ABSTRACT

Title of dissertation: Finite Frames and Graph Theoretical Uncertainty Principles

Paul J. Koprowski, Doctor of Philosophy, 2015

Dissertation directed by: Dr. John J. Benedetto
Department of Mathematics

The subject of analytical uncertainty principles is an important field within harmonic analysis, quantum physics, and electrical engineering. We explore uncertainty principles in the context of the graph Fourier transform, and we prove additive results analogous to the multiplicative version of the classical uncertainty principle. We establish additive uncertainty principles for finite Parseval frames. Lastly, we examine the feasibility region of simultaneous values of the norms of a graph differential operator acting on a function $f \in l^2(G)$ and its graph Fourier transform.

Finite Frames and Graph Theoretic Uncertainty Principles

by

Paul J. Koprowski

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2015

Advisory Committee:
Dr. John J. Benedetto, Chair/Advisor
Dr. Kenneth Berg,
Dr. Wojciech Czaja,
Dr. Patrick M. Fitzpatrick,
Dr. Kasso A. Okoudjou
Dr. Kaye Brubaker, Dean's Representative

Dedication

To the 2008 World Series Champion Philadelphia Phillies, for making that first semester of graduate school worth it.

Acknowledgments

First, and foremost, I would like to thank my advisor, John Benedetto, for guiding me to my position as number fifty three. I would also like to acknowledge the support of the Norbert Wiener Center for Harmonic Analysis and Applications. Thank you to professors Kasso Okoudjou and Wojtek Czaja for serving on both my preliminary oral exam, and my final oral exam. Also, thank you to the rest of my committee members: professors Berg, Brubaker, and Fitzpatrick. Thank you to Celeste for her work with graduate school forms, and, especially for her help securing the paperwork for my emergency committee member professor Schaefer, who I would also like to thank. Most importantly, I would like to acknowledge my family, my friends and my girlfriend Maia, who helped keep me sane during my prolonged slog toward a degree.

Table of Contents

List of Figures

Chapter 1: Introduction

1.1 Overview

The subject of analytical uncertainty principles is an important field within harmonic analysis, quantum physics, and electrical engineering. Graph theory is a well established field of mathematics and computer science. In the era of so called "big data" problems, such as searches of social networks and ever growing databases of digital information, applied science has joined pure mathematics in striving to understand the theory of graphs. Recent advances in analytic graph theory (namely the nascent field of Fourier analysis on graphs) have presented the challenge of determining what uncertainty principles exist, if any, within analytic graph theory. To determine some of the answers to this query, we turn to a diverse set of subjects including, but not limited to, linear operator theory, frames, wavelets, quantum physics, and signal processing. We shall provide a study of some of the existing results and techniques tied to these theories as they pertain to uncertainty principles. By doing so, we provide insight into the classical underpinnings of modern Fourier analysis on graphs in order to motivate the extension (when possible) of classical results to the analytic graph setting.

1.2 Outline of Thesis and Results

In chapter 2, we examine the classical uncertainty principle by providing exposition on classical Fourier analysis and on general linear operator theory as it pertains to Hilbert spaces. Both multiplicative and additive versions of the classical uncertainty are presented, the latter of which will be extended to the graph theoretical setting. Chapter 3 establishes the tools necessary to examine uncertainty principles in the discrete setting. Namely, we provide an introduction to key concepts in linear algebra and frame theory necessary for establishing the discrete Heisenberg uncertainty principle due to Grünbaum [24] and a finite frame uncertainty principle due to Lammers and Maeser [33]. In chapter 4, we define graphs, the graph Laplacian, and the graph Fourier transform. Chapter 5 motivates the techniques used in the main results of this study by providing exposition on the classical wavelet transform, as well as the spectral graph wavelet transform due to Hammond, Vandergheynst, and Gribonval in [25]. In chapter 6, we prove Theorems (6.2.1) and (6.2.2) which state (in analogy to the additive classical uncertainty principle) that the sum of the norms of a graph differential operator acting on a function $f \in l^2(G)$ and its graph Fourier transform are always bounded below by a positive constant based on the structure of the underlying graph G. We also prove (in analogy to the finite frame results in [33]) Theorems (6.3.2) and (6.3.3) which state that the sum of the norms of a graph differential operator acting on a Parseval frame E for $f \in l^2(G)$ and the graph Fourier transform of E are always bounded below by a positive constant based on the structure of the underlying graph G. We present the unit weighted complete graph, and compute specific lower and upper bounds whose existence is established in Theorems (6.2.1) and (6.3.2). In the final chapter, we provide exposition concerning the feasibility region for graph and spectral spreads due to Agaskar and Lu [1] and prove analogous results, via Proposition (7.3.1) and Theorem (7.3.4), for the differential feasibility region for simultaneous values of the norms of a graph differential operator acting on a function $f \in l^2(G)$ and its graph Fourier transform. We conclude chapter 7 by computing the specific values specified by Proposition (7.3.1) and Theorem (7.3.4) for the unit weighted complete graph. Lastly, we examine the differential feasibility region for the complete graph, and we compare the results to the Bell lab uncertainty principles in [35] and [47].

Chapter 2: The Fourier Transform and the Classical Uncertainty Principle

2.1 Introduction

Uncertainty principle inequalities play a fundamental role in Fourier analysis, as well as in quantum mechanics. The so called Heisenberg uncertainty principle derived in the works of Heisenberg [26], Pauli [38], Weyl [50], and Wiener has had great influence on science. Indeed, one of its fundamental consequences (no simultaneous exact knowledge of both the position and momentum of a quantum particle) has bridged the gap between physics and the lexicon of lay knowledge. We shall introduce the Fourier transform, and examine some of its properties. For in-depth treatment of the Fourier Transform in abstract settings see [9], [28], [29], or [22] among others. We shall use these properties to prove the classical uncertainty principle inequality, in a similar fashion to [6].

2.2 The Classical Uncertainty Principle

We begin with a Fourier centric version of the classical uncertainty principle inequality. This is a special case of the classical uncertainty principle for general

Hilbert spaces, which we shall formulate, and prove at the end of the chapter. Note that as a convention, $\int \cdot dt$, respectively, $\int \cdot d\gamma$, will denote an integral over all $t \in \mathbb{R}$, respectively, all $\gamma \in \widehat{\mathbb{R}} = \mathbb{R}$. We denote the space of functions $f : \mathbb{R} \to \mathbb{C}$ with finite Lebesgue integral of its modulus, respectively, modulus squared, as $L^1(\mathbb{R})$, respectively, $L^2(\mathbb{R})$. These are normed vector spaces with norms defined by the Lebesgue integral:

$$\|f(t)\|_{L^1(\mathbb{R})} = \int |f(t)|\, dt$$

and

$$\|f(t)\|_{L^2(\mathbb{R})} = \left(\int |f(t)|^2\, dt\right)^{1/2}.$$

We shall denote the space of Schwartz functions on $\mathbb{R}$ as $\mathcal{S}(\mathbb{R})$. The *Fourier Transform* of $f \in L^1(\mathbb{R})$ is defined as

$$\widehat{f}(\gamma) = \int_{\mathbb{R}} f(t)e^{-2\pi it\gamma}dt. \tag{2.1}$$

The formal *Fourier inversion formula* for $\widehat{f}$ is

$$f(t) = \int_{\widehat{\mathbb{R}}=\mathbb{R}} \widehat{f}(\gamma)e^{2\pi it\gamma}d\gamma. \tag{2.2}$$

There are several intriguing algebraic properties of the Fourier transform. In what follows, we examine the effect of the translation and the dilation of a function f on its Fourier transform, $\widehat{f}$. Notationally, for a fixed γ, we set

$$e_\gamma(t) = e^{2\pi it\gamma};$$

and, for a fixed u and a given function f, we set

$$(\tau_u f)(t) = f(t-u).$$

$\tau_u f$ is *translation* of f by u, and $e_\gamma(t)f(t)$ is*modulation* of f by γ.

Lemma 2.2.1 *For a function $f \in L^1(\mathbb{R})$ with Fourier transform $\widehat{f}$, the following hold.*

1. *For $u \in \mathbb{R}$, we have*

$$(\tau_u f)\widehat{\ }(\gamma) = e_{-u}\widehat{f}(\gamma). \tag{2.3}$$

2. *For $\lambda \in \widehat{\mathbb{R}}$, we have*

$$(e^{2\pi it\lambda} f(t))\widehat{\ }(\gamma) = \tau_\lambda \widehat{f}(\gamma). \tag{2.4}$$

Proof: We shall verify both (2.3) and (2.4) via direct calculation. For translation, we employ the substitution $y = t - u$ yielding:

$$\begin{aligned}
(\tau_u f(t))\widehat{\ }(\gamma) &= \int f(t-u)e^{-2\pi it\gamma}dt \\
&= \int f(y)e^{-2\pi i(y+u)\gamma}dy \\
&= e^{-2\pi iu\gamma}\int f(y)e^{-2\pi iy\gamma}dy \\
&= e_{-u}\widehat{f}(\gamma).
\end{aligned}$$

For the modulation case, we have

$$\begin{aligned}
(e_\lambda f(t))\widehat{\ }(\gamma) &= \int f(t)e^{-2\pi it(\gamma-\lambda)}dt \\
&= \widehat{f}(\gamma-\lambda) \\
&= (\tau_\lambda \widehat{f})(\gamma).
\end{aligned}$$

■

We shall show that the Fourier transform of a Gaussian is also a Gaussian.

Example 2.2.2 *Let $f(t) = e^{-st^2}$, $s > 0$. Upon differentiating $\widehat{f}(\gamma)$ with respect to γ we have*

$$\frac{d}{d\gamma}\left(\widehat{f}(\gamma)\right) = \int -2\pi i t e^{-st^2} e^{-2\pi i t\gamma} dt. \tag{2.5}$$

Noting that $\frac{d}{dt}f(t) = -2stf(t)$, we rewrite (2.5) as

$$\begin{aligned}
\widehat{f}'(\gamma) &= -2\pi i \int \frac{-1}{2s}\left(e^{-st^2}\right)' e^{-2\pi i t\gamma} dt \\
&= \frac{\pi i}{s}\left(\left(e^{-st^2} e^{-2\pi i t\gamma}\right)|_{-\infty}^{\infty} - (-2\pi i\gamma)\widehat{f}(\gamma)\right) \\
&= \frac{-2\pi^2\gamma}{s}\widehat{f}(\gamma).
\end{aligned}$$

Hence $\widehat{f}(\gamma)$ satisfies the differential equation $\widehat{f}'(\gamma) = \frac{-2\pi^2\gamma}{s}\widehat{f}(\gamma)$, which is solved by

$$\widehat{f}(\gamma) = Ce^{\frac{-\pi^2\gamma^2}{s}}.$$

To obtain C we set $\gamma = 0$, and we exploit the relationship

$$\left(\int e^{-st^2} dt\right)^2 = \int_{\mathbb{R}^2} e^{-s(t^2+y^2)} dA.$$

To calculate integral of the the two dimensional Gaussian we employ polar coordinates:

$$\begin{aligned}
\int_{\mathbb{R}^2} e^{-s(t^2+y^2)} dA &= \int_0^{2\pi}\int_0^{\infty} re^{-sr^2} dr d\theta \\
&= 2\pi\left(-\frac{-1}{2s}e^{-s0^2}\right) \\
&= \frac{\pi}{s}.
\end{aligned} \tag{2.6}$$

We conclude $C = \int e^{-st^2} dt = \sqrt{\frac{\pi}{s}}$, and that

$$\widehat{f}(\gamma) = \sqrt{\frac{\pi}{s}} e^{\frac{-\pi^2\gamma^2}{s}}.$$

Example 2.2.2 shows us that for any constant C, $Ce^{-\pi t^2}$ is an eigenfunction of the Fourier transform. This leaves us with a soupçon of suspicion that Gaussian functions have an intimate and important relationship with the Fourier transform. Indeed, a specific class of Gaussian functions turn out to be the minimizers of the Classical Uncertainty Principle Inequality.

Theorem 2.2.3 *(The Classical Uncertainty Principle Inequality). Let* $(t_0, \gamma_0) \in \mathbb{R} \times \widehat{\mathbb{R}}$. *Then*

$$\forall f \in \mathcal{L}^{\in}(\mathbb{R}), \ \ \|f\|^2_{L^2(\mathbb{R})} \leq 4\pi \, \|(t-t_0)f(t)\|_{L^2(\mathbb{R})} \left\|(\gamma-\gamma_0)\widehat{f}(\gamma)\right\|_{L^2(\widehat{\mathbb{R}})}, \tag{2.7}$$

with equality if $f(t) = Ce^{2\pi i t\gamma_0}e^{-s(t-t_0)^2}$, *for any* $C \in \mathbb{C}$ *and* $s > 0$.

Proof: First, consider the case when $t_0 = 0 = \gamma_0$. Integration by parts yields

$$\begin{aligned} \|f\|^4 &= \left(-\int t\frac{d}{dt}\,|f(t)|^2\,dt\right)^2 \\ &= \left(\int t\frac{d}{dt}\,|f(t)|^2\,dt\right)^2 \\ &\leq \left(\int |t| \left|\frac{d}{dt}\,|f(t)|^2\right| dt\right)^2. \end{aligned}$$

If $f(t) = a(t) + ib(t)$ for differentiable functions $a, b : \mathbb{R} \to \mathbb{R}$, then

$$\begin{aligned} \left|\frac{d}{dt}\,|f(t)|^2\right| &= 2\,|a(t)a'(t) + b(t)b'(t)| \\ &\leq 2\left((a(t)a'(t) + b(t)b'(t))^2 + (a(t)b'(t) - a'(t)b(t))^2\right)^{1/2} \\ &= \left|2\overline{f(t)}f'(t)\right|. \end{aligned}$$

Employing this inequality, followed by Hölder's inequality, and finally Plancherel's theorem, we obtain the desired result:

$$
\begin{aligned}
\|f\|^4 &\leq 4\left(\int \left|t\overline{f(t)}f'(t)\right| dt\right)^2 \\
&\leq 4\,\|tf(t)\|^2\,\|f'(t)\|^2 \\
&\leq 4\,\|tf(t)\|^2\,\|f'(t)\widehat{\ }\|^2 \\
&= 16\pi^2\,\|tf(t)\|^2\,\|\gamma f(\gamma)\widehat{\ }\|^2 \qquad (2.8)
\end{aligned}
$$

with (2.8) due to the fact that $\widehat{f'}(\gamma) = 2\pi i\gamma\widehat{f}(\gamma)$.

For the case of non-zero t_0 and γ_0, consider the function $g(t) = f(t+t_0)e^{-2\pi it\gamma_0}$. We then have

$$
\begin{aligned}
\int |f(s)|^2\,ds &= \int 1\,|f(t+t_0)|^2\,dt \\
&= \int \left|e^{-2\pi it\gamma_0}\right|^2 |f(t+t_0)|^2\,dt \\
&= \int |g(t)|^2\,dt. \qquad (2.9)
\end{aligned}
$$

Hence, $g \in L^2(\mathbb{R})$, and it has the same norm as f. Using this fact, and applying (2.8) to g yields

$$\|f(t)\|^2 = \|g(t)\|^2 \leq 4\pi\,\|tg(t)\|\,\|\gamma g(\gamma)\widehat{\ }\|\,.$$

To calculate $\widehat{g}(\gamma) = e^{2\pi it_0\gamma}\widehat{f}(\gamma+\gamma_0)$, we note that $g(t) = e_{-\gamma_0}(\tau_{-t_0}f(t))$. Hence we have

$$
\begin{aligned}
\widehat{g}(\gamma) &= [e_{-\gamma_0}(\tau_{-t_0}f(t))]\widehat{\ }(\gamma) \\
&= e_{t_0}(\tau_{-\gamma_0}f(t))\widehat{\ }(\gamma) \qquad (2.10) \\
&= e^{2\pi it_0\gamma}\widehat{f}(\gamma+\gamma_0),
\end{aligned}
$$

where (2.10) is due to lemma 2.2.1. Hence we have

$$\begin{aligned}\|\gamma\widehat{g}(\gamma)\|^2 &= \int \gamma^2 \left|\widehat{g}(\gamma)\right|^2 d\gamma \\ &= \int \gamma^2 \left|\widehat{f}(\gamma+\gamma_0)\right|^2 d\gamma \\ &= \int (\gamma-\gamma_0)^2 \left|\widehat{f}(\gamma)\right|^2 d\gamma \\ &= \left\|(\gamma-\gamma_0)\widehat{f}(\gamma)\right\|^2 .\end{aligned}$$

Similarly, we have that $\|tg(t)\|^2 = \|(t-t_0)f(t)\|^2$. Thus we are left with the desired result:

$$\|f(t)\|^2 \leq 4\pi \|(t-t_0)f(t)\| \left\|(\gamma-\gamma_0)\widehat{f}(\gamma)\right\| . \tag{2.11}$$

We shall show equality for $f(t) = Ce^{2\pi i t\gamma_0}e^{-s(t-t_0)^2}$. For simplicity of calculation, we assume, without loss of generality, that $C = 1$. We have that $\int e^{-s(t-t_0)^2}dt = \int e^{-st^2}dt = \sqrt{\frac{\pi}{s}}$, by (2.6), and it follows that $\|f(t)\|^2 = \int |e^{2\pi i t\gamma_0}|^2|e^{-s(t-t_0)^2}|^2 dt = \int e^{-2s(t-t_0)^2}dt = \sqrt{\frac{\pi}{2s}}$. We calculate the time spread of f via the substitution $y = t - t_0$ followed by integration by parts:

$$\begin{aligned}\|(t-t_0)f(t)\|^2 &= \int |t-t_0|^2 e^{-2s(t-t_0)^2} dt \\ &= \int yye^{-2sy^2} dy \\ &= \frac{1}{4s}\int e^{-2sy^2} dy \\ &= \frac{1}{4s}\sqrt{\frac{\pi}{2s}}, \end{aligned} \tag{2.12}$$

where the final equality is due to (2.6). For the frequency spread we, again, employ the substitution $y = \gamma - \gamma_0$ followed by integration by parts.

$$\begin{aligned}\left\|(\gamma-\gamma_0)\widehat{f}(\gamma)\right\|^2 &= \frac{\pi}{s}\int |\gamma-\gamma_0|^2 \left|e^{-2\pi t_0\gamma}\right|^2 e^{-2\pi(\gamma-\gamma_0)/s}d\gamma \\ &= \int yye^{-2\pi^2y^2/s}dy \\ &= \frac{\pi s}{4\pi^2 s}\int e^{-2\pi^2y^2/s}dy \\ &= \frac{1}{4\pi}\sqrt{\frac{\pi s}{2\pi^2}} = \frac{1}{4\pi}\sqrt{\frac{s}{2\pi}}, \end{aligned} \tag{2.13}$$

where we have employed (2.6) for the final equalities. Combining (2.12) and (2.13), and multiplying by $16\pi^2$ yields the desired result:

$$\|f(t)\|^4 = \frac{\pi}{2s} = 16\pi^2\frac{1}{4s}\sqrt{\frac{\pi}{2s}}\frac{1}{4\pi}\sqrt{\frac{s}{2\pi}} = 16\pi^2\|(t-t_0)f(t)\|^2\left\|(\gamma-\gamma_0)\widehat{f}(\gamma)\right\|^2.$$

■

Applying Cauchy's inequality to inequality (2.7) yields an additive corollary.

Corollary 2.2.4 *The following inequality holds:*

$$\forall f\in\mathcal{S}(\mathbb{R}),\quad \|f\|^2 \le 2\pi\left(\|tf(t)\|^2 + \left\|\gamma\widehat{f}(\gamma)\right\|^2\right). \tag{2.14}$$

Furthermore, the bound is sharp.

Proof: By Cauchy's inequality, and inequality (2.7) we have

$$\|f\|^2 \le 4\pi\|tf(t)\|\left\|\gamma\widehat{f}(\gamma)\right\| \le 2\pi\left(\|tf(t)\|^2 + \left\|\gamma\widehat{f}(\gamma)\right\|^2\right).$$

If $f(t) = e^{-\sqrt{\pi}t^2}$, then by Theorem 2.2.3 the left inequality is equality. Further, f is its own Fourier transform so

$$2\|tf(t)\|\left\|\gamma\widehat{f}(\gamma)\right\| = \|tf(t)\|^2 + \left\|\gamma\widehat{f}(\gamma)\right\|^2$$

as desired. ■

An additional corollary is of particular importance to our work, as we shall prove an analogous case in the graph setting.

Corollary 2.2.5 *The following inequality holds:*

$$\forall f \in \mathcal{S}(\mathbb{R}), \quad \|f\|^2 \leq \left(\|f'(t)\|^2 + \left\| \widehat{f}'(\gamma) \right\|^2 \right). \tag{2.15}$$

Furthermore, the bound is sharp.

Proof: If $f \in \mathcal{S}(\mathbb{R})$, then $f(t)$ is differentiable, and

$$(f'(t))\widehat{\ }(\gamma) = 2\pi i \gamma \widehat{f}(\gamma) \text{ and } ((-2\pi i t) f(t))\widehat{\ }(\gamma) = \widehat{f}'(\gamma).$$

Making the appropriate substitutions into inequality (2.14) and simplifying yields the desired result. ■

We shall continue to examine the classical uncertainty principle inequality using a general Hilbert space formulation. Some definitions will help consolidate notation.

Definition Let A, B be self-adjoint operators on a complex Hilbert space H with domains $D(A)$, and $D(B)$ respectively. Define the following:

1. The *domain* $D(AB)$, of AB, is defined as the set

   $$D(AB) = \{f \in D(B) \subset H : Bf \in D(A)\},$$

   and likewise for the domain of BA.

2. The *commutator* $[A, B]$, of A and B, is defined as $[A, B] = AB - BA$.

3. The *expected value* $E_f(A)$, of A at $f \in D(A) \subset H$,, is defined as $E_f(A) = \langle Af, f\rangle_H$.

4. The *variance* $\sigma_f^2(A)$, of A at $f \in D(A^2) \subset D(A)] \subset H$, is defined as $\sigma_f^2(A) = E_f(A^2) - (E_f(A))^2$.

We reformulate the classical uncertainty principle inequality in the context of the general operator notation.

Theorem 2.2.6 *Let A, B be self adjoint operators on a complex Hilbert space H. If $f \in D = D(A^2) \cap D(B^2) \cap D([A, B])$ with $\|f\|_H = 1$, we then have*

$$(E_f(i[A, B]))^2 \leq 4\sigma_f^2(A)\sigma_f^2(B).$$

Proof: The proof is a consequence of a few routine calculations. We have

$$\begin{aligned} E_f(i[A, B]) &= i(\langle ABf, f\rangle_H - \langle BAf, f\rangle_H) \\ &= i(\overline{\langle Af, Bf\rangle}_H - \langle Af, Bf\rangle_H) \\ &= i[(-i)\Im \langle Af, Bf\rangle_H - i\Im \langle Af, Bf\rangle] \\ &= 2\Im \langle Af, Bf\rangle_H , \end{aligned} \tag{2.16}$$

and

$$\begin{aligned} \|(B + iA)f\|_H^2 &= \langle (B + iA)f, (B + iA)f\rangle_H \\ &= \langle Bf, Bf\rangle_H + \langle iAf, iAf\rangle_H + \langle iAf, Bf\rangle_H + \langle Bf, iAf\rangle_H \\ &= \|Bf\|_H^2 + \|Af\|_H^2 - 2\Im \langle Af, Bf\rangle_H . \end{aligned} \tag{2.17}$$

Note that for any self adjoint operator C, $E_f(C)$ is real. Indeed, we have that

$$\overline{E_f(C)} = \langle f, Cf \rangle_H = \langle Cf, f \rangle_H = E_f(C),$$

and therefore, we have $\left(\Re \langle iAf, f \rangle_H\right)^2 = 0$ and $\left(\Im \langle iAf, f \rangle_H\right)^2 = \left(\langle Af, f \rangle_H\right)^2$. We conclude that

$$\begin{aligned} \left|\langle (B - iA)f, f \rangle_H\right|^2 &= \left|\langle Bf, f \rangle_H + i \langle Af, f \rangle_H\right|^2 \\ &= \left(\langle Bf, f \rangle_H\right)^2 + \left(\langle Af, f \rangle_H\right)^2. \end{aligned} \tag{2.18}$$

By the Cauchy-Schwarz Inequality we have

$$\begin{aligned} 0 &\leq \|(B + iA)f\|_H^2 - \left|\langle (B + iA)f, f \rangle_H\right|^2 \\ &= \left(\|Af\|_H^2 - \langle Af, f \rangle_H^2\right) + \left(\|Bf\|_H^2 - \langle Bf, f \rangle_H^2\right) - 2\Im \langle Af, Bf \rangle_H, \end{aligned} \tag{2.19}$$

where (2.19) is due to (2.17) and (2.18). Noting that the first two terms in (2.19) are the variance of A, respectively, B, at f, then rearranging, and applying (2.16) yields an additive inequality:

$$\begin{aligned} \sigma_f^2(A) + \sigma_f^2(B) &\geq 2\Im \langle Af, Bf \rangle_H \\ &= E_f(i[A, B]). \end{aligned} \tag{2.20}$$

Let $r > 0$, $s > 0$ and note that rA and sB are self adjoint operators. Applying (2.20) to rA and sB yields

$$r^2\sigma_f^2(A) + s^2\sigma_f^2(B) \geq rsE_f(i[A, B]).$$

Setting $r^2 = \sigma_f^2(B)$, $s^2 = \sigma_f^2(A)$, and squaring both sides yields the desired result:

$$\left(2\sigma_f^2(A)\sigma_f^2(B)\right)^2 = 4\sigma_f^4(A)\sigma_f^4(B) \geq \sigma_f^2(A)\sigma_f^2(B)\left(E(i[A, B])\right)^2$$

if and only if

$$\left(E_f(i[A,B])\right)^2 \leq 4\sigma_f^2(A)\sigma_f^2(B).$$

■

If we take $H = L^2(\mathbb{R})$, $A(f(t)) = tf(t)$, and $B(f(t)) = i(2\pi\gamma\widehat{f}(\gamma))\check{}(t)$ then Theorem 2.2.3 is direct consequence of Theorem 2.2.6. In quantum mechanics, pure quantum states are typically taken to be unit vectors in a Hilbert space H. If A, B are self adjoint operators on this space, then their eigenvalues are interpreted as the observable quantities of certain systems. The variances $\sigma_f(A)$, and $\sigma_f(B)$ are then representative of uncertainties in these observables. Interpreting A as the position operator and B as the momentum operator, Theorem (2.2.6) can be interpreted as the lay version of the Heisenberg uncertainty principle: no simultaneous knowledge of both position and momentum. The classical uncertainty inequalities set the table, so to speak, for the main work of this thesis. Given a space of functions, and a notion of a Fourier transform on that space, what can be said about the simultaneous properties of a function and its Fourier transform.

Chapter 3: Uncertainty Principles of the Discrete Fourier Transform

3.1 Introduction

As we saw in the previous chapter, the Heisenberg uncertainty principle states

$$\forall f \in \mathcal{S}(\mathbb{R}), \;\; \|f\|^2 \leq 4\pi \, \|tf(t)\| \, \|\gamma \widehat{f}(\gamma)\|.$$

In order to examine discrete versions of this uncertainty principle, we review key concepts of linear algebra, and we introduce the discrete Fourier transform (DFT) and some of its key properties. Having set up the necessary framework, we examine a discrete Heisenberg uncertainty principle from [24] and a discrete finite frame uncertainty principle from [33].

