Hermite-Gauss Quadrature with Generalized Hermite Weight Functions and Small Sample Sets for Sparse Polynomials

by
Brian-Tinh Duc Vu

A thesis submitted to the Department of Physics,

College of Natural Sciences and Mathematics

in partial fulfillment of the requirements for the degree of

Bachelor of Science
in Physics

Chair of Committee: Donna Stokes

Committee Member: Bernhard G. Bodmann

Committee Member: Donald J. Kouri

Committee Member: Lowell Wood

University of Houston
April 2020

DEDICATION/EPIGRAPH

Five years ago, if you had told me that I would be conducting research in advanced topics in mathematics, I would not have believed it. Among my closest friends and advisors it is no secret that I was once failing Advanced Placement Calculus, and that coming out of high school I doubted my ability to thrive in the quantitative sciences. I still have some of my old tests from that class, with scores like 59/100 which cause my face to turn red with shame.

But the moral of this story is that I kept the tests and, looking back on them from time to time, I faced the shame head-on. I cannot overstate how difficult it was to face my fears and embark on projects where the work was so theoretical and high-level. Today I still feel like I am able to contribute very little to what is being done by the giants in my field. But the very existence of this thesis proves that I am indeed moving forward on my journey to becoming a good physicist and mathematician, and that knowledge gives me enough strength to get out of bed and grapple with new, difficult ideas each and every day.

There are so many people to thank for my journey thus far at the University of Houston, including my friends and family, my professors, and Dr. Bodmann. But the person I want to honor the most is Dr. Kouri, who, on one fateful summer day while I was in high school, asked me to come to the University of Houston and join his research group. Dr. Kouri set a high bar for my undergraduate career by informing me of the Barry Goldwater Scholarship and telling me I had a good shot at receiving it if I worked hard enough. That goal guided my work ethic for the duration of my time at UH, and I have spent every moment of study aiming to become someone worthy of that award. As a result, I have received several accolades that not only include but also surpass the honor of being a Goldwater Scholar. I am thankful to Dr. Kouri for his mentorship and hope that my future PhD advisor will be just as inspirational and motivating as he was to me.

Brian-Tinh Duc Vu

April 1, 2020

ACKNOWLEDGMENTS

This research has been initiated through the support of the Office of Undergraduate Research at the University of Houston. We thank them for their financial support through the Summer Undergraduate Research Fellowship and Provost's Undergraduate Research Scholarship programs. B.G.B. was supported in part by the NSF grant DMS-1715735.

ABSTRACT

This thesis derives a Gaussian quadrature rule from a complete set of orthogonal lacunary polynomials. The resulting quadrature formula is exact for polynomials whose even part skips powers, with a set of sample values that is much smaller than the degree. The weight for these quadratures is a generalized Gaussian, whose negative logarithm is an even monomial; the powers of this monomial make up the even part of the polynomial to be integrated. We first present Rodrigues formulas for generalized Hermite polynomials (GHPs) that are complete and orthogonal with respect to the generalized Gaussian. From the Rodrigues formula for even GHPs we establish a three-term recursion relation and find the normalization constants. We present a slight modification to the Christoffel-Darboux identity and the Lagrange interpolation polynomials, and proceed to derive the roots, weights, and estimate of the error for the generalized Hermite-Gauss quadrature rule applied to sufficiently smooth functions. We illustrate the quadrature rule by applying it to two examples. Finally, we apply a major result from compressive sensing relating a matrix's coherence and sparse recovery guarantees to the quadrature setting.

TABLE OF CONTENTS

LIST OF FIGURES

1 Introduction

It is a remarkable fact in the theory of orthogonal polynomials that *interpolating* a polynomial of degree d on the real line requires at least $d+1$ sample values, while *integrating* it against a sufficiently regular weight function (or measure) can be done with about half the samples. In addition, the quadrature formulas that give the precise value of the integral for degree-limited polynomials can also serve as an estimate for integrals of a larger class of functions that are well approximated by their interpolants [20].

In general, a quadrature rule expresses an integral of a polynomial of the form $\int_a^b p(t)w(t)dt$ using a set of sampling points $\{t_j\}_{j=1}^k$ and their respective scaling factors $\{w_j\}_{j=1}^k$, such that

$$\int_a^b p(t)w(t)dt = \sum_{j=1}^k w_j p(t_j) \tag{1.1}$$

[12, 13, 14, 6]. Here, w is a weight function with sufficiently many moments, and we include the case of improper integrals with $a = -\infty$ and $b = +\infty$.

In this thesis we continue the tradition of using orthogonal polynomials to derive quadrature rules. The main result developed here shows that the number of sample points can be chosen far smaller than half the degree of the polynomial in a special case: if the non-zero coefficients of the polynomial occur at exponents that are odd or multiples of some even integer, so $p(t) = \sum_{l=0}^k c_{2ln}t^{2ln} + \sum_{l=0}^{kn} c_{2l+1}t^{2l+1}$ with $n \in \mathbb{N}$. In short, the even part of the polynomial is lacunary, it skips powers with a step size of $2n$. The weight for the quadrature formula is chosen to be a generalized Gaussian, $w(t) = e^{-t^{2n}/n}$. The development of a quadrature rule for the generalized Gaussian using generalized Hermite polynomials was inspired by [19], in which there is a discussion for finding the sampling nodes and weights corresponding to a weight function $w(x) = e^{-V(x)}$, where $V(x) = x^{2m} + \mathcal{O}(2m-1)$.

For the derivation of the quadrature rule we introduce generalized Hermite polynomials (GHPs). These polynomials are not the generalized Hermite polynomials of [5], which satisfy only a partial orthogonality condition. Like the polynomials to be integrated, the GHPs skip powers of t. That is, for some generalized Hermite polynomial indexed by k, the expanded form is $H_{2k}^{(n)}(t) = \sum_{l=0}^{k} a_l^{(k)} t^{2ln}$ or $H_{2k+1}^{(n)}(t) = \sum_{l=1}^{k+1} b_l^{(k)} t^{2ln-1}$ depending on whether the polynomial is even or odd (further details on GHPs are given in the following section). The absence of some of the exponents allows the sequence of polynomials to be orthogonal with respect to the chosen weight function.

Some treatments of weight functions start with computing entries of its Jacobi operators and subsequently finding the sampling nodes and weights [21]. From the GHPs we find the sampling nodes and weights using a method similar to that presented in [14]. We begin with the three-term recursion relation of the orthogonal polynomials. From the Rodrigues formula definition for the GHPs we can determine their recursion relation and subsequently compute the entries of the Jacobi matrix using exact expressions.

This thesis is organized as follows: After introducing the GHPs in Section 2, we derive two separate three-term recursion relations and normalization constants for the even and odd GHPs. We select sampling points for the quadrature rule based on the real and positive roots of the even GHP. We modify the Christoffel-Darboux identity and the Lagrange interpolation polynomials to accommodate the skipping of powers of t, and from these we derive expressions for the quadrature rule weights. We subsequently obtain an expression for the quadrature error and identify a class of functions for which the error goes to zero as $k \to \infty$. We also present two examples. The first example demonstrates that by applying the quadrature rule to a monomial we can calculate exact values of the gamma function for certain arguments. The second example numerically calculates the error bound of an integral approximation as $k \to \infty$. Finally, we apply techniques of compressive sensing to recover a linear combination of Hermite polynomials f when the sample nodes are not identical to roots of a Hermite polynomial. We address the requisite sparsity of solutions necessary for guaranteed sparse recovery.

2 Generalized Hermite polynomials (GHP)

The Hermite polynomials with the normalization constants appearing in the physics literature are generated by the Rodrigues formula [1]

$$H_k(t) = (-1)^k e^{t^2} \left(\frac{d}{dt}\right)^k e^{-t^2}.$$

This succinct formula can be expressed as two different formulas by considering the even and odd Hermite polynomials separately and noting that $\frac{d}{dt}e^{-t^2} = -2te^{-t^2}$:

$$H_{2k}(t) = e^{t^2} \left(\frac{d^2}{dt^2}\right)^k e^{-t^2}, \quad H_{2k+1}(t) = e^{t^2} \left(\frac{d^2}{dt^2}\right)^k 2te^{-t^2}.$$

With these formulas, we have expressed the Hermite polynomials in terms of the Laplacian instead of the derivative operator.

In this thesis, we use a singular Laplacian appearing in a Fourier-like transform studied earlier [23] to derive another class of orthonormal polynomials. In the context of the Fourier transform, the differential operator $\frac{d^2}{dt^2}$ is diagonalized by the Fourier transform. This is a direct consequence of the Fourier kernel $\frac{1}{\sqrt{2\pi}}e^{-i\omega t}$ being an eigenfunction of the differential operator $\frac{d^2}{dt^2}$. In previous work (*ibid.*), it was shown that the the operator $D_n = -\frac{d}{dt}\frac{1}{t^{2n-2}}\frac{d}{dt}$ is diagonalized by a generalized Fourier transform. This motivates us to define

$$H_{2k}^{(n)}(t) = (-1)^k e^{\frac{t^{2n}}{n}} D_n^k e^{-\frac{t^{2n}}{n}} \tag{2.1}$$

and

$$H_{2k+1}^{(n)}(t) = (-1)^k e^{\frac{t^{2n}}{n}} D_n^k (2t^{2n-1} e^{-\frac{t^{2n}}{n}}) \tag{2.2}$$

where the $H_{2k}^{(n)}$ and the $H_{2k+1}^{(n)}$ are the even and odd generalized Hermite polynomials (GHPs), respectively. We first claim that these formulas, in fact, do yield polynomials, and that these polynomials are of a particular form.

Theorem 2.1. *Define $H_{2k}^{(n)}$ by formula (2.1), where $n \in \mathbb{N}$ and k is a nonnegative integer, then $H_{2k}^{(n)}$ is a polynomial of degree $2kn$ that consists of only terms with powers $2ln$, where $l \in \{0, 1, \ldots, k\}$.*

Proof. We prove by induction, starting with $H_0^{(n)}(t) = 1$. Now suppose that $H_{2k}^{(n)}(t) = a_0 + a_1 t^{2n} + \ldots + a_k t^{2kn}$. From (2.1), it follows that $D_n^k e^{-\frac{t^{2n}}{n}} = (-1)^k (a_0 + a_1 t^{2n} + \ldots + a_k t^{2kn}) e^{-\frac{t^{2n}}{n}}$. Applying D_n to both sides yields another polynomial in powers of t^{2n} multiplying $e^{-\frac{t^{2n}}{n}}$ since

$$D_n e^{-\frac{t^{2n}}{n}} = (2 - 4t^{2n}) e^{-\frac{t^{2n}}{n}} \tag{2.3}$$

and

$$D_n(t^{2ln} e^{-\frac{t^{2n}}{n}}) = [c_1 t^{2(l-1)n} + c_2 t^{2ln} + c_3 t^{2ln}(2 - 4t^{2n})] e^{-\frac{t^{2n}}{n}}. \tag{2.4}$$

This is a polynomial multiplying the n-Gaussian ($e^{-\frac{t^{2n}}{n}}$) and combining terms gives another polynomial times the n-Gaussian. Furthermore, since $H_0^{(n)}$ is a constant, the first application of D_n yields a polynomial of degree $2n$, and each subsequent application of D_n increases the degree of the polynomial by $2n$ while maintaining terms of only t^{2ln}, by (2.4). Therefore, $H_{2k}^{(n)}(t)$ is a polynomial of degree $2kn$ consisting of a linear combination of monomials t^{2ln}, and the theorem is proved. □

Theorem 2.2. *Define $H_{2k+1}^{(n)}$ by formula (2.2), where $n \in \mathbb{N}$ and k is a nonnegative integer, then $H_{2k+1}^{(n)}$ is a polynomial of degree $2(k+1)n - 1$ that consists of terms with powers t^{2ln-1}, where $l \in \{1, \ldots, k+1\}$.*

Proof. Again, we prove by induction. Clearly, $H_1^{(n)}(t) = 2t^{2n-1}$ is a polynomial of the claimed form. Suppose that $H_{2k+1}^{(n)}$ is a polynomial with terms of t^{2ln-1}. Then $D_n^k(2t^{2n-1} e^{-\frac{t^{2n}}{n}})$ is a polynomial with terms of t^{2ln-1} multiplying the n-Gaussian. Applying D_n to this yields another polynomial multiplying the n-Gaussian with terms of the form t^{2ln-1} of degree greater than that of the previous polynomial by $2n$. Since $H_1^{(n)}$ is a polynomial of degree $2n - 1$, $H_{2k+1}^{(n)}$ is a polynomial of degree $2(k+1)n - 1$, proving the theorem. □

In the following subsection, we will show that the GHPs are orthogonal and complete. Afterwards we will show that two separate recursion relations exist for the generalized Hermites: one for the even polynomials and one for the odds. The recursion relation will allow for stable construction of the generalized Hermites for numerical purposes.

2.1 Orthogonality and completeness

Orthogonality for the GHPs is very similar to that of the standard Hermite polynomials, with the exception that the weight function is the appropriate generalized Gaussian for the GHPs, as opposed to the standard Gaussian for Hermite polynomials.

Theorem 2.3. *If m and m' are nonnegative integers such that $m \neq m'$, then*

$$\int_{-\infty}^{\infty} \left(H_m^{(n)}(t)\right)\left(H_{m'}^{(n)}(t)\right)e^{-\frac{t^{2n}}{n}}\,dt = 0.$$

Proof. We already know that the even and the odd GHPs are mutually orthogonal with respect to the weight function $e^{-\frac{t^{2n}}{n}}$, because with an even polynomial, an odd polynomial, and an even weight function, the integrand would be overall odd and vanish. Now consider the case of even GHP, where $m = 2k$ and $m' = 2k'$. Without loss of generality, assume $k' > k$. Then we have that

$$\begin{aligned}\int_{-\infty}^{\infty} &\left(H_{2k}^{(n)}(t)e^{-\frac{t^{2n}}{2n}}\right)\left(H_{2k'}^{(n)}(t)e^{-\frac{t^{2n}}{2n}}\right)dt \\ &= \int_{-\infty}^{\infty} \left(H_{2k}^{(n)}(t)e^{-\frac{t^{2n}}{2n}}\right)\left((-1)^{k'}e^{\frac{t^{2n}}{n}}D_n^{k'}e^{-\frac{t^{2n}}{n}}\right)e^{-\frac{t^{2n}}{2n}}\,dt \qquad (2.5)\\ &= (-1)^{k'}\int_{-\infty}^{\infty} H_{2k}^{(n)}(t)D_n^{k'}e^{-\frac{t^{2n}}{n}}\,dt. \qquad (2.6)\end{aligned}$$

When D_n acts on $H_{2k}^{(n)}(t)$ (since $H_{2k}^{(n)}(t)$ is a polynomial), it decreases the degree by t^{2n}. By successive integration by parts, this becomes

$$\begin{aligned}\int_{-\infty}^{\infty} &\left(H_{2k}^{(n)}(t)e^{-\frac{t^{2n}}{2n}}\right)\left(H_{2k'}^{(n)}(t)e^{-\frac{t^{2n}}{2n}}\right)dt \\ &= (-1)^{k'}\int_{-\infty}^{\infty} D_n^k\left(H_{2k}^{(n)}(t)\right)D_n^{k'-k}e^{-\frac{t^{2n}}{n}}\,dt.\end{aligned} \qquad (2.7)$$

$D_n^k H_{2k}^{(n)}(t)$ is a constant (call it C), so we need only to show that

$\int_{-\infty}^{\infty} D_n^{k'-k} e^{-\frac{t^{2n}}{n}} dt = 0$. Since we have that $\frac{d}{dt}\frac{1}{t^{2n-2}}\frac{d}{dt} D_n^{k'-k-1} e^{-\frac{t^{2n}}{n}}$ is integrable, we get

$$\int_{-\infty}^{\infty} \left(H_{2k}^{(n)}(t) e^{-\frac{t^{2n}}{2n}}\right)\left(H_{2k'}^{(n)}(t) e^{-\frac{t^{2n}}{2n}}\right) dt$$

$$= -(-1)^k C \int_{-\infty}^{\infty} \frac{d}{dt}\frac{1}{t^{2n-2}}\frac{d}{dt} D_n^{k'-k-1} e^{-\frac{t^{2n}}{n}} dt \tag{2.8}$$

$$= (-1)^{k+1} C \frac{1}{t^{2n-2}}\frac{d}{dt} D_n^{k'-k-1} e^{-\frac{t^{2n}}{n}} \Big|_{-\infty}^{\infty}$$

$$= 0. \tag{2.9}$$

Thus if $k \neq k'$, we get that the integral vanishes and so the even GHP are orthogonal. An identical analysis holds for the odd GHP. $\square$

To show that the GHP are complete, we use a standard approximation argument. The approximation properties of lacunary polynomials have already been explored, but mostly for uniform approximation of continuous functions on compact sets [17], see also [2]. Here, we treat the completeness in the Hilbert space $L^2(\mathbb{R}, e^{-t^{2n}/n} dt)$.

