

# Pointwise and uniform convergence of Fourier extensions

Marcus Webb<sup>\*1</sup>, Vincent Coppé<sup>†2</sup>, and Daan Huybrechs<sup>‡2</sup>

<sup>1</sup>Department of Mathematics, University of Manchester, United Kingdom

<sup>2</sup>Department of Computer Science, KU Leuven, Belgium

September 4, 2019

## Abstract

Fourier series approximations of continuous but nonperiodic functions on an interval suffer the Gibbs phenomenon, which means there is a permanent oscillatory overshoot in the neighbourhoods of the endpoints. Fourier extensions circumvent this issue by approximating the function using a Fourier series which is periodic on a larger interval. Previous results on the convergence of Fourier extensions have focused on the error in the  $L^2$  norm, but in this paper we analyze pointwise and uniform convergence of Fourier extensions (formulated as the best approximation in the  $L^2$ -norm). We show that the pointwise convergence of Fourier extensions is more similar to Legendre series than classical Fourier series. In particular, unlike classical Fourier series, Fourier extensions yield pointwise convergence at the endpoints of the interval. Similar to Legendre series, pointwise convergence at the endpoints is slower by an algebraic order of a half compared to that in the interior. The proof is conducted by an analysis of the associated Lebesgue function, and Jackson- and Bernstein-type theorems for Fourier extensions. Numerical experiments are provided. We conclude the paper with open questions regarding the regularized and oversampled least squares interpolation versions of Fourier extensions.

**Keywords** Fourier extension, Lebesgue function, Legendre polynomials on a circular arc, constructive approximation

**Mathematics Subject Classification (2010)** 42A10, 41A17, 65T40, 42C15

## 1 Introduction

The Fourier series of a periodic function converges spectrally fast with respect to the number of terms in the series, that is, with an algebraic order which increases with the number of available derivatives and exponentially fast for analytic functions. Furthermore, the truncated Fourier series can be approximated via the FFT in a fast and stable manner [40]. As such, it is the go-to approach to approximate a periodic function. However, when the function in question is nonperiodic, the situation is very different. Regardless of how smooth this function is, convergence is slow in the  $L^2$  norm and there is a permanent oscillatory overshoot close to the endpoints due to the Gibbs phenomenon [42].

Fourier extensions have been shown to be an effective means for the approximation of non-periodic functions while avoiding the Gibbs phenomenon [1, 4, 7, 18, 24, 25, 27]. The idea is as follows: For a function  $f \in L^2(-1, 1)$ , consider an approximant  $f_N$  given by

$$f_N(x) = \sum_{k=-n}^n c_k e^{\frac{i\pi}{T} kx}, \quad N = 2n + 1, \quad (1)$$

---

\*Email: [marcus.webb@manchester.ac.uk](mailto:marcus.webb@manchester.ac.uk). Website: <https://personalpages.manchester.ac.uk/staff/marcus.webb/>.

†Email: [vincent.coppe@cs.kuleuven.be](mailto:vincent.coppe@cs.kuleuven.be). Website: <http://people.cs.kuleuven.be/~vincent.coppe>.

‡Email: [daan.huybrechs@cs.kuleuven.be](mailto:daan.huybrechs@cs.kuleuven.be). Website: <http://people.cs.kuleuven.be/~daan.huybrechs>.

where the coefficients  $c_{-n}, \dots, c_n$  are chosen to minimise the error  $\|f - f_N\|_{L^2(-1,1)}$ , and  $T > 1$  is a user-determined parameter. This approximant  $f_N$  is the  $n$ th Fourier extension of  $f$  to the periodic interval  $[-T, T]$ . For the purposes of this paper, others kinds of Fourier extension, which might come from a discrete sampling of  $f$  or regularization, are a modification of this<sup>1</sup>.

There are many approximation schemes that avoid the Gibbs phenomenon. Chebyshev polynomial interpolants such as those implemented in the Chebfun [14, 36] and ApproxFun [31] software packages are extremely successful, so why consider Fourier extensions? First, discrete collocation versions of Fourier extensions sample the function on equispaced or near-equispaced grids, which in some situations are more natural than Chebyshev grids, which cluster near the endpoints [5]. Second, the approach generalises naturally to higher dimensions. If one has a function on a bounded subset  $\Omega \subset \mathbb{R}^d$ , then one can use multivariate Fourier series which are periodic on a  $d$ -dimensional bounding box containing  $\Omega$  [8, 19, 28]. Modifications of Fourier extensions which use discrete samples of a function are particularly relevant in this generalization, because the integrals defining the  $L^2(\Omega)$  norm can be difficult to compute.

Fourier extensions can be computed stably in  $\mathcal{O}(N \log^2(N))$  floating point operations, with the following important caveats ([21, 27, 24]). Computation of  $f_N$  is equivalent to inversion of the so-called prolate matrix [37], which is a Toeplitz matrix  $G \in \mathbb{R}^{N \times N}$  with entries  $G_{k,j} = \text{sinc}((k-j)\frac{\pi}{T})$ , with right-hand-side vector  $\mathbf{b} \in \mathbb{C}^N$  with entries  $b_k = (\frac{T}{2})^{\frac{1}{2}} \int_{-1}^1 e^{-\frac{i\pi}{T}kx} f(x) dx$  [27]. The prolate matrix is exponentially ill-conditioned [34, Eq. 63], so computation of the exact Fourier extension is practically impossible, even for moderately sized  $N$ . However, a truncated Singular Value Decomposition (SVD) solution is only worse than the exact solution (in the  $L^2(-1, 1)$  norm) by a small factor  $\mathcal{O}(\varepsilon^{\frac{1}{2}})$  in the limit as  $N \rightarrow \infty$ , where  $\varepsilon > 0$  is the truncation parameter [4, 3]. Furthermore, using an oversampled least squares interpolation in equispaced points in  $[-1, 1]$  can bring this down to  $\mathcal{O}(\varepsilon)$  for a sufficient oversampling rate [4, 3, 2]. At the heart of these facts is the observation that while the Fourier basis on  $[-T, T]$  does not form a Schauder basis for  $L^2(-1, 1)$ , it satisfies the weaker conditions of a *frame* [3].