3.2 Linear Algebra

Linear algebra is fundamental to our analysis on graphs. We shall provide the theory necessary for analyzing uncertainty principles associated with the DFT, and for proving our main graph theoretic results. We shall provide a minimalist overview of the necessary theory. For more complete and detailed treatments, see [34], [3], [43], or [30] among multitudes of others. For more abstract linear operator theory, see [2], [5], [41], or [40].

We shall restrict our considerations to the finite dimensional spaces $l^2(\mathbb{C}^N)$ and $l^2(\mathbb{R}^N)$, where we view each element $f \in l^2(\mathbb{C}^N)$, respectively, $f \in l^2(\mathbb{R}^N)$, as functions $f : \mathbb{Z}/(N\mathbb{Z}) \rightarrow \mathbb{C}^N$, respectively, $f : \mathbb{Z}/(N\mathbb{Z}) \rightarrow \mathbb{R}^N$. Notationally, we write $f[j]$ to denote f_j, the j^{th} element of the vector f. We restrict our attention to $l^2(\mathbb{C}^N)$ unless otherwise specified.

The set of linear operators, $l^2(\mathbb{C}^N) \rightarrow l^2(\mathbb{C}^d)$, may be represented by all complex valued matrices with d rows and N columns denoted $\mathbb{C}^{d\times N}$. If $N = d$, then this set consists of square matrices and it is denoted $\mathbb{C}^{N\times N}$. In this case, we may define several special properties. A square matrix M is said to be *diagonal* if the only non-zero entries of M are on the diagonal denoted by M_{jj} for $j = 0, ..., N-1$. In this case, any function e_k of the form

$$e_k[j] = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{otherwise,} \end{cases}$$

has the property that

$$Me_k = M_{kk}e_k = \lambda_k e_k.$$

We call the constant values λ_k for $k = 0, ..., N-1$ the *eigenvalues* associated with M, and the functions e_k the *eigenfunctions* associated with these values. More generally, we refer to $\{e_k\}_{k=0}^{N-1}$ as the *canonical basis* functions for $l^2(\mathbb{C}^N)$. If we have a square matrix A which is not diagonal, then eigenvalues λ_j with associated eigenfunctions χ_j satisfying

$$A\chi_j = \lambda_j\chi_j$$

still exist, and are of great importance to the discrete results discussed in this work.

We refer to the set of eigenvalues associated with a matrix A as the *spectra* of the operator A. We examine certain special types of square matrices, and the special properties that their respective spectra possess.

The monic *characteristic* polynomial $p(x)$ of a square matrix A is defined as

$$p_A(x) = p(x) = \det(xI - A) = c_0 + c_1 x + ... + c_{N-1} x^{N-1} + x^N,$$

and has the property that λ is an eigenvalue of A, if and only if $p(\lambda) = 0$. Since $\mathbb{C}$ is algebraically closed, we may factor $p(x)$ into linear terms:

$$p_A(x) = \prod_{j=0}^{N-1} (x - \lambda_j). \tag{3.1}$$

A celebrated theorem of matrix theory is the Cayley-Hamilton Theorem which ensures, in this setting, that $p_A(A) = 0_N$ which is the $N \times N$ matrix of all zeros. $p_A(x)$ is a monic polynomial of degree N, and it may possess repeated roots. If this is the case, there exists a monic *minimal polynomial* $m_A(x)$ of degree less than or equal to N such that $m_A(A) = 0_N$. In our setting, if we let λ_{j_l} for $l = 0, ..., N' \leq N$ denote the subset of distinct eigenvalues associated with A, then we have

$$m_A(x) = \prod_{l=0}^{N'} (x - \lambda_{j_l}). \tag{3.2}$$

For a square matrix A, we define the *transpose* A^t of A by setting

$$\left(A^t\right)_{jk} = A_{kj}.$$

We define the *conjugate transpose* A^* of A by setting

$$\left(A^*\right)_{jk} = \bar{A}_{kj}.$$

A matrix A for which $A^*A = AA^*$ is said to be *normal*. If $A^* = A$ then A is said to be *Hermitian*. If A is real and Hermitian then it is clear that $A^* = A^t = A$, and we say A is *symmetric*. All Hermitian matrices are normal. Indeed, if $A^* = A$ then $A^*A = A^2 = AA^*$. A key property of normal $N \times N$ matrix A (and hence Hermitian and real symmetric matrices) is that there exist N orthonormal eigenfunctions χ_j that diagonalize A. This property is the celebrated spectral theorem which we formulate in a manner similar to [3].

Theorem 3.2.1 *(Spectral Theorem) Let $A \in \mathbb{C}^{N\times N}$. There exists an orthonormal eigen basis for $\mathbb{C}^N$ associated with A if and only if A is normal.*

Hence we have that if χ is the $N \times N$ matrix whose columns are given by $\{\chi_j\}$, that is

$$\chi = [\chi_0, \chi_1, ..., \chi_{N-1}],$$

then

$$\chi^*\chi = I = \chi\chi^*, \tag{3.3}$$

and

$$D = \chi^* A \chi \text{ where } D = diag(\lambda_0, ..., \lambda_{N-1}).$$

More generally, a matrix U satisfying equation (3.3), that is $U^*U = I = UU^*$, is called *unitary*.

The spectral values of Hermitian matrices are key to our analysis of uncertainty principles in the discrete setting. If we let H be an $N \times N$ Hermitian matrix then all of the eigenvalues of H are real. Indeed, since H must also be normal, we may diagonlize H such that $H = UDU^*$ where $D = diag(\lambda_0, ..., \lambda_{N-1})$ and $H^* = H$ implies

$$\bar{D} = D^* = (U^*HU)^* = U^*H^*U = U^*HU = D$$

and hence the eigenvalues are all real. Hence, the spectral values $\{\lambda_j\}$ of H can be indexed such that

$$\lambda_0 \leq \lambda_1 \leq ... \leq \lambda_{N-1}.$$

For any Hermitian matrix H and for any non zero function f, we define the *Rayleigh quotient* as

$$R(H, f) = \frac{\langle f, Hf \rangle}{\langle f, f \rangle} = \frac{\langle Hf, f \rangle}{\langle f, f \rangle},$$

where $\langle\ ,\ \rangle$ is the standard inner product for $l^2(\mathbb{C}^N)$. Since H is also normal, it has an orthonormal set of eigenfunctions $\{\chi_j\}$ and therefore we can expand f in terms of this basis:

$$f = \sum_{j=0}^{N-1} \langle \chi_j, f \rangle \chi_j.$$

Upon applying this expansion, we calculate the Rayleigh quotient:

$$R(H, f) = \frac{\left\langle \sum_{j=0}^{N-1} \langle \chi_j, f \rangle \chi_j, \sum_{j=0}^{N-1} \langle \chi_j, f \rangle \lambda_j \chi_j \right\rangle}{\left\langle \sum_{j=0}^{N-1} \langle \chi_j, f \rangle \chi_j, \sum_{j=0}^{N-1} \langle \chi_j, f \rangle \chi_j \right\rangle} = \frac{\sum_{j=0}^{N-1} |\langle \chi_j, f \rangle|^2 \lambda_j}{\sum_{j=0}^{N-1} |\langle \chi_j, f \rangle|^2}.$$

We know that $\sum_{j=0}^{N-1} |\langle \chi_j, f \rangle|^2 = \|f\|^2$, and, hence, the Rayleigh quotient is minimized by taking $f = \chi_0$ and it is maximized by taking $f = \chi_{N-1}$, yielding the inequality

$$\lambda_0 \leq R(H, f) \leq \lambda_{N-1}. \tag{3.4}$$

Inequality (3.4) will be used for several proofs in future analysis. If we have $0 \leq \lambda_0$ for an Hermitian matrix H, then we say H is *positive semi-definite*. If there exists an

$d \times N$ matrix M such that $M^*M = H$, then for all unit norm functions $f \in l^2(\mathbb{C}^N)$ we have

$$R(H, f) = \langle f, Hf \rangle = \langle f, M^*Mf \rangle = \|Mf\|^2 \geq 0.$$

Hence any matrix of this form is positive semi-definite. If an Hermitian matrix H is positive semi-definite, then it has a diagonalization D with non-negative diagonal values, and a unitary matrix χ such that

$$H = \chi D \chi^* = \chi D^{1/2} D^{1/2} \chi^* = (D^{1/2}\chi^*)^*(D^{1/2}\chi^*).$$

Hence, H is positive semi-definite if and only if there exists a matrix M such that $M^*M = H$.

Turning to other methods of bounding the effect of linear operators on function in $l^2(\mathbb{C}^N)$, we define for any $N \times N$ square matrix A, the *operator* or *induced* norm of A:

$$\|A\|_{op} = \sup\{\frac{\|Af\|}{\|f\|} \text{ over all non-zero } f \in l^2(\mathbb{C}^N)\}.$$

If $A = (A_{jk})$ for $j, k = 0, ..., N-1$ then the *Euclidean* or *Frobenius* norm of A is given by

$$\|A\|_{fr} = \left(\sum_{j,k=0}^{N-1} |(A_{jk})|^2 \right)^{1/2}.$$

We denote the *trace* of a matrix A by

$$tr(A) = \sum_{j=0}^{N-1} A_{jj}.$$

It is straightforward to show that if A is $N \times d$ and B is $d \times N$ with complex values, we have $tr(AB) = tr(BA)$. This implies that the trace is invariant under cyclic permutations. We may conclude that the Frobenius norm of a square matrix A may also be given by

$$\|A\|_{fr}^2 = tr(A^*A) = tr(AA^*).$$

A matrix V if the form $V_{j,k} = \alpha_j^k$ for $j, k = 0, ..., N-1$ is called *Vandermonde* and is named for Alexandre-Theophile Vandermonde. See [31] for a more in depth treatment of such operators.

3.3 Finite Frames

We shall define a few of the properties of frames in the context $\mathbb{C}^d$ and $\mathbb{R}^d$ which are necessary for our analysis. This, however, does injustice to the rich and much more general theory of frames dating back to their introduction by Duffin and Schaeffer in 1952 [20], see also the article by Benedetto [7] and the two books by Christensen on the subject [13] and [14], the former providing a theoretical overview and the latter a more constructive approach. For a focused introduction to frames in the finite setting see chapter 1 in [12].

We begin by defining a frame for $\mathcal{H}^d$, where $\mathcal{H}^d$ is taken to be either $\mathbb{C}^d$ or $\mathbb{R}^d$. A set of $\{x_j\}_{j=0}^{N-1}$ functions in $l^2(\mathcal{H}^d)$ is a *frame* for $l^2(\mathcal{H}^d)$ if there exist positive constants A and B such that for all $f \in l^2(\mathcal{H}^d)$ the following inequality holds:

$$A \|f\|^2 \leq \sum_{j=0}^{N-1} |\langle f, x_j \rangle|^2 \leq B \|f\|^2 . \tag{3.5}$$

The supremum over all such A and the infimum over all such B satisfying (3.5) are the upper and lower *frame bounds*, respectively; and we refer to the property defined by (3.5) as the *frame condition*. In this finite setting, the existence of an upper bound B is trivially true. Hence, the frame condition is equivalent to the $\{x_j\}$ being a spanning set. Otherwise, $\{x_j\}$ would have a non-trivial orthogonal subspace contradicting the existence of a positive lower frame bound. If $A = B = C$ then the set $\{x_j\}$ is a *tight* frame or C-*tight* frame. If $C = 1$ then the frame is a *Parseval frame.* If $\|x_j\|$ is constant for all $j = 0, ..., N-1$ then $\{x_j\}$ is an *equal norm frame.* Of particular importance to our results is the matrix representation of the *synthesis operator* X which is the $d \times N$ matrix X whose columns are given by the N vectors in the frame $\{x_j\}$, that is,

$$X = \begin{bmatrix} x_0, & x_1, & \dots & , x_{N-1} \end{bmatrix}.$$

In this work, we abuse notation and refer to the frame $\{x_j\}$ and the matrix representation of the synthesis operator as merely the "frame." An important property of Parseval frames is that $XX^* = I_{d\times d}$, which we shall use frequently in our exposition and in our proofs.

3.4 The Discrete Fourier Transform Matrix

The discrete Fourier transform (DFT) is a fundamental tool in modern signal processing, solving partial differential equations, and performing convolutions. As with the linear algebra preliminaries, we shall only introduce some notation and fundamental properties of the DFT. For in depth treatment, see [9], [48], [51], or [11].

The $N \times N$ unitary Fourier transform matrix is given by

$$\mathcal{FT} = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & W_N^{(1)(1)} & \cdots & W_N^{(1)(N-1)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & W_N^{(N-1)(1)} & \cdots & W_N^{(N-1)(N-1)} \end{bmatrix}$$

where $W_N = e^{-2\pi i/N}$. The scaling by $\frac{1}{\sqrt{N}}$ ensures $\mathcal{FT}^* = \mathcal{FT}^{-1}$; and hence it is a unitary operator. By construction, $\mathcal{FT}$ is a Vandermonde matrix. In the ensuing chapters, we shall establish that the graph Fourier transform is not in general a Vandermonde matrix, and, as a result, the support theorems of [19] (which rely on the Vandermonde property of $\mathcal{FT}$) do not necessarily hold in the graph setting. For the purposes of this work, we refer to the unitary discrete Fourier transform matrix $\mathcal{FT}$ as the discrete Fourier transform.

3.5 A Discrete Heisenberg Uncertainty Principle

We introduce the discrete Heisenberg uncertainty principle due to Grünbaum [24]. Define operators Q and P to represent position and momentum operators affecting the so called "state" function $a \in \mathbb{C}^N$. We define

$$Q = \begin{bmatrix} q_0 & 0 & 0 & \dots & 0 \\ 0 & q_1 & 0 & \dots & 0 \\ 0 & 0 & \ddots & & 0 \\ 0 & 0 & \dots & q_{N-2} & 0 \\ 0 & 0 & \dots & 0 & q_{N-1} \end{bmatrix} \text{ and}$$

$$P = i \begin{bmatrix} 0 & 1 & 0 & 0 & \dots & -1 \\ -1 & 0 & 1 & 0 & \dots & 0 \\ 0 & -1 & 0 & 1 & \dots & 0 \\ & & \ddots & \ddots & \ddots & \\ 0 & 0 & \dots & -1 & 0 & 1 \\ 1 & 0 & \dots & 0 & -1 & 0 \end{bmatrix},$$

where the q_j are real numbers that we shall choose later. It is clear that P is self adjoint and that if we define a translation matrix

$$T = \begin{bmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ & & & \ddots & \ddots & \\ 0 & 0 & \dots & 0 & 0 & 1 \\ 1 & 0 & \dots & 0 & 0 & 0 \end{bmatrix}$$

then $P = i(T - T^*)$.

Theorem 3.5.1 *(Grünbaum [24]) For operator P, for operator Q with $q_j = \sin\left(\frac{2\pi j}{N}\right)$,* *and for a function* $a \in \mathbb{C}^N$*, the following inequality holds:*

$$\|Qa\|^2 \|Pa\|^2 = 4 \sum_{j=0}^{N-1} \sin^2\left(\frac{2\pi j}{N}\right) |a[j]|^2 \sum_{j=0}^{N-1} \sin^2\left(\frac{2\pi j}{N}\right) |\hat{a}[j]|^2$$
$$\geq \frac{1}{4}\left(\sum_{j=0}^{N-1}\left(\sin\left(\frac{2\pi j}{N}\right) + \sin\left(\frac{2\pi (j+1)}{N}\right)\right)(\bar{a}[j]a[j+1] - \bar{a}_[j+1]a[j])\right)^2$$
$$-\frac{1}{4}\left(\sum_{j=0}^{N-1}\left(\sin\left(\frac{2\pi j}{N}\right) - \sin\left(\frac{2\pi (j+1)}{N}\right)\right)(\bar{a}[j]a[j+1] + \bar{a}_[j+1]a[j])\right)^2.$$

Proof: Recall that the DFT is unitary, i.e., $\mathcal{FT}^*\mathcal{FT} = I_{N\times N}$. Hence, we have

$$\|Pa\|^2 = \langle Pa, Pa\rangle = \langle \mathcal{FT}^*\mathcal{FT}Pa, Pa\rangle = \langle \mathcal{FT}Pa, \mathcal{FT}Pa\rangle = \|\mathcal{FT}Pa\|^2,$$

so that $\|\mathcal{FT}Pa\|^2$ or, equivalently, $\|\mathcal{FT}(T - T^*)a\|^2$ will serve as the analog for $\|\gamma \hat{f}(\gamma)\|^2_{L^2(\mathbb{R})}$. Using the permutation property of T, it is easily shown for $\mathcal{FT}a = \hat{a} = [\hat{a}_0, ..., \hat{a}_{(N-1)}]'$ that

$$\mathcal{FT}Ta = \frac{1}{\sqrt{N}}\begin{bmatrix} \sum_{k=0}^{N-1} W_N^{(k-1)0} a_k \\ \sum_{k=0}^{N-1} W_N^{(k-1)1} a_k \\ \vdots \\ \sum_{k=0}^{N-1} W_N^{(k-1)(N-1)} a_k \end{bmatrix}$$
$$= \frac{1}{\sqrt{N}}\begin{bmatrix} W_N^{-0}\sum_{k=0}^{N-1} W_N^{(k)0} a_k \\ W_N^{-1}\sum_{k=0}^{N-1} W_N^{(k)1} a_k \\ \vdots \\ W_N^{-(N-1)}\sum_{k=0}^{N-1} W_N^{(k)(N-1)} a_k \end{bmatrix} = \begin{bmatrix} W_N^{-0}\hat{a}_0 \\ W_N^{-1}\hat{a}_1 \\ \vdots \\ W_N^{-(N-1)}\hat{a}_{N-1} \end{bmatrix},$$

and similarly we have

$$\mathcal{F}\mathcal{T}T^* a = \begin{bmatrix} W_N^0 \hat{a}_0 \\ W_N^1 \hat{a}_1 \\ \vdots \\ W_N^{N-1} \hat{a}_{N-1} \end{bmatrix}.$$

Therefore, the norm $\|Pa\|^2$ is given by

$$\begin{aligned}
\|Pa\|^2 &= \|\mathcal{F}\mathcal{T}(T - T^*)a\|^2 \\
&= \left\| \begin{bmatrix} (W_N^{-0} - W_N^0)\hat{a}_0 \\ (W_N^{-1} - W_N^1)\hat{a}_1 \\ \vdots \\ (W_N^{-(N-1)} - W_N^{(N-1)})\hat{a}_{N-1} \end{bmatrix} \right\|^2 \\
&= \sum_{j=0}^{N-1} \left|(W_N^{-j} - W_N^j)\hat{a}_j\right|^2 = 4\sum_{j=0}^{N-1} \sin^2\left(\frac{2\pi j}{N}\right) |\hat{a}_j|^2 . \qquad (3.6)
\end{aligned}$$

Equation (3.6) motivates our selection $q_j = \sin\left(\frac{2\pi j}{N}\right)$. With this choice, the following equality holds from direct calculation:

$$\|Qa\|^2 = \sum_{j=0}^{N-1} \sin^2\left(\frac{2\pi j}{N}\right) |a_j|^2 .$$

To establish the desired inequality, we shall use Cauchy's inequality, and the properties of the commutator C of Q and P, and of the anticommutator A of Q and

P. Since $QP = 1/2(A+C)$, the following holds:

$$\begin{aligned}
\|Qa\|^2 \|Pa\|^2 &\geq |\langle Qa, Pa\rangle|^2 \\
&= |\langle a, 1/2(A+C)a\rangle|^2 \\
&= \frac{1}{4} |\langle a, Aa\rangle - i \langle a, iCa\rangle|^2 \\
&= \frac{1}{4} \left(\langle a, Aa\rangle^2 - \langle a, Ca\rangle^2\right). \qquad (3.7)
\end{aligned}$$

We shall use inequality (3.7) to establish the desired discrete uncertainty principle. A and C are given by:

$$A = i \begin{bmatrix} 0 & q_0+q_1 & 0 & \dots & -q_0-q_{N-1} \\ -q_0-q_1 & 0 & q_1+q_2 & \dots & 0 \\ 0 & -q_1-q_2 & 0 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & q_{N-2}+q_{N-1} \\ q_0+q_{N-1} & 0 & \dots & -q_{N-2}-q_{N-1} & 0 \end{bmatrix} \text{ and}$$

$$C = i \begin{bmatrix} 0 & q_0-q_1 & 0 & \dots & q_{N-1}-q_0 \\ q_0-q_1 & 0 & q_1-q_2 & \dots & 0 \\ 0 & q_1-q_2 & 0 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & q_{N-2}-q_{N-1} \\ q_{N-1}-q_0 & 0 & \dots & q_{N-2}-q_{N-1} & 0 \end{bmatrix}.$$

Direct calculation now yields:

$$(\langle a, Aa\rangle)^2 = \left(\sum_{j=0}^{N-1} (q_j + q_{j+1})(\bar{a}_j a_{j+1} - \bar{a}_{j+1} a_j)\right)^2 \text{ and}$$

$$(\langle a, Ca\rangle)^2 = \left(\sum_{j=0}^{N-1} (q_j - q_{j+1})(\bar{a}_j a_{j+1} + \bar{a}_{j+1} a_j)\right)^2.$$

Setting $q_j = \sin\left(\frac{2\pi j}{N}\right)$ for $j = 0, ..., N-1$ and substituting into (3.7) yields the desired inequality:

$$4\sum_{j=0}^{N-1}\sin^2\left(\frac{2\pi j}{N}\right)|a_j|^2\sum_{j=0}^{N-1}\sin^2\left(\frac{2\pi j}{N}\right)|\hat{a}_j|^2$$
$$\geq \frac{1}{4}\left(\sum_{j=0}^{N-1}\left(\sin\left(\frac{2\pi j}{N}\right)+\sin\left(\frac{2\pi(j+1)}{N}\right)\right)(\bar{a}_j a_{j+1}-\bar{a}_{j+1}a_j)\right)^2$$
$$-\frac{1}{4}\left(\sum_{j=0}^{N-1}\left(\sin\left(\frac{2\pi j}{N}\right)-\sin\left(\frac{2\pi(j+1)}{N}\right)\right)(\bar{a}_j a_{j+1}+\bar{a}_{j+1}a_j)\right)^2.$$

■

This inequality is stated in terms of two linear operator's relationship with an arbitrary function. There is the question of what types of functions generate equality between

$$4\sum_{j=0}^{N-1}\sin^2\left(\frac{2\pi j}{N}\right)|a_j|^2\sum_{j=0}^{N-1}\sin^2\left(\frac{2\pi j}{N}\right)|\hat{a}_j|^2$$

and

$$\frac{1}{4}\left(\sum_{j=0}^{N-1}\left(\sin\left(\frac{2\pi j}{N}\right)+\sin\left(\frac{2\pi(j+1)}{N}\right)\right)(\bar{a}_j a_{j+1}-\bar{a}_{j+1}a_j)\right)^2$$
$$-\frac{1}{4}\left(\sum_{j=0}^{N-1}\left(\sin\left(\frac{2\pi j}{N}\right)-\sin\left(\frac{2\pi(j+1)}{N}\right)\right)(\bar{a}_j a_{j+1}+\bar{a}_{j+1}a_j)\right)^2.$$

As it turns out, if N is odd then there is a one dimensional subspace that generates equality, and if N is even there is a two dimensional space that generates equality. There is some work regarding these solutions in [23]. There are also similar results involving the DFT and frames due to Lammers and Maeser [33]. These are discussed throughout the rest of this chapter.

3.6 Finite time-frequency measures

Motivated by the additive classical uncertainty principle for $f \in \mathcal{S}(\mathbb{R})$, we introduce a discrete analog for $l^2(\mathbb{Z}/N\mathbb{Z})$ and justify the analog's use. For $f : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$, define the $N \times N$ difference operator $\mathbf{D} = I - T$, the $N \times N$ circulant difference operator by $\Delta = \mathbf{D}^*\mathbf{D}$, and the $N \times N$ modulation matrix M as a diagonal matrix where $M_{jj} = W_N^{-j}$, i.e.

$$\mathbf{D} = \begin{bmatrix} 1 & -1 & 0 & 0 & \dots & 0 \\ 0 & 1 & -1 & 0 & \dots & 0 \\ 0 & 0 & 1 & -1 & \dots & 0 \\ & & & \ddots & \ddots & \\ 0 & 0 & \dots & 0 & 1 & -1 \\ -1 & 0 & \dots & 0 & 0 & 1 \end{bmatrix},$$

$$\Delta = \begin{bmatrix} 2 & -1 & 0 & 0 & \dots & -1 \\ -1 & 2 & -1 & 0 & \dots & 0 \\ 0 & -1 & 2 & -1 & \dots & 0 \\ & & \ddots & \ddots & \ddots & \\ 0 & 0 & \dots & -1 & 2 & -1 \\ -1 & 0 & \dots & 0 & -1 & 2 \end{bmatrix} \text{ and}$$

$$M = \begin{bmatrix} W_N^0 & 0 & \dots & 0 \\ 0 & W_N^{-1} & \dots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \dots & W_N^{1-N} \end{bmatrix}.$$

It is straightforward to show that the following properties hold:

1. $\|\mathbf{D}f\|^2 = \langle \Delta f, f \rangle$,

2. $M^{-1} = M^*$,

3. $T\mathcal{F}\mathcal{T}^* = \mathcal{F}\mathcal{T}^* M$, and

4. $\mathcal{F}\mathcal{T}^*T^* = M\mathcal{F}\mathcal{T}^*$.

We define the $N \times N$ matrix $X = \mathcal{F}\mathcal{T}^*\Delta\mathcal{F}\mathcal{T}$. Using the aforementioned properties, it is straightforward to show that X is diagonal with real entries. In fact, we have $\Delta = \mathbf{D}^*\mathbf{D}$ is positive semidefinite, so it has real positive eigenvalues. Further, since $\Delta = 2I - T - T^{N-1}$, the Fourier transform matrix is a matrix of the orthonormal eigenfunctions of Δ. Hence, the diagonal entries of X, for $j = 0, ..., N-1$, are

$$\lambda_j 2 - W_N^j - W_N^{-j} = -2\cos(2\pi j/N) + 2 = 4\sin^2(\pi j/N) \in \mathbb{R}.$$

We shall use

$$\langle Xf, f\rangle = \langle \mathcal{F}\mathcal{T}^*\Delta\mathcal{F}\mathcal{T}f, f\rangle = \langle \mathcal{F}\mathcal{T}^*\mathbf{D}^*\mathbf{D}\mathcal{F}\mathcal{T}f, f\rangle = \langle \mathbf{D}\mathcal{F}\mathcal{T}f, \mathbf{D}\mathcal{F}\mathcal{T}f\rangle = \|\mathbf{D}\mathcal{F}\mathcal{T}f\|^2$$

for $f \in l^2(\mathbb{Z}/N\mathbb{Z})$ as a discrete analog of $\left\|\frac{d}{d\gamma}\widehat{f}(\gamma)\right\|^2$ on L^2 functions.

