NTNU
Norwegian University of
Science and Technology

Norms and Eigenvalues of Time-Frequency Localization Operators

Helge Knutsen

Master of Science in Physics and Mathematics
Submission date: June 2018
Supervisor: Eugenia Malinnikova, IMF

Norwegian University of Science and Technology
Department of Mathematical Sciences

Preface

This master's thesis represents the final submission at the study programme Industrial Mathematics, within Applied Physics and Mathematics at the Norwegian University of Science and Technology (NTNU). The thesis was written during the spring of 2018 under supervision of Professor Eugenia Malinnikova at the Department of Mathematics.

I would like to thank my supervisor Eugenia Malinnikova for her excellent guidance during this and last semester. My thesis work has certainly benefited from the many insightful conversations with her, in addition to her constant input and always inspiring feedback. I would also like to express my gratitude to her for introducing me to the research area of time-frequency analysis and for giving me the opportunity to participate on the BCAM meeting in Bilbao earlier this year.

Trondheim, Norway — Helge Knutsen
June 19, 2018

Abstract

In this report we study and compare two types of time-frequency localization operators, the first is based on composition of projections in time and frequency, and the second is Daubechies' localization operator. We provide a review of several uncertainty principles in time-frequency analysis and formulate these principles in terms of the operator norm of the localization operators.

Proceeding, the main focus is a particular kind of the Daubechies' localization operator. These operators are characterized by a window and a weight function, and with a Gaussian window and spherically symmetric weight we attain simple, explicit formulas for the eigenvalues. For such operators we consider the case when the weight takes the form of the characteristic function of some spherically symmetric subset of the time-frequency plane.

Based on the measure of the subset in question, we determine simple upper and lower bound estimates for the operator norm. For some specific examples of subsets we provide more accurate estimates for the operator norm. Notably, we consider the spherically symmetric Cantor set and derive precise asymptotics for the operator norm of the associated localization operator.

Sammendrag

I denne rapporten studerer vi og sammenlikner to typer tidsfrekvens-lokaliseringsoperatorer, den første er basert på komposisjon av projeksjoner i tid og frekvens, og den andre er Daubechies lokaliseringsoperator. Vi har en gjennomgang av flere av uskarphetsprinsippene i tidsfrekvensanalyse og formulerer disse prinsippene ved hjelp av operatornormen til lokaliseringsoperatorene.

Videre er hovedfokuset en bestemt type av Daubechies lokaliseringsoperator. Disse operatorene er karakterisert av en vindu- og en vektfunksjon, og med et Gaussisk vindu og sfærisk-symmetrisk vekt får vi enkle, eksplisitte formler for egenverdiene. For slike operatorer betrakter vi tilfellet hvor vekten er på formen til en karakterstisk funksjon av en sfærisk-symmetrisk undermengde av tids-frekvensplanet.

Basert på målet til den aktuelle undermengden bestemmer vi enkle øvre og nedre estimater for operatornormen. Mer presise estimater av operatornormen er gitt for enkelte spesifikke eksempler på undermengder. Blant annet betrakter vi den sfærisk symmetriske Cantor-mengden og utleder presise asymptoter for operatornormen til den korresponderende lokaliseringsoperatoren.

List of Figures

Contents

1 Introduction

An old and arguably one of the most important problems in signal analysis is the problem of localization in time and frequency. In applications, we often wish to analyze signals on different time-frequency domains, and we would therefore attempt to concentrate signals on these domains. For this purpose, we consider the aptly named time-frequency localization operators. Different approaches for how to construct such operators have been suggested either based on separate or joint time-frequency representations. In the 1960s a certain kind of localization operator was studied by Laundau, Pollack and Slepian (see [1],[2],[3]), which in its generality can be summarized as compositions of projections in frequency and time. In the 1980s Ingrid Daubechies presented an alternative family of operators, now based on a joint time-frequency representation[4]. We will consider and compare both classes of operators.

This report is divided into three main sections (Chapter 2-4). The first section, Chapter 2, contains what could be considered necessary background theory. In particular, we introduce the standard terminology of Fourier and Short-Time Fourier transform (STFT), which is our framework for performing time-frequency analysis. Relevant concepts and results from functional analysis are then covered before formally introducing the two classes of localization operators.

Note, however, that regardless of which localization operators we choose to work with, these operators will be subject to the fundamental barrier of time-frequency analysis, namely the uncertainty principles. Many versions of these principles exist, but all embody the notion that a signal cannot be highly localized in both time and frequency simultaneously. Since the optimal efficiency of any given localization operator is measured by its operator norm, it stands to reason that the uncertainty principles will produce non-trivial estimates of the operator norm. In Chapter 3 we review some of the classical uncertainty principles and formulate them in terms of the operator norm of the relevant localization operator.

At the end of the chapter we mention some more recent developments in the research area of uncertainty principles. Here we start to take into account some of the geometry of the time-frequency domains. Among the results discussed is Semyon Dyatlov's findings regarding projections onto fractal sets in time and frequency. From his 2017-notes[5] we obtain sequences of subsets $\{X_n\}_n$ such that the measure $|X_n| \to \infty$ as $n \to \infty$. However, this sequence is constructed such that the operator norm of the associated localization operator, that projects onto X_n in frequency and then onto X_n in time, tends to zero. As an illustrative example, Dyatlov considers a sequence of subsets based on the n-iterate mid-third Cantor set defined in an ever increasing interval.

Inspired by this model example, we investigate if similar behaviour can be observed for Daubechies' localization operator projecting onto a Cantor type fractal set in the time-frequency plane. In this context we will restrict to a certain subfamily of symmetric operators whose eigenfunctions are known and where we have simple formulas for the associated eigenvalues. Daubechies' localization operators are characterized by a window and weight function, and in her 1988-article[4] she derives simple expressions for the eigenvalues when we choose a Gaussian window and any spherically symmetric weight. For this reason, we will primarily focus on operators with spherically symmetric weights (in addition to the fixed Gaussian window), which in turn means we will consider a spherically symmetric Cantor set in the plane.

The entire Chapter 4 is in fact dedicated to this subfamily of Daubechies' localization operators. We start by restating Daubechies' result, and in section 4.1 we recapitulate the proof, which shows that the Hermite functions $\{H_k\}_k$ constitute the eigenfunctions of the localization operator. The associated eigenvalues $\{\lambda_k\}_k$ are given on integral form. In the subsequent sections, 4.2 and 4.3, we further restrict to the case when the weight equals the characteristic function of some spherically symmetric subset. These sections contain what could be considered the original research work of the report. Here our main objective is to determine or at least estimate the operator norm of the corresponding Daubechies' operator.

In section 4.2 we discuss some common properties of the eigenvalues $\{\lambda_k\}_k$ associated with localization on a spherically symmetric subset. The eigenvalues are utilized to estimate the operator norm, and to illustrate we consider two simple but important examples of subsets, namely a disk and a ring. More generally, we derive an upper and lower bound estimate for the operator norm based on the measure of the given subset. From the upper bound estimate it follows that when keeping the measure of the subset fixed, the optimal localization occurs when the subset takes the form of a disk in the plane. Afterwards, we consider a non-trivial example where the subset has infinite measure, but where we still have good control over the operator norm.

In section 4.3 we finally narrow in our focus on localization on the mid-third Cantor set. In the spherically symmetric context we distinguish between the distance regular and the measure regular Cantor set. Proceeding, we have chosen to focus on the latter version, i.e. the measure regular Cantor set. For the n-iterate Cantor set we derive *precise* asymptotic estimates for the operator norm. From here we are, similarly to Dyatlov, able to construct a sequence of iterates whose measure tends to infinity, but where the associated operator norm tends to zero.

2 Preliminaries

This chapter serves as a brief introduction to some of the fundamentals of localization operators in time-frequency analysis: The chapter is organized in three main sections. The first, section 2.1, provides the basic setup of Fourier and Short-Time Fourier transform as our separate and joint time-frequency representation, respectively. The second, section 2.2, covers some necessary background theory from functional analysis. This theory will be applied to the final section, section 2.3, where we introduce two approaches for how to construct time-frequency localization operators.

2.1 Fourier and Short-Time Fourier Transform

In this section we formally introduce our working-defnition of the Fourier transform and provide the standard analogy of time and frequency to accompany this defnition. From here we turn to the Short-Time Fourier transform as the the main focus and establish some key properties of this transform.

Throughout this report we will work with the following normalization for the Fourier transform. For a function $f : \mathbb{R}^d \to \mathbb{C}$ the *Fourier transform* evaluated at point $\omega \in \mathbb{R}^d$ is given by

$$\hat{f}(\omega) = \int_{\mathbb{R}^d} f(t) e^{-2\pi i \omega \cdot t} \mathrm{d}t, \tag{2.1}$$

where $\omega \cdot t = \sum_{j=1}^{d} \omega_j t_j$ denotes the standard Euclidean inner product in $\mathbb{R}^d$. Observe, in order to guarantee that the above transform is well-defined, certain restrictions are necessary on f. E.g., in order to be pointwise defined everywhere, it is sufficient to consider $f \in L^1(\mathbb{R}^d)$. However, the resulting function $\hat{f}$ is not necessarily integrable. If we no longer require a pointwise description, and are instead interested in control over the target space of the transform, one natural choice is to a assume f belongs to $L^2(\mathbb{R}^d)$. A standard density argument (see Chapter 1.1. in Gröchenig's book[6]) then shows that the Fourier transform can be expressed as a *unitary operator*

$$\mathcal{F} : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d), \quad f \mapsto \mathcal{F}f = \hat{f}, \tag{2.2}$$

whose inverse is

$$\mathcal{F}^{-1}\hat{f}(t) = \int_{\mathbb{R}^d} \hat{f}(\omega) e^{2\pi i \omega \cdot t} \mathrm{d}\omega. \tag{2.3}$$

Although other function spaces may be considered, unless otherwise stated, we will always assume $f \in L^2(\mathbb{R}^d)$.

In dimension one ($d = 1$) if we interpret f as an amplitude signal depending on *time*, then its Fourier transform $\hat{f}$ corresponds to a *frequency* representation of the signal. This analogy of frequency and time also extends to higher dimensions ($d > 1$), where f can be viewed as an amplitude signal from *multiple* time sources.

Notice, however, that the pair $(f, \hat{f})$ does not offer a joint description with respect to both frequency and time. Ideally, such a description would consist of precise knowledge of the frequencies present at any given time. One attempt of attaining a simultaneous time-frequency representation of f is by the means of the *Short-Time Fourier transform* (STFT).

The STFT is often referred to as the "windowed Fourier transform" as this transform relies on an additional fixed, non-zero function, $\phi : \mathbb{R}^d \to \mathbb{C}$, known as a *window function*. From this function ϕ we generate a family of coherent states $\{\phi_{\omega,t}\}_{\omega,t}$. These are functions labeled by points $(\omega, t) \in \mathbb{R}^d \times \mathbb{R}^d$, and they are obtained by performing a frequency modulation ω and a time translation t on ϕ such that

$$\phi_{\omega,t}(x) = e^{2\pi i\omega\cdot x}\phi(x - t). \tag{2.4}$$

The STFT of f with respect to the window function ϕ at point (ω, t) is then defined as the inner product of f with $\phi_{\omega,t}$, that is

$$\langle f, \phi_{\omega,t}\rangle = \int_{\mathbb{R}^d} f(x)\overline{\phi_{\omega,t}(x)}\mathrm{d}x, \tag{2.5}$$

which we will at convenience denote by $V_\phi f(\omega, t)$.

Similarly to the Fourier transform, certain restriction are necessary to impose such that the above inner product is well-defined for all points (ω, t). These are restrictions on the window and will depend on the function space the signal belongs to. In particular, suppose that $f \in L^p(\mathbb{R}^d)$ for some fixed $p \in [1, \infty[$. Then by Hölder's inequality, a natural restriction is to only consider ϕ in the dual of $L^p(\mathbb{R}^d)$, namely $L^q(\mathbb{R}^d)$ where $1/p + 1/q = 1$. Hence, in our case when $f \in L^2(\mathbb{R}^d)$, we will always presume $\phi \in L^2(\mathbb{R}^d)$. However, before proceeding, notice what happens to the inner product when ϕ is chosen to be the constant function equal to 1 (which is obviously not square integrable). In this case the STFT reduces to the regular Fourier transform, that is $V_1 f(\omega, t) = \hat{f}(\omega)$.

For other, more non-trivial, choices for ϕ it is evident that the STFT maps a function f of *one* d-dimensional variable, e.g. time, to a function of *two* d-dimensional variables ω, t, e.g. frequency and time. The domain of the transformed function is thus $\mathbb{R}^d \times \mathbb{R}^d$ which we refer to as the *phase space* or with the current analogy of time and frequency, the *time-frequency plane*.

One advantage of restricting both signals and windows to $L^2(\mathbb{R}^d)$ is what Gröchenig[6] refers to as the orthogonality relation.

Theorem 2.1. (Theorem 3.2.1: Orthogonality relation for the STFT) Suppose $f_1, f_2, \phi_1, \phi_2 \in L^2(\mathbb{R}^d)$. Then $V_{\phi_j} f_j \in L^2(\mathbb{R}^{2d})$ for $j = 1, 2$ and

$$\begin{aligned} \langle V_{\phi_1} f_1, V_{\phi_2} f_2 \rangle_{L^2(\mathbb{R}^{2d})} &= \iint_{\mathbb{R}^{2d}} V_{\phi_1} f_1(\omega, t) \overline{V_{\phi_2} f_2(\omega, t)} \mathrm{d}\omega \mathrm{d}t \\ &= \langle f_1, f_2 \rangle \overline{\langle \phi_1, \phi_2 \rangle}. \end{aligned} \tag{2.6}$$

Thus, with the current assumptions, the target space of the STFT is in fact a subspace of $L^2(\mathbb{R}^{2d})$. Furthermore, when both domain and target space are equipped with the standard L^2-norms, the STFT becomes a bounded, linear map such that

$$\|V_\phi f\|_2 = \|\phi\|_2 \|f\|_2 \quad \forall\ f \in L^2(\mathbb{R}^d).$$

In particular, if the window function ϕ is normalized, i.e. $\|\phi\|_2 = 1$, then the STFT becomes an *isometry* from $L^2(\mathbb{R}^d)$ onto some subspace of $L^2(\mathbb{R}^{2d})$, that is

$$\|V_\phi f\|_2 = \|f\|_2 \quad \forall\ f \in L^2(\mathbb{R}^d).$$

Another consequence of the orthogonality relation is that the original signal f can be recovered from the STFT. Take any $\gamma \in L^2(\mathbb{R}^d)$ such that $\langle \gamma, \phi \rangle \neq 0$, then the orthogonal projection of f onto any $g \in L^2(\mathbb{R}^d)$ is given by

$$\langle f, g \rangle = \frac{1}{\langle \gamma, \phi \rangle} \iint_{\mathbb{R}^{2d}} V_\phi f(\omega, t) \overline{V_\gamma g(\omega, t)} \mathrm{d}\omega \mathrm{d}t. \tag{2.7}$$

A canonical choice for γ is to set it equal to ϕ. Furthermore, if we assume that ϕ is normalized, then these projections read

$$\begin{aligned} \langle f, g \rangle &= \iint_{\mathbb{R}^{2d}} V_\phi f(\omega, t) \overline{V_\phi g(\omega, t)} \mathrm{d}\omega \mathrm{d}t \\ &= \iint_{\mathbb{R}^{2d}} \langle f, \phi_{\omega,t} \rangle \overline{\langle g, \phi_{\omega,t} \rangle} \mathrm{d}\omega \mathrm{d}t. \end{aligned} \tag{2.8}$$

Since any signal $f \in L^2(\mathbb{R}^d)$ is completely determined by such inner products $\langle f, g \rangle$, the right-hand side of both formula (2.7) and (2.8) provide a complete recovery from the STFT.

2.2 Elements of Functional Analysis

This section is meant as a brief exposition to some fundamental concepts and results from functional analysis. In particular, we cover the definitions of operator norm and spectrum of linear operators, with primary focus on the spectrum of self-adjoint, compact operators on separable Hilbert spaces (main results Theorem 2.3 and Corollary 2.1). Proceeding, we consider the Hilbert-Schmidt integral operators as a family of compact operators, and we provide a simple criterion (in Proposition 2.2) for self-adjointness. Needless to say, these notions will prove relevant once we finally introduce the time-frequency localization operators.

To begin with, we recall the definition of the operator norm: Let X and Y be two Banach spaces endowed with the norms $\|\cdot\|_X$ and $\|\cdot\|_Y$, respectively. For a linear operator $\mathcal{T} : X \to Y$ the *operator norm* is given by

$$\|\mathcal{T}\|_{\text{op}} = \sup_{\|f\|_X \leq 1} \|\mathcal{T}f\|_Y. \tag{2.9}$$

For the most part, we will consider the case when $X = Y = L^2(\mathbb{R}^d)$ equipped with the standard L^2-norm, for which the operator norm becomes

$$\|\mathcal{T}\|_{\text{op}} = \sup_{\|f\|_2 \leq 1} \|\mathcal{T}f\|_2, \quad \text{where} \quad \|\mathcal{T}f\|_2 = \sup_{\|g\|_2 \leq 1} |\langle \mathcal{T}f, g\rangle|. \tag{2.10}$$

It is well-known that the operator $\mathcal{T}$ is continuous with respect to the two norms $\|\cdot\|_X, \|\cdot\|_Y$ if and only if $\mathcal{T}$ is a bounded operator, that is the operator norm is bounded.

Proceeding, we make a formal definition of the spectrum: Let X be a Banach space over $\mathbb{C}$ and $\mathcal{T} : X \to X$ a bounded, linear operator. The *spectrum* of $\mathcal{T}$ consists precisely of all scalars $\lambda \in \mathbb{C}$ such that

$$\mathcal{T} - \lambda I \tag{2.11}$$

is non-invertible, where I denotes the identity operator on B. Notice that if the kernel of $T - \lambda I$ is nontrivial, then λ is an eigenvalue of $\mathcal{T}$. We refer to the set of eigenvalues as the *point spectrum* of $\mathcal{T}$.

With the possible exception of $\lambda = 0$, the next theorem establishes that if $\mathcal{T}$ is assumed to be a compact operator, the spectrum coincides with the point spectrum.

Theorem 2.2. (Theorem 8.25: Fredholm's alternative[7])
Suppose $\mathcal{T}\colon X \to X$ is a compact, linear operator on Banach space X over $\mathbb{C}$. Then for any non-zero scalar $\lambda \in \mathbb{C}$ either

(i) $\mathcal{T} - \lambda I$ is invertible, or

(ii) λ is an eigenvalue of $\mathcal{T}$.

Thus, whenever referring to the spectrum of a compact operator, we will in principle be dealing with its eigenvalues. In the context of $L^2(\mathbb{R}^d)$, observe that this Banach space is a well-known separable Hilbert space. In the next theorem we present a central and useful result regarding the eigenvalues and eigenfunctions of self-adjoint, compact operators on such spaces.

Theorem 2.3. (Theorem 7.30[7]) Let $\mathcal{T} : \mathcal{H} \to \mathcal{H}$ be a self-adjoint, compact operator on the separable Hilbert space $\mathcal{H}$. Then there exists a countable orthonormal basis $\{e_j\}_j$ for $\mathcal{H}$ such that e_j is an eigenvector of $\mathcal{T}$ for each j, i.e. $\mathcal{T}e_j = \lambda_j e_j$ for some sequence of real-valued scalars $\{\lambda_j\}_j$.[1]

From this theorem we make a simple conclusion on the operator norm:

Corollary 2.1. Let $\mathcal{T} : \mathcal{H} \to \mathcal{H}$ be as in Theorem 2.3. Let the eigenvalues of $\mathcal{T}$, say $\{\lambda_j\}_j$, be ordered such that $|\lambda_0| \geq |\lambda_j| \;\; \forall\, j$. Then the operator norm of $\mathcal{T}$ is given by

$$\|\mathcal{T}\|_{\mathrm{op}} = |\lambda_0|. \tag{2.12}$$

Proof. Let $\{e_j\}_j$ denote the orthonormal basis of $\mathcal{H}$ such that $\mathcal{T}e_j = \lambda_j e_j$. Note that any elements f, g in $\mathcal{H}$ can then be expressed

$$f = \sum_j \langle f, e_j\rangle e_j \quad \text{and} \quad g = \sum_j \langle g, e_j\rangle e_j,$$

where

$$\sum_j |\langle f, e_j\rangle|^2 = \|f\|^2_{\mathcal{H}} \quad \text{and} \quad \sum_j |\langle g, e_j\rangle|^2 = \|g\|^2_{\mathcal{H}}.$$

[1] In the infinite dimensional case, the proof relies on Zorn's Lemma.

Thus, we attain

$$\begin{aligned}|\langle \mathcal{T}f, g\rangle| &\leq \sum_{m,n} |\langle f, e_m\rangle||\langle g, e_n\rangle||\langle \mathcal{T}e_m, e_n\rangle| \\ &= \sum_{m} |\lambda_m||\langle f, e_m\rangle||\langle g, e_m\rangle| \\ &\leq |\lambda_0| \sum_{m} |\langle f, e_m\rangle||\langle g, e_m\rangle| \\ &\leq |\lambda_0|\|f\|_{\mathcal{H}}\|g\|_{\mathcal{H}} \text{ (by Cauchy-Schwarz' inequality).}\end{aligned}$$

By identity (2.10), we conclude that $\|\mathcal{T}\|_{\text{op}} \leq |\lambda_0|$. Since $|\langle \mathcal{T}e_0, e_0\rangle| = |\lambda_0|$, the inequality is indeed sharp.

□

In what follows, we will focus on a particular family of self-adjoint, compact operators, namely the self-adjoint Hilbert-Schmidt integral operators. We begin by introducing the notion of an integral transform.

For a function $f \in L^2(\mathbb{R}^d)$ we define an *integral transform* $\mathcal{T}$ on f by

$$\mathcal{T}f(x) = \int_{\mathbb{R}^d} K(x,y)f(y)\mathrm{d}y, \tag{2.13}$$

where $\mathcal{T}f$ is a new function of variable $x \in \mathbb{R}^n$ (here n is possibly different from d). We refer to the function $K(x,y)$ as the *integral kernel* of the transform. Notice that the idea of an integral transform should be somewhat familiar as we have already been exposed to them in section 2.1:

Example 2.1. Both the Fourier transform and the STFT represent a type of integral transform. In particular,

(i) The Fourier transform is an integral transform to a function in $x \in \mathbb{R}^d$, with integral kernel

$$K(x,y) = e^{-2\pi i x\cdot y}.$$

(ii) The STFT is an integral transform to a function in $x = (\omega, t) \in \mathbb{R}^{2d}$, with integral kernel

$$K(\omega, t, y) = e^{-2\pi i \omega\cdot y}\overline{\phi(y-t)}.$$

Recall that in the discussion of the STFT, certain restrictions were made on the integral kernel to guarantee a well-defined transform. This illustrates that for an arbitrary integral kernel transform (2.13) is not necessarily well-defined. As we shall see, the aforementioned Hilbert-Schmidt operators all represent well-defined integral transforms.

A *Hilbert-Schmidt integral operator* (or simply *Hilbert-Schmidt operator*) is a linear map $\mathcal{T} : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ which can be expressed as a integral transform according to (2.13), with the integral kernel $K \in L^2(\mathbb{R}^d \times \mathbb{R}^d)$.

Firstly, we verify that any integral transform with such a kernel is a bounded operator that maps to $L^2(\mathbb{R}^d)$. By Hölder's inequality,

$$|\mathcal{T} f(x)| \leq \int_{\mathbb{R}^d} |K(x,y)f(y)|\mathrm{d}y \leq \left(\int_{\mathbb{R}^d} |K(x,y)|^2 \mathrm{d}y \right)^{1/2} \|f\|_2,$$

which is well-defined for almost all x. From here,

$$\begin{aligned} \|\mathcal{T} f\|_2 &= \left(\int_{\mathbb{R}^d} |\mathcal{T} f(x)|^2 \mathrm{d}x \right)^{1/2} \\ &\leq \left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |K(x,y)|^2 \mathrm{d}y \right) \mathrm{d}x \right)^{1/2} \|f\|_2 = \|K\|_2 \|f\|_2. \end{aligned}$$

Hence, the map $\mathcal{T}$ is a bounded operator with target $L^2(\mathbb{R}^d)$ and operator norm

$$\|\mathcal{T}\|_{\mathrm{op}} \leq \|K\|_2. \tag{2.14}$$

The next proposition reveals these operators to be compact.

Proposition 2.1. Any Hilbert-Schmidt operator $\mathcal{T} : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ is compact.

Proof. In Bowers and Kalton's Proposition 7.38[7] this is shown for the 1-dimensional case ($d = 1$) when the integral kernels have compact support on $[a,b] \times [a,b]$. By the exact same procedure, we may extend to the d-dimensional case, to integral kernels with compact support on $[a,b]^d \times [a,b]^d$. This can again be generalized to arbitrary integral kernels $K \in L^2(\mathbb{R}^d \times \mathbb{R}^d)$.

We have that K can be written as a limit of compactly supported integral kernels $\{K_n\}_n$ that converges in the L^2-norm. Let $\{\mathcal{T}_n\}_n$ denote the corresponding Hilbert-Schmidt operators. Observe that if $\|K_n - K\|_2 \to 0$, then $\|\mathcal{T}_n - \mathcal{T}\|_2 \to 0$. Since limits of compact operators are indeed compact, we are done.

□

In the next proposition we characterize self-adjointness of Hilbert-Schmidt operators in terms of the integral kernel.

Proposition 2.2. Let $\mathcal{T} : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ be a Hilbert-Schmidt operator. Then $\mathcal{T}$ is self-adjoint if and only if the integral kernel K satisfies

$$K(x, y) = \overline{K(y, x)} \text{ for almost all } x, y \in \mathbb{R}^d. \tag{2.15}$$

Proof. By Cauchy-Schwarz' inequality, it is clear that for any $f, g \in L^2(\mathbb{R}^d)$

$$\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |K(x, y)||f(y)|\mathrm{d}y \right) |g(x)|\mathrm{d}x \leq \|K\|_2 \|f\|_2 \|g\|_2 < \infty.$$

Thus, the Fubini-Tonelli theorem applies such that the integration order in $\langle \mathcal{T} f, g \rangle$ can be exchanged to the effect

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

By inspection, the adjoint $\mathcal{T}^*$ is a Hilbert-Schmidt operator of the form

$$\mathcal{T}^* f(x) = \int_{\mathbb{R}^d} \overline{K(y, x)} f(y)\mathrm{d}x, \tag{2.16}$$

which coincides with $\mathcal{T}$ if and only if the integral kernels coincide.

□

2.3 Introduction to Localization Operators

In this section we finally introduce the time-frequency localization operators. We will distinguish between two kinds of localization operators depending on whether they are based on a separate or joint time-frequency representation. Section 2.3.1 focuses on the first kind, i.e. localization operators based on a separate time-frequency description, while section 2.3.2 focuses on the second kind.

2.3.1 Projections in Time and Frequency

When attempting to localize a signal f and its Fourier transform $\hat{f}$, there are two natural orthogonal projections to consider. The first projection, say π_T for some measurable set $T \subseteq \mathbb{R}^d$, is given by

$$\pi_T f(t) = \chi_T(t) f(t), \tag{2.17}$$

where $\chi_T(\cdot)$ denotes the characteristic function which is one for arguments in T and zero otherwise. Hence, this projection aims at and indeed does localize f in time on the set T. The other projection, say Q_Ω, localizes the signal on the (measurable) frequency band $\Omega \subseteq \mathbb{R}^d$ and is given by

$$Q_\Omega f(t) = \mathcal{F}^{-1}\{\chi_\Omega \hat{f}\}(t) = \int_\Omega \hat{f}(\omega) e^{2\pi i \omega \cdot t} \mathrm{d}\omega. \tag{2.18}$$

Combining these two projections by composition into a single operator,

$$Q_\Omega \pi_T \text{ or } \pi_T Q_\Omega : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d), \tag{2.19}$$

is the first attempt to construct an operator which aims at localizing a signal in *both* time and frequency. Note that, since orthogonal projections are self-adjoint, the two operators in (2.19) must be adjoints of each other, that is

$$(Q_\Omega \pi_T)^* = \pi_T Q_\Omega. \tag{2.20}$$

By the construction, it is clear that these two operators are both continuous with operator norm bounded by 1, regardless of T and Ω. If we further assume both T, Ω have finite measure, then the above localization operators are in fact Hilbert-Schmidt integral operators.