Theorem 2.4. *The polynomials $\{H_m^{(n)}\}_{m=0}^{\infty}$ form a complete set in $L^2(\mathbb{R}, e^{-t^{2n}/n} dt)$.*

Proof. It is enough to show that the closed linear span of the GHP contains each constant function, any even square-integrable function whose integral is zero, and any odd square-integrable function. Furthermore, after treating the constant function, we can specialize to even and odd functions whose support is compact and does not contain 0. In that case, an even function f is of the form $f(t) = t^n g(t^{2n})$ and an odd function f is of the form $f(t) = t^{n-1} g(t^{2n})$ where g has compact support excluding 0. After a change of coordinates, the result then follows from polynomials being dense in $L^2(\mathbb{R}^+, e^{-t} dt)$. $\square$

2.2 Three-term recursion relations

Because there are two Rodrigues formulas (one for the even polynomials and one for the odd polynomials), it follows that there must also be two separate recursion relations for the odd and even GHPs. We begin with the relation between the even polynomials.

Theorem 2.5. *Let k and n be nonnegative and positive integers, respectively, as they are in (2.1) and (2.2). The three-term recursion relation between the even GHPs is*

$$H_{2k+2}^{(n)}(t) = (\alpha_k t^{2n} + \beta_k) H_{2k}^{(n)}(t) + \gamma_k H_{2k-2}^{(n)}.$$

with the coefficients

$$\begin{aligned} \alpha_k &= 4, \\ \beta_k &= -8kn - 2, \\ \gamma_k &= -16k^2n^2 + 16kn^2 - 8kn. \end{aligned}$$

Proof. Before beginning the proof, it is convenient to evaluate the following expression

$$\begin{aligned} D_n\left(t^{2ln} e^{-\frac{t^{2n}}{n}}\right) &= [(-4l^2n^2 + 4ln^2 - 2ln)t^{2(l-1)n} \qquad (2.10) \\ &+ (8ln + 2)t^{2ln} - 4t^{2(l+1)n}]e^{-\frac{t^{2n}}{n}} \\ &= [c_-(l)t^{2(l-1)n} + c_0(l)t^{2ln} + c_+(l)t^{2(l+1)n}]e^{-\frac{t^{2n}}{n}}, \qquad (2.11) \end{aligned}$$

and introduce the notation

$$c_-(l) = -4l^2n^2 + 4ln^2 - 2ln, \tag{2.12}$$

$$c_0(l) = 8ln + 2, \tag{2.13}$$

$$c_+(l) = -4. \tag{2.14}$$

Let $H_{2k}^{(n)}$ be some arbitrary even GHP indexed by some nonnegative integer k, as defined by (2.1), such that when $n = 1$, the polynomial becomes the standard even Hermite polynomial H_{2k}. As shown before, $H_{2k}^{(n)}(t)$ is a linear combination of terms of the form t^{2ln}, where $l \in 0, 1, 2, ..., k$.

We write $H_{2k}^{(n)}(t)$ in terms of its coefficients a_l:

$$H_{2k}^{(n)}(t) = (-1)^k e^{\frac{t^{2n}}{n}} D_n^k e^{-\frac{t^{2n}}{n}} = \sum_{l=0}^{k} a_l t^{2ln}. \tag{2.15}$$

By rearrangement, (2.15) implies that $D_n^k(e^{-\frac{t^{2n}}{n}}) = (-1)^k \Big(\sum_{l=0}^{k} a_l t^{2ln} \Big) e^{-\frac{t^{2n}}{n}}$. We then write the next two generalized Hermite polynomials by successive operations on $H_{2k}^{(n)}(t)e^{-\frac{t^{2n}}{n}}$ by D_n:

$$\begin{aligned} H_{2k+2}^{(n)}(t) &= (-1)^{k+1} e^{\frac{t^{2n}}{n}} D_n^{k+1} e^{-\frac{t^{2n}}{n}} && (2.16)\\ &= (-1) e^{\frac{t^{2n}}{n}} D_n \Big[\Big(\sum_{l=0}^{k} a_l t^{2ln} \Big) e^{-\frac{t^{2n}}{n}} \Big] && (2.17)\\ &= (-1) \Big[\sum_{l=0}^{k} a_l [c_-(l) t^{2(l-1)n} + c_0(l) t^{2ln} + c_+ t^{2(l+1)n}] \Big]. && (2.18) \end{aligned}$$

$$\begin{aligned} H_{2k+4}^{(n)}(t) &= (-1)^{k+2} e^{\frac{t^{2n}}{n}} D_n^{k+2} e^{-\frac{t^{2n}}{n}} && (2.19)\\ &= e^{\frac{t^{2n}}{n}} D_n \Big[\Big(\sum_{l=0}^{k} a_l [c_-(l) t^{2(l-1)n} && (2.20)\\ &\qquad + c_0(l) t^{2ln} + c_+ t^{2(l+1)n}] \Big) e^{-\frac{t^{2n}}{n}} \Big] \\ &= \sum_{l=0}^{k} a_l \Big[c_-(l-1) c_-(l) t^{2(l-2)n} && (2.21)\\ &\qquad + c_-(l) \Big(c_0(l-1) + c_0(l) \Big) t^{2(l-1)n} \\ &\qquad + \Big(c_-(l) c_+ + [c_0(l)]^2 + c_-(l+1) c_+ \Big) t^{2ln} \\ &\qquad + c_+ \Big(c_0(l) + c_0(l+1) \Big) t^{2(l+1)n} \\ &\qquad + c_+^2 t^{2(l+2)n} \Big]. \end{aligned}$$

Next, we see that a three-term recursion relation exists between the even GHP, and that relation takes the form

$$H_{2k+4}^{(n)}(t) = (\alpha t^{2n} + \beta) H_{2k+2}^{(n)}(t) + \gamma H_{2k}^{(n)}(t) \tag{2.22}$$

with $\alpha, \beta, \gamma \in \mathbb{R}$. By substituting the polynomial forms of the GHP into the recursion relation and combining like powers of t, we find that:

$$\alpha = 4, \tag{2.23}$$

$$\beta = -8kn - 8n - 2, \tag{2.24}$$

$$\gamma = (16k^2n^2 - 48kn^2 + 8kn - 8n) + \frac{a_{k-1}}{a_k}(32n). \tag{2.25}$$

Recall that a_k and a_{k-1} are the coefficients of the highest and second-highest powers, respectively, in the polynomial $H_{2k}^{(n)}$. To further simplify the expression for γ, we define b_{k+1} and b_k to be the first and second coefficients of $H_{2k+2}^{(n)}$ and derive a relation between $\frac{b_k}{b_{k+1}}$ and $\frac{a_{k-1}}{a_k}$. Careful examination of (2.18), the polynomial form of $H_{2k+2}^{(n)}$, shows that

$$b_{k+1} = -c_+ a_k, \tag{2.26}$$

$$b_k = -c_+ a_{k-1} - c_0(k) a_k. \tag{2.27}$$

Define the ratio of the second and the first coefficients of a GHP $H_{2k}^{(n)}(t)$ to be Q_k (i.e. $Q_k = \frac{a_{k-1}}{a_k}$). The relation between the ratios then becomes

$$Q_{k+1} = Q_k + \frac{c_0(k)}{c_+} \tag{2.28}$$

$$= Q_k - 2kn - \frac{1}{2}. \tag{2.29}$$

Recognizing that $Q_0 = 0$ (since $H_0^{(n)}(t) = 1$), one can show that

$$Q_k = -\sum_{j=0}^{k-1} \left(2jn + \frac{1}{2}\right) \tag{2.30}$$

$$= -k(k-1)n - \frac{k}{2}, \tag{2.31}$$

and so

$$\gamma = (16k^2n^2 - 48kn^2 + 8kn - 8n) + Q_k(32n) \tag{2.32}$$

$$= -16k^2n^2 - 16kn^2 - 8kn - 8n. \tag{2.33}$$

Finally, replacing k by $k-1$ in (2.22), (2.24), and (2.33) proves the theorem. □

By the same method, one can derive the recursion relation between the odd GHP. Without proof, the odd GHP recursion relation is

$$H_{2k+3}^{(n)}(t) = (\alpha_k' t^{2n} + \beta_k') H_{2k+1}^{(n)}(t) + \gamma_k' H_{2k-1}^{(n)} \tag{2.34}$$

where the coefficients are defined as follows:

$$\alpha_k' = 4, \tag{2.35}$$

$$\beta_k' = -8kn - 8n + 2, \tag{2.36}$$

$$\gamma_k' = -16k^2n^2 - 16kn^2 + 8kn. \tag{2.37}$$

This recursion relation makes explicit computation of the polynomials straightforward and efficient. The polynomials can be stored as coefficients in an array of a computer program, and only addition, multiplication, and shift operations on the arrays are necessary to generate polynomials up to a desired index.

2.3 Normalization constants

While the GHPs are orthogonal with respect to the weight function $w^{(n)}(t) = e^{-\frac{t^{2n}}{n}}$, they are not orthonormal. We compute the required normalization constants.

Theorem 2.6. *Let $H^{(n)}_{2k}$ be as defined above and*

$$h^{(n)}_{2k}(t) = A_k H^{(n)}_{2k}(t)$$

with

$$A_k = \Big[\Big(2^{3k}k!n^k \prod_{j=1}^{k-1}[2jn+1]\Big)\Big(n^{\frac{1}{2n}-1}\Gamma(\frac{1}{2n})\Big)\Big]^{-\frac{1}{2}},$$

then $\{h^{(n)}_{2k}\}_{k=0}^{\infty}$ forms an orthonormal sequence in $L^2(\mathbb{R}, e^{-t^{2n}/n}dt)$.

Proof. All that remains is to fix the normalization of the GHPs. The derivation of the even polynomials' normalization constant begins the same way as the proof for the orthogonality of the GHP, only we let $m = 2k = 2k' = m'$. Split the integrand into the product of two generalized Hermite functions

$$\begin{aligned}
&\int_{-\infty}^{\infty}\Big(H^{(n)}_{2k}(t)e^{-\frac{t^{2n}}{2n}}\Big)\Big(H^{(n)}_{2k}(t)e^{-\frac{t^{2n}}{2n}}\Big)dt \\
&\quad = \int_{-\infty}^{\infty} H^{(n)}_{2k}(t)e^{-\frac{t^{2n}}{2n}}\Big((-1)^k e^{\frac{t^{2n}}{n}} D_n^k e^{-\frac{t^{2n}}{n}}\Big)e^{-\frac{t^{2n}}{2n}}\,dt \qquad (2.38)\\
&\quad = (-1)^k \int_{-\infty}^{\infty} H^{(n)}_{2k}(t) D_n^k e^{-\frac{t^{2n}}{n}}\,dt \qquad (2.39)
\end{aligned}$$

and perform successive integration by parts, just like in the proof for orthogonality:

$$\begin{aligned}
&\int_{-\infty}^{\infty}\Big(H^{(n)}_{2k}(t)e^{-\frac{t^{2n}}{2n}}\Big)\Big(H^{(n)}_{2k}(t)e^{-\frac{t^{2n}}{2n}}\Big)dt \\
&\qquad\qquad = (-1)^k \int_{-\infty}^{\infty} D_n^k\Big(H^{(n)}_{2k}(t)\Big)e^{-\frac{t^{2n}}{n}}\,dt.
\end{aligned} \qquad (2.40)$$

We recognize that $D_n^k H^{(n)}_{2k}(t)$ is a constant. Finding A_k^2 results from evaluating

$$I(n) = \int_{-\infty}^{\infty} e^{-\frac{t^{2n}}{n}}\,dt \qquad (2.41)$$

and

$$C = D_n^k H^{(n)}_{2k}(t) \qquad (2.42)$$

where

$$A_k = \frac{1}{\sqrt{(-1)^k I(n) C}}. \tag{2.43}$$

We consider $I(n)$ first. Using the improper integral definition of the gamma function, we find that

$$I(n) = n^{\frac{1}{2n}-1}\Gamma(\frac{1}{2n}). \tag{2.44}$$

Now we tackle the value of C. Consider the effect of D_n on the monomial t^{2kn}, where k is a nonnegative integer:

$$\begin{aligned} D_n t^{2kn} &= -\frac{d}{dt}\frac{1}{t^{2n-2}}\frac{d}{dt}t^{2kn} && (2.45) \\ &= -2kn\Big(2(k-1)n+1\Big)t^{2(k-1)n}. && (2.46) \end{aligned}$$

Repeated application of D_n yields

$$\begin{aligned} D_n^2 t^{2kn} &= (-1)^2\Big[2kn\Big(2(k-1)n+1\Big)\Big] \\ &\times \Big[\Big(2(k-1)n\Big)\Big(2(k-2)n+1\Big)\Big]t^{2(k-2)n} && (2.47) \\ &. \\ &. \\ &. \\ D_n^k t^{2kn} &= (-1)^k 2^k k! n^k [2(k-1)n+1][2(k-2)n+1] \\ &\times [2(1)n+1][2(0)n+1]t^{2(k-k)n} && (2.48) \\ &= (-1)^k 2^k k! n^k \prod_{j=1}^{k-1}[2jn+1]. && (2.49) \end{aligned}$$

Since only the leading coefficient $a_k = 4^k$ of $H_{2k}^{(n)}(t)$ remains after k applications of D_n,

$$C = 4^k D_n^k t^{2kn} \tag{2.50}$$

$$= 4^k(-1)^k 2^k k! n^k \prod_{j=1}^{k-1}[2jn+1] \tag{2.51}$$

$$= (-1)^k 2^{3k} k! n^k \prod_{j=1}^{k-1}[2jn+1]. \tag{2.52}$$

Therefore

$$A_k = \Big[\Big(2^{3k} k! n^k \prod_{j=1}^{k-1}[2jn+1]\Big)\Big(n^{\frac{1}{2n}-1}\Gamma(\frac{1}{2n})\Big)\Big]^{-\frac{1}{2}}, \tag{2.53}$$

which gives the normalization. Together with the orthogonality of the functions $H_{2k}^{(n)}$, this completes the proof. □

The odd GHP normalization constant can also be obtained through a similar means. Without proof, it is given as

$$B_k = \Big[\Big(2^{3k+2} k! n^k \prod_{j=1}^{k}[2(j+1)n-1]\Big)\Big(n^{1-\frac{1}{2n}}\Gamma(2-\frac{1}{2n})\Big)\Big]^{-\frac{1}{2}} \tag{2.54}$$

where

$$h_{2k+1}^{(n)}(t) = B_k H_{2k+1}^{(n)}(t). \tag{2.55}$$

Substituting $H_{2k}^{(n)}(t)$ for $h_{2k}^{(n)}(t)$ into the recursion relation of Theorem 2.5 yields a three-term recursion relation in terms of the $h_{2k}^{(n)}(t)$:

$$h_{2k+2}^{(n)}(t) = \frac{A_{k+1}}{A_k}\Big(\alpha_k t^{2n} + \beta_k\Big)h_{2k}^{(n)}(t) + \frac{A_{k+1}}{A_{k-1}}\gamma_k h_{2k-2}^{(n)}(t). \tag{2.56}$$

This relation will be used in the following sections.