In an attempt to further motivate using $\|\mathbf{D}f\|^2 + \|\mathbf{D}\mathcal{F}\mathcal{T}f\|^2$ as our discrete analog for the continuous case, let us consider $g \in l^2(\mathbb{Z}/N\mathbb{Z})$ for $N = n^2$ and $n \in \mathbb{Z}$. Given such a g and $1 >> \varepsilon > 0$, we can construct a smooth function $h \in L^2(\mathbb{R})$ so that

$$\left| \|g\|^2_{l^2(\mathbb{Z}/N\mathbb{Z})} - \|h\|^2_{L^2(\mathbb{R})} \right| \leq \varepsilon \|g\|^2_{l^2(\mathbb{Z}/N\mathbb{Z})}.$$

In fact, let $\lfloor x \rfloor$ be the floor function of x, fix $j \in \{\lfloor -N/2 \rfloor, ..., \lfloor (N-1)/2 \rfloor\}$, and, for $x \in I_j = [\frac{2j-1}{2n}, \frac{2j+1}{2n}]$, define $h(x)$ as follows:

$$h(x) = \sqrt{n}g(j)\chi_{[\frac{2j-1+\varepsilon}{2n}, \frac{2j+1-\varepsilon}{2n}]} * \phi_{\varepsilon/n}(x),$$

where $\phi_{\varepsilon/n}$ is a C^∞ mollifier with support in $[\frac{-\varepsilon}{2n}, \frac{\varepsilon}{2n}]$ and where $*$ denotes convolution. Then we have

$$(1-\varepsilon)\left|g(j)\right|^2 \leq \left\|h(x)\right\|^2_{L^2(I_j)} \leq \left|g(j)\right|^2.$$

If we define $h(x) = 0$ for $x \geq \lfloor (N-1)/2 \rfloor + 1/n$ and for $x \leq -\lfloor N/2 \rfloor - 1/n$, then summing over j yields

$$(1-\varepsilon)\left\|g\right\|^2_{l^2(\mathbb{Z}/N\mathbb{Z})} \leq \left\|h\right\|^2_{L^2(\mathbb{R})} \leq \left\|g\right\|^2_{l^2(\mathbb{Z}/N\mathbb{Z})}$$

and hence

$$\left| \left\|g\right\|^2_{l^2(\mathbb{Z}/N\mathbb{Z})} - \left\|h\right\|^2_{L^2(\mathbb{R})} \right| \leq \varepsilon \left\|g\right\|^2_{l^2(\mathbb{Z}/N\mathbb{Z})},$$

the desired estimate.

Note that

$$\begin{aligned}
\left\|\mathbf{D}g\right\|^2 &= \langle \Delta g, g \rangle \\
&= \langle \mathcal{FT} X \mathcal{FT}^* g, g \rangle \\
&= \langle X \mathcal{FT}^* g, \mathcal{FT}^* g \rangle \\
&= 4 \sum_{j=0}^{N-1} \sin^2(\frac{\pi j}{N}) \left|(\mathcal{FT}^* g)(j)\right|^2 \\
&= 4 \sum_{j=0}^{N-1} \sin^2(\frac{\pi j}{N}) \left|(\mathcal{FT} g)(j)\right|^2,
\end{aligned}$$

and that

$$\begin{aligned}
\|\mathbf{D}\mathcal{F}\mathcal{T}g\|^2 &= \langle \mathbf{D}\mathcal{F}\mathcal{T}g, \mathbf{D}^*\mathcal{F}\mathcal{T}^*g\rangle \\
&= \langle \mathcal{F}\mathcal{T}^*\Delta\mathcal{F}\mathcal{T}g, g\rangle \\
&= \langle Xg, g\rangle \\
&= 4\sum_{j=0}^{N-1}\sin^2(\frac{\pi j}{N})\,|(g)(j)|^2 .
\end{aligned}$$

Using these facts, and the approximations $\sin(x) \approx x$ for $-\pi/2 \leq x \leq \pi/2$ and $\|g\|_{l^2(\mathbb{Z}_N)} \approx \|h\|^2_{L^2(\mathbb{R})}$, we have the following:

$$\begin{aligned}
N(\|\mathbf{D}g\|^2 + \|\mathbf{D}\mathcal{F}\mathcal{T}g\|^2) &= 4N\left(\sum_{j=0}^{N-1}\sin^2(\frac{\pi j}{N})\,|(\mathcal{F}\mathcal{T}g)(j)|^2 + \sum_{j=0}^{N-1}\sin^2(\frac{\pi j}{N})\,|(g)(j)|^2\right) \\
&\approx 4N\left(\sum_{j=0}^{N-1}(\frac{\pi j}{N})^2\,|(\mathcal{F}\mathcal{T}g)(j)|^2 + \sum_{j=0}^{N-1}(\frac{\pi j}{N})^2\,|(g)(j)|^2\right) \\
&\approx 4N\left(\sum_{j=-\lfloor N/2\rfloor}^{\lfloor (N-1)/2\rfloor}(\frac{\pi j}{N})^2\left|\frac{\widehat{h}(j/n)}{\sqrt{n}}\right|^2 + \sum_{j=-\lfloor N/2\rfloor}^{\lfloor (N-1)/2\rfloor}(\frac{\pi j}{N})^2\left|\frac{h(j/n)}{\sqrt{n}}\right|^2\right) \\
&= 4N\frac{\pi^2}{N}\left(\sum_{j=-\lfloor N/2\rfloor}^{\lfloor (N-1)/2\rfloor}(\frac{j}{n})^2\left|\widehat{h}(j/n)\right|^2\frac{1}{n} + \sum_{j=-\lfloor N/2\rfloor}^{\lfloor (N-1)/2\rfloor}(\frac{j}{n})^2\,|h(j/n)|^2\,\frac{1}{n}\right) \\
&\approx (4/\pi^2)\left(\int \gamma^2\left|\widehat{h}(\gamma)\right|^2 d\gamma + \int x^2\,|h(x)|^2\,dx\right) \\
&= (4/\pi^2)\left(\left\|\gamma\widehat{h}(\gamma)\right\|^2 + \|xh(x)\|^2\right) \\
&\geq C\,\|h\|^4_{L^2(\mathbb{R})} \\
&\approx C\,\|g\|^4_{l^2(\mathbb{Z}/N\mathbb{Z})} .
\end{aligned}$$

Thus, it seems reasonable to assume this choice for measurement of $l^2(\mathbb{Z}/\mathbb{Z}N)$ will lead to interesting uncertainty principles.

3.7 Uncertainty Principles for Parseval frames

Define the matrix $E = [E_0, E_1, ..., E_{N-1}]$ to be a $d \times N$ matrix where the set of N $d-$vectors $\{E_k\}$ forms a Parseval frame for $\mathbb{C}^d$, i.e., $EE^* = I_{d\times d}$. We want to perform analysis on E using the the difference operator $\mathbf{D}$ as define in the previous section. Using the Frobenius norm for matrices, we have the following:

$$\begin{aligned} \|\mathbf{D}\mathcal{F}\mathcal{T}E\|^2 + \|\mathbf{D}E\|^2 &= tr(\mathbf{D}\mathcal{F}\mathcal{T}EE^*\mathcal{F}\mathcal{T}^*\mathbf{D}^*) + tr(\mathbf{D}EE^*\mathbf{D}^*) \\ &= 2tr\Delta \\ &= 4d. \end{aligned}$$

Hence, analysis on E with the difference operator $\mathbf{D}$ is not interesting as the Frobenius norms solely depend on the dimension d. Instead we shall analyze E^*. In this case, we have

$$\begin{aligned} \|\mathbf{D}\mathcal{F}\mathcal{T}E^*\|^2 + \|\mathbf{D}E^*\|^2 &= tr(\mathbf{D}\mathcal{F}\mathcal{T}E^*E\mathcal{F}\mathcal{T}^*\mathbf{D}^*) + tr(\mathbf{D}EE^*\mathbf{D}^*) \\ &= tr(\mathcal{F}\mathcal{T}^*\mathbf{D}^*\mathbf{D}\mathcal{F}\mathcal{T}E^*E) + tr(\mathbf{D}^*\mathbf{D}E^*E) \\ &= tr(XE^*E) + tr(\Delta E^*E). \end{aligned}$$

If $N = d$, then E is square and unitary. Hence, the frame is an orthonormal basis for $l^2(\mathbb{C}^d)$. The following lemma establishes a starting point for bounding the Frobenius norm of D acting on E^*.

Lemma 3.7.1 *For all equal normed Parseval frames E for $\mathbb{C}^d$, the following holds:* $\|\mathbf{D}\mathcal{F}\mathcal{T}E^*\|^2 = 2d$.

Proof: Since $EE^* = I_{d\times d}$ we know $tr(EE^*) = d$. Hence we have

$$\begin{aligned} d &= tr(E^*E) \\ &= \sum_{j=0}^{N-1} \|E_j\|^2 \\ &= N\|E_0\|^2 \end{aligned}$$

where the last equality is due to E being an equal norm Parseval frame. Hence, the j^{th} diagonal element of E^*E is given by $(E^*E)_{jj} = \|E_j\|^2 = \frac{d}{N}$. Noting that

$$\|\mathbf{D}\mathcal{F}\mathcal{T}E^*\|^2 = tr(\mathbf{D}\mathcal{F}\mathcal{T}E^*E\mathcal{F}\mathcal{T}^*\mathbf{D}^*) = tr(\mathcal{F}\mathcal{T}^*\Delta\mathcal{F}\mathcal{T}E^*E) = tr(XE^*E)$$

and that $2N = tr(\Delta) = tr(X)$, we can conclude that

$$\|\mathbf{D}\mathcal{F}\mathcal{T}E^*\|^2 = \frac{d}{N}tr(X) = 2N\frac{d}{N} = 2d,$$

as desired. ■

Corollary 3.7.2 *If E is a Unitary matrix, that is, if $d = N$ and E is a Parseval frame, then $\|\mathbf{D}\mathcal{F}\mathcal{T}E^*\|^2 + \|\mathbf{D}E^*\|^2 = 4d$.*

Proof: By assumption $E^* = E^{-1}$, hence we have $\|\mathbf{D}E^*\|^2 = tr(\mathbf{D}E^*E\mathbf{D}^*) = tr(\mathbf{D}\mathbf{D}^*) = 2d$. ■

We establish a minimization lemma that will be used to determine bounds for $\|\mathbf{D}E^*\|^2$.

Lemma 3.7.3 *Let $\{\alpha_j\}$ be a set of N real numbers $0 \le \alpha_j \le 1$ with $\sum_{j=1}^{N-1}\alpha_j = d$, and let $\{\lambda_j\}$ be the N eigenvalues of Δ (i.e., the diagonal values of X) ordered from smallest to largest. Then the sum $\sum_{j=0}^{N-1}\alpha_j\lambda_j$ is minimized, respectively, maximized,* *by setting* $\alpha_0, ..., \alpha_{d-1} = 1$*, respectively,* $\alpha_0, ..., \alpha_{d-1} = 0$*, and* $\alpha_d, ..., \alpha_{N-1} = 0$*, respectively,* $\alpha_d, ..., \alpha_{N-1} = 1$.

Proof: In the case of $N = d$ the lemma is trivially true, so assume $N > d$. Suppose $\alpha_0, ..., \alpha_{d-1} = 1$ and $\alpha_d, ..., \alpha_{N-1} = 0$ so that $\sum_{j=0}^{N-1} \alpha_j \lambda_j = \sum_{j=0}^{d-1} \alpha_j \lambda_j$. Let $m > d-1$. Then, for all $0 \leq k \leq d-1$, $\lambda_m \geq \lambda_k$, and thus for any $1 \geq \varepsilon > 0$ we have $\varepsilon(\lambda_m - \lambda_k) \geq 0$. Adding the sum $\sum_{j=0}^{N-1} \alpha_j \lambda_j$ to both sides of this inequality yields:

$$\sum_{j=0}^{d-1} \alpha_j \lambda_j - \varepsilon \lambda_k + \varepsilon \lambda_m \geq \sum_{j=0}^{d-1} \alpha_j \lambda_j.$$

Pulling the λ_k term out of the summation formula shows the desired inequality:

$$\sum_{j \leq d-1, j \neq k} \alpha_j \lambda_j + (1-\varepsilon)\lambda_k + \varepsilon \lambda_m \geq \sum_{j=0}^{d-1} \alpha_j \lambda_j.$$

Hence, any sum where $\alpha_j > 0$ for $j \geq d$ is greater than or equal to $\sum_{j=0}^{d-1} \lambda_j$. Similarly, setting the last d coefficients to 1 maximizes the sum. ■

Theorem 3.7.4 *(Lammers and Maeser [33]) For fixed dimension* d *and* $N \geq d \geq 2$*, there exist constants* $C(N,d) > 0$ *and* $B(N,d) > 0$ *so that for any equal norm Parseval frame* E *for* $\mathbb{C}^d$*, we have*

$$\begin{aligned} 2d + C(N,d) &\leq \|\mathbf{D}\mathcal{F}\mathcal{T}E^*\|^2 + \|\mathbf{D}E^*\|^2 \\ &\leq 2d + B(N,d) \\ &\leq 6d. \end{aligned}$$

Furthermore, the minimum, respectively, the maximum, occurs when E^* *is the* d *columns of the Fourier matrix corresponding to the* d *smallest, respectively, the* d *largest, eigenvalues of* Δ*. The constant* $C(N,d)$ *is the sum of those* d *smallest eigenvalues and* $B(N,d)$ *is the sum of those* d *largest eigenvalues.*

Proof: Due to Lemma 3.7.1, it suffices to find the minimizer of $\|\mathbf{D}E^*\|_{fr}^2$in order to find the lower limit. Expanding $\|\mathbf{D}E^*\|_{fr}^2$ we have:

$$\begin{aligned}\|\mathbf{D}E^*\|_{fr}^2 &= tr(\mathbf{D}E^*E\mathbf{D}^*)\\ &= tr(\Delta E^*E)\\ &= tr(\mathcal{FT}^*X\mathcal{FT}E^*E)\\ &= tr(X\mathcal{FT}E^*E\mathcal{FT}^*).\end{aligned}$$

We have $tr(\mathcal{FT}E^*E\mathcal{FT}^*) = tr(E\mathcal{FT}^*\mathcal{FT}E^*) = tr(I_{d\times d}) = d$. Since $E\mathcal{FT}^*$ is an equal norm Parseval frame, the diagonal elements of $\mathcal{FT}E^*E\mathcal{FT}^*$ are the norm squared of the columns of $E\mathcal{FT}^*$ and hence greater than or equal to zero. Further, we have $\|E^*\|_{op} = \|\mathcal{FT}\|_{op} = 1$. Therefore, each diagonal element of $\mathcal{FT}E^*E\mathcal{FT}^*$ is bounded between zero and one. By Lemma 3.7.3, we minimize $tr(X\mathcal{FT}E^*E\mathcal{FT}^*)$ if there exists a frame E such that $\mathcal{FT}E^*E\mathcal{FT}^*$ has canonical basis functions e_j for $\mathbb{C}^d$ in the d columns corresponding to the d smallest diagonals of X. Choosing the d rows of $\mathcal{FT}$ corresponding to those d values accomplishes this. Hence, $C(N,d)$ is the sum of the d smallest eigenvalues of Δ. Similarly, we want to maximize $\|\mathbf{D}E^*\|^2$ in order to find the upper limit. Via the same style argument we conclude that $B(N,d)$ is the sum of the d largest eigenvalues of Δ. ■

If we drop the assumption that we have an equal norm Parseval frame, and we only assume that E is Parseval, then we have a similar result also due to Lammers and Maeser.

Theorem 3.7.5 *(Lammers and Maeser [33]) For fixed dimension d and $N \geq d \geq 2$, there exist constants $L(N,d) > 0$ and $U(N,d) > 0$ so that for any Parseval frame*

E of $\mathbb{C}^d$,

$$\begin{aligned} L(N,d) &\leq \|\mathbf{D}\mathcal{F}\mathcal{T}E^*\|^2 + \|\mathbf{D}E^*\|^2 \\ &\leq U(N,d) \\ &\leq 8d. \end{aligned}$$

$L(N,d)$ is the sum of the d smallest eigenvalues of $\Delta + X$ and $U(N,d)$ is the sum of the d largest eigenvalues of $\Delta + X$.

Proof: Recall from the proof of Theorem 3.7.4 that

$$\begin{aligned} \|\mathbf{D}E^*\|^2 &= tr(\mathbf{D}E^*E\mathbf{D}^*) \\ &= tr(\Delta E^*E) \\ &= tr(\mathcal{F}\mathcal{T}^*X\mathcal{F}\mathcal{T}E^*E) \\ &= tr(X\mathcal{F}\mathcal{T}E^*E\mathcal{F}\mathcal{T}^*). \end{aligned}$$

We also have that

$$\begin{aligned} \|\mathbf{D}\mathcal{F}\mathcal{T}E^*\|^2 &= tr(\mathbf{D}\mathcal{F}\mathcal{T}E^*E\mathcal{F}\mathcal{T}^*\mathbf{D}^*) \\ &= tr(\Delta\mathcal{F}\mathcal{T}E^*E\mathcal{F}\mathcal{T}^*). \end{aligned}$$

Noting that since $\Delta + X$ is real and symmetric, there exists a unitary matrix U such that

$$U^*(\Delta + X)U = diag(\tilde{\lambda}_0, ..., \tilde{\lambda}_{N-1}) = \Lambda,$$

i.e., it diagonalizes $\Delta + X$. Combining these with the previous two equalities yields

$$\begin{aligned}
\|\mathbf{D}\mathcal{F}\mathcal{T}E^*\|^2 + \|\mathbf{D}E^*\|^2 &= tr(\Delta\mathcal{F}\mathcal{T}E^*E\mathcal{F}\mathcal{T}^*) + tr(X\mathcal{F}\mathcal{T}E^*E\mathcal{F}\mathcal{T}^*) \\
&= tr((\Delta + X)\mathcal{F}\mathcal{T}E^*E\mathcal{F}\mathcal{T}^*) \\
&= tr(\Lambda U^*\mathcal{F}\mathcal{T}E^*E\mathcal{F}\mathcal{T}^*U) \\
&= \sum_{j=0}^{N-1} \tilde{\lambda}_j \left(U^*\mathcal{F}\mathcal{T}E^*E\mathcal{F}\mathcal{T}^*U_{j,j}\right).
\end{aligned}$$

The operator $\Delta + X$ is also positive semidefinite as the linear combination of such operators is always positive semidefinite. Without loss of generality, assume $0 \leq \tilde{\lambda}_0 \leq ... \leq \tilde{\lambda}_{N-1}$. By the same minimizing arguments as in Theorem 3.7.4, choosing the first d rows of $U^*\mathcal{F}\mathcal{T}$ for E yields a Parseval frame that minimizes $\sum_{j=0}^{N-1} \tilde{\lambda}_j \left(U^*\mathcal{F}\mathcal{T}E^*E\mathcal{F}\mathcal{T}^*U_{j,j}\right)$. Similarly, choosing E to be the last d rows of $U^*\mathcal{F}\mathcal{T}$ maximizes $\sum_{j=0}^{N-1} \tilde{\lambda}_j \left(U^*\mathcal{F}\mathcal{T}E^*E\mathcal{F}\mathcal{T}^*U_{j,j}\right)$. ■

Chapter 4: Graph Theory

4.1 Introduction

Graph theory is a well established branch of mathematics with several comprehensive overviews of the material including [16], and [18]. The analysis of graphs is used in many applications in modern computing and information theory. [15] is a brief literature review of recent advances in graph theory. The Fourier transform on a graph has been defined using the spectrum of the graph Laplacian, see, e.g., [25], [46], [45], [44], [42], [39], [21], [17], and [1]. In this chapter, we introduce general graph theory definitions, establish notation, and define the graph Fourier transform and graph normalized Fourier transform.

4.2 Definitions

A graph $G = \{V, \mathbf{E} \subseteq V \times V, w\}$ consists of a set V of vertices, a set $\mathbf{E}$ of edges consisting of pairs of elements of V, and a weight function $w : V \times V \to \mathbb{R}^+$. For $u, v \in V$, $w(u, v) > 0$ if $(u, v) \in \mathbf{E}$ and is zero otherwise. If $w(u, v) = 1$ for all $(u, v) \in \mathbf{E}$, then we say G is "unit weighted." There is no restriction on the size of the set V, but we shall restrict our attention to $|V| = N < \infty$. We also assume that the set $\{v_j\}_{j=0}^{N-1} = V$ has an arbitrary, but fixed ordering.

For all graphs, we define the $N \times N$ *adjacency* matrix $A = (A_{m,n})$ component-wise as $(A_{m,n}) = w(v_m, v_n)$. If A is symmetric, that is, if $w(v_n, v_m) = (A_{n,m}) = (A_{m,n}) = w(v_m, v_n)$, then we say G is undirected. If a graph has loops, that is $w(v_j, v_j) > 0$ for some $v_j \in V$, then A has nonzero diagonal entries. Unless otherwise specified, we shall assume that our graphs are undirected and have no loops. The *degree* d of a vertex v_j is defined by $deg(v_j) = \sum_{n=0}^{N-1} w(v_j, v_n) = \sum_{n=0}^{N-1} (A_{j,n})$. We can then define a diagonal *degree matrix* $D = \text{diag}\,(deg(v_0), deg(v_1), ..., deg(v_{N-1}))$.

(a) (b)

Figure 4.1: The unit weighted graph G_a is shown in (a), and the graph G_b is shown in (b)

Example 4.2.1 *The graph G_a shown in Figure 4.1a has the set of vertices $V_a = \{0, 1\}$, the edge set $E_a = \{(0,1) = (1,0)\}$, and $w_a(u, v)$ defined by $w(0,1) = w(1,0) =$*

1. *The associated adjacency matrix is* $A_a = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. *The graph* G_b *with weight function* w_b *in Figure 4.1b is given by*

$$G_b = \{V_b = \{0, 1, 2, 3\}, E_b = \{(0, 1), (0, 2), (0, 3), (1, 2), (2, 3)\}, w_b\}$$

with the adjacency matrix

$$A_b = \begin{bmatrix} 0 & 1 & 2 & 3 \\ 1 & 0 & 1 & 0 \\ 2 & 1 & 0 & 1 \\ 3 & 0 & 1 & 0 \end{bmatrix}.$$

The degree matrix D_b *for the graph in Figure 4.1b is given by*

$$D_b = \begin{bmatrix} 6 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}.$$

4.3 The Graph Laplacian

There are two common choices for the graph Laplacian:

$$\begin{aligned} L &= D - A \\ \mathcal{L} &= I - D^{-1/2} A D^{-1/2}, \end{aligned}$$

where I is the $N \times N$ identity. L is defined as the *unnormalized graph Laplacian*, while $\mathcal{L}$ is defined as the *normalized graph Laplacian*. We shall refer to the unnormalized Laplacian L as the Laplacian, and to the normalized Laplacian $\mathcal{L}$ as the normalized Laplacian. Define the $|\mathbf{E}| \times |V|$ *incidence matrix* $M = (M_{k,j})$ with element $(M_{k,j})$ for edge $\mathbf{e}_k$ and vertex v_j by:

$$(M_{k,j}) = \begin{cases} 1, & \text{if } \mathbf{e}_k = (v_j, v_l) \text{ and } j < l \\ -1, & \text{if } \mathbf{e}_k = (v_j, v_l) \text{ and } j > l \\ 0, & \text{otherwise.} \end{cases}$$

Define the diagonal $|\mathbf{E}| \times |\mathbf{E}|$ *weight matrix* $W = diag(w(\mathbf{e}_0), w(\mathbf{e}_1), ..., w(\mathbf{e}_{|\mathbf{E}|-1}))$. Hence, if G is unit weighted, then $W = I_{|\mathbf{E}|\times|\mathbf{E}|}$. For any connected graph, the size of the edge set $\mathbf{E}$ is bounded as follows:

$$N - 1 \leq |\mathbf{E}| \leq \frac{N(N-1)}{2},$$

where the lower bound is attained by the path graph and the upper bound is attained by the complete graph.

Noting that $L = M^*WM = \left(W^{\frac{1}{2}}M\right)^* \left(W^{\frac{1}{2}}M\right)$, where $\cdot^*$ denotes the conjugate transpose of an operator $\cdot$, we conclude that L is real, symmetric, and positive semidefinite. By the spectral theorem (Theorem 3.2.1), L must have an orthonormal eigenbasis $\{\chi_l\}$ of eigenvectors with associated eigenvalues $\{\lambda_l\}$ ordered as $0 = \lambda_0 \leq \lambda_1 \leq \lambda_2 \leq ... \leq \lambda_{N-1}$. The kernel has dimension equal to the number of connected components of G. Indeed, any function that is constant and nonzero on connected vertices while zero on all other vertices is in the kernel of L. Hence, if G is connected, $\lambda_0 = 0$ has multiplicity 1. Let χ be the matrix whose l^{th} column is given by χ_l. Let Δ be the diagonalization of L, that is, $\chi^* L \chi = \Delta = diag(\lambda_0, ..., \lambda_{N-1})$. We shall use this set of eigenfunctions to define the graph Fourier transform.

Alternatively, after noting that

$$\mathcal{L} = D^{1/2} L D^{1/2} = \left(W^{\frac{1}{2}} M D^{1/2}\right)^* \left(W^{\frac{1}{2}} M D^{1/2}\right),$$

we may apply the spectral theorem to $\mathcal{L}$. Hence, $\mathcal{L}$ must have an orthonormal eigenbasis $\{F_l\}$ with associated eigenvalues $\{\mu_l\}$ ordered as $0 = \mu_0 < \mu_1 \leq \mu_2 \leq ... \leq \mu_{N-1}$. Let $\mathcal{F}$ be the matrix whose l^{th} column is given by F_l such that $\mathcal{F}$ diagonalizes $\mathcal{L}$. We shall use this set of eigenfunctions to define the normalized graph Fourier transform.

4.4 The Graph Fourier Transform

Functions $\widetilde{f} : V \to \mathbb{R}$ will be written notationally as vectors $f \in \mathbb{R}^N$ with $\widetilde{f}(v_j) = f[j]$ for $j = 0, ..., N-1$. We say $\widetilde{f} \in l^2(V)$, or, equivalently, $f \in l^2(V)$, if $\sum_{j=0}^{N-1} \left|\widetilde{f}(v_j)\right|^2 = \sum_{j=0}^{N-1} |f[j]|^2$. Given this space $l^2(V)$ of real-valued functions on the set V of vertices of the graph G, it is natural to define a Fourier transform based on the structure of G.