Fourier extensions which approximate a truncated SVD solution rather than the exact solution are called *regularized* Fourier extensions. An approximate SVD of the prolate matrix can be computed in  $\mathcal{O}(N \log^2(N))$  operations using the Fast Fourier Transform (FFT) and exploiting the so-called plunge region in the profile of its singular values [21]. This is a vast improvement on  $\mathcal{O}(N^3)$  operations for a standard SVD. Fast algorithms for regularized, oversampled least squares interpolation Fourier extensions were developed in [27], building on the work of Lyon [24].

Previous convergence results on Fourier extensions have focused on convergence in the  $L^2$  norm, because the Fourier extension by definition minimizes the error in the  $L^2$  norm over the approximation space. Convergence in  $L^2$  of algebraic order  $k$  for functions in the Sobolev space  $H^k(-1, 1)$  was proved by Adcock and Huybrechs [4, Thm. 2.1]. It follows immediately that convergence is superalgebraic for smooth functions. Exponential convergence in  $L^2$  and  $L^\infty$  norms for analytic functions was proved by Huybrechs for  $T = 2$  [18] and by Adcock et al. for general  $T > 1$  [4]. The proofs of exponential convergence appeal to connections between the Fourier extension problem and the sub-range Chebyshev polynomials [4], for which series approximations converge at an exponential rate which depends on analyticity in Bernstein ellipses in the complex plane. Regarding pointwise convergence of Fourier extensions for non-analytic functions, there are no proofs in the literature. Some numerical exploration of pointwise convergence appears in [9, Sec. 2], but a rigorous theoretical foundation is lacking.

## 1.1 Summary of new results

In this paper we prove that for  $f$  in the Hölder space  $C^{k,\alpha}([-1, 1])$ ,

$$f(x) - f_N(x) = \begin{cases} \mathcal{O}(N^{-k-\alpha} \log(N)) & \text{for } x \in [a, b] \subset (-1, 1) \\ \mathcal{O}(N^{\frac{1}{2}-k-\alpha}) & \text{for } x \in [-1, 1], \end{cases} \quad (2)$$

<sup>1</sup>Articles such as [4] refer to this type of Fourier extension as the *exact continuous* Fourier extension.

see Theorem 3.2. The factors of  $\log(N)$  and  $N^{\frac{1}{2}}$  come from bounds on the Lebesgue function associated with Fourier extension derived in Section 4, and the factor of  $N^{-k-\alpha}$  comes from a Jackson-type theorem proved for Fourier extensions derived in Section 5 on best uniform approximation by Fourier extensions.

This factor of  $N^{-k-\alpha}$  can be pessimistic if  $f$  is least regular at the boundary; in Section 5 we discuss how a weighted form of regularity (as opposed to Hölder regularity taken uniformly over the interval  $[-1, 1]$ ) might yield a more natural correspondence between regularity and convergence rate. This is precisely the case in best polynomial approximation on an interval, where weighted moduli of continuity have a tight correspondence with best approximation errors [11, Ch. 7, Thm. 7.7].

From equation (2), it is immediate that if  $f \in C^\alpha([-1, 1])$  where  $\alpha \in (0, 1)$ , then  $f_N$  converges to  $f$  uniformly in any subinterval  $[a, b] \subset (-1, 1)$ , and if  $\alpha > \frac{1}{2}$ , then we get uniform convergence over the whole interval  $[-1, 1]$ .

We also prove a local pointwise convergence result, which states that if  $f \in L^2(-1, 1)$ , but  $f$  is uniformly Dini–Lipschitz in a subinterval  $[a, b]$ , then the Fourier extension converges uniformly in compact subintervals of  $(a, b)$  (see Theorem 3.5). This is done by generalizing a localization theorem of Freud on convergence of orthogonal polynomial expansions in  $[-1, 1]$  (see Section 6).

A key insight of this paper is that the kernel associated with approximation by Fourier extension has an explicit formula which is related to the Christoffel–Darboux kernel of the *Legendre polynomials on a circular arc* (see Lemma 4.3). The asymptotics of these polynomials were derived by Krasovksy using Riemann–Hilbert analysis [22, 23, 10], which we use to derive asymptotics of the kernel. The Lebesgue function for Fourier extensions are estimated using these asymptotics in Theorem 4.1. We find that the Lebesgue function is  $\mathcal{O}(\log(N))$  in the interior of  $[-1, 1]$  and  $\mathcal{O}(N^{\frac{1}{2}})$  globally. This is just as with the Lebesgue function for Legendre series, and distinct from classical Fourier series which has a  $\mathcal{O}(\log N)$  Lebesgue function over the full periodic interval.

The results of this paper would become more interesting when they can be extended to regularized and oversampled interpolation versions of Fourier extensions, because as discussed above, these are the versions for which stable and efficient algorithms have been developed. The multivariate case is another direction this line of inquiry would ideally lead. We briefly discuss future research like this in Section 8.

The paper is structured as follows. Section 2 recounts the known results about convergence of Fourier extensions in the  $L^2$  norm. Section 3 gives new pointwise and uniform convergence theorems along with proofs which depend on results proved in the self-contained Sections 4, 5, and 6. Section 4 is on the Lebesgue function for Fourier extensions. Section 5 is on uniform best approximation for Fourier extensions, in which Jackson- and Bernstein-type theorems are proved. Section 6 is on an analogue of Freud’s localization theorem for Fourier extensions. Section 7 provides the reader with results from numerical experiments, and Section 8 provides discussion. The appendix contains a derivation of asymptotics of Legendre polynomials on a circular arc, on the arc itself, from the Riemann–Hilbert analysis of Krasovksy [22, 23, 10].

## 2 Convergence of Fourier extensions in $L^2$

In this section we summarise the already known results regarding convergence in the  $L^2$  norm.