3 The matrix equation

We proceed by setting up the matrix equation used in finding the roots of the GHPs. Tailoring the method employed by [12], [14], and [22] for our own purposes, we rearrange (2.56) and isolate the term containing t^{2n}:

$$t^{2n}h_{2j}^{(n)}(t) = -\frac{A_j\gamma_j}{A_{j-1}\alpha_j}h_{2j-2}^{(n)}(t) - \frac{\beta_j}{\alpha_j}h_{2j}^{(n)}(t) + \frac{A_j}{A_{j+1}\alpha_j}h_{2j+2}^{(n)}(t). \tag{3.1}$$

We then index j from 0 to some $k-1$, allowing us to generate a system of k equations from (3.1):

$$t^{2n}h_{0}^{(n)}(t) = -\frac{\beta_0}{\alpha_0}h_{0}^{(n)}(t) + \frac{A_0}{A_1\alpha_0}h_{2}^{(n)}(t) \tag{3.2}$$

$$t^{2n}h_{2}^{(n)}(t) = -\frac{A_1\gamma_1}{A_0\alpha_1}h_{0}^{(n)}(t) + -\frac{\beta_1}{\alpha_1}h_{2}^{(n)}(t) + \frac{A_1}{A_2\alpha_1}h_{4}^{(n)}(t) \tag{3.3}$$

$$t^{2n}h_{4}^{(n)}(t) = -\frac{A_2\gamma_2}{A_1\alpha_2}h_{2}^{(n)}(t) + -\frac{\beta_2}{\alpha_2}h_{4}^{(n)}(t) + \frac{A_2}{A_3\alpha_2}h_{6}^{(n)}(t) \tag{3.4}$$

.

.

.

$$t^{2n}h_{2k-4}^{(n)}(t) = -\frac{A_{k-2}\gamma_{k-2}}{A_{k-3}\alpha_{k-2}}h_{2k-6}^{(n)}(t) + -\frac{\beta_{k-2}}{\alpha_{k-2}}h_{2k-4}^{(n)}(t) + \frac{A_{k-2}}{A_{k-1}\alpha_{k-2}}h_{2k-2}^{(n)}(t) \tag{3.5}$$

$$t^{2n}h_{2k-2}^{(n)}(t) = -\frac{A_{k-1}\gamma_{k-1}}{A_{k-2}\alpha_{k-1}}h_{2k-4}^{(n)}(t) + -\frac{\beta_{k-1}}{\alpha_{k-1}}h_{2k-2}^{(n)}(t) + \frac{A_{k-1}}{A_{k}\alpha_{k-1}}h_{2k}^{(n)}(t) \tag{3.6}$$

Next, we define the matrices

$$\mathbf{h}^{(n)}(t) = \begin{bmatrix} h_0^{(n)}(t) \\ h_2^{(n)}(t) \\ h_4^{(n)}(t) \\ . \\ . \\ . \\ h_{2k-2}^{(n)}(t) \end{bmatrix} \tag{3.7}$$

and

$$\mathbf{T} = \begin{bmatrix} -\frac{\beta_0}{\alpha_0} & \frac{A_0}{A_1\alpha_0} & 0 & 0 & . & . & . & 0 \\ -\frac{A_1\gamma_1}{A_0\alpha_1} & -\frac{\beta_1}{\alpha_1} & \frac{A_1}{A_2\alpha_1} & 0 & . & . & . & 0 \\ 0 & -\frac{A_2\gamma_2}{A_1\alpha_2} & -\frac{\beta_2}{\alpha_2} & \frac{A_2}{A_3\alpha_2} & . & . & . & 0 \\ . & . & . & . & . & & & . \\ . & . & . & . & & . & & . \\ . & . & . & . & & & . & . \\ 0 & 0 & 0 & 0 & . & . & -\frac{A_{k-1}\gamma_{k-1}}{A_{k-2}\alpha_{k-1}} & -\frac{\beta_{k-1}}{\alpha_{k-1}} \end{bmatrix} \tag{3.8}$$

and condense (3.2) through (3.6) in the matrix equation

$$t^{2n}\mathbf{h}^{(n)}(t) = \mathbf{T}\mathbf{h}^{(n)}(t) + \frac{A_{k-1}}{A_k\alpha_{k-1}} h_{2k}^{(n)}(t)\hat{\mathbf{e}}_k \tag{3.9}$$

where $\hat{\mathbf{e}}_k$ is the unit vector with zeros in all but the k^{th} row. Let t_j be a root indexed by j of the polynomial $h_{2k}^{(n)}(t)$. By plugging in the root t_j, we point out that (3.9) reduces to

$$t_j^{2n}\mathbf{h}^{(n)}(t_j) = \mathbf{T}\mathbf{h}^{(n)}(t_j), \tag{3.10}$$

which is an eigenvalue problem with eigenvalue $\lambda_j = t_j^{2n}$. Solving for the set of eigenvalues λ_j will be a critical step in finding the set of roots t_j necessary for the quadrature rule.

But there is a minor complication; not all the roots of $h_{2k}^{(n)}(t)$ (or equivalently, $H_{2k}^{(n)}(t)$) are real. We can easily see this in the case that n is even, for which even if all the λ_j are positive, t_j^n can be negative, and t_j would subsequently be complex. This property is a byproduct of our formulation of the generalization of Hermite polynomials, particularly their property of consisting only of terms of t^{2ln}. We prefer not to use complex numbers in our Gaussian quadrature rule for real integrals. For this reason, in Section 4 we will determine which roots are important for quadrature rule sampling.

Alternatively, the matrix $\mathbf{T}$ can be formulated with the entries

$$T_{i,i} = -\frac{\beta_{i-1}}{4} \tag{3.11}$$

and

$$T_{i,i+1} = T_{i+1,i} = \frac{\sqrt{-\gamma_i}}{4} \tag{3.12}$$

since the matrix is symmetric.

Theorem 3.1. *The tridiagonal matrix* $\boldsymbol{T}$ *is symmetric, and its nonzero elements are given by (3.11) and (3.12).*

Proof. We only need to simplify the expressions for the off-diagonal elements of $\mathbf{T}$ in (3.8). Consider the ratio

$$\frac{A_{i-1}}{A_i} = \left[\frac{2^{3(i-1)}(i-1)!n^{i-1}\prod_{j=1}^{i-2}[2jn+1]\left(n^{\frac{1}{2n}-1}\Gamma\left(\frac{1}{2n}\right)\right)}{2^{3i}i!n^{i}\prod_{j=1}^{i-1}[2jn+1]\left(n^{\frac{1}{2n}-1}\Gamma\left(\frac{1}{2n}\right)\right)}\right]^{-\frac{1}{2}}, \tag{3.13}$$

$$= \sqrt{16i^2n^2 - 16in^2 + 8in}, \tag{3.14}$$

$$= \sqrt{-\gamma_i}. \tag{3.15}$$

Therefore,

$$T_{i,i+1} = \frac{A_{i-1}}{A_i\alpha_{i-1}} = \frac{\sqrt{-\gamma_i}}{4} \tag{3.16}$$

and

$$T_{i+1,i} = -\frac{A_i\gamma_i}{A_{i-1}\alpha_i} = \frac{\sqrt{-\gamma_i}}{4} \tag{3.17}$$

proving the theorem. □

4 Roots

The matrix $\mathbf{T}$ is a $k \times k$ tridiagonal matrix whose eigenvalues λ_j will provide us with the roots of $h_{2k}^{(n)}(t)$. We show that solving the eigenvalue equation (3.10) will yield k distinct and positive values of λ_j, of which we will take the positive roots

$$t_j = +\lambda_j^{\frac{1}{2n}} \tag{4.1}$$

for quadrature rule sampling.

Theorem 4.1. *Each even GHP $h_{2k}^{(n)}$ has k real, distinct, and positive roots $t_1, ..., t_k$.*

Proof. Before beginning with the proof inspired by a similar proof from [22], it is convenient to make a minor modification to the orthogonality relation. Since the integrand is even, we make the argument based on symmetry that

$$\int_{-\infty}^{\infty} h_{2k}^{(n)}(t)h_{2k'}^{(n)}(t)e^{-\frac{t^{2n}}{n}}dt = 2\int_{0}^{\infty} h_{2k}^{(n)}(t)h_{2k'}^{(n)}(t)e^{-\frac{t^{2n}}{n}}dt = 0 \tag{4.2}$$

which is the case if $k \neq k'$. Let $\tau = t^{2n}$, then

$$2\int_{0}^{\infty} h_{2k}^{(n)}(t)h_{2k'}^{(n)}(t)e^{-\frac{t^{2n}}{n}}dt = 2\int_{0}^{\infty} \tilde{h}_{2k}^{(n)}(\tau)\tilde{h}_{2k'}^{(n)}(\tau)e^{-\frac{\tau}{n}}\frac{\tau^{\frac{1}{2n}-1}}{2n}d\tau = 0. \tag{4.3}$$

Now we can consider an even GHP in τ on only the positive real line:

$$h_{2k}^{(n)}(t) = \sum_{l=0}^{k} a_l t^{2ln} = \sum_{l=0}^{k} a_l \tau^l = \tilde{h}_{2k}^{(n)}(\tau). \tag{4.4}$$

We proceed with a proof by contradiction. Let $k' = 0$ and $k \geq 1$ in (4.3):

$$2\int_{0}^{\infty} h_{0}^{(n)}\tilde{h}_{2k}^{(n)}(\tau)e^{-\frac{\tau}{n}}\frac{\tau^{\frac{1}{2n}-1}}{2n}d\tau = 0. \tag{4.5}$$

Since $2h_0^{(n)}$ and $e^{-\frac{\tau}{n}}\frac{\tau^{\frac{1}{2n}-1}}{2n}$ are positive for $\tau > 0$ (yet the integral vanishes), it follows that $\tilde{h}_{2k}^{(n)}$ must change sign at least once on the interval $(0, \infty)$. Let r be a natural number less than k that denotes the number of roots of $\tilde{h}_{2k}^{(n)}(\tau)$ on the positive real line for which $\tilde{h}_{2k}^{(n)}$ changes sign. These roots would have odd multiplicity (the powers $p_1, ..., p_r$ are odd integers), and so we factor $\tilde{h}_{2k}^{(n)}$:

$$\tilde{h}_{2k}^{(n)}(\tau) = (\tau - \tau_1)^{p_1}...(\tau - \tau_r)^{p_r}\psi(\tau). \tag{4.6}$$

ψ is then a polynomial of degree $k - (p_1 + ... + p_r)$ that by itself does not change sign for any $\tau > 0$.

Multiplying both sides by $(\tau-\tau_1)...(\tau-\tau_r)$, we get

$$\tilde{h}_{2k}^{(n)}(\tau)(\tau-\tau_1)...(\tau-\tau_r) = (\tau-\tau_1)^{p_1+1}...(\tau-\tau_r)^{p_r+1}\psi(\tau), \tag{4.7}$$

and note that neither side now ought not to change sign on the positive real line. Yet, by integrating both sides over the integrable singularity with respect to the weight function,

$$\int_0^\infty \tilde{h}_{2k}^{(n)}(\tau)(\tau-\tau_1)...(\tau-\tau_r)e^{-\frac{\tau}{n}}\frac{\tau^{\frac{1}{2n}-1}}{2n}d\tau = 0, \tag{4.8}$$

the integral vanishes due to the product

$$(\tau-\tau_1)...(\tau-\tau_r) = \sum_{l=0}^{r} c_l \tilde{h}_{2l}^{(n)}(\tau) \tag{4.9}$$

producing a polynomial that can be expressed as a linear combination of GHP indexed below $\tilde{h}_{2k}^{(n)}(\tau)$, since $r < k$. We have our contradiction. Therefore, $r = k$, and by taking $t_j = \tau_j^{\frac{1}{2n}}$ we prove the theorem. □

Corollary 4.1. *The $k \times k$ square matrix $\boldsymbol{T}$ has k positive eigenvalues.*

Proof. In Theorem 4.1, it was shown that $\tilde{h}_{2k}^{(n)}(\tau)$ was guaranteed k positive real roots $\tau_1, ..., \tau_k$. By (3.10) we know that these k roots are eigenvalues of the matrix $\mathbf{T}$. □

In practice, we apply the QR algorithm to $\mathbf{T}$ to find the eigenvalues $\lambda_j = \tau_j$. We are then able to find the real, positive roots of the GHP $t_j = \tau_j^{1/2n}$. The roots are then stored for later use as sampling points for the quadrature rule.

4.1 A Christoffel-Darboux identity

Before proceeding with the derivation of the quadrature weights, it is necessary to establish a general form of the Christoffel-Darboux identity. From (3.9) we replace t with s to produce a sister equation:

$$s^{2n}\mathbf{h}^{(n)}(s) = \mathbf{T}\mathbf{h}^{(n)}(s) + \frac{A_{k-1}}{A_k\alpha_{k-1}} h_{2k}^{(n)}(s)\hat{\mathbf{e}}_k. \tag{4.10}$$

Take the dot product of (3.9) with $\mathbf{h}^{(n)}(s)$ and (4.10) with $\mathbf{h}^{(n)}(t)$ and subtract one from the other to yield an identity similar to that of Christoffel-Darboux,

$$\begin{aligned}(t^{2n} - s^{2n})\langle\, \mathbf{h}^{(n)}(t), \mathbf{h}^{(n)}(s)\rangle \\ = \frac{A_{k-1}}{A_k\alpha_{k-1}}[h_{2k}^{(n)}(t)h_{2k-2}^{(n)}(s) - h_{2k}^{(n)}(s)h_{2k-2}^{(n)}(t)].\end{aligned} \tag{4.11}$$

We can also reduce the left side to its scalar form. This is the form of which we will make use,

$$\sum_{v=0}^{k-1} h_{2v}^{(n)}(t)h_{2v}^{(n)}(s) = \frac{A_{k-1}}{A_k\alpha_{k-1}} \frac{h_{2k}^{(n)}(t)h_{2k-2}^{(n)}(s) - h_{2k}^{(n)}(s)h_{2k-2}^{(n)}(t)}{(t^{2n} - s^{2n})} \tag{4.12}$$

5 Weights

We wish to establish a quadrature rule which is only dependent on the t_j, the roots specified in (4.1). With regular orthogonal polynomials, such a quadrature rule would take the form

$$\int_{-\infty}^{\infty} p(t)w(t)dt = \sum_{j=1}^{k} w_j p(t_j) \tag{5.1}$$

where w_j are the weights corresponding to the roots t_j and p is a suitably degree-limited polynomial. However, since the weight function $w^{(n)}(t) = e^{-\frac{t^{2n}}{n}}$ is even, both in the special $n = 1$ case and the general case, we know that any evaluation of (5.1) for which p is odd must vanish. (This is why we have omitted any discussion of the odd GHP since the section on their normalization constant.)