To motivate this definition, we recall from Equation (2.1), the *Fourier transform* on $L^1(\mathbb{R})$, viz.,

$$\widehat{f}(\gamma) = \int_{\mathbb{R}} f(t) e^{-2\pi i t\gamma} d\gamma,$$

and from (2.2), the formal *inverse Fourier transform*, viz.,

$$f(t) = \int_{\widehat{\mathbb{R}}} \widehat{f}(\gamma) e^{2\pi i t\gamma}\, d\gamma,$$

where $\widehat{\mathbb{R}} = \mathbb{R}$ is considered the frequency domain. The functions, $e^{2\pi i t\gamma}$, on $\mathbb{R}$ where $\gamma \in \widehat{\mathbb{R}}$, are the eigenfunctions of the Laplacian operator $\frac{d^2}{dt^2}$ since we have

$\frac{d^2}{dt^2}e^{2\pi it\gamma} = -4\pi^2\gamma^2 e^{2\pi it\gamma}$. If $\widehat{f} \in L^1(\widehat{\mathbb{R}})$, then the inverse Fourier transform is an expansion of the function f in terms of the eigenfunctions with coefficients $\widehat{f}(\gamma)$. With this in mind, we use the eigenvectors of the graph Laplacian to define the *graph Fourier transform* $\widehat{f}$ of $f \in l^2(V)$ as follows:

$$\forall l = 0, 1, ..., N-1, \quad \widehat{f}[l] = \langle \chi_l, f \rangle,$$

or, equivalently, $\widehat{f} = \chi^* f$. It is clear from the orthonormality of the basis, $\{\chi_l\}$, that $\chi^* = \chi^{-1}$. Thus, the *inverse graph Fourier transform* is given by

$$\chi\widehat{f} = \chi\chi^* f = If = f,$$

or, equivalently, $f[j] = \sum_{l=0}^{N-1} \langle \chi_l, f \rangle \chi_l[j]$.

Similarly, we define the *normalized graph Fourier transform* $\overset{*}{f}$ of $f \in l^2(V)$ as follows:

$$\forall l = 0, 1, ..., N-1, \quad \overset{*}{f}[l] = \langle F_l, f \rangle,$$

or, equivalently, $\overset{*}{f} = \mathcal{F}^* f$. It is clear from the orthonormality of the basis, $\{F_l\}$, that $\mathcal{F}^* = \mathcal{F}^{-1}$. Thus, the *inverse normalized graph Fourier transform* is given by

$$\mathcal{F}\overset{*}{f} = \mathcal{F}\mathcal{F}^* f = If = f,$$

or, equivalently, $f[j] = \sum_{l=0}^{N-1} \langle F_l, f \rangle F_l[j]$.

Example 4.4.1 *An interesting special case of the graph Fourier transform occurs when the graph is an unit weighted circulant graph as in Figure 4.2. The matrix for*

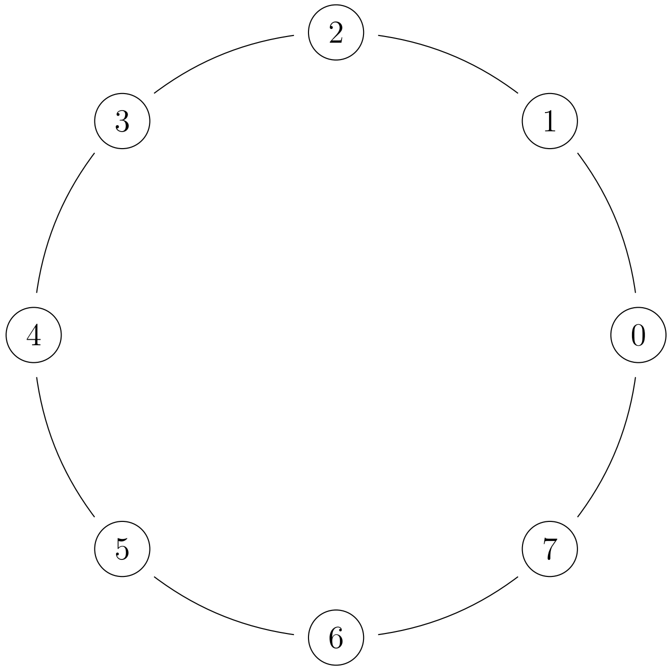

Figure 4.2: A unit weighted circulant graph with 8 vertices. The graph Laplacian associated with this graph is the circulant difference operator Δ.

the Laplacian is given by

$$L = \begin{bmatrix} 2 & -1 & 0 & & \cdots & & 0 & -1 \\ -1 & 2 & -1 & & & & & 0 \\ 0 & \ddots & \ddots & \ddots & & & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & & & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & & & \ddots & \ddots & \ddots & 0 \\ 0 & & & & & -1 & 2 & -1 \\ -1 & 0 & \cdots & \cdots & \cdots & 0 & -1 & 2 \end{bmatrix}.$$

The normalized graph Laplacian has the form

$$\mathcal{L} = \begin{bmatrix} 1 & -1/2 & 0 & & \cdots & & 0 & -1/2 \\ -1/2 & 1 & -1/2 & & & & & 0 \\ 0 & \ddots & \ddots & \ddots & & & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & & & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & & & \ddots & \ddots & \ddots & 0 \\ 0 & & & & & -1/2 & 1 & -1/2 \\ -1/2 & 0 & \cdots & \cdots & \cdots & 0 & -1/2 & 1 \end{bmatrix}.$$

Recall from Section 3.5, the $N \times N$ translation matrix T is defined by

$$(T_{i,j}) = \begin{cases} 1 & i = j-1 \\ 1 & i = N-1, j = 0 \\ 0 & \textit{otherwise.} \end{cases}$$

The Laplacian is given by

$$L = 2T^0 - T - T^{N-1},$$

where $T^0 = I$ is the $N \times N$ identity. Similarly, we have that

$$\mathcal{L} = D^{-1/2} 2T^0 D^{-1/2} - D^{-1/2} T D^{-1/2} - D^{-1/2} T^{N-1} D^{-1/2} = T^0 - \frac{1}{2}T - \frac{1}{2}T^{N-1}.$$

If $0 \leq j \leq N-1$, then an orthonormal eigenbasis for T^j is given by

$$\chi_l = \left(1/\sqrt{N}\right)[W^{0l}, W^{1l}, ..., W^{(N-1)l}]^*,$$

for $W = e^{-2\pi i/N}$ and $l = 0, 1, ..., N-1$. Indeed, we have $T^j \chi_l = W^{-jl}\chi_l$, and so χ_l is an eigenvector with the associated eigenvalue W^{-jl}. Therefore, L has the set $\{\chi_l\}$ of *orthonormal eigenvectors, with eigenvalues* $\lambda_l = -2\cos(2\pi l/N) + 2 = 4\sin^2(\pi l/N)$ *for* $l = 0, ..., N-1$.

Recall from Section 3.4, the unitary $N \times N$ *discrete Fourier transform (DFT) matrix is*

$$\mathcal{FT} = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & W_N^{(1)(1)} & \cdots & W_N^{(1)(N-1)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & W_N^{(N-1)(1)} & \cdots & W_N^{(N-1)(N-1)} \end{bmatrix}.$$

Therefore $\Lambda = (\mathcal{FT})^* PM$ *is the matrix whose columns are formed by the set* $\{\chi_l\}$ *reordered as* $\{\chi_{l_j}\}$ *for* $j = 0, ..., N-1$ *such that the columns are arranged in ascending order of their eigenvalues, and where* PM *is the permutation matrix that achieves this reordering. Hence, the graph Fourier transform associated with the circulant graph is given by*

$$\Lambda^* f = \begin{bmatrix} \langle \chi_{l_0}, f \rangle \\ \vdots \\ \langle \chi_{l_{N-1}}, f \rangle \end{bmatrix} = PM^*(\mathcal{FT})f.$$

Hence, we may view the graph Fourier transform as a permutation of the discrete Fourier transform. Since the normalized Laplacian has the same eigenvectors as the Laplacian for the circulant graph, the normalized graph Fourier transform $\mathcal{F}$ *can be viewed as a permutation of classical DFT as well.*

Graphs, similar to those in Example 4.4.1, provide an additional motivation for defining the graph Fourier transform by way of eigenvectors of the graph Laplacian. In fact, the DFT is essentially a special case of the graph Fourier transform.

Motivated by this example, we shall examine general uncertainty principles that arise from the graph setting in Chapter 6.

Chapter 5: The Spectral Graph Wavelet Transform

5.1 Introduction

Based on the work in [25], we define the classical Continuous Wavelet Transform (CWT) for functions $f \in L^2(\mathbb{R})$, and show how scaling can be accomplished in the Fourier domain. The motivation being that scaling the vertices of a graph is an ill defined operation, while scaling in the Fourier domain of the Graph Fourier transform can be defined in a fashion analogous to scaling elements of the spectrum of the CWT. The results from [25] serve as a motivation for the main results of this thesis. Specifically, manipulations in the Fourier domain will be a vital tool in the analysis in this chapter, and ensuing chapters.

5.2 Classical Wavelet Transforms

We introduce the classical wavelet transform. Wavelets have an interesting and varied history with origins in pure mathematics as well as many areas of physics and petroleum engineering. For history, the introduction to [8] has excellent insights. For a general overview see [27] or for a signal processing oriented analysis see [36] or [37].

Let $\psi(t)$ be a wavelet, that is a function such that the set of translations and dilations,

$$\left\{\psi_{s,a}(t) = \frac{1}{s}\psi\left(\frac{t-a}{s}\right)\right\}$$

for $s > 0$ and $a, s \in \mathbb{R}$ forms a spanning set for the set of L^2 functions on $\mathbb{R}$. For a given function $f \in L^2(\mathbb{R})$, the wavelet coefficient $W_f(s,a)$ at scale s, and location a for f is given by:

$$W_f(s,a) = \int_{\mathbb{R}} \frac{1}{s}\overline{\psi}\left(\frac{t-a}{s}\right) f(t)dt.$$

These are the coefficients for representing f as an expansion of the wavelet set. If the Fourier transform of ψ, given by

$$\widehat{\psi}(\gamma) = \int_{\mathbb{R}} \psi(t)e^{-2\pi i\gamma t}dt,$$

satisfies the admissibility condition

$$\int_{\mathbb{R}+} \frac{|\widehat{\psi}(\gamma)|^2}{\gamma} d\gamma = C_\psi < \infty,$$

then we can recover f from the following relation:

$$f(t) = \frac{1}{C_\psi}\int_{\mathbb{R}+}\int_{\mathbb{R}} W_f(s,a)\psi_{s,a}(t)\frac{1}{s}dads.$$

The aforementioned scaling problem for graphs makes it clear that we cannot use an analogous approach to wavelet transforms in the spatial vertex domain. Instead, we show how the wavelet coefficients can be recovered from scaling, and translating in the Fourier domain. We shall then define a graph wavelet transform based on this process.

Consider the case where the scale parameter s is discretized, and the translation parameter a is continuous. For a fixed scale s, the operator T^s is defined to be $T^s f(a) = W_f(s,a)$. For the wavelet ψ, define

$$\tilde{\psi}_s(t) = \frac{1}{s}\overline{\psi}\left(\frac{-t}{s}\right).$$

T^s can be represented as a convolution:

$$T^s f(a) = \int_{\mathbb{R}} \frac{1}{s}\overline{\psi}\left(\frac{t-a}{s}\right) f(t)dt = \int_{\mathbb{R}} \tilde{\psi}_s(a-t) f(t)dt = (\tilde{\psi}_s * f)(a).$$

Further, using the multiplicative property of the Fourier transform of convolutions, the fact that

$$\begin{aligned}
\widehat{\tilde{\psi}}_s(\gamma) &= \textstyle\int_{\mathbb{R}} \frac{1}{s}\overline{\psi}\left(\frac{-t}{s}\right) e^{-2\pi i\gamma t}dt \\
&= \textstyle\int_{\mathbb{R}} \overline{\psi}(u)e^{2\pi i s\gamma u}du \\
&= \overline{\textstyle\int_{\mathbb{R}} \psi(u)e^{-2\pi i s\gamma u}du} \\
&= \overline{\widehat{\psi}(s\gamma)},
\end{aligned}$$

and the Fourier inversion formula we have:

$$(T^s f)(x) = (\widehat{(T^s f)}(\gamma))^\vee = \int_{\mathbb{R}} \widehat{\tilde{\psi}}_s(\gamma)\widehat{f}(\gamma)e^{2\pi i\gamma t}d\gamma = \int_{\mathbb{R}} \overline{\left(\widehat{\psi}(s\gamma)\right)}\widehat{f}(\gamma)e^{2\pi i\gamma t}d\gamma.$$

Hence, we can define the spatial translation operator via a scaling in the Fourier domain followed by inversion. We now use this property to define the analogous "translation" operator in the graph setting by scaling in the graph Fourier domain and inverting using the invertibility of the graph Fourier transform.

5.3 Spectral Graph Wavelet Transform

Let $g : \mathbb{R}^+ \to \mathbb{R}^+$ be defined such that $g(0) = 0$ and $\lim_{t\to\infty} g(t) = 0$. We define the wavelet operator T_g on a given function f as follows:

$$\begin{aligned} T_g(f) &= \left(\begin{bmatrix} g(\lambda_0) & & \\ & \ddots & \\ & & g(\lambda_{N-1}) \end{bmatrix} \widehat{f} \right)^{\vee} \\ &= \chi \begin{bmatrix} g(\lambda_0) & & \\ & \ddots & \\ & & g(\lambda_{N-1}) \end{bmatrix} \chi^* f. \end{aligned}$$

In this setting, our translation in the Fourier domain is accomplished via the function g acting on the eigenvalues of the graph Laplacian which we denote

$$g(\mathcal{L}) = \begin{bmatrix} g(\lambda_0) & & \\ & \ddots & \\ & & g(\lambda_{N-1}) \end{bmatrix}.$$

The wavelet operators, at scale t, are defined as $T_g^t = g(t\mathcal{L})$ for $t > 0$. We define the spectral graph wavelets $\{\psi_{t,n}\}_{n=0}^{N-1}$ by applying T_g^t to the canonical basis elements

$\{e_n\}_{n=0}^{N-1}$:

$$\begin{aligned}
\psi_{t,n} &= \chi g(t\mathcal{L})\widehat{e}_n \\
&= \chi \begin{bmatrix} g(t\lambda_0) & & \\ & \ddots & \\ & & g(t\lambda_{N-1}) \end{bmatrix} \chi^* e_n \\
&= \chi \begin{bmatrix} g(t\lambda_0) & & \\ & \ddots & \\ & & g(t\lambda_{N-1}) \end{bmatrix} \begin{bmatrix} \bar{\chi}_0[n] \\ \vdots \\ \bar{\chi}_{N-1}[n] \end{bmatrix} \\
&= \chi \begin{bmatrix} g(t\lambda_0)\bar{\chi}_0[n] \\ \vdots \\ g(t\lambda_{N-1})\bar{\chi}_{N-1}[n] \end{bmatrix} \\
&= \begin{bmatrix} \sum_{l=0}^{N-1} g(t\lambda_l)\chi_l[0]\bar{\chi}_l[n] \\ \vdots \\ \sum_{l=0}^{N-1} g(t\lambda_l)\chi_l[N-1]\bar{\chi}_l[n] \end{bmatrix}.
\end{aligned}$$

This yields the summation formula $\psi_{t,n}[m] = \sum_{l=0}^{N-1} g(t\lambda_l)\chi_l[m]\bar{\chi}_l[n]$. The graph wavelet coefficients $W_f(t,n)$ are then found by taking the inner product with the function f:

$$\begin{aligned}
W_f(t,n) &= \langle \psi_{t,n}, f \rangle \\
&= \left(\chi \begin{bmatrix} g(t\lambda_0) & & \\ & \ddots & \\ & & g(t\lambda_{N-1}) \end{bmatrix} \chi^* e_n \right)^* f \\
&= \sum_{k=0}^{N-1} f[k] \sum_{l=0}^{N-1} g(t\lambda_l) \bar{\chi}_l[k] \chi_l[n] \\
&= \sum_{l=0}^{N-1} g(t\lambda_l) \chi_l[n] \sum_{k=0}^{N-1} f[k] \bar{\chi}_l[k] \\
&= \sum_{l=0}^{N-1} g(t\lambda_l) \chi_l[n] \widehat{f}[l].
\end{aligned}$$

Having established the notion of wavelets on graphs, we may now apply the methods to applications such as signal processing. However, this presents a computationally cumbersome challenge as the graph Laplacian, while sparse in many cases, scales with N^2. Nonetheless, we introduce an invertibility theorem from [25].

Theorem 5.3.1 *(Hammond, Vandergheynst, Gribonal [25]) If the spectral graph wavelet transform kernel g satisfies the admissibility condition*

$$\int_{\mathbb{R}^+} \frac{g^2(t)}{t} dt = C_g < \infty,$$

and if $g(0) = 0$, then

$$f[m] - \langle \chi_0, f \rangle \frac{1}{\sqrt{N}} = \frac{1}{C_g} \sum_{n=0}^{N-1} \int_{\mathbb{R}^+} W_f(t,n) \psi_{t,n}[m] \frac{dt}{t}.$$

.

By construction, the wavelets $\psi_{t,n}$ are orthogonal to the first eigenvector χ_0 (the $n \times 1$ constant column vector with value $1/\sqrt{N}$). Thus, for the complete reconstruction, the $\langle \chi_0, f \rangle \, 1/\sqrt{N}$ term must be somewhat artificially added back into the formula. Another note concerning the reconstruction formula is that it requires a continuous integral, despite the fact that the original space was entirely discrete. This leads to numerical concerns that must be addressed (see [25]). Theorem (5.3.1) demonstrates the motivation for our main results: it illuminates some of the important harmonic analysis properties of the classical DFT that still exist in the graph setting.

Chapter 6: Graph Theoretic Uncertainty Principles

6.1 Introduction

We extend the notion of discrete uncertainty principles such as those introduced in [24], and [19]. We show that for the graph setting, the cyclic structure of the discrete Fourier transform is no longer present for the graph Fourier transform. As a result, the support theorems (such as in [19] and [49]) are no longer guaranteed. Finally, we extend the frame uncertainty principle introduced by Lammers and Maeser in [33] to the graph Fourier transform and to the normalized graph Fourier transform.

6.2 A Graph Differential Uncertainty Principle

Recall from Corollary 2.2.4, the additive Heisenberg uncertainty principle:

$$\|f(t)\|^2 \leq 2\pi \left(\|tf(t)\|^2 + \|\gamma f(\gamma)\|^2\right). \tag{6.1}$$

For a function $f \in \mathcal{S}(\mathbb{R})$, the space of Schwartz functions on $\mathbb{R}$, Corallary 2.2.5 states that inequality (6.1) is equivalent to:

$$\|f(t)\|^2 \leq \left(\left\|\widehat{f}'(\gamma)\right\|^2 + \|f'(t)\|^2\right). \tag{6.2}$$

To achieve a graph analog of inequality (6.2), we must define the notion of a difference operator in the graph setting. To do this, we examine the following product: $g = W^{1/2}Mf = D_r f$, where $D_r = W^{1/2}M$. The function g is a function on the edges of the graph, where each value is the difference of the function f at the endpoints of the edge. Because of this property, it is common to define the function g as the derivative of f (see [1]). With this in mind, we establish a differential graph Fourier transform inequality of the form of (6.2).

Theorem 6.2.1 *Let G be a simple, connected, and undirected graph. Then, for any non-zero function $f \in l^2(V)$, the following inequalities hold:*

$$0 < \|f\|^2 \tilde{\lambda}_0 \leq \|D_r f\|^2 + \left\|D_r \widehat{f}\right\|^2 \leq \|f\|^2 \tilde{\lambda}_{N-1}, \tag{6.3}$$

where $0 < \tilde{\lambda}_0 \leq \tilde{\lambda}_1 \leq ... \leq \tilde{\lambda}_{N-1}$ are the ordered real eigenvalues of $L + \Delta$. Furthermore, the bounds are sharp.

Proof: Noting that

$$\begin{aligned}\|D_r f\|^2 &= \langle D_r f, D_r f\rangle \\ &= \langle f, \chi\Delta\chi^* f\rangle \\ &= \left\langle \widehat{f}, \Delta\widehat{f}\right\rangle\end{aligned}$$

and, similarly, that $\left\|D_r \widehat{f}\right\|^2 = \left\langle \widehat{f}, L\widehat{f}\right\rangle$, we have

$$\|D_r f\|^2 + \left\|D_r \widehat{f}\right\|^2 = \left\langle \widehat{f}, (L + \Delta)\widehat{f}\right\rangle.$$

Assuming $\tilde{\lambda}_0 > 0$, Inequality (3) follows directly from $L + \Delta$ being symmetric and positive semidefinite. Indeed, we have

$$0 < \|f\|^2 \tilde{\lambda}_0 \leq \left\langle \widehat{f}, (L + \Delta)\widehat{f}\right\rangle = \|D_r f\|^2 + \left\|D_r \widehat{f}\right\|^2 \leq \|f\|^2 \tilde{\lambda}_{N-1}$$

following directly from the properties of the Rayleigh quotient. To prove positivity of $\tilde{\lambda}_0$, note that for $\left\langle \widehat{f}, (L+\Delta)\widehat{f} \right\rangle = 0$ we must have $\langle h, \Delta h \rangle = 0 = \langle h, Lh \rangle$ for some $h \neq 0$. This is impossible as we have, for non-zero h, $\langle h, \Delta h \rangle = 0$ if and only if $h = c[1, 0, ..., 0]^*$ for some $c \neq 0$. This implies $\langle h, Lh \rangle = deg(v_0)c^2 > 0$ due to the connectivity of the graph. ■

A direct consequence of Theorem 6.2.1 is that for a constant function $f = c\chi_0$ ($c \neq 0$) we have $\|D_r c\widehat{\chi}_0\| > 0$. Hence, zero derivative in the graph domain implies a non-constant function in the graph Fourier domain.

Alternatively, if we consider the normalized Laplacian $\mathcal{L}$ we define a slightly different notion of the derivative in order to reflect the slightly different structure when using the normalized Laplacian. For a function $f \in l^2(G)$, define the normalized graph derivative as

$$D_{nr} = D^{1/2} D_r = D^{1/2} W^{1/2} M.$$

Let $\mathcal{D}$ be the diagonalization of $\mathcal{L}$. We establish a graph differential normalized Fourier transform inequality of the form of Theorem 6.2.1.

Theorem 6.2.2 *Let G be a simple, connected, and undirected graph. Then, for any non-zero function $f \in l^2(V)$, the following inequalities hold:*

$$0 < \|f\|^2 \tilde{\mu}_0 \leq \|D_{nr} f\|^2 + \left\| D_{nr} \overset{*}{f} \right\|^2 \leq \|f\|^2 \tilde{\mu}_{N-1}, \tag{6.4}$$

where $0 < \tilde{\mu}_0 \leq \tilde{\mu}_1 \leq ... \leq \tilde{\mu}_{N-1}$ are the ordered real eigenvalues of $\mathcal{L} + \mathcal{D}$. Furthermore, the bounds are sharp.

Proof: Noting that

$$\begin{aligned}\|D_{nr}f\|^2 &= \langle D_{nr}f, D_{nr}f\rangle \\ &= \langle f, \mathcal{F}\mathcal{D}\mathcal{F}^* f\rangle \\ &= \left\langle \overset{*}{f}, \mathcal{D}\overset{*}{f}\right\rangle\end{aligned}$$

and, similarly, that $\left\|D_{nr}\overset{*}{f}\right\|^2 = \left\langle \overset{*}{f}, \mathcal{L}\overset{*}{f}\right\rangle$, we have

$$\|D_{nr}f\|^2 + \left\|D_{nr}\overset{*}{f}\right\|^2 = \left\langle \overset{*}{f}, (\mathcal{L}+\mathcal{D})\overset{*}{f}\right\rangle.$$

Assuming $\tilde{\mu}_0 > 0$, Inequality (3) follows directly from $\mathcal{L}+\mathcal{D}$ being symmetric and positive semidefinite. Indeed, we have

$$0 < \|f\|^2 \tilde{\mu}_0 \leq \left\langle \overset{*}{f}, (\mathcal{L}+\mathcal{D})\overset{*}{f}\right\rangle = \|D_{nr}f\|^2 + \left\|D_{nr}\overset{*}{f}\right\|^2 \leq \|f\|^2 \tilde{\mu}_{N-1}$$

following directly from the properties of the Rayleigh quotient. To prove positivity of $\tilde{\mu}_0$, note that for $\left\langle \overset{*}{f}, (\mathcal{L}+\mathcal{D})\overset{*}{f}\right\rangle = 0$ we must have $\langle h, \mathcal{D}h\rangle = 0 = \langle h, \mathcal{L}h\rangle$ for some $h \neq 0$. This is impossible as we have, for non-zero h, $\langle h, \mathcal{D}h\rangle = 0$ if and only if $h = c[1, 0, ..., 0]^*$ for some $c \neq 0$. This implies $\langle h, \mathcal{L}h\rangle = \mathcal{L}_{00}c^2 = c^2 > 0$ due to the connectivity of the graph. ■

Theorems 6.2.1 and 6.2.2 establish a positive lower bound for the norms of differential operators acting on functions in $l^2(G)$ and their graph Fourier transform. The results do not address exactly what values are simultaneously possible in general. We examine the space of all feasible values in Chapter 7. The remainder of this chapter is dedicated to finding lower bounds for the differential operators acting on frames for $l^2(G)$, and calculating specific bound values for the class of complete graphs.

6.3 A Graph Frame Differential Uncertainty Principle

As a generalization of the work by Lammers and Maeser in [33], we show that the modified Laplacian operator $L+\Delta$ will dictate an additive uncertainty principle for frames. Let

$$E = \begin{bmatrix} \\ E_0 & E_1 & \dots & E_{N-1} \\ \\ \end{bmatrix}$$

be a $d \times N$ matrix whose columns form a Parseval frame for $\mathbb{C}^d$, i.e. $EE^* = I_{d\times d}$. If we let $\mathcal{S} = T^0 - T$, then $\mathcal{S}^* = T^0 - T^{N-1}$, and the classical Laplacian in the discrete setting is given by $L_c = \mathcal{S}^*\mathcal{S} = 2T^0 - T - T^{N-1}$. Let $\|\cdot\|_{fr}$ denote the Frobenius norm. Recall from chapter 3 that the following result holds.

Theorem 6.3.1 *(Lammers and Maeser [33]) For fixed dimension d and $N \geq d \geq 2$, the following inequalities hold for all $d \times N$ Parseval frames:*

$$\begin{aligned} 0 < G(N,d) &\leq \|\mathcal{S}\mathcal{F}\mathcal{T}E^*\|_{fr}^2 + \|\mathcal{S}E^*\|_{fr}^2 \\ &\leq H(N,d) \\ &\leq 8d. \end{aligned} \tag{6.5}$$

Furthermore, the minimum (maximum) occurs when columns of E^* the d orthonormal eigenfunctions corresponding to the d smallest (largest) eigenvalues of $L_c + \Delta_c$ where L_c is the classical Laplacian and Δ_c is its diagonalization. The constant $G(N,d)$ is the sum of those d smallest eigenvalues, and $H(N,d)$ is the sum of those d largest eigenvalues.

To extend the inequalities in Theorem 6.3.1 to the graph Fourier transform setting, we apply D_r to the frame's conjugate transpose E^* and to the graph Fourier transform $\chi^* E^*$, and then find bounds for the Frobenius norms.