Proposition 2.3. Let $T, \Omega \subseteq \mathbb{R}^d$ such that $|T|, |\Omega| < \infty$. Then the operators $Q_\Omega \pi_T$ and $\pi_T Q_\Omega$ are Hilbert-Schmidt integral operators of the form

$$Q_\Omega \pi_T f(x) = \int_{\mathbb{R}^d} K(x,t) f(t) \mathrm{d}t \tag{2.21}$$

and

$$\pi_T Q_\Omega f(x) = \int_{\mathbb{R}^d} \overline{K(t,x)} f(t) \mathrm{d}t, \tag{2.22}$$

where the integral kernel

$$K(x,t) = \chi_T(t) \int_\Omega e^{2\pi i (x-t)\cdot\omega} \mathrm{d}\omega \tag{2.23}$$

for which

$$\|K\|_2 = \left(\iint_{\mathbb{R}^{2d}} |K(x,t)|^2 \mathrm{d}x\mathrm{d}t \right)^{1/2} = \sqrt{|T||\Omega|}. \tag{2.24}$$

Proof. By definitions (2.17), (2.18), we have for any $f \in L^2(\mathbb{R}^d)$

$$\begin{aligned} Q_\Omega \pi_T f(x) &= \int_\Omega \left(\int_{\mathbb{R}^d} \chi_T(t) f(t) e^{-2\pi i \omega\cdot t} \mathrm{d}t \right) e^{2\pi i \omega \cdot x} \mathrm{d}\omega \\ &= \iint_{\mathbb{R}^{2d}} \chi_\Omega(\omega) \chi_T(t) f(t) e^{2\pi i (x-t)\cdot\omega} \mathrm{d}t\mathrm{d}\omega. \end{aligned}$$

Since both $|T|, |\Omega| < \infty$ by assumption, the integrand is easily verified to be in $L^1(\mathbb{R}^{2d})$. Hence, by the Fubini-Tonelli theorem, the integration order can be exchanged. The desired results (2.21), (2.23) follow after rearrangement. Result (2.22) is then evident by formula (2.16) for the adjoint of Hilbert-Schmidt operators combined with identity (2.20).

Finally, observe that

$$\int_\Omega e^{2\pi i (x-t)\cdot\omega} \mathrm{d}\omega = \mathcal{F}\{\chi_\Omega\}(t-x),$$

and since $\mathcal{F}$ is unitary, we obtain

$$\begin{aligned} \|K\|_2^2 &= \int_{\mathbb{R}^d} \chi_T(t) \int_{\mathbb{R}^d} |\mathcal{F}\{\chi_\Omega\}(t-x)|^2 \mathrm{d}x\mathrm{d}t \\ &= \|\mathcal{F}\{\chi_\Omega\}\|_2^2 \int_{\mathbb{R}^d} \chi_T(t)\mathrm{d}t = \|\chi_\Omega\|_2^2 \|\chi_T\|_2^2 = |\Omega||T|. \end{aligned}$$

□

Recall that by (2.1), any Hilbert-Schmidt operator is bounded by the norm of its integral kernel. Hence, by (2.24), we always have

$$\|Q_\Omega \pi_T\|_{\mathrm{op}}(= \|\pi_T Q_\Omega\|_{\mathrm{op}}) \leq \min\{\sqrt{|T||\Omega|}, 1\}. \tag{2.25}$$

The above estimate will prove particularly useful once we discuss the Donoho-Stark uncertainty principle in section 3.2.

Furthermore, by comparing the integral kernels of $Q_\Omega \pi_T$ and $\pi_T Q_\Omega$, it follows, by Proposition 2.2, that *neither* of these operators can be self-adjoint for sets $T, \Omega \subseteq \mathbb{R}^d$ of finite measure.[2] Nevertheless, by a simple $\mathcal{T}^*\mathcal{T}$-trick, we are able to construct self-adjoint, compact localization operators:

Consider the two compositions

$$(Q_\Omega \pi_T)^*(Q_\Omega \pi_T) = \pi_T Q_\Omega \pi_T \tag{2.26}$$

and

$$(\pi_T Q_\Omega)^*(\pi_T Q_\Omega) = Q_\Omega \pi_T Q_\Omega, \tag{2.27}$$

which are always self-adjoint. Since compositions of compact operators remain compact, we have that (2.26), (2.27) are self-adjoint, compact whenever $|T|, |\Omega| < \infty$. In this case, by Theorem 2.3, there exist an orthonormal basis $\{e_j\}_j$ for $L^2(\mathbb{R}^d)$ such that each e_j is an eigenfunction of $\pi_T Q_\Omega \pi_T$.[3]

Consider the subset of eigenfunctions $\{E_j\}_j \subseteq \{e_j\}_j$ whose associated eigenvalues are non-zero. It is easy to verify that any such eigenfunction must also be an eigenfunction of $\pi_T Q_\Omega$ and that these form an orthonormal basis for $L^2(T)$. Hence, the properties of $\pi_T Q_\Omega$ on $L^2(T)$ are essentially encoded in $\{E_j\}_j$ along with the associated eigenvalues.

Among the most natural choices for the time and frequency sets is when they take the form $T = [-M, M]$ and $\Omega = [-N, N]$ for some $M, N > 0$. For these particular choices of T and Ω, the eigenfunctions $\{E_j\}_j$ are more commonly referred to as the *prolate spheroidal wave functions*. In the 1960's these eigenfunctions were explicitly determined and extensively studied in a series of articles [1],[2],[3] by Landau, Pollak and Slepian.

[2] This could also be argued from the later presented Benedicks' Theorem (see section 3.4, Theorem 3.6).

[3] Similarly, we have such a set of eigenfunctions for the operator $Q_\Omega \pi_T Q_\Omega$.

2.3.2 Daubechies' Localization Operator

In this section we consider a different class of time-frequency localization operator, based on the joint representation produced by the STFT. This construction is motivated by the inner product (2.7), which shows how a time-dependent signal can be recovered from its phase space representation. In what follows, we will focus on the version stated in (2.8), where the window function ϕ is normalized.

When attempting to localize our time-dependent function f in both time and frequency, a natural approach is to modify the STFT of f before recovery by projections. Such a modification comes in the form of a multiplication by a *weight function*, say $F(\omega, t)$, with the intention of enhancing certain features of the phase space while diminishing others.

This process can be summarized as a sesquilinear functional $\mathscr{P}_{F,\phi}$ on the product $L^2(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$, defined by

$$\mathscr{P}_{F,\phi}(f, g) = \iint_{\mathbb{R}^{2d}} F(\omega, t)\langle f, \phi_{\omega,t}\rangle \overline{\langle g, \phi_{\omega,t}\rangle} \mathrm{d}\omega \mathrm{d}t. \tag{2.28}$$

Assuming $\mathscr{P}_{F,\phi}$ is a bounded functional, a duality argument[4] ensures the existence of a bounded, linear operator $P_{F,\phi} : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ such that

$$\mathscr{P}_{F,\phi}(f, g) = \langle P_{F,\phi} f, g\rangle. \tag{2.29}$$

The operator $P_{F,\phi}$ is our sought after time-frequency localization operator, which we will refer to as *Daubechies' localization operator* due to the fact that it was first introduced in her 1988-publication[4]. From the above definition we conclude that $P_{F,\phi}$ is characterized by the choice of weight F and window function ϕ. However, when the choice for window is either evident from the context or redundant in the discussion, we will usually drop the indexing ϕ and simply denote the operator by P_F.

In what follows, we will establish a few well-known but relevant properties of the localization operator under some reasonable restrictions on the weight function. To begin with, we consider two separate conditions on the weight function F such that the sesquilinear functional $\mathscr{P}_F$ is a bounded functional, for which the localization operator P_F is defined and continuous in the L^2-norm. Note that we always presume F to be a measurable function in the standard Lebesgue measure.

[4]This result follows from Riesz representation theorem for Hilbert spaces, see Theorem 7.16 (Riesz-Fréchet Theorem) in Bowers and Kalton[7].

Proposition 2.4. Let P_F denote Daubechies' localization operator with weight function F. Distinguish between the two cases:

(A) Suppose F is *bounded*, that is $\|F\|_\infty < \infty$, then $\|P_F\|_{\text{op}} \le \|F\|_\infty$.

(B) Suppose F is *integrable*, that is $\|F\|_1 < \infty$, then $\|P_F\|_{\text{op}} \le \|F\|_1$.

Proof. (A) By definition of P_F, we have for any $f, g \in L^2(\mathbb{R}^d)$

$$\begin{aligned}
|\langle P_F f, g\rangle| &= \Big| \iint_{\mathbb{R}^{2d}} F(\omega,t)\langle f, \phi_{\omega,t}\rangle \overline{\langle g, \phi_{\omega,t}\rangle} \mathrm{d}\omega \mathrm{d}t \Big| \\
&\le \|F\|_\infty \iint_{\mathbb{R}^{2d}} \big|\langle f, \phi_{\omega,t}\rangle \overline{\langle g, \phi_{\omega,t}\rangle}\big| \mathrm{d}\omega \mathrm{d}t \\
&\le \|F\|_\infty \|V_\phi f\|_2 \|V_\phi g\|_2 \text{ (by Cauchy-Schwarz).}
\end{aligned}$$

From the orthogonality relation in Theorem 2.1,

$$|\langle P_F f, g\rangle| \le \|F\|_\infty \|f\|_2 \|g\|_2.$$

Taking the supremum of all $\|f\|_2, \|g\|_2 \le 1$ produces the desired result.

(B) Once again by the definition of P_F,

$$\begin{aligned}
|\langle P_F f, g\rangle| &\le \iint_{\mathbb{R}^{2d}} \big|F(\omega,t)\langle f, \phi_{\omega,t}\rangle \overline{\langle g, \phi_{\omega,t}\rangle}\big| \mathrm{d}\omega \mathrm{d}t \\
&\le \|f\|_2 \|g\|_2 \|\phi\|_2^2 \iint_{\mathbb{R}^{2d}} \big|F(\omega,t)\big| \mathrm{d}\omega \mathrm{d}t \text{ (by Cauchy-Schwarz).}
\end{aligned}$$

The integral on the right-hand side is recognized as the L^1-norm of F, and since ϕ is assumed to be normalized, this concludes the proof.

□

Although both Proposition 2.4 (A) and (B) deal with continuity, observe that the assumptions on F in each case are fundamentally different. We can with ease construct unbounded integrable functions and conversely bounded functions which are not integrable. However, if both properties are present, the localization operator is evidently bounded by the minimum of the L^∞-norm and L^1-norm of F, that is

$$\|P_F\|_{\text{op}} \le \min\{\|F\|_\infty, \|F\|_1\}. \tag{2.30}$$

Assuming the weight is integrable, we show, similarly to the previous section, that the current localization operator becomes a Hilbert-Schmidt integral operator.

Proposition 2.5. Suppose $F \in L^1(\mathbb{R}^{2d})$. Then the associated Daubechies' localization operator P_F is a Hilbert-Schmidt operator with integral kernel

$$K_F(x,y) = \iint_{\mathbb{R}^{2d}} F(\omega,t)\phi_{\omega,t}(x)\overline{\phi_{\omega,t}(y)}\mathrm{d}\omega\mathrm{d}t. \tag{2.31}$$

Proof. Since F is integrable, it is evident by Cauchy-Schwarz' inequality that

$$\iint_{\mathbb{R}^{2d}} |F(\omega,t)| \left(\int_{\mathbb{R}^d} |f(y)\overline{\phi_{\omega,t}(y)}|\mathrm{d}y \int_{\mathbb{R}^d} |\phi_{\omega,t}(x)\overline{g(x)}|\mathrm{d}x \right) \mathrm{d}\omega\mathrm{d}t < \infty.$$

Hence, the Fubini-Tonelli theorem applies such that the integration order in $\langle P_F f, g\rangle$ can be exchanged to obtain

$$\begin{aligned}\langle P_F f, g\rangle &= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} K_F(x,y) f(y)\mathrm{d}y \right) \overline{g(x)}\mathrm{d}x \\ &= \langle \int_{\mathbb{R}^d} K_F(\cdot,y) f(y)\mathrm{d}y, g\rangle,\end{aligned}$$

where $K_F(x,y)$ coincides with (2.31). Since the above identity holds for all $g \in L^2(\mathbb{R}^d)$, we conclude that

$$P_F f(x) = \int_{\mathbb{R}^d} K_F(x,y) f(y)\mathrm{d}y \;\text{ for almost all } x \in \mathbb{R}^d.$$

It remains to show that $K_F \in L^2(\mathbb{R}^d \times \mathbb{R}^d)$. By Cauchy-Schwarz' inequality,

$$\begin{aligned}&\iint_{\mathbb{R}^{2d}} |F(\omega,t)| \Bigg[\iint_{\mathbb{R}^{2d}} |F(\omega',t')| \\ &\quad \cdot \left(\int_{\mathbb{R}^d} |\phi_{\omega,t}(x)\overline{\phi_{\omega',t'}(x)}|\mathrm{d}x \int_{\mathbb{R}^d} |\phi_{\omega',t'}(y)\overline{\phi_{\omega,t}(y)}|\mathrm{d}y \right) \mathrm{d}\omega'\mathrm{d}t' \Bigg] \mathrm{d}\omega\mathrm{d}t \\ &\quad \leq \|F\|_1^2 \|\phi\|_2^4 = \|F\|_1^2.\end{aligned}$$

Therefore, we may apply the Fubini-Tonelli theorem once more, now in the expression for $K_F(x,y)$, to yield $\|K_F\|_2 \leq \|F\|_1$.

□

A simple additional condition on F makes P_F self-adjoint.

Proposition 2.6. If F is a real-valued function in $L^1(\mathbb{R}^{2d})$, then P_F is a self-adjoint, compact operator.

Proof. By the previous proposition, we have that P_F is a compact operator, in particular, a Hilbert-Schmidt integral operator with integral kernel $K_F(x, y)$ according to (2.31). By criterion (2.15), we only require

$$K_F(x, y) = \overline{K_F(y, x)} \text{ for almost all } x \in \mathbb{R}^d$$

in order for P_F to be self-adjoint. Since F is real-valued, the latest identity follows readily.

$\square$

Observe that whenever the weight F is integrable, P_F becomes a *trace class operator*. More precisely, the localization operator P_F is a compact operator whose *trace*

$$\sum_j \langle P_F e_j, e_j \rangle$$

is always well-defined and independent of the choice of orthonormal basis $\{e_j\}_j$ for $L^2(\mathbb{R}^d)$.

Proposition 2.7. Suppose $F \in L^1(\mathbb{R}^{2d})$. Then the associated Daubechies' localization operator P_F is a trace class operator such that

$$\sum_j |\langle P_F e_j, e_j \rangle| \leq \|F\|_1 \tag{2.32}$$

and

$$\sum_j \langle P_F e_j, e_j \rangle = \iint_{\mathbb{R}^{2d}} F(\omega, t) \mathrm{d}\omega \mathrm{d}t, \tag{2.33}$$

for any orthonormal basis $\{e_j\}_j$ for $L^2(\mathbb{R}^d)$.

Proof. Let $\{e_j\}_j$ be any orthonormal basis for $L^2(\mathbb{R}^d)$. Then by the Monotone Convergence theorem,

$$\begin{aligned}
\sum_j |\langle P_F e_j, e_j \rangle| &\leq \sum_j \iint_{\mathbb{R}^{2d}} |F(\omega, t)| |\langle \phi_{\omega,t}, e_j \rangle|^2 \mathrm{d}\omega \mathrm{d}t \\
&= \iint_{\mathbb{R}^{2d}} |F(\omega, t)| \sum_j |\langle \phi_{\omega,t}, e_j \rangle|^2 \mathrm{d}\omega \mathrm{d}t \\
&= \iint_{\mathbb{R}^{2d}} |F(\omega, t)| \|\phi\|_2^2 \mathrm{d}\omega \mathrm{d}t \text{ (by Parseval's identity)} \\
&= \iint_{\mathbb{R}^{2d}} |F(\omega, t)| \mathrm{d}\omega \mathrm{d}t = \|F\|_1.
\end{aligned}$$

Since any partial sum $F(\omega,t)\sum_j |\langle \phi_{\omega,t}, e_j\rangle|^2$ is uniformly bounded in absolute value by $|F(\omega,t)|$, the summation and integration can be exchanged by Lebesque's Dominated Convergence theorem such that

$$\begin{aligned}\sum_j \langle P_F e_j, e_j\rangle &= \sum_j \iint_{\mathbb{R}^{2d}} F(\omega,t)|\langle \phi_{\omega,t}, e_j\rangle|^2 \mathrm{d}\omega \mathrm{d}t \\ &= \iint_{\mathbb{R}^{2d}} F(\omega,t) \sum_j |\langle \phi_{\omega,t}, e_j\rangle|^2 \mathrm{d}\omega \mathrm{d}t,\end{aligned}$$

which once again by Parseval's identity produces the desired result.

□

From these two latest propositions, we summarize the consequences on the spectrum in the subsequent corollary.

Corollary 2.2. Suppose F is a real-valued function in $L^1(\mathbb{R}^{2d})$, and let P_F denote the corresponding Daubechies' localization operator. Then there exists an orthonormal basis $\{e_j\}_j$ for $L^2(\mathbb{R}^d)$ such that for each j we have $P_F e_j = \lambda_j e_j$ for some sequence of real-valued scalars $\{\lambda_j\}_j$. This sequence of eigenvalues coincides with the point spectrum of P_F, which again coincides with the entire spectrum of P_F.

Furthermore, the sum of the eigenvalues of P_F is finite such that

$$\sum_j |\lambda_j| \le \|F\|_1 \tag{2.34}$$

and

$$\sum_j \lambda_j = \iint_{\mathbb{R}^{2d}} F(\omega,t)\mathrm{d}\omega \mathrm{d}t. \tag{2.35}$$

Proof. The first part of the corollary is a direct restatement of Theorem 2.2 and Theorem 2.3 with respect to Proposition 2.6. The second part regarding the sum is a special case of identities (2.32), (2.33) since $\langle P_F e_j, e_j\rangle = \lambda_j$ for each j.

□

3 Uncertainty Principles

The purpose of this chapter is to provide a brief survey of different aspects of the uncertainty principles in Fourier analysis. Further, we will formulate these principles in terms of the two types of localization operators introduced in section 2.3.1 and 2.3.2. In regular Fourier analysis the uncertainty principles all convey the idea that a signal and its Fourier transform cannot be well-localized simultaneously. With the time-frequency analogy presented in Chapter 2.1, a signal may not be concentrated in both time and frequency. Extending to simultaneous time-frequency representations, these principles find their analog. Hence, the uncertainty principles pose a fundamental obstacle when attempting to localize signal and its Fourier transform, whether it be for separate representations or simultaneous ones.

We will start by motivating in section 3.1 with perhaps the most recognized version of the uncertainty principles, namely Heisenberg's uncertainty principle. Much of this recognition can be argued from its frequent appearance in quantum mechanics and its direct implications for measurements of physical observables. Afterwards, we consider the classical Donoho-Stark uncertainty principle for the regular Fourier transform in section 3.2, before establishing the analog Lieb's uncertainty principle in section 3.3 for the STFT. In section 3.4 we present Benedicks' Theorem for the regular Fourier transform and Janssen's extension for the STFT. In the final section we briefly discuss some more recent results as a motivation for further research.

3.1 Heisenberg's Uncertainty Principle

Although Heisenberg's uncertainty principle can be generalized to a statement about self-adjoint operators on the Hilbert space, we will only consider the principle for a signal f and its Fourier transform $\hat{f}$ (both in the space $L^2(\mathbb{R}^d)$).

Theorem 3.1. (Heisenberg's Uncertainty Principle) Let $f \in L^2(\mathbb{R}^d)$, and let $a, b \in \mathbb{R}^d$ be arbitrary. Then

$$\int_{\mathbb{R}^d} |t-a|^2 |f(t)|^2 \mathrm{d}r \cdot \int_{\mathbb{R}^d} |\omega - b|^2 |\hat{f}(\omega)|^2 \mathrm{d}r \geq \frac{d^2 \|f\|_2^4}{16\pi^2}, \tag{3.1}$$

where $|t-a|^2 = \sum_{j=1}^d (t_j - a_j)^2$.

Since the function norm is invariant under any translation of the argument, the above theorem holds if and only if

$$\int_{\mathbb{R}^d} |t|^2 |f(t)|^2 \mathrm{d}r \cdot \int_{\mathbb{R}^d} |\omega|^2 |\hat{f}(\omega)|^2 \mathrm{d}r \geq \frac{d^2 \|f\|_2^4}{16\pi^2} \quad \forall\, f \in L^2(\mathbb{R}^d), \tag{3.2}$$

i.e. we may, without loss of generality, set a, b equal to zero.

In what follows, we provide a simple proof of Theorem 3.1 in the 1-dimensional case, which can easily be generalized to d dimensions. The proof, that we will consider, originates with Nicolaas G. de Bruijn in his 1967-publication[8] and involves the *Hermite functions*

$$H_k(t) = \frac{2^{1/4}}{\sqrt{k!}} \left(-\frac{1}{2\sqrt{\pi}} \right)^k e^{\pi t^2} \frac{\mathrm{d}^k}{\mathrm{d}t^k}(e^{-2\pi t^2}), \ k = 0, 1, 2, \dots, \tag{3.3}$$

which is a well-known orthonormal basis for $L^2(\mathbb{R})$ (see Folland's Chapter 1.7 point (vii)[9]). As an interesting side-note, the Hermite functions will also be essential in Chapter 4 when we discuss Daubechies' localization operators with a spherically symmetric weight.

In the current context, from Folland's Chapter 1.7[9], we obtain the following facts:

(i) If we set $H_{-1} \equiv 0$, we have the recursive relation

$$2\sqrt{\pi}t \cdot H_k(t) = \sqrt{k+1}H_{k+1}(t) + \sqrt{k}H_{k-1} \text{ for } k = 0, 1, 2, \dots, \tag{3.4}$$

(ii) Every H_k is an eigenfunction of the Fourier transform such that

$$\mathcal{F}H_k = (-i)^k H_k \text{ for } k = 0, 1, 2, \dots \tag{3.5}$$

Based on these two properties, we formulate the subsequent theorem.

Theorem 3.2. Let $f \in L^2(\mathbb{R})$. Then

$$\int_{\mathbb{R}} t^2 |f(t)|^2 \mathrm{d}t + \int_{\mathbb{R}} \omega^2 |\hat{f}(\omega)|^2 \mathrm{d}\omega = \frac{1}{2\pi} \sum_{k=0}^{\infty} (2k+1) |\langle f, H_k \rangle|^2. \tag{3.6}$$

In particular,

$$\int_{\mathbb{R}} t^2 |f(t)|^2 \mathrm{d}t + \int_{\mathbb{R}} \omega^2 |\hat{f}(\omega)|^2 \mathrm{d}\omega \geq \frac{\|f\|_2^2}{2\pi}, \tag{3.7}$$

where equality is realized precisely when f is a multiple of H_0.

Proof. Firstly, by the recursive relation (3.4),

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

Similarly, by the eigenvalue-equation (3.5) and the fact that $\mathcal{F}$ is unitary, we have

$$2\sqrt{\pi}\langle \omega\hat{f}(\omega), H_k(\omega)\rangle = i^{-(k+1)}\sqrt{k+1}\langle f, H_{k+1}\rangle + i^{-k+1}\sqrt{k}\langle f, H_{k-1}\rangle.$$

Now, apply Parseval's identity to these two latest formulas such that

$$\begin{aligned}\int_{\mathbb{R}} t^2|f(t)|^2\mathrm{d}t + \int_{\mathbb{R}} \omega^2|\hat{f}(\omega)|^2\mathrm{d}\omega &= \sum_{k=0}^{\infty}\left[|\langle tf(t), H_k(t)\rangle|^2 + |\langle \omega^2\hat{f}(\omega), H_k(\omega)\rangle|^2\right] \\ &= \frac{1}{4\pi}\sum_{k=0}^{\infty}\left[2(k+1)|\langle f, H_{k+1}\rangle|^2 + 2k|\langle f, H_{k-1}\rangle|^2\right] \\ &= \frac{1}{2\pi}\sum_{k=0}^{\infty}(2k+1)|\langle f, H_k\rangle|^2.\end{aligned}$$

As $2k+1 \geq 1$ for any positive k, the final inequality (3.7) is immediate. Equality holds if and only if the terms $(2k+1)|\langle f, H_k\rangle|^2$ are all zero for $k > 0$, which means f must be a multiple of H_0.

□

From here the 1-dimensional version of Theorem 3.1 follows by a simple dilation argument. Consider the dilation

$$g(t) = p^{-1/2}f(t/p) \text{ for any } p > 0. \tag{3.8}$$

Since $\|g\|_2 = \|f\|_2$, by Theorem 3.2, we must have

$$\begin{aligned}\frac{\|f\|_2^2}{2\pi} &\leq \int_{\mathbb{R}} t^2|g(t)|^2\mathrm{d}t + \int_{\mathbb{R}} \omega^2|\hat{g}(\omega)|^2\mathrm{d}\omega \\ &= p^2\int_{\mathbb{R}} t^2|f(t)|^2\mathrm{d}t + p^{-2}\int_{\mathbb{R}} \omega^2|\hat{f}(\omega)|^2\mathrm{d}\omega.\end{aligned} \tag{3.9}$$

Minimizing the right-hand side of (3.9) with respect to $p > 0$ produces the desired 1-dimensional Heisenberg's uncertainty principle

$$\int_{\mathbb{R}} t^2|f(t)|^2\mathrm{d}t \cdot \int_{\mathbb{R}} \omega^2|\hat{f}(\omega)|^2\mathrm{d}\omega \geq \frac{\|f\|_2^4}{16\pi^2}. \tag{3.10}$$

From Theorem 3.2, it is clear that equality of (3.10) is realized whenever f is a multiple of H_0, i.e. a multiple of the normalized Gaussian. Furthermore, a simple calculation reveals that equality holds for any dilation of the form

$$f(x) = cH_0(x/p), \text{ where } c \in \mathbb{C} \text{ and } p > 0. \tag{3.11}$$

Since any other dilation would keep inequality (3.9) strict for all $p > 0$, the functions in (3.11) are in fact the only solutions that minimize (3.10).