Therefore, we can rewrite (5.1) in terms of even and odd components of p:

$$\begin{aligned}
\int_{-\infty}^{\infty} p(t)w^{(n)}(t)dt &= \int_{-\infty}^{\infty} \Big(\frac{p(t)+p(-t)}{2} + \frac{p(t)-p(-t)}{2}\Big) w^{(n)}(t)dt \\
&= \int_{-\infty}^{\infty} \Big(\frac{p(t)+p(-t)}{2}\Big) w^{(n)}(t)dt \\
&= \frac{1}{2}\sum_{j=1}^{k} w_j \Big(p(t_j) + p(-t_j)\Big). \qquad (5.2)
\end{aligned}$$

This way, only the even component of p is singled out and integrated. We will develop a quadrature rule of the form (5.2) rather than (5.1) to accommodate functions that have nonzero odd and even components.

Our first main result states that we can integrate certain polynomials of degree up to $4kn-1$ by evaluating them at $2k$ points $\{\pm t_1, \pm t_2, \ldots, \pm t_k\}$.

Theorem 5.1. *Let p be a polynomial of the form*

$$p(t) = \sum_{l=0}^{2k-1} a_l t^{2ln} + \sum_{l=1}^{2kn} b_l t^{2l-1}$$

with real coefficients a_l for $l \in \{0, 1, \ldots, 2k-1\}$ and b_l for $l \in \{1, 2, \ldots, 2kn\}$. For the weights

$$w_j = \Big(\sum_{v=0}^{k-1} [h_{2v}^{(n)}(t_j)]^2\Big)^{-1} = \langle \mathbf{h}^{(n)}(t_j), \mathbf{h}^{(n)}(t_j)\rangle^{-1}$$

the quadrature rule (5.2) holds.

Proof. To derive the w_j, we first define an intermediate function $l_j^{(n)}(t)$

$$\begin{aligned}
l_j^{(n)}(t) &= \frac{h_{2k}^{(n)}(t)}{(t^{2n} - t_j^{2n}) \frac{d}{d\tau}\Big[h_{2k}^{(n)}(t)\Big]_{t=t_j}} \qquad (5.3) \\
&= \frac{(t^{2n} - t_1^{2n})...(t^{2n} - t_{j-1}^{2n})(t^{2n} - t_{j+1}^{2n})...(t^{2n} - t_k^{2n})}{(t_j^{2n} - t_1^{2n})...(t_j^{2n} - t_{j-1}^{2n})(t_j^{2n} - t_{j+1}^{2n})...(t_j^{2n} - t_k^{2n})} \qquad (5.4)
\end{aligned}$$

where we define the operator $\frac{d}{d\tau} = \frac{1}{2n} t^{1-2n} \frac{d}{dt}$ (since $\tau = t^{2n}$). We emphasize that factoring $h_{2k}^{(n)}(t)$

from (5.3) to (5.4) is only possible because of Theorem 4.1, hence its importance to the determination that the GHPs are suitable for the construction of a quadrature rule. The intermediate function has the property that

$$l_j^{(n)}(t_v) = \delta_{jv}. \tag{5.5}$$

We now define a Lagrange interpolating polynomial $L_{2k-2}^{(n)}(t)$:

$$L_{2k-2}^{(n)}(t) = \sum_{j=1}^{k} l_j^{(n)}(t) f(t_j). \tag{5.6}$$

This polynomial has degree $2(k-1)n$ and passes through the points $(t_j, f(t_j))$ (i.e. $L_{2k-2}^{(n)}(t_j) = f(t_j)$).

Observe that the zeros of the polynomial $h_{2k}^{(n)}(t)$ are also the zeros of the polynomial $L_{2k-2}^{(n)}(t) - f(t)$. Therefore, there exist polynomials $r(t)$ and $s(t)$ where

- $r(t)$ consists of terms $c_l t^{2ln}$, where $l \in [0, k-1]$, and
- $s(t)$ consists of terms $d_l t^{2ln-1}$, where $l \in [1, k]$,

such that

$$f(t) - L_{2k-2}^{(n)}(t) = h_{2k}^{(n)}(t)[r(t) + s(t)]. \tag{5.7}$$

Note that $r(t)$ and $s(t)$ are linear combinations of the even and odd generalized Hermite polynomials of degree less than $2kn$. Rearrange (5.7) and multiply by $w^{(n)}(t)$. Then, integrating over $\mathbb{R}$ yields

$$\begin{aligned} \int_{-\infty}^{\infty} f(t) w^{(n)}(t) dt &= \int_{-\infty}^{\infty} L_{2k-2}^{(n)}(t) w^{(n)}(t) dt \\ &\quad + \int_{-\infty}^{\infty} h_{2k}^{(n)}(t)[r(t) + s(t)] w^{(n)}(t) dt & (5.8) \\ &= \int_{-\infty}^{\infty} L_{2k-2}^{(n)}(t) w^{(n)}(t) dt \\ &= \sum_{j=1}^{k} f(t_j) \int_{-\infty}^{\infty} \frac{h_{2k}^{(n)}(t) w^{(n)}(t)}{(t^{2n} - t_j^{2n}) \frac{d}{d\tau}\left[h_{2k}^{(n)}(t)\right]_{t=t_j}} dt. & (5.9) \end{aligned}$$

The term containing $r(t)$ and $s(t)$ on the right of (5.8) evaluates to zero because $r(t)$ and $s(t)$ are linear combinations of the even and odd generalized Hermite polynomials of degree less than $2kn$, and $h_{2k}^{(n)}(t)$ is orthogonal to those polynomials.

Since $h_{2k}^{(n)}$ is even, for every positive root t_j there exists a negative root $-t_j$. Therefore, a similar derivation can be performed such that one receives a parallel result with that of (5.9):

$$\int_{-\infty}^{\infty} f(t)w^{(n)}(t)dt = \sum_{j=1}^{k} f(-t_j) \int_{-\infty}^{\infty} \frac{h_{2k}^{(n)}(t)w^{(n)}(t)}{(t^{2n}-t_j^{2n})\frac{d}{d\tau}\left[h_{2k}^{(n)}(t)\right]_{t=-t_j}} dt. \tag{5.10}$$

Add (5.9) and (5.10), and divide by 2:

$$\begin{aligned}\int_{-\infty}^{\infty} f(t)w^{(n)}(t)dt &= \frac{1}{2}\Bigg[\sum_{j=1}^{k} f(t_j) \int_{-\infty}^{\infty} \frac{h_{2k}^{(n)}(t)w^{(n)}(t)}{(t^{2n}-t_j^{2n})\frac{d}{d\tau}\left[h_{2k}^{(n)}(t)\right]_{t=t_j}} \\ &\qquad + f(-t_j) \int_{-\infty}^{\infty} \frac{h_{2k}^{(n)}(t)w^{(n)}(t)}{(t^{2n}-t_j^{2n})\frac{d}{d\tau}\left[h_{2k}^{(n)}(t)\right]_{t=-t_j}}\Bigg] \\ &= \frac{1}{2}\sum_{j=1}^{k}\Big(f(t_j)+f(-t_j)\Big) \\ &\qquad \times \int_{-\infty}^{\infty} \frac{h_{2k}^{(n)}(t)w^{(n)}(t)}{(t^{2n}-t_j^{2n})\frac{d}{d\tau}\left[h_{2k}^{(n)}(t)\right]_{t=t_j}}.\end{aligned} \tag{5.11}$$

Comparing with (5.2) we get that

$$w_j = \frac{1}{\frac{d}{d\tau}\left[h_{2k}^{(n)}(t)\right]_{t=t_j}} \int_{-\infty}^{\infty} \frac{h_{2k}^{(n)}(t)w^{(n)}(t)}{(t^{2n}-t_j^{2n})} dt. \tag{5.12}$$

But in its current form, (5.12) is impractical, since the problem of evaluating an integral from $-\infty$ to ∞ has been converted to evaluating k integrals from $-\infty$ to ∞.

To eliminate this problem, let $s = t_j$ in (4.12), the scalar form of the Christoffel-Darboux identity:

$$\sum_{v=0}^{k-1} h_{2v}^{(n)}(t)h_{2v}^{(n)}(t_j) = \frac{A_{k-1}}{A_k a_{k-1}} \frac{h_{2k}^{(n)}(t)h_{2k-2}^{(n)}(t_j)}{(t^{2n}-t_j^{2n})}. \tag{5.13}$$

Multiply by $w^{(n)}(t)$ and integrate from $-\infty$ to ∞,

$$1 = \frac{A_{k-1}}{A_k \alpha_{k-1}} h_{2k-2}^{(n)}(t_j) \int_{-\infty}^{\infty} \frac{h_{2k}^{(n)}(t) w^{(n)}(t)}{(t^{2n} - t_j^{2n})} dt, \tag{5.14}$$

$$= \frac{A_{k-1}}{A_k \alpha_{k-1}} h_{2k-2}^{(n)}(t_j) \frac{d}{d\tau} \left[h_{2k}^{(n)}(t) \right]_{t=t_j} w_j \tag{5.15}$$

and solve for w_j,

$$w_j = \frac{A_k \alpha_{k-1}}{A_{k-1}} \left[h_{2k-2}^{(n)}(t_j) \frac{d}{d\tau} \left[h_{2k}^{(n)}(t) \right]_{t=t_j} \right]^{-1}. \tag{5.16}$$

This expression for the weights is more practical than its integral form, but is still not ideal for quick computation due to the presence of the $\frac{d}{d\tau}$ operator. Again, go back to (4.12), the scalar form of the Christoffel-Darboux identity, but this time, take the limit as $s \to t$:

$$\begin{aligned}\sum_{v=0}^{k-1} [h_{2v}^{(n)}(t)]^2 &= \frac{A_{k-1}}{A_k \alpha_{k-1}} \lim_{s \to t} \Big[h_{2k}^{(n)}(t) \Big(\frac{h_{2k-2}^{(n)}(s) - h_{2k-2}^{(n)}(t)}{t^{2n} - s^{2n}} \Big) \\ &\quad + h_{2k-2}^{(n)}(t) \Big(\frac{h_{2k}^{(n)}(t) - h_{2k}^{(n)}(s)}{t^{2n} - s^{2n}} \Big) \Big] \qquad (5.17) \\ &= \frac{A_{k-1}}{A_k \alpha_{k-1}} \Big[- h_{2k}^{(n)}(t) \frac{d}{d\tau} \Big(h_{2k-2}^{(n)}(t) \Big) \\ &\quad + h_{2k-2}^{(n)}(t) \frac{d}{d\tau} \Big(h_{2k}^{(n)}(t) \Big) \Big]. \qquad (5.18)\end{aligned}$$

Evaluate $t = t_j$,

$$\sum_{v=0}^{k-1} [h_{2v}^{(n)}(t_j)]^2 = \frac{A_{k-1}}{A_k \alpha_{k-1}} \Big[h_{2k-2}^{(n)}(t_j) \frac{d}{d\tau} \Big(h_{2k}^{(n)}(t) \Big)_{t=t_j} \Big] = \frac{1}{w_j}, \tag{5.19}$$

and solve for w_j to prove the theorem:

$$w_j = \Big(\sum_{v=0}^{k-1} [h_{2v}^{(n)}(t_j)]^2 \Big)^{-1} = \langle \mathbf{h}^{(n)}(t_j), \mathbf{h}^{(n)}(t_j) \rangle^{-1}. \tag{5.20}$$

□

The form (5.20) for the weights is much more usable since there are no integral or derivative operators necessary for the computation. The brunt of the computation is the evaluation of the even GHPs indexed from 0 to k at t_j.

Equation (5.20) shows an expression for the weights of the generalized Hermite-Gauss quadrature rule that is very similar to those shown for standard orthogonal polynomials. The expression suggests that to calculate the roots and weights for a generalized weight function we can look to the standard orthogonal polynomials and generalize them such that they are compatible with the weight function of interest, as we have done with the generalized Gaussian.

6 Hermite interpolation and the quadrature error

We retrieved the weights by deriving the Lagrange interpolation polynomial from the GHPs. However, to obtain the error in the quadrature rule, we must derive a type of Hermite interpolation polynomial in our setting. The Hermite interpolation polynomial seeks to approximate a function $f(t)$ by matching its values at points $\{t_j\}_{j=1}^{k}$ and at the same matching the values of its derivative. In our case, if $y(t)$ is the Hermite interpolation polynomial, then

$$y(t_j) = f(t_j) \tag{6.1}$$

and

$$\left(t^{1-2n}\frac{d}{dt}y(t)\right)_{t=t_j} = \left(t^{1-2n}\frac{d}{dt}f(t)\right)_{t=t_j} \tag{6.2}$$

[15, compare]. Since the GHPs are complete, we know we can approximate $f(t)$ as a linear combination of the GHPs, and therefore, there is a series of polynomials

$$p_k(t) = \sum_{l=0}^{k} a_l t^{2ln} \tag{6.3}$$

indexed by k that converges in $L^2(\mathbb{R}, e^{-t^{2n}/n})$ to a given even f as $k \to \infty$.

Theorem 6.1. *If f is integrable with respect to the n-Gaussian and its even part is $2k$-times* *continuously differentiable on $(0,\infty)$, then the approximation given by the quadrature rule (5.2) has an error*

$$E = \frac{\left(\frac{t^{1-2n}}{2n}\frac{d}{dt}\right)^{2k}\Big(f(t)\Big)_{t=\xi}}{c_k^2(2k)!}$$

where c_k is the leading coefficient for $h_{2k}^{(n)}(t)$ and ξ is a value between the smallest and largest of the sampling points t_j.

Proof. To simplify notation, we abbreviate $\frac{d}{d\tau} \equiv \frac{t^{1-2n}}{2n}\frac{d}{dt}$.

Suppose values of $f(t)$ and $\frac{d}{d\tau}\Big(f(t)\Big)$ are known for $t = t_1, t_2, ..., t_k$. Assume a polynomial $y(t)$ of degree $4kn - 2n$ such that $y(t_j) = f(t_j)$ and $\frac{d}{d\tau}\Big(y(t)\Big)_{t=t_j} = \frac{d}{d\tau}\Big(f(t)\Big)_{t=t_j}$. Define a polynomial $y(t)$ which takes the form:

$$y(t) = \sum_{j=1}^{k} p_j(t)f(t_j) + \sum_{j=1}^{k} \bar{p}_j(t)\frac{d}{d\tau}\Big(f(t)\Big)_{t=t_j}. \tag{6.4}$$

$p_j(t)$ and $\bar{p}_j(t)$ are polynomials of maximum degree $4kn-2n$, to be determined. $y(t)$ is the Hermite interpolation polynomial, which not only seeks to match $f(t)$ at the roots, but also $\frac{d}{d\tau}\Big(f(t)\Big)$ at the roots.

