UNCERTAINTY PRINCIPLES IN HILBERT SPACES

SAY SONG GOH†
DEPARTMENT OF MATHEMATICS
NATIONAL UNIVERSITY OF SINGAPORE
10 KENT RIDGE CRESCENT, SINGAPORE 119260
REPUBLIC OF SINGAPORE
AND
CHARLES A. MICCHELLI†‡
DEPARTMENT OF MATHEMATICS AND STATISTICS
STATE UNIVERSITY OF NEW YORK
THE UNIVERSITY AT ALBANY
ALBANY, NY 12222
USA

ABSTRACT. In this paper we provide several generalizations of inequalities bounding the commutator of two linear operators acting on a Hilbert space which relate to the Heisenberg uncertainty principle and time/frequency analysis of periodic functions. We develop conditions that ensure these inequalities are sharp and apply our results to concrete examples of importance in the literature.

2000 *Mathematics Subject Classification.* 47B47, 47B25, 41A50, 42A05.

Key words and phrases. Uncertainty inequalities, commutator, Hilbert space.

†Supported in part by the Wavelets Strategic Research Programme, National University of Singapore, under a grant from the National Science and Technology Board and the Ministry of Education, Singapore.

‡Supported in part by the US National Science Foundation under grants DMS-9504780, DMS-9973427.

1. Introduction

Uncertainty principles have attracted wide interest. First formulated as a principle of quantum mechanics by Heisenberg in 1927 [28] and later Weyl's reinterpretation [36] as an inequality led to operator theoretic extensions of various types. Recent information on these ideas can be found in the articles and books [1], [5], [8], [9], [15], [16], [28]. For other perspectives in the study of uncertainty, see [7], [12]. Our goal here is to explore uncertainty inequalities on three fronts. First, we study inequalities bounding the commutator of two linear operators acting on a Hilbert space, see Section 2. Existing inequalities of this type assume that the two operators are self-adjoint. Here a general inequality is derived without this assumption. We make various improvements of the upper bound of our inequality and available inequalities by introducing two other linear operators with certain commuting properties and applying approximation techniques. The analysis is enriched by examples on periodic functions and finite-dimensional Hilbert spaces.

Next, we examine, both in some generality as well as in specific cases, the issue of equality and asymptotic equality in the new inequalities, see Section 3. A necessary and sufficient condition for equality is obtained. Our results in Section 3 are motivated by the well-known observation that the classical Heisenberg inequality is sharp for the Gaussian function, see for example [8], while the recent version for periodic functions introduced in [2] and later used in time/frequency analysis of wavelet representations [10], [23], [26], [27], [29] was shown in [27] to be only asymptotically sharp. It is this relationship of uncertainty inequalities to time/frequency analysis that sparked our interest in the subject. In the periodic case analyzed, we consider a more general setup and characterize situations for which equality is possible. In situations where equality cannot be attained, we obtain sufficient conditions for asymptotic equality. Functions that give equality or asymptotic equality in the respective inequalities are also identified.

The third idea that we explore here is rooted in the uncertainty principles arising from the dichotomy between band and time limited functions. This subject is of fundamental importance in electrical engineering where it has stimulated enormous interest in the mathematical analysis of signals [17]–[19], [30]–[34]. The uncertainty inequality in [18] gives a precise description of the concentration of a function and its Fourier transform when it has total energy one. Specifically, this result describes the common field of values of two appropriate quadratic forms given the third is normalized to one. In another context, the interplay among three quadratic forms appears in optimal estimation and n-widths [21], [22], [25], see also [3]. Here we look into the application of these ideas to uncertainty inequalities that involve three quadratic forms, see Section 4. We study the region formed by two of these quadratic forms when the third takes the value one. This geometric problem is investigated for several situations and uncertainty inequalities. It gives general results on self-adjoint linear operators as well as information on regions arising from the Heisenberg uncertainty principle and the periodic inequality applied to various classes of periodic functions including real-valued even trigonometric polynomials.

We begin with a review of the classical Heisenberg inequality. Let $\mathcal{L}_2(\mathbb{R})$ be the space of complex-valued, square-integrable functions over the real line $\mathbb{R}$ with inner product

$$\langle f, g\rangle_{\mathbb{R}} := \int_{\mathbb{R}} f(t)\overline{g(t)}\, dt, \quad f, g \in \mathcal{L}_2(\mathbb{R})$$

and corresponding norm $\|\cdot\|_{\mathbb{R}}$. We let $\mathcal{W}_2^1(\mathbb{R})$ denote the space of absolutely continuous functions f in $\mathcal{L}_2(\mathbb{R})$ such that $f' \in \mathcal{L}_2(\mathbb{R})$. The Heisenberg uncertainty principle concerns the position operator M and the momentum operator D given by the equations

$$(Mf)(t) := tf(t), \qquad (Df)(t) := if'(t), \qquad t \in \mathbb{R}. \tag{1.1}$$

The domain of the operator M consists of all functions $f \in \mathcal{L}_2(\mathbb{R})$ such that $Mf \in \mathcal{L}_2(\mathbb{R})$ and the domain of the operator D is $\mathcal{W}_2^1(\mathbb{R})$. The uncertainty principle says if $f \neq 0$ then

$$\frac{1}{4} \leq \nu_{\mathbb{R}}(f) \tag{1.2}$$

where

$$\nu_{\mathbb{R}}(f) := \frac{\left(\|Mf\|_{\mathbb{R}}^2 - \frac{|\langle Mf, f\rangle_{\mathbb{R}}|^2}{\|f\|_{\mathbb{R}}^2}\right)\left(\|Df\|_{\mathbb{R}}^2 - \frac{|\langle Df, f\rangle_{\mathbb{R}}|^2}{\|f\|_{\mathbb{R}}^2}\right)}{\|f\|_{\mathbb{R}}^4}. \tag{1.3}$$

A weaker version of (1.2) is the inequality

$$\frac{1}{4} \leq \frac{\|Mf\|_{\mathbb{R}}^2 \|Df\|_{\mathbb{R}}^2}{\|f\|_{\mathbb{R}}^4} \tag{1.4}$$

valid whenever $f,\ Mf,\ Df \in \mathcal{L}_2(\mathbb{R})$ (with $f \neq 0$), see for example [28]. Moreover equality in (1.4) is achieved if and only if $f(t) = ae^{bt}e^{-ct^2}$, $t \in \mathbb{R}$, for some $a \in \mathbb{C}\backslash\{0\}$, $b \in \mathbb{C}$ and $c \in \mathbb{R}_+\backslash\{0\}$.

It is well known that inequality (1.2) can be viewed as a consequence of an inequality for *self-adjoint* linear operators in a complex Hilbert space. Indeed, let $\mathcal{H}$ be a complex Hilbert space with inner product $\langle\cdot,\cdot\rangle$ and corresponding norm $\|\cdot\|$. Suppose that A and B are *any* linear operators with domain and range in $\mathcal{H}$. (The operators A and B can be bounded or unbounded.) We define an operator $[A, B] : \mathcal{D}(AB) \cap \mathcal{D}(BA) \longrightarrow \mathcal{H}$ by the formula

$$[A, B] := AB - BA$$

where $\mathcal{D}(AB)$ and $\mathcal{D}(BA)$ are the domains of the operators AB and BA respectively. The operator $[A, B]$ is called the *commutator* of A and B. Since we require the adjoint of A and B we will always assume throughout the paper that A and B are densely defined on $\mathcal{H}$. Fixing notation, we define for a linear operator A with domain and range in $\mathcal{H}$ and any $x \in \mathcal{D}(A)\backslash\{0\}$ the quantity

$$\Delta_x(A) := \left(\|Ax\|^2 - \frac{|\langle Ax, x\rangle|^2}{\|x\|^2}\right)^{1/2}.$$

We note that

$$\Delta_x(A) = \left\| Ax - \frac{\langle Ax, x\rangle}{\|x\|^2} x \right\|$$

and $\Delta_x(A) = 0$ if and only if x is an eigenvector of A. The following theorem on the commutator is a generalization of the Heisenberg uncertainty principle, see for instance [8], [15], [16].

Theorem 1.1. *Let A and B be any self-adjoint linear operators with domain and range in the same complex Hilbert space $\mathcal{H}$. For any nonzero $x \in \mathcal{D}(AB) \cap \mathcal{D}(BA)$ there holds*

$$|\langle [A,B]x,\, x \rangle| \le 2\Delta_x(A)\Delta_x(B). \tag{1.5}$$

For $\mathcal{H} = \mathcal{L}_2(\mathbb{R})$, $A = M$ and $B = D$, we have that $[A,B] = -iI$ and so we conclude that inequality (1.2) follows from (1.5) when $f \in \mathcal{D}(MD) \cap \mathcal{D}(DM)$. The range of validity of this inequality is extended to all functions such that f, Mf and $Df \in \mathcal{L}_2(\mathbb{R})$ by an approximation argument, see for example [28].

In [2], an uncertainty principle for the space $\mathcal{L}_2[0,2\pi)$ of complex-valued, square-integrable 2π-periodic functions was proposed. To describe this result we use

$$\langle f,g\rangle_{2\pi} := \frac{1}{2\pi}\int_{[0,2\pi)} f(t)\overline{g(t)}\,dt, \quad f,g \in \mathcal{L}_2[0,2\pi)$$

for the inner product on $\mathcal{L}_2[0,2\pi)$, $\|\cdot\|_{2\pi}$ for the corresponding norm and consider the operator

$$(Sf)(t) := e^{it}f(t), \quad t \in [0,2\pi)$$

which acts as a shift operator on the Fourier coefficients of $f \in \mathcal{L}_2[0,2\pi)$. Let $\mathcal{W}_2^1[0,2\pi)$ be the space of absolutely continuous functions f in $\mathcal{L}_2[0,2\pi)$ such that $f' \in \mathcal{L}_2[0,2\pi)$. It was observed in [2] that for any $f \in \mathcal{W}_2^1[0,2\pi)\backslash\{0\}$ there follows the inequality

$$\frac{1}{4}|\langle Sf,f\rangle_{2\pi}|^2 \le \left(\|f\|_{2\pi}^2 - \frac{|\langle Sf,f\rangle_{2\pi}|^2}{\|f\|_{2\pi}^2}\right)\left(\|Df\|_{2\pi}^2 - \frac{|\langle Df,f\rangle_{2\pi}|^2}{\|f\|_{2\pi}^2}\right). \tag{1.6}$$

If $\langle Sf,f\rangle_{2\pi} \ne 0$ this inequality can be rewritten as

$$\frac{1}{4} \le \nu_{2\pi}(f) \tag{1.7}$$

where

$$\nu_{2\pi}(f) := \frac{\left(\|f\|_{2\pi}^2 - \frac{|\langle Sf,f\rangle_{2\pi}|^2}{\|f\|_{2\pi}^2}\right)\left(\|Df\|_{2\pi}^2 - \frac{|\langle Df,f\rangle_{2\pi}|^2}{\|f\|_{2\pi}^2}\right)}{|\langle Sf,f\rangle_{2\pi}|^2}.$$

Note that S is not self-adjoint and so Theorem 1.1 does not apply to this case. This suggests the possibility of a result more general than Theorem 1.1.

2. Inequality

In this section, we shall obtain various improvements of the uncertainty principle embodied in Theorem 1.1. To this end we let A^* and B^* be the adjoint of the operators A and B respectively. Throughout the paper, we shall use the notation

$$\mathcal{D}(A|B) := \mathcal{D}(AB) \cap \mathcal{D}(BA) \cap \mathcal{D}(A^*) \cap \mathcal{D}(B^*).$$

One can verify for any $x \in \mathcal{D}(A)\backslash\{0\}$ that

$$\min_{a\in\mathbb{C}} \|Ax - ax\| = \Delta_x(A) \tag{2.1}$$

and the minimum value is attained uniquely at $a = \dfrac{\langle Ax, x\rangle}{\|x\|^2}$.

Theorem 2.1. *Let A and B be linear operators with domain and range in the same complex Hilbert space $\mathcal{H}$. For any nonzero $x \in \mathcal{D}(A|B)$ there holds*

$$|\langle\, [A,B]x,\, x\,\rangle| \le \Delta_x(A)\Delta_x(B^*) + \Delta_x(A^*)\Delta_x(B). \tag{2.2}$$

Proof: First, observe that for any nonzero $x \in \mathcal{D}(A|B)$,

$$\langle\, [A,B]x,\, x\,\rangle = \langle ABx, x\rangle - \langle BAx, x\rangle = \langle Bx, A^*x\rangle - \langle Ax, B^*x\rangle.$$

This implies that

$$\begin{aligned} |\langle\, [A,B]x,\, x\,\rangle| &\le |\langle Ax, B^*x\rangle| + |\langle Bx, A^*x\rangle| \\ &\le \|Ax\|\,\|B^*x\| + \|A^*x\|\,\|Bx\|. \end{aligned} \tag{2.3}$$

For any $a, b \in \mathbb{C}$, we replace A and B above by $A - aI$ and $B - bI$ respectively to obtain that

$$|\langle\, [A,B]x,\, x\,\rangle| \le \|(A - aI)x\|\,\|(B^* - \overline{b}I)x\| + \|(A^* - \overline{a}I)x\|\,\|(B - bI)x\|.$$

Using (2.1), we conclude that

$$\begin{aligned} \min_{a,b\in\mathbb{C}} \{\|Ax - ax\|\,\|B^*x - \overline{b}x\| &+ \|A^*x - \overline{a}x\|\,\|Bx - bx\|\} \\ &= \Delta_x(A)\Delta_x(B^*) + \Delta_x(A^*)\Delta_x(B) \end{aligned}$$

from which (2.2) follows. ■

Corollary 2.2. *Let A and B be any normal linear operators with domain and range in the same complex Hilbert space $\mathcal{H}$. For any nonzero $x \in \mathcal{D}(A|B)$ we have that*

$$|\langle\, [A,B]x,\, x\,\rangle| \le 2\|Ax\|\,\|Bx\| \tag{2.4}$$

and in addition there holds

$$|\langle\, [A,B]x,\, x\,\rangle| \le 2\Delta_x(A)\Delta_x(B). \tag{2.5}$$

Proof: Since A and B are normal operators, we have for any $x \in \mathcal{D}(A) \cap \mathcal{D}(B)$ that

$$\|Ax\| = \|A^*x\|, \qquad \|Bx\| = \|B^*x\|.$$

Thus (2.4) and (2.5) follow from (2.3) and (2.2) respectively. ■

The uncertainty principle (1.6) is a consequence of Corollary 2.2 since the operator S is normal and the operator D is self-adjoint. In general, inequality (2.5) need not hold. For example on the Hilbert space $\mathbb{C}^2$, we consider the matrices

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{2.6}$$

and choose $x = (1,0)^T$. Then the left hand side of (2.5) is one while the right hand side is zero.

The results in Theorem 2.1 and Corollary 2.2 rely upon inequality (2.3) and the fact that the identity operator commutes with *any* operator. Let us try to improve upon this idea. Starting with operators A and B, we try to find two operators U and V that yield better bounds on $|\langle\, [A,B]x,\, x\,\rangle|$. We shall use the operators U and V to reduce the upper bound in (2.3). To this end, we require that $x \in \mathcal{D}(A|B) \cap \mathcal{D}(A|V) \cap \mathcal{D}(B|U) \cap \mathcal{D}(U|V)$ and that

$$[A,V]x = [B,U]x = [U,V]x = 0. \tag{2.7}$$

These conditions imply the equation

$$\langle\, [A,B]x,\, x\,\rangle = \langle\, Bx - Vx,\, A^*x - U^*x\,\rangle - \langle\, Ax - Ux,\, B^*x - V^*x\,\rangle$$

and consequently the inequality

$$|\langle\, [A,B]x,\, x\,\rangle| \le \|Ax - Ux\|\,\|B^*x - V^*x\| + \|A^*x - U^*x\|\,\|Bx - Vx\|. \tag{2.8}$$

There arises the problem of minimizing (2.8) subject to the condition (2.7). We see two reasons that this is a formidable task. Firstly, the identification of all operators U and V such that $[A,V]x = [B,U]x = 0$ is difficult and the difficulty is compounded by the additional requirement that $[U,V]x = 0$. Secondly, the objective function on the right hand side of inequality (2.8) is complicated. Thus we are forced to explore less than optimal strategies to make effective use of (2.8). In this connection, we choose U to be a multiple of the identity, proceed as in the proof of Theorem 2.1 to derive the bound

$$|\langle\, [A,B]x,\, x\,\rangle| \le \|Bx - Vx\|\Delta_x(A^*) + \|B^*x - V^*x\|\Delta_x(A) \tag{2.9}$$

and then replace (2.9) by

$$|\langle\, [A,B]x,\, x\,\rangle| \le \left\{\|Bx - Vx\|^2 + \|B^*x - V^*x\|^2\right\}^{1/2}\left\{\Delta_x^2(A^*) + \Delta_x^2(A)\right\}^{1/2} \tag{2.10}$$

where both inequalities are valid whenever $[A,V]x = 0$.

To provide insight into the use of these inequalities we return to the Hilbert space $\mathcal{L}_2[0,2\pi)$. Let $\mathcal{W}_\infty^1[0,2\pi)$ be the space of complex-valued, absolutely continuous 2π-periodic functions f such that $f' \in \mathcal{L}_\infty[0,2\pi)$. For a function $w \in \mathcal{W}_\infty^1[0,2\pi)\backslash\{0\}$, we apply Corollary 2.2 to the multiplication operator M_w defined by

$$M_w f := wf, \quad f \in \mathcal{L}_2[0,2\pi)$$

and the derivative operator D to conclude for any $f \in \mathcal{W}_2^1[0,2\pi)$ with $\langle M_{w'}f, f\rangle_{2\pi} \ne 0$ that

$$\frac{1}{4} \le \nu_{2\pi}(f;w) \tag{2.11}$$

where

$$\nu_{2\pi}(f;w) := \frac{\left(\|M_w f\|_{2\pi}^2 - \dfrac{|\langle M_w f, f\rangle_{2\pi}|^2}{\|f\|_{2\pi}^2}\right)\left(\|Df\|_{2\pi}^2 - \dfrac{|\langle Df, f\rangle_{2\pi}|^2}{\|f\|_{2\pi}^2}\right)}{|\langle M_{w'}f, f\rangle_{2\pi}|^2}. \tag{2.12}$$

To improve upon this inequality we use the operators aI and M_h where $a \in \mathbb{C}$ and $h \in \mathcal{L}_\infty[0,2\pi)$ as the operators U and V in (2.8). It is worth mentioning that if V is a bounded linear operator which commutes with M_w where $w \in \mathcal{L}_\infty[0,2\pi)$, then $V = M_h$ for some $h \in \mathcal{L}_\infty[0,2\pi)$, see [11]. With this choice of operators, (2.8) becomes

$$|\langle M_{w'}f, f\rangle_{2\pi}| \leq \left(\|Df - M_h f\|_{2\pi} + \|Df - M_{\overline{h}} f\|_{2\pi}\right) \|M_w f - af\|_{2\pi}. \tag{2.13}$$

Since (2.10) is valid whenever $[A,V]x = 0$, we note that for a fixed $f \in \mathcal{W}_2^1[0,2\pi)\backslash\{0\}$, h can be *any* 2π-periodic function such that $hf \in \mathcal{L}_2[0,2\pi)$. Based on this observation, we use (2.10) to obtain the following result.