For the d-dimensional version of the uncertainty principle, we consider the d-dimensional Hermite functions, say

$$\eta_k(t) = \prod_{j=1}^{d} H_{k_j}(t_j), \text{ for } k = (k_1, \ldots, k_d) \in (\mathbb{N} \cup 0)^d, \tag{3.12}$$

which serves as an orthonormal basis for $L^2(\mathbb{R}^d)$. Then by Theorem 3.2 and Parseval's identity, it follows that for each $j = 1, \ldots, d$

$$\int_{\mathbb{R}^d} t_j^2 |f(t)|^2 \mathrm{d}t + \int_{\mathbb{R}^d} \omega_j^2 |\hat{f}(\omega)|^2 \mathrm{d}\omega = \frac{1}{2\pi} \sum_{k \in (\mathbb{N} \cup \{0\})^d} (2k_j + 1) |\langle f, \eta_k \rangle|^2.$$

Since the above identity is bounded from below by $\|f\|_2^2/(2\pi)$ for each j, we may apply a similar dilation argument as in the 1-dimensional case to conclude

$$\int_{\mathbb{R}^d} t_j^2 |f(t)|^2 \mathrm{d}t \cdot \int_{\mathbb{R}^d} \omega_j^2 |\hat{f}(\omega)|^2 \mathrm{d}\omega \geq \frac{\|f\|_2^4}{16\pi^2} \text{ for } j = 1, \ldots, d.$$

Based on this result and by Cauchy-Schwarz' inequality for Euclidean vectors in $\mathbb{R}^d$, we finally obtain the d-dimensional Heisenberg's uncertainty principle

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

A careful analysis of the above calculation reveals that the two *inequalities* are both reduced to *equalities* precisely when the signal f can be written as

$$f(t) = c\prod_{j=1}^{d} H_0(t_j/p) \;\; \forall\, c \in \mathbb{C} \text{ and } p > 0.$$

Proceeding, we will express inequality (3.1) on its more recognized form. Based on said inequality, it is natural to introduce the following concept: The *dispersion* of a signal $f \in L^2(\mathbb{R}^d)$ about point $a \in \mathbb{R}^d$ is defined as

$$(\Delta_a f)^2 = \frac{1}{\|f\|_2^2}\int_{\mathbb{R}^d} |t-a|^2|f(t)|^2 \mathrm{d}t. \tag{3.13}$$

Note that the dispersion does not have to be finite. Nevertheless, this quantity measures to what extent the graph of the signal deviates from the point $t = a$. If $(\Delta_a f)^2$ is small, then f is concentrated near a. While for a larger dispersion, the signal is more spread out. In the case when the dispersion is finite, it is straightforward to verify that the minimal dispersion occurs at $t = \bar{a}$, where

$$\bar{a} = \frac{1}{\|f\|_2^2}\int_{\mathbb{R}^d} t|f(t)|^2 \mathrm{d}t. \tag{3.14}$$

In terms of probability distributions, observe that for any normalized signal, i.e. $\|f\|_2 = 1$, the function $|f|^2$ can be regarded as a probability density function over $\mathbb{R}^d$. Then $\bar{a}$ represents the *expectation value* of $|f|^2$, and the associated dispersion $(\Delta_{\bar{a}} f)^2$ corresponds to the *variance* (or equivalently, $\Delta_{\bar{a}} f$ corresponds to the standard deviation).

With these notions, for any non-trivial signal $f \in L^2(\mathbb{R}^d)$, Heisenberg's uncertainty principle in Theorem 3.1 reads

$$\Delta_a f \cdot \Delta_b \hat{f} \geq \frac{d}{4\pi}. \tag{3.15}$$

In particular, if f is normalized, the above inequality conveys that the product of the standard deviation of f and that of $\hat{f}$ is greater than the threshold $d/(4\pi)$. Since the standard deviation is the (most) common metric by which we quantify the uncertainty or spread in our measurements, the appeal of Heisenberg's principle in applications is not unfounded.

We conclude this section with a motivational example of one of the most infamous consequences of the uncertainty principle:

Example 3.1. (Position and Momentum in Quantum Mechanics) Note that this example is not meant as a comprehensive introduction to the fundamentals of quantum mechanics (see instead Chapter 1 in Griffiths's book [10]).

Nevertheless, in quantum mechanics the position $q \in \mathbb{R}^d$ of a particle is encoded in a wave function, say Ψ. For a bound state, Ψ is a normalized element of $L^2(\mathbb{R}^d)$, where $|\Psi|^2$ represents the probability density function of the position of said particle. Similarly, the there exists a wave function $\Phi \in L^2(\mathbb{R}^d)$ for the momentum coordinates $p \in \mathbb{R}^d$, where $|\Phi|^2$ is the probability density function of the momentum.

Let $\Delta q, \Delta p$ denote the standard deviations of q, p, respectively. By the interpretation of the wave functions in terms of probability distributions, it is evident that $\Delta q, \Delta p$ must coincide with the corresponding standard deviations of Ψ, Φ. As it turns out, the momentum representation Φ is the Fourier transform of the position representation Ψ. Observe, however, that the Fourier transform between Ψ and Φ is normalized somewhat differently than in (2.1). In particular, this Fourier transform includes a non-zero physical constant h, also known as Planck's constant, that determines the physical scale of the position and momentum. In terms of (2.1), we obtain

$$\Phi(p) = h^{-d/2} \cdot \hat{\Psi}(p/h), \tag{3.16}$$

such that $\Delta p = h \cdot \Delta\hat{\Psi}$. By Heisenberg's uncertainty principle (3.15),

$$\Delta q \cdot \Delta p \geq h \frac{d}{4\pi}. \tag{3.17}$$

Hence, we have the remarkable result that the *position and momentum cannot be determined precisely simultaneously.* This example shows that the uncertainty principle is not merely of mathematical or theoretical interest, but also that the principle manifests itself directly in nature.

3.2 Donoho-Stark's Uncertainty Principle

In the previous section we introduced the notion of dispersion (and in return the standard deviation) as a measurement of the spread of the signal graph. Donoho-Stark's uncertainty principle, however, is formulated in terms of the measure of specific time set T and frequency band Ω. In order to evaluate to what extent a signal is concentrated on a specific set, we invoke the following definition:

For a fixed $\epsilon \in [0, 1]$ we say that a signal $f \in L^2(\mathbb{R}^d)$ is *at most ϵ-supported* outside a subset $E \subseteq \mathbb{R}^d$ if

$$\|f - \chi_E f\|_2 \leq \epsilon \|f\|_2. \tag{3.18}$$

Alternatively, the function f is said to be *at least* $(1 - \epsilon)$*-supported* on E.

In terms of the two projections π_T, Q_Ω from (2.17), (2.18), f is at least $(1 - \epsilon_T)$-supported on T and $\hat{f}$ is at least $(1 - \epsilon_\Omega)$-supported on Ω precisely when

$$\|f - \pi_T f\|_2 \leq \epsilon_T \|f\|_2,$$

and

$$\|f - Q_\Omega f\|_2 \leq \epsilon_\Omega \|f\|_2,$$

respectively[5]. With this new terminology, we are ready to present Donoho-Stark's uncertainty principle, attributed to David L. Donoho and Philip B. Stark for their findings in the 1989-paper[11].

Theorem 3.3. (Donoho-Stark's Uncertainty Principle)
Suppose $0 \neq f \in L^2(\mathbb{R}^d)$ is at least $(1 - \epsilon_T)$-supported on $T \subseteq \mathbb{R}^d$, and suppose $\hat{f}$ is at least $(1 - \epsilon_\Omega)$-supported on $\Omega \subseteq \mathbb{R}^d$. Then

$$|T||\Omega| \geq (1 - \epsilon_T - \epsilon_\Omega)^2. \tag{3.19}$$

Proof. Recall that by inequality (2.25), we already have a lower bound estimate for $|T||\Omega|$, namely the operator norm (squared) $\|Q_\Omega \pi_T\|_{\text{op}}^2$. Hence, it is sufficient to show that $\|Q_\Omega \pi_T f\|_2$ is bounded from below by $(1 - \epsilon_T - \epsilon_\Omega)\|f\|_2$. Observe, by the reverse triangle inequality,

$$\|Q_\Omega \pi_T f\|_2 \geq \|f\|_2 - \|f - Q_\Omega \pi_T f\|_2. \tag{3.20}$$

[5] The latest inequality follows from $\mathcal{F}$ being a unitary operator.

Furthermore, by the triangle inequality and the fact that $\|Q_\Omega\|_{\text{op}} \leq 1$,

$$\begin{aligned}\|f - Q_\Omega \pi_T f\|_2 &= \|f - Q_\Omega f + Q_\Omega(f - \pi_T f)\|_2 \\ &\leq \|f - Q_\Omega f\|_2 + \|Q_\Omega\|_{\text{op}} \|f - \pi_T f\|_2 \\ &\leq (\epsilon_\Omega + \epsilon_T)\|f\|_2,\end{aligned}$$

which combined with observation (3.20) yields the desired result.

□

As noted by Donoho and Stark, the key ingredient in the proof is the operator norm of $Q_\Omega \pi_T$ (or similarly $\pi_T Q_\Omega$). Inequality (3.19) thus poses a limitation for how efficiently the aforementioned operator can concentrate a signal on the set T in time and Ω in frequency.

3.3 Lieb's Uncertainty Principle

In this section we return focus to the STFT as our joint time-frequency representation, and consider the analog of Donoho-Stark's uncertainty principle for the STFT, namely Lieb's uncertainty principle. For this signal representation, the natural localization operators will be Daubechies' time-frequency localization operators. In particular, we will study the localization operators when the weight equals the characteristic function of some subset of the time-frequency plane. More precisely, let $U \subseteq \mathbb{R}^{2d}$ be measurable and $\phi \in L^2(\mathbb{R}^d)$ be a window function, then the localization operator $P_{U,\phi}$ is defined by

$$\langle P_{U,\phi} f, g\rangle = \iint_U \langle f, \phi_{\omega,t}\rangle \langle \phi_{\omega,t}, g\rangle \mathrm{d}\omega \mathrm{d}t \quad \forall\ f, g \in L^2(\mathbb{R}^d). \tag{3.21}$$

However, as opposed to Donoho-Stark, the proof of Lieb's uncertainty principle does not rely directly on the localization operators to be introduced. Nevertheless, these operators will indeed be impacted by the uncertainty principle, similar to that of the previous section.

Initially, consider a weaker version of the uncertainty principle:

Proposition 3.1. Suppose $f, \phi \in L^2(\mathbb{R}^d)$ are non-zero. If $U \subseteq \mathbb{R}^{2d}$ is measurable and $\epsilon \in [0, 1]$ is such that

$$\iint_U |V_\phi f(\omega, t)|^2 \mathrm{d}t \mathrm{d}\omega \geq (1 - \epsilon)\|V_\phi f\|_2^2,$$

then

$$|U| \geq 1 - \epsilon. \tag{3.22}$$

Proof. By Cauchy-Schwarz' inequality,

$$|V_\phi f(\omega,t)| = |\langle f, \phi_{\omega,t}\rangle| \leq \|f\|_2\|\phi\|_2 = \|V_\phi f\|_2 \quad \forall\ (\omega,t)\in\mathbb{R}^{2d}.$$

Thus, by monotonicity of the integral,

$$(1-\epsilon)\|V_\phi f\|_2^2 \leq \iint_U |V_\phi f(\omega,t)|^2 \mathrm{d}t\mathrm{d}\omega \leq |U|\|V_\phi f\|_\infty^2 \leq |U|\|V_\phi f\|_2^2.$$

□

A sharper estimate of the above inequality is attributed to Elliot H. Lieb for the discoveries in his 1989-paper[12]. Before presenting this, we require the following estimate:

Lemma 3.4. (Lieb's Inequality) Suppose $f,\phi\in L^2(\mathbb{R}^d)$ and $p\in[2,\infty[$, then

$$\iint_{\mathbb{R}^{2d}} |V_\phi f(\omega,t)|^p \mathrm{d}t\mathrm{d}\omega \leq \left(\frac{2}{p}\right)^d (\|f\|_2\|\phi\|_2)^p. \tag{3.23}$$

Proof. See proof of Theorem 1 (a) in Lieb's paper[12]. Alternatively, see proof of Theorem 3.3.2 in Gröchenig[6].

□

From here it becomes easy to prove Lieb's uncertainty principle.

Theorem 3.5. (Lieb's Uncertainty Principle) Suppose $f,\phi\in L^2(\mathbb{R}^d)$ are non-zero. If $U\subseteq\mathbb{R}^{2d}$ is measurable and $\epsilon\in[0,1]$ is such that

$$\iint_U |V_\phi f(\omega,t)|^2 \mathrm{d}t\mathrm{d}\omega \geq (1-\epsilon)\|V_\phi f\|_2^2,$$

then

$$|U| \geq (1-\epsilon)^{\frac{p}{p-2}} \left(\frac{p}{2}\right)^{\frac{2d}{p-2}} \quad \forall\ p>2. \tag{3.24}$$

Proof. By Hölder's inequality with exponents $p/2$ and $p/(p-2)$ for $p>2$,

$$\begin{aligned}
(1-\epsilon)\|V_\phi f\|_2^2 &\leq \iint_{\mathbb{R}^{2d}} \chi_U(\omega,t)|V_\phi f(\omega,t)|^2 \mathrm{d}t\mathrm{d}\omega \\
&\leq \left(\iint_{\mathbb{R}^{2d}} |V_\phi f(\omega,t)|^{2\cdot\frac{p}{2}} \mathrm{d}t\mathrm{d}\omega\right)^{\frac{p}{2}} \left(\iint_{\mathbb{R}^{2d}} \chi_U(\omega,t)^{\frac{p}{p-2}} \mathrm{d}t\mathrm{d}\omega\right)^{\frac{p-2}{p}} \\
&= \left(\iint_{\mathbb{R}^{2d}} |V_\phi f(\omega,t)|^p \mathrm{d}t\mathrm{d}\omega\right)^{\frac{2}{p}} |U|^{\frac{p-2}{p}} \\
&\leq \left(\frac{2}{p}\right)^{\frac{2d}{p}} \|V_\phi f\|_2^2 |U|^{\frac{p-2}{p}} \text{ (by Lemma 3.4).}
\end{aligned}$$

The final inequality is obtained by rearrangement. □

From the uncertainty principle presented in Proposition 3.1 and improved in Theorem 3.5, it is evident that no signal can be mostly concentrated in an arbitrarily small part of the phase space. As the measure of the region in question tends to zero, so must the support of the signal in that region. Note that the localization operator $P_{U,\phi}$ defined by (3.21) attempts to concentrate signals on the set U in phase space. Hence, the effective localization of $P_{U,\phi}$ on U will indeed be limited by the uncertainty principle.

We also make a few remarks of how the operator norm of $P_{U,\phi}$ is affected by the measure of U. By the continuity conditions on $P_{U,\phi}$ we always require the norm to be bounded from above by

$$\|P_{U,\phi}\|_{\mathrm{op}} \leq \min\{|U|, 1\}. \tag{3.25}$$

In addition, recall that the operator norm satisfies

$$\|P_{U,\phi}\|_{\mathrm{op}} \geq \langle P_{U,\phi} f, f \rangle = \iint_U |V_\phi f(\omega, t)|^2 \mathrm{d}t\mathrm{d}\omega \quad \forall \ \|f\|_2 = 1. \tag{3.26}$$

Hence, if for some non-zero $f \in L^2(\mathbb{R})$ the conditions of Lieb's principle are satisfied for a fixed $\epsilon \in [0, 1]$ on subset U, then we attain a lower bound for the norm

$$\|P_{U,\phi}\|_{\mathrm{op}} \geq (1 - \epsilon). \tag{3.27}$$

3.4 Benedicks' Theorem

So far we have considered uncertainty principles for sets of small measure in either the separate or joint time-frequency representation. From these principles we can deduce lower bounds for the measure of the support of the signal $f \neq 0$ in either representation. In particular, set $\epsilon = 0$ in Theorem 3.3 and 3.5. Then for the regular Fourier transform let $\mathrm{supp} f = T$ and $\mathrm{supp} \hat{f} = \Omega$, from which we must have

$$|T||\Omega| \geq 1. \tag{3.28}$$

Similarly, for the STFT with window $\phi \neq 0$ and $U = \mathrm{supp} V_\phi f$, we must have

$$|U| \geq \lim_{p \to 2^+} \left(\frac{p}{2}\right)^{\frac{2d}{p-2}} = e^d. \tag{3.29}$$

A natural question to ask is whether these inequalities are sharp. And if not, a followup question would be if a signal can be entirely concentrated on some finite time and frequency sets or concentrated on a finite part of the phase space. As it turns out, neither of these propositions are true.

Already in a preprint from 1974 Michael Benedicks had shown that no non-trivial function may be concentrated on a finite time and frequency band (see [13]). In 1997 it was conjectured by G.B. Folland and A. Sitaram in [14] that a similar proposition should hold for joint time-frequency representations (more precisely the Wigner distribution[6]). The conjecture was proven the following year by A.J.E.M. Janssen in [15].

Proceeding, we will present and prove both of these uncertainty principles. Initially, we consider Benedicks' classical theorem for the separate time-frequency representation.

Theorem 3.6. (Benedicks' Theorem) Consider a function $f \in L^2(\mathbb{R}^d)$ with Fourier transform $\hat{f}$. Let $T = \text{supp} f$ and $\Omega = \text{supp} \hat{f}$. If $|T||\Omega| < \infty$, then $f(t) = 0$ almost everywhere (i.e. $f = 0$).

Proof. Without loss of generality, assume $|T| < 1$ since we may replace f by the dilation $f_a(x) = f(ax)$ for some $a > 0$.

Now, consider the 1-periodization

$$\begin{aligned} \int_{\mathbb{R}^d} \chi_\Omega(\omega) \mathrm{d}\omega &= \sum_{n \in \mathbb{Z}^d} \int_{[0,1]^d} \chi_\Omega(\omega + n) \mathrm{d}\omega \\ &= \int_{[0,1]^d} \sum_{n \in \mathbb{Z}^d} \chi_\Omega(\omega + n) \mathrm{d}\omega = |\Omega| < \infty. \end{aligned} \tag{3.30}$$

Hence, for almost all $\omega_0 \in \mathbb{R}^d$ we have $\chi_\Omega(\omega_0 + n) \neq 0$ only for a finite number of $n \in \mathbb{Z}$. Since $\hat{f} = \hat{f} \cdot \chi_\Omega$, it follows that

(A) for almost all $\omega_0 \in \mathbb{R}^d$, $\hat{f}(\omega_0 + n) \neq 0$ only for a finite number of $n \in \mathbb{Z}^d$.

By a similar argument for T, we obtain that

(B) for almost all $x \in \mathbb{R}^d$, $f(x + n) \neq 0$ only for a finite number $n \in \mathbb{Z}^d$.

[6] The Wigner distribution is an alternative simultaneous time-frequency representation to the STFT and can easily we expressed in terms of the STFT.

Define the function

$$\widetilde{f}_{\omega_0}(x) := \sum_{n\in\mathbb{Z}^d} e^{-2\pi i\omega_0\cdot(x-n)} f(x-n), \tag{3.31}$$

which by (B) is absolutely convergent for almost all $x \in \mathbb{R}^d$. We claim that $\widetilde{f}_{\omega_0}$ satisfies the following properties for almost all $\omega_0 \in \mathbb{R}^d$

(i) $\widetilde{f}_{\omega_0} \in L^1(\mathbb{T}^d)$, where $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ denotes the d-torus,

(ii) $\widetilde{f}_{\omega_0}$ has Fourier coefficients $\hat{f}_{\omega_0}(n) = \hat{f}(\omega_0 + n)$ for $n \in \mathbb{Z}^d$, and

(iii) the support of $\widetilde{f}_{\omega_0}$ in $\mathbb{T}^d$ has measure strictly less than 1, i.e. $C = \mathrm{supp}\widetilde{f}_{\omega_0}$, where $|C| < 1$.

Begin by considering the L^1-norm of $\widetilde{f}_{\omega_0}$ over $\mathbb{T}^d$, that is

$$\begin{aligned}\int_{\mathbb{T}^d} |\widetilde{f}_{\omega_0}(x)|\mathrm{d}x &= \int_{[0,1]^d} \Big|\sum_{n\in\mathbb{Z}^d} e^{-2\pi i\omega_0\cdot(x-n)} f(x-n)\Big|\mathrm{d}x \\ &\leq \int_{[0,1]^d} \sum_{n\in\mathbb{Z}^d} |f(x-n)|\mathrm{d}x = \sum_{n\in\mathbb{Z}^d} \int_{[0,1]^d} |f(x-n)|\mathrm{d}x \\ &= \int_{\mathbb{R}^d} |f(x)|\mathrm{d}x = \|f\|_1.\end{aligned}$$

Since $|T| < \infty$ and $f \in L^2(\mathbb{R}^d)$ implies $f \in L^1(\mathbb{R}^d)$, claim (i) follows. The second claim (ii) is verified by explicit computation of the Fourier coefficients on the d-cube $[0,1]^d$

$$\begin{aligned}\hat{f}_{\omega_0}(m) &= \int_{[0,1]^d} \widetilde{f}_{\omega_0}(x) e^{-2\pi i m\cdot x}\mathrm{d}x \\ &= \int_{[0,1]^d} \sum_{n\in\mathbb{Z}^d} e^{-2\pi i\omega_0\cdot(x-n)} f(x-n) e^{-2\pi i m\cdot x}\mathrm{d}x \\ &= \sum_{n\in\mathbb{Z}^d} \int_{[0,1]^d} e^{-2\pi i\omega_0\cdot(x-n)} f(x-n) e^{-2\pi i m\cdot x}\mathrm{d}x \\ &= \sum_{n\in\mathbb{Z}^d} e^{2\pi i m\cdot n} \int_{n+[0,1]^d} f(x) e^{-2\pi i(\omega_0+m)\cdot x}\mathrm{d}x \\ &= \int_{\mathbb{R}^d} f(x) e^{-2\pi i(\omega_0+m)\cdot x}\mathrm{d}x = \hat{f}(\omega_0+m) \text{ for } m \in \mathbb{Z}^d.\end{aligned}$$

Note that the exchange of summation and integration in line three is valid by the computations for claim (i) and subsequently Lebesque's Dominated Convergence Theorem. As a curiosity, observe that by expressing function (3.31) in terms its Fourier coefficients, we arrive at *Poisson summation formula* (see Chapter 1.4 in [6])

$$\sum_{n\in\mathbb{Z}^d} e^{-2\pi i\omega_0\cdot(x-n)} f(x-n) = \sum_{n\in\mathbb{Z}^d} e^{-2\pi i n\cdot x}\hat{f}(\omega_0+n). \tag{3.32}$$

For the final claim, it is sufficient to recognize that $|C| \leq |T| < 1$.

Combine observations (A) and (iii) to conclude that for almost all $\omega_0 \in \mathbb{R}^d$ the function $\widetilde{f}_{\omega_0}$ is a trigonometric polynomial which is zero on a set of positive measure. Since the zero function is the only trigonometric polynomial which is zero on set a positive measure, it follows that $\widetilde{f}_{\omega_0} = 0$. From (ii), this means $\hat{f}(\omega_0+n) = 0$ for almost all $\omega_0 \in \mathbb{R}^d$, i.e. $\hat{f} = 0$.

□

As mentioned earlier, a similar property holds for the STFT and can be summarized as follows:

Theorem 3.7. Let $V_\phi f$ be the STFT of a signal f with window $\phi \neq 0$ both in $L^2(\mathbb{R}^d)$. If $|\text{supp}V_\phi f| < \infty$, then $f = 0$.

The proof will be based on Janssen's proof in [15]. Note that in contrast to Janssen, we will prove this property directly for the STFT, instead of transformation via the Wigner distribution.

The idea of the proof is inspired by the following observation: If $V_\phi f$ has finite support, and if it can be shown that so does its Fourier transform, then we are in a position to apply Benedicks' theorem to obtain $V_\phi f = 0$. However, instead of considering $V_\phi f$ directly, we consider a family of functions in $L^1(\mathbb{R}^{2d})$, say $\{\chi_{\nu,s}\}_{\nu,s\in\mathbb{R}^d}$, all containing the factor $V_\phi f$. This family is chosen such that it is easy to determine the support of the Fourier transforms and such that $\chi_{\nu,s} = 0$ for all ν, s only if $V_\phi f = 0$. The challenge then becomes how to construct such a family.

We start with an alternative to formula (2.1) in [15]:

Lemma 3.8. Let $a, b, \nu, s \in \mathbb{R}^d$ and let $f_1, f_2, g_1, g_2 \in L^2(\mathbb{R}^d)$. Then

$$\begin{aligned}&\iint_{\mathbb{R}^{2d}} V_{f_1}f_2(\omega,t)\overline{V_{g_1}g_2(\omega+\nu,t+s)}e^{-2\pi i(a\cdot t-b\cdot\omega)}\mathrm{d}t\mathrm{d}\omega\\ &= e^{2\pi i b\cdot\nu}V_{g_2}f_2(a-\nu,b)\overline{V_{g_1}f_1(a,b+s)}.\end{aligned} \tag{3.33}$$

Proof. Define

$$h_1(x) := (g_1)_{a,b+s}(x) = g_1(x - (b+s))e^{2\pi i a\cdot x}$$

and

$$h_2(x) := (g_2)_{a-\nu,b}(x) = g_2(x-b)e^{2\pi i(a-\nu)\cdot x},$$

which are both in $L^2(\mathbb{R}^d)$. Explicit computation then yields

$$V_{h_1}h_2(\omega,t) = e^{-2\pi i(b\cdot\nu - a\cdot t + b\cdot\omega)}V_{g_1}g_2(\omega+\nu, t+s),$$

from which

$$\begin{aligned}
&e^{2\pi i b\cdot\nu}\iint_{\mathbb{R}^{2d}} V_{f_1}f_2(\omega,t)\overline{V_{g_1}g_2(\omega+\nu,t+s)}e^{2\pi i(a\cdot t - b\cdot\omega)}\mathrm{d}t\mathrm{d}\omega\\
&=\iint_{\mathbb{R}^{2d}} V_{f_1}f_2(\omega,t)\overline{V_{h_1}h_2(\omega,t)}\mathrm{d}t\mathrm{d}\omega\\
&=\langle f_2,h_2\rangle\overline{\langle f_1,h_1\rangle}.
\end{aligned}$$

Since it is straightforward to show

$$\langle f_2, h_2\rangle = V_{g_2}f_2(a-\nu, b)$$

and

$$\langle f_1, h_1\rangle = V_{g_1}f_1(a, b+s),$$

we are done.

□

From this formula it becomes easy to construct a sufficient family $\{\chi_{\nu,s}\}_{\nu,s}$.

Proof. (Theorem 3.7) For $\nu, s \in \mathbb{R}^d$ define the function

$$\chi_{\nu,s}(\omega,t) := V_\phi f(\omega,t)\overline{V_f\phi(\omega+\nu,t+s)},$$

which is evidently in $L^1(\mathbb{R}^{2d})$. From formula (3.33), the Fourier transform of $\chi_{\nu,s}$ can be written

$$\hat{\chi}_{\nu,s}(-b,a) = e^{-2\pi i b\cdot\nu}V_\phi f(a-\nu,b)\overline{V_f\phi(a,b+s)}.$$

Since both $\chi_{\nu,s}$ and $\hat{\chi}_{\nu,s}$ contain to some shift of the factor $V_\phi f$, it follows that

$$|\mathrm{supp}\chi_{\nu,s}|, |\mathrm{supp}\hat{\chi}_{\nu,s}| \leq |\mathrm{supp}V_\phi f| < \infty.$$

Thus, by Benedicks' theorem, we conclude that $\chi_{\nu,s} = 0$. Since ν, s were arbitrary, we must have $V_\phi f = 0$ or $V_f\phi = 0$, which only occurs if $\phi = 0$ or $f = 0$.

□

Observe that Donoho-Stark and Lieb's uncertainty principle represent local and indeed quantitative uncertainty principles in the sense that they provide estimates for the concentration of a signal on sets of small measure. In contrast, Benedicks' theorem with its extensions can be regarded as more global and qualitative uncertainty principles where no estimates are produced and instead describe the behaviour of a signal as a whole. Nevertheless, we are able to draw a simple conclusion on the operator norm of the localization operators $\pi_T Q_\Omega$ and $P_{U,\phi}$, which we list as two corollaries.

Corollary 3.1. Let $T, \Omega \subseteq \mathbb{R}^d$ be the time set and frequency set, respectively. Assume that $|T||\Omega| < \infty$. Then the operator norm of $\pi_T Q_\Omega$ is strictly smaller than 1, i.e.

$$\|\pi_T Q_\Omega\|_{\text{op}} < 1. \tag{3.34}$$

Proof. Note that for any linear operator $\mathcal{T} : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$, we have that

$$\|\mathcal{T}\mathcal{T}^*\|_{\text{op}} = \|\mathcal{T}\|_{\text{op}}^2,$$

where $\mathcal{T}^*$ denotes the adjoint of $\mathcal{T}$. Since $\pi_T Q_\Omega \pi_T = (\pi_T Q_\Omega)(\pi_T Q_\Omega)^*$, it is sufficient to show that $\|\pi_T Q_\Omega \pi_T\|_{\text{op}} < 1$ in order to prove the statement.