We will be using the same definition for $l_j^{(n)}(t)$ as before:

$$l_j^{(n)}(t) = \frac{h_{2k}^{(n)}(t)}{(t^{2n} - t_j^{2n})\frac{d}{d\tau}\Big(h_{2k}^{(n)}(t)\Big)_{t=t_j}}. \tag{6.5}$$

In order for $y(t)$ to have the properties we want it to have, we must fulfill the following properties:

$$p_i(t_j) = \delta_{ij}, \qquad \bar{p}_i(t_j) = 0, \tag{6.6}$$

$$\frac{d}{d\tau}\Big(p_i(t)\Big)_{t=t_j} = 0, \qquad \frac{d}{d\tau}\Big(\bar{p}_i(t)\Big)_{t=t_j} = \delta_{ij}. \tag{6.7}$$

Recognize that $[l_i^{(n)}(t)]^2$ is a polynomial of degree $4kn-4n$ having the property that $[l_i^{(n)}(t)]^2 = \delta_{ij}$, and so it follows that

$$p_i(t) = r_i(t)[l_i^{(n)}(t)]^2, \qquad \bar{p}_i(t) = s_i(t)[l_i^{(n)}(t)]^2, \tag{6.8}$$

where $r_i(t)$ and $s_i(t)$ are polynomials with degree $2n$. From the above four conditions we deduce the following properties for $r_i(t)$ and $s_i(t)$:

$$r_i(t_i) = 1, \qquad s_i(t_i) = 0, \tag{6.9}$$

$$\frac{d}{d\tau}\Big(r_i(t)\Big)_{t=t_i} + 2\frac{d}{d\tau}\Big(l_i^{(n)}(t)\Big)_{t=t_i} = 0, \qquad \frac{d}{d\tau}\Big(s_i(t)\Big)_{t=t_i} = 1. \tag{6.10}$$

Then, it follows that

$$r_i(t) = 1 - 2\frac{d}{d\tau}\Big(l_i^{(n)}(t)\Big)_{t=t_i}(t^{2n} - t_i^{2n}), \qquad s_i(t) = t^{2n} - t_i^{2n}. \tag{6.11}$$

Then,

$$p_j(t) = \Big[1 - 2\frac{d}{d\tau}\Big(l_j^{(n)}(t)\Big)_{t=t_j}(t^{2n} - t_j^{2n})\Big][l_j^{(n)}(t)]^2 \tag{6.12}$$

and

$$\bar{p}_j(t) = \Big(t^{2n} - t_j^{2n}\Big)[l_j^{(n)}(t)]^2. \tag{6.13}$$

Define the error function

$$E(t) = f(t) - y(t) \tag{6.14}$$

and notice that $E(t)$, $h_{2k}^{(n)}(t)$, and $\frac{d}{d\tau}E(t)$ all vanish at each of the roots $t_1, ..., t_k$. Define the auxiliary function

$$F(t) = f(t) - y(t) - K[h_{2k}^{(n)}(t)]^2 \tag{6.15}$$

which has the same properties of vanishing at the roots $t_1, ..., t_k$, and let K be such that $F(t)$ also vanishes at an additional point $t = \bar{t}$. Since $F(t)$ vanishes at the $k+1$ points $t_1, ..., t_k$, and $\bar{t}$, $\frac{d}{d\tau}F(t)$ must vanish at at least k intermediate points by Rolle's Theorem. But $\frac{d}{d\tau}F(t)$ also vanishes at the

k points $t_1, ..., t_k$, since

$$\left(\frac{d}{d\tau}F(t)\right)_{t=t_j} = \left(\frac{d}{d\tau}f(t)\right)_{t=t_j} - \left(\frac{d}{d\tau}y(t)\right)_{t=t_j} - \left(\frac{d}{d\tau}K[h_{2k}^{(n)}(t)]^2\right)_{t=t_j} \tag{6.16}$$

$$= \left(\frac{d}{d\tau}K[h_{2k}^{(n)}(t)]^2\right)_{t=t_j} \tag{6.17}$$

$$= \left(\frac{d}{d\tau}(K)\,[h_{2k}^{(n)}(t)]^2\right)_{t=t_j} + \left(K\frac{d}{d\tau}\left([h_{2k}^{(n)}(t)]^2\right)\right)_{t=t_j} \tag{6.18}$$

$$= \left(\frac{d}{d\tau}(K)\,[h_{2k}^{(n)}(t)]^2\right)_{t=t_j} + \left(2Kh_{2k}^{(n)}(t)\frac{d}{d\tau}[h_{2k}^{(n)}(t)]\right)_{t=t_j} \tag{6.19}$$

$$= 0. \tag{6.20}$$

Therefore, it vanishes at least $2k$ times on the positive real line. Thus, $\frac{d^2}{d\tau^2}F(t)$ vanishes at least $2k-1$ times, $\frac{d^3}{d\tau^3}F(t)$ at least $2k-2$ times, and so on. Hence, $\frac{d^{2k}}{d\tau^{2k}}F(t)$ vanishes at least once on the interval $[t_1, t_k]$. Recall that $y(t)$ is a polynomial of degree $4kn-2n$, and so $\frac{d^{2k}}{d\tau^{2k}}y(t) = 0$. If at the value ξ the function $\frac{d^{2k}}{d\tau^{2k}}F(t)$ vanishes, then

$$0 = \frac{d^{2k}}{d\tau^{2k}}\Big(F(t)\Big)_{t=\xi} = \frac{d^{2k}}{d\tau^{2k}}\Big(f(t)\Big)_{t=\xi} - Kc_k^2(2k)! \tag{6.21}$$

where c_k is the leading coefficient for $h_{2k}^{(n)}(t)$. Rearrange for K:

$$K = \frac{\frac{d^{2k}}{d\tau^{2k}}\Big(f(t)\Big)_{t=\xi}}{c_k^2(2k)!}. \tag{6.22}$$

Since $F(\bar{t}) = 0$, it follows that

$$E(\bar{t}) \equiv f(\bar{t}) - y(\bar{t}) = K[h_{2k}^{(n)}(\bar{t})]^2. \tag{6.23}$$

By suppressing the bars [15], it follows that

$$E(t) = \frac{\frac{d^{2k}}{d\tau^{2k}}\Big(f(t)\Big)_{t=\xi}}{c_k^2(2k)!}[h_{2k}^{(n)}(t)]^2, \tag{6.24}$$

where ξ is somewhere between the smallest and largest of the roots $t_1, ..., t_k$.

As of now, this is the error in the Hermite interpolation polynomial. However, we need the error in the Lagrange interpolation polynomial. We will show that they are one and the same. We first write

$$f(t) = y(t) + E(t) \tag{6.25}$$

and take the integral of both sides after multiplying by the weight function:

$$\int_{-\infty}^{\infty} f(t)w^{(n)}(t)dt = \int_{-\infty}^{\infty} y(t)w^{(n)}(t)dt + \int_{-\infty}^{\infty} E(t)w^{(n)}(t)dt. \tag{6.26}$$

Evaluate the first integral on the right,

$$\begin{aligned}
\int_{-\infty}^{\infty} y(t)w^{(n)}(t)dt &= \sum_{j=1}^{k} f(t_j) \int_{-\infty}^{\infty} p_j(t)w^{(n)}(t)dt \\
&\quad + \sum_{j=1}^{k} \frac{d}{d\tau}\Big(f(t)\Big)_{t=t_j} \int_{-\infty}^{\infty} \bar{p}_j(t)w^{(n)}(t)dt && (6.27)\\
&= \sum_{j=1}^{k} f(t_j)H_j + \sum_{j=1}^{k} \frac{d}{d\tau}\Big(f(t)\Big)_{t=t_j} \bar{H}_j, && (6.28)
\end{aligned}$$

where H_j and $\bar{H}_j$ are the Hermite weights. We work out the first set of weights:

$$\begin{aligned}
H_j &= \int_{-\infty}^{\infty} p_j(t)w^{(n)}(t)dt && (6.29)\\
&= \int_{-\infty}^{\infty} \left[1 - 2\frac{d}{d\tau}\Big(l_j^{(n)}(t)\Big)_{t=t_j}(t^{2n} - t_j^{2n})\right][l_j^{(n)}(t)]^2 w^{(n)}(t)dt && (6.30)\\
&= \int_{-\infty}^{\infty} [l_j^{(n)}(t)]^2 w^{(n)}(t)dt. && (6.31)
\end{aligned}$$

We then work out the second set of weights:

$$\bar{H}_j = \int_{-\infty}^{\infty} \bar{p}_j(t) w^{(n)}(t) dt \tag{6.32}$$

$$= \int_{-\infty}^{\infty} \left(t^{2n} - t_j^{2n} \right) [l_j^{(n)}(t)]^2 w^{(n)}(t) dt \tag{6.33}$$

$$= 0. \tag{6.34}$$

We then obtain a quadrature rule based on Hermite polynomial interpolation, which bears a striking resemblance to the quadrature rule from Lagrange polynomial interpolation:

$$\int_{-\infty}^{\infty} f(t) w^{(n)}(t) dt = \sum_{j=1}^{k} f(t_j) H_j + \int_{-\infty}^{\infty} E(t) w^{(n)}(t) dt. \tag{6.35}$$

The quadrature rule derived from Hermite interpolation is the same as the quadrature rule derived from Lagrange interpolation. We show this by showing that $H_j = w_j$. In (6.35), the Hermite interpolation quadrature rule, let $f(t) = l_i^{(n)}(t)$:

$$\int_{-\infty}^{\infty} l_i^{(n)}(t) w^{(n)}(t) dt = \sum_{j=1}^{k} l_i^{(n)}(t_j) H_j + \int_{-\infty}^{\infty} E(t) w^{(n)}(t) dt \tag{6.36}$$

$$= H_i + \int_{-\infty}^{\infty} E(t) w^{(n)}(t) dt \tag{6.37}$$

$$= H_i \tag{6.38}$$

$$= w_i. \tag{6.39}$$

(The fact that $\frac{d^{2k}}{d\tau^{2k}} l_i^{(n)}(t) = 0$ and the definition of w_i produce the last two lines.) And so "both" quadrature rules are really the same rule. Therefore, the error of the Lagrange-derived rule is the same as the error of the Hermite-derived rule.

Define E as the evaluation of the integral containing $E(t)$ in (6.26),

$$E = \frac{\frac{d^{2k}}{d\tau^{2k}}\Big(f(t)\Big)_{t=\xi}}{c_k^2(2k)!}\int_{-\infty}^{\infty}[h_{2k}^{(n)}(t)]^2 w^{(n)}(t)dt \tag{6.40}$$

$$= \frac{\frac{d^{2k}}{d\tau^{2k}}\Big(f(t)\Big)_{t=\xi}}{c_k^2(2k)!}, \tag{6.41}$$

thus proving the theorem. $\square$

Therefore, the final statement of the generalized Hermite-Gauss quadrature rule is

$$\int_{-\infty}^{\infty} f(t)e^{-\frac{t^{2n}}{n}}dt = \frac{1}{2}\sum_{j=1}^{k} w_j\Big(f(t_j)+f(-t_j)\Big) + \frac{\left(\frac{t^{1-2n}}{2n}\frac{d}{dt}\right)^{2k}\Big(f(t)\Big)_{t=\xi}}{c_k^2(2k)!}. \tag{6.42}$$

We conclude with more concrete error bounds for a class of functions that includes polynomials, the bandlimited distributions.

Definition 6.1. *Define for any compactly supported function $\phi \in C^\infty(\mathbb{R})$, the Sobolev norm $\|\phi\|_N = \max\{\phi^{(k)}(x) : x \in \mathbb{R}, 0 \le k \le N\}$. A function f is a distribution of order N if N is the smallest integer so that for each compact set K there is $C_K \ge 0$ such that $|\int_K f\phi dx| \le C_K\|\phi\|_N$ holds for each $\phi \in C^\infty(\mathbb{R})$ with support in K.*

A function f that is analytic in the entire complex plane and satisfies that for some $\gamma, \Omega > 0$ the growth condition

$$|f(z_1+iz_2)| \le \gamma(1+|z_1+iz_2|)^N e^{\Omega|z_2|}$$

holds for all $z_1, z_2 \in \mathbb{R}$ is called an Ω-bandlimited distribution.

Using integration by parts, we see that polynomials of degree N are indeed bandlimited distributions of degree N for any $\Omega > 0$. However, such distributions of degree N can be more general. For example, multiplying a polynomial of degree N with a smooth function whose Fourier transform has support in $[-\Omega, \Omega]$ retains the degree [18, p. 159] and gives an Ω-bandlimited distribtution. This is a consequence of the next theorem.

Theorem 6.2. *[18, Theorem 7.23] Let f be an entire function that is an Ω-bandlimited distribution, then there is a sequence of signed measures μ_j, $j \in \{0, 1, \ldots, N+1\}$ with support in $[-\Omega, \Omega]$ such that*

$$f(z) = \sum_{j=0}^{N+1} \int \frac{d^j}{dt^j} e^{izt} d\mu_j(t) \, .$$

We obtain a growth bound for the derivatives of such distributions.

Theorem 6.3. *Let f be an entire function that is an Ω-bandlimited distribution of degree N, then there is a constant C_f such that for $n \in \mathbb{N}$, the n-th derivative of f is bounded by*

$$|f^{(n)}(z)| \leq C_f \Omega^n (n + |z|)^{N+1} \, .$$

Proof. From the representation of f,

$$f^{(n)}(z) = \sum_{j=0}^{N+1} \int \frac{d^n}{dz^n} (iz)^j e^{izt} d\mu_j(t) \, .$$

Using the commutation relation $\frac{d}{dz} z = (z+1) \frac{d}{dz}$ repeatedly and estimating the integrand gives

$$|f^{(n)}(z)| \leq \sum_{j=0}^{N+1} \int (n + |z|)^j |t|^n d|\mu_j|(t) \leq \Omega^n \sum_{j=0}^{N+1} (n + |z|)^j \|\mu_j\|_{TV} \, ,$$

where $\|\mu_j\|_{TV}$ is the total variation norm of the signed measure μ_j. Next, replacing the summation index j by $N+1$ and summing the total variation norms gives the claimed bound with $C_f = \sum_{j=0}^{N+1} \|\mu_j\|_{TV}$. □

This permits us to conclude an error bound for the quadrature of functions that are related to bandlimited distributions.