Theorem 2.3. *For any $f \in \mathcal{W}_2^1[0,2\pi)\backslash\{0\}$ and $w \in \mathcal{W}_\infty^1[0,2\pi)\backslash\{0\}$, there holds*

$$|\langle M_{w'}f, f\rangle_{2\pi}| \leq 2\left\{\|M_w f\|_{2\pi}^2 - \frac{|\langle M_w f, f\rangle_{2\pi}|^2}{\|f\|_{2\pi}^2}\right\}^{1/2} \times \tag{2.14}$$
$$\left\{\|Df\|_{2\pi}^2 - \frac{1}{2\pi}\int_{\{t: f(t)\neq 0,\, t\in[0,2\pi)\}} \frac{(\mathrm{Re}\{\overline{f}(t)(Df)(t)\})^2}{|f(t)|^2}\,dt\right\}^{1/2}.$$

Proof: Specializing (2.10) to the case at hand gives the inequality

$$|\langle M_{w'}f, f\rangle_{2\pi}| \leq \sqrt{2}\left\{\|M_w f\|_{2\pi}^2 - \frac{|\langle M_w f, f\rangle_{2\pi}|^2}{\|f\|_{2\pi}^2}\right\}^{1/2} \gamma(h)^{1/2}$$

where

$$\gamma(h) := \|p - hf\|_{2\pi}^2 + \|q - \overline{h}f\|_{2\pi}^2,$$

$p = q = Df$ and h is a 2π-periodic function such that $hf \in \mathcal{L}_2[0,2\pi)$. Let χ be the characteristic function of the set $\{t : f(t) \neq 0,\, t \in [0,2\pi)\}$. Since

$$\gamma(h) = \|p\|_{2\pi}^2 + \|q\|_{2\pi}^2 - \frac{1}{2}\left\|\frac{(\overline{f}p + f\overline{q})\chi}{f}\right\|_{2\pi}^2 + 2\left\|\left(fh - \frac{\overline{f}p + f\overline{q}}{2\overline{f}}\right)\chi\right\|_{2\pi}^2, \tag{2.15}$$

we choose

$$h := \begin{cases} \dfrac{\overline{f}p + f\overline{q}}{2|f|^2}, & \text{if } f \neq 0, \\ 0, & \text{otherwise} \end{cases} = \begin{cases} \dfrac{\mathrm{Re}\{\overline{f}(Df)\}}{|f|^2}, & \text{if } f \neq 0, \\ 0, & \text{otherwise} \end{cases}$$

to conclude that the minimum of the quantity $\gamma(h)$ is given by

$$\min_h \gamma(h) = 2\|Df\|_{2\pi}^2 - 2\left\|\frac{(\mathrm{Re}\{\overline{f}(Df)\})\chi}{f}\right\|_{2\pi}^2$$

which leads us to the result. ■

The improvement of (2.14) over (2.5) as an inequality is essentially given by

$$\frac{|\langle Df, f\rangle_{2\pi}|^2}{\|f\|_{2\pi}^2} \leq \frac{1}{2\pi}\int_{\{t: f(t)\neq 0,\, t\in[0,2\pi)\}} \frac{(\mathrm{Re}\{\overline{f}(t)(Df)(t)\})^2}{|f(t)|^2}\,dt.$$

If f is real-valued or of the form $e^{ij\cdot}$ where $j \in \mathbb{Z}$, equality holds in this inequality and (2.14) yields the same estimate as (2.5). There are functions for which (2.14) provides a much smaller upper bound than (2.5). Indeed, for $f(t) = c + e^{it}$ where $c \geq 2$, we have that

$$\frac{|\langle Df, f\rangle_{2\pi}|^2}{\|f\|_{2\pi}^2} = \frac{1}{c^2+1} < \frac{c^2+2}{2(c+1)^2} \leq \frac{1}{2\pi}\int_{\{t: f(t)\neq 0,\, t\in[0,2\pi)\}} \frac{(\mathrm{Re}\{\overline{f}(t)(Df)(t)\})^2}{|f(t)|^2}\, dt.$$

We shall now shift the focus of our study of inequality (2.8) back to a general Hilbert space. To this end, we take the operators U and V to be of the form

$$U = \sum_{k=0}^{n} a_k B^k, \qquad V = bI$$

where n is a positive integer and $a_0, \ldots, a_n, b \in \mathbb{C}$. Let x be an element in

$$\mathcal{D}_n(A|B) := \mathcal{D}(AB) \cap \mathcal{D}(BA) \cap \mathcal{D}(A^*) \cap \mathcal{D}(B^n) \cap \mathcal{D}((B^*)^n).$$

Note that $\mathcal{D}_1(A|B) = \mathcal{D}(A|B)$. Specializing (2.8) to the case at hand results in the inequality

$$|\langle\, [A,B]x,\, x\,\rangle| \leq \|Ax - \sum_{k=0}^{n} a_k B^k x\|\, \|B^*x - \overline{b}x\| + \|A^*x - \sum_{k=0}^{n} \overline{a}_k (B^*)^k x\|\, \|Bx - bx\|.$$

The minimum values of $\|Bx - bx\|$ and $\|B^*x - \overline{b}x\|$ are both attained at $b = \dfrac{\langle Bx, x\rangle}{\|x\|^2}$. Hence, we obtain the inequality

$$|\langle\, [A,B]x,\, x\,\rangle| \leq \|Ax - \sum_{k=0}^{n} a_k B^k x\|\, \Delta_x(B^*) + \|A^*x - \sum_{k=0}^{n} \overline{a}_k (B^*)^k x\|\, \Delta_x(B). \tag{2.16}$$

This suggests that we examine the quantity

$$\delta_n(A|B)(x) := \min_{a_0,\ldots,a_n\in\mathbb{C}} \left\{ \|Ax - \sum_{k=0}^{n} a_k B^k x\|\, \Delta_x(B^*) + \|A^*x - \sum_{k=0}^{n} \overline{a}_k (B^*)^k x\|\, \Delta_x(B) \right\}$$

since the estimate (2.16) takes the form

$$|\langle\, [A,B]x,\, x\,\rangle| \leq \delta_n(A|B)(x).$$

The minimum above exists because this problem can be recast as the best approximation of the vector (Ax, A^*x) in $\mathcal{H}\times\mathcal{H}$ relative to the norm

$$\|(u,u')\| := \|u\|\Delta_x(B^*) + \|u'\|\Delta_x(B), \quad (u,u') \in \mathcal{H}\times\mathcal{H}$$

by the subspace consisting of all *real* linear combinations of the vectors $(B^k x, (B^*)^k x)$, $i(B^k x, -(B^*)^k x)$, $k = 0, \ldots, n$, on $\mathcal{H}\times\mathcal{H}$. By definition, whenever $x \in \mathcal{D}_{n+1}(A|B)$ we have $\delta_{n+1}(A|B)(x) \leq \delta_n(A|B)(x)$. At least for linear operators on a *finite* dimensional Hilbert space this sequence is eventually constant.

Proposition 2.4. *Let $\mathcal{H}$ be a complex Hilbert space of dimension m. For any linear operators A and B with domain and range in $\mathcal{H}$, $x \in \mathcal{H}$ and $n \geq m-1$, we have that $\delta_n(A|B)(x) = \delta_{m-1}(A|B)(x)$.*

Proof: By the Cayley-Hamilton theorem, see for instance [13], there are complex constants $\alpha_0, \dots, \alpha_{m-1}$ such that

$$B^m = \alpha_0 I + \alpha_1 B + \cdots + \alpha_{m-1} B^{m-1}.$$

Hence for any $a_0, \dots, a_n \in \mathbb{C}$, $n \geq m-1$, there exist complex constants $\gamma_0, \dots, \gamma_{m-1}$ such that

$$Ax - \sum_{k=0}^{n} a_k B^k x = Ax - \sum_{k=0}^{m-1} \gamma_k B^k x$$

and

$$A^* x - \sum_{k=0}^{n} \overline{a}_k (B^*)^k x = A^* x - \sum_{k=0}^{m-1} \overline{\gamma}_k (B^*)^k x.$$

This proves the result. ■

To benefit from inequality (2.16) we should minimize the right hand side of (2.16) over $a_0, \dots, a_n \in \mathbb{C}$. However, this problem seems to be formidable. As in (2.10), we apply the Cauchy-Schwarz inequality to the right hand side of (2.16) to obtain the inequality

$$\begin{aligned} |\langle [A,B]x,\, x \rangle| \quad &\leq \quad \left\{ \|Ax - \sum_{k=0}^{n} a_k B^k x\|^2 + \|A^* x - \sum_{k=0}^{n} \overline{a}_k (B^*)^k x\|^2 \right\}^{1/2} \\ &\quad \times \left\{ \Delta_x^2(B) + \Delta_x^2(B^*) \right\}^{1/2}. \end{aligned} \tag{2.17}$$

To resolve the minimum of the right hand side of (2.17), we require the following lemma. To this end, for any elements $y_1, \dots, y_m, y_1', \dots, y_m'$ in $\mathcal{H}$, we introduce the quantity

$$\mathcal{G}\left((y_1, y_1'), \dots, (y_m, y_m')\right) := \det\left(\langle y_i, y_j \rangle + \overline{\langle y_i', y_j' \rangle}\right)_{i,j=1}^{m}.$$

Lemma 2.5. *If $v_0, \dots, v_n, v_0', \dots, v_n'$ are elements in a complex Hilbert space $\mathcal{H}$ then the quantity $\mathcal{G}((v_0, v_0'), \dots, (v_n, v_n'))$ is nonzero if and only if there do not exist complex constants $\lambda_0, \dots, \lambda_n$ such that $\sum_{k=0}^{n} |\lambda_k| \neq 0$ and*

$$\sum_{k=0}^{n} \overline{\lambda}_k v_k = \sum_{k=0}^{n} \lambda_k v_k' = 0.$$

Moreover, when $\mathcal{G}((v_0, v_0'), \dots, (v_n, v_n'))$ is nonzero, for any u, u' in $\mathcal{H}$ the quantity

$$\|u - \sum_{k=0}^{n} a_k v_k\|^2 + \|u' - \sum_{k=0}^{n} \overline{a}_k v_k'\|^2$$

has a unique minimum over $a_0, \dots, a_n \in \mathbb{C}$ at $\widetilde{a}_0, \dots, \widetilde{a}_n \in \mathbb{C}$ defined by the equations

$$\sum_{k=0}^{n} \widetilde{a}_k \left(\langle v_k, v_\ell \rangle + \overline{\langle v_k', v_\ell' \rangle}\right) = \langle u, v_\ell \rangle + \overline{\langle u', v_\ell' \rangle}, \quad \ell = 0, \dots, n$$

and the value of the minimum is

$$\frac{\mathcal{G}\left((u,u'),(v_0,v_0'),\dots,(v_n,v_n')\right)}{\mathcal{G}\left((v_0,v_0'),\dots,(v_n,v_n')\right)}.$$

Proof: The proof follows by differentiation and subsequent simplifications. ■

Before we apply Lemma 2.5 to (2.17), let us comment on the original problem of minimizing the right hand side of (2.16). We shall discuss this problem in the more general setting of minimizing the quantity

$$F(a) := \|u - \sum_{k=0}^{n} a_k v_k\|\sigma + \|u' - \sum_{k=0}^{n} \overline{a}_k v_k'\|\sigma' \tag{2.18}$$

over $a := (a_0, \dots, a_n)^T \in \mathbb{C}^{n+1}$ where $v_0, \dots, v_n, v_0', \dots, v_n', u, u'$ are elements in $\mathcal{H}$ and σ, σ' are positive constants. The minimum of $F(a)$ need not be unique. For instance, for the case when $\mathcal{H} = \mathbb{C}$, $n = 0$, $v_0 = v_0' = 1$, $u = 1$, $u' = 0$ and $\sigma = \sigma' = 1$, the minimum of $F(a_0) = |1 - a_0| + |a_0|$ is attained at all $\widetilde{a}_0 \in [0,1]$.

The following lemma characterizes all the minimum points of $F(a)$.

Lemma 2.6. *Suppose that both $v_0, \dots, v_n$ and $v_0', \dots, v_n'$ are linearly independent elements in a complex Hilbert space $\mathcal{H}$ and u, u' are any two elements in $\mathcal{H}$. For $0 \le \rho \le \infty$, define the $(n+1) \times 1$ vector $a(\rho) = (a_0(\rho), \dots, a_n(\rho))^T$ by*

$$a(\rho) := \begin{cases} (G + \rho\overline{G}')^{-1}(b + \rho\overline{b}'), & \text{if } 0 \le \rho < \infty, \\ (\overline{G}')^{-1}\overline{b}', & \text{if } \rho = \infty \end{cases}$$

where G, G' are $(n+1) \times (n+1)$ matrices and b, b' are $(n+1) \times 1$ vectors given by

$$G := (\langle v_k, v_\ell \rangle)_{\ell,k=0}^{n}, \qquad G' := (\langle v_k', v_\ell' \rangle)_{\ell,k=0}^{n}$$

and

$$b := (\langle u, v_0 \rangle, \dots, \langle u, v_n \rangle)^T, \qquad b' := (\langle u', v_0' \rangle, \dots, \langle u', v_n' \rangle)^T$$

respectively.

(a) *If $\widetilde{a}$ is a minimum point of $F(a)$ in (2.18) with*

$$\|u - \sum_{k=0}^{n} \widetilde{a}_k v_k\| \cdot \|u' - \sum_{k=0}^{n} \overline{\widetilde{a}}_k v_k'\| \neq 0, \tag{2.19}$$

then there exists $\rho \in (0,\infty)$ such that $\widetilde{a} = a(\rho)$ and

$$\rho = \frac{\sigma' \|u - \sum_{k=0}^{n} a_k(\rho) v_k\|}{\sigma \|u' - \sum_{k=0}^{n} \overline{a_k(\rho)} v_k'\|}. \tag{2.20}$$

On the other hand, if there exists $\rho \in (0,\infty)$ *satisfying*

$$\|u - \sum_{k=0}^{n} a_k(\rho) v_k\| \cdot \|u' - \sum_{k=0}^{n} \overline{a_k(\rho)} v'_k\| \neq 0 \tag{2.21}$$

and (2.20)*, then* $a(\rho)$ *is a minimum point of* $F(a)$*.*

(b) *If* $\widetilde{a}$ *is a minimum point of* $F(a)$ *with* $\|u - \sum_{k=0}^{n} \widetilde{a}_k v_k\| = 0$*, then* $\widetilde{a} = a(0)$*.*

(c) *If* $\widetilde{a}$ *is a minimum point of* $F(a)$ *with* $\|u' - \sum_{k=0}^{n} \overline{\widetilde{a}}_k v'_k\| = 0$*, then* $\widetilde{a} = a(\infty)$*.*

Proof: If $\widetilde{a}$ is a minimum point of $F(a)$ that satisfies (2.19) then by differentiation and further manipulations we see that

$$\sum_{k=0}^{n} \widetilde{a}_k \left(\langle v_k, v_\ell \rangle + \rho \overline{\langle v'_k, v'_\ell \rangle} \right) = \langle u, v_\ell \rangle + \rho \overline{\langle u', v'_\ell \rangle}, \quad \ell = 0, \dots, n$$

where

$$\rho = \frac{\sigma' \|u - \sum_{k=0}^{n} \widetilde{a}_k v_k\|}{\sigma \|u' - \sum_{k=0}^{n} \overline{\widetilde{a}}_k v'_k\|}.$$

These equations can be expressed as the matrix equation

$$(G + \rho \overline{G}')\widetilde{a} = b + \rho \overline{b}'.$$

Since the elements $v_0, \dots, v_n$ are linearly independent and $0 < \rho < \infty$, the matrix $G + \rho \overline{G}'$ is invertible and the vector $\widetilde{a}$ is uniquely determined by $\widetilde{a} = a(\rho)$. Moreover, the constant ρ satisfies equation (2.20). Conversely, if $a(\rho)$ for some $\rho \in (0,\infty)$ is a vector that satisfies (2.20) and (2.21), then using the convexity of the function $F(a)$ we conclude that $a(\rho)$ minimizes $F(a)$. This proves part (a) of the lemma.

For part (b), observe that if $\|u - \sum_{k=0}^{n} \widetilde{a}_k v_k\| = 0$ then

$$\sum_{k=0}^{n} \widetilde{a}_k \langle v_k, v_\ell \rangle = \langle u, v_\ell \rangle, \quad \ell = 0, \dots, n$$

which leads to the equation $\widetilde{a} = G^{-1}(b) = a(0)$. The proof of part (c) is similar. ∎

It should be mentioned that $F(a)$ has a unique minimum point $\widetilde{a}$ such that (2.19) holds if and only if (2.20) has a unique solution for $\rho \in (0,\infty)$ satisfying (2.21). This can be deduced from part (a) of Lemma 2.6. To illustrate this observation, let us revisit the example preceding Lemma 2.6 that deals with the function $F(a_0) = |1 - a_0| + |a_0|$, $a_0 \in \mathbb{C}$. We have already seen that $F(a_0)$ does not have a unique minimum point. This is again reflected by the fact that (2.20) holds for all $\rho \in (0,\infty)$.