In the final remarks of section 2.3.1 we concluded that there exists an orthonormal eigenbasis $\{E_j\}_j$ for $L^2(T)$ such that

$$\pi_T Q_\Omega E_j = \pi_T Q_\Omega \pi_T E_j = \lambda_j E_j$$

for some sequence of real scalars $\{\lambda_j\}_j$. Then $\|\pi_T Q_\Omega \pi_T\|_{\text{op}}$ equals the modulus of the largest eigenvalue, say $|\lambda_0|$. Assume therefore, by contradiction, that $|\lambda_0| = 1$. However, $\|\pi_T Q_\Omega E_0\|_2 = 1$ only if $\text{supp}\hat{E}_0 \subseteq \Omega$ and since we already have $\text{supp}E_0 \subseteq T$, we arrive at a contradiction to Benedicks' theorem.

□

Similarly, we have for Daubechies' localization operator:

Corollary 3.2. Suppose $U \subseteq \mathbb{R}^{2d}$ has finite measure. Then the operator norm of the localization operator $P_{U,\phi}$ is strictly smaller than 1, i.e.

$$\|P_{U,\phi}\|_{\text{op}} < 1. \tag{3.35}$$

Proof. The proof is analogous to proof of $\|\pi_T Q_\Omega \pi_T\|_{\text{op}} < 1$ in Corollary 3.1.

□

3.5 Further Results

In this section we briefly mention some more recent results regarding the uncertainty principles, which shows that this is a research area in constant development. In contrast to the previous versions presented, these results take into account some of the *geometry* of the time-frequency domains and not only their *measure*. Note, however, that for the most part the proofs will be omitted and only referenced. Nevertheless, this section serves as reminder that the geometry of the time-frequency domains is not irrelevant when it comes to localization.

Theorem 3.9. (The Paneyah-Logvinenko-Sereda Theorem)
Let E be a measurable subset of $\mathbb{R}^d$ such that for some constants $r, \gamma > 0$

$$|E \cap B| \geq \gamma |B| \quad \forall \text{ balls } B \subseteq \mathbb{R}^d \text{ of radius } r. \tag{3.36}$$

Let $B_\delta(0) \subseteq \mathbb{R}^d$ denote the closed ball of radius $\delta > 0$ centered at the origin. Then for any signal $f \in L^2(\mathbb{R}^d)$ with $\text{supp}\hat{f} \subseteq B_\delta(0)$,

$$\|f\|_2^2 \leq C \|\chi_E f\|_2^2, \tag{3.37}$$

for some constant $C = C(r, \gamma, \delta, d) \geq 1$. In particular, if $\text{supp} f \subseteq \mathbb{R}^d \backslash E$, then f is the zero function, i.e. $f = 0$.

Proof. See the proof of Theorem 10.7 in Muscalu and Schlag's book[16]. Although they impose the symmetry condition $r = \delta^{-1}$, this is not a strict requirement as it only reduces the number of parameters in C.

□

Note that it is easy to construct examples where both E and $\mathbb{R}^d \backslash E$ has infinite measure. Therefore, a time set with infinite measure cannot in and of itself guarantee complete localization. Proceeding, we present a simple upper bound estimate for the operator norm of the relevant localization operator $\pi_{\mathbb{R}^d \backslash E} Q_\Omega$ for $\Omega \subseteq B_\delta(0)$.

Corollary 3.3. Let E be as in Theorem 3.9 with the same constants $\gamma, r > 0$, in addition to the constants $\delta > 0$ and $C = C(r, \gamma, \delta, d) \geq 1$. Suppose Ω is a measurable subset of $B_\delta(0)$. Then the localization operator $\pi_{\mathbb{R}^d \backslash E} Q_\Omega$ satisfies

$$\|\pi_{\mathbb{R}^d \backslash E} Q_\Omega\|_{\text{op}} \leq \sqrt{1 - \frac{1}{C}} < 1. \tag{3.38}$$

Proof. By Theorem 3.9, for any $f \in L^2(\mathbb{R}^d)$, we must have

$$\|Q_\Omega f\|_2^2 \leq C\|\chi_E Q_\Omega f\|_2^2 = C\Big(\|Q_\Omega f\|_2^2 - \|\pi_{\mathbb{R}^d\setminus E} Q_\Omega f\|_2^2\Big).$$

Result (3.38) then follows after rearrangement combined with the fact that $\|Q_\Omega f\|_2 \leq \|f\|_2$.

□

So far we have not provided any estimates of the constant $C = C(r, \gamma, \delta, d)$ in Theorem 3.9. As it turns out, the proof presented in [16] does not contain a sharp estimate. However, this proof is based on the 2001-paper[17] by Oleg Kovrijkine, where some effort is made to estimate C. In the 1-dimensional case Kovrijkine shows that up to some absolute and unknown constant $\beta > 0$, the optimal estimate is given by

$$C(r, \gamma, \delta, d = 1) = \left(\frac{\gamma}{\beta}\right)^{\beta(4r\delta+1)}. \tag{3.39}$$

More recently, Alexander Reznikov has in his 2010-publication[18] determined sharp constants for specific subsets E in the 1-dimensional case. We illustrate by one of the simpler examples in the article:

Example 3.2. (Union of Equidistant Intervals) Suppose $E \subseteq \mathbb{R}$ is of the form

$$E(R) = \bigcup_{n\in\mathbb{Z}} [n - R, n + R] \quad \text{for } R \in]0, 1/2[. \tag{3.40}$$

We will now outline a short argument for how to derive the constant C in Theorem 3.9 for any signal f with $\text{supp}\hat{f} \subseteq [-\frac{1}{2}, \frac{1}{2}]$:

Firstly, by a similar 1-periodization as in Benedicks' theorem,

$$\int_E |f(x)|^2 \mathrm{d}x = \int_{-1/2}^{1/2} \chi_{[-R,R]}(x) \sum_{n\in\mathbb{Z}} |f(x + n)|^2 \mathrm{d}x.$$

Secondly, express the two factors $\chi_{[-R,R]}(x)$ and $\sum_{n\in Z} |f(x + n)|^2$ in terms of their Fourier series on $[-\frac{1}{2}, \frac{1}{2}]$ with the standard orthonormal basis $\{e^{-2\pi ikx}\}_{k\in\mathbb{Z}}$, from which we obtain

$$\|\chi_E f\|_2^2 = \sum_{k\in\mathbb{Z}} \widehat{|f|^2}(k) \int_{-R}^{R} e^{-2\pi ikx} \mathrm{d}x. \tag{3.41}$$

By the Convolution Theorem, $\widehat{|f|^2} = \hat{f} * \hat{\overline{f}}$, where

$$f * g(t) = \int_{\mathbb{R}} f(x)g(t-x)\mathrm{d}x$$

denotes the standard 1-dimensional convolution. Since $\mathrm{supp}\hat{f} \subseteq [-\frac{1}{2}, \frac{1}{2}]$, it follows that $\mathrm{supp}\widehat{|f|^2} \subseteq [-1, 1]$ with $\widehat{|f|^2}(\pm 1) = 0$. Hence, the sum (3.41) reduces to

$$\|\chi_E f\|_2^2 = 2R\widehat{|f|^2}(0) = 2R\|f\|_2^2.$$

From here we conclude that the constant in Theorem 3.9 is given by $C = C(R) = (2R)^{-1}$. Since this constant is obviously sharp, Corollary 3.3 yields

$$\|\pi_{\mathbb{R}\setminus E(R)} Q_{[-\frac{1}{2},\frac{1}{2}]}\|_{\mathrm{op}} = \sqrt{1-2R}. \tag{3.42}$$

Observe that even though the measure of $\mathbb{R}\setminus E$ is infinite for any $R \in]0, 1/2[$, the operator norm $\|\pi_{\mathbb{R}\setminus E(R)} Q_{[-\frac{1}{2},\frac{1}{2}]}\|_{\mathrm{op}} \to 0$ as $R \to 1/2$. In addition, notice that the quantity $1 - 2R$ coincides with the length of each interval in $\mathbb{R}\setminus E$.

For the STFT there exists a somewhat similar result to Theorem 3.9, which prohibits complete localization on a certain family of subsets (possibly of infinite measure) but now in the time-frequency plane. This was discovered by Carmen Fernández and Antonio Galbis in their 2010-paper[19]. In said paper they focus on the following subsets:

A measurable set $U \subseteq \mathbb{R}^{2d}$ is said to be *thin at infinity* if

$$\lim_{|x|\to\infty} |U \cap B_R(x)| = 0 \quad \text{for some (for all) } R > 0, \tag{3.43}$$

where $B_R(x)$ denotes the closed ball of radius R centered at $x \in \mathbb{R}^{2d}$.

With this definition in mind, we formulate the next theorem.

Theorem 3.10. Suppose $U \subseteq \mathbb{R}^{2d}$ is a thin set at infinity, and let $\phi \in L^2(\mathbb{R}^d)$ be a fixed, non-zero window function. Then there exists a constant $C \geq 1$ such that

$$\|f\|_2^2 \leq C \iint_{\mathbb{R}^{2d}\setminus U} |V_\phi f(\omega, t)|^2 \mathrm{d}\omega \mathrm{d}t \quad \forall\ f \in L^2(\mathbb{R}^d). \tag{3.44}$$

In particular, if $\mathrm{supp} V_\phi f \subseteq U$, then $f = 0$.

Proof. See proof of Theorem 4.1 in [19]. □

Similarly to Corollary 3.3, we may rephrase the above theorem as a statement about the relevant localization operator, namely Daubechies' localization operator $P_{U,\phi}$.

Corollary 3.4. Let $U \subseteq \mathbb{R}^{2d}$ and $\phi \in L^2(\mathbb{R}^d)$ be as in Theorem 3.10 with the same constant $C \geq 1$. Presume ϕ is normalized, then Daubechies' localization operator $P_{U,\phi}$ satisfies

$$\|P_{U,\phi}\|_{\mathrm{op}} \leq 1 - \frac{1}{C} < 1. \tag{3.45}$$

Proof. By an analogous argument as in Corollary 3.3, we rearrange (3.44) such that

$$\langle P_{U,\phi} f, f \rangle = \iint_U |V_\phi f(\omega, t)|^2 \mathrm{d}\omega \mathrm{d}t \leq \left(1 - \frac{1}{C}\right) \|f\|_2^2 \quad \forall\ f \in L^2(\mathbb{R}^d).$$

By Cauchy-Schwarz' inequality on $L^2(U)$,

$$|\langle P_{U,\phi} f, g \rangle| \leq \max\{\langle P_{U,\phi} f, f \rangle, \langle P_{U,\phi} g, g \rangle\} \quad \forall\ f, g \in L^2(\mathbb{R}^d),$$

from which the statement follows.

□

Note that Fernández and Galbis proof does not offer any sharp estimates of the constant involved. Nevertheless, for specific subsets it is possible to estimate the operator norm of $P_{U,\phi}$ more accurately. In section 4.2.3 we consider an example when the subset in question U has infinite measure. This example is, to some extent, meant to mirror Example 3.2 for the separate time-frequency representation, but now for the joint representation. Note, however, that in the example U is not thin at infinity. This goes to show that the notion of thin at infinity is a sufficient condition for inequality (3.44) and *not* a necessary one.

Returning to the separate time-frequency representation, in notes from 2017 recent contributions have been made by Semyon Dyatlov[5] on the topic of uncertainty principles in 1 dimension. The primary focus in these notes is what Dyaltov refers to a the *fractal uncertainty principle*. In particular, this means that the time and frequency sets take the form of fractal sets or exhibit a prescribed regularity close to it. One specific fractal set that is studied in detail is the *Cantor set*. We conclude this section by outlining in what sense Dyatlov measures localization on the Cantor set in time and frequency.

Example 3.3. (Dyatlov's Localization on the Cantor Set) A formal definition of the Cantor set can be found in section 4.3. For the purpose of this example, recall that the Cantor set is a compact subset of $\mathbb{R}$ which is obtained by a nested iterative scheme $C_0 \supseteq C_1 \supseteq C_2 \supseteq \dots$ While the Cantor set itself has zero measure, each iterate does not. In particular, if the Cantor set $C(M)$ is based in the interval $[0, M]$, then the n-iterate $C_n(M)$ consists of 2^n disjoint intervals in $[0, M]$ each with measure $3^{-n}M$. Therefore, instead of projecting onto the Cantor set directly, we project onto one of the iterates.

Although projecting onto $C_n(M)$ is essentially a two-parameter problem (in terms of n and M), Dyatlov reduces this to a one-parameter problem. This in the sense that the iterate n is not chosen independently from M. More precisely, if I_j denotes a single interval in $C_n(M)$, we consider n such that $|I_j| \sim 1/M$. Since $|I_j| = 3^{-n}M$, this means

$$3^n \sim M^2. \tag{3.46}$$

Hence, for any iterate $C_n(M)$ that satisfies condition (3.46), we must have

$$|C_n(M)| \sim \left(\frac{2}{\sqrt{3}}\right)^n \sim M^{2\ln 2/\ln 3 - 1}. \tag{3.47}$$

Let $\{X_n := C_n(M(n))\}_n$ denote such family of n-iterates based in an ever increasing interval, and consider the corresponding localization operator $\pi_{X_n} Q_{X_n}$. Then by Theorem 2.1.1 in [5],[7] there exists constants $\alpha, \beta > 0$ such that the operator norm is bounded by

$$\|\pi_{X_n} Q_{X_n}\|_{\mathrm{op}} \le \alpha e^{-\beta n} \quad \forall\ n = 0, 1, 2, \dots \tag{3.48}$$

Since X_n satisfies the measure scale (3.47), it follows that the operator norm $\|\pi_{X_n} Q_{X_n}\|_{\mathrm{op}} \to 0$ as $|X_n| \to \infty$.

[7] Observe that by reading Dyatlov's notes (see Chapter 2.1 [5]), it may appear that the problem is framed somewhat differently than what is presented above. In particular, Dyatlov only considers Cantor sets based in $[0, 1]$, where the approximation sets are characterized by the single parameter $h \in [0, 1]$, e.g., representing the length of the intervals included in the iterate. However, this parameter is also encoded in the Fourier transform (not unlike the normalization in quantum mechanics) such that

$$\mathcal{F}_h : L^2(\mathbb{R}) \to L^2(\mathbb{R}), \quad \text{where} \quad \mathcal{F}_h f(\omega) = \frac{1}{\sqrt{h}} \hat{f}(\omega/h).$$

Since for any h-dependent sets $X(h), Y(h) \subseteq [0, 1]$ we have

$$\|\chi_{X(h)} \mathcal{F}_h \chi_{Y(h)}\|_{\mathrm{op}} = \|\chi_{X(h)/\sqrt{h}} \mathcal{F} \chi_{Y(h)/\sqrt{h}}\|_{\mathrm{op}} = \|\pi_{X(h)/\sqrt{h}} Q_{Y(h)/\sqrt{h}}\|_{\mathrm{op}},$$

we can easily translate to the approach of n-iterate Cantor sets in $[0, M]$ with restrictions (3.46), (3.47).

4 Spherically Symmetric Weight

In this chapter we will focus on a particular class of Daubechies' time-frequency localization operators. As the title suggests, we will consider operators with a spherically symmetric weight function F. More precisely, let $r_j^2 = \omega_j^2 + t_j^2$ denote the radius squared for each time-frequency dimension $j = 1, 2, \ldots, d$. Collect these coordinates into a single vector $r^2 = (r_1^2, r_2^2, \ldots, r_d^2)$. In this notation, we consider

$$F(\omega, t) = \mathscr{F}(r^2). \tag{4.1}$$

Note that the weight $F \in L^1(\mathbb{R}^{2d})$ precisely when $\mathscr{F} \in L^1(\mathbb{R}_+^d)$, where $\|F\|_1 = \pi^d \|\mathscr{F}\|_1$. Although such weights can be analyzed in a multidimensional phase space, we will consider the two dimensional time-frequency plane, i.e. when $d = 1$. Furthermore, unless otherwise stated, we always presume the weight to be real-valued, non-negative and integrable.

In general, it is difficult to determine expressions for either the eigenvalues or the eigenfunctions of a localization operator, given an arbitrary window ϕ and weight F. This prospect remains true even when the weight is chosen to be spherically symmetric. Nevertheless, for certain windows the symmetry of the weight can be exploited to produce explicit expressions for both the eigenvalues and eigenfunctions of the associated localization operator.

One such window is the *normalized Gaussian*, i.e.

$$\phi(x) = 2^{1/4} e^{-\pi x^2}, \tag{4.2}$$

which was investigated by Daubechies in her pioneering 1988-paper[4]. For this canonical choice for ϕ, explicit expressions for the eigenvalues are derived in section IV of the article, where the *Hermite functions* H_k for $k = 0, 1, 2, \ldots$ (see definition (3.3)) are shown to constitute eigenfunctions of the operator.[8] Since the Hermite functions form an orthonormal basis for $L^2(\mathbb{R})$, they provide a complete description of the eigenfunctions and subsequently the spectrum of said operator.

[8] Due to the choice of normalization for the Fourier transform, both the Gaussian and the Hermite functions are normalized differently than in [4]. The normalization is chosen in accordance with Folland[9]. If h_k denotes the k-th Hermite function in [4], this relates to H_k in (3.3) by $H_k(x) = \frac{2^{1/4}}{\sqrt{2^k k!}} h_k(\sqrt{2\pi} x)$.

By recognizing that the normalized Gaussian coincides with H_0, it is natural to ask if an arbitrary H_j as window can produce similar results. The 2017-project report[20] attempted to answer that question, and after correcting for a misprint in one of the formulas the project work was based on,[9] it was shown that the Hermite functions still constitute the eigenbasis. In addition, general formulas for the eigenvalues with H_j as window were derived and expressed in terms of the Laguerre polynomials.

As remarked in the project report, the complexity of these expressions is positively correlated with the order of the Hermite function chosen as window. This follows from the fact that the effective order of the Laguerre polynomials increases for increasing j with H_j as window. Thus, the simplest expressions are obtained with H_0 as window, i.e. the Gaussian window.

Throughout the remainder of this chapter we will only consider the case when the window ϕ equals the normalized Gaussian. It is well-known that the Gaussian minimizes Heisenberg's uncertainty principle for the regular Fourier transform (see Theorem 3.2). In addition, certain Gaussians (including the normalized Gaussian) are shown in [12] to minimize Lieb's inequality in Lemma 3.4. Since the Gaussian window indeed provides optimal resolution for the STFT, it is among the most popular choices for windows, which is further substantiated by the aforementioned result on the eigenvalues.

4.1 Formulas for Eigenvalues with Gaussian Window

For completeness, we derive the formulas for the eigenvalues of the time-frequency localization operator with a spherically symmetric weight and the Gaussian window.

Theorem 4.1. Let $P_{F,\phi}$ denote the localization operator with weight $F(\omega, t) = \mathscr{F}(r^2)$ and window ϕ equals the normalized Gaussian in (4.2). Then the eigenvalues of $P_{F,\phi}$ are given by

$$\lambda_k = \int_0^\infty \mathscr{F}\left(\frac{r}{\pi}\right)\frac{r^k}{k!}e^{-r}\mathrm{d}r, \text{ for } k = 0, 1, 2, \ldots, \tag{4.3}$$

such that

$$P_{F,\phi}H_k = \lambda_k H_k, \tag{4.4}$$

where H_k denotes the k-th Hermite function.

[9] In January of 2018 a misprint in Folland's book[9] was discovered and corrected for in an additional note[21]. This correction proved to simplify the subsequent calculations.

In order to prove the above statement, we require the following lemma:

Lemma 4.2. Let H_k denote the k-th Hermite function and $\phi = H_0$. Then

$$V_\phi H_k(\omega, t) = \sqrt{\frac{\pi^k}{k!}}(t - i\omega)^k e^{-\pi i\omega\cdot t} \cdot e^{-\pi(\omega^2+t^2)/2}. \tag{4.5}$$

Proof. By definition (3.3),

$$\begin{aligned} V_\phi H_k(\omega, t) &= \int_{\mathbb{R}} H_k(x)\overline{H_0(x-t)}e^{-2\pi i\omega x}\mathrm{d}x \\ &= \frac{2^{1/2}}{\sqrt{k!}}\left(-\frac{1}{2\sqrt{\pi}}\right)^k e^{-\pi t^2}\int_{\mathbb{R}} e^{2\pi(t-i\omega)x}\left(\frac{\mathrm{d}^k}{\mathrm{d}x^k}e^{-2\pi x^2}\right)\mathrm{d}x. \end{aligned}$$

After integration by parts k times, we obtain

$$\begin{aligned} V_\phi H_k(\omega, t) &= \frac{2^{1/2}}{\sqrt{k!}}\left(\frac{1}{2\sqrt{\pi}}\right)^k e^{-\pi t^2}\int_{\mathbb{R}}\left(\frac{\mathrm{d}^k}{\mathrm{d}x^k}e^{-2\pi(t-i\omega)x}\right)e^{-2\pi x^2}\mathrm{d}x \\ &= 2^{1/2}\sqrt{\frac{\pi^k}{k!}}(t-i\omega)^k e^{-\pi t^2}\int_{\mathbb{R}} e^{-2\pi(x^2+(t-i\omega)x)}\mathrm{d}x. \end{aligned} \tag{4.6}$$

It is well-known that the integral

$$\int_{\mathbb{R}} e^{-2\pi(x^2+ax)}\mathrm{d}x = \frac{1}{2^{1/2}}e^{-a^2/2} \quad \forall\ a \in \mathbb{C}.$$

Apply the above identity to (4.6) for $a = t - i\omega$ to complete the proof.

□

With this estimate, we are able to derive the formulas for the eigenvalues.

Proof. (Theorem 4.1) It is sufficient to verify

$$\langle P_{F,\phi}H_k, H_l\rangle = \lambda_k\delta_{k,l}, \tag{4.7}$$

where $\delta_{k,l}$ denotes the Kronecker-delta which is one for $k = l$ and zero otherwise. By Lemma 4.2, we have

$$\begin{aligned} \langle P_{F,\phi}H_k, H_l\rangle &= \iint_{\mathbb{R}^2} F(\omega, t)V_\phi H_k(\omega, t)\overline{V_\phi H_l(\omega, t)}\mathrm{d}t\mathrm{d}\omega \\ &= \sqrt{\frac{\pi^{k+l}}{k!l!}}\iint_{\mathbb{R}^2} F(\omega, t)(t-i\omega)^k(t+i\omega)^l e^{-\pi(\omega^2+t^2)}\mathrm{d}t\mathrm{d}\omega. \end{aligned}$$

Now, rewrite the integral over phase space in terms of radius $r > 0$ and angle θ, such that $t + i\omega = re^{i\theta}$. Since F is spherically symmetric, the above integral factorizes into an angular and radial integral. The angular integral is of the form $\int_0^{2\pi} e^{i\theta(l-k)} \mathrm{d}\theta$ which is non-zero precisely when $k = l$. Thus for the non-zero case, we obtain

$$\langle P_{F,\phi} H_k, H_k \rangle = 2\pi \int_0^\infty \mathscr{F}(r^2) \frac{(\pi r^2)^k}{k!} e^{-\pi r^2} r \mathrm{d}r.$$

The final formula (4.3) follows by the substitution $s = \pi r^2$.

□

4.2 Localization on Spherically Symmetric Set

In this section we consider the case when the weight $\mathscr{F}$ equals the characteristic function of some subset $E \subseteq \mathbb{R}_+$, i.e.

$$\mathscr{F}(r) = \chi_E(r). \tag{4.8}$$

To avoid confusion, notice that due to the definition of $\mathscr{F}$ in (4.1), any point $r \in E$ corresponds precisely to all points in the plane whose radius *squared* equals r. In total, the set E is identified with the following subset of the plane

$$\mathcal{E} = \{(\omega, t) \in \mathbb{R}^2 \mid \omega^2 + t^2 \in E\}. \tag{4.9}$$

Although the induced localization operator aims at concentrating signals on the set $\mathcal{E}$, we will denote the operator in question by P_E as a matter of convenience in the spherically symmetric context. For this family of operators, we will attempt to derive upper and lower bound estimates for the operator norm based on the measure of E in Proposition 4.2 and 4.3, respectively.

From Theorem 4.1, the eigenvalue corresponding to the k-th Hermite function is given by

$$\lambda_k = \int_{\pi \cdot E} \frac{r^k}{k!} e^{-r} \mathrm{d}r, \text{ for } k = 0, 1, 2, \ldots, \tag{4.10}$$

where $\pi \cdot E := \{x \in \mathbb{R}_+ \mid x/\pi \in E\}$. Since the above integrands will appear frequently, we define, for simplicity, the functions

$$f_k(r) := \frac{r^k}{k!} e^{-r}, \; r \geq 0, \text{ for } k = 0, 1, 2, \ldots \tag{4.11}$$

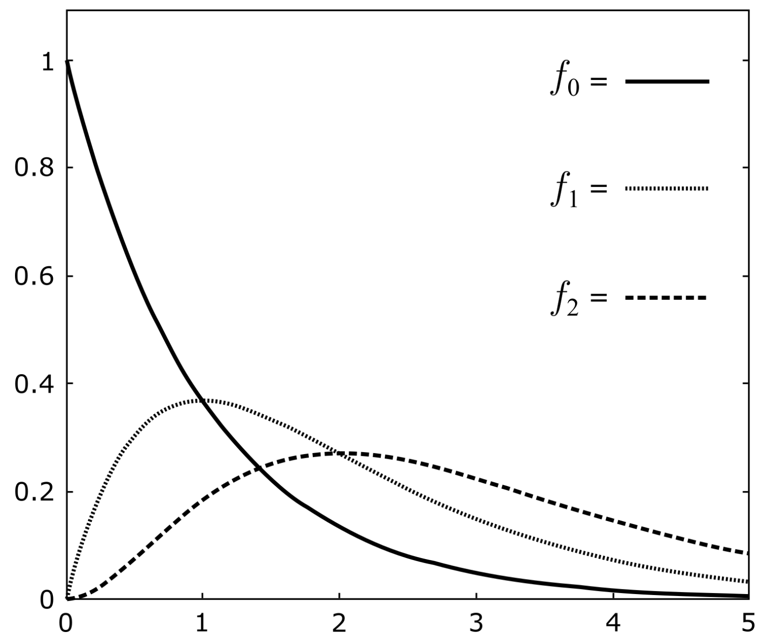

Figure 1: Plot of the first three gamma distributed integrands f_0, f_1, f_2, where $f_k(r) = \frac{r^k}{k!}e^{-r}$ for $r \geq 0$ and $k = 0, 1, 2, \ldots$

As an interesting sidenote, by inspection, we recognize f_k to be the probability density of a gamma distributed variable with expectation value and variance equal to $k+1$. Hence, the study of these eigenvalues $\{\lambda_k\}_k$ could alternatively be presented as a study of the probability distributions $\{f_k\}_k$ (without the terminology of time-frequency localization operators attached). For this reason, we begin by making a few remarks on these integrands.

From the perspective of probability distributions or even from the perspective of localization operators, it should come as no surprise that every function f_k is normalized, i.e. $\|f_k\|_1 = 1$ for $k = 0, 1, 2, \ldots$ Differentiating $f_k(r)$ reveals a single local (and global) maximum at point $r = k$ such that

$$f_k(r) \leq f_k(k) = \frac{k^k}{k!}e^{-k} \quad \forall\ r \geq 0. \tag{4.12}$$

Therefore, we must have that $f_k(r)$ is monotonically increasing for $r \leq k$. Conversely, $f_k(r)$ is monotonically decreasing for $r > k$. This behaviour is illustrated in Figure 1 for the first three integrands f_0, f_1, f_2. In the next lemma we consider symmetric points about this maximum.

