}$ is a subsequence of $\{f_n(q_i)\}_{n\in S_i}$, therefore $\{f_n(q_i)\}_{n\in S}$ converges to y_i. Moreover, we notice that if $q_i < q_j$ then, by the monotonicity of f_n, we have $f_n(q_i) \leq f_n(q_j)$ for every $n \in S$. Taking the limit for $n \to +\infty$ we conclude that $y_i \leq y_j$.

Up to now we built a subsequence of $\{f_n(q_1)\}_{n\in\mathbb{N}}$ that is pointwise convergent over rational numbers. Through the density of $\mathbb{Q}$ in $\mathbb{R}$ it is immediate to create a monotonically increasing function g over $\mathbb{R}$. Indeed, if $q \in \mathbb{Q}$ we simply let $g(q) = y_q$, while if $r \in \mathbb{R}\backslash\mathbb{Q}$ we let $g(r) = \sup_{\substack{q\in\mathbb{Q}\\ q\leq r}} g(q)$. We point out that the supremum is finite because y_q are bounded from above. Moreover, it is clear from the definition that g is monotonically increasing.

The proof is still not complete because, up to now, we only know that the subsequence $\{f_n\}_{n\in S}$ is pointwise converging to g only over rational numbers. However, we can show that if g is continuous in $r \in \mathbb{R}$, then $\lim_{\substack{n\to+\infty\\ n\in S}} f_n(x) = g(x)$. If g is continuous in r, given

$\varepsilon > 0$ there exist $\delta > 0$ such that $|g(y) - g(x)| < \varepsilon$ if $|y - x| < \delta$. In particular, we can pick $q_1, q_2 \in \mathbb{Q}$ such that $r - \delta < q_1 < r < q_2 < r + \delta$. Since g is monotonically increasing this implies that $0 \leq g(r) - g(q_1) < \varepsilon$ and $0 \leq g(q_2) - g(r) < \varepsilon$. Moreover, for n sufficiently large, we have that $|f_n(q_1) - g(q_1)| < \varepsilon$ and $|f_n(q_2) - g(q_2)| < \varepsilon$. These, together with the fact that f_n are monotonically increasing leads to:

$$f_n(q_1) \leq f(x) \leq f_n(q_2) \implies$$
$$-2\varepsilon < f_n(q_1) - g(q_1) + g(q_1) - g(x) \leq f_n(x) - g(x) \leq f_n(q_2) - g(q_2) + g(q_2) - g(x) < 2\varepsilon,$$

which means that $f_n(x)$ converges to $g(x)$.

In order to conclude, we have to deal with the points where g is not continuous. It is well known that a monotonic function on an open interval is continuous except possibly on a countable subset $J \subset \mathbb{R}$ (see [26, Section 6.1 Theorem 1] or [27, Theorem 4.30]). Since J is at most countable we con consider an enumeration $\{j_1, j_2, \ldots\}$. If we repeat the same argument we used at the beginning of the proof, we see that there exist a countable subset $P \subseteq S$ such that $\{f_n\}_{n \in P}$ is pointwise converging to g in $\mathbb{R} \setminus J$ and to some values in J. Taking f as the limit function of $\{f_n\}_{n \in P}$ the proof is complete. $\square$

Bibliography

[1] J. Bell. *Trace Class Operators and Hilbert-Schmidt Operators*. 2016.

[2] Haim Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Vol. 2. 3. Springer, 2011.

[3] N. G. de Bruijn. "Uncertainty principles in Fourier analysis". In: *Inequalities (Proc. Sympos. Wright-Patterson Air Force Base, Ohio, 1965)*. Academic Press, New York, 1967, pp. 57–71.

[4] John B Conway. *A course in functional analysis*. Vol. 96. Springer, 2019.

[5] Ingrid Daubechies. "Time-frequency localization operators: a geometric phase space approach". In: *IEEE Transactions on Information Theory* 34.4 (1988), pp. 605–612.

[6] G. B Folland. *Harmonic analysis in phase space*. 122. Princeton university press, 1989.