As another illustration of Lemma 2.6, consider the case when both $v_0, \ldots, v_n$ and $v'_0, \ldots, v'_n$ are orthonormal vectors with u and u' lying outside the linear spans of $\{v_0, \ldots, v_n\}$ and $\{v'_0, \ldots, v'_n\}$ respectively. Then $a(\rho)$ is given by

$$a(\rho) = \frac{1}{1+\rho}(b + \rho\overline{b}')$$

and (2.21) holds for every $\rho \in (0, \infty)$. Some computations show that (2.20) is equivalent to the quartic equation

$$(1+\rho)^2(\rho^2\beta' - \mu\beta) + (1-\mu)\rho^2\|b - \overline{b}'\|_2^2 = 0$$

where

$$\beta := \|u\|^2 - \|b\|_2^2 > 0, \qquad \beta' := \|u'\|^2 - \|b'\|_2^2 > 0, \qquad \mu := (\sigma'/\sigma)^2.$$

(The symbol $\|\cdot\|_2$ denotes the Euclidean norm on $\mathbb{C}^{n+1}$.) This quartic equation has a unique solution for $\rho \in (0, \infty)$ because the third derivative of its left hand side is always positive on $(0, \infty)$. Consequently, $F(a)$ has a unique minimum point $a(\widetilde{\rho})$ and its minimum value is given by

$$\left(\beta + \frac{\widetilde{\rho}^2}{(1+\widetilde{\rho})^2}\|b - \overline{b}'\|_2^2\right)^{1/2}\sigma + \left(\beta' + \frac{1}{(1+\widetilde{\rho})^2}\|b - \overline{b}'\|_2^2\right)^{1/2}\sigma'$$

where $\widetilde{\rho}$ is the unique solution of (2.20). For the special case of $\sigma = \sigma'$, $\widetilde{\rho}$ takes the value $\sqrt{\beta/\beta'}$.

In view of the complicated analysis arising from Lemma 2.6, we prefer to utilize Lemma 2.5 in our study. We shall now return to Lemma 2.5 to obtain variations of the uncertainty principle (2.2). If either $x, Bx, \ldots, B^n x$ or $x, B^*x, \ldots, (B^*)^n x$ are linearly independent elements in $\mathcal{H}$, then we conclude from Lemma 2.5 that

$$\begin{aligned}
&\min_{a_0,\ldots,a_n\in\mathbb{C}}\left\{\|Ax - \sum_{k=0}^{n} a_k B^k x\|^2 + \|A^*x - \sum_{k=0}^{n} \overline{a}_k (B^*)^k x\|^2\right\} \\
&= \frac{\mathcal{G}\left((Ax, A^*x), (x,x), (Bx, B^*x), \ldots, (B^n x, (B^*)^n x)\right)}{\mathcal{G}\left((x,x), (Bx, B^*x), \ldots, (B^n x, (B^*)^n x)\right)}. \qquad (2.22)
\end{aligned}$$

Using (2.22), we obtain the following theorem.

Theorem 2.7. *With A and B as in Theorem 2.1, let x be an element in $\mathcal{D}_n(A|B)$ where n is a positive integer such that either $x, Bx, \ldots, B^n x$ or $x, B^*x, \ldots, (B^*)^n x$ are linearly independent. Then*

$$\begin{aligned}
|\langle [A,B]x, x\rangle| \leq\ & \left\{\frac{\mathcal{G}\left((Ax, A^*x), (x,x), (Bx, B^*x), \ldots, (B^n x, (B^*)^n x)\right)}{\mathcal{G}\left((x,x), (Bx, B^*x), \ldots, (B^n x, (B^*)^n x)\right)}\right\}^{1/2} \\
& \times \left\{\Delta_x(B)^2 + \Delta_x(B^*)^2\right\}^{1/2}. \qquad (2.23)
\end{aligned}$$

Depending on the operators A and B and the element x, the upper bound in (2.23) can be significantly smaller than that in (2.2). For instance, consider the matrices A and B defined in (2.6) on the Hilbert space $\mathbb{C}^2$. For $x = (x_1, x_2)^T \in \mathbb{C}^2$ with x_1 and x_2 both nonzero, both sides of (2.23) with $n = 1$ equal $||x_1|^2 - |x_2|^2|$. This is strictly less than $(|x_1|^4 + |x_2|^4)/(|x_1|^2 + |x_2|^2)$ which is the right hand side of (2.2).

For the rest of this section, we shall deal with the case when A and B are both self-adjoint which is the setting of the original uncertainty inequality (1.5). In this case, it is worthwhile to note for all $x \in \mathcal{D}(A|B)$ that

$$\langle\, [A,B]x,\, x\,\rangle = -2i\,\mathrm{Im}\{\langle Ax, Bx\rangle\} \tag{2.24}$$

which is purely imaginary. This can also be seen by the formula

$$\|Ax + iBx\|^2 = \|Ax\|^2 + \|Bx\|^2 + i\langle\, [A,B]x,\, x\,\rangle. \tag{2.25}$$

The following result is an immediate consequence of Theorem 2.7.

Corollary 2.8. *With A and B as in Theorem* 1.1*, let x be an element in $\mathcal{D}_n(A|B)$ where n is a positive integer such that $x, Bx, \ldots, B^n x$ are linearly independent. Then*

$$|\langle\, [A,B]x,\, x\,\rangle| \le \sqrt{2}\left\{\frac{\mathcal{G}\left((Ax,Ax),(x,x),(Bx,Bx),\ldots,(B^n x, B^n x)\right)}{\mathcal{G}\left((x,x),(Bx,Bx),\ldots,(B^n x, B^n x)\right)}\right\}^{1/2} \Delta_x(B). \tag{2.26}$$

Since the right hand side of (2.26) is the value of the minimum of the expression

$$\|Ax - \sum_{k=0}^{n} a_k B^k x\|^2 + \|Ax - \sum_{k=0}^{n} \overline{a}_k B^k x\|^2 \tag{2.27}$$

over $a_0, \ldots, a_n \in \mathbb{C}$ and the right hand side of (1.5) is the value of the function in (2.27) at $a_0 = \dfrac{\langle Ax, x\rangle}{\|x\|^2}$ and $a_1 = \cdots = a_n = 0$, inequality (2.26) is an improvement of the original inequality (1.5). However, to achieve this gain we must assume that $x, Bx, \ldots, B^n x$ are linearly independent. Can one still achieve a *concrete* quantitative improvement if weaker conditions are assumed? In this connection, we observe the following fact. For $a_0, \ldots, a_n \in \mathbb{C}$, we let $a_k = d_k + ie_k$ where $d_k, e_k \in \mathbb{R}$ for $k = 0, \ldots, n$ and define

$$B_R x := \sum_{k=0}^{n} d_k B^k x, \qquad B_I x := \sum_{k=0}^{n} e_k B^k x.$$

Using (2.25), we see that (2.27) equals

$$2\|Ax - B_R x\|^2 + 2\|B_I x\|^2.$$

Hence, we conclude that the minimum of (2.27) over $a_0, \ldots, a_n \in \mathbb{C}$ is equal to its minimum over $a_0, \ldots, a_n \in \mathbb{R}$. As a result, it suffices to consider only real constants $a_0, \ldots, a_n$ in (2.16) so that (2.16) and (2.17) reduce to the inequality

$$|\langle\, [A,B]x,\, x\,\rangle| \le 2\|Ax - \sum_{k=0}^{n} a_k B^k x\|\, \Delta_x(B). \tag{2.28}$$

The best bound in (2.28) is obtained by minimizing the right hand side of this inequality over all $a_0, \ldots, a_n \in \mathbb{R}$. As in Proposition 2.4 when $\dim \mathcal{H} = m$, the value of the minimum remains constant for $n \ge m-1$. As a means to understand the benefit of this bound we provide conditions on B so that the bound for $n = m-1$ actually gives equality in (2.28). To this end, we note that the left hand side of this inequality is zero when $\mathcal{H}$ is a *real* Hilbert space and $x \in \mathcal{H}$. This is due to equation (2.24). The next proposition gives a sufficient condition on B so that the minimum of the upper bound in (2.28) over $a_0, \ldots, a_n \in \mathbb{R}$ is likewise zero.

Proposition 2.9. *Let B be a $m \times m$ symmetric real matrix that has distinct eigenvalues $\lambda_1, \ldots, \lambda_m$. Suppose that $v_1, \ldots, v_m$ are the corresponding orthonormal eigenvectors of B and x is an element in $\mathbb{R}^m$ of the form $x = \mu_1 v_1 + \cdots + \mu_m v_m$ where $\mu_1, \ldots, \mu_m$ are nonzero real constants. Then the elements $x, Bx, \ldots, B^{m-1}x$ are linearly independent and for any $m \times m$ self-adjoint real matrix A, we have that*

$$\delta_{m-1}(A|B)(x) = 2 \min_{a_0, \ldots, a_{m-1} \in \mathbb{R}} \|Ax - \sum_{k=0}^{m-1} a_k B^k x\| \Delta_x(B) = 0.$$

Proof: If $\sum_{k=0}^{m-1} \alpha_k B^k x = 0$ for some $\alpha_0, \ldots, \alpha_{m-1} \in \mathbb{R}$, then taking the inner product with $v_1, \ldots, v_m$ generates an invertible system of linear equations that yields the linear independence of $x, Bx, \ldots, B^{m-1}x$. Since Ax lies in $\mathbb{R}^m$, it follows that $\delta_{m-1}(A|B)(x) = 0$. ■

In general, for any self-adjoint linear operators A and B in $\mathcal{H}$ and $x \in \mathcal{H}$, it is not always true that $\delta_n(A|B)(x)$ converges to the quantity $|\langle [A, B]x, x \rangle|$ as n tends to infinity. For instance, consider a complex Hilbert space $\mathcal{H}$ whose dimension is at least three. Let x, y, z be three linearly independent elements of $\mathcal{H}$ with $\langle x, y \rangle \neq 0$. Suppose that A and B are projection operators given by

$$Aw := \frac{\langle w, y \rangle}{\|y\|^2} y, \qquad Bw := \frac{\langle w, z \rangle}{\|z\|^2} z$$

where $w \in \mathcal{H}$. Then for $n \geq 2$,

$$\delta_n(A|B)(x) = \delta_1(A|B)(x) > 0.$$

On the other hand, the quantity $|\langle [A, B]x, x \rangle|$ sometimes equals zero. For example, when $\langle Ax, Bx \rangle$ takes a real value and in particular, when $\mathcal{H}$ is a real Hilbert space.

It is of interest to provide general examples in which inequality (2.28) provides, after minimizing over $a_0, \ldots, a_n \in \mathbb{R}$, bounds that successively *decrease* in n to zero at $n = m - 1$ (and hence to the lower bound in (2.28)). We do this in the next observation.

Proposition 2.10. *Let B, $\lambda_1, \ldots, \lambda_m$ and x be as in Proposition 2.9. If $A = f(B)$ for a real-valued function f satisfying the condition that $f', \ldots, f^{(m-1)}$ are positive on an interval $[c, d]$ that contains $\lambda_1, \ldots, \lambda_m$, then*

$$\langle [A, B]x, x \rangle = 0 = \delta_{m-1}(A|B)(x) < \delta_{m-2}(A|B)(x) < \cdots < \delta_0(A|B)(x). \tag{2.29}$$

Proof: For $0 \leq n \leq m - 1$, by writing $p_n(t) := \sum_{k=0}^{n} a_k t^k$, we observe that

$$\min_{a_0, \ldots, a_n \in \mathbb{R}} \|Ax - \sum_{k=0}^{n} a_k B^k x\|^2 = \min_{p_n \in \mathcal{P}_n} \left\{ \sum_{\ell=1}^{m} \mu_\ell^2 \left(f(\lambda_\ell) - p_n(\lambda_\ell) \right)^2 \right\} \tag{2.30}$$

where $\mathcal{P}_n$ is the space of all real polynomials of degree n or less. The unique solution $\widetilde{a} = (\widetilde{a}_0, \dots, \widetilde{a}_n)^T$ to the minimization problem (2.30) is given by the matrix equation $L\widetilde{a} = b$ where

$$L := \left(\sum_{\ell=1}^{m} \lambda_\ell^k \mu_\ell^2 \lambda_\ell^j\right)_{j,k=0}^{n}, \qquad b := \left(\sum_{\ell=1}^{m} f(\lambda_\ell)\mu_\ell^2\lambda_\ell^0, \dots, \sum_{\ell=1}^{m} f(\lambda_\ell)\mu_\ell^2\lambda_\ell^n\right)^T.$$

Using the Cauchy-Binet formula, see for instance [14], the leading coefficient $\widetilde{a}_n$ of $\widetilde{p}_n$ is given by

$$\widetilde{a}_n = \frac{\sum \cdots \sum_{\ell_1<\ell_2<\cdots<\ell_{n+1}} \mu_{\ell_1}^2 \cdots \mu_{\ell_{n+1}}^2 V(\lambda_{\ell_1}, \dots, \lambda_{\ell_{n+1}})^2 \, [\lambda_{\ell_1}, \dots, \lambda_{\ell_{n+1}}]f}{\sum \cdots \sum_{\ell_1<\ell_2<\cdots<\ell_{n+1}} \mu_{\ell_1}^2 \cdots \mu_{\ell_{n+1}}^2 V(\lambda_{\ell_1}, \dots, \lambda_{\ell_{n+1}})^2}$$

where $[\lambda_{\ell_1}, \dots, \lambda_{\ell_{n+1}}]f$ denotes the $(n+1)$th divided difference of f at $\lambda_{\ell_1}, \dots, \lambda_{\ell_{n+1}}$ and $V(\lambda_{\ell_1}, \dots, \lambda_{\ell_{n+1}})$ is the determinant of the $(n+1)\times(n+1)$ Vandermonde matrix $(\lambda_{\ell_j}^{k-1})_{j,k=1}^{n+1}$. For $1 \le n \le m-1$, the assumption that $f^{(n)} > 0$ on $[c, d]$ assures that all the divided difference $[\lambda_{\ell_1}, \dots, \lambda_{\ell_{n+1}}]f$ appearing in the numerator of the above expression are positive. Hence, $\widetilde{a}_n > 0$ and $\widetilde{p}_n$ is of exact degree n. Consequently, for $0 \le n_1 < n_2 \le m-1$, it is not possible for $\delta_{n_1}(A|B)(x)$ and $\delta_{n_2}(A|B)(x)$ to be the same. This yields the inequalities in (2.29). ∎

By the same argument, the above result can be extended to (complex) normal matrices.

Proposition 2.11. *Let B be a $m \times m$ normal complex matrix that has distinct eigenvalues $\lambda_1, \dots, \lambda_m$. Suppose that $v_1, \dots, v_m$ are the corresponding orthonormal eigenvectors of B and x is an element in $\mathbb{C}^m$ of the form $x = \mu_1 v_1 + \cdots + \mu_m v_m$ where $\mu_1, \dots, \mu_m$ are nonzero complex constants. Let $\mathcal{V}$ be a compact convex set in $\mathbb{C}$ whose interior contains $\lambda_1, \dots, \lambda_m$ and f be a complex-valued function that is continuous on $\mathcal{V}$ and analytic in the interior of $\mathcal{V}$. If $A = f(B)$ and for $1 \le n \le m-1$ the convex hull of $\{f^{(n)}(z) : z \in \mathcal{V}\}$ does not contain the origin then* (2.29) *holds.*

A version of this result may be formulated for bounded normal operators in Hilbert spaces using the spectral decomposition of the operators. In this regard, we first note that the Weierstrass approximation theorem implies for any continuous function f on the spectrum of B with $A = f(B)$, it follows that

$$\lim_{n\to\infty} \delta_n(A|B)(x) = \langle\, [A, B]x,\, x\,\rangle = 0.$$

Likewise, when f satisfies the conditions of Proposition 2.11 for $m = \infty$ and $\mathcal{V}$ contains the spectrum of B in its interior, the sequence $\delta_n(A|B)(x)$, $n \ge 0$, is strictly decreasing.

Let us now return to the issue of using (2.28) to obtain alternative bounds for the quantity $\langle\, [A, B]x,\, x\,\rangle$ under conditions that are weaker than the linear independence of $x, Bx, \dots, B^n x$. Our results are based on the following observation. Let x be a nonzero element in $\mathcal{D}(A)$ and $v_1, \dots, v_n$ be nonzero elements in $\mathcal{H}$. We define elements $u_1, \dots, u_n$ in

$\mathcal{H}$ recursively by the formulas

$$\begin{aligned} u_1 &:= Ax - \frac{\langle Ax, x\rangle}{\|x\|^2}x, \\ u_{k+1} &:= u_k - \frac{\mathrm{Re}\{\langle u_k, v_k\rangle\}}{\|v_k\|^2}v_k, \quad k = 1, \dots, n-1. \end{aligned} \tag{2.31}$$

It turns out that for $k = 1, \dots, n-1$,

$$\|u_{k+1}\|^2 = \min_{a\in\mathbb{R}} \|u_k - av_k\|^2 = \|u_k\|^2 - \frac{(\mathrm{Re}\{\langle u_k, v_k\rangle\})^2}{\|v_k\|^2}$$

and the minimum is attained at $a = \dfrac{\mathrm{Re}\{\langle u_k, v_k\rangle\}}{\|v_k\|^2}$. Furthermore,

$$\left\| u_1 - \sum_{k=1}^{n} \frac{\mathrm{Re}\{\langle u_k, v_k\rangle\}}{\|v_k\|^2} v_k \right\|^2 = \Delta_x(A)^2 - \sum_{k=1}^{n} \frac{(\mathrm{Re}\{\langle u_k, v_k\rangle\})^2}{\|v_k\|^2}. \tag{2.32}$$

In the above iterative procedure, we seek a real multiple of v_k that optimally approximates u_k in the sense of minimizing the quantity $\|u_k - av_k\|$ over all *real* values of a. We continue the process with the residue of the approximation u_{k+1} and the element v_{k+1}. In view of (2.28), we set $v_k := B^k x$ for $k = 1, \dots, n$ to benefit from (2.32).

Theorem 2.12. *With A and B as in Theorem 1.1, if x is an element in $\mathcal{D}_n(A|B)$ where n is a positive integer such that $x \notin \mathrm{Ker}(B^n)$ and $u_1, \dots, u_n$ are defined in (2.31) with $v_k := B^k x$ for $k = 1, \dots, n$, then*

$$|\langle [A, B]x, x\rangle| \le 2 \left(\Delta_x(A)^2 - \sum_{k=1}^{n} \frac{(\mathrm{Re}\{\langle u_k, B^k x\rangle\})^2}{\|B^k x\|^2} \right)^{1/2} \Delta_x(B). \tag{2.33}$$

Proof: Since $x \notin \mathrm{Ker}(B^n)$, all the elements $x, Bx, \dots, B^n x$ are nonzero. With $u_1, \dots, u_n$ as given in (2.31), we choose $a_0 := \dfrac{\langle Ax, x\rangle}{\|x\|^2}$ and

$$a_k := \frac{\mathrm{Re}\{\langle u_k, B^k x\rangle\}}{\|B^k x\|^2}, \quad k = 1, \dots, n.$$

Then by (2.31) and (2.32), (2.33) follows from (2.28). ∎

In our next result, instead of considering $x \notin \mathrm{Ker}(B^n)$ for some $n \ge 1$, we make the stronger assumption that x is nonzero and not an eigenvector of any of the operators $B, \dots, B^n$. Consequently, we are assured that

$$v_k := B^k x - \frac{\langle B^k x, x\rangle}{\|x\|^2} x \ne 0, \quad k = 1, \dots, n \tag{2.34}$$

and the iterative procedure (2.31) is again applicable. However, a preliminary manipulation on the right hand side of (2.28) has to be made to suit this choice of $v_1, \dots, v_n$.