### 2.1 Exponential convergence

As is discussed in [18, 1], the Fourier extension  $f_N$  in equation (1) is a polynomial in the mapped variable  $t = m(x)$ , where

$$m(x) = 2 \frac{\cos\left(\frac{\pi}{T}x\right) - \cos\left(\frac{\pi}{T}\right)}{1 - \cos\left(\frac{\pi}{T}\right)} - 1.$$

This change of variables transforms the Fourier extension problem into two series expansions in modified Jacobi polynomials [18]. Since exponential convergence in this setting is dictated by

Bernstein ellipses in the complex plane, which are defined to be the closed contours,

$$\mathcal{B}(\rho) = \left\{ \frac{1}{2} \left( \rho e^{i\theta} + \frac{1}{\rho} e^{-i\theta} \right) : \theta \in [-\pi, \pi] \right\}, \quad \rho > 1,$$

it makes sense to consider the mapped contours,

$$\mathcal{D}(\rho) := m^{-1}(\mathcal{B}(\rho)), \quad (3)$$

as a candidate for determining the rate of exponential convergence for Fourier extensions. They are indeed the relevant contours, as was proven in the following theorem.

**Theorem 2.1** (Adcock–Huybrechs [18, Thm. 3.14],[4, Thm. 2.3]). *If  $f$  is an analytic function in  $\mathcal{D}(\rho^*)$  and continuous on  $\mathcal{D}(\rho^*)$  itself, then*

$$\|f - f_N\|_{L^2(-1,1)} = \mathcal{O}(\rho^{-n}) \|f\|_{L^\infty(\mathcal{D}(\rho))},$$

where  $\rho < \min\{\rho^*, \cot^2(\frac{\pi}{4T})\}$  and  $N = 2n + 1$ . The constant in the big  $\mathcal{O}$  depends only on  $T$ .

Note that there is a  $T$ -dependent upper limit on the rate of exponential convergence.

## 2.2 Algebraic convergence

For functions in the Sobolev space  $H^k(-1,1)$  of  $L^2(-1,1)$  functions whose  $k$ th weak derivatives are in  $L^2(-1,1)$ , we have algebraic convergence of order  $k$ .

**Theorem 2.2.** [Adcock–Huybrechs [1, Thm. 2.1]] *If  $f \in H^k(-1,1)$ , then*

$$\|f - f_N\|_{L^2(-1,1)} = \mathcal{O}(N^{-k}) \|f\|_{H^k(-1,1)},$$

where the constant in the big  $\mathcal{O}$  depends only on  $k$  and  $T$ .

**Corollary 2.3.** *If  $f$  is smooth then  $f_N \rightarrow f$  superalgebraically in the  $L^2(-1,1)$  norm.*

## 2.3 Subalgebraic convergence

This elementary result says that Fourier extensions converge in the  $L^2$  norm for  $L^2$  functions.

**Proposition 2.4.** *If  $f \in L^2(-1,1)$ , then*

$$\|f - f_N\|_{L^2(-1,1)} \rightarrow 0 \text{ as } N \rightarrow \infty.$$

*Proof.* Let  $g \in L^2(-T,T)$  be the function that is equal to  $f$  inside  $[-1,1]$  and zero in the complement. Let  $g(x) = \sum_{k=-\infty}^{\infty} c_k e^{\frac{i\pi}{T} kx}$  be its Fourier series, and for all odd integers  $N = 2n + 1$ , define  $t_N(x) = \sum_{k=-n}^n c_k e^{\frac{i\pi}{T} kx}$ . Then following the definitions of  $f_N$ ,  $g$  and  $t_N$ , we have  $\|f - f_N\|_{L^2(-1,1)} \leq \|f - t_N\|_{L^2(-1,1)} = \|g - t_N\|_{L^2(-T,T)} \rightarrow 0$  as  $N \rightarrow \infty$ .  $\square$

## 3 Pointwise and uniform convergence

We prove pointwise convergence rates for functions in various Hölder spaces. For  $k = 0, 1, 2, \dots$  and  $\alpha \in [0, 1]$ , the Hölder space  $C^{k,\alpha}([-1,1])$  is the space,

$$C^{k,\alpha}([-1,1]) := \left\{ f \in C^k([-1,1]) : |f^{(k)}|_{C^\alpha([-1,1])} < \infty \right\},$$

where

$$|g|_{C^\alpha([-1,1])} := \sup_{x,y \in [-1,1]} \frac{|g(x) - g(y)|}{|x - y|^\alpha}.$$

It is a Banach space when endowed with the norm  $\|f\|_{C^{k,\alpha}([-1,1])} = \|f\|_{C^k([-1,1])} + |f^{(k)}|_{C^\alpha([-1,1])}$  [15]. For all  $\alpha \in [0, 1]$ , we have  $C^\alpha([-1,1]) := C^{0,\alpha}([-1,1])$ .

### 3.1 Exponential convergence

The pointwise convergence result for analytic functions is the same as Theorem 2.1. In fact, Theorem 2.1 is a corollary of the following theorem.

**Theorem 3.1** (Huybrechs [18, Theorem 3.14], Adcock–Huybrechs [4, Theorem 2.11], [1, Theorem 2.3]). *If  $f$  is analytic inside of the mapped Bernstein ellipse,  $\mathcal{D}(\rho^*)$  (see equation (3)) and continuous on  $\mathcal{D}(\rho^*)$  itself, then*

$$\|f - f_N\|_{L^\infty(-1,1)} = \mathcal{O}(\rho^{-n})\|f\|_{L^\infty(\mathcal{D}(\rho))},$$

where  $\rho < \min\{\rho^*, \cot^2(\frac{\pi}{4T})\}$  and  $N = 2n + 1$ . The constant in the big  $\mathcal{O}$  depends only on  $T$ .

### 3.2 Algebraic convergence

Pointwise convergence for Hölder continuous functions is as follows.