Lemma 4.3.

$$f_k(k-r) \leq f_k(k+r) \quad \forall\, r \in [0,k] \text{ for } k = 1,2,3,\dots \tag{4.13}$$

Proof. Consider the fraction

$$\delta_k(r) = \frac{f_k(k-r)}{f_k(k+r)} = \left(\frac{k-r}{k+r}\right)^k e^{2r},$$

which is a well-defined function for all $r \in [0,k]$. Further, observe that inequality (4.13) holds if and only if $\delta_k(r) \leq 1$ for all $r \in [0,k]$. Consider therefore its derivative,

$$\delta_k'(r) = -\frac{2r^2e^{2r}}{(k+r)^2}\left(\frac{k-r}{k+r}\right)^{k-1} \leq 0 \quad \forall\, r \in [0,k].$$

Since $\delta_k(0) = 1$, we conclude from latest result that $\delta_k(r) \leq 1$ whenever $r \in [0,k]$.

□

This latest lemma will prove especially relevant when we discuss the operator norm for localizing on the spherically symmetric Cantor set in section 4.3.3. More generally, Lemma 4.3 reveals a simple feature regarding the distribution of the L^1-norm of f_k. In particular, a larger portion of the norm is located right of the maximum rather than left of the maximum, that is

$$\|\chi_{[0,k[}f_k\|_1 \leq \|\chi_{[k,2k[}f_k\|_1 < \|\chi_{[k,\infty[}f_k\|_1 \text{ for } k = 0,1,2,\dots$$

Now, consider the difference between two subsequent integrands

$$f_k(r) - f_{k+1}(r) = \frac{r^k}{k!}e^{-r} - \frac{r^{k+1}}{(k+1)!}e^{-r} = \frac{r^k e^{-r}}{(k+1)!}(k+1-r). \tag{4.14}$$

It is clear that f_k, f_{k+1} intersect only at the point $r = k+1$, which, as it happens, is the maximum of $f_{k+1}(r)$. The above difference is therefore negative precisely when $r > k+1$. Thus, depending on the choice of subset $E \subseteq \mathbb{R}_+$, the integral of f_k over $\pi \cdot E$ can be made larger or smaller than the integral of f_{k+1}.

4.2.1 Two Examples: Disk and Ring

In this section we consider two natural examples of spherically symmetric subsets, namely a disk and a ring.[10] Both of these subsets are examples of *connected* subsets, and as we shall see, it becomes easy to compute the eigenvalues of the associated localization operator in each case.

Example 4.1. (Localization on a Disk) Suppose the weight equals the characteristic function of a disk of radius $R > 0$ centered at the origin, that is

$$\mathscr{F}(r) = \chi_{[0,R^2[}(r). \tag{4.15}$$

Besides the non-zero constant function, which is definitely not integrable, this is perhaps the simplest choice for the weight. From formula (4.10) and definition (4.11), the eigenvalues of $P_{[0,R^2[}$ read

$$\lambda_k(R) = \int_0^{\alpha(R)} f_k(r)\mathrm{d}r, \text{ for } k = 0, 1, 2, \ldots, \tag{4.16}$$

where $\alpha(R) = \pi R^2$, and we have written $\lambda_k = \lambda_k(R)$ to emphasize the dependency on the radius R of the disk. By integration by parts on the above integral and inductive reasoning, it follows that

$$\lambda_k(R) = 1 - e^{-\alpha(R)} \sum_{n=0}^{k} \frac{(\alpha(R))^n}{n!}, \text{ for } k = 0, 1, 2, \ldots \tag{4.17}$$

It is straightforward to verify that

$$\lambda_k(R) - \lambda_{k+1}(R) = e^{-\alpha(R)} \frac{(\alpha(R))^{k+1}}{(k+1)!} > 0,$$

from which we obtain the ordering

$$\lambda_0(R) > \lambda_1(R) > \cdots > \lambda_k(R) > \lambda_{k+1}(R) > \ldots \tag{4.18}$$

Consequently, the operator norm of $P_{[0,R^2[}$ is given by

$$\|P_{[0,R^2[}\|_{\text{op}} = \lambda_0(R) = 1 - e^{-\alpha(R)} = 1 - e^{-\pi R^2}. \tag{4.19}$$

[10]Here, the word "ring" is not used in the algebraic sense, rather it refers to the geometrical object also known as an *annulus*.

As an initial inspection, observe that $\|P_{[0,R^2[}\|_{\text{op}} < 1$ for all $R > 0$, which is in accordance with Corollary 3.2 to Benedicks' theorem. Furthermore, by condition (3.27), we have

$$\|P_{[0,R^2[}\|_{\text{op}} = 1 - \epsilon_{\min} \tag{4.20}$$

Recall from section 3.3, for a any signal $f \in L^2(\mathbb{R})$, the quantity $(1 - \epsilon_{\min})$ represents the maximum portion of $\|V_\phi f\|_2^2 (= \|f\|_2^2)$ that can be concentrated on the subset in question. In the current context the subset is a disk in phase space of radius $R > 0$. By comparing equation (4.19) and (4.20), it becomes clear that at most $(1 - e^{-\pi R^2})$ of $\|V_\phi f\|_2^2$ can be concentrated on the disk, and this is realized precisely for the normalized Gaussian $f = \phi$ $(= H_0)$. For a numerical example, suppose $\pi R^2 = 1$, then at most $(1 - e^{-1}) \approx 0.632 = 63.2\%$ of $\|V_\phi f\|_2^2$ can be concentrated here.

The above example is an important one, not only because it is illustrative with regard to the eigenvalues and the uncertainty principle but because of formula (4.17). With this result, we can easily deduce that

$$\int_a^b f_k(r)\mathrm{d}r = e^{-a}\sum_{n=0}^{k}\frac{a^n}{n!} - e^{-b}\sum_{n=0}^{k}\frac{b^n}{n!} \quad \forall\ a, b \in \mathbb{R}_+. \tag{4.21}$$

The latest identity is relevant as any open subset $E \subseteq \mathbb{R}_+$ can be written as a countable union of disjoint intervals, i.e. $E = \cup_n]a_n, b_n[$. Thus, in principle, every eigenvalue can be computed such that we exchange the integral (4.10) for a linear combination of sums of the form (4.21). However, evaluating or estimating the sum representation is not necessarily easier than considering the integrals themselves directly. This is especially true when the number of intervals becomes large or when we want to compare the eigenvalues to determine their ordering.

In the next example we still let $E \subseteq \mathbb{R}_+$ consist only of a single interval, i.e. we consider a ring in phase space. Here we show that under the right conditions any λ_k can be made the largest eigenvalue.

Example 4.2. (Localization on a Ring: Asymptotic Estimate) Suppose the weight function equals the characteristic function of a ring with inner radius $R \geq 0$ and measure 1, that is

$$\mathscr{F}(r) = \chi_{[R^2, R^2+\pi^{-1}[}(r). \tag{4.22}$$

By formula (4.10), the eigenvalues then read

$$\lambda_k(R) = \int_{\alpha(R)}^{\alpha(R)+1} f_k(r) \mathrm{d}r \text{ for } k = 0, 1, 2, \ldots, \tag{4.23}$$

where we once again have written $\lambda_k = \lambda_k(R)$, now to indicate the dependency on the *inner* radius $R \geq 0$. Note that for $R = 0$, the ring reduces to a disk of area 1, which, by Example 4.1, means λ_0 is the largest eigenvalue. However, for $R > 0$, this is not always the case.

Assume $\alpha(R) \in [k, k+1]$ for some $k \in \mathbb{N} \cup \{0\}$. By the difference (4.14) and subsequently the monotonicity of the integrals in (4.23), we must have the ordering

$$\lambda_0(R) \leq \lambda_1(R) \leq \lambda_2(R) \leq \cdots \leq \lambda_k(R),$$

in addition to

$$\lambda_{k+1}(R) \geq \lambda_{k+2}(R) \geq \lambda_{k+3}(R) \geq \ldots$$

Thus, by process of elimination, the largest eigenvalue under these conditions is either $\lambda_k(R)$ or $\lambda_{k+1}(R)$. In particular, if $\alpha(R) = k$, then $\lambda_k(R)$ is the largest eigenvalue. In any case, the operator norm of $P_{[R^2, R^2+\pi^{-1/2}[}$ for the ring of measure 1 can be written as

$$\|P_{[R^2, R^2+\pi^{-1}[}\|_{\text{op}} = \max\{\lambda_k(R) \mid k \in [\alpha(R)-1, \alpha(R)+1]\}. \tag{4.24}$$

Proceeding, we will not derive an exact formula for the operator norm, rather we present a simple upper and lower bound estimate. These estimates will be based on a zero order approximation of the integrands f_k.

Note that since the eigenvalues are obtained by integrating over an interval $[\alpha(R), \alpha(R)+1]$ of measure 1, the eigenvalue $\lambda_k(R)$ will be bounded from above by the *maximum* value of f_k on said interval. Similarly, λ_k will be bounded from below by the *minimum* value of f_k on the interval in question. For small k, this is a rather crude estimate, while for larger k these approximations become increasingly more accurate. This can be argued from the fact that the for any fixed $r > 0$, we have that $\lim_{k\to\infty} |f_k'(r)| = 0$.

Let $\lfloor\cdot\rfloor$ denote the floor function which rounds down any number to the nearest integer. Now, define $n(R) := \lfloor\alpha(R)\rfloor$. Then by result (4.24) and the difference (4.14), the operator norm is always bounded by

$$\begin{aligned}\|P_{[R^2,R^2+\pi^{-1}[}\|_{\text{op}} &\le \max_{r\ge 0} f_{n(R)}(r) \\ &= f_{n(R)}(n(R)) = \frac{n(R)^{n(R)}}{n(R)!}e^{-n(R)} \text{ (by (4.12))}. \end{aligned} \tag{4.25}$$

For the lower bound estimate, note that the interval $[\alpha(R), \alpha(R)+1]$ is contained in $[n(R), n(R)+2] =: I_{n(R)}$. Hence, the operator norm is bounded from below by the minimal value of any f_k on $I_{n(R)}$. In particular, the operator norm must be greater than the minimal value of $f_{n(R)+1}$ on $I_{n(R)}$, which, by Lemma (4.13), means

$$\|P_{[R^2,R^2+\pi^{-1}[}\|_{\text{op}} \ge f_{n(R)+1}(n(R)) = \frac{n(R)^{n(R)+1}}{(n(R)+1)!}e^{-n(R)}. \tag{4.26}$$

By *Stirling's asymptotic formula* (see Example 7.4.10 p.301 in [22])

$$\sqrt{2\pi}\cdot n^{n+1/2}e^{-n} \le n! \le e^{\frac{1}{12n}}\sqrt{2\pi}\cdot n^{n+1/2}e^{-n} \text{ for } n = 1, 2, 3, \ldots, {}^{11} \tag{4.27}$$

the two bounds (4.25), (4.26) can be expressed more conveniently.

Proposition 4.1. The operator norm of $P_{[R^2,R^2+\pi^{-1/2}[}$ satisfies the bounds

$$\begin{aligned}\frac{1}{\sqrt{2\pi}}n(R)^{-1/2}\left(1+\frac{1}{n(R)}\right)^{-1}e^{-\frac{1}{12n(R)}} &\le \|P_{[R^2,R^2+\pi^{-1}[}\|_{\text{op}} \\ &\le \frac{1}{\sqrt{2\pi}}n(R)^{-1/2} \quad \forall\ R \ge 0. \end{aligned} \tag{4.28}$$

Proof. For $n(R) = 0$, inequality (4.28) reads $0 \le \|P_{[R^2,R^2+\pi^{-1}[}\|_{\text{op}} \le \infty$, which is a trivial statement for all bounded (and unbounded) linear operators. For $n(R) > 0$, simply apply the lower and upper bound version of Stirling's formula on $n(R)!$ in (4.25) and (4.26), respectively. □

In terms of $\alpha(R)$, note that

$$n(R) = \alpha(R)\Big(1 - \sigma(R)\Big)$$

[11] The lower bound also holds for $n = 0$.

for some positive function $\sigma(R) \leq 1$, where $\sigma(R) = \mathcal{O}(\alpha(R)^{-1}) = \mathcal{O}(R^{-2})$ as $R \to \infty$. Combine this observation with Proposition 4.1, from which we deduce that

$$\begin{aligned} \|P_{[R^2,R^2+\pi^{-1}[}\|_{\mathrm{op}} &= \frac{1}{\sqrt{2\pi}}\alpha(R)^{-1/2} + \mathcal{O}(R^{-3}) \\ &= \frac{1}{\pi\sqrt{2}}R^{-1} + \mathcal{O}(R^{-3}) \quad \text{as } R \to \infty. \end{aligned} \tag{4.29}$$

Observe that the estimates for the operator norm in Example 4.2 can easily be generalized to rings with measure less than 1, where we only adjust by multiplying by the measure of the ring. In particular, consider the operator $P_{[R^2,R^2+a\pi^{-1}[}$ with $a \in]0,1[$. Then

$$\begin{aligned} \frac{1}{\sqrt{2\pi}}n(R)^{-1/2}\left(1+\frac{1}{n(R)}\right)^{-1} e^{-\frac{1}{12n}}a &\leq \|P_{[R^2,R^2+a\pi^{-1}[}\|_{\mathrm{op}} \\ &\leq \frac{1}{\sqrt{2\pi}}n(R)^{-1/2}a \quad \forall\ R \geq 0. \end{aligned} \tag{4.30}$$

Similarly to what was done at the end of Example 4.1, for any $f \in L^2(\mathbb{R})$ we may compute the largest portion of $\|V_\phi f\|_2^2$ that can be concentrated on a ring in phase space of measure 1. By the latest results, for a large inner radius $R > 0$, this portion corresponds to $\approx \pi^{-1}R^{-1}/\sqrt{2}$.

To give a numerical example, suppose the inner radius is $R = 10$, then at most $\approx 0.0225 = 2.25\%$ of $\|V_\phi f\|_2^2$ can be concentrated here. Compare this number to the disk of the same measure, where at most $\approx 63.2\%$ of $\|V_\phi f\|_2^2$ could be concentrated there. Thus, replacing the disk of measure 1 with a ring with inner radius $R = 10$ and measure 1 reduces the best possible localization on the subset to a fraction ($\approx 1/28$). As the inner radius R increases, this reduction will become more extreme.

4.2.2 Bounds for the Operator Norm

Before presenting the upper and lower bound estimate for the operator norm P_E, we make a few (necessary) remarks on the integrand f_k. Recall that in the introductory remarks we observed that $f_k(r)$ contains a single maximum at $r = k$, from which the integrand is monotonically increasing and decreasing for $r \leq k$ and $r \geq k$, respectively. Thus, we obtain the following statement:

Lemma 4.4. Let E be a measurable subset of $\mathbb{R}_+$. Then there exists an interval $I_k \subseteq \mathbb{R}_+$ with measure $|I_k| = |E|$ such that

$$\int_E f_k(r)\mathrm{d}r \leq \int_{I_k} f_k(r)\mathrm{d}r, \text{ for } k = 0, 1, 2, \ldots \tag{4.31}$$

Proof. For a more thorough discussion on integral bounds on monotonic functions, see Appendix A. Without loss of generality, we may assume $|E| < \infty$. Based on the preceding remarks, if E is bounded, the above inequality is a special case of Corollary A.2. Assume therefore E is unbounded. Let $I_k(s) = [s, s+|E|[$. By continuity, there exists an $s_0 \in [0, k]$ such that

$$\int_{I_k(s)} f_k(r)\mathrm{d}r \leq \int_{I_k(s_0)} f_k(r)\mathrm{d}r \quad \forall\ s \geq 0. \tag{4.32}$$

Further, since $\|f_k\|_1 = 1$, for every $\epsilon > 0$ there exists $N > 0$ such that $\|\chi_{[N,\infty[} f_k\|_1 \leq \epsilon$. Thus, we may apply Corollary A.2 to the set $E \cap [0, N[$ to produce an interval of measure $|E \cap [0, N[|$ contained in $I_k(s_0)$, for which we conclude

$$\int_E f_k(r)\mathrm{d}r \leq \int_{I_k(s_0)} f_k(r)\mathrm{d}r + \epsilon.$$

□

It is evident that in order to maximize the integral of f_k over some interval $I_k(s) = [s, s+|E|[$, the interval must contain the maximum point k. However, any such interval will never be centered at k. Instead, for an interval $I_k(s_0)$ that satisfies (4.32), we always have

$$|I_k(s_0) \cap [0, k]| < |I_k(s_0) \cap [k, \infty[|.$$

This qualitative observation can be traced back to Lemma 4.3. Nevertheless, as it turns out, the next *upper* bound estimate for the operator norm only requires Lemma 4.4.

Proposition 4.2. Let E be a measurable subset of $\mathbb{R}_+$. Then

$$\int_0^{|E|} f_0(r)\mathrm{d}r = 1 - e^{-|E|} \geq \int_E f_k(r)\mathrm{d}r \quad \text{for } k = 0, 1, 2, \ldots \tag{4.33}$$

Proof. Since f_k is normalized and positive, we may assume $|E| = L < \infty$. By Lemma 4.4, we may further assume E is an interval of finite length $L > 0$, i.e. $E = [s, s+L[$ for some $s \geq 0$. From here we consider the following quantity

$$g_k(L, s) = \int_0^L f_0(r)\mathrm{d}r - \int_s^{s+L} f_k(r)\mathrm{d}r.$$

Note that the proposition holds if and only if $g_k(L,s) \geq 0$ for all $L, s \geq 0$. Since this holds trivially for $k = 0$, we will presume $k > 0$ in the subsequent calculations. Consider the derivative of g_k with respect to L, which, by the Fundamental Theorem of Calculus, becomes

$$\frac{\partial g_k}{\partial L}(L,s) = f_0(L) - f_k(s+L) = e^{-L}\left(1 - e^{-s}\frac{(s+L)^k}{k!}\right).$$

It is evident that the derivative equals zero only when $L = L_0$, where L_0 solves the equation

$$e^{-s}\frac{(s+L_0)^k}{k!} = 1. \tag{4.34}$$

The second derivative is given by

$$\frac{\partial^2 g_k}{\partial L^2}(L,s) = -e^{-L}\left(1 - e^{-s}\frac{(s+L)^k}{k!} + e^{-s}\frac{(s+L)^{k-1}}{(k-1)!}\right),$$

which evaluated at the solution of (4.34) yields

$$\frac{\partial^2 g_k}{\partial L^2}(L = L_0, s) = -e^{-(L_0+s)}\frac{(s+L_0)^{k-1}}{(k-1)!} < 0 \quad \forall\ s \geq 0.$$

Thus, by the second derivative test, the point $L = L_0$ represents a maximum for $g_k(L,s)$ for any fixed $s \geq 0$. The other possible extrema (minima) occur at $L = 0$ and when $L \to \infty$. Since both of these instances can be shown to be greater or equal to zero, either by direct evaluation or the normalization-condition $\|f_k\|_1 = 1$, we have that $g_k(L,s) \geq 0$ for all $L, s \geq 0$.

□

Recall that by formula (4.10) for the eigenvalues, the above proposition contains an upper bound estimate for the operator norm of P_E. In particular,

$$\|P_E\|_{\text{op}} \leq \int_0^{\pi|E|} f_0(r)\mathrm{d}r = 1 - e^{-\pi|E|} = \|P_{[0,|E|]}\|_{\text{op}}. \tag{4.35}$$

From here it is clear that the best localization on some spherically symmetric subset $E \subseteq \mathbb{R}_+$ with measure $|E|$ is obtained when we consider a disk of radius $\sqrt{|E|}$ centered at the origin. As remarked in Example 4.1, the best concentration on such a disk is obtained when the signal equals the normalized Gaussian. This latest result then corresponds well with the fact that a Gaussian signal and window minimize Lieb's inequality in Lemma 3.4.

In the next proposition we present a *lower* bound estimate for the norm.

Proposition 4.3. Let E be a measurable subset $\mathbb{R}_+$ contained in the interval $[0, L]$. Then

$$\|P_E\|_{\text{op}} \geq e^{-\pi(L-|E|)} \|P_{[0,|E|]}\|_{\text{op}}. \tag{4.36}$$

Proof. We always have

$$\|P_E\|_{\text{op}} \geq \lambda_0 = \int_{\pi \cdot E} f_0(r) \mathrm{d}r.$$

Since the integrand $f_0(r) = e^{-r}$ is monotonically decreasing on $[0, L]$, we may apply Proposition A.1 to the above integral such that

$$\int_{\pi \cdot E} f_0(r) \mathrm{d}r \geq \int_{\pi(L-|E|)}^{\pi L} f_0(r) \mathrm{d}r = e^{-\pi(L-|E|)} \int_0^{\pi |E|} f_0(r) \mathrm{d}r.$$

□

Note that both of these estimates for the norm are quite generic as they only require the measure and some upper bound for E. To summarize, if $E \subseteq [0, L]$ is measurable, then

$$\|P_{[0,|E|]}\|_{\text{op}} e^{-\pi(L-|E|)} \leq \|P_E\|_{\text{op}} \leq \|P_{[0,|E|]}\|_{\text{op}}. \tag{4.37}$$

4.2.3 Set of Infinite Measure: Equidistant Intervals

Now, consider the case when the subset $E \subseteq \mathbb{R}_+$ has infinite measure. For such subsets, it should be evident that the estimates of the previous section provide no useful information regarding the operator norm $\|P_E\|_{\text{op}}$. In particular, we obtain $\|P_E\|_{\text{op}} \in [0, 1]$, which, by Proposition 2.4 (A), holds for *any* localization operator with weight F (not only spherically symmetric ones) such that $\|F\|_\infty \leq 1$.

In what follows, we will consider a non-trivial example of a subset with infinite measure where it is possible to attain more precise estimates of the operator norm. The subset is constructed such as to resemble the time set in Example 3.2 for the separate time-frequency representation.

Example 4.3. (Equidistant Intervals) Suppose $E \subseteq \mathbb{R}_+$ is of the form

$$E(s) = \bigcup_{n=0}^{\infty} \frac{1}{\pi} \cdot [n, n+s] \quad \text{for } s \in [0, 1]. \tag{4.38}$$

Our goal is to show that there exists constants $0 \le C_1 \le C_2$ such that the associated localization operator P_E satisfies

$$C_1 s \le \|P_E\|_{\mathrm{op}} \le C_2 s \quad \forall\ s \in [0,1]. \tag{4.39}$$

We begin by examining the eigenvalues, which, by formula (4.10), becomes

$$\lambda_k(s) = \int_{\pi \cdot E(s)} f_k(r)\mathrm{d}r = \sum_{n=0}^{\infty} \int_n^{n+s} f_k(r)\mathrm{d}r \quad \text{for } k = 0, 1, 2, \dots, \tag{4.40}$$

where λ_k (as always) refers to the eigenvalue associated with the k-th Hermite function. For each integral over $[n, n+s]$, we apply the same zero order approximation for the integrands f_k as in Example 4.2. Since $[n, n+s] \subseteq [n, n+1]$, we consider the maximum of $f_k(r)$ for $r \in [n, n+1]$, which yields

$$\lambda_0(s) \le s \sum_{n=0}^{\infty} f_0(n) = s \sum_{n=0}^{\infty} e^{-n} = \frac{s}{1 - e^{-1}} \tag{4.41}$$

and

$$\lambda_k(s) \le s \left(f_k(k) + \sum_{n=0}^{\infty} f_k(n) \right) \quad \text{for } k = 1, 2, 3, \dots \tag{4.42}$$

We now claim that the following inequality holds

$$f_k(k) + \sum_{n=0}^{\infty} f_k(n) \le \sum_{n=0}^{\infty} f_0(n) \quad \text{for } k = 1, 2, 3, \dots \tag{4.43}$$

For $k = 1$, we compute the series explicitly by

$$\begin{aligned} \sum_{n=0}^{\infty} f_k(n) = \sum_{n=0}^{\infty} n e^{-n} &= \left(-\frac{\partial}{\partial \beta}\right) \sum_{n=0}^{\infty} e^{-n\beta} \Big|_{\beta=1} \\ &= \left(-\frac{\partial}{\partial \beta}\right) \frac{1}{1 - e^{-\beta}} \Big|_{\beta=1} = \frac{e}{(e-1)^2}, \end{aligned}$$

from which it is straightforward to verify (4.43). For $k \ge 2$, we compare the series with the integral over $\mathbb{R}_+$, that is

$$\sum_{n \ne k} f_k(n) \le \int_0^{\infty} f_k(r)\mathrm{d}r = 1.$$

Thus,

$$f_k(k) + \sum_{n=0}^{\infty} f_k(n) \leq 1 + 2f_k(k) \leq 1 + 2f_2(2) = 1 + 4e^{-2} \quad \text{for } k = 2, 3, \ldots$$

With this common upper bound, claim (4.43) readily follows for all $k \geq 1$. Combine the recent calculations with the bounds (4.41)-(4.43) to conclude that for any $k \geq 1$

$$\lambda_k(s) \leq s \max\left\{\frac{e}{(e-1)^2} + e^{-1}, 1 + 4e^{-2}\right\} < s\frac{e}{e-1}, \tag{4.44}$$

and

$$\lambda_k(s) \leq s\frac{e}{e-1} \quad \text{for every } k = 0, 1, 2, \ldots \tag{4.45}$$

Note that the right-hand side of the above inequality represent the upper bound estimate of the operator norm $\|P_E\|_{\mathrm{op}}$. For a lower bound estimate of the operator norm, it is sufficient to find a lower bound estimate for any of the eigenvalues. In particular, since $f_0(r)$ is monotonically decreasing,

$$\lambda_0(s) \geq s \sum_{n=1}^{\infty} f_0(n) = s\left(\frac{e}{e-1} - 1\right) = \frac{s}{e-1}. \tag{4.46}$$

We summarize with a proposition.

Proposition 4.4. Let $E(s) \subseteq \mathbb{R}_+$ be as in (4.38) with $s \in [0, 1]$. Then the operator norm of $P_{E(s)}$ satisfies the bounds

$$(C-1)s \leq \|P_{E(s)}\|_{\mathrm{op}} \leq \min\{Cs, 1\} \quad \forall\, s \in [0,1] \quad \text{with } C = \frac{e}{e-1}. \tag{4.47}$$

Further, there exists $s_0 > 0$ such that

$$\|P_{E(s)}\|_{\mathrm{op}} = (1 - e^{-s})C \quad \forall\, 0 < s < s_0. \tag{4.48}$$

Proof. The first part is just a restatement of identities (4.45) and (4.46). For the second part, note that

$$\lambda_0(s) = \sum_{n=0}^{\infty} \int_n^{n+s} e^{-r} \mathrm{d}r = (1 - e^{-s}) \sum_{n=0}^{\infty} e^{-n} = (1 - e^{-s})C.$$

By inequality (4.44), there exists a constant $0 < C_0 < C$ such that $\lambda_k(s) \leq C_0 s$ for any $k, s > 0$. Since $1 - e^{-s} \nearrow s$ as $s \to 0$, it then follows that some $s_0 > 0$ with property (4.48) must exist.

□

4.3 Localization on Cantor Set

As the headline suggests, we will in this section consider the case when weight equals or, more precisely, approaches the characteristic function of a Cantor set in the spherically symmetric sense. In measure theory, the Cantor set is a well-known example of a set with the same cardinality as the reals, i.e. uncountable but with zero measure. For this reason, no positive part of the signal may be concentrated on such a set. However, these type of sets are constructed by an iterative scheme, where each iterate has positive measure. Hence, instead of localizing on the Cantor set directly, we consider the n-iterate of measure $\epsilon_n > 0$, and compare the asymptotic behaviour to other sets with the same measure. Naturally, we will compare the operator norm of the n-iterate to the disk $[0, \epsilon_n]$ and the ring $[R, R + \epsilon_n]$ for some $R > 0$.