Theorem 6.3.2 *For any graph G as in Theorem 6.2.1, the following inequalities hold for all $d \times N$ Parseval frames E:*

$$\sum_{j=0}^{d-1} \tilde{\lambda}_j \leq \|D_r \chi^* E^*\|_{fr}^2 + \|D_r E^*\|_{fr}^2 \leq \sum_{j=N-d}^{N-1} \tilde{\lambda}_j, \tag{6.6}$$

where $\left\{\tilde{\lambda}_j\right\}$ is the ordered set of real, non-negative eigenvalues of $L + \Delta$. Furthermore, these bounds are sharp.

Proof: Writing out the Frobenius norms as trace operators yield:

$$\begin{aligned} \|D_r \chi^* E^*\|_{fr}^2 + \|D_r E^*\|_{fr}^2 = tr(E \chi D_r^* D_r \chi^* E^*) \\ + tr(D_r E^* E D_r^*). \end{aligned} \tag{6.7}$$

Using the invariance of the trace when reordering products, we have $\|D_r \chi^* E^*\|_{fr}^2 +$ $\|D_r E^*\|_{fr}^2$

$$\begin{aligned} &= tr(L \chi^* E^* E \chi) + tr(L E^* E) \\ &= tr(L \chi^* E^* E \chi) + tr(\chi \Delta \chi^* E^* E) \\ &= tr((L + \Delta) \chi^* E^* E \chi). \end{aligned}$$

The operator $\Delta + L$ is real, symmetric, and positive semidefinite. By the spectral theorem, it has an orthonormal eigenbasis P that, upon conjugation, diagonalizes $\Delta + L$:

$$P^*(\Delta + L)P = \tilde{\Delta} = \text{diag}(\tilde{\lambda}_0, \tilde{\lambda}_1, ..., \tilde{\lambda}_{N-1}).$$

Hence, we have

$$\begin{aligned}\|D_r \chi E^*\|_{fr}^2 + \|D_r E^*\|_{fr}^2 &= tr((\Delta + L)\chi^* E^* E \chi) \\ &= tr(P\tilde{\Delta}P^* \chi^* E^* E \chi) \\ &= tr(\tilde{\Delta}P^* \chi^* E^* E \chi P) \\ &= \sum_{j=0}^{N-1} (K^* K_{j,j})\, \tilde{\lambda}_j,\end{aligned}$$

where $K = E\chi P$. The matrix K is a Parseval frame because unitary transformations of Parseval frames are Parseval frames . Therefore, $tr(K^*K) = tr(KK^*) = d$. K^*K is also the product of matrices with operator norm ≤ 1. Therefore, each of the entries, $(K^*K_{j,j})$, satisfies $0 \leq (K^*K_{j,j}) \leq 1$. Hence, minimizing (maximizing) $\sum_{j=0}^{N-1} (K^*K_{jj})\, \tilde{\lambda}_j$ is achieved if

$$(K^*K_{j,j}) = \begin{cases} 1 & j < d\ (j \geq N-d) \\ 0 & j \geq d\ (j < N-d). \end{cases}$$

Choosing E to be the first (last) d rows of $(\chi P)^*$ accomplishes this. The positivity of the bounds follows from the proof of Theorem 6.2.1 ■

A similar result holds for the normalized graph Laplacian.

Theorem 6.3.3 *For any graph G as in Theorem 6.2.1, the following inequalities hold for all $d \times N$ Parseval frames E:*

$$\sum_{j=0}^{d-1} \tilde{\mu}_j \leq \|D_{nr}\mathcal{F}^* E^*\|_{fr}^2 + \|D_{nr} E^*\|_{fr}^2 \leq \sum_{j=N-d}^{N-1} \tilde{\mu}_j, \tag{6.8}$$

where $\{\tilde{\mu_j}\}$ is the ordered set of real, non-negative eigenvalues of $\mathcal{L}+\mathcal{D}$. Furthermore, these bounds are sharp.

Proof: Writing out the Frobenius norms as trace operators yield:

$$\|D_{nr}\mathcal{F}^*E^*\|_{fr}^2 + \|D_{nr}E^*\|_{fr}^2 = tr(E\mathcal{F}D_{nr}^*D_{nr}\mathcal{F}^*E^*) + tr(D_{nr}E^*ED_{nr}^*). \tag{6.9}$$

Using the invariance of the trace when reordering products, we have $\|D_{nr}\mathcal{F}^*E^*\|_{fr}^2 + \|D_{nr}E^*\|_{fr}^2$

$$\begin{aligned} &= tr(\mathcal{L}\mathcal{F}^*E^*E\mathcal{F}) + tr(\mathcal{L}E^*E) \\ &= tr(\mathcal{L}\mathcal{F}^*E^*E\mathcal{F}) + tr(\mathcal{F}\mathcal{D}\mathcal{F}^*E^*E) \\ &= tr((\mathcal{L}+\mathcal{D})\mathcal{F}^*E^*E\mathcal{F}). \end{aligned}$$

The operator $\mathcal{D}+\mathcal{L}$ is real, symmetric, and positive semidefinite. By the spectral theorem, it has an orthonormal eigenbasis P_n that, upon conjugation, diagonalizes $\mathcal{D}+\mathcal{L}$:

$$P_n^*(\mathcal{D}+\mathcal{L})P_n = \tilde{\mathcal{D}} = \text{diag}(\tilde{\mu}_0, \tilde{\mu}_1, ..., \tilde{\mu}_{N-1}).$$

Hence, we have

$$\begin{aligned} \|D_{nr}\mathcal{F}E^*\|_{fr}^2 + \|D_{nr}E^*\|_{fr}^2 &= tr((\mathcal{D}+\mathcal{L})\mathcal{F}^*E^*E\mathcal{F}) \\ &= tr(P_n\tilde{\mathcal{D}}P_n^*\mathcal{F}^*E^*E\mathcal{F}) \\ &= tr(\tilde{\mathcal{D}}P_n^*\mathcal{F}^*E^*E\mathcal{F}P_n) \\ &= \sum_{j=0}^{N-1}\left((K_n^*K_n)_{j,j}\right)\tilde{\mu}_j, \end{aligned}$$

where $K_n = E\mathcal{F}P_n$. The matrix K_n is a Parseval frame because unitary transformations of Parseval frames are Parseval frames . Therefore, $tr(K_n^*K_n) = tr(K_nK^*) = d$. $K_n^*K_n$ is also the product of matrices with operator norm ≤ 1. Therefore, each of the entries, $\left((K_n^* K_n)_{j,j}\right)$, satisfies $0 \leq \left((K_n^* K_n)_{j,j}\right) \leq 1$. Hence, minimizing (maximizing) $\sum_{j=0}^{N-1} \left((K_n^* K_n)_{j,j}\right) \tilde{\mu}_j$ is achieved if

$$\left((K_n^* K_n)_{j,j}\right) = \begin{cases} 1 & j < d \ (j \geq N-d) \\ \\ 0 & j \geq d \ (j < N-d). \end{cases}$$

Choosing E to be the first (last) d rows of $(\mathcal{F}P_n)^*$ accomplishes this. The positivity of the bounds follows from the proof of Theorem 6.2.1 ■

6.4 The Complete Graph

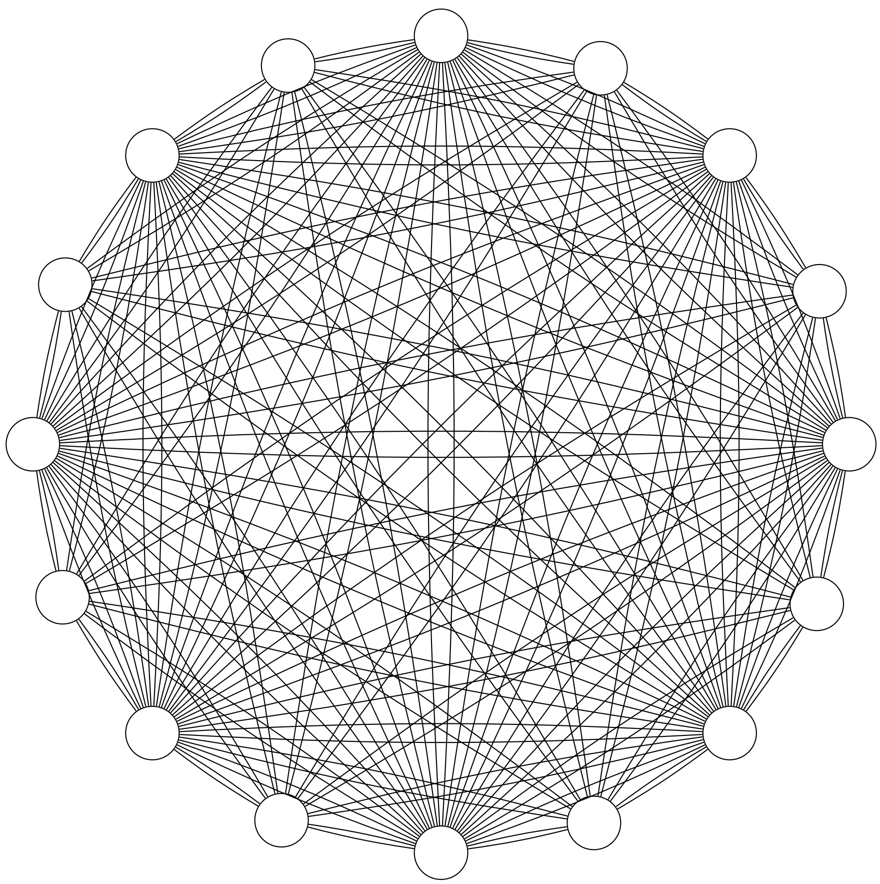

Figure 6.1: A unit weighted complete graph with 16 vertices.

Unit weighted graphs for which every vertex is connected directly to every other vertex, as in Figure 6.1, are referred to as *complete graphs*. A complete graph with N vertices has graph Laplacian

$$L = \begin{bmatrix} N-1 & -1 & -1 & & \cdots & & -1 & -1 \\ -1 & N-1 & -1 & & & & & -1 \\ -1 & \ddots & \ddots & \ddots & & & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & & & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & & & \ddots & \ddots & \ddots & 0 \\ -1 & & & & & -1 & N-1 & -1 \\ -1 & -1 & \cdots & \cdots & \cdots & -1 & -1 & N-1 \end{bmatrix} = NI - O$$

where O is an $N \times N$ matrix each of whose elements is 1. Noting that

$$L^2 = (NI - O)^2 = N^2 I - 2NO + O^2 = N^2 I - 2NO + NO = N^2 I - NO,$$

and that

$$LNI = N^2 I - NO,$$

we have

$$L(L - NI) = L^2 - NL = N^2 I - NO - N^2 I + NO = 0.$$

Hence, the minimal polynomial $m(x)$ for L is given by $m(x) = x(x - N)$. Since the zero eigenvalue has multiplicity one, the characteristic polynomial is $c(x) = x(x - N)^{N-1}$. As is the case with all connected graphs, the eigenspace associated with the null eigenvalue is the span of the constant vector $\chi_0 = \left(1/\sqrt{N}\right)[1, ..., 1]^*$. Let $\chi_1 = \left(1/\sqrt{2}\right)[1, -1, 0, ..., 0]$. Then $\langle \chi_0, \chi_1 \rangle = 0$ and $L\chi_1 = N\chi_1$. Upon solving for the $N-2$ remaining orthonormal eigenfunctions χ_l for $l = 2, ..., N-1$, we define the complete graph Fourier transform matrix $\chi_c^* = [\chi_0, \chi_1, \chi_2, ..., \chi_{N-1}]^*$. We then have $\widehat{\chi_1} = [0, 1, 0, ..., 0]^*$, and

$$|supp(\chi_1)|\,|supp(\widehat{\chi_1})| = 2 < N$$

for $N \geq 3$; and we see that the support theorems in [19] do not hold for graphs. The cyclic structure of the $\mathcal{FT}$ matrix is not necessarily present in the graph setting. Namely, the $\mathcal{FT}$ matrix is a Vandermonde matrix, while the graph Fourier transform matrix is merely unitary.

For $N > 2$, the eigenvalues associated with $L + \Delta$ are given by $\tilde{\lambda}_0 = N - \sqrt{N}$, $\tilde{\lambda}_1 = N + \sqrt{N}$, and $\tilde{\lambda}_j = 2N$ for $d \leq j \leq N - 1$. Let

$$v_0 = \begin{bmatrix} \sqrt{N} + 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \sqrt{N} e_0 + C$$

where e_0 is the canonical first basis vector, and C is the N vector whose elements are all 1. Using the fact that $LC = 0$, and that $\Delta e_0 = 0$, it is straightforward to show

$$(L+\Delta)v_0 = \sqrt{N}\begin{bmatrix} N-1 \\ -1 \\ \vdots \\ -1 \end{bmatrix} + n\begin{bmatrix} 0 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} \sqrt{N}(N-1) \\ N-\sqrt{N} \\ \vdots \\ N-\sqrt{N} \end{bmatrix}$$

$$= (N-\sqrt{N})\begin{bmatrix} \sqrt{N}+1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = (N-\sqrt{N})v_0.$$

Hence, v_0 is associated with the eigenvalue $N-\sqrt{N}$. If we let

$$v_1 = \begin{bmatrix} 1-\sqrt{N} \\ 1 \\ \vdots \\ 1 \end{bmatrix} = C - \sqrt{N}e_0,$$

then

$$(L+\Delta)v_1 = (N+\sqrt{N})v_1,$$

and we associate v_1 with the eigenvalue $N+\sqrt{N}$. Suppose some v_j is orthogonal to both v_0 and v_1, that is $\langle v, v_0 \rangle = 0 = \langle v, v_0 \rangle$. These conditions imply

$$\sum_{k=0}^{N-1} v[k] = -\sqrt{N} v[0]$$

and

$$\sum_{k=0}^{N-1} v[k] = \sqrt{N} v[0]$$

which is only satisfied by

$$\sum_{k=0}^{N-1} v[k] = 0 = v[0]. \tag{6.10}$$

Applying the modified Laplacian $(L + \Delta)$ to v yields

$$\begin{aligned} (L+\Delta)v &= (L+\Delta) \begin{bmatrix} 0 \\ v[1] \\ \vdots \\ v[N-1] \end{bmatrix} \\ &= \begin{bmatrix} -\sum_{k=1}^{N-1} v[k] \\ Nv[1] - \sum_{k=1}^{N-1} v[k] \\ \vdots \\ Nv[N-1] - \sum_{k=1}^{N-1} v[k] \end{bmatrix} + Nv \\ &= 2Nv, \end{aligned}$$

where the final equality is due to (6.10). Since the orthogonal compliment of the closed linear span of v_0 and v_1 has dimension $N-2$, we conclude that $2N$ is the largest eigenvalue of $L + \Delta$ with multiplicity $N - 2$. Due to these calculations and the results of Theorem 6.2.1, we have, for $N > 2$, that

$$\|f\|^2 (N - \sqrt{N}) \leq \|D_r f\|^2 + \left\|D_r \widehat{f}\right\|^2 \leq \|f\|^2 2N.$$

If we note that

$$\sum_{k=0}^{d-1} \tilde{\lambda}_k = (N-\sqrt{N})+(N+\sqrt{N})+\sum_{k=2}^{d-1} 2N = 2N+\sqrt{N}-\sqrt{N}+2N(d-2) = 2N(d-1),$$

and apply Theorem 6.3.2, then we have, for all $d \times N$ Parseval frames E, that

$$2N(d-1) \leq \|D_r \chi^* E^*\|_{fr}^2 + \|D_r E^*\|_{fr}^2 \leq \begin{cases} 2Nd & \text{if } N > d+1 \\ 2Nd - N & \text{N=d+1} \end{cases}.$$

Chapter 7: Feasibility Results

7.1 Introduction

In [1], the authors define the notion of spread in the spectral and graph domains using the analytic properties of the graph Fourier transform. The eigenvalues and eigenfunctions of the graph Laplacian play a central role in determining what values of spread are feasible. We shall examine this result in great detail and note its similarity to the results in [47].

7.2 A Spectral Graph Uncertainty Principle

In this section, we shall give a spectral graph analogy to the classical uncertainty principle due to Heisenberg, Pauli, Weyl, and Wiener due to Agaskar, and Lu in [1]. Further, we shall explore the intriguing apparent connection between diffusion processes on graphs and the uncertainty bounds.

Recall that the normalized Laplacian $\mathcal{L} = I - D^{-1/2}AD^{-1/2}$ has non-negative eigenvalues $0 = \mu_0 < \mu_1 \leq \mu_2 \leq ... \leq \mu_{N-1}$ with the associated orthonormal eigenbasis $\{F_j\}$. Recalling that for a connected graph that the eigenspace of μ_0 has dimension 1, we have (up to a choice of sign) that

$$F_0 = \left[\sqrt{\frac{deg(v_0)}{\sum_{l=0}^{N-1} deg(v_l)}}, \sqrt{\frac{deg(v_1)}{\sum_{l=0}^{N-1} deg(v_l)}}, ..., \sqrt{\frac{deg(v_{N-1})}{\sum_{l=0}^{N-1} deg(v_l)}}\right]'.$$

Indeed, for

$$\mathcal{L}F_0 = F_0 - D^{-1/2}AD^{-1/2}F_0$$

we have that $D^{-1/2}F_0 = 1/\sqrt{\sum_{j=0}^{N-1} deg(v_j)}[1,1,...,1]'$ and that

$$D^{-1/2}A[1,1,...,1]' = \begin{bmatrix} \sum_{j=0}^{N-1} deg(v_0)^{-1/2}a_0[j] \\ \sum_{j=0}^{N-1} deg(v_1)^{-1/2}a_1[j] \\ \vdots \\ \sum_{j=0}^{N-1} deg(v_{N-1})^{-1/2}a_{N-1}[j] \end{bmatrix} = \begin{bmatrix} deg(v_0)^{1/2} \\ deg(v_1)^{1/2} \\ \vdots \\ deg(v_{N-1})^{1/2}, \end{bmatrix}.$$

Hence, $\mathcal{L}F_0 = F_0 - F_0 = 0F_0$ so F_0 is (up to sign change) a unit norm eigenfunction for $\mu_0 = 0$.

In the classical setting, we define the time spread of a nonzero function $f \in L^2(\mathbb{R})$ about a point $t_0 \in \mathbb{R}$ as

$$\Delta^2_{t,t_0} := \frac{1}{\|f\|^2}\int_{\mathbb{R}} (t-t_0)^2 \, |f(t)|^2 \, dt.$$

To generalize the notion of time spread to the graph setting we introduce the following definition.

Let $u_0 \in V$ for a graph G. Define the distance to u_0, $d(u_0, v)$ as the shortest length of a path connecting v and u_0. Let $P_{u_0} := diag(d(u_0,v_0), d(u_0,v_1), ..., d(u_0,v_{N-1}))$.

For a non zero function $f : G \to \mathbb{R}$, the *graph spread* of f about u_0 is

$$\begin{aligned}\Delta^2_{g,u_0}(f) &= \frac{1}{\|f\|^2}\sum_{j=0}^{N-1} d(u_0, v_j)^2 f(v_j)^2 \\ &= \frac{1}{\|f\|^2}\left\langle f, P^2 f\right\rangle \end{aligned} \tag{7.1}$$

The analogy between the time spread and the graph spread is straight forward: they normalize the function and then measure the distance from a central node multiplied by the value of the function. In order to define the spectral spread for a graph, a less direct approach is necessary. In the classical setting the spectral or frequency spread is given by

$$\Delta^2_{\gamma}(f) := \frac{1}{\|f\|^2}\int \gamma^2 \left|\hat{f}(\gamma)\right|^2 d\gamma. \tag{7.2}$$

We have chosen $\gamma_0 = 0$ due to the symmetry of the Fourier transform of real valued functions. We have by integration by parts

$$\left(-\frac{d^2}{dt^2}f(t)\right)\hat{}(\gamma) = 4\pi^2\gamma^2\hat{f}(\gamma).$$

Hence, we can replace (7.2) with

$$\Delta^2_{\gamma}(f) = \frac{C}{\|f(t)\|^2}\int f(t)\frac{-d^2}{dt^2}f(t)dt$$

where $C = 4\pi^2$. It is with this formulation that we now form a graph theoretic analog.

For a nonzero function $f \in l^2(G)$, we define its *spectral spread* of f as

$$\Delta^2_s(f) = \frac{1}{\|f\|^2}\langle f, \mathcal{L}f\rangle = \frac{1}{\|f\|^2}\sum_{j=0}^{N-1}\lambda_j\left|\hat{f}[j]\right|^2$$

where the second equality is due to the fact that $\mathcal{L} = \mathcal{F}\Lambda\mathcal{F}^*$ and the definition of the normalized graph Fourier transform.

In the classical setting, not all pairs of $(\Delta_t^2, \Delta_\gamma^2)$ are achievable: we must have

$$\forall f \in L^2(\mathbb{R})\ \Delta_t^2(f)\Delta_\gamma^2(f) \geq \frac{1}{4}.$$

We shall show that in the graph setting the allowable pairs of graph and spectral spread are confined to a bounded, convex region in the first quadrant of $\mathbb{R}^2$. Define the *feasibility region* D_{u_0} as follows:

$$D_{u_0} = \left\{(s,g) : \Delta_s^2(f) = s \text{ and } \Delta_{g,u_0}^2(f) = g \text{ for some } f \neq 0 \in l^2(G)\right\}.$$

We shall prove the some key properties of the feasibility region.

Proposition 7.2.1 *(Agaskar and Lu [1]) Let D_{u_0} be the feasibility region for a connected graph G with N vertices. Then, the following properties hold.*

a) D_{u_0} is a closed subset subset of $[0, \mu_{N-1}] \times [0, \mathcal{E}_G^2(u_0)]$ where $\mathcal{E}_G^2(u_0) := \max_{v \in V} dist(u_0, v)$ is called the eccentricity of u_0.

b) $g = 0$ and $s = 1$ is the only point on the horizontal axis in D_{u_0}. $s = 0$ and $g = \left\langle F_0, P_{u_0}^2 F_0 \right\rangle$ is the only point on the vertical axis in D_{u_0}.

c) The points $(1, \mathcal{E}_G^2(u_0))$, and $(\mu_{N-1}, \left\langle F_{N-1} P_{u_0}^2 F_{N-1} \right\rangle)$ belong to D_{u_0}.

d) If $N \geq 3$ then D_{u_0} is a convex region.

Proof:

a) Let $(s,g) \in D_{u_0}$ for some f. Then for $\tilde{f} = \frac{f}{\|f\|}$, $\Delta_s^2(f) = s = \Delta_s^2(\tilde{f})$ and similarly for $\Delta_{g,u_o}^2(f)$ so D_{u_0} is in the image of the unit sphere. By definition, the image of the unit sphere is in D_{u_0}. Therefore D_{u_0} is a closed compact set as it is the image of a closed compact set under a continuous transformation. WLOG we shall assume f is on the unit sphere. We have that $\mu_0 \leq \langle f, \mathcal{L}f \rangle \leq \mu_{N-1}$, and that $0 \leq \left\langle f, P_{u_0}^2 f \right\rangle \leq \mathcal{E}_G^2(u_0)$ by the Rayleigh inequalities.

b) If $g = 0$ then $\left\langle f, P_{u_0}^2 f \right\rangle = 0 = \sum_{j=0}^{N-1} \mathrm{dist}^2(u_0, v_j) \left| f[j] \right|^2$. For a connected graph, $\mathrm{dist}^2(u_0, v_j) = 0$ if and only if $v_j = u_0$. Such an f must be the canonical j^{th} basis vector or its negative. Regardless of this choice of sign, we must then have $s = \langle f, \mathcal{L}f \rangle = \mathcal{L}_{jj} = 1$. If $s = 0$ for some f, then f is in the eigenspace of $\mu_0 = 0$. Hence, we have $g = \left\langle F_0, P_{u_0}^2 F_0 \right\rangle$.

c) Let $\mathrm{dist}(u_0, v_j) = \mathcal{E}_G(u_0)$, then setting f equal to the canonical j^{th} basis vector yields the coordinates $\left(1, \left\langle f, P_{u_0}^2 f \right\rangle\right) = \left(1, \left(P_{u_0}^2\right)_{jj}\right) = \left(1, \mathcal{E}_G^2(u_0)\right)$. If $g = \left\langle F_{N-1}, P_{u_0}^2 F_{N-1} \right\rangle$, then $s = \langle F_{N-1}, \mathcal{L}F_{N-1} \rangle = \mu_{N-1}$.

d) Showing that the feasibility region is convex is equivalent to showing the following proposition.

Proposition 7.2.2 *(Agaskar and Lu [1]) Let f_1 and f_2 be functions on a graph G with $N \geq 3$ vertices such that*

$$\langle f_i, f_i \rangle = 1, \ \ \langle f_i, \mathcal{L}f_i \rangle = s_i, \ \ \text{and} \ \left\langle f_i, P_{u_0}^2 f_i \right\rangle = g_i \ \text{for} \ i = 1, 2. \tag{7.3}$$

Then for any $\beta \in [0, 1]$, we can always find a function f on the graph satisfying

$$\langle f, f \rangle = 1, \ \ \langle f, \mathcal{L}f \rangle = s, \ \ \text{and} \ \left\langle f, P_{u_0}^2 f \right\rangle = g \tag{7.4}$$

where $s := \beta s_1 + (1 - \beta) s_2$ and $g := \beta g_1 + (1 - \beta) g_2$. That is to say, any line segment in $\mathbb{R}^2$ connecting (s_1, g_1) with (s_2, g_2) is in D_{u_0}.

In order to prove the proposition, we shall formulate this as a problem in Sym_N, the space of $N \times N$ symmetric matrices. Every function f on $l^2(G)$ can be mapped onto a symmetric, rank one $N \times N$ matrix M by setting $M = ff^*$ (i.e. M is the grammian of f). Further, if the following properties hold for a rank one matrix M (which can be decomposed as $M = ff^*$), then 7.4 holds for f:

1. $1 = \langle f, f \rangle = tr(f^* f) = tr(ff^*) = tr(M)$,
2. $s = \langle f, \mathcal{L}f \rangle = tr(f^* \mathcal{L} f) = tr(\mathcal{L} f f^*) = tr(\mathcal{L}M)$, and
3. $g = \left\langle f, P_{u_0}^2 f \right\rangle = tr(f^* P_{u_0}^2 f) = tr(P_{u_0}^2 f f^*) = tr(P_{u_0}^2 M) = g_i$.

In general, finding a rank one matrix satisfying these conditions is not necessarily an easy problem, but the following theorem shall help us find such a matrix.