Theorem 2.13. *With A and B as in Theorem 1.1, suppose that x is a nonzero element in $\mathcal{D}_n(A|B)$ where n is a positive integer such that x is not an eigenvector of any of $B, \ldots, B^n$. We define elements $u_1, \ldots, u_n$ in $\mathcal{H}$ by (2.31) with $v_1, \ldots, v_n$ as in (2.34). Then we conclude that*

$$|\langle [A,B]x,\, x\,\rangle| \le 2\left(\Delta_x(A)^2 - \sum_{k=1}^{n} \frac{(\mathrm{Re}\{\langle u_k, v_k\rangle\})^2}{\|v_k\|^2}\right)^{1/2} \Delta_x(B).$$

Proof: First, we fix $a_1, \ldots, a_n \in \mathbb{R}$ and choose

$$a_0 := \frac{\langle Ax - \sum_{k=1}^{n} a_k B^k x, x\rangle}{\|x\|^2}$$

to minimize the right hand side of (2.28) and obtain the inequality

$$|\langle [A,B]x,\, x\,\rangle| \le 2\left\| \left(Ax - \frac{\langle Ax, x\rangle}{\|x\|^2}x\right) - \sum_{k=1}^{n} a_k \left(B^k x - \frac{\langle B^k x, x\rangle}{\|x\|^2}x\right)\right\| \Delta_x(B).$$

With u_k and v_k as in (2.31) and (2.34) respectively, choose $a_k := \dfrac{\mathrm{Re}\{\langle u_k, v_k\rangle\}}{\|v_k\|^2}$ for $k = 1, \ldots, n$ and apply (2.32) to complete the proof. ∎

As a further remark, we note that for any nonzero $x \in \mathcal{D}(A|B)$, it is possible to obtain an improvement of the upper bound in (1.5) *without* imposing any additional condition on x.

Corollary 2.14. *With A and B as in Theorem 1.1, for any nonzero $x \in \mathcal{D}(A|B)$ we have that*

$$\begin{aligned} &|\langle [A,B]x,\, x\,\rangle| && (2.35)\\ \le\ & 2\left(\Delta_x(A)^2\Delta_x(B)^2 - \left(\mathrm{Re}\left\{\left\langle Ax - \frac{\langle Ax, x\rangle}{\|x\|^2}x, Bx - \frac{\langle Bx, x\rangle}{\|x\|^2}x\right\rangle\right\}\right)^2\right)^{1/2}. \end{aligned}$$

Proof: If x is an eigenvector of B, then $\Delta_x(B) = 0$ which forces both sides of (2.35) to be zero. If x is not an eigenvector of B, then Theorem 2.13 yields the result. ∎

3. Equality

We shall now address the issue of equality in the uncertainty inequalities of the last section. The general inequality (2.8) is central in all these inequalities. The following theorem gives necessary and sufficient conditions for equality to hold in (2.8).

Theorem 3.1. *Let A, B, U, V be linear operators with domain and range in a Hilbert space $\mathcal{H}$ and x be a nonzero element in $\mathcal{D}(A|B) \cap \mathcal{D}(A|V) \cap \mathcal{D}(B|U) \cap \mathcal{D}(U|V)$ that satisfies (2.7).*

(a) *If x is not in the kernel of any of the operators $A - U$, $A^* - U^*$, $B - V$, $B^* - V^*$, then equality holds in (2.8) if and only if there exist nonzero constants $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{C}$ where $\dfrac{\alpha_2\beta_2}{\alpha_1\beta_1} < 0$ such that x lies in the kernel of the operators*

$$S_{\alpha_1,\alpha_2} := \alpha_1(A - U) + \alpha_2(B^* - V^*), \qquad T_{\beta_1,\beta_2} := \beta_1(A^* - U^*) + \beta_2(B - V). \qquad (3.1)$$

(b) *If x is in the kernel of at least one of the operators $A-U$, A^*-U^*, $B-V$, B^*-V^*, then equality holds in* (2.8) *if and only if there exist constants $\alpha_1,\alpha_2,\beta_1,\beta_2\in\mathbb{C}$ where $|\alpha_1|+|\alpha_2|\neq 0$ and $|\beta_1|+|\beta_2|\neq 0$ such that x lies in the kernel of the operators S_{α_1,α_2} and T_{β_1,β_2} in* (3.1).

Proof: By examining the derivation of (2.8), we see that equality holds in (2.8) if and only if

$$\begin{aligned} & |\langle Bx-Vx, A^*x-U^*x\rangle - \langle Ax-Ux, B^*x-V^*x\rangle| \\ = \; & |\langle Bx-Vx, A^*x-U^*x\rangle| + |\langle Ax-Ux, B^*x-V^*x\rangle| \end{aligned} \tag{3.2}$$

and there exist constants $\alpha_1,\alpha_2,\beta_1,\beta_2\in\mathbb{C}$ where $|\alpha_1|+|\alpha_2|\neq 0$ and $|\beta_1|+|\beta_2|\neq 0$ such that

$$\alpha_1(Ax-Ux)+\alpha_2(B^*x-V^*x)=0, \qquad \beta_1(A^*x-U^*x)+\beta_2(Bx-Vx)=0. \tag{3.3}$$

Observe that (3.3) holds if and only if x lies in the kernel of the operators S_{α_1,α_2} and T_{β_1,β_2} in (3.1).

First, we prove part (a) of the result. To this end, observe that if x is not in the kernel of the operators $A-U$, A^*-U^*, $B-V$, B^*-V^*, then all of the quantities $\|Ax-Ux\|^2$, $\|A^*x-U^*x\|^2$, $\|Bx-Vx\|^2$ and $\|B^*x-V^*x\|^2$ are positive. As observed above, equality in (2.8) implies that (3.2) and (3.3) are satisfied for some constants $\alpha_1,\alpha_2,\beta_1,\beta_2\in\mathbb{C}$ where $|\alpha_1|+|\alpha_2|\neq 0$ and $|\beta_1|+|\beta_2|\neq 0$. We note that all the constants α_1, α_2, β_1, β_2 are nonzero. Indeed, for instance, if $\alpha_1=0$ then by (3.3) we conclude that

$$\alpha_2(B^*x-V^*x)=0.$$

Since $B^*x-V^*x\neq 0$, this implies that $\alpha_2=0$, a contradiction. Moreover, because $\alpha_1,\beta_1\neq 0$, (3.3) can be rewritten as

$$Ax-Ux=-\frac{\alpha_2}{\alpha_1}(B^*x-V^*x), \qquad A^*x-U^*x=-\frac{\beta_2}{\beta_1}(Bx-Vx). \tag{3.4}$$

Substituting (3.4) into (3.2), we obtain that

$$\left|\overline{\frac{\beta_2}{\beta_1}}\|Bx-Vx\|^2-\frac{\alpha_2}{\alpha_1}\|B^*x-V^*x\|^2\right| = \left|\frac{\beta_2}{\beta_1}\right|\|Bx-Vx\|^2+\left|\frac{\alpha_2}{\alpha_1}\right|\|B^*x-V^*x\|^2$$

and using the fact that $\|Bx-Vx\|^2\|B^*x-V^*x\|^2>0$ it follows that $\dfrac{\alpha_2\beta_2}{\alpha_1\beta_1}<0$.

For the converse direction, one can show that (3.2) and (3.3) are satisfied and hence equality holds in (2.8).

For part (b) of the result, suppose that x is in the kernel of at least one of the operators $A-U$, A^*-U^*, $B-V$, B^*-V^*. Then (3.2) is satisfied automatically. This gives the desired equivalent condition for equality in (2.8). ■

Corollary 3.2. *Let A and B be any self-adjoint linear operators with domain and range in the same complex Hilbert space $\mathcal{H}$ and x be a nonzero element in $\mathcal{D}(A|B)$. Either x is an eigenvector of A or B and equality holds in* (1.5) *or equality holds in* (1.5) *if and* *only if there exists a nonzero constant* $\mu \in \mathbb{R}$ *such that* x *is an eigenvector of the operator* $S_{1,-i\mu} := A - i\mu B$.

Proof: With the choice of $U := aI$ and $V := bI$ where $a := \dfrac{\langle Ax, x\rangle}{\|x\|^2}$ and $b := \dfrac{\langle Bx, x\rangle}{\|x\|^2}$ and the self-adjointness of A and B, the first assertion follows from Theorem 3.1 or the fact that x is an eigenvector of A or B if and only if $\Delta_x(A)\Delta_x(B) = 0$. The case when x is neither an eigenvector of A nor B can be handled by following the proof of Theorem 3.1. ∎

Let us apply Theorem 3.1 to the Hilbert space $\mathcal{L}_2[0, 2\pi)$ and the operators M_w, D, aI and M_h in (2.13). For the analysis of equality in (2.13), the following fact is central. If $f \in \mathcal{W}_2^1[0, 2\pi)\backslash\{0\}$ and $g \in \mathcal{L}_\infty[0, 2\pi)$ satisfy the differential equation

$$gf + f' = 0, \tag{3.5}$$

then there exists $n \in \mathbb{Z}$ such that

$$\int_{[0,2\pi)} g(t)\, dt = 2\pi i n \tag{3.6}$$

and f is everywhere given by the formula

$$f(t) = c \exp\left\{-\int_0^t g(s)\, ds\right\}, \quad t \in [0, 2\pi) \tag{3.7}$$

where $c \in \mathbb{C}\backslash\{0\}$. On the other hand, if $g \in \mathcal{L}_\infty[0, 2\pi)$ satisfies (3.6) for some $n \in \mathbb{Z}$, then any f of the form (3.7) is in $\mathcal{W}_2^1[0, 2\pi)$ and satisfies (3.5).

Proposition 3.3. *Suppose that* M_w, D, aI, M_h *are the operators in* (2.13) *and equality holds in* (2.13) *for some* $f \in \mathcal{W}_2^1[0, 2\pi)$ *not in the kernel of* $M_w - aI$, $M_{\overline{w}} - \overline{a}I$, $D - M_h$, $D - M_{\overline{h}}$. *Then there exists a constant* $b \in \mathbb{C}\backslash\{0\}$ *and real-valued* $v \in \mathcal{W}_\infty^1[0, 2\pi)\backslash\{0\}$ *such that*

$$w = a + bv. \tag{3.8}$$

Moreover, if h *is not real-valued, then* w *can only be of the form*

$$w = a + b\,\mathrm{Im}\{h\}. \tag{3.9}$$

Proof: Using part (a) of Theorem 3.1 and (3.7), there exist $\alpha, \beta \in \mathbb{C}\backslash\{0\}$ with $\alpha\beta < 0$ such that for every $t \in [0, 2\pi)$,

$$c_1 \exp\left\{-\int_0^t \left(\frac{w(s) - a}{i\alpha} - \frac{\overline{h}(s)}{i}\right) ds\right\} = c_2 \exp\left\{-\int_0^t \left(\frac{\overline{w}(s) - \overline{a}}{i\beta} - \frac{h(s)}{i}\right) ds\right\}$$

where $c_1, c_2 \in \mathbb{C}\backslash\{0\}$. This implies that w is of the form (3.8). For h which is not real-valued, we can further conclude that w takes the form (3.9). ∎

In [27], it was shown for inequality (1.6) that the only functions in $\mathcal{L}_2[0, 2\pi)$ which attain equality are of the form ce^{ikt}, $c \in \mathbb{C}$ and $k \in \mathbb{Z}$. This can be obtained as a consequence of Proposition 3.3. Indeed, if $f \in \mathcal{L}_2[0, 2\pi)$ is not such a function but attains equality in (1.6), then by considering $a := \dfrac{\langle Sf, f\rangle_{2\pi}}{\|f\|_{2\pi}^2}$ and $h := \dfrac{\langle Df, f\rangle_{2\pi}}{\|f\|_{2\pi}^2}$, we conclude from Proposition 3.3

that $e^{it} = a + bv(t)$ a.e. on $[0, 2\pi)$ for some $b \in \mathbb{C}\backslash\{0\}$ and real-valued $v \in \mathcal{W}^1_\infty[0, 2\pi)\backslash\{0\}$. By differentiating this equation with respect to t, we obtain a contradiction.

For the operators M_w, D, aI, M_h as in Proposition 3.3 with w of the form (3.8) or (3.9), we can find functions $f \in \mathcal{L}_2[0, 2\pi)$ such that equality holds in (2.13). Indeed, the following proposition characterizes the extremal functions for w of the form (3.8).

Proposition 3.4. *Let M_w, D, aI, M_h be as in Proposition 3.3 where h is real-valued, $\int_{[0,2\pi)} h(t)\,dt = 2\pi n$ for some $n \in \mathbb{Z}$ and w is of the form (3.8) with a real-valued $v \in \mathcal{W}^1_\infty[0, 2\pi)\backslash\{0\}$ that satisfies $\int_{[0,2\pi)} v(t)\,dt = 0$. For $\mu \in \mathbb{R}\backslash\{0\}$, define*

$$f_\mu(t) := \exp\left\{-\int_0^t \left(\frac{v(s)}{\mu} + ih(s)\right) ds\right\}, \quad t \in [0, 2\pi). \tag{3.10}$$

Then equality holds in (2.13) for some $f \in \mathcal{W}^1_2[0, 2\pi)$ not in the kernel of $M_w - aI$, $D - M_h$ if and only if f is a constant multiple of f_μ for some $\mu \in \mathbb{R}\backslash\{0\}$.

Proof: The result is a consequence of Theorem 3.1 and the basic principle described in (3.5)–(3.7). ∎

Taking $a = 0$ in Proposition 3.4 yields the specialized case of $w = v$ in which f_μ, $\mu \in \mathbb{R}\backslash\{0\}$, of (3.10) attains equality in (2.13). The extremal functions form a rich class of functions that yield equality in inequality (2.14). Indeed, consider a function f_μ of the form (3.10) for some real-valued function h. We recall that the right hand side of (2.14) is obtained by minimizing the quantity $\gamma(\widetilde{h})$ in (2.15) over all 2π-periodic functions $\widetilde{h}$ satisfying $\widetilde{h} f_\mu \in \mathcal{L}_2[0, 2\pi)$ and the minimum is attained at a real-valued function $\widetilde{h}$. Thus the value of the right hand side of (2.14) is at most that of the right hand side of (2.13) which already equals to $|\langle M_{w'} f_\mu, f_\mu\rangle_{2\pi}|$. In fact, in this case a direct computation reveals that both sides of (2.14) are independent of the function h in (3.10).

For w of the form (3.9), we can also use Theorem 3.1 and the idea described in (3.5)–(3.7) to identify a family of functions that attain equality in (2.13).

Let us turn our attention to the possibility of *asymptotic* equality. We begin with a remark that under the hypothesis of Corollary 2.2 the condition of asymptotic equality in inequality (2.4) can be given. We say (2.4) is asymptotically equal provided that

$$\sup\left\{\frac{|\langle\, [A, B]x,\, x\,\rangle|}{\|Ax\|\,\|Bx\|} : x \in \mathcal{D}(A|B),\ \|Ax\|\,\|Bx\| > 0\right\} = 2.$$

Following carefully the proof of Theorem 2.1 we see that (2.4) is asymptotically optimal when A and B are self-adjoint if and only if there is a sequence $\{x_n\}_{n\geq 1}$ in $\mathcal{D}(A|B)$ such that $\lim_{n\to\infty} \tau(A, B|x_n) = \pm i$ where

$$\tau(A, B|x) := \frac{\langle Ax, Bx\rangle}{\|Ax\|\,\|Bx\|}, \quad x \in \mathcal{D}(A|B), \quad \|Ax\|\,\|Bx\| > 0.$$

For normal operators, the case of asymptotic optimality is not as simple. We cover a part of this situation in the next result.

Theorem 3.5. *Let A_1, A_2, B be self-adjoint linear operators with domain and range in a Hilbert space $\mathcal{H}$ such that A_1 and A_2 commute and set $A := A_1 + iA_2$. Then inequality* (2.4) *is asymptotically equal if and only if there is a sequence $\{x_n\}_{n\geq 1}$ in $\mathcal{D}(A|B)$ such that $\delta = (0,1)$ and $\lim_{n\to\infty} \tau(A_2, B|x_n) = \pm i$ or $\delta = (1,0)$ and $\lim_{n\to\infty} \tau(A_1, B|x_n) = \pm i$ or $\delta_1\delta_2 > 0$, $\delta_1^2 + \delta_2^2 = 1$ and $\lim_{n\to\infty} \tau(A_\ell, B|x_n) = \pm i$ for $\ell = 1,2$ where*

$$\delta_\ell := \lim_{n\to\infty} \frac{\|A_\ell x_n\|}{\|Ax_n\|}, \quad \ell = 1,2$$

and $\delta := (\delta_1, \delta_2)$.

Proof: The proof follows from the formula

$$\frac{\langle\, [A,B]x,\, x\,\rangle}{\|Ax\|\,\|Bx\|} = 2\left\{\frac{\|A_2x\|}{\|Ax\|}\operatorname{Im}\{\tau(A_2,B|x)\} - i\frac{\|A_1x\|}{\|Ax\|}\operatorname{Im}\{\tau(A_1,B|x)\}\right\}.$$

■

We now return to the Hilbert space $\mathcal{L}_2[0,2\pi)$ and contrast the difference between the restrictive conditions on the function $w \in \mathcal{L}_\infty[0,2\pi)$ in Proposition 3.4 for which equality occurs in (2.11) with the fact that asymptotic equality occurs for a large class of weight functions. To make this point, we are content with the following setup. We introduce two classes of functions. For every $\beta > 0$ we denote by $\mathcal{S}_\beta(\mathbb{R})$ the set of all functions defined on $\mathbb{R}$ such that f'' is absolutely continuous on $\mathbb{R}$ with $f^{(j)} \in \mathcal{L}_2(\mathbb{R})$, $j = 0,1,2,3$ and for which there exists a constant $c > 0$ such that $|f(t)| \leq c(1+|t|)^{-\beta}$, $t \in \mathbb{R}$. The second class we require is defined to be

$$\mathcal{T}[0,2\pi) := \{w : w \in \mathcal{L}_\infty[0,2\pi),\ \sum_{k\in\mathbb{Z}} |k|^3 |\widehat{w}(k)| < \infty\}$$

where

$$\widehat{w}(k) := \frac{1}{2\pi}\int_{[0,2\pi)} w(t)e^{-ikt}\,dt, \quad k \in \mathbb{Z}$$

are the Fourier coefficients of $w \in \mathcal{L}_\infty[0,2\pi)$.