[7] Gerald B Folland and Alladi Sitaram. "The uncertainty principle: a mathematical survey". In: *Journal of Fourier analysis and applications* 3 (1997), pp. 207–238.

[8] Dennis Gabor. "Theory of communication. Part 1: The analysis of information". In: *Journal of the Institution of Electrical Engineers-part III: radio and communication engineering* 93.26 (1946), pp. 429–441.

[9] Loukas Grafakos. *Classical fourier analysis*. Vol. 2. Springer, 2008.

[10] Karlheinz Gröchenig. *Foundations of time-frequency analysis*. Springer Science & Business Media, 2001.

[11] Littlewood J. E. Hardy G. H. and G. Pólya. *Inequalities*. Cambridge University Press, 1952.

[12] W. Heisenberg. "Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik". In: *Zeitschrift für Physik* 43 (1927).

[13] Morris W. Hirsch. *Differential topology*. Vol. 33. Springer Science & Business Media, 2012.

[14] Helge Knutsen. "Norms and Eigenvalues of Time-Frequency Localization Operators".

[15] Steven G Krantz and Harold R Parks. *A primer of real analytic functions*. Springer Science & Business Media, 2002.

[16] Steven G Krantz and Harold R Parks. *Geometric integration theory.* Springer Science & Business Media, 2008.

[17] H. J. Landau and H. O. Pollak. “Prolate spheroidal wave functions, Fourier analysis and uncertainty—II”. In: *Bell System Technical Journal* 40.1 (1961), pp. 65–84.

[18] Henry J Landau and Henry O Pollak. “Prolate spheroidal wave functions, fourier analysis and uncertainty—III: the dimension of the space of essentially time-and band-limited signals”. In: *Bell System Technical Journal* 41.4 (1962), pp. 1295–1336.

[19] E. H Lieb. “Integral bounds for radar ambiguity functions and Wigner distributions”. In: *Journal of Mathematical Physics* 31.3 (1990), pp. 594–599.

[20] Elliott H. Lieb and Michael Loss. *Analysis.* Vol. 14. American Mathematical Soc., 2001.

[21] Boris Mityagin. “The zero set of a real analytic function”. In: *arXiv preprint arXiv:1512.07276* (2015).

[22] F. Nicola and L. Rodino. *Global Pseudo-Differential Calculus on Euclidean Spaces.* Vol. 4. Springer, 2010.

[23] Fabio Nicola and Paolo Tilli. “The Faber–Krahn inequality for the short-time Fourier transform”. In: *Inventiones mathematicae* 230.1 (2022), pp. 1–30.

[24] Fabio Nicola and Paolo Tilli. “The norm of time-frequency localization operators”. In: *arXiv preprint arXiv:2207.08624* (2022).

[25] Michael Reed, Barry Simon, et al. *I: Functional analysis.* Vol. 1. Gulf Professional Publishing, 1980.

[26] Halsey Lawrence Royden and Patrick Fitzpatrick. *Real analysis.* Vol. 32. Macmillan New York, 1988.

[27] Walter Rudin et al. *Principles of mathematical analysis.* Vol. 3. McGraw-hill New York, 1976.

[28] Arthur Sard. “The measure of the critical values of differentiable maps”. In: (1942).

[29] D. Slepian. “Prolate spheroidal wave functions, Fourier analysis and uncertainty—IV: extensions to many dimensions; generalized prolate spheroidal functions”. In: *Bell System Technical Journal* 43.6 (1964), pp. 3009–3057.

[30] D. Slepian and H. O. Pollak. “Prolate spheroidal wave functions, Fourier analysis and uncertainty—I”. In: *Bell System Technical Journal* 40.1 (1961), pp. 43–63.

[31] Man Wah Wong. *Wavelet transforms and localization operators.* Vol. 136. Springer Science & Business Media, 2002.

[32] Kehe Zhu. *Analysis on Fock spaces.* Vol. 263. Springer Science & Business Media, 2012.