Theorem 6.4. *Let f be a function that is integrable with respect to the n-Gaussian weight and such that $f(t) + f(-t) = 2g(t^{2n})$ where g is an entire function and an Ω-bandlimited distribution* *of degree N, then for any $k \in \mathbb{N}$,*

$$\left| \int_{-\infty}^{\infty} f(t) e^{-\frac{t^{2n}}{n}} dt - \frac{1}{2} \sum_{j=1}^{k} w_j \Big(f(t_j) + f(-t_j) \Big) \right| \leq \frac{C_f \Omega^{2k} (2k + t_k^{2n})^{N+1}}{c_k^2 (2k)!}. \tag{6.43}$$

Proof. We take the error for the generalized Hermite-Gauss quadrature as stated in Equation (6.42). We then define $g(\tau) = (f(\tau^{1/2n}) + f(-\tau^{1/2n}))/2$ which extends to an entire function by assumption. If g is an Ω-bandlimited distribution of degree N, then the growth bound from the preceding theorem applies,

$$|g^{(2k)}(z)| \leq C_g \Omega^{2k} (2k + |z|)^{N+1} .$$

The error bound is maximized for $z = t_k^{2n}$. □

Theorem 6.5. *The spectral radius $\rho(\boldsymbol{T})$ of the tridiagonal matrix $\boldsymbol{T}$ grows at most polynomially with k.*

Proof. It is a well-known fact that the spectral radius of a matrix is bounded by any matrix norm [16]. We will show the the Frobenius norm of the matrix grows polynomially, and therefore the spectral radius must grow at most polynomially. The Frobenius norm of the matrix $\mathbf{T}$ is

$$||\mathbf{T}||_F \equiv \sqrt{\sum_{i=1}^{k} \sum_{j=1}^{k} |T_{i,j}|^2} \tag{6.44}$$

$$= \left[\sum_{i=1}^{k} \left| -\frac{\beta_{i-1}}{4} \right|^2 + 2 \sum_{i=1}^{k-1} \left| \frac{\sqrt{-\gamma_i}}{4} \right|^2 \right]^{\frac{1}{2}} \tag{6.45}$$

$$= \frac{1}{4} \left[\beta_{k-1}^2 + \sum_{i=1}^{k-1} \left(\beta_{i-1}^2 - 2\gamma_i \right) \right]^{\frac{1}{2}} \tag{6.46}$$

$$= \frac{1}{4} \left(4k - 24kn + 24k^2 n + 32kn^2 - 64k^2 n^2 + 32k^3 n^2 \right)^{\frac{1}{2}} \tag{6.47}$$

which is a polynomial in k. Therefore, since $\rho(\mathbf{T}) \leq ||\mathbf{T}||_F$, the spectral radius grows at most polynomially in k. □

In the context of the error expression in (6.43), since the largest eigenvalue λ_k grows at most polynomially, the condition for the error to vanish as $k \to \infty$ occurs when Ω^{2k} decreases more rapidly than c_k^{2k}. We conclude a class of functions for which the error vanishes as the number of sample points increases.

Corollary 6.1. *Let f be a function that is integrable with respect to the n-Gaussian weight and such that $f(t) + f(-t) = 2g(t^{2n})$ where g is an entire function and an Ω-bandlimited distribution of degree N. If $\Omega < 2/n$, then the quadrature error goes to zero as $k \to \infty$.*

Proof. We recall that the normalized polynomial $h_{2k}^{(n)}(t) = A_k H_{2k}^{(n)}(t)$ has a leading coefficient of

$$c_k = 4^k A_k = 2^{\frac{k}{2}} \left[\left(k! n^k \prod_{j=1}^{k-1} [2jn + 1] \right) \left(n^{\frac{1}{2n} - 1} \Gamma(\frac{1}{2n}) \right) \right]^{-\frac{1}{2}}. \tag{6.48}$$

Consider

$$c_k^2 (2k)! = \frac{2^k (2k)!}{\left(k! n^k \prod_{j=1}^{k} [2(j-1)n + 1] \right) \left(n^{\frac{1}{2n} - 1} \Gamma(\frac{1}{2n}) \right)} \tag{6.49}$$

and

$$c_{k+1}^2 (2(k+1))! = \frac{2^{k+1} (2k+2)!}{\left((k+1)! n^{k+1} \prod_{j=1}^{k+1} [2(j-1)n + 1] \right) \left(n^{\frac{1}{2n} - 1} \Gamma(\frac{1}{2n}) \right)} \tag{6.50}$$

as subsequent terms in a sequence. If we consider the ratio between the two and take its limit as $k \to \infty$,

$$\lim_{k \to \infty} \frac{c_{k+1}^2 (2(k+1))!}{c_k^2 (2k)!} = \lim_{k \to \infty} \frac{2(2k+1)(2k+2)}{(k+1)n(2kn+1)} = \frac{4}{n^2}, \tag{6.51}$$

we find that for $n = 1$, the sequence diverges, and for at least $n \geq 3$, the sequence converges exponentially to zero, since for large enough k, the following term is the product of the previous term and a factor approaching $\frac{4}{n^2}$. However, by assumption $\Omega^2 \frac{n^2}{4} < 1$, so we obtain that $\Omega^{2k} / (c_k^2 (2k)!) \to 0$. □

7 Examples

In this section we provide two examples of integrals that can be approximated using generalized Hermite-Gauss quadrature. The first of these examples demonstrates an application of Theorem

5.1, where the approximation of the integral given by the quadrature rule turns out to be exact for a sufficiently large and finite k. The second example demonstrates that, for sufficiently small Ω, the error of the quadrature approximation goes to zero as k increases for cases where the quadrature rule is inexact (Corollary 6.1).

7.1 Example 1: monomials

Consider the integral

$$I = \int_{-\infty}^{\infty} t^{2q} e^{-\frac{t^{2n}}{n}} dt \tag{7.1}$$

where $q = ln$ and $l \in \{0, 1, \ldots, 2k-1\}$, as stated in Theorem 5.1. This integral is compatible with the generalized Hermite-Gauss quadrature rule and, through a change of variables, the improper integral formulation of the gamma function

$$\Gamma(z) = \int_0^{\infty} x^{z-1} e^{-x} dx. \tag{7.2}$$

If we define $x = \frac{t^{2n}}{n}$ and $z = \frac{1}{n}\left(q + \frac{1}{2}\right)$, we get two methods for calculating I:

$$I = n^{z-1}\Gamma(z) = \sum_{j=1}^{k} w_j t_j^{2q}. \tag{7.3}$$

Solving for $\Gamma(z)$ and substituting $q = nz - \frac{1}{2}$:

$$\Gamma(z) = n^{1-z} \sum_{j=1}^{k} w_j t_j^{2nz-1} \tag{7.4}$$

where $z = l + \frac{1}{2n}$.

Figure 1 is a plot that contains these calculations for various n and compares them to the gamma function.

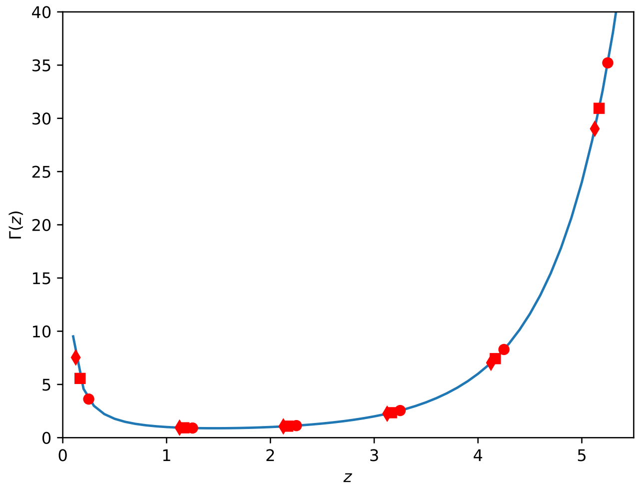

Figure 1: The plot of $\Gamma(z)$ (blue line) compared to the values computed using the quadrature rule for $l \in \{0, 1, 2, 3, 4, 5\}$ and $n \in \{2, 3, 4\}$. $n = 2$ is plotted using circular markers, $n = 3$ using square markers, and $n = 4$ using diamond markers. For very specific values of z the quadrature is exact and so the calculation for $\Gamma(z)$ at those points is also exact.

7.2 Example 2: error bound of successive approximations

Consider the integral

$$I = \int_{-\infty}^{\infty} (3t^{4n} + t^{2n} + 4) \frac{\sin(\Omega t^{2n})}{\Omega t^{2n}} e^{-\frac{t^{2n}}{n}} dt \tag{7.5}$$

where $\Omega = \frac{1}{n}$. Since the integrand meets the condition that $\Omega < \frac{2}{n}$, by Corollary 6.1, the error should vanish as $k \to \infty$. To explicitly calculate the example, take $n = 3$. Then we are able to generate Figs 2 and 3 for the quadrature approximation of the integral I_k and the error of each approximation E_k, respectively. The expression for E_k is given by (6.43).

8 Quadrature in the setting of compressive sensing

So far, we have developed a theory for integrating lacunary polynomials using Gaussian sampling methods. We have shown that for a given polynomial, the number of sample points to integrate that polynomial exactly can be made far fewer than half its order. The underlying scheme of Gauss quadrature is polynomial interpolation - the creation of Lagrange polynomials that match

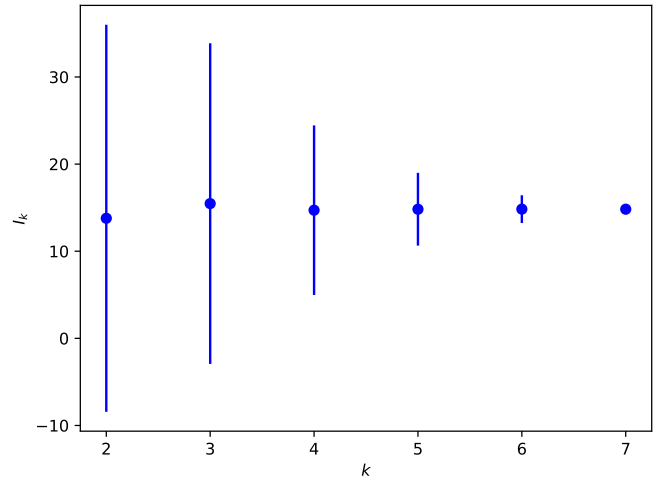

Figure 2: Successive approximations of the integral I in Example 2 and their corresponding errors. Since $\Omega = \frac{1}{n} < \frac{2}{n}$, the error decreases exponentially. For $n = 3$ and $k = 7$, the quadrature approximation yields 14.82.

the target function at specific arguments, particularly at the roots of an orthogonal polynomial. From the error expression we have shown that these sample points are necessarily ideal points for the integration of the target function. This begs the question: what if samples at the roots of the orthogonal polynomial are unavailable? That is, what if we are left with knowledge of the function's values at points that are not the roots of some Hermite polynomial? Is one doomed to sustain some amount of error because Gauss quadrature is no longer an option?

Not quite. At least, not if we assume that our function is sparse in the orthogonal polynomial basis, meaning that it is comprised of a linear combination of only a few Hermite polynomials, and most Hermite polynomials have no contribution. In this section we will demonstrate that in the setting of compressive sensing we can integrate an unknown polynomial exactly with a number of samples less than half its order even if the samples are not at the Gauss quadrature nodes. First, we will establish some major definitions and results that are of interest to us in the field of compressive sensing. Next, we will show that our ability to recover a sparse function in the Hermite polynomial space will largely depend on error expressions in numerical quadrature. This occurs when we compute the coherence, a property of the sensing matrix which ensures sparse

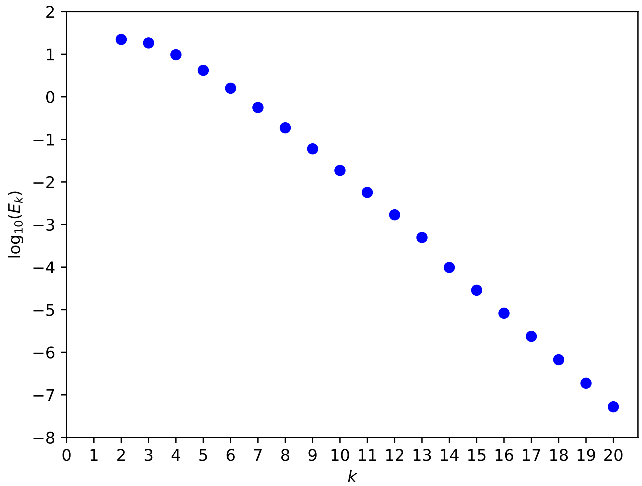

Figure 3: Asymptotically, the error E_k in Example 2 vanishes as $k \to \infty$. We have plotted $\log_{10}(E_k)$ vs. k.

recovery with l_1-minimization techniques. We tackle in detail the uniform sample sparse recovery problem and provide an example for which recovery of the polynomial (and its integral) is possible given sparsity in the Hermite polynomial basis and a number of measurements of the unknown polynomial that is less than the order of the polynomial. Moreover, these measurements do not lie on the Gauss quadrature nodes. Finally, we extend the lessons gleaned from the uniform sample sparse recovery problem to Gauss quadrature and generalized Gauss quadrature, and we compare coherence results.

8.1 Definitions and theorems

Compressive sensing aims to solve the following underdetermined linear system

$$\mathbf{m} = \mathbf{\Phi}\mathbf{a} \tag{8.1}$$

where $\mathbf{m} \in \mathbb{R}^M$, $\mathbf{a} \in \mathbb{R}^N$, and $\mathbf{\Phi} \in \mathbb{R}^{M\times N}$ such that $M < N$. [3] and [8] showed that given knowledge that the solution $\mathbf{a}$ is sparse is some basis, it could be reconstructed with fewer samples than required by the Nyquist-Shannon sampling theorem. $\mathbf{\Phi}$ is called the *sensing matrix*, and $\mathbf{m}$

is the set of measurements that we have on the solution. Before developing our own results in applying compressive sensing to Gauss quadrature, we wish to provide background on a few key concepts. The following results are summarized from [9, 10].

Definition 8.1. *A signal* $\mathbf{a}$ *is called k-sparse if it has at most k nonzero values when represented in a basis* $\mathbf{\Psi}$*. If a signal* $\mathbf{a}$ *is k-sparse we say that* $\mathbf{a} \in \Sigma_k$*.*

The idea behind solving the underdetermined system is that we do not want one set of measurements to correspond to two more more possible solutions. That is, for each $\mathbf{m}$ there ought to exist a unique sparse solution $\mathbf{a} \in \Sigma_k$ to the underdetermined linear system. Let us mathematize this idea by saying for $\mathbf{a}, \mathbf{a}' \in \Sigma_k$ where $\mathbf{a} \neq \mathbf{a}'$, we want

$$\mathbf{\Phi a} \neq \mathbf{\Phi a}'. \tag{8.2}$$

Otherwise, it would be very difficult to differentiate between the two solutions based on the measurements alone. Observe that if we are indeed in that predicament, then

$$\mathbf{\Phi}(\mathbf{a} - \mathbf{a}') = 0 \tag{8.3}$$

with $\mathbf{a} - \mathbf{a}' \in \Sigma_{2k}$ in general. Therefore we can formalize a condition on the sensing matrix $\mathbf{\Psi}$ that will be beneficial to us: that its null space $\mathcal{N}(\mathbf{\Phi})$ contains no vectors in Σ_{2k}. We follow up on this idea with the definition of the *spark* of a matrix.

Definition 8.2. *The spark of a matrix* $\mathbf{\Phi}$*, denoted spark*$(\mathbf{\Phi})$*, is the smallest number of columns of* $\mathbf{\Phi}$ *that are linearly dependent.*

Now, we present a key result which links the spark of a matrix and the uniqueness of a solution $\mathbf{a}$ to a set of measurements $\mathbf{m}$.