Theorem 3.6. *Suppose $s, Ms \in \mathcal{S}_\beta(\mathbb{R})$ for some $\beta > 1/2$ with s nontrivial, $w, w', |w|^2 \in \mathcal{T}[0,2\pi)$ with $w'(0) \neq 0$ and s_ρ is defined by*

$$s_\rho(t) := \sum_{j\in\mathbb{Z}} s(\rho j)e^{ijt}, \quad t \in [0,2\pi) \tag{3.11}$$

for $\rho > 0$. Then there is a constant $c > 0$ such that for all $\rho > 0$,

$$|\nu_{2\pi}(s_\rho; w) - \nu_{\mathbb{R}}(s)| \leq c\rho$$

where the quantities $\nu_{\mathbb{R}}(s)$ and $\nu_{2\pi}(s_\rho; w)$ are defined as in (1.3) *and* (2.12) *respectively.*

A consequence of this fact is that

$$\lim_{\rho\to 0} \nu_{2\pi}(s_\rho; w) = \frac{1}{4}$$

when $s(t) = e^{-t^2}$, $t \in \mathbb{R}$ and $w, w', |w|^2 \in \mathcal{T}[0, 2\pi)$ with $w'(0) \neq 0$. That is, inequality (2.11) is asymptotically sharp for this class of weight functions.

The proof of Theorem 3.6 is a consequence of the following lemma.

Lemma 3.7. *Suppose $s \in \mathcal{S}_\beta(\mathbb{R})$ for some $\beta > 1/2$ and $w \in \mathcal{T}[0, 2\pi)$. Then there is a constant $c > 0$ such that for $\rho > 0$, s_ρ defined by (3.11) satisfies*

$$\left| \rho \langle M_w s_\rho, s_\rho \rangle_{2\pi} - w(0)\|s\|_{\mathbb{R}}^2 - iw'(0)\rho\langle s', s\rangle_{\mathbb{R}} - \frac{1}{2} w''(0)\rho^2 \|s'\|_{\mathbb{R}}^2 \right| \leq c\rho^3.$$

Proof: Under the hypothesis of the lemma, we have that

$$\langle M_w s_\rho, s_\rho \rangle_{2\pi} = \sum_{k\in\mathbb{Z}} \sum_{j\in\mathbb{Z}} \overline{s(\rho j)} s(\rho(j-k)) \widehat{w}(k). \tag{3.12}$$

We claim that there exists a constant $d > 0$ such that for all $\rho > 0$ and $k \in \mathbb{Z}$ there holds

$$\left| \rho \sum_{j\in\mathbb{Z}} \overline{s(\rho j)} s(\rho(j-k)) - \|s\|_{\mathbb{R}}^2 + k\rho\langle s', s\rangle_{\mathbb{R}} + \frac{(k\rho)^2}{2} \|s'\|_{\mathbb{R}}^2 \right| \leq d(1 + |k|^3)\rho^3. \tag{3.13}$$

Before proving this inequality, let us observe that the lemma follows from it. To see this, we use the formulas

$$w(0) = \sum_{k\in\mathbb{Z}} \widehat{w}(k), \qquad w''(0) = -\sum_{k\in\mathbb{Z}} k^2 \widehat{w}(k),$$

multiply both sides of (3.13) by $|\widehat{w}(k)|$, sum over all k in $\mathbb{Z}$ and apply formula (3.12).

The proof of inequality (3.13) relies on the Poisson summation formula applied to the function $v_{k,\rho} := \overline{s(\rho\,\cdot)} s(\rho(\cdot - k))$. Indeed, since $s \in \mathcal{S}_\beta(\mathbb{R})$, we can find a positive constant $c(k, \rho)$ for which the inequality

$$|v_{k,\rho}(t)| \leq c(k, \rho)(1 + |t|)^{-2\beta}, \quad t \in \mathbb{R}$$

holds. Furthermore, there exists a constant $c_0 > 0$ such that for all $\rho > 0$, $k \in \mathbb{Z}$,

$$|\widehat{v}_{k,\rho}(t)| \leq c_0 \rho^2 |t|^{-3}, \quad t \in \mathbb{R}\backslash\{0\}. \tag{3.14}$$

Hence by the Poisson summation formula, we have that

$$\sum_{j\in\mathbb{Z}} \overline{s(\rho j)} s(\rho(j-k)) = \sum_{j\in\mathbb{Z}} \widehat{v}_{k,\rho}(2\pi j).$$

Consequently, we obtain from (3.14) another constant $c_1 > 0$ such that for all $\rho > 0$, $k \in \mathbb{Z}$,

$$\left| \rho \sum_{j\in\mathbb{Z}} \overline{s(\rho j)} s(\rho(j-k)) - \int_{\mathbb{R}} \overline{s(t)} s(t - \rho k)\, dt \right| \leq c_1 \rho^3.$$

Again using the fact that $s \in \mathcal{S}_\beta(\mathbb{R})$ and applying Taylor's formula with remainder to the real and imaginary parts of the function $\Phi := \int_{\mathbb{R}} \overline{s(t)} s(t - \cdot)\, dt$, we conclude that there exists a constant $c_2 > 0$ such that for all $\rho > 0$, $k \in \mathbb{Z}$,

$$\left| \int_{\mathbb{R}} \overline{s(t)} s(t - \rho k)\, dt - \|s\|_{\mathbb{R}}^2 + k\rho\langle s', s\rangle_{\mathbb{R}} + \frac{(k\rho)^2}{2} \|s'\|_{\mathbb{R}}^2 \right| \leq c_2 |k|^3 \rho^3.$$

Combining this inequality with the one above gives (3.13). ■

Proof of Theorem 3.6: We apply Lemma 3.7 successively to the pair of functions (i) w, s_ρ, (ii) w', s_ρ, (iii) $|w|^2$, s_ρ, (iv) 1, s_ρ, (v) 1, $(Ms)_\rho$. The reason for applying the lemma to the last pair of functions is based upon the observation that

$$\|Ds_\rho\|_{2\pi} = \rho^{-1}\|(Ms)_\rho\|_{2\pi}.$$

In addition, we use the Poisson summation formula on the function $u_\rho := \cdot\,|s(\rho\,\cdot)|^2$ to conclude that there exists a constant $c > 0$ such that for all $\rho > 0$,

$$\left|\rho^2\langle Ds_\rho, s_\rho\rangle_{2\pi} + \langle Ms, s\rangle_{\mathbb{R}}\right| \le c\rho^3.$$

The theorem follows from all the estimates obtained. ■

4. Region

The basic uncertainty inequality (2.2) as given by Theorem 2.1 involves *eight* quadratic forms, namely $\langle\, [A,B]x,\, x\,\rangle$, $\|Ax\|^2$, $\langle Ax, x\rangle$, $\|A^*x\|^2$, $\|Bx\|^2$, $\langle Bx, x\rangle$, $\|B^*x\|^2$ and $\|x\|^2$. Inequality (2.2) represents only one constraint on these eight quantities. It would be of interest to describe the totality of all eight tuples arising from vectors in $\mathcal{H}$. Our attempt to understand this problem is based on the following result from Proposition 2.1 in [22].

Theorem 4.1. *Let $\mathcal{X}$ be a linear space over the reals and $\langle\cdot,\cdot\rangle_i$, $i = 0,1,2$, be quadratic functionals defined on $\mathcal{X}$ such that $\langle\cdot,\cdot\rangle_0$ is positive definite. If $\dim\mathcal{X} \ge 3$ then the set*

$$\mathcal{M} := \{(\langle x,x\rangle_1, \langle x,x\rangle_2) : \langle x,x\rangle_0 = 1,\, x \in \mathcal{X}\}$$

is convex.

Proposition 2.1 in [22] requires that the quadratic functionals $\langle\cdot,\cdot\rangle_i$, $i = 1,2$ are nonnegative on $\mathcal{X}$. However, the proof does not use this restriction. For $\dim\mathcal{X} = 2$, the region $\mathcal{M}$ is an ellipse, see the discussion preceding Proposition 4.8. Theorem 4.1 is restricted to three quadratic forms ensuring that $\mathcal{M}$ is *convex*. This fails to be the case for more than three quadratic forms, as observed in [21].

We shall apply Theorem 4.1 to uncertainty inequalities thereby obtaining more information about the quadratic forms appearing in them. However, we are restricted to dealing with cases of the uncertainty inequalities that involve only *three* quadratic forms. Specifically, we shall consider inequality (2.3) under the setting of self-adjoint linear operators $A = M_w$ and $B = D$ where w is real-valued for both the Hilbert spaces $\mathcal{L}_2(\mathbb{R})$ and $\mathcal{L}_2[0,2\pi)$ as well as inequality (1.7) for real-valued functions in $\mathcal{L}_2[0,2\pi)$.

Before we proceed, we remark that Theorem 4.1 plays an important role in optimal estimation and n-widths, see [21], [22], [25]. An especially interesting case is discussed in [18] where the region

$$\left\{\left(\int_{|t|>T}|f(t)|^2\,dt, \int_{|t|>\sigma}|\widehat{f}(t)|^2\,dt\right) : \int_{\mathbb{R}}|f(t)|^2\,dt = 1\right\}$$

is described. Another application of Theorem 4.1 is to the construction of shape preserving spline interpolation given in [**6**].

We begin our use of Theorem 4.1 with the following general observation. (Recall from (2.24) that for self-adjoint linear operators A and B, the quantity $i\langle\, [A,B]x,\, x\,\rangle$ is always real.)

Theorem 4.2. *Let A and B be any self-adjoint linear operators with domain and range in the same complex Hilbert space $\mathcal{H}$. Suppose for every $\mu > 0$ there exists $x_\mu \in \mathcal{D}(A|B)$ such that $x_\mu \notin \mathrm{Ker}(B)$ and*

$$Ax_\mu - i\mu Bx_\mu = 0. \tag{4.1}$$

Then the region

$$\{(\|Ax\|^2, \|Bx\|^2) : i\langle\, [A,B]x,\, x\,\rangle = 1,\, x \in \mathcal{D}(A|B)\} \tag{4.2}$$

contains the boundary of the hyperbolic region

$$\mathcal{R} := \{(u,v) : uv \geq 1/4,\, u,v > 0\}. \tag{4.3}$$

Proof: For every $\mu > 0$, we use (4.1) to obtain

$$ABx_\mu = -i\mu^{-1}A^2x_\mu, \qquad BAx_\mu = i\mu B^2x_\mu.$$

Furthermore, since A and B are self-adjoint, it follows that

$$i\langle\, [A,B]x_\mu,\, x_\mu\,\rangle = \mu^{-1}\|Ax_\mu\|^2 + \mu\|Bx_\mu\|^2 = 2\mu\|Bx_\mu\|^2.$$

As a result, the assumption that $x_\mu \notin \mathrm{Ker}(B)$ assures that $r_\mu := i\langle\, [A,B]x_\mu,\, x_\mu\,\rangle > 0$. Thus we may define an element y_μ in $\mathcal{D}(A|B)$ by $y_\mu := r_\mu^{-1/2}x_\mu$. By straightforward calculations, we obtain

$$i\langle\, [A,B]y_\mu,\, y_\mu\,\rangle = 1, \qquad \|Ay_\mu\|^2 = \mu/2, \qquad \|By_\mu\|^2 = 1/(2\mu).$$

Hence, the region in (4.2) contains the set

$$\big\{\big(\|Ay_\mu\|^2,\, \|By_\mu\|^2\big) : \mu > 0\big\} = \{(\mu/2, 1/(2\mu)) : \mu > 0\}$$

which is the entire boundary of the hyperbolic region $\mathcal{R}$ in (4.3). ■

It is worthwhile to note from the above proof that equality holds in inequality (2.3) for all x_μ, $\mu > 0$, satisfying the hypothesis of Theorem 4.2. In addition, equality holds in inequality (2.2) since (2.2) is an improvement of (2.3). This conclusion can also be deduced from Corollary 3.2 or the observation that (4.1) forces the real quantities $\langle Ax_\mu, x_\mu\rangle$ and $\langle Bx_\mu, x_\mu\rangle$ to be zero.

Theorem 4.2 has provided us with sufficient conditions for the hyperbola $uv = 1/4$ to be included in the region in (4.2). If the region in (4.2) is convex, then this implies that it fills up all of the hyperbolic region $\mathcal{R}$. This is the case by Theorem 4.1 whenever the space concerned is a real Hilbert space with dimension at least three and $i[A,B]$ is a nonnegative operator.

In our first example, we deal with the Hilbert space $\mathcal{L}_2(\mathbb{R})$, the multiplication operator M_w given by $M_wf := wf$, $f \in \mathcal{L}_2(\mathbb{R})$, where w is a nontrivial real-valued differentiable function on $\mathbb{R}$ and the derivative operator D in (1.1). Both M_w and D are self-adjoint linear operators. We study the geometric problem associated with inequality (2.3) which involves only three quadratic forms in this case and simplifies to

$$|\langle M_{w'}f, f\rangle_{\mathbb{R}}| \leq 2\|M_w f\|_{\mathbb{R}}\|Df\|_{\mathbb{R}}, \quad f \in \mathcal{D}(M_w|D). \tag{4.4}$$

The region that we are concerned with is

$$\mathcal{Q}_{\mathbb{R}} := \{(\|M_w f\|_{\mathbb{R}}^2, \|Df\|_{\mathbb{R}}^2) : \langle M_{w'}f, f\rangle_{\mathbb{R}} = 1,\ f \in \mathcal{D}(M_w|D)\} \tag{4.5}$$

and the above inequality shows that it lies inside the hyperbolic region $\mathcal{R}$ in (4.3). A question of interest is whether $\mathcal{Q}_{\mathbb{R}}$ *coincides* with $\mathcal{R}$. To this end, we impose some additional conditions on the function w. These conditions are stated in the following proposition.

Proposition 4.3. *Let $W : \mathbb{R} \longrightarrow \mathbb{R}$ be an everywhere differentiable function where $w := W'$ is nontrivial and absolutely continuous on $\mathbb{R}$ such that $W''(t) > 0$, $t \in \mathbb{R}$, a.e.. Furthermore, suppose for every $\mu > 0$, the functions $e^{-W/\mu}$, $w^2 e^{-W/\mu}$, $w' e^{-W/\mu}$ are all in $\mathcal{L}_2(\mathbb{R})$. Then the region $\mathcal{Q}_{\mathbb{R}}$ in* (4.5) *is exactly the hyperbolic region $\mathcal{R}$ in* (4.3).

Proof: For $\mu > 0$, define the real-valued function $f_\mu := e^{-W/\mu}$. By straightforward calculations, we see that the hypothesis of Theorem 4.2 is satisfied. Then it follows from the proof of Theorem 4.2 that the set

$$\{(\|M_w f\|_{\mathbb{R}}^2, \|Df\|_{\mathbb{R}}^2) : \langle M_{w'}f, f\rangle_{\mathbb{R}} = 1,\ f : \mathbb{R} \longrightarrow \mathbb{R},\ f \in \mathcal{D}(M_w|D)\} \tag{4.6}$$

contains the boundary of $\mathcal{R}$. Now, $\langle M_{w'}\cdot, \cdot\rangle_{\mathbb{R}}$ is positive definite on the *real* linear space of all real-valued functions in $\mathcal{D}(M_w|D)$. We apply Theorem 4.1 to conclude that the set in (4.6) is convex and hence contains the region $\mathcal{R}$.

On the other hand, as observed earlier from (4.4), the set $\mathcal{Q}_{\mathbb{R}}$ which contains the set in (4.6) lies inside $\mathcal{R}$. This implies that $\mathcal{Q}_{\mathbb{R}}$ is exactly $\mathcal{R}$. ∎

There are many examples of W that satisfy the assertions in Proposition 4.3. Indeed, we may take W to be polynomials of the form

$$W(t) = a_0 + a_1 t + \sum_{j=1}^{n} a_{2j} t^{2j}, \quad t \in \mathbb{R}$$

where n is a positive integer, $a_0, a_1 \in \mathbb{R}$, $a_{2j} \geq 0$ for $1 \leq j < n$ and $a_{2n} > 0$. The special case of $W(t) = t^2/2$, $t \in \mathbb{R}$ is the setting of the classical Heisenberg uncertainty principle (1.2). Proposition 4.3 then gives a complete description of the set

$$\left\{ \left(\int_{\mathbb{R}} t^2 |f(t)|^2\, dt, \int_{\mathbb{R}} |f'(t)|^2\, dt \right) : \int_{\mathbb{R}} |f(t)|^2\, dt = 1,\ f \in \mathcal{D}(M|D) \right\}.$$

We find the application of Theorem 4.2 to the case of 2π-periodic functions inadequate for reasons which we do not describe. The next theorem applies perfectly to the case of 2π-periodic functions.