**Theorem 3.2.** *If  $f \in C^{k,\alpha}([-1,1])$  where  $k \geq 0$  and  $\alpha \in [0,1]$ , then for all  $[a,b] \subset (-1,1)$ ,*

$$\|f - f_N\|_{L^\infty(a,b)} = \mathcal{O}(N^{-\alpha-k} \log N) |f^{(k)}|_{C^\alpha([-1,1])}.$$

The constant in the big  $\mathcal{O}$  depends on  $a, b, k, \alpha$ , and  $T$ . Over the whole interval,  $[-1,1]$ , we have

$$\|f - f_N\|_{L^\infty(-1,1)} = \mathcal{O}(N^{\frac{1}{2}-\alpha-k}) |f^{(k)}|_{C^\alpha([-1,1])}.$$

The constant in the big  $\mathcal{O}$  depends on  $k, \alpha$ , and  $T$ .

We lose a half order of algebraic convergence at the endpoints, something that we could not possibly see in classical Fourier series because a periodic interval has no endpoints.

**Corollary 3.3.** *If  $f$  is smooth then  $f_N \rightarrow f$  superalgebraically in  $L^\infty(-1,1)$ .*

### 3.3 Subalgebraic convergence

The loss of a half order of algebraic convergence at the endpoints predicted by Theorem 3.2 means that we require at least Hölder continuity with order greater than a half in order to guarantee uniform convergence.

**Theorem 3.4.** *If  $f \in C^\alpha([-1,1])$ , where  $\alpha > \frac{1}{2}$ , then*

$$\|f - f_N\|_{L^\infty(-1,1)} \rightarrow 0 \text{ as } N \rightarrow \infty.$$

In order to guarantee local, pointwise convergence, there is a weak local continuity condition which can be employed as follows. A function  $f$  is *uniformly Dini–Lipschitz* in  $[a,b]$  if [42],

$$\lim_{\delta \searrow 0} \sup_{\substack{x,y \in [a,b] \\ |x-y| < \delta}} |(f(x) - f(y)) \log \delta| = 0. \quad (4)$$

This is a very weak condition, weaker than the Hölder condition for any  $\alpha > 0$ , but it is sufficient for convergence of Fourier extensions in the interior of  $[-1,1]$ .

**Theorem 3.5.** *If  $f \in L^2(-1,1)$  is uniformly Dini–Lipschitz in  $[a,b] \subseteq [-1,1]$ , then*

$$\|f - f_N\|_{L^\infty(c,d)} \rightarrow 0 \text{ as } N \rightarrow \infty,$$

for all  $[c,d] \subset (a,b)$ .

*Remark 3.6.* This theorem is stronger than it might appear at first. It says that even if a function is in  $L^2(-1,1)$ , and can have for example jump discontinuities, we will still have pointwise convergence in regions where  $f$  is Dini–Lipschitz. However, the localization theorem (Theorem 6.1) which we use to prove this result, does not give any indication of the *rate* of convergence.

### 3.4 Proofs of the results of this section

For each odd positive integer  $N = 2n + 1$ , let  $P_N$  be the orthogonal projection from  $L^2(-1, 1)$  onto the subspace  $\mathcal{H}_N$ ,

$$\mathcal{H}_N = \text{span}\{e^{\frac{i\pi}{T}kx}\}_{k=-n}^n.$$

Then  $f_N = P_N(f)$ , since  $f_N$  minimizes the  $L^2(-1, 1)$  distance between  $f$  and  $\mathcal{H}_N$ . Let  $\{e_k\}_{k=1}^N$  be *any* orthonormal basis for  $\mathcal{H}_N \subset L^2(-1, 1)$ . Then the kernel,

$$K_N(x, y) = \sum_{k=1}^N e_k(x) \overline{e_k(y)},$$

satisfies

$$P_N f(x) = \int_{-1}^1 K_N(x, y) f(y) dy,$$

for all  $f \in L^2(-1, 1)$ . The *Lebesgue function* for the projection  $P_N$  at a point  $x \in [-1, 1]$  is the  $L^1$  norm of the kernel at  $x$ ,

$$\Lambda(x; P_N) = \int_{-1}^1 |K_N(x, y)| dy.$$

The *best approximation error functional* on  $\mathcal{H}_N$  is defined for all  $f \in C([-1, 1])$  by

$$E(f; \mathcal{H}_N) = \inf_{r_N \in \mathcal{H}_N} \|f - r_N\|_{L^\infty(-1, 1)}. \quad (5)$$

The importance of  $\Lambda(x; P_N)$  and  $E(f; \mathcal{H}_N)$  are encapsulated in Lebesgue's Lemma, which states that for any  $f \in C([-1, 1])$ ,

$$|f(x) - P_N(f)(x)| \leq (1 + \Lambda(x; P_N))E(f; \mathcal{H}_N), \quad (6)$$

for all  $x \in [-1, 1]$  [11, Ch. 2, Prop. 4.1], [32, Thm. 2.5.2].

Now we can proceed to prove the pointwise convergence results stated above. The proofs depend on the content of Sections 4, 5 and 6, which consist of self-contained results.

**Lemma 3.7.** *Let  $f \in C([-1, 1])$ . Then for all closed subsets  $[a, b] \subset (-1, 1)$ , we have*

$$\|f - P_N(f)\|_{L^\infty(a, b)} = \mathcal{O}(\log N)E(f; \mathcal{H}_N),$$

where the constant in the big  $\mathcal{O}$  depends on  $a, b$  and  $T$ . Over the whole interval  $[-1, 1]$ , we have

$$\|f - P_N(f)\|_{L^\infty(-1, 1)} = \mathcal{O}(N^{\frac{1}{2}})E(f; \mathcal{H}_N),$$

where the constant in the big  $\mathcal{O}$  depends only on  $T$ .