Theorem 8.1. *For any vector* $\mathbf{m} \in \mathbb{R}^M$*, there exists at most one signal* $\mathbf{a} \in \Sigma_k$ *(at most one k-sparse signal) such that* $\mathbf{m} = \mathbf{\Phi a}$ *if and only if spark*$(\mathbf{\Phi}) > 2k$*.*

Proof. We prove both directions of the theorem. First, assume that for any $\mathbf{m} \in \mathbb{R}^M$, there exists at most one signal $\mathbf{a} \in \Sigma_k$ such that $\mathbf{m} = \mathbf{\Phi a}$. For the sake of contradiction, suppose $spark(\mathbf{\Phi}) \leq 2k$. Then there exists some set of at most $2k$ columns that are linearly dependent, which implies that there exists an $\mathbf{h} \in \mathcal{N}(\mathbf{\Phi})$ where $\mathbf{h} \in \Sigma_{2k}$. Since $\mathbf{h} \in \Sigma_{2k}$ we can write $\mathbf{h} = \mathbf{a} - \mathbf{a}'$, where $\mathbf{a}, \mathbf{a}' \in \Sigma_k$. Also, since $\mathbf{h} \in \mathcal{N}(\mathbf{\Phi})$ we have that $\mathbf{\Phi}(\mathbf{a} - \mathbf{a}') = 0$ and hence $\mathbf{\Phi a} = \mathbf{\Phi a}'$. This contradicts our assumption that there exists a unique signal $\mathbf{a} \in \Sigma_k$ such that $\mathbf{m} = \mathbf{\Phi a}$. Therefore, it must be that $spark(\mathbf{\Phi}) > 2k$.

Now, suppose that $spark(\mathbf{\Phi}) > 2k$. Assume that for some $\mathbf{m}$, there exist vectors $\mathbf{a}, \mathbf{a}' \in \Sigma_k$ such that $\mathbf{m} = \mathbf{\Phi a} = \mathbf{\Phi a}'$. Therefore $\mathbf{\Phi}(\mathbf{a} - \mathbf{a}') = 0$. Let $\mathbf{h} = \mathbf{a} - \mathbf{a}'$, and write this as $\mathbf{\Phi h} = 0$. Since $spark(\mathbf{\Phi}) > 2k$, all sets of columns up to $2k$ in size are linearly independent, and therefore it must be that $\mathbf{h} = 0$. This implies that $\mathbf{a} = \mathbf{a}'$. $\square$

Remark 8.1 (Measurements and sparsity)**.** *Clearly, since* $\mathbf{\Phi} \in \mathbb{R}^{M \times N}$*,* $spark(\mathbf{\Phi})$ *is upper-bounded such that its maximum value is* $M + 1$*. Therefore, for our sensing matrix to be useful, the theorem requires that* $M + 1 > 2k$*, or* $M \geq 2k$*.*

The spark gives a characterization for when sparse recovery is possible in the case of exactly sparse vectors, but when dealing with *approximately* sparse vectors, we must introduce somewhat more restrictive conditions on the null space of $\mathbf{\Phi}$. Moreover, for practical purposes, an even stronger condition is imposed on the sensing matrix which makes sparse recovery more robust to noise. These two conditions are called the *null space property* (NSP) and *restricted isometry property* (RIP), respectively. Beyond this general description of these two properties, this thesis will not cover the NSP or the RIP. Further reading can be done in [4, 7, 10].

In practice, the spark, NSP, and RIP are difficult to compute, and therefore we turn to the coherence.

Definition 8.3. *The coherence of a matrix* $\mathbf{\Phi}$*, denoted* $\mu(\mathbf{\Phi})$*, is the largest absolute inner product* *between any two distinct normalized columns φ_i, φ_j of $\mathbf{\Phi}$:*

$$\mu(\mathbf{\Phi}) = \max_{1 \le i,j \le N} \frac{|\langle \varphi_i, \varphi_j \rangle|}{||\varphi_i||_2 ||\varphi_j||_2}.$$

We present a useful and well-established lemma, with the proof given by [11], to develop our understanding of the relationship between the spark and coherence of $\mathbf{\Phi}$.

Lemma 8.1. *For any matrix $\mathbf{\Phi}$,*

$$spark(\mathbf{\Phi}) \ge 1 + \frac{1}{\mu(\mathbf{\Phi})}.$$

Proof. First, modify the matrix $\mathbf{\Phi}$ by normalizing its columns to be of unit l_2-norm, and obtain $\tilde{\mathbf{\Phi}}$. This operation preserves both the spark and the coherence. The entries of the resulting Gram matrix,

$$\mathbf{G} = \tilde{\mathbf{\Phi}}^T \tilde{\mathbf{\Phi}}, \tag{8.4}$$

satisfy the following properties:

- $G_{i,i} = 1$, for $1 \le i \le N$, and
- $|G_{i,j}| \le \mu(\mathbf{\Phi})$, for $1 \le i, j \le N$, where $i \ne j$.

Consider an arbitrary leading minor from G of size $p \times p$, built by choosing a subgroup of p columns from $\tilde{\mathbf{\Phi}}$ and computing their sub-Gram matrix. From the Gershgorin disk theorem, if this minor is diagonally-dominant, i.e. if

$$\sum_{j \ne i} |G_{i,j}| < |G_{i,i}| \tag{8.5}$$

for every i, then this submatrix of G is positive definite, and so those p columns from $\tilde{\mathbf{\Phi}}$ are linearly independent. The expression

$$1 > (p-1)\mu(\mathbf{\Phi}) \tag{8.6}$$

implies positive-definiteness of every $p \times p$ minor. Thus, by rearrangement, $p = 1 + \frac{1}{\mu(\mathbf{\Phi})}$ is the smallest number of columns that *might* lead to linear dependence, and thus

$$spark(\mathbf{\Phi}) \geq 1 + \frac{1}{\mu(\mathbf{\Phi})} \tag{8.7}$$

proving the lemma. □

Remark 8.2 (Gershgorin disk theorem)**.** *The expression* $1 > (p-1)\mu(\mathbf{\Phi})$ *was obtained by applying the Gershgorin disk theorem to an arbitrary* $p \times p$ *positive-definite sub-Gram matrix. The left-hand side was obtained by recognizing that* $\tilde{\mathbf{\Phi}}$ *is a normalization of the sensing matrix, so that* $\langle \varphi_i, \varphi_i \rangle = 1$. *The right-hand side recognizes that there are* $(p-1)$ *off-diagonal elements in each row of the sub-Gram matrix, and that each one has a maximum value of* $\mu(\mathbf{\Phi})$.

We develop the final piece of background that will be useful in our approach to applying compressive sensing to quadrature.

Theorem 8.2. *If*

$$k < \frac{1}{2}\left(1 + \frac{1}{\mu(\mathbf{\Phi})}\right),$$

then for each measurement vector $\mathbf{m} \in \mathbb{R}^M$ *there exists at most one signal* $\mathbf{a} \in \Sigma_k$ *such that* $\mathbf{m} = \mathbf{\Phi}\mathbf{a}$.

Proof. From Theorem 8.1 we are guaranteed a unique solution $\mathbf{a} \in \Sigma_k$ to our measurement vector $\mathbf{m} \in \mathbb{R}^M$ if $spark(\mathbf{\Phi}) > 2k$. Also, from Lemma 8.1 we place a lower bound on the spark: $spark(\mathbf{\Phi}) \geq 1+\frac{1}{\mu(\mathbf{\Phi})}$. Therefore, in relating the coherence to the sparsity of $\mathbf{a}$ for which we can guarantee recovery we get the following inequality:

$$spark(\mathbf{\Phi}) \geq 1 + \frac{1}{\mu(\mathbf{\Phi})} \geq 2k. \tag{8.8}$$

Which leads to the following criterion that we wish for our sensing matrix to fulfill so we can guarantee sparse recovery:

$$k < \frac{1}{2}\left(1 + \frac{1}{\mu(\mathbf{\Phi})}\right). \tag{8.9}$$

$\square$

8.2 Sparse recovery with uniform samples

Now that we have developed the necessary background, we present a general problem which can be solved with techniques from compressive sensing. Suppose that we are presented a function f that is a linear combination of Hermite functions:

$$f(x) = \sum_{n=0}^{N-1} b_n H f_n(x). \tag{8.10}$$

where the Hermite function $Hf_j(x) = H_j(x)e^{-\frac{x^2}{2}}$. (We use the regular Hermite polynomials to simplify our problem, but in general the problem can be posed in terms of any set of orthogonal polynomials, including the generalized Hermite polynomials.) However, we are not given the N coefficients in the Hermite polynomial basis. Rather, we are given $M < N$ uniform samples of f starting from the origin, $\mathbf{m} \in \mathbb{R}^M$, such that its elements

$$m_i = f((i-1)\Delta x) \tag{8.11}$$

with $i = 1, 2, ..., M$ and $\Delta x > 0$ fixed. Our goal is to recover the vector of coefficients $\mathbf{a} \in \mathbb{R}^N$, with its elements as the coefficients of f:

$$a_j = b_{j-1} \tag{8.12}$$

with $j = 1, 2, ..., N$. The sensing matrix is

$$\mathbf{\Phi} = \begin{bmatrix} Hf_0(0) & Hf_1(0) & . \ . \ . & Hf_{N-1}(0) \\ Hf_0(\Delta x) & Hf_1(\Delta x) & . \ . \ . & Hf_{N-1}(\Delta x) \\ Hf_0(2\Delta x) & Hf_1(2\Delta x) & . \ . \ . & Hf_{N-1}(2\Delta x) \\ . & . & . \ . \ . & . \\ . & . & . \ . \ . & . \\ . & . & . \ . \ . & . \\ Hf_0((M-1)\Delta x) & Hf_1((M-1)\Delta x) & . \ . \ . & Hf_{N-1}((M-1)\Delta x) \end{bmatrix}. \tag{8.13}$$

Since the Hermite functions are orthogonal on the interval $[0, \infty)$, we can expect that $\mu(\mathbf{\Phi})$ is minimized as M grows larger and Δx gets smaller. We formalize this notion with a theorem.

Theorem 8.3. *The coherence of $\mathbf{\Phi}$, $\mu(\mathbf{\Phi})$, is upper-bounded by the error of a left Riemann sum, i.e.*

$$\mu(\mathbf{\Phi}) \leq \max_{1 \leq i,j \leq N} \frac{\left|\frac{1}{2}M(\Delta x)^2 \frac{d}{dx}\left[Hf_i(x)Hf_j(x)\right]_{x=c} + \tau\right|}{||\varphi_i||_2||\varphi_j||_2} \tag{8.14}$$

where $c = \operatorname{argmax}_{c \geq 0} \frac{d}{dx}\left[Hf_i(x)Hf_j(x)\right]$, and τ is the area unaccounted for by the left Riemann sum scheme due to $M < \infty$.

Proof. Consider the error of interpolating an arbitrary function f over an interval $[a, b]$ using left-sided rectangles of equal width. [15, Section 2.6] states that if f is approximated by a polynomial p of degree n which coincides with f at the $n+1$ distinct points $x_0, x_1, ..., x_n$, then the error E is given by

$$E(x) = f(x) - p(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}\prod_{i=0}^{n}(x - x_i) \tag{8.15}$$

where ξ is a value located in the interval $[a, b]$. Interpolation of f based on its value on the left end of the interval with a zero-order polynomial (a constant) results in

$$E(x) = f(x) - f(a) = f'(\xi)(x - a). \tag{8.16}$$

Integrating both sides over the interval yields the quadrature error for a single rectangle:

$$E_q = \frac{f'(\xi)}{2}(b-a)^2. \tag{8.17}$$

When we compute the inner product between any two columns of $\mathbf{\Phi}$, we effectively compute a left Riemann sum of the product of two Hermite functions. This inner product is

$$\begin{aligned}
|\langle \varphi_i, \varphi_j \rangle| &= \left| \sum_{n=0}^{M-1} Hf_{i-1}(n\Delta x) Hf_{j-1}(n\Delta x) \right| &(8.18)\\
&= \left| \int_0^\infty Hf_i(x)Hf_j(x)dx - \sum_{n=0}^{M-1} \frac{\frac{d}{dx}\left[Hf_i(x)Hf_j(x)\right]_{x=\xi_n}}{2} \left((n+1)\Delta x - n\Delta x\right)^2 - \tau \right| &(8.19)\\
&= \left| \sum_{n=0}^{M-1} \frac{\frac{d}{dx}\left[Hf_i(x)Hf_j(x)\right]_{x=\xi_n}}{2} (\Delta x)^2 + \tau \right|. &(8.20)
\end{aligned}$$

where $\tau = \int_{M\Delta x}^{\infty} Hf_i(x)Hf_j(x)dx$. Let $c = \text{argmax}_{c\geq 0} \frac{d}{dx}\left[Hf_i(x)Hf_j(x)\right]$. Then

$$|\langle \varphi_i, \varphi_j \rangle| \leq \left| \frac{M}{2} \frac{d}{dx} \left[Hf_i(x)Hf_j(x)\right]_{x=c} (\Delta x)^2 + \tau \right|, \tag{8.21}$$

which, when substituted into the expression for $\mu(\mathbf{\Phi})$, proves the theorem. □

Corollary 8.1. *Let*

$$\mu' = \max_{1\leq i,j\leq N} \frac{\left|\frac{1}{2}M(\Delta x)^2 \frac{d}{dx}\left[Hf_i(x)Hf_j(x)\right]_{x=c} + \tau\right|}{||\varphi_i||_2 ||\varphi_j||_2}.$$

If

$$k < \frac{1}{2}\left(1 + \frac{1}{\mu'}\right),$$

then for each measurement vector $\mathbf{m} \in \mathbb{R}^M$ *there exists at most one signal* $\mathbf{a} \in \Sigma_k$ *such that* $\mathbf{m} = \mathbf{\Phi}\mathbf{a}$.

Proof. This follows from $\mu \leq \mu'$, and hence $\frac{1}{2}\left(1+\frac{1}{\mu'}\right) \leq \frac{1}{2}\left(1+\frac{1}{\mu}\right)$. If $k < \frac{1}{2}\left(1+\frac{1}{\mu'}\right)$, then our recovery condition is guaranteed by Theorem 8.2. □

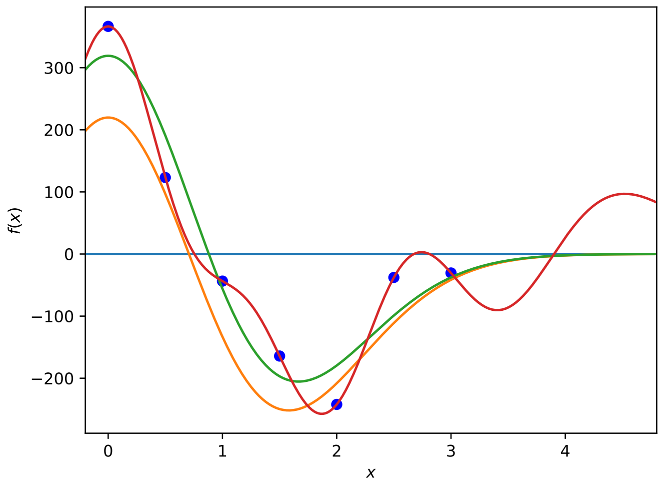

Figure 4: Three iterations (in order: orange, green, red) of orthogonal matching pursuit to recover a 3-sparse function f with $M = 7$ uniform samples (blue points) and $N = 17$. In this example, f is a polynomial of degree 12.