Theorem 4.4. *With A and B as in Theorem 4.2, suppose for every $\mu \in \mathbb{R}\backslash\{0\}$ there exists $x_\mu \in \mathcal{D}(A|B)$ such that $x_\mu \notin \operatorname{Ker}(B)$ and (4.1) holds. In addition, assume that the set $\operatorname{Ker}(B)\backslash\operatorname{Ker}(A)$ is nonempty. Then the region*

$$\{(i\langle [A,B]x,\, x\,\rangle, \|Bx\|^2) : \|Ax\|^2 = 1,\, x \in \mathcal{D}(A|B)\} \tag{4.7}$$

contains the boundary of the parabolic region

$$\mathcal{S} := \{(u,v) : v \geq u^2/4,\, u \in \mathbb{R}\}. \tag{4.8}$$

Proof: The proof of this result is similar to that of Theorem 4.2. The elements x_μ, $\mu \in \mathbb{R}\backslash\{0\}$, allow us to conclude that the region in (4.7) contains the set

$$\{(2/\mu, 1/\mu^2) : \mu \in \mathbb{R}\backslash\{0\}\} = \{(u,v) : v = u^2/4,\, u \in \mathbb{R}\backslash\{0\}\}.$$

By considering any element of $\operatorname{Ker}(B)\backslash\operatorname{Ker}(A)$, we see that the vertex $(0,0)$ of the parabola $v = u^2/4$ also lies inside the region. ■

Now, we shall apply Theorem 4.4 to the Hilbert space $\mathcal{L}_2[0,2\pi)$ and the self-adjoint operators M_w and D where w is a real-valued function in $\mathcal{W}^1_\infty[0,2\pi)\backslash\{0\}$. In this case, inequality (2.3) becomes

$$|\langle M_{w'}f, f\rangle_{2\pi}| \leq 2\|M_w f\|_{2\pi}\|Df\|_{2\pi}, \quad f \in \mathcal{W}^1_2[0,2\pi). \tag{4.9}$$

Proposition 4.5. *If $w \in \mathcal{W}^1_\infty[0,2\pi)\backslash\{0\}$ is real-valued with $\displaystyle\int_{[0,2\pi)} w(t)\,dt = 0$, then the region*

$$\mathcal{Q}_{2\pi} := \{(\langle M_{w'}f, f\rangle_{2\pi}, \|Df\|^2_{2\pi}) : \|M_w f\|^2_{2\pi} = 1,\, f \in \mathcal{W}^1_2[0,2\pi)\}$$

is exactly the parabolic region $\mathcal{S}$ in (4.8).

Proof: For $\mu \in \mathbb{R}\backslash\{0\}$, define the real-valued function

$$f_\mu(t) := \exp\left\{-\int_0^t \frac{w(s)}{\mu}\,ds\right\}, \quad t \in [0,2\pi).$$

Furthermore, note that the set $\operatorname{Ker}(D)\backslash\operatorname{Ker}(M_w)$ contains constant functions. Then proceeding as in the proof of Proposition 4.3, we deduce from Theorem 4.4, Theorem 4.1 and inequality (4.9) that $\mathcal{Q}_{2\pi} = \mathcal{S}$. ■

Examples of w that satisfy the hypothesis of Proposition 4.5 include trigonometric polynomials of the form

$$w(t) = \sum_{j=1}^n (a_j \cos jt + b_j \sin jt), \quad t \in \mathbb{R}$$

where n is a positive integer, $a_j, b_j \in \mathbb{R}$ for $1 \leq j \leq n$ and $|a_n| + |b_n| \neq 0$. Although the proofs of Propositions 4.3 and 4.5 are similar, there are subtle differences in the details. For instance, one deals with only the case $\mu > 0$ in the proof of Proposition 4.3 whereas both cases $\mu > 0$ and $\mu < 0$ are considered to establish Proposition 4.5.

In the remainder of this section our attention turns to inequality (1.7) and we restrict ourselves to the Hilbert space $\mathcal{T}_\infty$ of all real-valued even functions in $\mathcal{W}_2^1[0,2\pi)$. In this case, the functional $\nu_{2\pi}(f)$ in (1.7) reduces to

$$\nu_{2\pi}(f) = \frac{(\|f\|_{2\pi}^4 - |\langle Sf, f\rangle_{2\pi}|^2)\|Df\|_{2\pi}^2}{|\langle Sf, f\rangle_{2\pi}|^2\|f\|_{2\pi}^2} \tag{4.10}$$

where the quantity $\langle Sf, f\rangle_{2\pi}$ simplifies to

$$\langle Sf, f\rangle_{2\pi} = \frac{1}{2\pi}\int_{[0,2\pi)} (\cos t) f^2(t)\, dt.$$

In (4.10) there appear three quadratic forms. We introduce the set

$$\mathcal{A}_\infty := \{(\langle Sf, f\rangle_{2\pi}, \|Df\|_{2\pi}^2) : \|f\|_{2\pi}^2 = 1,\ f \in \mathcal{T}_\infty\}$$

and note by Theorem 4.1 that $\mathcal{A}_\infty$ is a convex set. As a preliminary remark, we observe that $\mathcal{A}_\infty$ is contained in the vertical strip $[-1,1]\times\mathbb{R}_+$ and the only boundary point of the strip which lies in $\mathcal{A}_\infty$ is the origin. Clearly, $\mathcal{A}_\infty$ is symmetric about the vertical axis. Let θ be the function defined by the formula

$$\theta(u) := \frac{u^2}{4(1-u^2)}, \quad |u| < 1.$$

Then the uncertainty inequality (1.7) says that $\mathcal{A}_\infty$ lies in the region above the graph of θ, that is, $\mathcal{A}_\infty \subseteq \mathcal{U}$ where

$$\mathcal{U} := \{(u,v) : v \geq \theta(u),\, |u| < 1\}.$$

This inclusion captures inequality (1.7). Nevertheless, the only boundary point of $\mathcal{U}$ in $\mathcal{A}_\infty$ is again the origin, see our remarks after Proposition 3.3. To actually identify the boundary of $\mathcal{A}_\infty$, and hence the set itself, we are led to consider the function

$$\rho_\infty(u) := \inf\{\|Df\|_{2\pi}^2 : f \in \mathcal{C}_\infty(u)\} \tag{4.11}$$

where

$$\mathcal{C}_\infty(u) := \{f : \|f\|_{2\pi}^2 = 1,\ \langle Sf, f\rangle_{2\pi} = u,\ f \in \mathcal{T}_\infty\}$$

for $|u| < 1$. By the definition of this function, we see that $\mathcal{A}_\infty$ lies above its graph, that is, $\mathcal{A}_\infty \subseteq \mathcal{V}_\infty$ where

$$\mathcal{V}_\infty := \{(u,v) : v \geq \rho_\infty(u),\, |u| < 1\}.$$

We claim that $\mathcal{V}_\infty = \mathcal{A}_\infty$. This fact as well as more information about the function ρ_∞ appears in the next result.

Theorem 4.6. *The infimum* (4.11) *is attained and the function* ρ_∞ *is continuous, convex and even on* $(-1,1)$. *It is strictly increasing on* $[0,1)$ *and* $\lim_{u\to 1^-}\rho_\infty(u) = \infty$. *We also have*

$$\mathcal{V}_\infty = \mathcal{A}_\infty.$$

Thus the graph of ρ_∞ *is the boundary of* $\mathcal{A}_\infty$.

Before we prove the theorem, let us make sure that $\rho_\infty(u) < \infty$ for $|u| < 1$. For this purpose we consider the function

$$f_u(t) := \frac{\sqrt{1-u^2}}{\sqrt{1+u^2-2u\cos t}}, \quad t \in \mathbb{R}.$$

Since

$$f_u^2(t) = \frac{1-u^2}{|1-ue^{it}|^2} = \sum_{n\in\mathbb{Z}} u^{|n|} e^{int}, \quad t \in \mathbb{R}, \tag{4.12}$$

it follows that $f_u \in \mathcal{C}_\infty(u)$. Differentiating both sides of the first equation in (4.12) and simplifying yields the formula

$$(f_u'(t))^2 = \frac{u^2(1-u^2)\sin^2 t}{|1-ue^{it}|^6}, \quad t \in \mathbb{R}.$$

By Cauchy's residue theorem, we have that

$$\|Df_u\|_{2\pi}^2 = \frac{u^2}{2(1-u^2)^2}.$$

Thus we obtain $\mathcal{C}_\infty(u) \neq \emptyset$ if and only if $|u| < 1$ and in this case

$$\theta(u) \le \rho_\infty(u) \le \frac{u^2}{2(1-u^2)^2} \tag{4.13}$$

where the lower bound follows from the uncertainty inequality (1.7). We are ready to prove Theorem 4.6.

Proof of Theorem 4.6: First we show that for $|u| < 1$, the infimum in the definition of $\rho_\infty(u)$ is attained. Let $\{f_k\}_{k\ge1}$ be a minimizing sequence. Thus we have that $f_k \in \mathcal{T}_\infty$,

$$\lim_{k\to\infty} \|Df_k\|_{2\pi}^2 = \rho_\infty(u), \tag{4.14}$$

$$\|f_k\|_{2\pi}^2 = 1 \tag{4.15}$$

and

$$\langle Sf_k, f_k\rangle_{2\pi} = u.$$

From (4.15) there is a subsequence $\{f_{k'}\}$ which converges weakly to a function $f \in \mathcal{L}_2[0,2\pi)$. Let P_N be the orthogonal projection onto the subspace $\{e^{ij\cdot} : |j| \le N\}$ and observe for any $g \in \mathcal{W}_2^1[0,2\pi)$ that

$$\|g\|_{2\pi}^2 - \|P_N g\|_{2\pi}^2 = \|g - P_N g\|_{2\pi}^2 \le \frac{\|Dg\|_{2\pi}^2}{(N+1)^2}. \tag{4.16}$$

By (4.14) and (4.15), specializing (4.16) to $g = f_k$ gives

$$1 - \|P_N f_k\|_{2\pi}^2 \le \frac{2\rho_\infty(u)}{(N+1)^2}$$

for sufficiently large k. From the weak convergence of $\{f_{k'}\}$ to f it follows that

$$1 - \|P_N f\|_{2\pi}^2 \le \frac{2\rho_\infty(u)}{(N+1)^2}.$$

Letting $N \to \infty$ we get that $\|f\|_{2\pi}^2 = 1$. Similarly, for any $g \in \mathcal{W}_2^1[0, 2\pi)$, we have that

$$|\langle Sg, g\rangle_{2\pi} - \langle SP_N g, P_N g\rangle_{2\pi}| \leq \frac{\|Dg\|_{2\pi}^2}{N^2}$$

from which it follows that $\langle Sf, f\rangle_{2\pi} = u$. Finally, weak convergence ensures that $\|Df\|_{2\pi}^2 \leq \rho_\infty(u)$ which proves that the infimum is attained and

$$(u, \rho_\infty(u)) = (\langle Sf, f\rangle_{2\pi}, \|Df\|_{2\pi}^2)$$

with $f \in \mathcal{T}_\infty$ and $\|f\|_{2\pi}^2 = 1$. Thus the graph of ρ_∞ is in $\mathcal{A}_\infty$.

Let us show that all points of $\mathcal{V}_\infty$ are in $\mathcal{A}_\infty$. To this end, we prove that ρ_∞ is convex. For every u_1, u_2 in $(-1, 1)$ and λ in $[0, 1]$, the convexity of $\mathcal{A}_\infty$ implies that the point $\lambda(u_1, \rho_\infty(u_1)) + (1-\lambda)(u_2, \rho_\infty(u_2))$ is in $\mathcal{A}_\infty$. Thus there is a $g \in \mathcal{T}_\infty$ such that

$$\|Dg\|_{2\pi}^2 = \lambda\rho_\infty(u_1) + (1-\lambda)\rho_\infty(u_2), \qquad \langle Sg, g\rangle_{2\pi} = \lambda u_1 + (1-\lambda)u_2,$$

and $\|g\|_{2\pi}^2 = 1$. Hence we conclude that

$$\rho_\infty(\lambda u_1 + (1-\lambda)u_2) \leq \|Dg\|_{2\pi}^2 = \lambda\rho_\infty(u_1) + (1-\lambda)\rho_\infty(u_2)$$

and so ρ_∞ is convex. Consequently, ρ_∞ is also continuous on $(-1, 1)$. Now, choose any point $(u, v) \in \mathcal{V}_\infty$ which is not on the graph of ρ_∞ so that $v > \rho_\infty(u)$. By the continuity of ρ_∞ and the fact that $\lim_{t\to 1^-} \rho_\infty(t) = \infty$, see inequality (4.13), there exists $u' \in (u, 1)$ such that $v = \rho_\infty(u')$. Therefore the point (u, v) is on the line segment joining $(-u', \rho_\infty(u'))$ and $(u', \rho_\infty(u'))$ and so must be in $\mathcal{A}_\infty$. That is, we have that $\mathcal{V}_\infty = \mathcal{A}_\infty$.

There remains to prove that ρ_∞ is strictly increasing on $[0, 1)$. We do this by considering another extremal problem, namely

$$\widetilde{\rho}_\infty(u) := \inf\{\|Df\|_{2\pi}^2 : \|f\|_{2\pi}^2 = 1,\ \langle Sf, f\rangle_{2\pi} \geq u,\ f \in \mathcal{T}_\infty\}.$$

As with the extremal problem (4.11) defining ρ_∞, the infimum for this one is attained as well. Clearly, $\widetilde{\rho}_\infty$ is nondecreasing and $\widetilde{\rho}_\infty \leq \rho_\infty$. We shall show that $\rho_\infty = \widetilde{\rho}_\infty$. Suppose $u \in [0, 1)$ and $h \in \mathcal{T}_\infty$, $\|h\|_{2\pi}^2 = 1$, $\langle Sh, h\rangle_{2\pi} \geq u$ with $\widetilde{\rho}_\infty(u) = \|Dh\|_{2\pi}^2$. If we can show $\langle Sh, h\rangle_{2\pi} = u$, then $\rho_\infty(u) \leq \widetilde{\rho}_\infty(u)$ from which we conclude that $\rho_\infty = \widetilde{\rho}_\infty$. Suppose to the contrary that $\langle Sh, h\rangle_{2\pi} > u$. Then $h \in \mathcal{W}_2^2[0, 2\pi)$ and there exists a constant $\alpha \in \mathbb{R}$ such that $h'' + \alpha h = 0$ a.e. on $[0, 2\pi)$. Hence $\alpha = \|Dh\|_{2\pi}^2$ and we set $\alpha = \mu^2$ for some $\mu \in \mathbb{R}$. Since h is 2π-periodic, it follows that $\mu \in \mathbb{Z}$ and using the fact that h is even gives us that $h(t) = \sqrt{2}\cos\mu t$, $t \in [0, 2\pi)$ if $\mu \neq 0$ and $h(t) = 1$, $t \in [0, 2\pi)$ when $\mu = 0$. In both cases $\langle Sh, h\rangle_{2\pi} = 0$, but that contradicts the assumption that $u \geq 0$. This leads us to the desired conclusion that indeed $\rho_\infty = \widetilde{\rho}_\infty$. To finish the proof we will now show that ρ_∞ is strictly increasing. Suppose $\rho_\infty(u_1) = \rho_\infty(u_2)$ with $u_1 \neq u_2$ and for definiteness that $u_1 > u_2$. Choose $f_1 \in \mathcal{C}_\infty(u_1)$ such that $\rho_\infty(u_1) = \|Df_1\|_{2\pi}^2$. But then $\langle Sf_1, f_1\rangle_{2\pi} > u_2$ and so $\widetilde{\rho}_\infty(u_2) = \|Df_1\|_{2\pi}^2$. This is a contradiction to what we have already established and the proof is complete. ■

For a fixed $u \in (-1,1)$, whenever $\rho_\infty(u) = \|Df\|_{2\pi}^2$ with $f \in \mathcal{C}_\infty(u)$ we conclude that $f \in \mathcal{W}_2^2[0,2\pi)$ and there are real constants α and β such that

$$f''(t) + (\alpha + \beta \cos t) f(t) = 0, \tag{4.17}$$

$t \in [0,2\pi)$, a.e.. Indeed, let $a_0 + 2\sum_{j=1}^{\infty} a_j \cos jt$ be the Fourier series of f. For $N \geq 1$, consider $g_N \in \mathcal{C}_\infty(u)$ whose Fourier series is of the form $b_0 + 2\sum_{j=1}^{N} b_j \cos jt + 2\sum_{j=N+1}^{\infty} a_j \cos jt$. We apply the method of Lagrange multipliers and set $\nabla H = 0$ where

$$H(b_0, \ldots, b_N) := \|Dg_N\|_{2\pi}^2 - \alpha_N(\|g_N\|_{2\pi}^2 - 1) - \beta_N(\langle Sg_N, g_N\rangle_{2\pi} - u)$$

to derive a system of linear equations in $a_0, \ldots, a_N$. The sequence $\{(\alpha_N, \beta_N)\}_{N\geq 1}$ in $\mathbb{R}^2$ is bounded since $\|f\|_{2\pi}^2 = 1$. By passing onto a converging subsequence with limit $(\alpha, \beta) \in \mathbb{R}^2$, we obtain (4.17).

The solution of (4.17) is related to the *Mathieu function* which is defined by the equation

$$y''(t) + (a - 2q\cos 2t) y(t) = 0, \tag{4.18}$$

see for instance [4], [20]. Properties of the Mathieu function have been intensively investigated because of its importance in a number of practical applications, for example, determining the vibrational modes of a stretched membrane with elliptical boundary. For us, it is clear that the function f above is the unique (up to a sign) even (real) solution to (4.17) normalized so that $\|f\|_{2\pi}^2 = 1$ (for α and β given). Moreover, $f(t) = y(t/2)$ where y is the Mathieu function corresponding to (4.18) when $a = 4\alpha$, $q = -2\beta$. The parameters α, β, and hence a, q, must be interrelated to ensure that f is 2π-periodic. A choice of q will lead to an infinite number of admissible values of a which are called the *characteristic numbers* of the Mathieu equation, see for instance [4], [20]. In addition, a and q must be chosen so that $\langle Sf, f\rangle_{2\pi} = u$. We are not sure whether or not these requirements specify a and q *uniquely*. In any case, it follows that $\rho_\infty(u) = \alpha + \beta u$.

Returning to the quantity $\nu_{2\pi}(f)$ in (4.10), we see that

$$\frac{1}{4} = \inf\{\nu_{2\pi}(f) : f \in \mathcal{T}_\infty\} = \inf\left\{\frac{1-u^2}{u^2}\rho_\infty(u) : |u| < 1\right\}$$

which gives an indication of the asymptotic behavior of ρ_∞ near $u = 1$. The benefit of the function ρ_∞ is that it leads to an improvement of the uncertainty inequality (1.7) in the form

$$\rho_\infty(\langle Sf, f\rangle_{2\pi}) \leq \|Df\|_{2\pi}^2$$

which can be contrasted to (1.7) rewritten as

$$\theta(\langle Sf, f\rangle_{2\pi}) \leq \|Df\|_{2\pi}^2$$

for $f \in \mathcal{T}_\infty$ and $\|f\|_{2\pi} = 1$.