*Proof.* By Lebesgue's Lemma, given in equation (6), it suffices to show that  $\sup_{x \in [a, b]} \Lambda(x; P_N) = \mathcal{O}(\log N)$ , and  $\sup_{x \in [-1, 1]} \Lambda(x; P_N) = \mathcal{O}(N^{\frac{1}{2}})$ . This is proved in Theorem 4.1.  $\square$

*Proof of Theorem 3.2.* By Lemma 3.7, it suffices to show that for  $f \in C^{k, \alpha}([-1, 1])$ , we have  $E(f; \mathcal{H}_N) = \mathcal{O}(N^{-k-\alpha}) \|f\|_{C^\alpha([-1, 1])}$ . This follows from Lemma 5.1 and Theorem 5.3.  $\square$

*Proof of Theorem 3.4.* This follows from Theorem 3.2 with  $k = 0$ , because  $N^{\frac{1}{2}-\alpha} \log N \rightarrow 0$  as  $N \rightarrow \infty$  for all  $\alpha > \frac{1}{2}$ .  $\square$

*Proof of Theorem 3.5.* The following proof is an analogue of a proof of Freud for polynomial approximation ([16, Thm. IV.5.6]). Define the functions  $f_1$  and  $f_2$  by

$$f_1(x) = \begin{cases} f(x) & \text{for } x \in [a, b] \\ f(a) & \text{for } x \in [-1, a) \\ f(b) & \text{for } x \in (b, 1], \end{cases}$$

and  $f_2 = f - f_1$ . Since  $f_2$  vanishes in  $[a, b]$  and is in  $L^2(-1, 1)$ , we have by Theorem 6.1 that  $P_N(f_2) \rightarrow 0$  uniformly in all subintervals  $[c, d] \subset (a, b)$ . It is clear by the definition of  $f_1$  and the definition of Dini–Lipschitz continuity in equation (4) that  $f_1$  is also uniformly Dini–Lipschitz in  $[-1, 1]$ . By Lemma 3.7,

$$\|f_1 - P_N(f_1)\|_{L^\infty(c,d)} = \mathcal{O}(\log N)E(f_1; \mathcal{H}_N).$$

By Lemma 5.2 and Theorem 5.3,  $E(f_1; \mathcal{H}_N) = o(1/\log N)$ . This proves that  $P_N(f_1) \rightarrow f_1$  uniformly on all subintervals  $[c, d] \subset (a, b)$ . Now, since  $f = f_1 + f_2$ , we have proved the result.  $\square$

## 4 The Lebesgue function of Fourier extensions

Recall from Section 3 that the kernel associated with the Fourier extension operator  $P_N$  is the bivariate function on  $[-1, 1] \times [-1, 1]$ ,

$$K_N(x, y) = \sum_{k=1}^N e_k(x)\overline{e_k(y)},$$

where  $\{e_k\}_{k=1}^N$  is any orthonormal basis for  $\mathcal{H}_N$ . We call this kernel the *prolate kernel*, because one particular choice of orthonormal basis is the discrete prolate spheroidal wave functions (DPSWFs). These functions, denoted  $\{\xi_{k,N}\}_{k=1}^N$ , are the  $N$  eigenfunctions of a time-band-limiting operator; specifically, there exist eigenvalues  $\{\lambda_{k,N}\}_{k=1}^N$  such that

$$\int_{-1}^1 \xi_{k,N}(y) \frac{\sin\left(\frac{N\pi}{T}(x-y)\right)}{\sin\left(\frac{\pi}{T}(x-y)\right)} dy = \lambda_{k,N} \xi_{k,N}(x),$$

for  $k = 1, \dots, N$ . DPSWFs play an important role in the analysis of perfectly bandlimited and nearly timelimited periodic signals, which was pioneered by Landau, Pollak and Slepian in the 1970s [34]. More recently, they have also been shown to be important for the computation of Fourier extensions, because the regularized version of Fourier extensions projects onto the DP-SWFs  $\xi_{k,N}$  with eigenvalues  $\lambda_{k,N} > \varepsilon$  for a given tolerance  $\varepsilon > 0$  [4, 3]. This is discussed in Section 8.

The key outcome of this section is a proof of the following theorem.

**Theorem 4.1** (Lebesgue function bounds). *(i) For each closed interval  $[a, b] \subset (-1, 1)$ , the Lebesgue function satisfies*

$$\sup_{x \in [a,b]} \Lambda(x; P_N) = \mathcal{O}(\log N).$$

*(ii) Over the whole interval,  $[-1, 1]$ , we have*

$$\sup_{x \in [-1,1]} \Lambda(x; P_N) = \mathcal{O}(N^{\frac{1}{2}}).$$

This will be proved by finding asymptotic formulae for the prolate kernel  $K_N$ . The reader can verify that  $K_N$  is invariant under a change of orthonormal basis for  $\mathcal{H}_N$ , so a suitable choice of basis is desired. We have found that rather than the DPSWF basis, a basis related to orthogonal polynomials on the unit circle have been more amenable to analysis. For  $N = 2n + 1$ , recall the definition of the  $N$ -dimensional space  $\mathcal{H}_N$ ,

$$\mathcal{H}_N = \text{span} \left\{ e^{\frac{i\pi}{T} kx} \right\}_{k=-n}^n.$$

Any function  $r_N \in \mathcal{H}_N$  is of the form

$$r_N(x) = e^{-\frac{i\pi}{T} nx} p_{2n}(e^{\frac{i\pi}{T} x}),$$

where  $p_{2n}$  is a polynomial of degree  $2n$ . Using this idea we prove the following lemma.

**Lemma 4.2** (Orthonormal basis for  $\mathcal{H}_N$ ). *Let  $\{\Pi_k(z)\}_{k=0}^\infty$  be the (normalized) orthogonal polynomials on the unit circle with respect to the weight*

$$f(\theta) = 2T \cdot \chi_{[-\frac{\pi}{T}, \frac{\pi}{T}]}(\theta), \quad \theta \in [-\pi, \pi],$$

i.e. for  $j, k = 0, 1, 2, \dots$ ,

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \overline{\Pi_k(e^{i\theta})} \Pi_j(e^{i\theta}) f(\theta) d\theta = \delta_{j,k}.$$

Then the set

$$\left\{ e^{-\frac{i\pi}{T}nx} \cdot \Pi_k \left( e^{\frac{i\pi}{T}x} \right) \right\}_{k=0}^{2n}$$

forms an orthonormal basis for  $\mathcal{H}_N$ .