Theorem 7.2.3 *(Barvinok [4]) Suppose that $R \geq 0$ and $M \geq R + 2$. Let $\mathcal{H} \subset Sym_N$ be an affine subspace such that codim($\mathcal{H}$) $\leq \binom{R+2}{2}$. If $S_+^N \cap \mathcal{H}$ is nonempty and bounded (where S_+^N is the set of $N \times N$ positive semidefinite symmetric matrices), then there exists a matrix $M \in S_+^N \cap \mathcal{H}$ of rank less than or equal to R.*

Proof of proposition: Let f_1 and f_2 satisfy equation 7.3. Under the mapping $M_i = f_i f_i^*$ for $i = 1, 2$, each M_i satisfies

$$tr(M_i) = 1, \; tr(M_i) = s_i, \; \text{and } tr(P_{u_0}^2).$$

By construction, each M_i is symmetric positive definite. For any $\beta \in [0,1]$, let $M' = \beta M_1 + (1-\beta)M_2$. Clearly, $M' \in S_+^N$ by the convexity of S_+^N and if we let $s = \beta s_1 + (1-\beta)s_2$ and $g = \beta g_1 + (1-\beta)g_2$ then

$$M' \in \mathcal{H} = \left\{ M \in \mathrm{Sym}_N : tr(M) = 1, tr(\mathcal{L}M) = s, \text{ and } tr(P_{u_0}^2 M) = g \right\}.$$

By the linear independence of I, $\mathcal{L}$, and $P_{u_0}^2$, we have that $\mathcal{H}$ is an affine subspace of Sym_N with codimension 3. Hence, we have that $S_+^N \cap \mathcal{H} \neq \emptyset$. Noting that any element of $S_+^N \cap \mathcal{H}$ has nonnegative eigenvalues which must sum to 1, the boundedness of this subspace is straightforward to show:

$$\forall M \in S_+^N \cap \mathcal{H}, \;\; \|M\|_{fr} = tr(M^2) \leq tr(M) = 1.$$

By theorem 7.2.3, we conclude that there exists a matrix M of rank one that can be decomposed as $M = ff^*$ with f satisfying equation 7.4 as desired. ■

The boundedness and convexity of the feasibility region suggests that it can be characterized by its lower and upper boundary. We shall explore the lower boundary of the region, however, we could do the same analysis on the upper boundary. We refer to the lower boundary as the *uncertainty curve*, and define it as follows:

$$\begin{aligned}\forall s \in [0, \mu_{N-1}], \; \gamma_{u_0}(s) &= \min_f \Delta_{g,u_0}^2(f) \text{ subject to } \Delta_s^2(f) = s \\ &= \min_f \left\langle f, P_{u_0}^2 f \right\rangle \text{ subject to } \|f\|^2 = 1 \text{ and } \langle f, \mathcal{L}f \rangle = s.\end{aligned}$$

For fixed $s \in [0, \mu_{N-1}]$, we say that a function f' with $\Delta_s^2(f') = s$ attains the uncertainty curve if $\forall f$ with $\Delta_s^2(f) = s$ we have

$$\Delta_{G,u_0}(f') \leq \Delta_{G,u_0}(f).$$

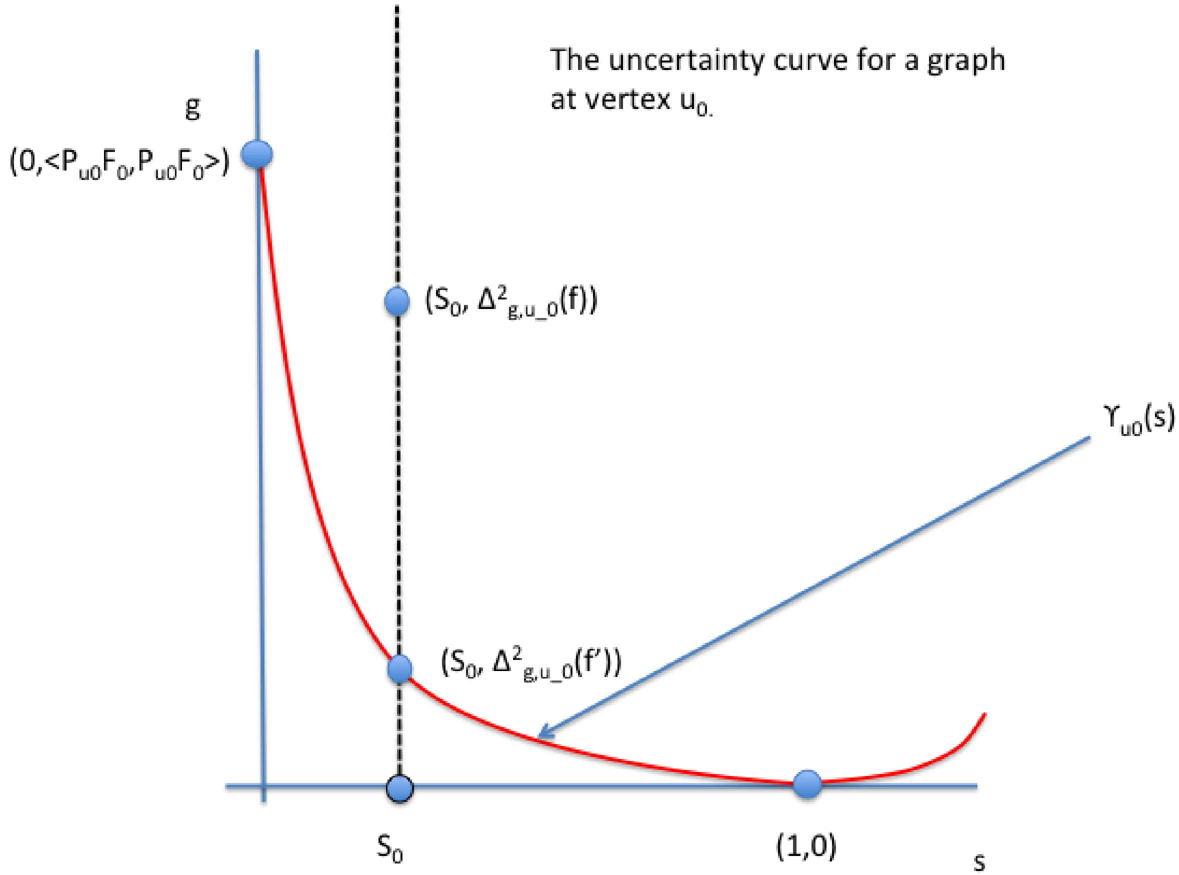

Figure 7.1: The uncertainty curve (red) for a graph G about the vertex u_0

See Figure 7.1 for a graphical representation of the uncertainty curve. Finding a unit norm f' that becomes an N dimensional constrained optimization problem which is solved by Lagrange multipliers [32]. Upon differentiating

$$\Lambda = \left\langle f, P_{u_0}^2 f \right\rangle + \alpha \left(\langle f, \mathcal{L} f \rangle - s \right) + \beta \left(\|f\|^2 - 1 \right)$$

with respect the the N variables in f we have for an optimal solution f' that

$$2P_{u_0}^2 f' = 2\alpha\mathcal{L}f' + 2\beta f'$$

$$\iff \qquad \left(P_{u_0}^2 - \alpha\mathcal{L}\right) f' = \beta f'.$$

If we fix $\alpha \in \mathbb{R}$ and define the linear operator $M(\alpha) = P_{u_0}^2 - \alpha\mathcal{L}$ then we see that any optimal solution must be an eigenfunction for $M(\alpha)$. Define $q(\alpha)$ to be the minimal eigenvalue associated with $M(\alpha)$, and let $S(\alpha)$ be its associated eigenspace. We shall show that any unit norm eigenfunction $g \in S(\alpha)$ lies on $\boldsymbol{\gamma}_{u_0}$.

Proposition 7.2.4 *(Agaskar and Lu [1]) For all $\alpha \in \mathbb{R}$ and for any unit norm eigenfunction $\tau \in S(\alpha)$, we have $\left(\langle \tau, \mathcal{L}\tau \rangle, \left\langle \tau, P_{u_o}^2 \tau \right\rangle\right)$ lies on $\boldsymbol{\gamma}_{u_0}$.*

Proof: Fix $\alpha \in \mathbb{R}$, then for any arbitrary unit norm f we have

$$\begin{aligned} \langle f, M(\alpha) f\rangle &= \left\langle f, (P_{u_0}^2 - \alpha\mathcal{L}) f\right\rangle \\ &= \left\langle f, P_{u_0}^2 f\right\rangle - \alpha \langle f, \mathcal{L}f\rangle . \end{aligned}$$

The Rayleigh quotient for $M(\alpha)$ is bounded sharply below by $q(\alpha)$, hence we conclude for any unit normed $\tau \in S(\alpha)$ that

$$\begin{aligned} \left\langle \tau, P_{u_0}^2 \tau\right\rangle - \alpha \langle \tau, \mathcal{L}\tau\rangle &= q(\alpha) \\ &\leq \left\langle f, P_{u_0}^2 f\right\rangle - \alpha \langle f, \mathcal{L}f\rangle . \end{aligned}$$

Upon restricting f to $\langle f, \mathcal{L}f\rangle = s$ we have

$$\left\langle f, P_{u_0}^2 f\right\rangle - s \geq \left\langle \tau, P_{u_0}^2 \tau\right\rangle - s \quad \iff \quad \left\langle f, P_{u_0}^2 f\right\rangle \geq \left\langle \tau, P_{u_0}^2 \tau\right\rangle .$$

Therefore $\left(\langle \tau, \mathcal{L}\tau \rangle, \left\langle \tau, P_{u_o}^2 \tau \right\rangle\right)$ lies on $\boldsymbol{\gamma}_{u_0}$ as desired. ■

Proposition 7.2.4 guarantees that unit norm eigenfunctions associated with $q(\alpha)$ lie on the uncertainty curve. We shall show the converse is true: for all (s, g) lying on γ_{u_0} there exists an $\alpha \in (-\infty, \infty)$ and a unit normed eigenfunction $\tau \in S(\alpha)$ such that $\left(\langle \tau, \mathcal{L}\tau \rangle, \langle \tau, P_{u_o}^2 \tau \rangle\right) = (s, g)$. To establish this result, we shall rely on the following two functions:

$$h_+(\alpha) := \max_{\tau \in S(\alpha): \|\tau\|=1} \langle \tau, \mathcal{L}\tau \rangle \tag{7.5}$$

$$h_-(\alpha) := \min_{\tau \in S(\alpha): \|\tau\|=1} \langle \tau, \mathcal{L}\tau \rangle$$

which measure the maximal and, respectively, the minimal spectral spread that can be achieved by eigenfunctions in $S(\alpha)$.

Lemma 7.2.5 *(Agaskar and Lu [1]) The following properties hold for $h_+(\alpha)$ and $h_-(\alpha)$.*

a) $h_\pm(\alpha)$ are increasing functions.

b) As α tends to infinity, $h_\pm(\alpha)$ limit to μ_{N-1}, and as α tends to negative infinity $h_\pm(\alpha)$ limit to 0.

c) On any finite interval $[a, b]$, the functions differ on at most a finite number of points denoted by $\mathcal{B} = \{b_1, ..., b_k\}$ for some $k \geq 0$. For all $\alpha \notin \mathcal{B}$, the following holds: $h_+(\alpha) = h_-(\alpha) = -q'(\alpha)$.

Proof:

a) For $\alpha_1 < \alpha_2$, we take any $g_1 \in S(\alpha_1)$ and $g_2 \in S(\alpha_2)$, and we, again, employ the Rayleigh quotient for symmetric matrices: $\langle g_2, M(\alpha_1) g_2 \rangle \geq q(\alpha_1) =$

$\langle g_1, M(\alpha_1) g_1 \rangle$. Similarly, we have $-\langle g_2, M(\alpha_2) g_2 \rangle = -q(\alpha_2) \geq -\langle g_1, M(\alpha_2) g_1 \rangle$.

Combining the inequalities yields

$$\langle g_2, (M(\alpha_1) - M(\alpha_2)) g_2 \rangle \geq \langle g_1, (M(\alpha_1) - M(\alpha_2)) g_1 \rangle . \tag{7.6}$$

Noting that $M(\alpha_1) - M(\alpha_2) = (\alpha_2 - \alpha_1) \mathcal{L}$, and plugging into (7.6) yields

$$\langle g_2, \mathcal{L} g_2 \rangle \geq \langle g_1, \mathcal{L} g_1 \rangle$$

Upon specializing to the unit norm eigenfunctions that attain the maximization in (7.5) we have

$$h_+(\alpha_2) = \langle g_2, \mathcal{L} g_2 \rangle \geq \langle g_1, \mathcal{L} g_1 \rangle = h_+(\alpha_1)$$

Similarly, we have that $h_-(\alpha_2) \geq h_-(\alpha_1)$.

b) Let $\alpha \in \mathbb{R}$, then we clearly have

$$h_+(\alpha) \geq h_-(\alpha) \geq 0$$

by the positive semidefinite property of $\mathcal{L}$. Let $v \in S(\alpha)$ be unit normed. Recall that the eigenfunction F_0 is associated with $\mu_0 = 0$ and hence $\langle F_0, \mathcal{L} F_0 \rangle = 0$. The following inequality holds by Rayleigh's inequality:

$$\langle v, M(\alpha) v \rangle \leq \langle F_0, M(\alpha) F_0 \rangle = \left\langle F_0, P_{u_0}^2 F_0 \right\rangle + 0 = \left\langle F_0, P_{u_0}^2 F_0 \right\rangle . \tag{7.7}$$

For $\alpha < 0$, subtracting $\left\langle v, P_{u_0}^2 v \right\rangle$ from both sides of 7.7 and multiplying by $-\frac{1}{\alpha}$ yields

$$\begin{aligned} \langle v, \mathcal{L} v \rangle &\leq -\frac{1}{\alpha} \left(\left\langle F_0, P_{u_0}^2 F_0 \right\rangle - \left\langle v, P_{u_0}^2 v \right\rangle \right) \\ &= -\frac{1}{\alpha} \left\langle (F_0 - v), P_{u_0}^2 (F_0 - v) \right\rangle \\ &\leq -\frac{\mathcal{E}_G^2(u_0)}{\alpha}, \end{aligned}$$

where the final inequality is due to the eccentricity $\mathcal{E}_G^2(u_0)$ being a global bound for the graph spread. As $\alpha \to -\infty$ we squeeze $h_\pm(\alpha)$ to zero:

$$0 \leq h_-(\alpha) \leq h_+(\alpha) \leq -\frac{\mathcal{E}^2(u_0)}{\alpha} \to 0,$$

which proves $\lim_{\alpha \to -\infty} h_\pm(\alpha) = 0$.

For the limit as $\alpha \to \infty$, recall that $\langle F_{N-1}, \mathcal{L}F_{N-1}\rangle = \mu_{N-1}$ and hence we have

$$\begin{aligned}\langle v, M(\alpha)v\rangle &\leq \langle F_{N-1}, M(\alpha)F_{N-1}\rangle \\ &= \left\langle F_{N-1}, P_{u_0}^2 F_{N-1}\right\rangle - \alpha\mu_{N-1} \\ &\leq \mathcal{E}_G^2(u_0) - \alpha\mu_{N-1}.\end{aligned}$$

Dividing both sides by $-\alpha$, and solving for $\langle v, \mathcal{L}v\rangle$ yields:

$$\langle v, \mathcal{L}, v\rangle \geq \mu_{N-1} - \frac{1}{\alpha}\left(\mathcal{E}_G^2(u_0) - \left\langle v, P_{u_0}^2 v\right\rangle\right).$$

Noting that $\mu_{N-1} \geq \langle v, \mathcal{L}v\rangle$ by the Rayleigh inequality and taking $\alpha \to \infty$ yields the desired result:

$$\mu_{N-1} \geq \lim_{\alpha\to\infty} h_+(\alpha) \geq \lim_{\alpha\to\infty} h_-(\alpha) \geq \mu_{N-1}.$$

c) We use eigenvalue perturbation results such as those in [34] to establish that $q(\alpha)$ is analytic for $[a,b] \cap (\mathcal{A})^c$ where $\mathcal{A}$ is a finite subset of $[a,b]$. $M(\alpha)$ is real, it is linear in α, hence analytic, and it is symmetric. By Theorem 2 on page 404 of [34], there exist N analytic functions $\mu_0(\cdot), ..., \mu_{N-1}(\cdot)$ and N analytic vector valued functions $\omega_0(\cdot), ..., \omega_{N-1}(\cdot)$ such that

$$\forall \alpha \in \mathbb{R}\ M(\alpha)\omega_j(\alpha) = \mu_j(\alpha)\omega_j(\alpha) \tag{7.8}$$

and

$$\langle \omega_j(\alpha), \omega_k(\alpha) \rangle = \begin{cases} 0 & \text{if } j \neq k \\ \\ 1 & \text{if } j = k \end{cases}.$$

Let $[a, b]$ be an arbitrary finite interval in $\mathbb{R}$, and fix $\alpha_0 \in (a, b)$. If $S(\alpha_0)$ is one dimensional, then exactly one eigenvalue function $\mu_j(\alpha_0)$ equals $q(\alpha_0)$. By the analycity of all the eigenvalue functions, there exists some δ ball about α_0, such that if $|\alpha - \alpha_0| < \delta$ then $\mu_j(\alpha) < \mu_k(\alpha)$ for $k \neq j$. Hence, $q(\alpha) = \mu_j(\alpha)$ for $\alpha \in (\alpha_0 - \delta, \alpha_0 + \delta)$ and therefore $q(\alpha)$ is analytic on the δ ball.

If $S(\alpha_0)$ has dimension greater than one, then more than one eigenvalue function from $\mu_l(\alpha_0)$ for $l = 0, ..., N-1$ attains the value $q(\alpha_0)$. In this case, $q(\alpha)$ may not be analytic in any neighborhood of α_0. For instance, if two of the eigenvalue functions cross at exactly α_0, then there is no derivative for $q(\alpha)$ at α_0. Define $p_j(\alpha)$ for $j = 0, ..., d \leq N-1$ as the distinct eigenvalue functions of $M(\alpha)$, and let n_j be the multiplicity of each function. For $[a, b] \subset \mathbb{R}$, define

$$\mathcal{A} = \bigcup_{0 \leq i < j \leq d} \{\alpha \in [a, b] : p_i(\alpha) = p_j(\alpha)\}.$$

As defined, $\mathcal{A}$ has finite order. Indeed, if $|\mathcal{A}| = \infty$ then at least two of the p_j's would be equal on an infinite set of points on the interval, and therefore would be equal on the interval because both are analytic.

To conclude the proof, we shall relate $q(\alpha)$ to $H_\pm(\alpha)$. For fixed $\alpha_0 \in [a, b]$, we without loss of generality, assume the first $k+1$ distinct eigenvalue functions p_i for $i = 0, ..., k$ intersect at α_0, and are minimal. That is to say, $p_i(\alpha_0) = q(\alpha_0)$. The associate eigenfunction functions are denoted by $\omega_{i,j}(\alpha)$ for $i =$

$0, ..., k$, and $j = 1, ..., n_i$. Hence, $\omega_{ij}(\alpha_0)$ form an orthonormal basis for $S(\alpha_0)$. Therefore, if $\tau \in S(\alpha_0)$ is unit normed, we have

$$\tau = \sum_{i=0}^{k} \sum_{j=1}^{n_i} c_{ij} \omega_{ij}(\alpha_0).$$

The coefficients $c_{ij} = \langle \tau, \omega_{ij}(\alpha) \rangle$ and therefore $\sum_{i=0}^{k} \sum_{j=1}^{n_i} c_{ij}^2 = \|\tau\|^2 = 1$. We define the analytic function $\tau(\alpha)$ such that $\tau(\alpha_0) = \tau$ as follows:

$$\tau(\alpha) = \sum_{i=0}^{k} \sum_{j=1}^{n_i} c_{ij} \omega_{ij}(\alpha).$$

Applying $M(\alpha)$ to $\tau(\alpha)$ yields

$$\begin{aligned} M(\alpha)\tau(\alpha) &= \sum_{i=0}^{k} \sum_{j=1}^{n_i} c_{ij} M(\alpha) \omega_{ij}(\alpha) \\ &= \sum_{i=0}^{k} \sum_{j=1}^{n_i} c_{ij} p_i(\alpha) \omega_{ij}(\alpha). \end{aligned} \tag{7.9}$$

We apply the product rule to differentiate equation (7.9) which yields

$$M'(\alpha)\tau(\alpha) + M(\alpha)\tau'(\alpha) = \sum_{i=0}^{k} \sum_{j=1}^{n_i} c_{ij} p_i'(\alpha) \omega_{ij}(\alpha) + \sum_{i=0}^{k} \sum_{j=1}^{n_i} c_{ij} p_i(\alpha) w_{ij}'(\alpha). \tag{7.10}$$

Evaluating equation (7.10) at α_0 yields

$$M'(\alpha_0)\tau(\alpha_0) + M(\alpha_0)\tau'(\alpha_0) = \sum_{i=0}^{k} \sum_{j=1}^{n_i} c_{ij} p_i'(\alpha_0) \omega_{ij}(\alpha_0) + \sum_{i=0}^{k} \sum_{j=1}^{n_i} c_{ij} p_i(\alpha_0) w_{ij}'(\alpha_0). \tag{7.11}$$

Noting that $M'(\alpha) = -\mathcal{L}$, $p_i(\alpha_0) = q(\alpha_0)$,

$$\langle \tau(\alpha_0), M(\alpha_0)\tau'(\alpha_0) \rangle = \langle M(\alpha_0)\tau(\alpha_0), \tau'(\alpha_0) \rangle = q(\alpha_0) \langle \tau(\alpha_0), \tau'(\alpha_0) \rangle,$$

and that

$$\left\langle \tau(\alpha_0), \sum_{i=0}^{k} \sum_{j=1}^{n_i} c_{ij} p_i'(\alpha_0) \omega_{ij}(\alpha_0) \right\rangle = \sum_{i=0}^{k} \sum_{j=1}^{n_i} c_{ij}^2 p_i',$$

we have that the inner product of $\tau(\alpha_0)$ with the left and right hand sides of equation (7.11) yields

$$
\begin{aligned}
-\langle \tau(\alpha_0), \mathcal{L}\tau(\alpha_0)\rangle + q(\alpha_0)\langle \tau(\alpha_0), \tau'(\alpha_0)\rangle = & \sum_{i=0}^{k}\sum_{j=1}^{n_i} c_{ij}^2 p_i' \\
& + q(\alpha_0)\sum_{i=0}^{k}\sum_{j=1}^{n_i} c_{ij}\langle \tau(\alpha_0)\omega_{ij}(\alpha_0)\rangle .
\end{aligned} \tag{7.12}
$$

The second summands on the LHS and RHS of equation (7.12) are equal, hence we have

$$\langle \tau(\alpha_0), \mathcal{L}\tau(\alpha_0)\rangle = -\sum_{i=0}^{k}\sum_{j=1}^{n_i} c_{ij}^2 p_i'.$$

Since $\tau(\alpha_0) = \tau \in S(\alpha_0)$ was arbitrary and unit normed, and since the dimension of $S(\alpha_0)$ is finite, maximizing (respectively minimizing) $\langle \tau, \mathcal{L}\tau\rangle$ is achieved by maximizing (respectively minimizing) over the $p_i(\alpha_0)$'s. Hence we have

$$H_+(\alpha_0) = \max_{0\le i\le k} -p_i'(\alpha_0),$$

and

$$H_-(\alpha_0) = \min_{0\le i\le k} -p_i'(\alpha_0).$$

Since all of the $p_i(\alpha)$ are distinct (except at α_0) in some neighborhood $\mathcal{N}$ covering α_0 and small enough that $\mathcal{N}\cap\mathcal{A} = \alpha_0$ or $\emptyset$, there exist $l, m \in \{0, ..., k\}$ such that

$$q(\alpha) = \begin{cases} p_l(\alpha) & \alpha \le \alpha_0 \\ p_m(\alpha) & \alpha \ge \alpha_0. \end{cases}$$

If for some $j \ne m$, $p_j'(\alpha_0) < p_m'(\alpha_0)$ then $p_j(\alpha) < p_m(\alpha)$ for some $\alpha \in \mathcal{N} \cap [\alpha_0, \infty)$. This contradicts the fact that $q(\alpha) = p_q(\alpha)$ on this interval. Similarly, if there exists some $j \neq l$, with $p'_j(\alpha_0) > \rho'_l(\alpha_0)$ there is a contradiction on $\mathcal{N} \cap (-\infty, \alpha_0]$. Hence, we have

$$H_+(\alpha) = -\rho'_q(\alpha) = -m'(\alpha) \text{ for } \alpha \in \mathcal{N} \cap [\alpha_0, \infty),$$

and

$$H_-(\alpha) = -\rho'_l(\alpha) = -m'(\alpha) \text{ for } \alpha \in \mathcal{N} \cap (-\infty, \alpha_0].$$

Right and left continuity follow from p_m and p_l having continuous derivatives. If $k = 0$, that is, if only one of the p_i functions aligns with m at α_0, then $q(\alpha)$ is analytic on $\mathcal{N}$ and we have $H_-(\alpha) = H_+(\alpha) = -m'(\alpha)$ on $\mathcal{N}$. If we denote $\mathcal{B}$ as the set of $\alpha \in [a, b]$ for which $H_-(\alpha) \neq H_+(\alpha)$, we must have $\mathcal{B} \subseteq \mathcal{A}$ and therefore $\mathcal{B}$ is a finite set. ■

Having established lemma 7.2.5, we shall use it to prove that the eigenspace $S(\alpha)$ associated with the minimal eigenvalue $q(\alpha)$ of $M(\alpha)$ precisely characterizes the uncertainty curve $\gamma(s)$ for $s \in (0, \mu_{n-1})$.

Theorem 7.2.6 *(Agaskar and Lu [1]) A function $f \in l^2(G)$ with $\Delta_s^2(f) \in (0, \mu_{N-1})$ achieves the uncertainty curve if and only if it is a nonzero eigenfunction in $S(\alpha)$ for some $\alpha \in \mathbb{R}$.*

Proof: The "if" direction was established in proposition 7.2.4. To show the other direction, we shall establish that for any function $f \in l^2(G)$ that achieves the uncertainty curve, there is an α and a unit norm $v \in S(\alpha)$ such that $\langle v, \mathcal{L}v \rangle = \Delta_s^2(f)$. As before, we assume f has unit norm (as we can normalize any function without affecting its spreads). Having also assumed f lies on the uncertainty curve, and being guaranteed that v lies on the curve by Proposition 7.2.4, we have $\Delta^2_{G,u_0} = \left\langle f, P^2_{u_0} f \right\rangle = \left\langle v, P^2_{u_0} \right\rangle$, and hence

$$\begin{aligned}\langle f, M(\alpha) f \rangle \Delta^2_{G,u_0}(f) - \alpha \Delta^2_s(f) \\ &= \Delta^2_{G,u_0}(v) - \alpha \Delta^2_s(v) \\ &= \langle v, M(\alpha) v \rangle \\ &= q(\alpha).\end{aligned}$$

Therefore, f must also be a unit vector in $S(\alpha)$. In order to complete that proof, we shall show that for any $s \in (0, \mu_{N-1})$ there is an α and a unit norm eigenfunction $v \in S(\alpha)$ such that $\langle v, \mathcal{L} v \rangle = s$.