To begin with, we recapitulate the construction of the *mid-third Cantor set* on the real line, starting at zero. Start with an interval $C_0(R) = [0, R]$. Then remove the interior mid-third interval to produce the first iterate $C_1(R) = R \cdot [0, 1/3] \cup R \cdot [2/3, 1]$. Similarly, remove the interior mid-third interval to each of the two intervals $R \cdot [0, 1/3], R \cdot [2/3, 1]$ to produce $C_2(R)$. Continue this procedure inductively such that the n-iterate mid-third Cantor set becomes

$$C_n(R) = \bigcup_{a_1,\ldots,a_n=0,2} R \cdot \left[\sum_{j=1}^{n} \frac{a_j}{3^j}, \sum_{j=1}^{n} \frac{a_j}{3^j} + \frac{1}{3^n} \right], \text{ for } n = 0, 1, 2, \ldots \tag{4.49}$$

See Figure 2 for an illustration of this process. It is clear from the above description that $C_{n+1}(R) \subseteq C_n(R)$ for each $n = 0, 1, 2, \ldots$ The mid-third Cantor set $C(R)$ on the interval $[0, R]$ is then defined as the intersection of all the n-iterates, i.e.

$$C(R) = \bigcap_{n=0}^{\infty} C_n(R). \tag{4.50}$$

Since the n-iterate has measure

$$|C_n(R)| = \left(\frac{2}{3} \right)^n R, \tag{4.51}$$

the Cantor set itself will evidently have zero measure. Alternatively, we can consider the mid-third Cantor set on the interval $[s, s + R]$ by performing an s-translation on the n-iterates in (4.49).

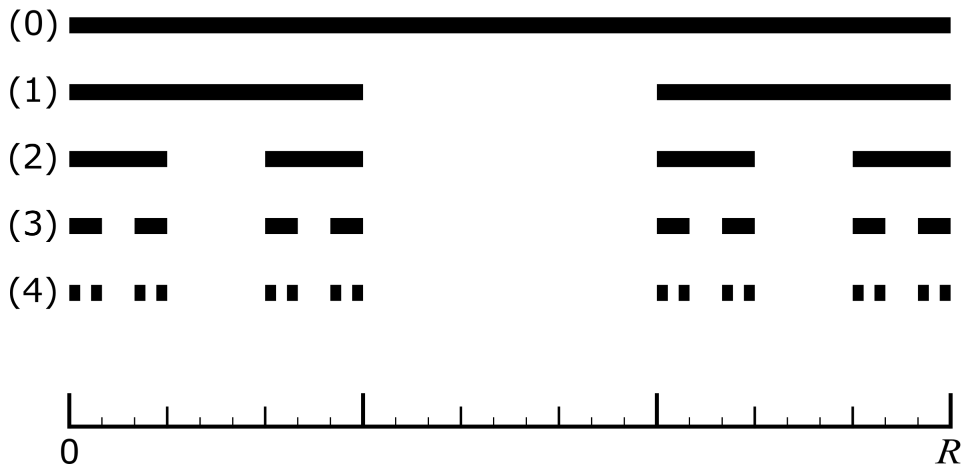

Figure 2: Illustration of the first 4 iterates of the mid-third Cantor set on the interval $[0, R]$. Here (n) refers to the n-iterate $C_n(R)$ for $n = 0, 1, 2, 3, 4$.

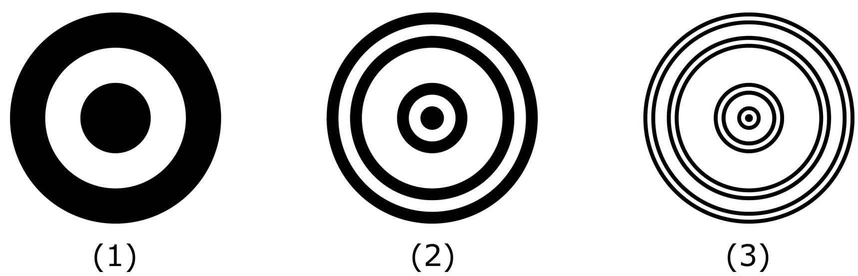

Figure 3: Illustration of the first 3 iterates of the mid-third distance-regular Cantor set on the disk. Note that the distances between the rings are the same as for the real line case. Here (n) refers to the n-iterate $\mathscr{C}_n^{(0)}$ for $n = 1, 2, 3$.

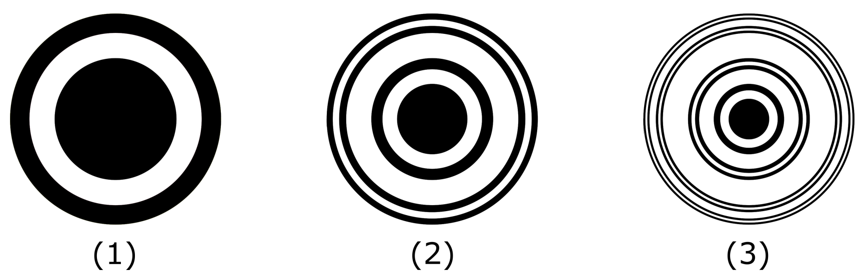

Figure 4: Illustration of the first 3 iterates of the mid-third measure-regular Cantor set on the disk. Note that the measure of each ring is the same. Here (n) refers to the n-iterate $\mathscr{C}_n$ for $n = 1, 2, 3$.

For the disk of radius $R > 0$ centered at the origin, there are two natural approaches to construct a similar mid-third Cantor set:

(i) The first scheme is to naively remove the interior part on the mid-third radius to produce the first iterate and then proceed inductively. Then $C_n(R)$ from the real line corresponds to the different radii included in the n-iterate for the disk, i.e. we identify $C_n(R)$ with the following subset of the plane

$$\mathscr{C}_n^{(0)}(R) = \{(\omega, t) \in \mathbb{R}^2 \mid \sqrt{\omega^2 + t^2} \in C_n(R)\}. \tag{4.52}$$

The procedure is illustrated in Figure 3, which reveals one caveat. Although the distances between each ring in $\mathscr{C}_n^{(0)}(R)$ coincides with the distances between the intervals in $C_n(R)$, the measures of each ring do not. The rings near the center of the disk will obviously have smaller measure than the rings close to the periphery. This behaviour does not resemble the real line case where the intervals have a fixed length.

(ii) Another possibility is to remove spherically symmetric parts of the disk in such a way that the remaining rings all have the same measure. This can be achieved by considering the set $C_n(R^2)$ for the radius *squared* (instead of $C_n(R)$ for the radius). To reiterate the comment at the beginning of section 4.2 regarding the spherically symmetric set $E \subseteq \mathbb{R}_+$, any point $r^2 \in C_n(R^2)$ corresponds to all points $(\omega, t) \in \mathbb{R}^2$ such that $\omega^2 + t^2 = r^2$. Hence, we identify $C_n(R^2)$ with the following part of the plane

$$\mathscr{C}_n(R) = \{(\omega, t) \in \mathbb{R}^2 \mid \omega^2 + t^2 \in C_n(R^2)\}. \tag{4.53}$$

With this transformation from points in $\mathbb{R}_+$ to rings in $\mathbb{R}^2$, it is evident that the intervals in $C_n(R^2)$ corresponds to rings in $\mathscr{C}_n(R)$ all with the same measure (but with different thickness). This latest scheme is showcased in Figure 4.

Based on these two approaches there is a clear trade-off between correct distances and correct measures when attempting to mimic the Cantor set on the real line now for a disk. Regularity in the distances between the rings, according to the real line case, will result in irregularity in the measures. Conversely, keeping the measures fixed will inevitably produce different distances between the rings than on the real line.

In what follows, we have prioritized regularity in the measures over regularity in distances and have therefore chosen to consider the second Cantor set construction for the disk. More precisely, we attempt to concentrate signals on the set $\mathscr{C}_n(R) \subseteq \mathbb{R}^2$ according to (4.53), which means we consider weights of the form

$$\mathscr{F}(r) = \chi_{C_n(R^2)}(r), \text{ for } R > 0 \text{ and } n = 0, 1, 2, \dots \tag{4.54}$$

With the same notation as in the previous section, the localization operators in question are $P_{C_n(R^2)}$. In this context we will attempt to derive asymptotic estimates for the operator norm for increasing iterates.

Observe that by (4.51), the measure of $C_n(R^2)$ equals $|C_n(R^2)| = (2/3)^n R^2$, from which Proposition 4.2 yields the upper bound

$$\mu_n(R) := 1 - \exp\left[-\left(\frac{2}{3}\right)^n \alpha(R)\right] \geq \|P_{C_n(R^2)}\|_{\mathrm{op}}, \text{ for } n = 0, 1, 2, \dots \tag{4.55}$$

Similarly to that of Example 4.1, we have used the abbreviation $\alpha(R) := \pi R^2$ and will continue to do so throughout the remainder of this chapter. Recall that the quantity $\mu_n(R)$ represents the best possible localization on any spherically symmetric subset $E \subseteq \mathbb{R}_+$ with measure $|E| = |C_n(R^2)|$. Thus, the ratio $\|P_{C_n(R^2)}\|_{\mathrm{op}}/\mu_n(R) \in [0, 1]$ is relevant as it measures to what extent the localization on $C_n(R^2) \subseteq \mathbb{R}_+$ (i.e. localization on $\mathscr{C}_n(R)$ in the plane) coincides with the *optimal* localization on a set of the *same* measure.

In order to produce sharper estimates, we will naturally consider the eigenvalues of the localization operator, namely $\lambda_k = \lambda_k(\mathscr{C}_n(R))$ for $k = 0, 1, 2, \dots$ Based on formula (4.10), the eigenvalues can be expressed

$$\lambda_k(\mathscr{C}_n(R)) = \int_{\pi \cdot C_n(R^2)} f_k(r)\mathrm{d}r = \int_{C_n(\alpha(R))} f_k(r)\mathrm{d}r, \tag{4.56}$$

with integrand f_k as in (4.11).

Note, however, that from the inductive scheme in (4.49), the number of intervals in $C_n(\cdot)$ grows as 2^n. Hence, estimating the eigenvalue $\lambda_k(\mathscr{C}_n(R))$ directly via the above integral or applying formula (4.21) for each interval seem quite impractical.

Instead, our main approach is to consider the effect locally of increasing from one iterate to the next. In particular, this means we initially consider the integral of f_k over a *single* interval, say $[s, s + 3L]$ for $s \geq 0$ and $L > 0$. Then we attempt to determine the relative area left under the curve f_k once the mid-third of the interval is removed, i.e. we wish to understand the function

$$\mathcal{A}_k(s, 3L) := \left[\int_s^{s+L} f_k(r)\mathrm{d}r + \int_{s+2L}^{s+3L} f_k(r)\mathrm{d}r\right] \Big/ \int_s^{s+3L} f_k(r)\mathrm{d}r. \tag{4.57}$$

By formula (4.21), this latest quantity can be written

$$\mathcal{A}_k(s,3L) = \left[\sum_{n=0}^{k} \frac{1}{n!}\big(s^n - e^{-L}(s+L)^n + e^{-2L}(s+2L)^n - e^{-3L}(s+3L)^n\big)\right] \cdot \Big/ \left[\sum_{n=0}^{k} \frac{1}{n!}\big(s^n - e^{-3L}(s+3L)^n\big)\right]. \tag{4.58}$$

From this expression, we note that $\mathcal{A}_k(s,3L)$ is independent of the starting point s precisely when $k=0$. Based on this simple observation, we derive upper and lower bound estimates (including an exact expression) for the first eigenvalue $\lambda_0(\mathscr{C}_n(R))$ in section 4.3.1. In the remaining subsections, we attempt to generalize these estimates to include the operator norm.

4.3.1 Estimates for the First Eigenvalue λ_0

We begin by making a few remarks on the relative area left under the curve f_0 when removing the mid-third of an interval $[s, s+3L]$, i.e. the quantity $\mathcal{A}_0(s,3L)$. By identity (4.58), $\mathcal{A}_0(s,3L)$ reads

$$\mathcal{A}_0(s,3L) = \frac{1 - e^{-L} + e^{-2L} - e^{-3L}}{1 - e^{-3L}}, \tag{4.59}$$

which, as previously noted, is independent of s. For simplicity, we will always write $\mathcal{A}_0(s,3L) =: \mathcal{A}_0(3L)$. Differentiating with respect to L, reveals

$$\frac{\partial}{\partial L}\big[\mathcal{A}_0(3L)\big] = \frac{e^{-L}\big(1+e^{-L}\big)\big(1-e^{-L}\big)^3}{\big(1-e^{-3L}\big)^2} > 0 \;\; \forall\, L > 0. \tag{4.60}$$

Hence, we conclude

$$\begin{aligned}\mathcal{A}_0(3L) &\geq \lim_{L\to 0} \mathcal{A}_0(3L)\\ &= \lim_{L\to 0} \frac{1 - e^{-L} + e^{-2L} - e^{-3L}}{1 - e^{-3L}} = \frac{2}{3} \text{ (by L'Hôpital's rule).}\end{aligned} \tag{4.61}$$

Recall that the $(n+1)$-iterate of the Cantor set is constructed by removing the mid-third of each interval included in the n-iterate. By (4.56) the eigenvalue $\lambda_0(\mathscr{C}_n(R))$ is obtained by integrating over the set $C_n(\alpha(R))$, which consists of 2^n intervals each of length $\alpha(R)/3^n$. Since $\mathcal{A}_0(s,\cdot)$ is independent of the starting point s, we attain the recursive relation

$$\lambda_0(\mathscr{C}_{n+1}(R)) = \mathcal{A}_0(\alpha(R)/3^n)\lambda_0(\mathscr{C}_n(R)) \text{ for } n = 0,1,2,\ldots \tag{4.62}$$

From the above relation, the first eigenvalue of the n-iterate of the Cantor set on a disk of radius $R > 0$ can readily be expressed as

$$\begin{aligned}\lambda_0(\mathscr{C}_n(R)) &= \lambda_0(\mathscr{C}_0(R)) \prod_{j=0}^{n-1} \mathcal{A}_0(\alpha(R)/3^j) \\ &= \left(1 - e^{-\alpha(R)}\right) \prod_{j=0}^{n-1} \mathcal{A}_0(\alpha(R)/3^j), \end{aligned} \tag{4.63}$$

where the final equality follows from formula (4.17) for the disk. Note that by result (4.61), we acquire the simple lower bound estimate

$$\lambda_0(\mathscr{C}_n(R)) \geq \left(\frac{2}{3}\right)^n \left(1 - e^{-\alpha(R)}\right). \tag{4.64}$$

In addition, we have the upper bound estimate $\mu_n(R)$ from (4.55).

To gain an impression of the sharpness of these two estimates, we consider the quotient of the lower and upper bound for large n. That is,

$$\left(\frac{2}{3}\right)^n \frac{1 - e^{-\alpha(R)}}{\mu_n(R)} \to \frac{1 - e^{-\alpha(R)}}{\alpha(R)} \quad \text{as } n \to \infty, \tag{4.65}$$

where we have utilized the Taylor series expansion $e^{-x} = 1 - x + \mathcal{O}(x^2)$ in the denominator. For small initial radii R the above limit is close to 1, while for large R the fraction tends to $\alpha(R)^{-1} (= (\pi R^2)^{-1})$. Thus, for small R, it is clear that both upper and lower bound represent quite sharp estimates for $\lambda_0(\mathscr{C}_n(R))$. In contrast, as R increases, the sharpness of either of these estimates remain somewhat ambiguous. Since $\mathcal{A}_0(\alpha(R)/3^j) \to 1$ as $R \to \infty$, the lower bound estimate will indeed deviate substantially from $\lambda_0(\mathscr{C}_n(R))$ for sufficiently large R and n. It is unclear whether the same is true for the upper bound estimate.

Proceeding, we will derive an improved asymptotic upper and lower bound estimate for the eigenvalue $\lambda_0(\mathscr{C}_n(R))$ (see Proposition 4.6). For this purpose, we shall express the relative area left under the curve f_0, i.e. $\mathcal{A}_0(3L)$, somewhat more conveniently.

Lemma 4.5. Suppose $s \geq 0$ and $L > 0$. Let $\mathcal{A}_0(3L)$ denote the relative area left under the curve f_0 once the mid-third of the interval $[s, s+3L]$ is removed, which is defined by (4.57). Then we have the identity

$$\mathcal{A}_0(3L) = \frac{\left(1 + e^{-2L}\right)\left(1 - e^{-L}\right)}{1 - e^{-3L}}. \tag{4.66}$$

Proof. Simply factorize expression (4.59) to obtain (4.66). □

With this new identity, we write down a more explicit expression for the eigenvalue $\lambda_0(\mathscr{C}_n(R))$ than in (4.63).

Proposition 4.5. Let $\lambda_0(\mathscr{C}_n(R))$ be the first eigenvalue to the localization operator $P_{C_n(R^2)}$. Then

$$\lambda_0(\mathscr{C}_n(R)) = \left(1 - e^{-\alpha(R)/3^n}\right) \prod_{j=1}^{n} \left(1 + e^{-2\alpha(R)/3^j}\right) \text{ for } n = 0, 1, 2, \ldots \quad (4.67)$$

Proof. Apply Lemma 4.5 to result (4.63) such that

$$\begin{aligned} \lambda_0(\mathscr{C}_n(R)) &= \left(1 - e^{-\alpha}\right) \prod_{j=1}^{n} \left(1 + e^{-2\alpha/3^j}\right) \left(\frac{1 - e^{-\alpha/3^j}}{1 - e^{-\alpha/3^{j-1}}}\right) \\ &= \left(1 - e^{-\alpha}\right) \left[\prod_{j=1}^{n} \left(1 + e^{-2\alpha/3^j}\right)\right] \left[\prod_{k=1}^{n} \left(\frac{1 - e^{-\alpha/3^k}}{1 - e^{-\alpha/3^{k-1}}}\right)\right] \quad (4.68) \end{aligned}$$

Observe that the product over k is a *telescoping product* such that, after cancellation, only the initial denominator and final numerator remain. That is,

$$\prod_{k=1}^{n} \left(\frac{1 - e^{-\alpha/3^k}}{1 - e^{-\alpha/3^{k-1}}}\right) = \frac{1 - e^{-\alpha/3^n}}{1 - e^{-\alpha}}.$$

Combine this with (4.68) to produce (4.67). □

We are now ready to formulate the improved asymptotic estimate for the first eigenvalue $\lambda_0(\mathscr{C}_n(R))$.

Proposition 4.6. Let $\mu_n(R)$ and $\lambda_0(\mathscr{C}_n(R))$ be as in equation (4.55) and (4.56), respectively. Then there exist positive, finite constants $c_1 \leq c_2$ such that for each $n = 0, 1, 2, \ldots$

$$\begin{aligned} c_1\Big(2\alpha(R) + 1\Big)^{-\ln 2/\ln 3} &\leq 2^{-n}\lambda_0(\mathscr{C}_n(R))/\left(1 - e^{-\alpha(R)/3^n}\right) \\ &\leq c_2\Big(2\alpha(R) + 1\Big)^{-\ln 2/\ln 3} \quad \forall\ \alpha(R) \in [0, 3^n/2]. \quad (4.69) \end{aligned}$$

In particular,

$$\begin{aligned} c_1\Big(2\alpha(R) + 1\Big)^{-\ln 2/\ln 3} &\leq \lim_{n\to\infty} \frac{\lambda_0(\mathscr{C}_n(R))}{\mu_n(R)} \\ &\leq c_2\Big(2\alpha(R) + 1\Big)^{-\ln 2/\ln 3} \quad \forall\ R > 0. \quad (4.70) \end{aligned}$$

Proof. By equation (4.67), the ratio

$$\lim_{n\to\infty} \frac{\lambda_0(\mathscr{C}_n(R))}{\mu_n(R)} = \lim_{n\to\infty} \left(\frac{1 - e^{-\alpha/3^n}}{1 - e^{-\alpha(2/3)^n}} \right) \prod_{j=1}^{n} \left(1 + e^{-2\alpha/3^j}\right)$$
$$= \lim_{n\to\infty} \prod_{j=1}^{n} \frac{1}{2} \left(1 + e^{-2\alpha(R)/3^j}\right).$$

Note that for each j in the above product, the factors converge to 1 at a geometric rate. Thus, for any fixed $\alpha(R) > 0$, the limit will indeed be defined and positive. For any n-iterate, we consider the product of the n first factors, which we exchange for a sum

$$\prod_{j=1}^{n} \frac{1}{2} \left(1 + e^{-2\alpha(R)/3^j}\right) = \exp\left[-n\ln(2) + \sum_{j=1}^{n} \ln\left(1 + e^{-2\alpha(R)/3^j}\right)\right].$$

Observe, the statement follows once we show that there exist finite constants $\tilde{c}_1 \le \tilde{c}_2$ such that

$$\tilde{c}_1 \le \sum_{j=1}^{n} \ln\left(1 + e^{-x/3^j}\right) - \left(n - \frac{\ln(x+1)}{\ln(3)}\right)\ln(2) \le \tilde{c}_2 \text{ for } x \in [0, 3^n]. \quad (4.71)$$

The proof of (4.71) is based on two claims

(i) $\sup_{y\in[0,1]} \sum_{j=1}^{\infty} \left[\ln\left(1 + y^{1/3^j}\right) - y^{1/3^j}\ln(2)\right] =: \beta < \infty$, and

(ii) there exist finite constants $\gamma_1 \le \gamma_2$ such that

$$\gamma_1 \le \sum_{j=1}^{n} e^{-x/3^j} - \left(n - \frac{\ln(x+1)}{\ln(3)}\right) \le \gamma_2 \text{ for } x \in [0, 3^n].$$

Although it is not particularly difficult to verify each claim, the arguments are somewhat technical. Details and precise arguments for claim (i) and (ii) are left in Appendix B.1 and B.2, respectively. Once these two claims are established, it becomes trivial to prove (4.71).

Since $\ln(1+y) - y\ln(2) \ge 0$ for all $y \in [0,1]$, claim (i) yields

$$0 \le \sum_{j=1}^{n} \ln\left(1 + e^{-x/3^j}\right) - \ln(2)\sum_{j=1}^{n} e^{-x/3^j} \le \beta.$$

Now for $x \in [0, 3^n]$, apply claim (ii) to the above inequality such that

$$\gamma_1 \ln(2) \leq \sum_{j=1}^{n} \ln\left(1 + e^{-x/3^j}\right) - \left(n - \frac{\ln(x+1)}{\ln(3)}\right) \ln(2) \leq \beta + \gamma_2 \ln(2).$$

Since any constants $\tilde{c}_1 \leq \gamma_1 \ln(2)$ and $\tilde{c}_2 \geq \beta + \gamma_2 \ln(2)$ work, we are done.

□

By close inspection of the above proof, it becomes clear that restriction $\alpha(R) \in [0, 3^n/2]$ in (4.69) is not a strict condition. This restriction is a result of the condition $x \in [0, 3^n]$ in claim (ii). If we had instead considered $x \in [0, 3^n M]$ for some fixed $M > 1$, we could have shown a similar identity as in claim (ii), but with different constants $\gamma_1(M) \leq \gamma_2(M)$. Since these constants would have to be updated according to our choice of M, we would not gain any additional information about the n-iterates where $\alpha(R) \geq 3^n/2$. More naturally is to consider n-iterates such that $\alpha(R) \ll 3^n/2$.

4.3.2 Upper Bound for the Second Eigenvalue $\lambda_1 \leq \lambda_0$

In Example 4.1 it was established that the localization operator for a disk produces eigenvalues $\{\lambda_k\}_k$ which are ordered according to the order of the associated eigenfunctions, namely the Hermite functions $\{H_k\}_k$. Since the subset $\mathscr{C}_0(R)$ equals a disk of radius $R > 0$ centered at the origin, we must have

$$\lambda_0(\mathscr{C}_0(R)) > \lambda_1(\mathscr{C}_0(R)) > \cdots > \lambda_k(\mathscr{C}_0(R)) > \lambda_{k+1}(\mathscr{C}_0(R)) > \ldots$$

Proceeding to the n-iterate, it remains an open question what are the necessary conditions under which $\lambda_0(\mathscr{C}_n(R))$ continues to be the largest eigenvalue. In this section we consider the second eigenvalue $\lambda_1(\mathscr{C}_n(R))$ and present a simple proof that this eigenvalue is always bounded by $\lambda_0(\mathscr{C}_n(R))$ for $n = 0, 1, 2, \ldots$ The statement itself is listed as Proposition 4.7.

The idea is to consider the relative area left under the curve f_1 for a single interval $[s, s + 3R]$ after the mid-third is removed. We then compare this area to the relative area under the curve f_0, from which we attain a constant ordering presented in the next lemma.

Lemma 4.6. Let $\{\mathcal{A}_k\}_k$ be given by (4.57). Then

$$\mathcal{A}_1(s, 3L) \leq \mathcal{A}_0(3L) \quad \forall\ s, L \geq 0. \tag{4.72}$$

Proof. By formula (4.58), $\mathcal{A}_1(s,3L)$ can be written

$$\mathcal{A}_1(s,3L) = \Big[(1+s) - e^{-L}(1+s+L) + e^{-2L}(1+s+2L) \\ - e^{-3L}(1+s+3L)\Big] \Big/ \Big[(1+s) - e^{-3L}(1+s+3L)\Big].$$

Differentiating with respect to s, yields, after factorization,

$$\frac{\partial \mathcal{A}_1}{\partial s}(s,3L) = \frac{Le^{-L}\left(1-e^{-2L}\right)\left(1-e^{-L}\right)^2}{\left((1+s)-e^{-3L}(1+s+3L)\right)^2} > 0 \ \ \forall\ L>0,\ \ s \geq 0.$$

Hence,

$$\mathcal{A}_1(s,3L) \leq \lim_{s\to\infty} \mathcal{A}_1(s,3L) = \frac{1-e^{-L}+e^{-2L}-e^{-3L}}{1-e^{-3L}},$$

where we identify, by (4.59), the right-hand side as $\mathcal{A}_0(3L)$.

□

From here it becomes easy to verify the ordering of the first two eigenvalues.

Proposition 4.7. Let $\{\lambda_k(\mathscr{C}_n(R))\}_k$ be the eigenvalues of $P_{C_n(R^2)}$ given by (4.56). Then

$$\lambda_1(\mathscr{C}_n(R)) \leq \lambda_0(\mathscr{C}_n(R)) \ \ \forall\ R>0 \text{ and } n = 0,1,2,\ldots \tag{4.73}$$

Proof. By a similar argument to how the relation (4.62) was acquired, we now apply Lemma 4.6 to obtain the recursive *inequality*

$$\lambda_1(\mathscr{C}_{n+1}(R)) \leq \mathcal{A}_0(\alpha(R)/3^n)\lambda_1(\mathscr{C}_n(R)), \text{ for } n = 0,1,2,\ldots,$$

from which we have

$$\begin{aligned}\lambda_1(\mathscr{C}_n(R)) &\leq \lambda_1(\mathscr{C}_0(R)) \prod_{j=1}^{n-1} \mathcal{A}_0(\alpha(R)/3^j) \\ &\leq \lambda_0(\mathscr{C}_0(R)) \prod_{j=1}^{n-1} \mathcal{A}_0(\alpha(R)/3^j) = \lambda_0(\mathscr{C}_n(R)).\end{aligned}$$

The final inequality follows by the constant ordering of the eigenvalues for the disk.

□

The crucial point of Lemma 4.6 is that the relative area under f_1 is less than the relative area under f_0 *regardless* of the starting point $s \geq$ and the length $3L > 0$ of the initial interval $[s, s+3L]$. Should the same constant ordering hold when we exchange f_1 with f_k for some $k > 1$, we would arrive at a similar conclusion as in Proposition 4.7, namely $\lambda_k(\mathscr{C}_n(R)) \leq \lambda_0(\mathscr{C}_n(R))$. However, as it turns out, for $k > 1$ the ordering of the relative areas is no longer independent of both s and L.