Corollary 8.1 demonstrates the weakness of uniform sampling. The quadrature error of left Riemann sums is controlled merely by a single derivative. For higher order polynomials μ' is poorly bounded. This is especially true for the generalized Hermite polynomials, whose powers increase by $2n$. Nevertheless, we present the sensing matrix for a uniformly sampled function which is a linear combination of generalized Hermite polynomials:

$$\mathbf{\Phi}^{(n)} = \begin{bmatrix} Hf_0^{(n)}(0) & Hf_2^{(n)}(0) & \cdots & Hf_{2N-2}^{(n)}(0) \\ Hf_0^{(n)}(\Delta x) & Hf_2^{(n)}(\Delta x) & \cdots & Hf_{2N-2}^{(n)}(\Delta x) \\ Hf_0^{(n)}(2\Delta x) & Hf_2^{(n)}(2\Delta x) & \cdots & Hf_{2N-2}^{(n)}(2\Delta x) \\ \vdots & \vdots & \cdots & \vdots \\ Hf_0^{(n)}((M-1)\Delta x) & Hf_2^{(n)}((M-1)\Delta x) & \cdots & Hf_{2N-2}^{(n)}((M-1)\Delta x) \end{bmatrix}. \tag{8.22}$$

where $Hf_{2l}^{(n)}(t) = H_{2l}^{(n)}(t)e^{-\frac{t^{2n}}{2n}}$.

Figure 4 demonstrates an example of using an orthogonal matching pursuit algorithm to recover a linear combination of Hermite functions from sparse measurements. The algorithm was taken from [11].

8.3 Sparse recovery with Gauss quadrature nodes

Looking at the possibility of sparse recovery of f with uniform samples, where f is a linear combination of Hermite functions, demonstrated that there is a link between quadrature error and when sparse recovery is guaranteed. This link manifests as the upper bound of the coherence of the sensing matrix $\mu(\mathbf{\Phi}) \leq \mu'$, and in the case of uniform sampling, the coherence is bounded by the error of taking left Riemann sums.

The coherence is a metric which evaluates how effectively we pick our samples, and this idea motivates us to experiment with sparse recovery by sampling at the roots of the normalized Hermite polynomial $h_M(x)$:

$$\mathbf{\Phi} = \begin{bmatrix} h_0(x_0) & h_1(x_0) & . \;\; . \;\; . & h_{N-1}(x_0) \\ h_0(x_1) & h_1(x_1) & . \;\; . \;\; . & h_{N-1}(x_1) \\ h_0(x_2) & h_1(x_2) & . \;\; . \;\; . & h_{N-1}(x_2) \\ . & . & . \;\; . \;\; . & . \\ . & . & . \;\; . \;\; . & . \\ . & . & . \;\; . \;\; . & . \\ h_0(x_{M-1}) & h_1(x_{M-1}) & . \;\; . \;\; . & h_{N-1}(x_{M-1}) \end{bmatrix} \tag{8.23}$$

where $\{x_j\}_{j=0}^{M-1}$ are the M roots of the the M^{th} normalized Hermite polynomial, $h_M(x)$. Note that we have abandoned our use of the Hermite functions in the sensing matrix, and that now we are preferring to use the normalized Hermite polynomials, which are orthonormal on the measure $e^{-x^2}dx$. If $f(x) = \sum_{n=0}^{N-1} b_n h_n(x)$ is our function of interest, then $\mathbf{m} \in \mathbb{R}^M$ is our measurement vector with $m_i = f(x_{i-1})$ and $\mathbf{a} \in \mathbb{R}^N$ our unknown coefficient vector with $a_j = b_{j-1}$. Our matrix equation is then, as usual, $\mathbf{m} = \mathbf{\Phi a}$. However, we ought to take advantage of the discrete inner product naturally provided to us by Gauss quadrature, and thus, we write an $M \times M$ matrix containing the square roots of the quadrature weights on the diagonal and zeros for the off-diagonal elements:

$$\mathbf{W} = \begin{bmatrix} w_0 & 0 & 0 & . & . & . & 0 \\ 0 & w_1 & 0 & . & . & . & 0 \\ 0 & 0 & w_2 & . & . & . & 0 \\ . & . & . & . & & & . \\ . & . & . & & . & & . \\ . & . & . & & & . & . \\ 0 & 0 & 0 & . & . & . & w_{M-1} \end{bmatrix}^{\frac{1}{2}} \tag{8.24}$$

Then, we take both sides of the matrix equation and multiply by $\mathbf{W}$ on the left to obtain

$$\mathbf{Wm} = \mathbf{W\Phi a} \tag{8.25}$$

where $\mathbf{W\Phi}$ is our new sensing matrix and $\mathbf{Wm}$ are our weighted measurements. Like in the previous section, we bound the coherence of this sensing matrix with the error of the quadrature.

Theorem 8.4. *The coherence of* $\mathbf{W\Phi}$*,* $\mu(\mathbf{W\Phi})$*, is upper-bounded by the error of Gauss quadrature with the nodes and weights of* $h_M(x)$*, i.e.*

$$\mu(\mathbf{W\Phi}) \leq \max_{1\leq i,j \leq N} \frac{\left| \frac{1}{c_M^2 (2M)!} \left(\frac{d}{dx}\right)^{2M} [h_i(x)h_j(x)]_{x=\xi} \right|}{||\varphi_i||_2 ||\varphi_j||_2} \tag{8.26}$$

where $\xi = \mathrm{argmax}_{\xi\in\mathbb{R}} \left(\frac{d}{dx}\right)^{2M} [h_i(x)h_j(x)]$.

Proof. Observe that the inner product $|\langle\varphi_i, \varphi_j\rangle|$ is simply an application of the Hermite-Gauss quadrature rule of order M to distinct Hermite polynomials. The theorem follows easily from a similar argument made for Theorem 8.3, where instead, the error in the numerator is substituted with that of Gauss quadrature. $\square$

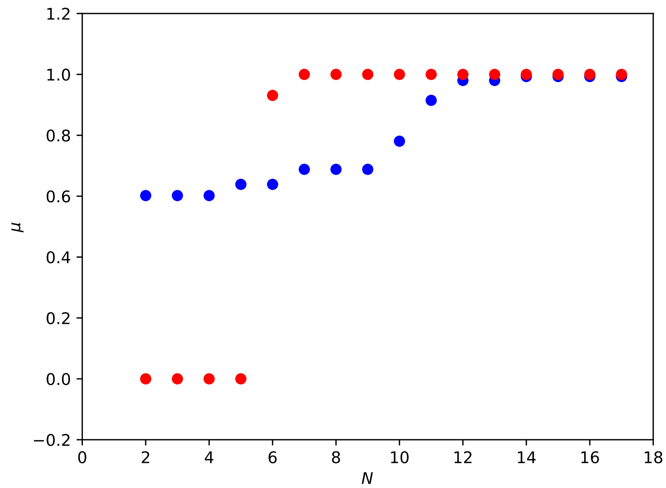

Figure 5: Comparison of coherence values between sensing matrices of uniform (blue) and Gaussian (red) sampling schemes for varying N and fixed M. These values were computed for $M = 5$ measurements and upwards of $N = 17$ polynomials. Note that in the compressive sensing regime (underdetermined system where $N > M$) the Gaussian sample scheme performs poorer (higher coherence) than the uniform sample scheme.

With this formulation it is easy to incorrectly assume that the presence of higher-order derivatives more strongly controls the coherence of the sensing matrix. Actually computing the coherence of the weighted sensing matrix reveals that in fact these Gaussian sample nodes are a terrible choice for compressive sensing - much worse than sampling uniformly. Figure 5 plots the values of the coherence of both the uniform and Gaussian samples for increasing N, the maximum order orthogonal polynomial we include in our basis, and fixed M, . We can see that in the underdetermined case, when $N > M$, Gaussian sampling proves to be absolutely useless for sparse recovery.

The reason Gaussian sampling performs so poorly with compressive sensing is because the three-term recursion relation guarantees that all our orthogonal columns with order $n > M$ are linear combinations of the polynomials with order less than M, with at least the $(M+1)^{st}$ and $(M+2)^{nd}$ columns appearing simply as rescalings of existing columns. Take, for example, a general three-term recursion relation centered about H_M:

$$H_{M+1}(x) = (\alpha_M x + \beta_M)H_M(x) + \gamma_M H_{M-1}(x). \tag{8.27}$$

Let $\{x_j\}_{j=1}^{M}$ be the roots of H_M. Evaluating the three-term relation at the roots yields

$$\begin{aligned} H_{M+1}(x_j) &= (\alpha_M x + \beta_M) H_M(x_j) + \gamma_M H_{M-1}(x_j) \\ &= \gamma_M H_{M-1}(x_j). \end{aligned} \tag{8.28}$$

A similar argument holds for H_{M+2} and H_{M-2}, where we can use the Rodrigues formula implementing the Laplacian operator to write one as a rescaling of the other.

Now in the compressive sensing setting it is easy to see that sampling at the Gaussian nodes causes the column corresponding to H_{M+1} to look exactly like H_{M-1}. Our sensing matrix cannot distinguish between these two elements in our basis. This causes our sensing matrix to have maximum coherence, since there is now a linear dependency that involves a rescaling of an already exisiting column in our matrix. Therefore, sampling at the roots of an orthogonal polynomial results in poor results for sparse recovery guarantees. In the context of Gaussian quadrature, we incorrectly assume that Gaussian quadrature is "more accurate" than standard Riemann sums. This is true in the case where the function we are trying to integrate is of order $2M+1$ (exact approximation), or in the case where the function's coefficients, when expressed as a linear combination of orthogonal polynomials, decay sufficiently rapidly. In the upper bound for the sensing matrix's coherence, neither of these things are true, and therefore uniform samples work better than Gauss samples.

Our exercise in placing upper bounds for the coherence in the previous two sensing matrices teaches us that compressive sensing techniques repurpose the error imposed by polynomial interpolation in our reconstruction of the target function. In the case of classical quadrature, we interpolate the function at a number of samples using polynomials and integrate over the interpolation to get an estimate of the value of the definite integral. This is true of the left Riemann sum, Gaussian, and generalized Gaussian quadratures. However, we showed that in the Gaussian case, the samples poorly approximate the integral of the square of orthogonal polynomials of order higher than the number of measurements. This quadrature error manifests in the coherence of the sensing matrix, which is maximized for $N > M$.

In compressive sensing, these interpolation errors translate not to errors in our computation of the definite integral, but restrictions in the sparsity of our recoverable solutions. Moreover, sparse recovery does not discriminate based on the order of the terms; it only cares if they are nonzero. Hence, with compressive sensing it becomes possible to reconstruct (and subsequently integrate) the unknown function, even if that function is a polynomial of order that far exceeds the number of samples.

9 Conclusions

We have presented Rodrigues formulas for the even and odd generalized Hermite polynomials. We demonstrated that these polynomials have the property of being orthogonal, complete, and normalizable, and we subsequently developed separate three-term recursion relations for the even and odd polynomials. We recognized that any odd component of the function f integrated over the infinite real line with respect to the generalized Gaussian would vanish, and we used that as justification for needing to include only the even polynomials in the matrix equation and Christoffel-Darboux identity. From the tridiagonal matrix in the matrix equation we used the QR algorithm to find the eigenvalues and subsequently the positive real roots of the generalized Hermite polynomial $h_{2k}^{(n)}$, and from the Christoffel-Darboux identity we derived an expression for the weights.

We then determined an expression for the quadrature rule error. In the special case where the function f in the integrand consists only of terms of t shared by $h_{2k}^{(n)}$, the generalized quadrature rule can yield an exact result. Moreover, for Ω-bandlimited distributions where $\Omega < 2/n$, the quadrature approximation error goes to zero as $k \to \infty$. We concluded our discussion by providing two applications of the generalized quadrature rule; the first demonstrated the exactness of the quadrature rule in the context of calculating exact values for the gamma function, while the second numerically verified the exponentially decreasing error bound of the integral approximation. Finally, we demonstrated sparse recovery for linear combinations of Hermite functions and showed that in computing the coherence of the sensing matrix, quadrature error limits sparse recovery guarantees.

Bibliography

[1] Abramowitz, M., and Stegun, I. A. *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*, 10th ed. National Bureau of Standards, Washington D.C., 1972.

[2] Borwein, P., and Erdélyi, T. Generalizations of müntz's theorem via a remez-type inequality for müntz spaces. *Journal of the American Mathematical Society* (1997), 327–349.

[3] Candès, E. J., Romberg, J. K., and Tao, T. Stable signal recovery from incomplete and inaccurate measurements. *Comm. Pure Appl. Math. 59* (2006), 1207–1223.

[4] Candès, E. J., and Tao, T. Decoding by linear programming. *IEEE Trans. Inform. Theory 51*, 12 (2005), 4203–4215.

[5] Chihara, T. S. *Generalized Hermite Polynomials*. PhD thesis, Purdue University, West Lafayette, 1955.

[6] Chihara, T. S. *An Introduction to Orthogonal Polynomials*. Gordon and Breach Science Publishers, New York-London-Paris, 1978.

[7] Donoho, A., Dahmen, W., and DeVore, R. Compressed sensing and best k-term approximation. *J. Amer. Math. Soc. 22*, 1 (2009), 211–231.

[8] Donoho, D. L. Compressed sensing. *IEEE Transactions on Information Theory 52*, 4 (2006), 1289–1306.

[9] Donoho, D. L., and Elad, M. Optimally sparse representation in general (nonorthogonal) dictionaries via l1 minimization. *Proc. Natl. Acad. Sci. 100*, 5 (2003), 2197–2202.

[10] Duarte, M. F. *Introduction to Compressive Sensing*. OpenStax CNX, 2015.

[11] Elad, M. *Sparse and Redundant Representations*. Springer, New York, 2012.

[12] Gautschi, W. Construction of Gauss-Christoffel quadrature formulas. *Math. Comp. 22* (1968), 251–270.

[13] Gautschi, W. Orthogonal polynomials: Applications and computation. *Acta Numer. 5* (1996), 45–119.

[14] Golub, G. H., and Welsch, J. H. Calculation of Gauss quadrature rules. *Math. Comp. 23* (1969), 221–230.

[15] Hildebrand, F. B. *Introduction to Numerical Analysis*. Dover Publications, Inc., New York, 1987.

[16] Horn, R. A., and Johnson, C. R. *Matrix Analysis*, 2nd ed. Cambridge University Press, Cambridge, 2013.

[17] Korevaar, J., and Dixon, M. *Lacunary Polynomial Approximation*. Birkhäuser Basel, Basel, 1978, pp. 479–489.

[18] Rudin, W. *Functional Analysis*, 2nd ed. McGraw-Hill, New York, 1992.

[19] Townsend, A., Trogdon, T., and Olver, S. Fast computation of Gauss quadrature nodes and weights on the whole real line. *IMA J. Numer. Anal. 36* (2016), 337–358.

[20] Trefethen, L. N. Six myths of polynomial interpolation and quadrature. *Math. Today (Southend-on-Sea) 47* (2011), 184–188.

[21] Trogdon, T., and Olver, S. A Riemann-Hilbert approach to Jacobi operators and Gaussian quadrature. *IMA J. Numer. Anal. 36* (2016), 174–196.

[22] Wilf, H. S. *Mathematics for the Physical Sciences*. Dover Publications, Inc., New York, 1978.

[23] Williams, C. L., Bodmann, B. G., and Kouri, D. J. Fourier and beyond: Invariance properties of a family of integral transforms. *J. Fourier Anal. Appl. 23* (2017), 660–678.