*Proof.* By the observation immediately preceding this lemma, the set forms a basis for  $\mathcal{H}_N$  because  $\{\Pi_k\}_{k=0}^{2n}$  forms a basis for polynomials of degree  $2n$ . We need only show its orthonormality with respect to the inner product on  $\mathcal{H}_N$  induced by  $L^2(-1, 1)$ . Let  $j, k \in \{0, \dots, 2n\}$ . Then, making the change of variables  $\theta = \frac{\pi}{T}x$ , we have

$$\begin{aligned} \int_{-1}^1 \overline{e^{-\frac{i\pi}{T}nx} \cdot \Pi_j \left( e^{\frac{i\pi}{T}x} \right)} e^{-\frac{i\pi}{T}nx} \cdot \Pi_k \left( e^{\frac{i\pi}{T}x} \right) dx &= \int_{-1}^1 \overline{\Pi_j \left( e^{\frac{i\pi}{T}x} \right)} \Pi_k \left( e^{\frac{i\pi}{T}x} \right) dx \\ &= \int_{-\frac{\pi}{T}}^{\frac{\pi}{T}} \overline{\Pi_j(e^{i\theta})} \Pi_k(e^{i\theta}) \frac{T}{\pi} d\theta \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \overline{\Pi_j(e^{i\theta})} \Pi_k(e^{i\theta}) f(\theta) d\theta. \end{aligned}$$

By the orthonormal relationship between  $\Pi_k$  and  $\Pi_j$  on the unit circle, the basis is orthonormal on  $[-1, 1]$ .  $\square$

The Christoffel-Darboux formula for orthogonal polynomials on the unit circle states that,

$$\sum_{k=0}^{N-1} \overline{\Pi_k(\zeta)} \Pi_k(z) = \frac{\overline{\Pi_N^*(\zeta)} \Pi_N^*(z) - \overline{\Pi_N(\zeta)} \Pi_N(z)}{1 - \bar{\zeta}z}, \quad z, \zeta \in \mathbb{C}, \quad \bar{\zeta}z \neq 1,$$

where  $\Pi_N^*(z) = z^N \overline{\Pi(\bar{z}^{-1})}$  (which is also a polynomial of degree  $N$ ) [35, Thm 11.42]. On the unit circle itself, where  $z = e^{i\theta}$ ,  $\zeta = e^{i\phi}$ , this reduces, after some elementary manipulations, to

$$\sum_{k=0}^{N-1} \overline{\Pi_k(e^{i\phi})} \Pi_k(e^{i\theta}) = e^{i\frac{N-1}{2}(\theta-\phi)} \cdot \text{Imag} \left( \frac{e^{-i\frac{N}{2}\phi} \cdot \Pi_N(e^{i\phi}) \cdot e^{-i\frac{N}{2}\theta} \cdot \Pi_N(e^{i\theta})}{\sin\left(\frac{\theta-\phi}{2}\right)} \right). \quad (7)$$

From this general formula for orthogonal polynomials on the unit circle, we prove the following lemma regarding the prolate kernel.

**Lemma 4.3** (Prolate kernel formula). *For all  $x, y \in [-1, 1]$ ,*

$$K_N(x, y) = \text{Imag} \left( \frac{\overline{e^{-\frac{i\pi}{T}\frac{N}{2}y} \cdot \Pi_N \left( e^{\frac{i\pi}{T}y} \right)} \cdot e^{-\frac{i\pi}{T}\frac{N}{2}x} \cdot \Pi_N \left( e^{\frac{i\pi}{T}x} \right)}{\sin\left(\frac{\pi}{2T}(x-y)\right)} \right).$$

The formula in fact holds for all  $x, y \in [-T, T]$ .

*Remark 4.4.* Setting  $T = 1$  in this formula returns the Dirichlet kernel of classical Fourier series, because  $\Pi_N(z) = z^N$  for the trivial weight  $f(\theta) \equiv 1$ .



*Proof.* By the fact that  $\left\{e^{-\frac{i\pi}{T}nx} \cdot \Pi_k\left(e^{\frac{i\pi}{T}x}\right)\right\}_{k=0}^{2n}$  is an orthonormal basis for  $\mathcal{H}_N$ , from Lemma 4.2, we have that

$$\begin{aligned} K_N(x, y) &= \sum_{k=0}^{2n} \overline{e^{-\frac{i\pi}{T}ny} \Pi_k\left(e^{\frac{i\pi}{T}y}\right)} e^{-\frac{i\pi}{T}nx} \Pi_k\left(e^{\frac{i\pi}{T}x}\right) \\ &= e^{\frac{i\pi}{T}n(y-x)} \sum_{k=0}^{N-1} \overline{\Pi_k\left(e^{\frac{i\pi}{T}y}\right)} \Pi_k\left(e^{\frac{i\pi}{T}x}\right). \end{aligned}$$

The proof is completed by considering the Christoffel-Darboux formula for orthogonal polynomials on the unit circle in equation (7) (note that  $\frac{N-1}{2} = n$ ).  $\square$

Now, to ascertain asymptotics of the prolate kernel, it is sufficient to ascertain asymptotics of the orthogonal polynomials  $\{\Pi_k(z)\}_{k=0}^{\infty}$ . These polynomials have been studied before in the literature, and are known as the Legendre polynomials on a circular arc [26].

**Theorem 4.5.** *Let  $\{\Pi_k\}_{k=0}^{\infty}$  be the (normalized) orthogonal polynomials on the unit circle with respect to the weight  $f(\theta) = 2T \cdot \chi_{[-\frac{\pi}{2T}, \frac{\pi}{2T}]}(\theta)$ , and for  $x \in [-1, 1]$  define the variable  $\eta \in [0, \pi]$  by*

$$\eta = \cos^{-1}\left(\frac{\sin\left(x\frac{\pi}{2T}\right)}{\sin\left(\frac{\pi}{2T}\right)}\right).$$