We now turn to analogs of the above results for trigonometric polynomials. We denote by $\mathcal{T}_n$ the subspace of trigonometric polynomials of degree at most n, $n \geq 1$, in $\mathcal{T}_\infty$ and set

$$\mathcal{C}_n(u) := \mathcal{C}_\infty(u) \cap \mathcal{T}_n, \quad |u| < 1.$$

Proposition 4.7. *$\mathcal{C}_n(u) \neq \emptyset$ if and only if $|u| \leq \cos \dfrac{\pi}{2(n+1)}$.*

Proof: Every $f \in \mathcal{T}_n$ can be written in the form

$$f(t) = p(\cos t), \quad t \in [0, 2\pi) \tag{4.19}$$

where $p \in \mathcal{P}_n$, the space of algebraic polynomials of degree at most n with real coefficients. Consequently, we obtain that

$$\|f\|_{2\pi}^2 = \frac{1}{\pi} \int_{-1}^{1} \frac{|p(u)|^2}{\sqrt{1-u^2}}\, du \tag{4.20}$$

and

$$\langle Sf, f\rangle_{2\pi} = \frac{1}{\pi} \int_{-1}^{1} \frac{u|p(u)|^2}{\sqrt{1-u^2}}\, du. \tag{4.21}$$

It is a result of Chebyshev, see for example [35], that the maximum value of $\langle Sf, f\rangle_{2\pi}$ for $f \in \mathcal{T}_n$, $\|f\|_{2\pi}^2 = 1$, is $\cos \dfrac{\pi}{2(n+1)}$ and is attained uniquely (except for a sign) by the polynomial

$$p_+(u) := \sqrt{n+1}\, \frac{T_{n+1}(u)}{T'_{n+1}(\xi_1)(u-\xi_1)}, \quad |u| \leq 1. \tag{4.22}$$

The proof is based on the Gauss quadrature formula

$$\frac{1}{\pi} \int_{-1}^{1} \frac{q(u)}{\sqrt{1-u^2}}\, du = \frac{1}{n+1} \sum_{j=1}^{n+1} q(\xi_j) \tag{4.23}$$

where $\xi_j := \cos \dfrac{2j-1}{2(n+1)}\pi$, $j = 1, \dots, n+1$, are the zeros of the $(n+1)$-st Chebyshev polynomial T_{n+1} defined by

$$T_{n+1}(\cos t) := \cos(n+1)t, \quad t \in [0, 2\pi)$$

and $q \in \mathcal{P}_{2n+1}$, see [35]. Specializing (4.23) yields

$$\langle Sf, f\rangle_{2\pi} = \frac{1}{n+1} \sum_{j=1}^{n+1} \xi_j |p(\xi_j)|^2 \tag{4.24}$$

and

$$\|f\|_{2\pi}^2 = \frac{1}{n+1} \sum_{j=1}^{n+1} |p(\xi_j)|^2. \tag{4.25}$$

The polynomial p_+ in (4.22) has been chosen so that the corresponding $f_+ \in \mathcal{T}_n$ in (4.19) satisfies $\|f_+\|_{2\pi}^2 = 1$ and $\langle Sf_+, f_+\rangle_{2\pi} = \cos\dfrac{\pi}{2(n+1)}$. Likewise, the polynomial

$$p_-(u) := \sqrt{n+1}\frac{T_{n+1}(u)}{T'_{n+1}(\xi_{n+1})(u-\xi_{n+1})}, \quad |u| \le 1 \tag{4.26}$$

leads to an $f_- \in \mathcal{T}_n$ such that $\|f_-\|_{2\pi}^2 = 1$ and $\langle Sf_-, f_-\rangle_{2\pi} = -\cos\dfrac{\pi}{2(n+1)}$, which minimizes the value of $\langle Sf, f\rangle_{2\pi}$ over all $f \in \mathcal{T}_n$, $\|f\|_{2\pi}^2 = 1$. Moreover for $f = c_- f_- + c_+ f_+$, we have that $\|f\|_{2\pi}^2 = c_-^2 + c_+^2$ and $\langle Sf, f\rangle_{2\pi} = \xi_1 c_+^2 + \xi_{n+1} c_-^2$ which means that for *any* $|u| \le \cos\dfrac{\pi}{2(n+1)}$, there is indeed an $f \in \mathcal{T}_n$ with $\|f\|_{2\pi}^2 = 1$ and $\langle Sf, f\rangle_{2\pi} = u$ obtained by adjusting c_- and c_+ above. This proves the result. ■

Formulas (4.19)–(4.21), (4.24) and (4.25) suggest that we can also express $\|Df\|_{2\pi}^2$ in terms of the quantities $p(\xi_j)$, $j = 1, \ldots, n+1$. For this purpose we first observe that

$$\|Df\|_{2\pi}^2 = \sum_{j=1}^{n+1}\sum_{k=1}^{n+1} R_{jk}\, p(\xi_j)p(\xi_k) \tag{4.27}$$

where

$$R_{jk} := \frac{1}{\pi}\int_{-1}^{1} \ell'_j(u)\ell'_k(u)\sqrt{1-u^2}\,du, \quad j,k = 1,\ldots,n+1$$

and

$$\ell_j(u) := \frac{T_{n+1}(u)}{T'_{n+1}(\xi_j)(u-\xi_j)}, \quad j = 1,\ldots,n+1,\ u \in \mathbb{R}.$$

Using the Christoffel-Darboux formula for Chebyshev polynomials, see for instance [35], we have for $j = 1, \ldots, n+1$ that

$$\ell_j(u) = \frac{2}{n+1}{\sum_{k=0}^{n}}' T_k(\xi_j)T_k(u), \quad u \in \mathbb{R}$$

(where the prime over the summation indicates that the first term is halved). Consequently, we have that

$$\ell'_j(u) = \frac{2}{n+1}\sum_{k=1}^{n} kT_k(\xi_j)v_{k-1}(u), \quad u \in \mathbb{R}$$

where v_k is the Chebyshev polynomial of the second kind, orthonormal relative to the weight function $\dfrac{2}{\pi}\sqrt{1-u^2}$ on $[-1,1]$, see [35]. Hence we have for $j,k = 1,\ldots,n+1$ that

$$R_{jk} = \frac{2}{(n+1)^2}\sum_{r=1}^{n} r^2\, T_r(\xi_j)T_r(\xi_k). \tag{4.28}$$

A direct computation confirms the (explicit) formula

$$R_{jk} = \frac{1}{8(n+1)^2}\left\{V_n\left(\frac{(j+k-1)\pi}{n+1}\right) + V_n\left(\frac{(j-k)\pi}{n+1}\right)\right\}$$

where

$$V_n(t) := \frac{-n^2 \sin(n+\frac{3}{2})t + (2n^2+2n-1)\sin(n+\frac{1}{2})t - (n+1)^2 \sin(n-\frac{1}{2})t}{\sin^3 \frac{t}{2}}, \quad t \in \mathbb{R}.$$

Recall that the matrix

$$U := \frac{\sqrt{2}}{\sqrt{n+1}} (T_r(\xi_j))_{r=0,j=1}^{n,n+1}$$

has the property that $UU^T = \operatorname{diag}\{2,1,\dots,1\}$ and the matrix $R := (R_{jk})_{j,k=1}^{n+1}$ can be alternatively written in the form

$$R = \frac{1}{n+1} U^T \Lambda U$$

where $\Lambda := \operatorname{diag}\{0,1,\dots,n^2\}$.

The analog of $\mathcal{A}_\infty$ for $\mathcal{T}_n$ is the set

$$\mathcal{A}_n := \{(\langle Sf, f\rangle_{2\pi}, \|Df\|_{2\pi}^2) : \|f\|_{2\pi}^2 = 1,\ f \in \mathcal{T}_n\}.$$

According to Proposition 4.7, $\mathcal{A}_n$ lies in the vertical strip $I_n \times \mathbb{R}_+$ where

$$I_n := \left[-\cos\frac{\pi}{2(n+1)}, \cos\frac{\pi}{2(n+1)}\right].$$

In addition, since $\|Df\|_{2\pi}^2 \le n^2 \|f\|_{2\pi}^2$ for $f \in \mathcal{T}_n$ (Bernstein inequality for $\mathcal{L}_2[0,2\pi)$), we have that $\mathcal{A}_n$ lies in the rectangle $I_n \times J_n$ where $J_n := [0, n^2]$ and the only boundary points of $I_n \times J_n$ in $\mathcal{A}_n$ are $(0,0)$, $(0,n^2)$, $\left(\cos\frac{\pi}{2(n+1)}, \|Df_+\|_{2\pi}^2\right)$ and $\left(-\cos\frac{\pi}{2(n+1)}, \|Df_-\|_{2\pi}^2\right)$.

To describe the boundary completely we find it convenient to *reflect* the set $\mathcal{A}_n$ along the line $v = u$ and consider the set

$$\mathcal{B}_n := \{(\|Df\|_{2\pi}^2, \langle Sf, f\rangle_{2\pi}) : \|f\|_{2\pi}^2 = 1,\ f \in \mathcal{T}_n\}.$$

Since $\dim \mathcal{T}_n = n+1$, it follows from Theorem 4.1 that the sets $\mathcal{A}_n$ and $\mathcal{B}_n$ are convex for $n \ge 2$. When $n = 1$, we can determine them precisely. Specifically, the conditions $f \in \mathcal{T}_1$, $\|Df\|_{2\pi}^2 = u$, $0 \le u \le 1$, $\|f\|_{2\pi}^2 = 1$ imply that $f(t) = a_0 + 2a_1 \cos t$, $t \in [0, 2\pi)$, $a_0^2 = 1 - u$ and $a_1^2 = u/2$. Thus we have that $\langle Sf, f\rangle_{2\pi} = \pm\sqrt{2u(1-u)}$, that is, we conclude that $\mathcal{B}_1$ is the ellipse

$$\left\{(u,v) : \left(u - \frac{1}{2}\right)^2 + \frac{v^2}{2} = \frac{1}{4},\ u, v \in \mathbb{R}\right\}.$$

This is to be expected. Indeed, for *any* two 2×2 real symmetric matrices A and B, the image of the unit circle under the mapping $x \longrightarrow (x^T A x, x^T B x)$ is an ellipse or a line, see [6] for a special case. To specify the ellipse we let

$$\alpha := \frac{a_{11} + a_{22}}{2}, \qquad \beta := \frac{b_{11} + b_{22}}{2}$$

and

$$\gamma := \frac{1}{2}\{(a_{11} - a_{22})b_{12} - (b_{11} - b_{22})a_{12}\}$$

where

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}, \qquad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{12} & b_{22} \end{pmatrix}.$$

The equation of the ellipse (or line) is given by

$$\gamma^2 = (b_{12}(u-\alpha) - a_{12}(v-\beta))^2 + \frac{1}{4}\left((b_{22}-b_{11})(u-\alpha) + (a_{11}-a_{22})(v-\beta)\right)^2.$$

Returning to the general case $n \geq 2$, we provide the following description of the boundary of $\mathcal{B}_n$ which requires the function $\tau_n : J_n \longrightarrow I_n$ defined by

$$\tau_n(u) := \max\{\langle Sf, f\rangle_{2\pi} : \|f\|_{2\pi}^2 = 1,\ \|Df\|_{2\pi}^2 = u,\ f \in \mathcal{T}_n\}. \tag{4.29}$$

By considering functions of the form $f(t) = a_0 + 2a_n \cos nt$, $t \in [0, 2\pi)$, we see that for every $u \in J_n$ there exists an $f \in \mathcal{T}_n$ such that $\|f\|_{2\pi}^2 = 1$ and $\|Df\|_{2\pi}^2 = u$. Therefore the function τ_n is well defined and nonnegative.

Proposition 4.8. *If $n \geq 2$, then $\tau_n : J_n \longrightarrow I_n$ is a continuous concave function and*

$$\mathcal{B}_n = \{(u,v) : |v| \leq \tau_n(u),\ u \in J_n\}.$$

Thus the curves $(u, \tau_n(u))$, $(u, -\tau_n(u))$, $u \in J_n$, provide the boundary of $\mathcal{B}_n$.

Proof: If $(u,v) \in \mathcal{B}_n$, then by definition (4.29) we have that $v \leq \tau_n(u)$. Choose an $f \in \mathcal{T}_n$ such that $u = \|Df\|_{2\pi}^2$ and $v = \langle Sf, f\rangle_{2\pi}$ and let $g = f(\pi - \cdot)$. Then $(\|Dg\|_{2\pi}^2, \langle Sg, g\rangle_{2\pi}) = (u, -v)$ and so we also have $-v \leq \tau_n(u)$, that is, $|v| \leq \tau_n(u)$. Conversely, given $(u,v) \in J_n \times I_n$ such that $|v| \leq \tau_n(u)$, let $f \in \mathcal{T}_n$ be chosen so that $\langle Sf, f\rangle_{2\pi} = \tau_n(u)$, $\|Df\|_{2\pi}^2 = u$, $\|f\|_{2\pi}^2 = 1$. Consequently, the trigonometric polynomial $g = f(\pi - \cdot)$ gives $(u, -\tau_n(u)) \in \mathcal{B}_n$. Therefore, by the convexity of the set $\mathcal{B}_n$, we conclude that every (u, v') with $|v'| \leq \tau_n(u)$ is in $\mathcal{B}_n$, and in particular, $(u,v) \in \mathcal{B}_n$.

The concavity of the function τ_n follows as in the proof of the convexity of ρ_∞ given in Theorem 4.6. This gives the continuity of τ_n in the interior of J_n. For the continuity of τ_n at the end-points of J_n we use the estimate

$$\tau_n(u) \leq n^2 \min\{\sqrt{u}, \sqrt{n^2 - u}\}$$

which can be verified by direct computation and the fact that τ_n vanishes at the end-points of J_n. ∎

There are alternative ways to describe the boundary of $\mathcal{B}_n$. To this end, we write any $f \in \mathcal{T}_n$ in its Fourier series expansion

$$f(t) = \sum_{j=-n}^{n} a_j e^{ijt}, \quad t \in \mathbb{R}$$

where $a_j = \overline{a_j} = a_{-j}$, $j = -n, \ldots, n$. We then have that

$$\langle Sf, f\rangle_{2\pi} = 2\sum_{j=0}^{n-1} a_j a_{j+1} = b^T \Gamma b$$

where $b := (a_0, \sqrt{2}a_1, \dots, \sqrt{2}a_n)^T \in \mathbb{R}^{n+1}$ and Γ is the $(n+1) \times (n+1)$ matrix

$$\Gamma := \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & 0 & \cdots & \cdots & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{2} & \cdots & \cdots & 0 \\ 0 & \frac{1}{2} & 0 & \ddots & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \cdots & \ddots & \ddots & 0 & \frac{1}{2} \\ 0 & 0 & \cdots & 0 & \frac{1}{2} & 0 \end{pmatrix}.$$

Also, $\|Df\|_{2\pi}^2 = b^T \Lambda b$ where Λ is the diagonal matrix $\Lambda = \operatorname{diag}(0, 1, \dots, n^2)$ and $\|f\|_{2\pi}^2 = b^T b$. Suppose (u, v) is a boundary point of $\mathcal{B}_n$. Thus there is a $c \in \mathbb{R}^{n+1}$ with $c^T c = 1$ and $\zeta, \eta, \kappa \in \mathbb{R}$ where $|\zeta| + |\eta| \neq 0$ such that for *every* $b \in \mathbb{R}^{n+1}$ with $b^T b = 1$ there holds

$$\zeta b^T \Lambda b + \eta b^T \Gamma b \leq \kappa$$

while

$$\zeta c^T \Lambda c + \eta c^T \Gamma c = \kappa$$

and

$$u = c^T \Lambda c, \qquad v = c^T \Gamma c.$$

Hence c is an eigenvector of the matrix $\zeta\Lambda + \eta\Gamma$ with maximum eigenvalue κ. The matrix $\zeta\Lambda + \eta\Gamma$ is a real symmetric tridiagonal matrix and when $\eta \neq 0$ it is *unreduced* in the terminology of [24]. Thus, the lemma on p. 124 of [24] implies that it has simple eigenvalues, a fact that persists even when $\eta = 0$.

So far, we have observed that a point on the boundary of $\mathcal{B}_n$ is of the form $(c^T \Lambda c, c^T \Gamma c)$ where c is the unique (up to a sign) eigenvector corresponding to the largest eigenvalue of $\zeta\Lambda + \eta\Gamma$ with $c^T c = 1$. The argument which yields that conclusion is "reversible", that is, given any $\zeta, \eta \in \mathbb{R}$ with $|\zeta| + |\eta| \neq 0$, the point $(c^T \Lambda c, c^T \Gamma c)$, where c is the eigenvector corresponding to the largest eigenvalue of the matrix $\zeta\Lambda + \eta\Gamma$ such that $c^T c = 1$, is a boundary point of $\mathcal{B}_n$. Thus, in this fashion, the boundary of $\mathcal{B}_n$ can be parametrized by ζ and η.

The pictures in Figure 1 were obtained for us by Mr Kok Hoong Wai of the Centre for Wavelets, Approximation and Information Processing, National University of Singapore, using this method. These pictures suggest that the boundary curve can be described analytically for any n. We will discuss the case $n = 2$ in detail and make some remarks about the general case which may be helpful to resolve it.