Given $s \in (0, \mu_{N-1})$, parts (b) and (c) of Lemma 7.2.5 ensure that there exist $a < b$ such that $h_-(a) < s < h_+(b)$ and that on the interval $[a, b]$ there exists at most one point $\beta \in [a, b]$ at which $h_-(\beta) < h_+(\beta)$. The interval $[h_-(a), h_+(b)]$ can be written as the union of three subintervals:

$$[h_-(a), h_+(b)] = [h_-(a), h_-(\beta)) \cup [h_-(\beta), h_+(\beta)] \cup (h_+(\beta), h_+(b)].$$

Thus s must belong to one of these three intervals. If s is in the first or third subinterval, the continuity of $h_-(\alpha)$ and $h_+(\alpha)$, respectively, on these intervals guarantees some α hits the value s on one of these intervals. By the construction of the $h_\pm$, this also guarantees a v achieving the uncertainty curve exists. This leaves the less straight forward case of $s \in [h_-(\beta), h_+(\beta)]$. In order to show that there exists

$\tau \in S(\beta)$, we introduce

$$\tau_+ = \text{argmax}_{z\in S(\beta),\|z\|=1} \langle z, \mathcal{L}z\rangle$$

and

$$\tau_- = \text{argmax}_{z\in S(\beta),\|z\|=1} \langle z, \mathcal{L}z\rangle ,$$

and define for $\theta \in [0, \pi/2]$ the vector valued function

$$y(\theta) = \frac{\cos\theta\tau_+ + \sin\theta\tau_-}{(1+\sin(2\theta)\langle\tau_+, \tau_-\rangle)^{1/2}}.$$

The assumption that $h_-(\beta) \neq h_+(\beta)$ ensure that the denominator is nonzero. The denominator has norm squared given by

$$\begin{aligned}\|\cos\theta\tau_+ + \sin\theta\tau_-\|^2 &= 1 + 2\cos\theta\sin\theta\langle\tau_+, \tau_-\rangle \\ &= 1 + \sin(2\theta)\langle\tau_+, \tau_-\rangle\end{aligned}$$

so $\|y(\theta)\| = 1$. Further, as the composition of continuous functions, $y(\theta)$ is continuous, and, as the linear combination of elements of $S(\beta)$, $y(\theta) \in S(\beta)$. By continuity, the intermediate value theorem, and the fact that $\langle y(\pi/2), \mathcal{L}y(\pi/2)\rangle = h_-(\beta)$ and $\langle y(0), \mathcal{L}y(0)\rangle = h_+(\beta)$, we have that there exists $\theta_0 \in [0, \pi/2]$ such that $\langle y(\theta_0), \mathcal{L}y(\theta_0)\rangle = s$. This completes the characterization of the uncertainty curve. ■

7.3 Vertex Frequency Difference Operator Feasibility Region

We extend the concept of the feasibility region for graph and spectral spreads from [1], and the feasible regions discussed in the Bell labs uncertainty papers ([47], [35]). Define the *difference operator feasibility region* FR as follows:

$$FR = \left\{(x, y) : \|D_r f\|^2 = x \text{ and } \left\|D_r\widehat{f}\right\|^2 = y \text{ for some unit normed } f \neq 0 \in l^2(G)\right\}.$$

We shall prove the some key properties of the difference operator feasibility region.

Proposition 7.3.1 *Let FR be the difference operator feasibility region for a simple, and connected graph G with N vertices. Then, the following properties hold.*

a) FR is a closed subset subset of $[0, \lambda_{N-1}] \times [0, \lambda_{N-1}]$ where λ_{N-1} is the maximal eigenvalue of the Laplacian L

b) $y = 0$ and $x = \frac{1}{N}\sum_{j=0}^{N-1} \lambda_j$ is the only point on the horizontal axis in FR. $x = 0$ and $y = L_{0,0}$ is the only point on the vertical axis in FR.

c) FR is in the half plane defined by $x + y \geq \tilde{\lambda}_0 > 0$ with equality if and only if $\widehat{f}$ is in the eigenspace associated with $\tilde{\lambda}_0$.

d) If $N \geq 3$ then FR is a convex region.

Proof: Recall that

$$\|D_r f\|^2 = \langle f, Lf \rangle = \left\langle \widehat{f}, \Delta \widehat{f} \right\rangle,$$

and that

$$\left\|D_r \widehat{f}\right\|^2 = \left\langle \widehat{f}, L\widehat{f} \right\rangle.$$

Note that the operation $f \to \widehat{f}$ is an isomorphism of the unit ball in $l^2(G)$. Hence, for the entirety of this proof we rely on the fact that if a unit normed $g \in l^2(G)$ (respectively a unique unit normed $g \in l^2(G)$) achieves a value in the feasibility region for $\langle g, \Delta g \rangle$, and for $\langle g, Lg \rangle$ then there exists a unique unit normed $f \in l^2(G)$ (respectively a unique unit normed $f \in l^2(G)$) that achieves the same values for $\|D_r f\|^2$ and $\left\|D_r \widehat{f}\right\|^2$ respectively. Namely, $f = \chi g$ achieves the values.

a) By the properties of the Rayleigh quotient, we have any unit normed $g \in l^2(G)$

$$0 = \lambda_0 \leq \langle g, \Delta g \rangle \leq \lambda_{N-1}$$

with maximal equality if $g = [0, ..., 0, 1]'$, and that

$$0 = \lambda_0 \leq \langle g, Lg \rangle \leq \lambda_{N-1}$$

with maximal equality if g is in the eigenspace associated with λ_{N-1} for L. Hence, $FR \subset [0, \lambda_{N-1}] \times [0, \lambda_{N-1}]$. It is closed because FR is the image of a continuous mapping from the closed unit ball of $l^2(G)$ into $\mathbb{R}^2$.

b) $\langle g, Lg \rangle = 0$ if and only if $g = \pm\frac{1}{\sqrt{N}}[1, ..., 1]'$. Hence, we have

$$x = \langle g, \Delta g \rangle = \frac{1}{N}[1, ..., 1] \begin{bmatrix} \lambda_0 \\ \lambda_1 \\ \vdots \\ \lambda_{N-1} \end{bmatrix} = \frac{1}{N}\sum_{j=0}^{N-1} \lambda_j.$$

$x = 0$ if and only if $g = \pm[1, 0, ..., 0]'$. Hence, we have $y = L_{0,0}$, which is the degree of the first vertex of G.

c) This follows directly from Theorem 6.2.1.

d) To show that FR is convex we prove the following equivalent property in a similar fashion to Proposition 7.2.2 originally shown in [1].

Proposition 7.3.2 *Let g_1 and g_2 be functions on a graph G with $N \geq 3$ vertices such that*

$$\langle g_i, g_i \rangle = 1, \;\; \langle g_i, \Delta g_i \rangle = x_i, \;\; \textit{and} \;\; \langle g_i, Lg_i \rangle = y_i \;\textit{for}\; i = 1, 2. \tag{7.13}$$

Then for any $\beta \in [0,1]$, we can always find a function g on the graph satisfying

$$\langle g, g \rangle = 1, \;\; \langle g, \Delta g \rangle = x, \;\; and \;\; \langle g, Lg \rangle = y \tag{7.14}$$

where $x = \beta x_1 + (1-\beta)x_2$ and $y = \beta x_1 + (1-\beta)x_2$. That is to say, any line segment in $\mathbb{R}^2$ connecting (x_1, y_1) with (x_2, y_2) is in FR.

In order to prove the proposition, we shall formulate this as a problem in Sym_N, the space of $N \times N$ symmetric matrices. Every function $g \in l^2(G)$ can be mapped onto a symmetric, rank one $N \times N$ matrix M by setting $M = gg^*$. Further, if the following properties hold for a rank one matrix M (which can be decomposed as $M = gg^*$), then (7.14) holds for g:

1. $1 = \langle g, g \rangle = tr(g^* g) = tr(gg^*) = tr(M)$,
2. $x = \langle g, \Delta g \rangle = tr(g^* \Delta g) = tr(\Delta g g^*) = tr(\Delta M)$, and
3. $y = \langle g, Lg \rangle = tr(g^* Lg) = tr(Lgg^*) = tr(LM)$.

Again, we refer to Theorem (7.2.3) from [4] in order to show that such a rank one symmetric positive semi-definite exists.

Proof of proposition: Let g_1 and g_2 satisfy equation (7.13). Under the mapping $M_i = g_i g_i^*$ for $i = 1, 2$, each M_i satisfies

$$tr(M_i) = 1, \;\; tr(\Delta M_i) = x_i, \;\; \text{and } tr(LM_i) = y_i.$$

By construction, each M_i is symmetric positive semi-definite. For any $\beta \in [0,1]$, let $M' = \beta M_1 + (1-\beta)M_2$. Clearly, $M' \in S_+^N$ by the convexity of S_+^N

and if we let $x = \beta x_1 + (1-\beta)x_2$ and $y = \beta y_1 + (1-\beta)y_2$ then

$$M' \in \mathcal{H} = \{M \in \mathrm{Sym}_N : tr(M) = 1, tr(\Delta M) = x, \text{ and } tr(LM) = y\}.$$

By the linear independence of I, $\mathcal{L}$, and Δ, we have that $\mathcal{H}$ is an affine subspace of Sym_N with codimension 3. Hence, we have that $S_+^N \cap \mathcal{H} \neq \emptyset$. Noting that any element of $S_+^N \cap \mathcal{H}$ has nonnegative eigenvalues which must sum to 1, the boundedness of this subspace is straightforward to show:

$$\forall M \in S_+^N \cap \mathcal{H}, \;\; \|M\|_{fr} = tr(M^2) \leq tr(M) = 1.$$

By Theorem (7.2.3), we conclude that there exists a matrix M of rank one that can be decomposed as $M = gg^*$ with g satisfying equation (7.13) as desired. ■

We now turn our attention the lower boundary of FR: the differential uncertainty curve (DUC) (see Figure (7.2)) $\omega(x)$ is defined as

$$\forall x \in [0, \lambda_{N-1}], \; \omega(x) = \inf_{g \in l^2(G)} \langle g, Lg \rangle \text{ subject to } \langle g, \Delta g \rangle = x.$$

Given a fixed $x \in [0, \lambda_{N-1}]$, we say g' attains the DUC if for all g with $\langle g, \Delta \rangle = x$ we have

$$\langle g', Lg' \rangle \leq \langle g, Lg \rangle.$$

We shall show that for all $x \in [0, \lambda_{N-1}]$, there exists a function attaining the DUC. In fact, we shall show that certain eigenfunctions of the matrix valued function $K(\alpha) = L - \alpha\Delta$ will attain the DUC for every value of x. Hence, we shall show that

$$\omega(\alpha) = \min_{g \in l^2(G)} \langle g, Lg \rangle \text{ subject to } \langle g, \lambda g \rangle = x \text{ and } \langle g, g \rangle = 1.$$

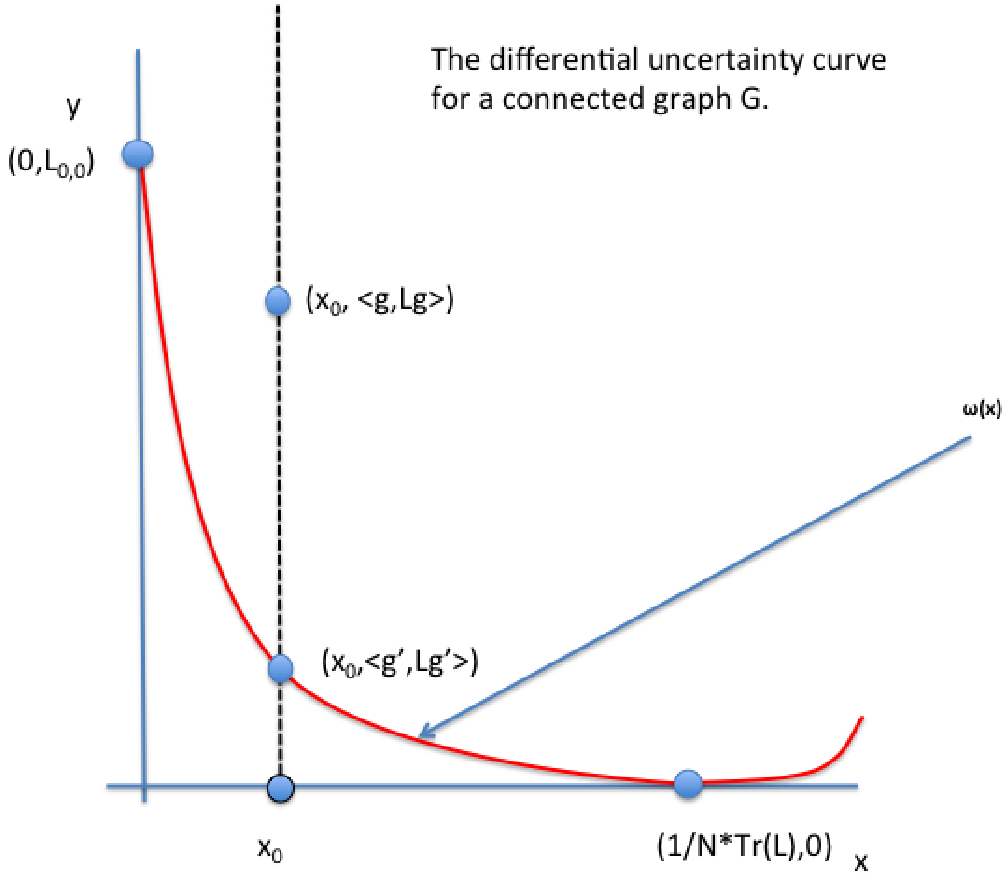

Figure 7.2: The differential uncertainty curve (red) for a connected graph G

We begin classifying $\omega(x)$ by motivating the use of the operator $K(\alpha)$. Finding values that attain the differential uncertainty curve amounts to solving a quadratically contrained convex optimization problem. We achieve this by defining the following Lagrangian function, and setting its gradient equal to zero. Define the DUC Lagrangian as

$$\Gamma(g, \alpha, \beta) = \langle g, Lg \rangle - \alpha(\langle g, \Delta g \rangle) - \beta(\langle g, g \rangle - 1).$$

Upon taking the gradient with respect to g and setting the gradient equal to zero, we have for some optimal g' that

$$\nabla_g(\Gamma(g, \alpha, \beta))(g') = 2Lg' - 2\alpha\Delta g' - 2\beta g' = 0$$

and

$$K(\alpha)g' = (L - \alpha\Delta)g' = \beta g'.$$

Thus the minimizer of the quadratically constrained problem is an eigenfunction of the operator $K(\alpha)$. Define $m(\alpha)$ to be the minimal eigenvalue of $K(\alpha)$, and define $\sigma(\alpha)$ to be its associated eigenspace. We shall prove that a function g attains the DUC if and only if it is $\sigma(\alpha)$. In order to prove this, we shall set up some tools for proving this.

We shall rely heavily on analysis of the following two functions:

$$H_+(\alpha) = \max_{g \in \sigma(\alpha):\|g\|=1} \langle g, \Delta g \rangle \tag{7.15}$$

$$H_-(\alpha) = \min_{g \in \sigma(\alpha):\|g\|=1} \langle g, \Delta g \rangle$$

which measure the maximal and, respectively, the minimal spectral spread that can be achieved by eigenfunctions in $S(\alpha)$.

Lemma 7.3.3 *The following properties hold for $H_+(\alpha)$ and $H_-(\alpha)$.*

a) For all $\alpha \in \mathbb{R}$) $H_+(\alpha)$, and $H_-(\alpha)$ are increasing functions increasing functions.

b) As α tends to infinity, $H_\pm(\alpha)$ limit to λ_{N-1}, and as α tends to negative infinity $H_\pm(\alpha)$ limit to 0.

c) *On any finite interval* $[a, b]$*, the functions differ on at most a finite number of points denoted by* $\Sigma = \{b_1, ..., b_k\}$ *for some* $k \geq 0$*. For all* $\alpha \notin \Sigma$*, the following holds:* $H_+(\alpha) = H_-(\alpha) = -m'(\alpha)$*.*

Proof:

a) For $\alpha_1 < \alpha_2$, we take any $\nu_1 \in \sigma(\alpha_1)$ and $\nu_2 \in \sigma(\alpha_2)$, and we have, by the Rayleigh quotient for symmetric matrices:

$$\langle \nu_2, K(\alpha_1)\nu_2 \rangle \geq m(\alpha_1) = \langle \nu_1, K(\alpha_1)\nu_1 \rangle .$$

Similarly, we have

$$-\langle \nu_2, K(\alpha_2)\nu_2 \rangle = -m(\alpha_2) \geq -\langle \nu_1, K(\alpha_2)\nu_1 \rangle .$$

Combining the inequalities yields

$$\langle \nu_2, (K(\alpha_1) - K(\alpha_2))\nu_2 \rangle \geq \langle \nu_1, (K(\alpha_1) - K(\alpha_2))\nu_1 \rangle . \tag{7.16}$$

Noting that $K(\alpha_1) - K(\alpha_2) = (\alpha_2 - \alpha_1)\,\Delta$, and plugging into (7.16) yields

$$\langle \nu_2, \Delta\nu_2 \rangle \geq \langle \nu_1, \Delta\nu_1 \rangle$$

Upon specializing to the unit norm eigenfunctions that attain the maximization in (7.15) we have

$$H_+(\alpha_2) = \langle \nu_2, \Delta\nu_2 \rangle \geq \langle \nu_1, \Delta\nu_1 \rangle = H_+(\alpha_1)$$

Similarly, upon specializing to the unit norm eigenfunctions that attain the minimum in (7.15) we have that $H_-(\alpha_2) = \langle \nu_2, \Delta\nu_2 \rangle \geq \langle \nu_1, \Delta\nu_1 \rangle = H_-(\alpha_1)$.

b) Let $\alpha \in \mathbb{R}$, then we clearly have

$$H_+(\alpha) \geq H_-(\alpha) \geq 0$$

by the positive semidefinite property of Δ. Let $\nu \in \sigma(\alpha)$ be unit normed. Recall that the canonical first basis vector e_0 spans the null space of Δ and hence $\langle e_0, \Delta e_0 \rangle = 0$. For any unit norm $\nu \in \sigma(\alpha)$, we have $\langle \nu, L\nu \rangle \geq 0$, and if $\alpha < 0$, we have $-\alpha \langle \nu, \Delta\nu \rangle \geq 0$. Thus by the properties the Rayleigh quotient we have

$$0 \leq -\alpha \langle \nu, \Delta\nu \rangle \leq \langle \nu, K(\alpha)\nu \rangle \leq \langle e_0, K(\alpha)e_0 \rangle = L_{o,o} + 0 = L_{o,o}. \tag{7.17}$$

Multiplying (7.17) by $-\frac{1}{\alpha}$ yields

$$0 \leq \langle \nu, \Delta\nu \rangle \leq -\frac{1}{\alpha} L_{0,0}.$$

Since this is valid for all $\nu \in \sigma(\alpha)$ we have

$$0 \leq H_-(\alpha) \leq H_+(\alpha) \leq -\frac{L_{0,0}}{\alpha}$$

As $\alpha \to -\infty$, we squeeze $H_\pm(\alpha)$ to zero as desired.

For the limit as $\alpha \to \infty$, recall that the last canonical eigenfunction e_{N-1} is in the eigenspace of λ_{N-1} for Δ. Hence, $\langle e_{N-1}, \Delta e_{N-1} \rangle = \lambda_{N-1}$, and we have

$$\begin{aligned} \langle \nu, K(\alpha)\nu \rangle &\leq \langle e_{N-1}, K(\alpha)e_{N-1} \rangle \\ &= \langle e_{N-1}, Le_{N-1} \rangle - \alpha\lambda_{N-1} \\ &= L_{N-1,N-1} - \alpha\lambda_{N-1}. \end{aligned}$$

Adding $(\alpha \langle \nu, \Delta\nu \rangle - L_{N-1,N-1})$ to both sides yields

$$\langle \nu, L\nu \rangle - L_{N-1,N-1} \leq \alpha \left(\langle \nu, \Delta\nu \rangle - \lambda_{N-1} \right) \leq 0 \tag{7.18}$$

where the last inequality in (7.18) is due to $\alpha > 0$ and the properties of the Rayleigh quotient. Taking the absolute value of both sides, and dividing by α yields

$$\left| \frac{\langle \nu, L\nu \rangle - L_{N,N}}{\alpha} \right| \geq |\langle \nu, \Delta\nu \rangle - \lambda_{N-1}| \geq 0.$$

The desired result follow from taking $\alpha \to \infty$.

c) We use eigenvalue perturbation results such as those in [34] to establish that $m(\alpha)$ is analytic for $[a,b] \cap (\Upsilon)^c$ where Υ is a finite subset of $[a,b]$. $K(\alpha)$ is real, it is linear in α, hence analytic, and it is symmetric. By Theorem 2 on page 404 of [34], there exist N analytic functions $\xi_0(\cdot), ..., \xi_{N-1}(\cdot)$ and N analytic vector valued functions $w_0(\cdot), ..., w_{N-1}(\cdot)$ such that

$$\forall \alpha \in \mathbb{R} \ K(\alpha) w_j(\alpha) = \xi_j(\alpha) w_j(\alpha) \tag{7.19}$$

and

$$\langle w_j(\alpha), w_k(\alpha) \rangle = \begin{cases} 0 & \text{if } j \neq k \\ 1 & \text{if } j = k \end{cases}.$$

Let $[a,b]$ be an arbitrary finite interval in $\mathbb{R}$, and fix $\alpha_0 \in (a,b)$. If $\sigma(\alpha_0)$ is one dimensional, then exactly one eigenvalue function $\xi_j(\alpha_0)$ equals $m(\alpha_0)$. By the analycity of all the eigenvalue functions, there exists some δ ball about α_0, such that if $|\alpha - \alpha_0| < \delta$ then $\xi_j(\alpha) < \xi_k(\alpha)$ for $k \neq j$. Hence, $m(\alpha) = \xi_j(\alpha)$ for $\alpha \in (\alpha_0 - \delta, \alpha_0 + \delta)$ and therefore $m(\alpha)$ is analytic on the δ ball.

In $\sigma(\alpha_0)$ has dimension greater than one, then more than one eigenvalue function from $\xi_l(\alpha_0)$ for $l = 0, ..., N-1$ attains the value $m(\alpha_0)$. In this case, $q(\alpha)$ may not be analytic in any neighborhood of α_0. For instance, if two of the eigenvalue functions cross at exactly α_0, then there is no derivative for $q(\alpha)$ at α_0. Define $\rho_j(\alpha)$ for $j = 0, ..., d \leq N-1$ as the distinct eigenvalue functions of $K(\alpha)$, and let n_j be the multiplicity of each function. For $[a, b] \subset \mathbb{R}$, define

$$\Upsilon = \bigcup_{0 \leq i < j \leq d} \{\alpha \in [a, b] : \rho_i(\alpha) = \rho_j(\alpha)\}.$$

As defined, Υ has finite order. Indeed, if $|\Upsilon| = \infty$ then at least two of the ρ_j's would be equal on an infinite set of points on the interval, and therefore would be equal on the interval because both are analytic.

To conclude the proof, we shall relate $m(\alpha)$ to $H_\pm(\alpha)$. For fixed $\alpha_0 \in [a, b]$, we without loss of generality, assume the first $k+1$ distinct eigenvalue functions ρ_i for $i = 0, ..., k$ intersect at α_0, and are minimal. That is to say, $\rho_i(\alpha_0) = m(\alpha_0)$. The associate eigenfunction functions are denoted by $w_{i,j}(\alpha)$ for $i = 0, ..., k$, and $j = 1, ..., n_i$. Hence, $w_{ij}(\alpha_0)$ form an orthonormal basis for $\sigma(\alpha_0)$. Therefore, if $\nu \in \sigma(\alpha_0)$ is unit normed, we have

$$\nu = \sum_{i=0}^{k} \sum_{j=1}^{n_i} c_{ij} w_{ij}(\alpha_0).$$

The coefficients $c_{ij} = \langle \nu, w_{ij}(\alpha) \rangle$ and therefore $\sum_{i=0}^{k} \sum_{j=1}^{n_i} c_{ij}^2 = \|\nu\|^2 = 1$. We define the analytic function $\nu(\alpha)$ such that $\nu(\alpha_0) = \nu$ as follows:

$$\nu(\alpha) = \sum_{i=0}^{k} \sum_{j=1}^{n_i} c_{ij} w_{ij}(\alpha).$$

Applying $K(\alpha)$ to $\nu(\alpha)$ yields

$$\begin{aligned} K(\alpha)\nu(\alpha) &= \sum_{i=0}^{k}\sum_{j=1}^{n_i} c_{ij} K(\alpha) w_{ij}(\alpha) \\ &= \sum_{i=0}^{k}\sum_{j=1}^{n_i} c_{ij} \rho_i(\alpha) w_{ij}(\alpha). \end{aligned} \tag{7.20}$$

We apply the product rule to differentiate equation (7.20) which yields

$$K'(\alpha)\nu(\alpha) + K(\alpha)\nu'(\alpha) = \sum_{i=0}^{k}\sum_{j=1}^{n_i} c_{ij}\rho_i'(\alpha) w_{ij}(\alpha) + \sum_{i=0}^{k}\sum_{j=1}^{n_i} c_{ij}\rho_i(\alpha) w_{ij}'(\alpha). \tag{7.21}$$

Evaluating equation (7.21) at α_0 yields

$$K'(\alpha_0)\nu(\alpha_0) + K(\alpha_0)\nu'(\alpha_0) = \sum_{i=0}^{k}\sum_{j=1}^{n_i} c_{ij}\rho_i'(\alpha_0) w_{ij}(\alpha_0) + \sum_{i=0}^{k}\sum_{j=1}^{n_i} c_{ij}\rho_i(\alpha_0) w_{ij}'(\alpha_0). \tag{7.22}$$

Noting that $K'(\alpha) = -\Delta$, $\rho_i(\alpha_0) = m(\alpha_0)$,

$$\langle \nu(\alpha_0), K(\alpha_0)\nu'(\alpha_0)\rangle = \langle K(\alpha_0)\nu(\alpha_0), \nu'(\alpha_0)\rangle = m(\alpha_0)\langle \nu(\alpha_0), \nu'(\alpha_0)\rangle,$$

and that

$$\left\langle \nu(\alpha_0), \sum_{i=0}^{k}\sum_{j=1}^{n_i} c_{ij}\rho_i'(\alpha_0) w_{ij}(\alpha_0) \right\rangle = \sum_{i=0}^{k}\sum_{j=1}^{n_i} c_{ij}^2 \rho_i',$$

we have that the inner product of $\nu(\alpha_0)$ with the left and right hand sides of equation (7.22) yields

$$\begin{aligned} -\langle \nu(\alpha_0), \Delta\nu(\alpha_0)\rangle + m(\alpha_0)\langle \nu(\alpha_0), \nu'(\alpha_0)\rangle &= \sum_{i=0}^{k}\sum_{j=1}^{n_i} c_{ij}^2 \rho_i' \\ &\quad + m(\alpha_0)\sum_{i=0}^{k}\sum_{j=1}^{n_i} c_{ij}\langle \nu(\alpha_0) w_{ij}(\alpha_0)\rangle. \end{aligned} \tag{7.23}$$

The second summands on the LHS and RHS of equation (7.23) are equal, hence we have

$$\langle \nu(\alpha_0), \Delta\nu(\alpha_0)\rangle = -\sum_{i=0}^{k}\sum_{j=1}^{n_i} c_{ij}^2 \rho_i'.$$

Since $\nu(\alpha_0) = \nu \in \sigma(\alpha_0)$ was arbitrary and unit normed, and since the dimension of $\sigma(\alpha_0)$ is finite, maximizing (respectively minimizing) $\langle \nu, \Delta\nu\rangle$ is achieved by maximizing (respectively minimizing) over the $\rho_i(\alpha_0)$'s. Hence we have

$$H_+(\alpha_0) = \max_{0\le i\le k} -\rho_i'(\alpha_0),$$

and

$$H_-(\alpha_0) = \min_{0\le i\le k} -\rho_i'(\alpha_0).$$

Since all of the $\rho_i(\alpha)$ are distinct (except at α_0) in some neighborhood $\mathcal{N}$ covering α_0 and small enough that $\mathcal{N} \cap \Upsilon = \alpha_0$ or $\emptyset$, there exist $l, m \in \{0, ..., k\}$ such that

$$m(\alpha) = \begin{cases} \rho_l(\alpha) & \alpha \le \alpha_0 \\ \\ \rho_m(\alpha) & \alpha \ge \alpha_0. \end{cases}$$

If for some $j \neq m$, $\rho_j'(\alpha_0) < \rho_m'(\alpha_0)$ then $\rho_j(\alpha) < \rho_m(\alpha)$ for some $\alpha \in \mathcal{N} \cap [\alpha_0, \infty)$. This contradicts the fact that $m(\alpha) = \rho_m(\alpha)$ on this interval. Similarly, if there exists some $j \neq l$, with $\rho_j'(\alpha_0) > \rho_l'(\alpha_0)$ there is a contradiction on $\mathcal{N} \cap (-\infty, \alpha_0]$. Hence, we have

$$H_+(\alpha) = -\rho_m'(\alpha) = -m'(\alpha) \text{ for } \alpha \in \mathcal{N} \cap [\alpha_0, \infty),$$

and

$$H_-(\alpha) = -\rho_l'(\alpha) = -m'(\alpha) \text{ for } \alpha \in \mathcal{N} \cap (-\infty, \alpha_0].$$

Right and left continuity follow from ρ_m and ρ_l having continuous derivatives. If $k = 0$, that is, if only one of the ρ_i functions aligns with m at α_0, then $m(\alpha)$ is analytic on $\mathcal{N}$ and we have $H_-(\alpha) = H_+(\alpha) = -m'(\alpha)$ on $\mathcal{N}$. If we denote Σ as the set of $\alpha \in [a, b]$ for which $H_-(\alpha) \neq H_+(\alpha)$, we must have $\Sigma \subseteq \Upsilon$ and therefore Σ is a finite set. ■

We now show that vectors in $\sigma(\alpha)$ characterize the DUC.