We illustrate by a simple counter-example:

Example 4.4. (Counter-example to Lemma 4.6 for $k > 1$)
Consider the relative area under the curve f_k for $k > 1$ with the initial interval $[0, 3L]$, i.e. consider

$$\mathcal{A}_k(0, 3L) = \left[\int_0^L f_k(r)\mathrm{d}r + \int_{2L}^{3L} f_k(r)\mathrm{d}r\right] \Big/ \int_0^{3L} f_k(r)\mathrm{d}r.$$

Further, let the length L tend to zero. By L'Hôpital's rule and the Fundamental Theorem of Calculus,

$$\begin{aligned}\lim_{L\to 0} \mathcal{A}_k(0, 3L) &= \lim_{L\to 0} \Big(f_k(L) - 2f_k(2L) + 3f_k(3L)\Big) \Big/ \Big(3f_k(3L)\Big) \\ &= \lim_{L\to 0} \Big(e^{-L} - 2^{k+1}e^{-2L} + 3^{k+1}e^{-3L}\Big) \Big/ \Big(3^{k+1}e^{-3L}\Big) \\ &= \frac{1 - 2^{k+1} + 3^{k+1}}{3^{k+1}}.\end{aligned}$$

Note that for $k = 0, 1$ this limit equals $2/3$, while for $k > 1$ it is clear that this limit is strictly greater than $2/3$. Therefore, by continuity of $\mathcal{A}_k$, for any $k > 1$ there exists $L_k > 0$ such that

$$\mathcal{A}_k(0, 3L) > \mathcal{A}_0(3L) \ \ \forall\, 0 < L < L_k.$$

4.3.3 Upper Bound Estimate for the Operator Norm

In this section we finally determine an *upper* bound estimate for the operator norm of $P_{C_n(R^2)}$ for each $n = 0, 1, 2, \ldots$ Since every eigenvalue $\lambda_k(\mathscr{C}_n(R))$ satisfies

$$\lambda_k(\mathscr{C}_n(R)) \leq \|P_{C_n(R^2)}\|_{\mathrm{op}} \text{ for } k = 0, 1, 2, \ldots, \tag{4.74}$$

we already have a *lower* bound estimate, e.g., the lower bounds estimates for $\lambda_0(\mathscr{C}_n(R))$ in section 4.3.1.

Returning to the upper bound, for a small initial radius $R > 0$, it is clear that $\lambda_0(\mathscr{C}_n(R))$ represents the largest eigenvalue. By "small" we mean any radius R such that $\alpha(R) = \pi R^2 \leq 1$, where the monotonicity of the integrals (4.56) guarantees such an ordering. This notion can further be extended, by Proposition 4.7, to include $\alpha(R) \leq 2$.

However, for a larger radius, we have not verified whether $\lambda_0(\mathscr{C}_n(R))$ remains the largest eigenvalue or not. In particular, the counter-example 4.4 shows that we cannot utilize the relative areas $\{\mathcal{A}_k(s, 3L)\}_k$ *directly* to determine which eigenvalue is the largest for an arbitrary initial radius $R > 0$. Nevertheless, from these relative areas we will be able to show that the operator norm $\|P_{C_n(R^2)}\|_{\text{op}} \leq 2\lambda_0(\mathscr{C}_n(R))$ for all radii $R > 0$ and $n = 0, 1, 2, \ldots$ (see Proposition 4.9).

Initially, we compare the relative areas $\mathcal{A}_k(s, 3L)$ and $\mathcal{A}_0(3L)$ when $s \geq k$.

Proposition 4.8. Let $\{\mathcal{A}_k\}_k$ be given by (4.57). Then

$$\mathcal{A}_k(s, 3L) \leq \mathcal{A}_0(3L) \quad \forall\ s \geq k,\ L > 0 \text{ and } k = 0, 1, 2, \ldots \tag{4.75}$$

Proof. See Appendix C. □

Thus, for any shifted n-iterate Cantor set $C_n(\alpha(R)) + s + k$ for $s \geq 0$, we obtain the recursive inequality

$$\begin{aligned} \int_{C_{n+1}(\alpha(R))+s+k} f_k(r)\mathrm{d}r \leq & \mathcal{A}_0(\alpha(R)/3^n) \\ & \cdot \int_{C_n(\alpha(R))+s+k} f_k(r)\mathrm{d}r, \text{ for } n = 0, 1, 2, \ldots \end{aligned} \tag{4.76}$$

Furthermore, by Proposition 4.2,

$$\begin{aligned} \int_{C_0(\alpha(R))+s+k} f_k(r)\mathrm{d}r &= \int_{s+k}^{\alpha(R)+s+k} f_k(r)\mathrm{d}r \\ &\leq \int_0^{\alpha(R)} f_0(r)\mathrm{d}r = \lambda_0(\mathscr{C}_0(R)), \end{aligned}$$

such that

$$\begin{aligned} \int_{C_n(\alpha(R))+s+k} f_k(r)\mathrm{d}r &\leq \int_{C_0(\alpha(R))+s+k} f_k(r)\mathrm{d}r \cdot \prod_{j=0}^{n-1} \mathcal{A}_0(\alpha(R)/3^j) \\ &\leq \lambda_0(\mathscr{C}_0(R)) \prod_{j=0}^{n-1} \mathcal{A}_0(\alpha(R)/3^j) = \lambda_0(\mathscr{C}_n(R)). \end{aligned} \tag{4.77}$$

In the next lemma we relate the integrals of f_k over the shifted n-iterate Cantor sets to the non-shifted iterates.

Lemma 4.7. Let $R > 0$. Then for every fixed $k, n = 0, 1, 2, \dots$, we have

(A) $\displaystyle\int_{C_n(\alpha(R))\cap[k,\infty[} f_k(r)\mathrm{d}r \leq \int_{C_n(\alpha(R))+k} f_k(r)\mathrm{d}r$ and

(B) $\displaystyle\int_{C_n(\alpha(R))\cap[0,k]} f_k(r)\mathrm{d}r \leq \int_{C_n(\alpha(R))+k} f_k(r)\mathrm{d}r.$

Proof. For the first case (A), note that the integrand $f_k(r)$ is monotonically decreasing. For this reason, it suffices to verify the following inequality

$$|C_n(\alpha(R)) \cap [k, r]| \leq |(C_n(\alpha(R)) + k) \cap [k, r]| \quad \forall\, r \geq k. \tag{4.78}$$

Introduce the function

$$G_n(x) := \begin{cases} 0, & x < 0 \\ |C_n(\alpha(R)) \cap [0, x]|, & x \geq 0 \end{cases}$$

which is a non-normalized version of the n-iterate Cantor function (on the interval $[0, \alpha(R)]$). With this notation, the left-hand side of inequality (4.78) can be written

$$|C_n(\alpha(R)) \cap [k, r]| = G_n(r) - G_n(k)$$

and by shift-invariance, the right-hand side

$$|(C_n(\alpha(R)) + k) \cap [k, r]| = G_n(r - k).$$

Claim (4.78) can then be rephrased as

$$G_n(r) \leq G_n(k) + G_n(r - k) \quad \forall\, r \geq k, \tag{4.79}$$

i.e. it is sufficient to show that G_n is *subadditive*. This property is shown in Appendix D.

For the second case (B), consider the subset

$$\mathcal{R}_{n,k} := \{r \geq k \mid 2k - r \in C_n(\alpha(R)) \cap [0, k]\},$$

which denotes the reflection of elements in $C_n(\alpha(R)) \cap [0, k]$ about the point k. By Lemma 4.3, we have

$$\int_{C_n(\alpha(R))\cap[0,k]} f_k(r)\mathrm{d}r \leq \int_{\mathcal{R}_{n,k}} f_k(r)\mathrm{d}r.$$

Similarly to (A), it suffices to show that

$$|\mathcal{R}_{n,k} \cap [k,r]| \leq G_n(r-k) \quad \forall\, r \geq k. \tag{4.80}$$

By symmetry of the Cantor set, the reflected set $\mathcal{R}_{n,k}$ satisfies

$$\begin{aligned}|\mathcal{R}_{n,k} \cap [k,r]| &= |C_n(\alpha(R)) \cap [\alpha(R)-k, \alpha(R)+r-2k]| \\ &= G_n(\alpha(R)+r-2k) - G_n(\alpha(R)-k).\end{aligned}$$

Claim (4.80) then follows by subadditivity of G_n.

□

From this latest lemma and formula (4.56), we have that each eigenvalue $\lambda_k(\mathscr{C}_n(R))$ satisfies

$$\begin{aligned}\lambda_k(\mathscr{C}_n(R)) &= \int_{C_n(\alpha(R))\cap[0,k]} f_k(r)\mathrm{d}r + \int_{C_n(\alpha(R))\cap[k,\infty]} f_k(r)\mathrm{d}r \\ &\leq 2\int_{C_n(\alpha(R))+k} f_k(r)\mathrm{d}r.\end{aligned}$$

By inequality (4.77), this means

$$\lambda_k(\mathscr{C}_n(R)) \leq 2\lambda_0(\mathscr{C}_n(R)) \quad \text{for } k,n = 0,1,2,\dots \tag{4.81}$$

Since the operator norm of $P_{C_n(R^2)}$ equals its largest eigenvalue, and since these eigenvalues all share a common upper bound, we conclude in the following statement.

Proposition 4.9. Let $\lambda_0(\mathscr{C}_n(R))$ denote the first eigenvalue of the localization operator $P_{C_n(R^2)}$. Then the operator norm is bounded from above by

$$\|P_{C_n(R^2)}\|_{\mathrm{op}} \leq 2\lambda_0(\mathscr{C}_n(R)) \text{ for } n = 0,1,2,\dots \tag{4.82}$$

Hence, by the above inequality, the asymptotic estimates for $\lambda_0(\mathscr{C}_n(R))$ in Proposition 4.6 must also hold for the operator norm (although possibly with different constants).

Corollary 4.1. Let $\mu_n(R)$ be as in equation (4.55). Then there exist positive, finite constants $\beta_1 \leq \beta_2$ such that for each $n = 0,1,2,\dots$ the operator norm of $P_{C_n(R^2)}$ is bounded by

$$\begin{aligned}\beta_1\Big(2\alpha(R)+1\Big)^{-\ln 2/\ln 3} &\leq 2^{-n}\|P_{C_n(R^2)}\|_{\mathrm{op}}\big/\big(1-e^{-\alpha(R)/3^n}\big) \\ &\leq \beta_2\Big(2\alpha(R)+1\Big)^{-\ln 2/\ln 3} \quad \forall\, \alpha(R) \in [0,3^n/2]. \end{aligned} \tag{4.83}$$

In particular,

$$\beta_1\Big(2\alpha(R)+1\Big)^{-\ln 2/\ln 3} \leq \lim_{n\to\infty} \frac{\|P_{C_n(R^2)}\|_{\text{op}}}{\mu_n(R)} \leq \beta_2\Big(2\alpha(R)+1\Big)^{-\ln 2/\ln 3} \quad \forall\ R>0. \tag{4.84}$$

Proof. Simply combine the estimates of Proposition 4.6 with Proposition 4.9.
□

Assume now that the area of the initial disk $\alpha(R) \geq 1$. For such $\alpha(R)$ we may express identity (4.83) on a simpler and perhaps more comprehensible form.

Corollary 4.2. There exists positive, finite constants $\gamma_1 \leq \gamma_2$ such that for each $n = 0, 1, 2, \ldots$ the operator norm of $P_{C_n(R^2)}$ is bounded by

$$\gamma_1\left(\frac{2}{3}\right)^n R^{2-2\ln 2/\ln 3} \leq \|P_{C_n(R^2)}\|_{\text{op}} \leq \gamma_2\left(\frac{2}{3}\right)^n R^{2-2\ln 2/\ln 3} \quad \forall\ \pi R^2 \in [1, 3^n/2] \tag{4.85}$$

Proof. Based on identity (4.83), it is sufficient to determine positive, finite constants $a \leq b$ and $c \leq d$ such that

(i) $(1-e^{-x})/x \in [a,b] \quad \forall\ x \in [0,1/2]$

(ii) $1 + 1/x \in [c,d] \quad \forall\ x \geq 1$.

Since it is straightforward to show that conditions (i) and (ii) are satisfied whenever $a \leq 2(1-e^{-1/2}) < 1 \leq b$ and $c \leq 1 < 2 \leq d$, we are done.
□

Notice that we could have chosen a different threshold $\alpha(R) \geq \epsilon$ for some fixed $0 < \epsilon < 1$, but then we would have to update the constants $\gamma_1 \leq \gamma_2$ accordingly.

4.3.4 Comparisons and Concluding Remarks

We begin by comparing localization on the n-iterate Cantor set $\mathscr{C}_n(R)$ with localization on a ring of same measure and with inner radius $R > 0$.

Example 4.5. (Comparison Between the n-iterate Cantor Set and the Ring) From Example 4.2, we have that the operator norm of a ring with inner radius $R > 0$ and measure $a \leq 1$ is of the form

$$\|P_{[R^2,R^2+a\pi^{-1}[}\|_{\mathrm{op}} = \frac{1}{\pi\sqrt{2}}R^{-1}a + a \cdot \mathcal{O}(R^{-3}) \quad \text{as } R \to \infty. \tag{4.86}$$

Recall that the measure of the n-iterate spherically symmetric Cantor set $\mathscr{C}_n(R)$ is given by $(2/3)^n\alpha(R) =: a(n)$, which is less than 1 whenever

$$n \geq \frac{\ln(\alpha(R))}{\ln(3/2)} =: N(R). \tag{4.87}$$

Hence, for all such n we may compare the operator norm $\|P_{C_n(R^2)}\|_{\mathrm{op}}$ with the asymptotic estimate (4.86) for the operator norm of a the ring with measure $a(n)$ and inner radius R.

From (4.85), we obtain

$$\|P_{C_n(R^2)}\|_{\mathrm{op}} \leq \gamma R^{-2\ln 2/\ln 3}a(n) \quad \forall\ \alpha(R) \in [1, 3^n/2].$$

Thus, for large $R \gg 0$, we have that

$$\begin{aligned}\|P_{C_n(R^2)}\|_{\mathrm{op}} \Big/ \|P_{[R^2,R^2+a(n)\pi^{-1}}\|_{\mathrm{op}} &\leq \mathcal{O}(\alpha(R)^{1/2-\ln 2/\ln 3}) \\ &= \mathcal{O}(R^{1-2\ln 2/\ln 3}) \quad \forall\ n \geq N(R). \end{aligned} \tag{4.88}$$

Notice that once we require $n \geq N(R)$, the restriction $\alpha(R)(= \pi R^2) \leq 3^n/2$ in (4.85) is automatically satisfied. Since $\ln 2/\ln 3 \approx 0.6309\cdots > 1/2$, it is clear that the right-hand side of inequality (4.88) tends to zero as $R \to \infty$.

By Corollary 4.1, it is evident that localization on the spherically symmetric Cantor set is in essence a two-parameter problem, in terms of iterate n and initial radius R (or equivalently the measure of the initial disk $\alpha(R) = \pi R^2$). Similarly to what was done in Example 3.3, we reduce the number of parameters by making the choice of iterate n dependent on $\alpha(R)$.

Example 4.6. (Comparison with Dyatlov's Cantor Set Construction)
In Example 3.3 we considered the operator $\pi_{C_n(M)}Q_{C_n(M)}$, first projecting onto $C_n(M)$ in frequency and then projecting onto $C_n(M)$ in time. In terms of the time-frequency plane, this operator attempts to concentrate a signal in the following region of the plane $C_n(M)\times C_n(M) \subseteq \mathbb{R}^2$. Recall that for the regular n-iterate in $[0, M]$, Dyatlov restricts to iterates consisting of 2^n intervals I_j such that

(i) $|I_j| \sim \dfrac{1}{M}$, which in returns means

(ii) $|C_n(M) \times C_n(M)| \sim \left(\dfrac{4}{3}\right)^n$.

While these two conditions are equivalent for the Cartesian product, they refer to different aspects of localization on the n-iterate in time and frequency. In particular, condition (i) refers to the measure of any single interval included in the n-iterate in either time or frequency, and (ii) is the total measure of the relevant region of the phase space. For the spherical symmetric n-iterate $\mathscr{C}_n(R)$, we treat these conditions separately, which yields two natural options for how to make n and $\alpha(R)$ dependent.

(I) Consider n-iterates $\mathscr{C}_n(R)$ such that the rings included satisfy a similar measure scale condition as the intervals in (i). More precisely, let I_j denote any ring in $\mathscr{C}_n(R)$ and restrict to n such that $|I_j| \sim 1/\alpha(R)$. This leads to

$$\alpha(R) \sim 3^{\frac{n}{2}}. \tag{4.89}$$

Consider a family of n-iterates $\{E_n^{(\mathrm{I})} = C_n(R^2)\}_n$ in $\mathbb{R}_+$ such that the above condition (4.89) holds.[12] Since $3^{\frac{n}{2}} \leq 3^n/2$ whenever $n \geq 2$, we may apply estimate (4.83) such that

$$\|P_{E_n^{(\mathrm{I})}}\|_{\mathrm{op}} = \mathcal{O}\left(\left(\frac{2}{3}\right)^{\frac{n}{2}}\right) = \mathcal{O}\left(\exp\left[-\frac{1}{2}\ln\left(\frac{3}{2}\right)n\right]\right) \quad \text{as } n \to \infty. \tag{4.90}$$

(II) Let instead the measure of $\mathscr{C}_n(R)$ be the same as in (ii), which yields the condition

$$\alpha(R) \sim 2^n. \tag{4.91}$$

[12] Recall that the subset E_n corresponds to the spherically symmetric set in the plane $\mathcal{E}_n = \{(\omega, t) \in \mathbb{R}^2 \mid \omega^2 + t^2 \in E_n\}$.

Once again, consider a family of n-iterates $\{E_n^{(\mathrm{II})} = C_n(R^2)\}_n$ in $\mathbb{R}_+$ but now such that condition (4.91) is satisfied. Similarly, $2^n \leq 3^n/2$ whenever $n \geq 2$, from which we apply estimate (4.83) to produce

$$\begin{aligned}\|P_{E_n^{(\mathrm{II})}}\|_{\mathrm{op}} &= \mathcal{O}\left(\left(2^{2-\ln 2/\ln 3}/3\right)^n\right) \\ &= \mathcal{O}\Big(\exp\Big[-\Big(\ln 3 + (\ln 2)^2/\ln 3 - 2\ln 2\Big)n\Big]\Big) \quad \text{as } n \to \infty. \end{aligned} \tag{4.92}$$

Observe that in both cases (I) and (II), we end up with a situation where $\|P_{E_n^{(\cdot)}}\|_{\mathrm{op}} \to 0$ as $|E_n^{(\cdot)}| \to \infty$. In addition, we have precise estimates for the rate at which $\|P_{E_n^{(\cdot)}}\|_{\mathrm{op}}$ tends to zero, i.e. we know the exponential coefficient $\beta^{(\cdot)} > 0$ in $\|P_{E_n^{(\cdot)}}\|_{\mathrm{op}} = \mathcal{O}(e^{-\beta^{(\cdot)}n})$ for large n.

At this point it should come as no surprise that having explicit expressions for the eigenvalues at our disposal is a great advantage when attempting to estimate the operator norm. For Daubechies' localization operator with a spherically symmetric weight and Gaussian window, we have such expressions for the eigenvalues (on integral form, see (4.3)). By virtue of the numerous examples provided in this chapter, it is clear that the main challenge then becomes how to determine the largest eigenvalue or at least how to determine a common upper bound. Although not without computational effort, in the same examples we have been able to produce precise asymptotic estimates for the operator norm.

It is worth noting that in the computations for the mid-third spherically symmetric Cantor set, the specific regularity of the iterates has been of great benefit. Based on this observation, it seems relevant to ask whether similar results could hold for subsets with the same type of regularity. One natural extension could be to consider more general Cantor set constructions, e.g., the mid-$\frac{1}{L}$ Cantor set $C_n(R, L)$ for $L > 1$, where we remove the mid-interior $\frac{1}{L}$ inductively.

Appendices

A Monotonic Functions and Integral Bounds

The purpose of this section is to provide formal proofs of the rather intuitive statements regarding integral bounds of monotonic functions (either monotonically increasing or decreasing) integrated over some measurable set. The idea can be summarized as follows:

Proposition A.1. Let $f : [a, b] \to \mathbb{R}_+$ be an integrable, monotonically increasing function, and let E be a measurable subset of $[a, b]$ with measure $|E|$. Then

$$\int_a^{a+|E|} f(x)\mathrm{d}x \leq \int_E f(x)\mathrm{d}x \leq \int_{b-|E|}^{b} f(x)\mathrm{d}x. \tag{A.1}$$

Proof. We only show the first inequality in (A.1) as the proof for the second inequality is almost completely analogous. Since f is monotonically increasing, we have that

$$\begin{aligned}
\int_E f(x)\mathrm{d}x &= \int_{E\cap[a,a+|E|]} f(x)\mathrm{d}x + \int_{E\setminus[a,a+|E|]} f(x)\mathrm{d}x \\
&\geq \int_{E\cap[a,a+|E|]} f(x)\mathrm{d}x + \inf_{x\in E\setminus[a,a+|E|]} f(x)|E\setminus[a, a+|E|]| \\
&\geq \int_{E\cap[a,a+|E|]} f(x)\mathrm{d}x + f(a+|E|)|E\setminus[a, a+|E|]|.
\end{aligned}$$

Similarly, we obtain

$$\int_a^{a+|E|} f(x)\mathrm{d}x \leq \int_{E\cap[a,a+|E|]} f(x)\mathrm{d}x + f(a+|E|)|[a, a+|E|]\setminus E|.$$

By symmetry, $|E\setminus[a, a+|E|]| = |[a, a+|E|]\setminus E|$ such that the difference

$$\int_E f(x)\mathrm{d}x - \int_a^{a+|E|} f(x)\mathrm{d}x \geq 0.$$

□

We obtain the reverse inequality if the function is monotonically decreasing.

Corollary A.1. Let $g : [a, b] \to \mathbb{R}_+$ be an integrable, monotonically decreasing function, and let E be a measurable subset of $[a, b]$ with measure $|E|$. Then

$$\int_a^{a+|E|} g(x)\mathrm{d}x \geq \int_E g(x)\mathrm{d}x \geq \int_{b-|E|}^b g(x)\mathrm{d}x. \tag{A.2}$$

Proof. Simply recognize that $g(x) = f(-x)$ for some monotonically increasing function f.

□

Corollary A.2. Let $f : [a, b] \to \mathbb{R}_+$ be an integrable function, and let E be a measurable subset of $[a, b]$. Suppose there exist a point $x_0 \in [a, b]$ such that f is monotonically increasing in $[a, x_0]$ and decreasing in $[x_0, b]$. Then there exists an interval I with $|I| = |E|$ such that

$$\int_E f(x)\mathrm{d}x \leq \int_I f(x)\mathrm{d}x. \tag{A.3}$$

Proof. Apply inequality (A.1) to $E \cap [a, x_0[$ and inequality (A.2) to $E \cap]x_0, b]$ to produce the desired interval.

□

B Details on Series

In this section we present precise arguments for the two claims (i), (ii) set forth in Proposition 4.6, thus completing the proof. Details on claim (i) can be found in section B.1 while claim (ii) in section B.2.

B.1 Logarithmic Series

We shall prove the following statement:

Proposition B.1.

$$\sup_{y\in[0,1]} \sum_{j=1}^{\infty} \left[\ln\left(1 + y^{1/3^j}\right) - y^{1/3^j} \ln(2)\right] < \infty. \tag{B.1}$$

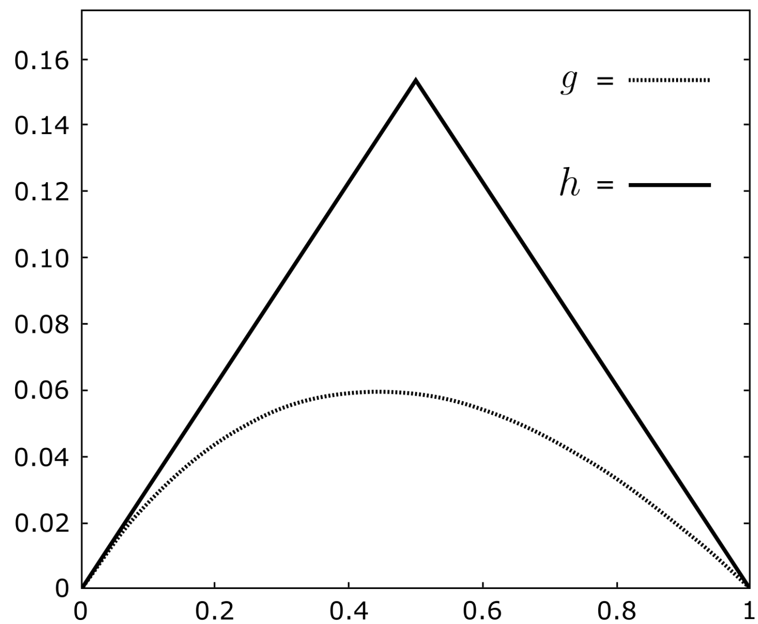

Figure 5: Plot of the difference $g(y)$ given by (B.2) for $y \in [0, 1]$. The linear spline h from (B.3) which bounds g from above is plotted alongside.

Proof. Initially, consider only the function

$$g(y) := \ln(1 + y) - y \ln(2) \text{ for } y \in [0, 1]. \tag{B.2}$$

It is easily verified that $g(y) \geq 0$ for all $y \in [0, 1]$. Furthermore, the slope of the function is bounded by the slope of the endpoints $y = 0, 1$, i.e.

$$|g'(y)| \leq \max\{|g'(0)|, |g'(1)|\} = 1 - \ln(2) =: a \;\; \forall\, y \in [0, 1].$$

Our function g may therefore be bounded from above by a linear spline, say h, defined by

$$h(y) := a \cdot \begin{cases} y, \;\; y \in [0, 1/2] \\ (1 - y), \;\; y \in]1/2, 1]. \end{cases} \tag{B.3}$$

Both the function g and the linear spline h are depicted in Figure 5.

Further, the sum in (B.1) is bounded from above by

$$\sum_{j=1}^{\infty} h\left(y^{1/3^j}\right) \;\text{ for } y \in [0, 1]. \tag{B.4}$$

Hence, it is sufficient to show that there exists a common upper bound for (B.4) for all $y \in [0,1]$.

Since the sum (B.1) (and (B.4)) is zero for $y = 0$, we may, without loss of generality, assume $y > 0$. In addition, for any $y \in]0,1]$ we have that $y^{1/3^j} \nearrow 1$ as $j \to \infty$. Thus, for any fixed $y \in]0,1]$ there exists a *smallest* $j_0 \in \mathbb{N}$ such that $y^{1/3^j} \geq 1/2$ for all $j \geq j_0$. With this notion, we *split* the sum to the effect

$$\begin{aligned}\sum_{j=1}^{\infty} h\left(y^{1/3^j}\right) &= \sum_{j=1}^{j_0-1} h\left(y^{1/3^j}\right) + \sum_{j=j_0}^{\infty} h\left(y^{1/3^j}\right) \\ &= a\sum_{j=1}^{j_0-1} y^{1/3^j} + a\sum_{j=j_0}^{\infty}\left(1 - y^{1/3^j}\right) \qquad \text{(B.5)}\end{aligned}$$

Consider each of these two sums separately.