When $n = 2$, the eigenvalues of the 3×3 tridiagonal matrix $\zeta\Lambda + \eta\Gamma$ are given by the cubic equation

$$P(\lambda|\zeta, \eta) := -\lambda^3 + 5\zeta\lambda^2 + \left(\frac{3}{4}\eta^2 - 4\zeta^2\right)\lambda - 2\zeta\eta^2 = 0. \tag{4.30}$$

The general theory says that $\zeta\Lambda + \eta\Gamma$ has three distinct real eigenvalues and so there are three real solutions to (4.30). For $\zeta \neq 0$ and $\eta \neq 0$, the corresponding unique (up to a sign)

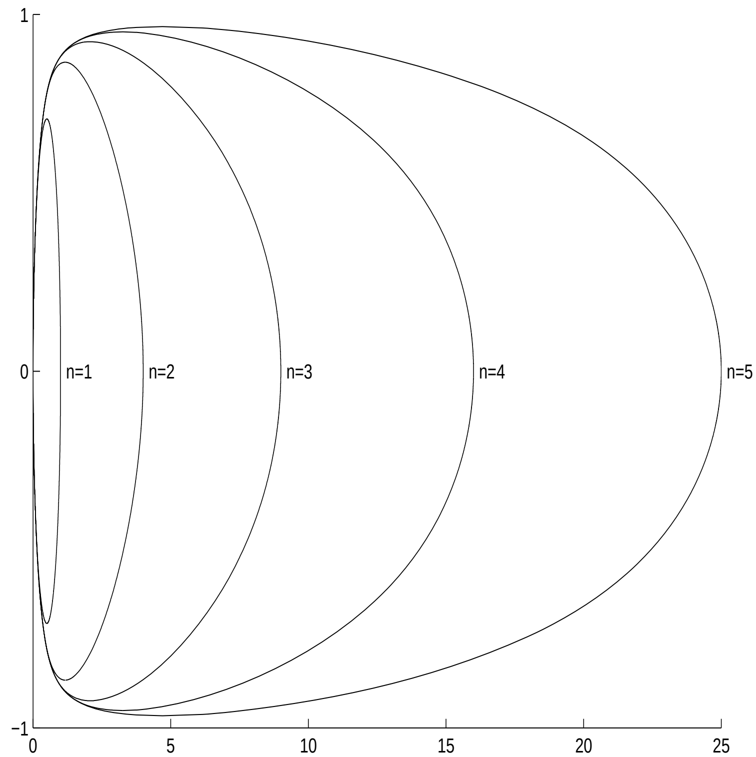

FIGURE 1. Boundary curves of the regions $\mathcal{B}_n$ for $n = 1, \ldots, 5$.

eigenvector $c = (c_1, c_2, c_3)^T$ with $c^T c = 1$ is given by

$$c_1 = \frac{\sqrt{2}\eta(\lambda - 4\zeta)}{\sqrt{Q(\lambda|\zeta, \eta)}}, \qquad c_2 = \frac{2\lambda(\lambda - 4\zeta)}{\sqrt{Q(\lambda|\zeta, \eta)}}, \qquad c_3 = \frac{\eta\lambda}{\sqrt{Q(\lambda|\zeta, \eta)}}$$

where

$$Q(\lambda|\zeta, \eta) := 4\lambda^2(\lambda - 4\zeta)^2 + \eta^2\lambda^2 + 2\eta^2(\lambda - 4\zeta)^2.$$

This eigenvector then produces the point $(u, v) = (c^T \Lambda c, c^T \Gamma c)$ described by the equations

$$u = \frac{4\lambda^2(\lambda - 4\zeta)^2 + 4\eta^2\lambda^2}{Q(\lambda|\zeta, \eta)}, \qquad v = \frac{4\eta\lambda(\lambda - 4\zeta)^2 + 2\eta\lambda^2(\lambda - 4\zeta)}{Q(\lambda|\zeta, \eta)}. \tag{4.31}$$

When λ is the largest eigenvector of $\zeta\Lambda + \eta\Gamma$, (u, v) is a point on the boundary of $\mathcal{B}_2$. Therefore (4.30) and (4.31) provide a parametrization of points on the boundary of $\mathcal{B}_2$, albeit a complex one. Alternatively, these points can also be obtained from the eigenvectors associated with the largest and smallest eigenvalues of the matrix $\Omega := \Lambda + \eta\Gamma$ for $\eta > 0$. First, we consider some exceptional cases. When $\zeta = 0$ and $\eta \neq 0$, (4.30) gives the eigenvalues $\lambda = 0, \pm(\sqrt{3}/2)\eta$. Consequently, we obtain the points $(7/6, \pm\sqrt{3}/2)$, $(8/3, 0)$. For $\eta > 0$, the first two points correspond to the largest and smallest eigenvalues of $\eta\Gamma$ and they lie on the boundary of $\mathcal{B}_2$. (These points can also be identified from the polynomials p_+ and p_- in (4.22) and (4.26) with the aid of the formulas (4.24), (4.25), (4.27) and (4.28).) As for the remaining point $(8/3, 0)$, it corresponds to the middle eigenvalue of $\eta\Gamma$ and it lies in the interior of $\mathcal{B}_2$. Similarly, the case when $\zeta \neq 0$ and $\eta = 0$ gives two boundary points $(0,0)$, $(4,0)$ and one interior point $(1,0)$.

Now, let us return to the situation when $\zeta \neq 0$ and $\eta \neq 0$ and derive two distinct explicit representations of the boundary of $\mathcal{B}_2$. As mentioned above, it suffices to consider the largest and smallest eigenvalues of Ω when $\eta > 0$. Specializing (4.30) to the case that $\zeta = 1$ leads to the equation

$$P(\lambda|1,\eta) = -\lambda^3 + 5\lambda^2 + \left(\frac{3}{4}\eta^2 - 4\right)\lambda - 2\eta^2 = 0. \tag{4.32}$$

Consequently, we conclude that

$$\eta^2 = \frac{4\lambda(\lambda-1)(\lambda-4)}{3\lambda-8} \tag{4.33}$$

and deduce that

$$\lambda \in (-\infty,0) \cup (1,8/3) \cup (4,\infty).$$

Recall that for a fixed $\eta > 0$, (4.32) has three real solutions. By observing the behavior of $P(\cdot|1,\eta)$ at $\lambda = 0,1,4$ and $\pm\infty$, we see that each of the intervals $(-\infty,0)$, $(1,8/3)$, $(4,\infty)$ contains one and only one eigenvalue.

Substituting (4.33) into (4.31) for $\zeta = 1$, we obtain the representation

$$u = \frac{p(\lambda)}{r(\lambda)}, \qquad v^2 = \frac{q(\lambda)}{r^2(\lambda)} \tag{4.34}$$

where

$$p(\lambda) := \lambda(7\lambda^2 - 24\lambda + 32), \qquad q(\lambda) := \lambda(\lambda-1)(\lambda-4)(3\lambda-8)^3$$

and

$$r(\lambda) := 6\lambda^3 - 39\lambda^2 + 80\lambda - 32.$$

On the other hand, given any $\lambda \in (-\infty,0)\cup(1,8/3)\cup(4,\infty)$, we can use (4.33) to find $\eta > 0$ such that (4.32) holds. Furthermore, setting $\lambda = 0,4,\infty,1,8/3$ respectively, (4.34) gives the boundary points $(0,0)$, $(4,0)$, $(7/6,\pm\sqrt{3}/2)$ and interior points $(1,0)$, $(8/3,0)$ identified earlier on. Hence (4.34) taken over all λ in $(-\infty,0]\cup[4,\infty]$ forms a parametric representation of the boundary of $\mathcal{B}_2$. Figure 2 shows the boundary of $\mathcal{B}_2$ plotted using (4.34) with $\lambda \in (-\infty,0]\cup[4,\infty]$. The curve inside $\mathcal{B}_2$ is obtained by varying λ over the interval $[1,8/3]$ which corresponds to the middle eigenvalue of $\Lambda + \eta\Gamma$. This figure was also generated for us by Mr Kok Hoong Wai.

It is also possible to obtain an algebraic equation in u and v^2 for the boundary of $\mathcal{B}_2$ by eliminating λ from (4.34). In this connection, we compute the resultant of the polynomials

$$f(\lambda) := p(\lambda) - ur(\lambda), \qquad g(\lambda) := q(\lambda) - v^2 r^2(\lambda), \qquad \lambda \in \mathbb{C}$$

to obtain the equation

$$\begin{aligned}-2304v^6 - 4384u^2v^4 + 4352uv^4 + 2048v^4 \\ -225u^4v^2 + 192u^3v^2 + 3392u^2v^2 - 2560uv^2 - 1024v^2 \\ -27u^6 + 351u^5 - 1764u^4 + 4256u^3 - 4864u^2 + 2048u = 0.\end{aligned} \tag{4.35}$$

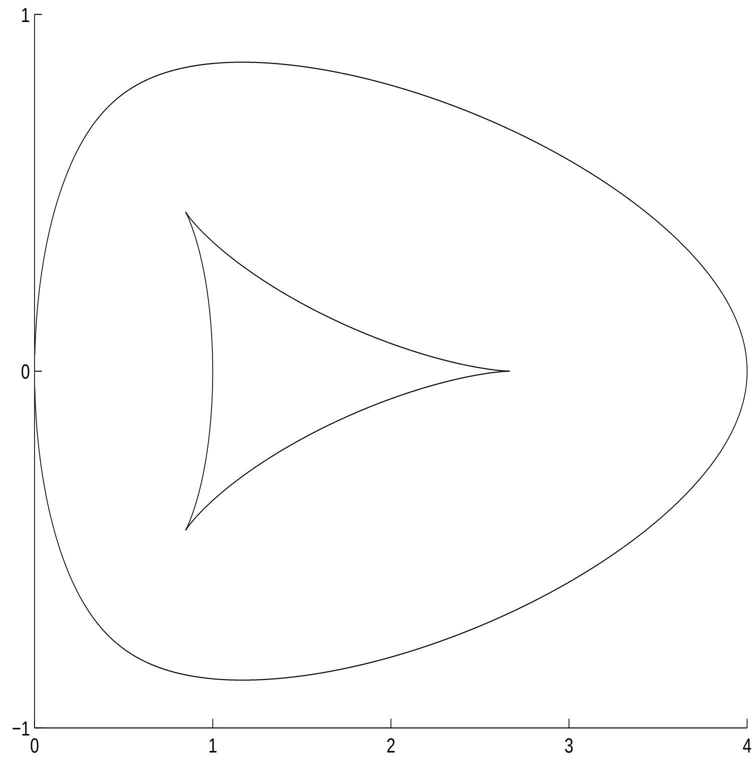

FIGURE 2. Algebraic curve that contains the boundary of the region $\mathcal{B}_2$.

It is easily verified that the points $(0,0)$, $(4,0)$, $(7/6, \pm\sqrt{3}/2)$, $(1,0)$, $(8/3,0)$ also satisfy this equation. Therefore the boundary curve of $\mathcal{B}_2$ satisfies (4.35). In addition, we have confirmed computationally that the graph of (4.35) is the same as that of (4.34) shown in Figure 2.

For the general case, we use the $(n+1)\times(n+1)$ matrix $\Omega := \Lambda + \eta\Gamma$ for $\eta > 0$ and let $P_j(\lambda|\eta)$, $j = 0, \dots, n$, be the determinant of the $(j+1)$-th principal minor of $\Omega - \lambda I$. Since $\Omega - \lambda I$ is a tridiagonal matrix, it follows that

$$P_j(\lambda|\eta) = (j^2 - \lambda)P_{j-1}(\lambda|\eta) - \frac{1}{4}\eta^2 P_{j-2}(\lambda|\eta), \quad j = 1, \dots, n$$

where we set $P_0(\lambda|\eta) := -\lambda$, $P_{-1}(\lambda|\eta) := 2$. Using this recursion it follows that

$$P_n(\lambda|\eta) = \sum_{j=0}^{\lceil n/2 \rceil} \eta^{2j} Q_{jn}(\lambda)$$

where $\deg Q_{jn} \le n + 1 - 2j$, $j = 0, \dots, \lceil n/2 \rceil$ with $\lceil n/2 \rceil :=$ "least integer $\ge n/2$".

If λ is an eigenvalue of Ω, that is,

$$P_n(\lambda|\eta) = 0, \tag{4.36}$$

then some computations show that the corresponding unique (up to a constant) eigenvector $b = (a_0, \sqrt{2}a_1, \ldots, \sqrt{2}a_n)^T$ is given by the formula

$$a_j = \frac{(-\eta/2)^{n-j} P_{j-1}(\lambda|\eta)}{\sqrt{2}P_{n-1}(\lambda|\eta)}, \quad j = 0, \ldots, n.$$

Therefore, if λ is either the largest or smallest eigenvalue of Ω, the point (u, v) given by

$$u = \frac{b^T \Lambda b}{b^T b} = \frac{\sum_{j=1}^{n} j^2 (\eta/2)^{2n-2j} P_{j-1}^2(\lambda|\eta)}{\sum_{j=0}^{n}{}' (\eta/2)^{2n-2j} P_{j-1}^2(\lambda|\eta)} \tag{4.37}$$

and

$$v = \frac{b^T \Gamma b}{b^T b} = -\frac{\sum_{j=0}^{n-1} (\eta/2)^{2n-2j-1} P_{j-1}(\lambda|\eta) P_j(\lambda|\eta)}{\sum_{j=1}^{n}{}' (\eta/2)^{2n-2j} P_{j-1}^2(\lambda|\eta)} \tag{4.38}$$

is on the boundary of $\mathcal{B}_n$ and conversely every boundary point has this form.

Note that u, v^2 and $P_n(\lambda|\eta)$ are only functions of η^2. Following the case $n = 2$, we can first eliminate η^2 from (4.36) and (4.37) and then from (4.36) and (4.38) to obtain two polynomial equations relating u, v^2, λ. Then additional elimination of λ from these two equations produce an algebraic curve in u, v^2 which contains the boundary of $\mathcal{B}_n$.

We conjecture that the boundary curves also have a rational parametrization over the set $(-\infty, 0] \cup [n^2, \infty]$. In fact, if λ_0, λ_n are the smallest and largest eigenvalues of Ω for $\eta > 0$, then

$$\lambda_0 \le \min\left\{0, \frac{1}{2} - \frac{\eta}{\sqrt{2}}\right\}, \qquad \lambda_n \ge \max\left\{n^2, n^2 - n + \frac{\eta + 1}{2}\right\}.$$

Hence by the continuity of λ_0 and λ_n in η we see that the range of these functions over $\mathbb{R}_+$ are respectively $(-\infty, 0]$ and $[n^2, \infty)$.

References

[1] J. J. Benedetto, Frame decompositions, sampling, and uncertainty principle inequalities, in Wavelets: Mathematics and Applications, J. J. Benedetto and M. W. Frazier, eds., CRC Press, 1994, pp. 247–304.

[2] E. Breitenberger, Uncertainty measures and uncertainty relations for angle observables, Found. Phys., 15(1985), pp. 353–364.

[3] L. Brickman, On the field of values of a matrix, Proc. Amer. Math. Soc., 12(1961), pp. 61–66.

[4] E. A. Coddington and N. Levinson, Theory of Ordinary Differential Equations, McGraw-Hill, 1955.

[5] H. Dym and H. P. McKean, Fourier Series and Integrals, Academic Press, 1972.

[6] A. Edelman and C. A. Micchelli, Admissible slopes for monotone and convex interpolation, Numer. Math., 51(1987), pp. 441–458.

[7] C. L. Fefferman, The uncertainty principle, Bull. Amer. Math. Soc., 9(1983), pp. 129–206.

[8] G. B. Folland, Harmonic Analysis in Phase Space, Annals of Math. Studies, no. 122, Princeton University Press, 1989.

[9] G. B. Folland and A. Sitaram, The uncertainty principle: a mathematical survey, J. Fourier Anal. Appl., 3(1997), pp. 207–238.
[10] S. S. Goh and C. H. Yeo, Uncertainty products of local periodic wavelets, Adv. Comput. Math., 13(2000), pp. 319–333.
[11] P. R. Halmos, A Hilbert Space Problem Book, 2nd edition, Springer-Verlag, 1982.
[12] V. Havin and B. Jöricke, The Uncertainty Principle in Harmonic Analysis, Springer-Verlag, 1994.
[13] R. A. Horn and C. R. Johnson, Matrix Analysis, Cambridge University Press, 1985.
[14] S. Karlin, Total Positivity, vol. 1, Stanford University Press, 1968.
[15] J. R. Klauder and E. C. G. Sudarshan, Fundamentals of Quantum Optics, W. A. Benjamin, 1968.
[16] E. Kreyszig, Introductory Functional Analysis with Applications, John Wiley & Sons, 1978.
[17] H. J. Landau, An overview of time and frequency limiting, in Fourier Techniques and Applications, J. F. Price, ed., Plenum Press, 1985, pp. 201-220.
[18] H. J. Landau and H. O. Pollak, Prolate spheroidal wave functions, Fourier analysis and uncertainty–II, Bell System Tech. J., 40(1961), pp. 65–84.
[19] H. J. Landau and H. O. Pollak, Prolate spheroidal wave functions, Fourier analysis and uncertainty–III: the dimension of the space of essentially time- and band-limited signals, Bell System Tech. J., 41(1962), pp. 1295–1336.
[20] N. W. McLachlan, Theory and Application of Mathieu Functions, Oxford University Press, 1951.
[21] A. A. Melkman and C. A. Micchelli, Optimal estimation of linear operators in Hilbert spaces from inaccurate data, SIAM J. Numer. Anal., 16(1979), pp. 87–105.
[22] C. A. Micchelli, Optimal estimation of linear operators from inaccurate data: a second look, Numer. Algorithms, 5(1993), pp. 375–390.
[23] F. J. Narcowich and J. D. Ward, Wavelets associated with periodic basis functions, Appl. Comput. Harmon. Anal., 3(1996), pp. 40–56.
[24] B. N. Parlett, The Symmetric Eigenvalue Problem, Prentice Hall, 1980.
[25] A. Pinkus, n-widths in Approximation Theory, Springer-Verlag, 1982.
[26] J. Prestin and E. Quak, Time frequency localization of trigonometric Hermite operators, in Approximation Theory VIII, Vol. 2: Wavelets and Multilevel Approximation, C. K. Chui and L. L. Schumaker, eds., World Scientific, 1995, pp. 343–350.
[27] J. Prestin and E. Quak, Optimal functions for a periodic uncertainty principle and multiresolution analysis, Proc. Edinburgh Math. Soc., 42(1999), pp. 225–242.
[28] J. F. Price, Uncertainty principles and sampling theorems, in Fourier Techniques and Applications, J. F. Price, ed., Plenum Press, 1985, pp. 25–44.
[29] K. Selig, Trigonometric wavelets and the uncertainty principle, in Approximation Theory, Math. Research, vol. 86, M. W. Müller, M. Felten and D. H. Mache, eds., Akademie Verlag, 1995, pp. 293–304.
[30] D. Slepian, Prolate spheroidal wave functions, Fourier analysis and uncertainty–IV: extensions to many dimensions; generalized prolate spheroidal functions, Bell System Tech. J., 43(1964), pp. 3009–3057.
[31] D. Slepian, On bandwidth, Proc. IEEE, 64(1976), pp. 292–300.
[32] D. Slepian, Prolate spheroidal wave functions, Fourier analysis and uncertainty–V: the discrete case, Bell System Tech. J., 57(1978), pp. 1371–1429.
[33] D. Slepian, Some comments on Fourier analysis, uncertainty and modeling, SIAM Rev., 25(1983), pp. 379–393.
[34] D. Slepian and H. O. Pollak, Prolate spheroidal wave functions, Fourier analysis and uncertainty–I, Bell System Tech. J., 40(1961), pp. 43–64.
[35] G. Szegö, Orthogonal Polynomials, 4th edition, Amer. Math. Soc. Coll. Publ., vol. 23, American Mathematical Society, 1975.
[36] H. Weyl, The Theory of Groups and Quantum Mechanics, 2nd edition, English Translation by H. P. Robertson, E. P. Dutton & Co., 1931.