There exists a constant  $\delta > 0$  such that for  $x \in [-1 + \delta, 1 - \delta]$ ,

$$\begin{aligned} \Pi_N\left(e^{\frac{i\pi}{T}x}\right) &= \frac{e^{\frac{i\pi}{T}\frac{N}{2}x}}{\sqrt{2T \sin\left(\frac{\pi}{2T}\right)}} \left( e^{-\frac{i\pi}{4T}} \left( \frac{\sin\left((1+x)\frac{\pi}{2T}\right)}{\sin\left((1-x)\frac{\pi}{2T}\right)} \right)^{\frac{1}{4}} \cos\left(N\eta - \frac{\pi}{4}\right) \right. \\ &\quad \left. - e^{\frac{i\pi}{4T}} \left( \frac{\sin\left((1-x)\frac{\pi}{2T}\right)}{\sin\left((1+x)\frac{\pi}{2T}\right)} \right)^{\frac{1}{4}} \sin\left(N\eta - \frac{\pi}{4}\right) \right) + \mathcal{O}(N^{-1}), \end{aligned} \quad (8)$$

and for  $x \in [1 - \delta, 1]$ ,

$$\begin{aligned} \Pi_N\left(e^{\frac{i\pi}{T}x}\right) &= \frac{e^{\frac{i\pi}{T}\frac{N}{2}x}}{\sqrt{2T \sin\left(\frac{\pi}{2T}\right)}} \left( \frac{\pi}{2} N\eta \right)^{\frac{1}{2}} \left( e^{-\frac{i\pi}{4T}} \left( \frac{\sin\left((1+x)\frac{\pi}{2T}\right)}{\sin\left((1-x)\frac{\pi}{2T}\right)} \right)^{\frac{1}{4}} J_0(N\eta) \right. \\ &\quad \left. - e^{\frac{i\pi}{4T}} \left( \frac{\sin\left((1-x)\frac{\pi}{2T}\right)}{\sin\left((1+x)\frac{\pi}{2T}\right)} \right)^{\frac{1}{4}} J_1(N\eta) \right) + \mathcal{O}(N^{-\frac{1}{2}}). \end{aligned} \quad (9)$$

The constants in the big  $\mathcal{O}$  depend only on  $T$  and  $\delta$ . The asymptotics for  $x \in [-1, -1 + \delta]$  are found by using the relation  $\Pi_N\left(e^{-\frac{i\pi}{T}x}\right) = \overline{\Pi_N\left(e^{\frac{i\pi}{T}x}\right)}$ .

In terms of magnitude with respect to  $N$ , we have

$$\Pi_N\left(e^{\frac{i\pi}{T}x}\right) = \begin{cases} \mathcal{O}(1) & \text{for } x \in [-1 + \delta, 1 - \delta], \\ \mathcal{O}(N^{\frac{1}{2}}) & \text{for } x \in [-1, -1 + \delta] \cup [1 - \delta, 1]. \end{cases} \quad (10)$$

*Remark 4.6.* The asymptotic order of  $\Pi_N\left(e^{\frac{i\pi}{T}x}\right)$  with respect to  $N$  in equation (10) is the same as for the  $N$ th (normalized) Legendre polynomial in  $[-1, 1]$  [35, Thm. 8.21.6]. Further discussion on how Legendre series approximations compare to Fourier extensions lies in subsection 8.1.

*Proof.* This result follows directly from Lemma A.1, because if we take  $\alpha = \pi - \pi/T$  and  $f_\alpha(\theta) \equiv 1$ , then the polynomials  $\Pi_N(z) = (2T)^{-\frac{1}{2}} \phi_N(-z, \alpha)$  satisfy the orthonormality conditions that define

$\Pi_N$  as in Lemma 4.2. To obtain the asymptotic formula above, make the change of variables  $\theta = \frac{\pi}{T}x + \pi$  in the asymptotic formulae for  $\phi_N(z, \alpha)$ . Be careful to note that the endpoint with explicit formula given above ( $x = 1$ ), corresponds to  $\theta = 2\pi - \alpha$ , which is not the endpoint with explicit formula given in Lemma A.1 ( $\theta = \alpha$ ). This was done to shorten the expressions for the asymptotics at the endpoints.

To complete the proof we must prove equation (10). For  $x \in [-1 + \delta, 1 - \delta]$ , all of the terms are clearly bounded above by  $\mathcal{O}(\delta^{-\frac{1}{4}}) = \mathcal{O}(1)$ . Now let  $x \in [1 - \delta, 1]$ . We have  $\eta^2 \leq \frac{\pi^2}{4}(1 - \cos(\eta))$  for all  $\eta \in [0, \frac{\pi}{2}]$  and  $x \sin(\frac{\pi}{2T}) \leq \sin(x\frac{\pi}{2T}) \leq x\frac{\pi}{2T}$  for all  $x \in [0, 1]$ . Assuming  $\delta < \frac{1}{2}$ , we have  $x, y \in [0, 1]$  and  $\eta, \lambda \in [0, \frac{\pi}{2}]$ , and hence  $\eta^2 \leq \frac{\pi^2}{4}(1 - x)$ . Since  $1 - x \in [0, 1]$  we have  $1 - x \leq \sin((1 - x)\frac{\pi}{2T}) / \sin(\frac{\pi}{2T})$ , so  $\eta^2 \leq \sin((1 - x)\frac{\pi}{2T}) \frac{\pi^2}{4 \sin(\frac{\pi}{2T})}$ . This implies that

$$\left( \frac{\eta^2 \sin((1 + x)\frac{\pi}{2T})}{\sin((1 - x)\frac{\pi}{2T})} \right)^{\frac{1}{4}} = \mathcal{O}(1),$$

uniformly for all  $x \in [0, 1]$ . Note also that Bessel functions are uniformly bounded in absolute value by 1 (see [13, Eq. 10.14.1]). This makes it clear that  $\Pi_N(e^{\frac{i\pi}{T}x}) = \mathcal{O}(N^{\frac{1}{2}})$  for  $x \in [1 - \delta, 1]$ .

For  $x \in [-1, -1 + \delta]$ , use the relation  $\Pi_N(e^{-\frac{i\pi}{T}x}) = \overline{\Pi_N(e^{\frac{i\pi}{T}x})}$ .  $\square$

We now have the required results to prove Theorem 4.1.