Theorem 7.3.4 *A unit normed function $f \in l^2(G)$ with $\|D_r f\|^2 = x \in (0, \lambda_{N-1})$ achieves the uncertainty curve if and only if $\widehat{f}$ is a nonzero eigenfunction in $\sigma(\alpha)$ for some $\alpha \in \mathbb{R}$.*

Proof: As before, it suffices to show that if a unit normed $\eta \in l^2(G)$ satisfying $\langle \eta, \Delta\eta \rangle = x \in (0, \lambda_{N-1})$ achieves the DUC if and only if $\eta \in \sigma(\alpha)$ for some $\alpha \in \mathbb{R}$.

For this "if" direction, fix $\alpha \in \mathbb{R}$. Then for any arbitrary unit norm $\eta \in l^2(G)$ we have

$$\langle \eta, K(\alpha)\eta \rangle = \langle \eta, L\eta \rangle - \alpha \langle \eta, \Delta\eta \rangle .$$

The Rayleigh quotient for $K(\alpha)$ is bounded sharply below by $m(\alpha)$, hence we conclude that for any unit normed $\nu \in \sigma(\alpha)$ that

$$\begin{aligned} \langle \nu, L\nu \rangle - \alpha \langle \nu, \Delta\nu \rangle &= m(\alpha) \\ &\leq \langle \eta, L\eta \rangle - \alpha \langle \eta, \Delta\eta \rangle . \end{aligned}$$

Upon restricting η to $\langle \eta, \Delta\eta \rangle = x$ we have

$$\langle \eta, L\eta \rangle - x \geq \langle \nu, L\nu \rangle - x \quad \Longleftrightarrow \quad \langle \eta, L\eta \rangle \geq \langle \nu, L\nu \rangle .$$

Hence, any unit normed $\nu \in \sigma(\alpha)$ achieves the DUC.

To prove the "only if" direction, it suffices to show that for any function $\eta \in l^2(G)$ that achieves the DUC, there is an α and a unit norm $v \in \sigma(\alpha)$ such that $\langle \nu, \Delta\nu \rangle = \langle \eta, \Delta\eta \rangle = x$. Indeed, having also assumed η lies on the uncertainty curve, and being guaranteed that such a ν lies on the curve by the "if" direction of this proof, we have $\langle \eta, L\eta \rangle = \langle \nu, L\nu \rangle$, and hence

$$\begin{aligned} \langle \eta, K(\alpha)\eta \rangle &= \langle \eta, L\eta \rangle - \alpha x \\ &= \langle \nu, L\nu \rangle - \alpha x \\ &= \langle \nu, K(\alpha)\nu \rangle \\ &= q(\alpha). \end{aligned}$$

Therefore, η must also be a unit vector in $\sigma(\alpha)$.

We complete the proof by showing that for any $x \in (0, \mu_{N-1})$ there is an α and a unit norm eigenfunction $\nu \in \sigma(\alpha)$ such that $\langle \nu, \Delta\nu \rangle = s$.

Given $x \in (0, \mu_{N-1})$, parts (b) and (c) of Lemma 7.3.3 ensure that there exist $a' < b'$ such that $H_-(a') < s < H_+(b')$ and that there are $a < b$ with $a' \le a < b \le b'$ such that on the interval $[a, b]$ there exists at most one point $\beta \in [a, b]$ at which $H_-(\beta) < H_+(\beta)$. The interval $[H_-(a), H_+(b)]$ can be written as the union of three subintervals:

$$[H_-(a), H_+(b)] = [H_-(a), H_-(\beta)) \cup [H_-(\beta), H_+(\beta)] \cup (H_+(\beta), H_+(b)].$$

Thus x must belong to one of these three intervals. If x is in the first or third subinterval, the continuity of $H_-(\alpha)$ and $H_+(\alpha)$, respectively, on these intervals guarantees for some α_-, respectively, α_+ that $H_(\alpha) = x$, respectively, $H_+(\alpha_+) = x$, on one of these intervals. By the construction of the $H_\pm$ functions, this also guarantees a ν achieving the uncertainty curve exists.

It remains to be shown that such an α and ν exist for $x \in [H_-(\beta), H_+(\beta)]$. We introduce

$$\nu_+ = \text{argmax}_{z\in\sigma(\beta),\|z\|=1} \langle z, \Delta z\rangle$$

and

$$\nu_- = \text{argmin}_{z\in\sigma(\beta),\|z\|=1} \langle z, \Delta z\rangle ,$$

and define for $\theta \in [0, \pi/2]$ the vector valued function

$$\upsilon(\theta) = \frac{\cos\theta\nu_+ + \sin\theta\nu_-}{(1 + \sin(2\theta)\langle\nu_+, \nu_-\rangle)^{1/2}}.$$

The assumption that $H_-(\beta) \neq H_+(\beta)$ ensure that the denominator is nonzero. The numerator has norm squared given by

$$\begin{aligned}\|\cos\theta\nu_+ + \sin\theta\nu_-\|^2 &= 1 + 2\cos\theta\sin\theta\langle\nu_+, \nu_-\rangle \\ &= 1 + \sin(2\theta)\langle\nu_+, \nu_-\rangle\end{aligned}$$

so $\|\upsilon(\theta)\| = 1$. Further, as the composition of continuous functions, $\upsilon(\theta)$ is continuous, and, as the linear combination of elements of $\sigma(\beta)$, $\upsilon(\theta) \in \sigma(\beta)$. By continuity, the intermediate value theorem, and the fact that $\langle\upsilon(\pi/2), \Delta\upsilon(\pi/2)\rangle = H_-(\beta)$ and $\langle\upsilon(0), \Delta\upsilon(0)\rangle = H_+(\beta)$, we have that there exists $\theta_0 \in [0, \pi/2]$ such that $\langle\upsilon(\theta_0), \Delta\upsilon(\theta_0)\rangle = x$. ■

7.4 The Complete Graph Revisited

We shall compute the differential feasibility region for the complete graph, and compare it to the analogous feasibility results in the Bell labs paper [35]. We begin our analysis by analyzing the eigenspace of $K(\alpha)$.

Proposition 7.4.1 *Let G be the unit weighted complete graph with N vertices. For all $\alpha \neq 0 \in \mathbb{R}$, if $K(\alpha) = L - \alpha\Delta$ where L is the graph Laplacian for G and Δ is its diagonalization, then $K(\alpha)$ has an $N-2$ degree eigenspace associated with the eigenvalue $N(1-\alpha)$.*

Proof: $K(\alpha)$ is a block matrix of the form

$$K(\alpha) = \left[\begin{array}{c|c} N-1 & -\mathbf{1}_{N-1}^t \\ \hline -\mathbf{1}_{N-1} & C(\alpha) \end{array}\right],$$

where $\mathbf{1}_{N-1}$ is the $(N-1)\times 1$ constant function of all ones, and $C(\alpha)$ is the circulant matrix with $N-1-N\alpha$ on the diagonal and -1 at every other coordinate, i.e.,

$$C(\alpha) = N(1-\alpha)I_{N-1\times N-1} - O_{N-1\times N-1}.$$

Let $V \subset \mathbb{R}^{N-1}$ be the orthogonal compliment of $span(\mathbf{1}_{N-1})$ in $\mathbb{R}^{N-1}$. Then for all $f \in V$ we have that

$$C(\alpha)f = N(1-\alpha)f - O_{N-1\times N-1}f = N(1-\alpha)f.$$

V has dimension $N-2$ and may be embedded in $l^2(G)$ via the mapping

$$f \mapsto \begin{bmatrix} 0 \\ f \end{bmatrix}.$$

and we denote this space as $\tilde{V}$. Hence, we have that

$$K(\alpha)\begin{bmatrix}0\\ \\ f\end{bmatrix} = N(1-\alpha)\begin{bmatrix}0\\ \\ f\end{bmatrix}$$

and the eigenspace $ES(\alpha)$ associated with $N(1-\alpha)$ has at least dimension $N-2$ as it properly contains $\tilde{V}$. Let a and b denote the remaining two eigenvalues. Let v_a be an eigenvector associated with a and orthogonal to all $v \in \tilde{V}$. Then v_a is of the form

$$v_a = c[x1...1]'$$

for some real constant c. Without loss of generality, we set $c = 1$ and we have

$$K(\alpha)v_a = \begin{bmatrix}(N-1)x-(N-1)\\ -x+(1-\alpha N)\\ \vdots\\ -x+(1-\alpha N)\end{bmatrix} = av_a.$$

Therefore, we must have $a = -x + (1-\alpha N)$. Solving the quadratic resulting from equality in the first coordinate

$$x^2 - (2 - N(\alpha+1))x - (N-1) = 0$$

yields

$$x = \frac{2 - N(\alpha+1) \pm \sqrt{(N(\alpha+1)-2)^2 + 4(N-1)}}{2}.$$

We conclude that

$$a = 1 - \alpha N - \frac{2 - N(\alpha+1) + \sqrt{(N(\alpha+1)-2)^2 + 4(N-1)}}{2}$$

and

$$b = 1 - \alpha N - \frac{2 - N(\alpha+1) - \sqrt{(N(\alpha+1)-2)^2 + 4(N-1)}}{2}.$$

We conclude that $ES(\alpha)$ has dimension $N-2$ as desired. ■

From the proof of Proposition 7.4.1 we find that the minimal eigenvalue of $K(\alpha)$ is

$$\lambda_{min}(\alpha) = -\frac{-N(\alpha+1) + \sqrt{(N(\alpha+1)-2)^2 + 4(N-1)}}{2} - \alpha(N), \qquad (7.24)$$

for all $\alpha \neq 0$. When $\alpha = 0$ the minimum eigenvalue is zero, so we may conclude that equation (7.24) holds for all $\alpha \in \mathbb{R}$.

Let $[x(\alpha), 1, ..., 1]'$ with

$$x(\alpha) = \frac{2 - N(\alpha+1) \pm \sqrt{(N(\alpha+1)-2)^2 + 4(N-1)}}{2}$$

be vector valued eigenfunction associated with $\lambda_{min}(\alpha)$ for all $\alpha \in \mathbb{R}$. Upon apply the Rayleigh quotient to this vector we find that the DUC is the lower boundary of the ellipse with coordinates

$$\left(\frac{N(N-1)}{x(\alpha)^2 + (N-1)}, \frac{(x(\alpha)-1)^2(N-1)}{x(\alpha)^2 + (N-1)}\right).$$

The differential feasibility regions for various values of N are displayed in Figure 7.3. There is an intriguing connection between these regions and the feasibility regions in [35] and [47]. Figure 7.4 displays the feasibility region for time and band limited functions. We see that the region is convex, and is a similar but rotated shape to the differential uncertainty region for complete graphs. We leave further rigorous study of this apparent relationship to future work.

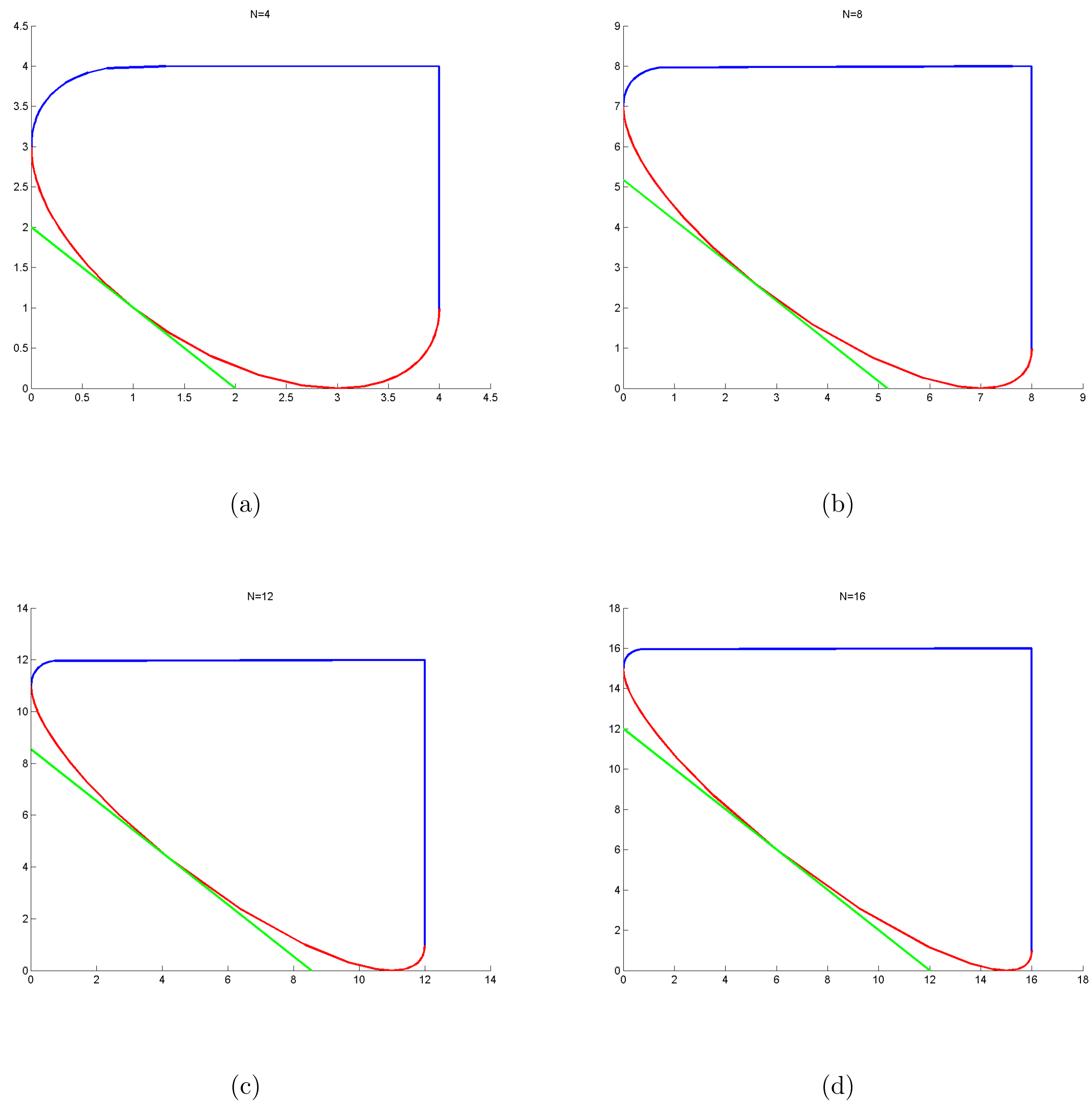

Figure 7.3: The complete graph differential feasibility regions for various values of N. The red curve is the differential uncertainty curve, the blue is the remaining differential feasibility region boundary, and the green line is the line $x+y = N-\sqrt{N}$

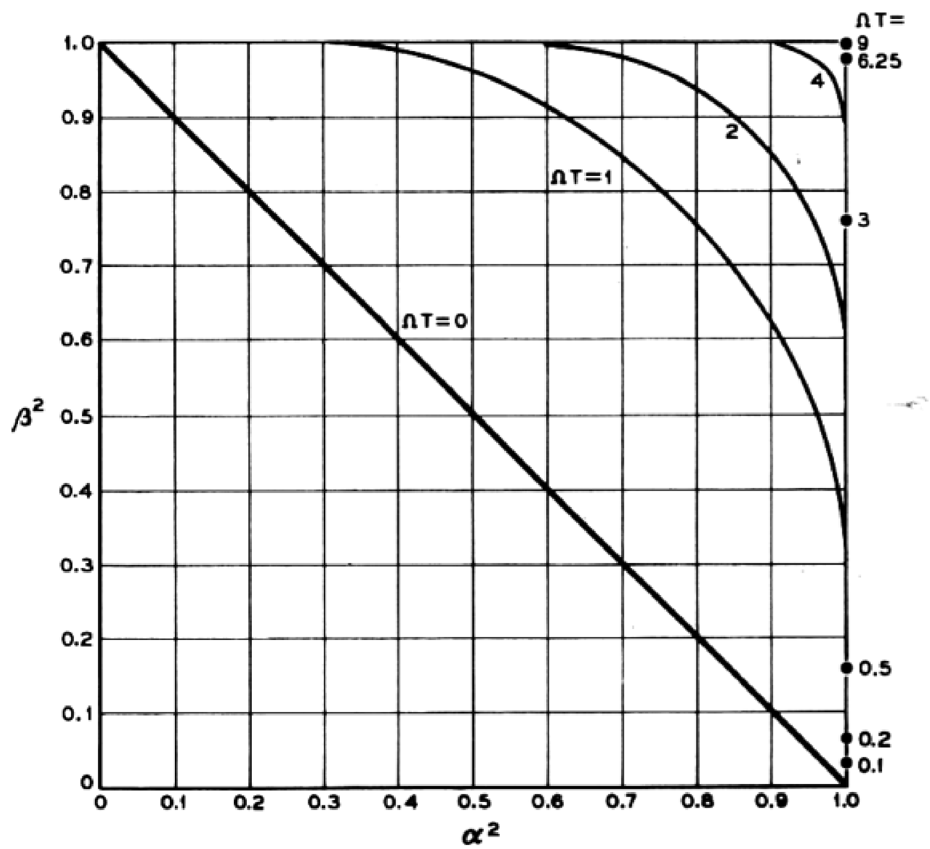

Figure 7.4: The feasibility region in [35]. α and β are the norms of possible time limited and band limited functions derived from $f \in L^2(\mathbb{R})$.

Bibliography

[1] A. Agaskar and Y. M. Lu. A spectral graph uncertainty principle. *Information Theory, IEEE Transactions on*, 59(7):4338–4356, 2013.

[2] N. Akhiezer and I. Glazman. *Theory of Linear Operators in Hilbert Space*. Dover Books on Mathematics Series. Dover Publications, 1993.

[3] S. Axler. *Linear Algebra Done Right*. Undergraduate Texts in Mathematics. Springer International Publishing, 2014.

[4] A. Barvinok. A remark on the rank of positive semidefinite matrices subject to affine constraints. *Discrete and Computational Geometry*, 25:2001, 2001.

[5] J. Benedetto and W. Czaja. *Integration and Modern Analysis*. Birkhäuser Advanced Texts Basler Lehrbücher. Birkhäuser, 2010.

[6] J. J. Benedetto. *Uncertainty principle inequalities and spectrum estimation*. Springer, 1990.

[7] J. J. Benedetto. Irregular sampling and frames. *Wavelets: A Tutorial in Theory and Applications*, 2:445–507, 1992.

[8] J. J. Benedetto. *Wavelets: Mathematics and Applications*. Studies in Advanced Mathematics. Taylor & Francis, 1993.

[9] J. J. Benedetto. *Harmonic Analysis and Applications*, volume 23. CRC Press, 1996.

[10] J. J. Benedetto and P. J. Koprowski. Graph theoretical uncertainty principles. *In preparation*, 2015.

[11] W. Briggs and V. Henson. *The DFT: An Owners' Manual for the Discrete Fourier Transform*. Society for Industrial and Applied Mathematics, 1995.

[12] P. Casazza and G. Kutyniok. *Finite Frames: Theory and Applications*. Applied and Numerical Harmonic Analysis. Birkhäuser Boston, 2012.

[13] O. Christensen. *An Introduction to Frames and Riesz Bases*. Applied and Numerical Harmonic Analysis. Birkhäuser Boston, 2003.

[14] O. Christensen. *Frames and Bases: An Introductory Course*. Applied and Numerical Harmonic Analysis. Birkhäuser, 2008.

[15] F. Chung. Graph theory in the information age. *Notices of the AMS*, 57(6):726–732, 2010.

[16] F. R. Chung. *Spectral Graph Theory*, volume 92. American Mathematical Soc., 1997.

[17] P. V. David I Shuman, Benjamin Ricaud. Vertex-frequency analysis on graphs. *preprint*, 2013.

[18] R. Diestel. *Graph Theory*. Electronic library of mathematics. Springer, 2006.

[19] D. L. Donoho and P. B. Stark. Uncertainty principles and signal recovery. *SIAM Journal on Applied Mathematics*, 49(3):906–931, 1989.

[20] R. J. Duffin and A. C. Schaeffer. A class of nonharmonic fourier series. *Transactions of the American Mathematical Society*, pages 341–366, 1952.

[21] V. N. Ekambaram, G. Fanti, B. Ayazifar, and K. Ramchandran. Wavelet regularized graph semi supervised learning. *Global Conference on Signal and Information Processing*, 2013.

[22] G. B. Folland. *A course in abstract harmonic analysis*. CRC press, 1994.

[23] F. A. Grünbaum. The eigenvectors of the discrete fourier transform: A version of the hermite functions. *Journal of Mathematical Analysis and Applications*, 88(2):355–363, 1982.

[24] F. A. Grünbaum. The Heisenberg inequality for the discrete Fourier transform. *Applied and Computational Harmonic Analysis*, 15(2):163–167, 2003.

[25] D. K. Hammond, P. Vandergheynst, and R. Gribonval. Wavelets on graphs via spectral graph theory. *Applied and Computational Harmonic Analysis*, 30(2):129–150, 2011.

[26] W. Heisenberg. Über den anschaulichen inhalt der quantentheoretischen kinematik und mechanik. *Zeitschrift für Physik*, 43(3-4):172–198, 1927.

[27] E. Hernandez and G. Weiss. *A First Course on Wavelets*. Studies in Advanced Mathematics. CRC Press, 1996.

[28] E. Hewitt and K. A. Ross. *Abstract harmonic analysis*, volume 1. Springer Science & Business Media, 1994.

[29] E. Hewitt and K. A. Ross. *Abstract harmonic analysis*, volume 2. Springer Science & Business Media, 1994.

[30] K. Hoffman and R. Kunze. *Linear algebra*. Prentice-Hall mathematics series. Prentice-Hall, 1971.

[31] R. Horn and C. Johnson. *Matrix Analysis*. Matrix Analysis. Cambridge University Press, 2012.

[32] J. H. Hubbard and B. B. Hubbard. Vector calculus, linear algebra, and differential forms: a unified approach. *AMC*, 10:12, 2002.

[33] M. Lammers and A. Maeser. An uncertainty principle for finite frames. *Journal of Mathematical Analysis and Applications*, 373(1):242–247, 2011.

[34] P. Lancaster and M. Tismenetsky. *The theory of matrices : with applications*. Computer science and applied mathematics. Academic Press, Orlando, 1985. Includes index.

[35] H. J. Landau and H. O. Pollak. Prolate spheroidal wave functions, fourier analysis and uncertainty II. *Bell System Technical Journal*, 40(1):65–84, 1961.

[36] S. Mallat. *A Wavelet Tour of Signal Processing*. Wavelet Tour of Signal Processing. Elsevier Science, 1999.

[37] S. Mallat. *A Wavelet Tour of Signal Processing: The Sparse Way*. Elsevier Science, 2008.

[38] W. Pauli. Über den zusammenhang des abschlusses der elektronengruppen im atom mit der komplexstruktur der spektren. *Zeitschrift für Physik A Hadrons and Nuclei*, 31(1):765–783, 1925.

[39] M.-T. Pham, G. Mercier, and J. Michel. Wavelets on graphs for very high resolution multispectral image texture segmentation. In *Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International*, pages 2273–2276. IEEE, 2014.

[40] H. Royden and P. Fitzpatrick. *Real Analysis*. Prentice Hall, 2010.

[41] W. Rudin. *Functional Analysis*. International series in pure and applied mathematics. McGraw-Hill, 2006.

[42] A. Sandryhaila and J. Moura. Discrete signal processing on graphs: Frequency analysis. *Signal Processing, IEEE Transactions on*, 62(12), 2014.

[43] D. Serre. *Matrices: Theory and Applications*. Graduate Texts in Mathematics. Springer New York, 2010.

[44] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst. The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. *Signal Processing Magazine, IEEE*, 30(3):83–98, 2013.

[45] D. I. Shuman, B. Ricaud, and P. Vandergheynst. A windowed graph Fourier transform. In *Statistical Signal Processing Workshop (SSP), 2012 IEEE*, pages 133–136. Ieee, 2012.

[46] D. I. Shuman, B. Ricaud, and P. Vandergheynst. Vertex-frequency analysis on graphs. *arXiv preprint arXiv:1307.5708*, 2013.

[47] D. Slepian. Prolate spheroidal wave functions, Fourier analysis, and uncertainty V: The discrete case. *Bell System Technical Journal*, 57(5):1371–1430, 1978.

[48] J. Smith. *Mathematics of the Discrete Fourier Transform (DFT): With Audio Applicaitons*. BookSurge Publishing, 2007.

[49] T. Tao. An uncertainty principle for cyclic groups of prime order. *Math. Res. Let.*, 12(1):121–127, 2005.

[50] H. Weyl. Quantenmechanik und gruppentheorie. *Zeitschrift für Physik*, 46(1-2):1–46, 1927.

[51] W. Wong. *Discrete Fourier Analysis*. Pseudo-Differential Operators. Springer Basel, 2011.