Observe that the first sum is empty whenever $y^{1/3} \geq 1/2$. For the non-empty case, introduce the variable $z := y^{1/3^{j_0-1}} \in]0, 1/2[$. Then the first sum can be expressed

$$\sum_{j=1}^{j_0-1} y^{1/3^j} = \sum_{j=0}^{j_0-2} z^{3^j} \leq \sum_{j=0}^{\infty} z^{3^j} \qquad \text{(B.6)}$$

Since $z^{3^j} \leq z^j$ for any $j = 0, 1, 2, \ldots$, by direct comparison with the *geometric series*, we obtain

$$\sum_{j=0}^{\infty} z^{3^j} \leq \sum_{j=0}^{\infty} z^j = \frac{1}{1-z} \leq \frac{1}{1-1/2} = 2 \quad \forall\ z \in [0, 1/2]. \qquad \text{(B.7)}$$

Similarly to the first sum, introduce the variable $\tilde{z} := y^{1/3^{j_0}} \in [1/2, 1]$, now for the second sum, such that

$$\begin{aligned}\sum_{j=j_0}^{\infty}\left(1 - y^{1/3^j}\right) &= \sum_{j=0}^{\infty}\left(1 - \tilde{z}^{1/3^j}\right) \\ &\leq \sum_{j=0}^{\infty}\left(1 - (1/2)^{1/3^j}\right) \quad \forall\ \tilde{z} \in [1/2, 1]. \qquad \text{(B.8)}\end{aligned}$$

Inserting these two upper bounds (B.7), (B.8) into (B.5), yields

$$\sum_{j=1}^{\infty} h\left(y^{1/3^j}\right) \leq 2a + a\sum_{j=0}^{\infty}\left(1 - (1/2)^{1/3^j}\right) \quad \forall\ y \in [0, 1].$$

It only remains to verify that the right-hand side series converges. By the *ratio test*,

$$\lim_{j\to\infty} \frac{1-(1/2)^{1/3^{j+1}}}{1-(1/2)^{1/3^j}} = \frac{1}{3} < 1,$$

where we have identified $(1/2)^x = e^{-x\ln(2)}$ and applied the Taylor series expansion $e^{-x} = 1 - x + \mathcal{O}(x^2)$ in both numerator and denominator. From here we conclude that the series is in fact convergent.

□

B.2 Exponential Series

We shall prove the following statement:

Proposition B.2. There exists finite constants $\gamma_1 \leq \gamma_2$ such that

$$\gamma_1 \leq \sum_{j=1}^{n} e^{-x/3^j} - \left(n - \frac{\ln(x+1)}{\ln(3)}\right) \leq \gamma_2 \text{ for } x \in [0, 3^n]. \tag{B.9}$$

Proof. Initially, observe that for any fixed $x \geq 0$, we have that $x/3^j \leq 1$ precisely when $j \geq \ln(x)/\ln(3)$. Let $\lfloor\cdot\rfloor$ denote the floor function which rounds down any number to the nearest integer. Define

$$j_0(x) := \max\left\{\left\lfloor\frac{\ln(x)}{\ln(3)}\right\rfloor, 0\right\} \tag{B.10}$$

and consider the split

$$\sum_{j=1}^{n} e^{-x/3^j} = \sum_{j=1}^{j_0(x)} e^{-x/3^j} + \sum_{j=j_0(x)+1}^{n} e^{-x/3^j} \text{ for } x \in [0, 3^n]. \tag{B.11}$$

Note that the first sum is possibly empty, while the second sum is *always* non-empty except when $x = 3^n$. Nonetheless, since $e^{-1} \in [0, 1/2]$, we attain, by the same reasoning as in equations (B.6) and (B.7), the following bounds for the first sum

$$0 \leq \sum_{j=1}^{j_0(x)} e^{-x/3^j} \leq 2. \tag{B.12}$$

For the second sum, we utilize the fact that $1 - y \leq e^{-y} \leq 1$ for all $y \geq 0$. Thus, we acquire the upper bound

$$\sum_{j=j_0(x)+1}^{n} e^{-x/3^j} \leq n - j_0(x), \tag{B.13}$$

in addition to the lower bound

$$\begin{aligned}\sum_{j=j_0(x)+1}^{n} e^{-x/3^j} &\geq \sum_{j=j_0(x)+1}^{n} \left(1 - \frac{x}{3^j}\right) \\ &= n - j_0(x) - \frac{x}{3^{j_0(x)+1}} \sum_{j=0}^{n-j_0(x)-1} 3^{-j}.\end{aligned}$$

By comparison with the geometric series and by the fact that $x/3^{j_0(x)+1} \leq 1$, we further conclude

$$\begin{aligned}\sum_{j=j_0(x)+1}^{n} e^{-x/3^j} &\geq n - j_0(x) - \sum_{j=0}^{\infty} 3^{-j} \\ &= n - j_0(x) - \frac{3}{2}. \end{aligned} \tag{B.14}$$

Now, insert these three estimates (B.12)-(B.14) into (B.11) such that

$$-\frac{3}{2} \leq \sum_{j=1}^{n} e^{-x/3^j} - (n - j_0(x)) \leq 2 \ \ \forall\, x \in [0, 3^n].$$

Finally, the desired inequality (B.9) follows once we include the bounds

$$\frac{\ln(x+1)}{\ln(3)} - 1 \leq j_0(x) \leq \frac{\ln(x+1)}{\ln(3)}.$$

□

C Technical Proof of Proposition 4.8

In this section we provide a formal proof to Proposition 4.8, which, as it turns out, is a rather long and technical proof. Note that all the symbols correspond to the symbols used in the aforementioned proposition.

Proof. In the current context we will deal directly with the integral definition of $\mathcal{A}_k(s, 3L)$, namely definition (4.57):

$$\mathcal{A}_k(s,3L) = \left[\int_s^{s+L} f_k(r)\mathrm{d}r + \int_{s+2L}^{s+3L} f_k(r)\mathrm{d}r\right] \Big/ \int_s^{s+3L} f_k(r)\mathrm{d}r.$$

Differentiate said definition with respect to s, which, by the Fundamental Theorem of Calculus, yields

$$\frac{\partial \mathcal{A}_k}{\partial s}(s,3L) = N_k(s,L) \Big/ \left[\int_s^{s+3L} f_k(r)\mathrm{d}r\right]^2,$$

where

$$\begin{aligned} N_k(s,L) = \Big(f_k(s+L) - f_k(s) + f_k(s+3L) - f_k(s+2L)\Big) \int_s^{s+3L} f_k(r)\mathrm{d}r \\ - \Big(f_k(s+3L) - f_k(s)\Big)\left[\int_s^{s+3L} f_k(r)\mathrm{d}r + \int_{s+2L}^{s+3L} f_k(r)\mathrm{d}r\right]. \end{aligned} \tag{C.1}$$

From identity (4.58), it is clear that $\lim_{s\to\infty} \mathcal{A}_k(s,3L) = \mathcal{A}_0(\cdot, 3L)$. Hence, in order to prove the proposition, it is sufficient to show that $N_k(s,L) \geq 0$ for all $s \geq k,\ L > 0$ and $k = 1, 2, 3, \ldots$ [13]

Proceeding, we shall express $N_k(s,L)$ in (C.1) more conveniently. Firstly, observe that the integral over $[s, s{+}3L]$ can be split into three separate integrals over $[s, s+L],\ [s+L, s+2L]$ and $[s+2L, s+3L]$, respectively. On this form, certain terms cancel such that

$$\begin{aligned} N_k(s,L) = \Big(f_k(s+L) - f_k(s+2L)\Big)\left[\int_s^{s+L} + \int_{s+L}^{s+2L} + \int_{s+2L}^{s+3L}\right] f_k(r)\mathrm{d}r \\ - \Big(f_k(s) - f_k(s+3L)\Big)\int_{s+L}^{s+2L} f_k(r)\mathrm{d}r. \end{aligned}$$

[13] The proposition holds trivially for $k = 0$.

Secondly, we rewrite $N_k(s,L)$ as a single integral over $[0,L]$ by shifting the integrands $f_k(r)$ accordingly. After rearrangement, we obtain

$$\begin{aligned}N_k(s,L) = \int_0^L \Bigg\{ &\bigg[f_k(r+s)\Big(f_k(s+L) - f_k(s+2L)\Big) \\ &-f_k(r+s+L)\Big(f_k(s) - f_k(s+L)\Big)\bigg] \\ &-\bigg[f_k(r+s+L)\Big(f_k(s+2L) - f_k(s+3L)\Big) \\ &-f_k(r+s+2L)\Big(f_k(s+L) - f_k(s+2L)\Big)\bigg]\Bigg\}\mathrm{d}r.\end{aligned}$$

Based on the division into square brackets $[\dots]$, we introduce the function

$$\begin{aligned}\Phi_k(r,s,L) := f_k(r+s)&\Big(f_k(s+L) - f_k(s+2L)\Big) \\ -f_k(r+s+L)&\Big(f_k(s) - f_k(s+L)\Big),\end{aligned} \tag{C.2}$$

such that

$$N_k(s,L) = \int_0^L \Big(\Phi_k(r,s,L) - \Phi_k(r,s+L,L)\Big)\mathrm{d}r. \tag{C.3}$$

Observe that if the difference $\Phi_k(r,s,L) - \Phi_k(r,s+L,L)$ is positive for all $r \in [0,L]$ and $s \geq k$, then so is $N_k(s,L)$. Furthermore, this difference is *positive* if the derivative of $\Phi_k(r,s,L)$ with respect to s is *negative* whenever $s \geq k$. In what follows, we will verify that this particular partial derivative is indeed negative (where we assume the same restrictions on the variables r, s, L and k as outlined above).

Begin by rearranging the terms in definition (C.2) such that

$$\begin{aligned}\Phi_k(r,s,L) =&\Big[f_k(r+s)f_k(s+L) - f_k(r+s+L)f_k(s)\Big] \\ &+\Big[f_k(r+s+L)f_k(s+L) - f_k(r+s)f_k(s+2L)\Big].\end{aligned}$$

Consider the terms in each square bracket $[\dots]$ and define the function

$$\begin{aligned}\Psi_k(r,s,L,y) :=& f_k(r+s+y)f_k(s+L) \\ &- f_k(r+s+L-y)f_k(s+2y) \text{ for } y \in \{0,L\},\end{aligned} \tag{C.4}$$

where

$$\Phi_k(r,s,L) = \Psi_k(r,s,L,0) + \Psi_k(r,s,L,L). \tag{C.5}$$

The key observation here is that the arguments of $f_k(\cdot)$ in each term of Ψ_k add to a *fixed* value, namely

(i) $\quad 2a := 2s + r + L + y.$

In addition, define the corrections to each argument

(ii) $\quad \epsilon := \dfrac{L - r - y}{2},$

(iii) $\quad \delta := \dfrac{L + r - 3y}{2}.$[14]

Based on the quantities (i)-(iii), we now express Ψ_k on the form

$$\begin{aligned}\Psi_k(r, s, L, y) &= f_k(a - \epsilon) f_k(a + \epsilon) - f_k(a - \delta) f_k(a + \delta) \\ &= \frac{1}{(k!)^2} e^{-2a} \left[\left(a^2 - \epsilon^2 \right)^k - \left(a^2 - \delta^2 \right)^k \right] \text{ (by definition (4.11)).}\end{aligned}$$

Since $a = a(s)$ is the only quantity that depends on the parameter s, we easily differentiate the latest result to obtain

$$\begin{aligned}\frac{\partial \Psi_k}{\partial s}(r, s, L, y) = \frac{2}{(k!)^2} e^{-2t} \Big[&\left(a^2 - \epsilon^2 \right)^{k-1} \left(-a^2 + \epsilon^2 + ka \right) \\ &- \left(a^2 - \delta^2 \right)^{k-1} \left(-a^2 + \delta^2 + ka \right) \Big]. \end{aligned} \tag{C.6}$$

By (i)-(iii) and that $r \leq L$, we always have the inequalities

$$|\epsilon(y)| \leq |\delta(y)| \leq a(s, y) \quad \forall\ s \geq k \text{ and } y \in \{0, L\}. \tag{C.7}$$

From this ordering, it is clear that expression (C.6) is negative whenever the factor $-a^2 + \epsilon^2 + ka$ is negative.

By inspection of (i) and (ii), note that since $s \geq k$, we must have

$$a(s, y) - k \geq a(s, y) - s = \frac{r + L + y}{2} \geq |\epsilon(y)| \text{ for } y \in \{0, L\}.$$

Completing the square then reveals

$$-a^2 + \epsilon^2 + ka = -\left(a - \frac{k}{2} \right)^2 + \epsilon^2 + \left(\frac{k}{2} \right)^2 \leq 0.$$

[14] For simplicity with regard to notation, we have avoided writing these three quantities as functions of the variables s, y, k, L and r. Nevertheless, they should always be thought of as functions of such variables (e.g. $a = a(s, y, k, L, r)$).

In total, by (C.5), this means

$$\frac{\partial \Phi_k}{\partial s}(r,s,L) \leq 0 \ \ \forall\ s \geq k,\ \ L > 0 \text{ and } r \in [0,L],$$

for which we are done.

□

D Cantor Function

This section completes the proof of Lemma 4.7, where we formally introduce the family of n-iterate Cantor functions $\{\mathcal{G}_n\}_n$ and show that every member of said family is subadditive. This is based heavily on and structured similarly to the short 1996-paper by Jozef Doboš [23], which proves this exact property by induction.

Consider the n-iterate Cantor set $C_n(1)$ on the interval $[0,1]$ according to definition (4.49). Then we may define a corresponding map $\mathcal{G}_n : \mathbb{R} \to [0,1]$ by

$$\mathcal{G}_n(x) = \frac{1}{|C_n(1)|} \cdot \begin{cases} 0,\ x \leq 0, \\ |C_n(1) \cap [0,x]|,\ x > 0 \ \text{ for } n = 0,1,2,\dots, \end{cases} \tag{D.1}$$

which we refer to as the (standard) n-iterate Cantor function. From the above expression, we may think of $\mathcal{G}_n(x)$ as the *amount* of the n-iterate contained in the interval $]-\infty, x]$. Thus, every $\mathcal{G}_n$ is obviously monotonically increasing. Further, as $n \to \infty$, the sequence $\mathcal{G}_n$ converges uniformly to some continuous function $\mathcal{G}$, known as the *Cantor function*. In Figure 6 the first few iterates $\mathcal{G}_n$ for $n = 0,1,2,3$ are plotted. In the subsequent Figure 7 an approximation of the Cantor function is shown.

For the purpose of verifying subadditivity, the sequence $\{\mathcal{G}_n\}_n$ can alternatively be expressed recursively by

$$\mathcal{G}_0(x) = \begin{cases} 0,\ x < 0, \\ x,\ x \in [0,1], \\ 1,\ x > 1, \end{cases} \tag{D.2}$$

and

$$\mathcal{G}_{n+1}(x) = \frac{1}{2} \cdot \begin{cases} \mathcal{G}_n(3x),\ x \leq \frac{2}{3}, \\ 1 + \mathcal{G}_n(3x-2),\ x \geq \frac{1}{3}. \end{cases} \tag{D.3}$$

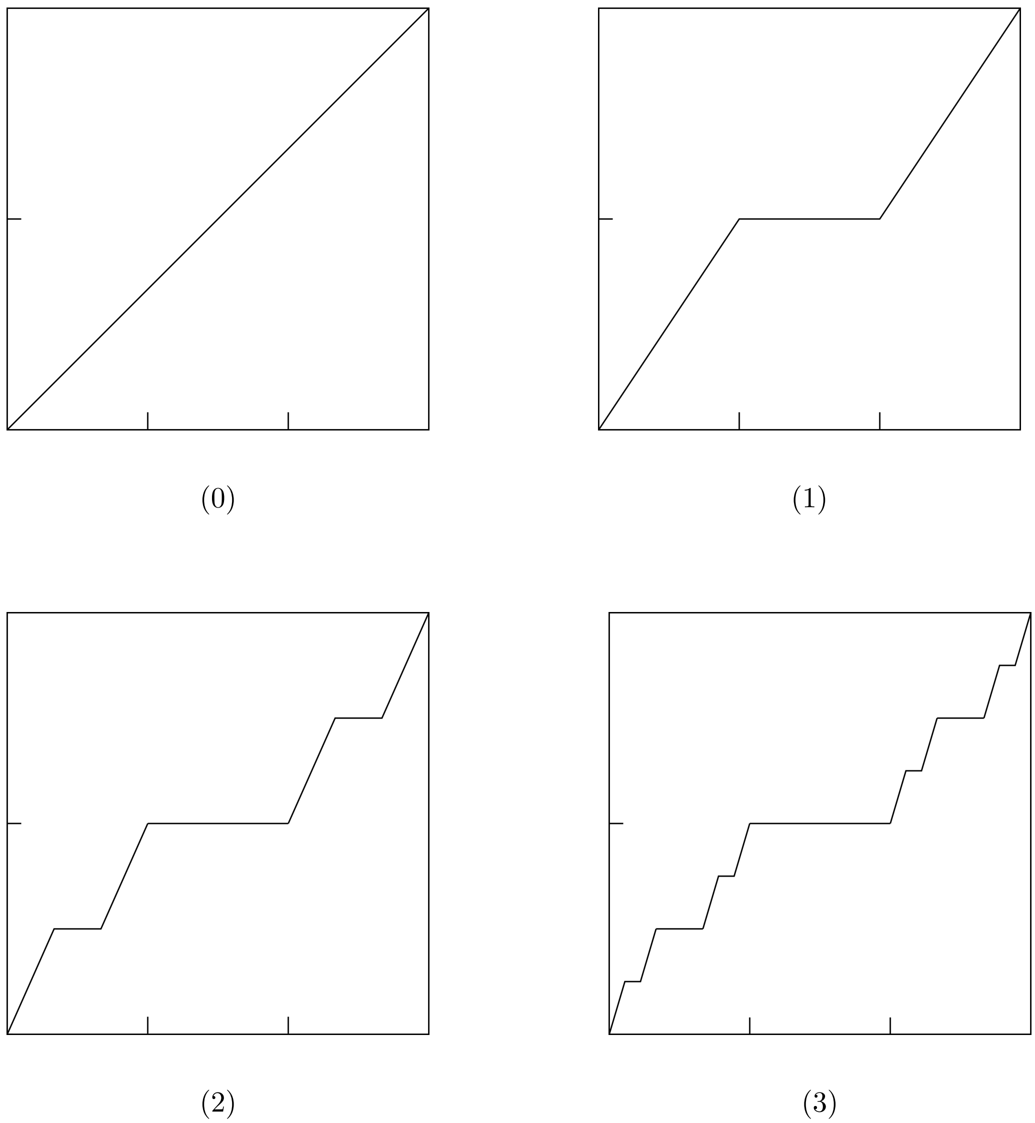

Figure 6: Illustration of the first iterates of the Cantor function on the interval $[0, 1]$. Here (n) refers to the n-iterate $\mathcal{G}_n$ for $n = 0, 1, 2, 3$.

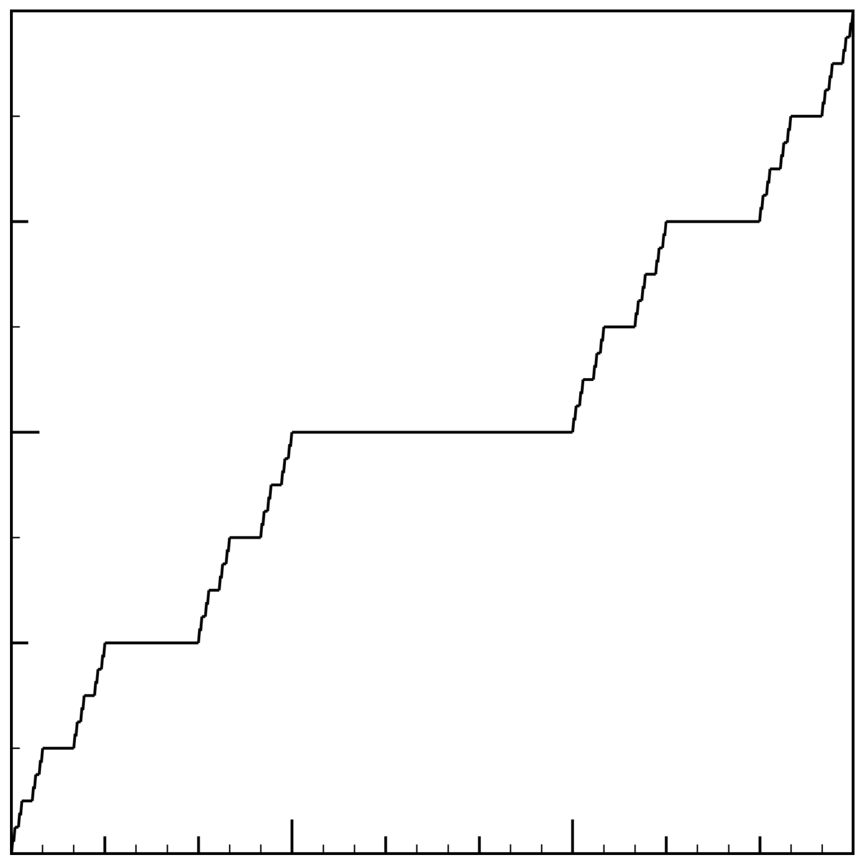

Figure 7: Plot of the Cantor function $\mathcal{G}(x) = \lim_{n\to\infty} \mathcal{G}_n(x)$ on the interval $[0, 1]$. This function has also been referred to by the more intriguing name the "Devil's staircase".

Note that there is some overlap in the split expression in (D.3), namely for arguments $x \in [\frac{1}{3}, \frac{2}{3}]$. However, as $\mathcal{G}_n(3x) = 1 + \mathcal{G}_n(3x - 2)$ for such x, (D.3) is indeed well-defined, and in the overlap we can apply either form at convenience. This observation will prove useful with regard to subadditivity.

Proposition D.1. The standard n-iterate Cantor function $\mathcal{G}_n$ is subadditive, i.e.

$$\mathcal{G}_n(a + b) \leq \mathcal{G}_n(a) + \mathcal{G}_n(b) \quad \forall\, a, b \in \mathbb{R} \text{ and } n = 0, 1, 2, \ldots \tag{D.4}$$

Proof. We will prove the statement by induction. Since the base case $n = 0$ is trivial, assume the subadditivity holds for $n = m$. We will show that subadditivity must also holds for $n = m + 1$. Without loss of generality, assume $a \leq b$ and consider the following possibilities:

(i) $a \leq 0$. Evident as $\mathcal{G}_{m+1}$ is monotonically increasing.

(ii) $a \geq \frac{1}{3}$. Then $\mathcal{G}_{m+1}(a+b) \leq 1 = \frac{1}{2} + \frac{1}{2} \leq \mathcal{G}_{m+1}(a) + \mathcal{G}_{m+1}(b)$.

So far we have not required the induction hypothesis. The next two cases, however, utilize said hypothesis.

(iii) $b \leq \frac{1}{3}$. Since $a, b, a+b \leq \frac{2}{3}$, we must have

$$\begin{aligned}\mathcal{G}_{m+1}(a+b) &= \frac{1}{2}\mathcal{G}_m(3a+3b)\\ &\leq \frac{1}{2}\mathcal{G}_m(3a) + \frac{1}{2}\mathcal{G}_m(3b) = \mathcal{G}_{m+1}(a) + \mathcal{G}_{m+1}(b).\end{aligned}$$

(iv) $0 \leq a \leq \frac{1}{3} \leq b$. Since $a + b \geq \frac{1}{3}$, we attain

$$\begin{aligned}\mathcal{G}_{m+1}(a+b) &= \frac{1}{2} + \frac{1}{2}\mathcal{G}_m(3a+3b-2)\\ &\leq \frac{1}{2}\mathcal{G}_m(3a) + \frac{1}{2} + \frac{1}{2}\mathcal{G}_m(3b-2) = \mathcal{G}_{m+1}(a) + \mathcal{G}_{m+1}(b).\end{aligned}$$

Since these four cases (i)-(iv) cover every possible value $a \leq b$ can take in $\mathbb{R}$, we are done.

□

References

[1] David Slepian and Henry O. Pollak. Prolate spheroidal wave functions, Fourier analysis and uncertainty - I. *Bell System Technical Journal*, 40:43–63, January 1961.

[2] Henry J. Landau and Henry O. Pollak. Prolate spheroidal wave functions, Fourier analysis and uncertainty - II. *Bell System Technical Journal*, 40:65–84, January 1961.

[3] Henry J. Landau and Henry O. Pollak. Prolate spheroidal wave functions, Fourier analysis and uncertainty - III: The dimensions of the space of essentially time- and bandlimited signals. *Bell System Technical Journal*, 41:1295–1336, July 1962.

[4] Ingrid Daubechies. Time-frequency localization operators: A geometric phase space approach. *IEEE Transactions on Information Theory*, 34(4):605–612, July 1988.

[5] Semyon Dyatlov. Notes on fractal uncertainty principle version 0.5 (october 4, 2017). `http://math.mit.edu/~dyatlov/files/2017/fupnotes.pdf`, 2017.

[6] Karlheinz Gröchenig. *Foundations of time-frequency analysis.* Applied and numerical harmonic analysis. Birkhäuser, Boston, 2001.

[7] Adam Bowers and Nigel J. Kalton. *An Introductory Course in Functional Analysis*. Universitext. Springer New York, New York, NY, 2014.

[8] Nicolaas. G. de Bruijn. *Uncertainty principles in Fourier analysis.* In O. Shisha (Ed.), Inequalities : Proceedings of a symposium held at Wright-Patterson air force base, Ohio, August 19-27, 1965: 57-71. New York: Academic Press, 1967.

[9] Gerald B. Folland. *Harmonic analysis in phase space*, volume 122 of *Annals of mathematics studies.* Princeton University Press, Princeton, N.J, 1989.

[10] David J Griffiths. *Introduction to quantum mechanics.* Prentice Hall, Upper Saddle River, N.J, 1995.

[11] David L. Donoho and Philip B. Stark. Uncertainty principles and signal recovery. *SIAM Journal on Applied Mathematics*, 49(3):906–931, June 1989.

[12] Elliott H. Lieb. Integral bounds for radar ambiguity functions and Wigner distributions. *Journal of Mathematical Physics*, 31(3):594–599, March 1990.

[13] Michael Benedicks. On Fourier transforms of functions supported on sets of finite Lebesgue measure. *Journal of Mathematical Analysis and Applications*, 106(1):180–183, 1985.

[14] Gerald Folland and Alladi Sitaram. The uncertainty principle: A mathematical survey. *Journal of Fourier Analysis and Applications*, 3(3):207–238, May 1997.

[15] A. Janssen. Proof of a conjecture on the supports of Wigner distributions. *Journal of Fourier Analysis and Applications*, 4(6):723–726, November 1998.

[16] Camil Muscalu and Wilhelm Schlag. *Classical and multilinear harmonic analysis : Vol. 1*, volume vol. 137 of *Cambridge studies in advanced mathematics*. Cambridge University Press, Cambridge, 2013.

[17] Oleg Kovrijkine. Some results related to the Logvinenko-Sereda theorem. *Proceedings of the American Mathematical Society*, 129(10):3037–3047, April 2001.

[18] Alexander Reznikov. Sharp constants in the Paneyah–Logvinenko–Sereda theorem. *Comptes rendus - Mathématique*, 348(3):141–144, January 2010.

[19] Carmen Fernández and Antonio Galbis. Annihilating sets for the short time Fourier transform. *Advances in Mathematics*, 224(5):1904–1926, August 2010.

[20] Helge Knutsen. *Spectral Properties of Time-Frequency Localization Operators*. Projet report for Masters Programme in Industrial Mathematics. NTNU, Trondheim, Desember 2017.

[21] Helge Knutsen. *Comments and corrections to error in Project report: Spectral Properties of Time-Frequency Localization Operators*. NTNU, Trondheim, January 2018.

[22] Germund Dahlquist. *Numerical methods*. Prentice-Hall series in automatic computation. Prentice-Hall, Englewood Cliffs, N.J, 1974.

[23] Jozef Doboš. The standard Cantor function is subadditive. *Proceedings of the American Mathematical Society*, 124(11):3425–3426, November 1996.