*Proof of Theorem 4.1 part (i).* Let  $[a, b] \subset (-1, 1)$  and choose  $\tau > 0$  sufficiently small so that  $[a - \tau, b + \tau] \subset (-1, 1)$ . Applying the first part of Theorem 4.5 gives us

$$\Pi_N(e^{\frac{i\pi}{T}x}) = \mathcal{O}(1), \quad x \in [a - \tau, b + \tau]. \quad (11)$$

For the proof of part (i) we need to bound the integral  $\int_{-1}^1 |K_N(x, y)| dy$  uniformly for  $x \in [a, b]$ . We do so by dividing the interval  $[-1, 1]$  into the following subsets:

$$\begin{aligned} I_1 &= \{y \in [-1, 1] : |y - x| \leq N^{-1}\} \\ I_2 &= \{y \in [-1, 1] : N^{-1} < |y - x| \leq \tau\} \\ I_3 &= \{y \in [-1, 1] : \tau < |y - x|\}. \end{aligned}$$

We will obtain estimates for the kernel for  $x \in [a, b]$  and  $y$  in each of  $I_1$ ,  $I_2$  and  $I_3$ , and then estimate the associated integral over each of  $I_1$ ,  $I_2$  and  $I_3$ .

For  $N > 1/\tau$ , we have  $I_1$  and  $I_2$  are nonempty and are contained within  $[a - \tau, b + \tau] \subset (-1, 1)$ . By equation (11),

$$\Pi_N(e^{\frac{i\pi}{T}y}) = \mathcal{O}(1), \quad y \in I_1 \cup I_2.$$

For  $y \in I_1$ , we have

$$K_N(x, y) = e^{\frac{i\pi}{T}n(y-x)} \sum_{k=0}^{N-1} \overline{\Pi_k(e^{\frac{i\pi}{T}y})} \Pi_k(e^{\frac{i\pi}{T}x}) = \mathcal{O}(N).$$

This implies

$$\int_{I_1} |K_N(x, y)| dy \leq \mathcal{O}(N) \int_{I_1} dy = \mathcal{O}(1),$$

because  $|I_1| \leq 2N^{-1}$ .

By Lemma 4.3,

$$K_N(x, y) = \text{Imag} \left( \frac{e^{-\frac{i\pi}{T} \frac{N}{2} y} \cdot \Pi_N(e^{\frac{i\pi}{T}y}) \cdot e^{-\frac{i\pi}{T} \frac{N}{2} x} \cdot \Pi_N(e^{\frac{i\pi}{T}x})}{\sin(\frac{\pi}{2T}(x - y))} \right).$$

Note that since the sine function is concave in  $[0, \pi]$ , we have  $|\sin(\frac{\pi}{2T}(x-y))| \geq \sin(\frac{\pi}{2T})|x-y|$  for  $x, y \in [-1, 1]$ . Therefore, for all  $y \in [-1, 1]$ ,

$$|K_N(x, y)| \leq \mathcal{O}(1) \frac{1}{|x-y|} \left| \Pi_N \left( e^{\frac{i\pi}{T}y} \right) \right|$$

For  $y \in I_2$ , this can be reduced to  $|K_N(x, y)| \leq \mathcal{O}(1) \frac{1}{|x-y|}$ . Therefore,

$$\int_{I_2} |K_N(x, y)| dy \leq \mathcal{O}(1) \int_{I_2} \frac{1}{|x-y|} dy \leq \mathcal{O}(1) \int_{N^{-1}}^{\tau} \frac{1}{s} ds = \mathcal{O}(\log(N)).$$

For  $y \in I_3$ , since  $|x-y|^{-1} < \tau^{-1} = \mathcal{O}(1)$ , we have  $|K_N(x, y)| \leq \mathcal{O}(1) \left| \Pi_N \left( e^{\frac{i\pi}{T}y} \right) \right|$ . Therefore,

$$\int_{I_3} |K_N(x, y)| dy \leq \mathcal{O}(1) \int_{I_3} \left| \Pi_N \left( e^{\frac{i\pi}{T}y} \right) \right| dy \leq \mathcal{O}(1) \left( \int_{-1}^1 \left| \Pi_N \left( e^{\frac{i\pi}{T}y} \right) \right|^2 dy \right)^{\frac{1}{2}} = \mathcal{O}(1).$$

This proves that  $\Lambda(x; P_N) = \mathcal{O}(\log(N))$  uniformly for  $x \in [a, b]$ .  $\square$

*Proof of Theorem 4.1 part (ii).* Let  $\delta \in (0, \frac{1}{4})$  be sufficiently small so that Theorem 4.5 applies to the intervals  $[-1+2\delta, 1-2\delta]$  and  $[1-2\delta, 1]$ . Using part (i) of the present theorem, we have that for all  $x \in [-1+\delta, 1-\delta]$  the Lebesgue function satisfies  $\Lambda(x; P_N) = \mathcal{O}(\log(N)) = \mathcal{O}(N^{\frac{1}{2}})$  uniformly in such  $x$ . Now, since  $\Pi_N \left( e^{-\frac{i\pi}{T}x} \right) = \overline{\Pi_N \left( e^{\frac{i\pi}{T}x} \right)}$ , it follows that  $K_N(-x, y) = \overline{K_N(x, -y)}$ , so that  $\Lambda(-x; P_N) = \Lambda(x; P_N)$ . Therefore, to complete the proof we need only show that  $\Lambda(x; P_N) = \mathcal{O}(N^{\frac{1}{2}})$  uniformly for  $x \in [1-\delta, 1]$ . For such  $x$ , we divide the interval  $[-1, 1]$  into the following subsets:

$$\begin{aligned} I_1 &= \{y \in [-1, 1] : |y-x| \leq N^{-1} \text{ or } |1-y| \leq N^{-1}\} \\ I_2 &= \{y \in [-1, 1] : N^{-1} < |y-x| \leq \delta \text{ and } |1-y| > N^{-1}\} \\ I_3 &= \{y \in [-1, 1] : \delta < |y-x| \text{ and } |1-y| > N^{-1}\}. \end{aligned}$$

By Theorem 4.5,

$$\Pi_N \left( e^{\frac{i\pi}{T}x} \right) = \mathcal{O}(N^{\frac{1}{2}}), \quad x \in [-1, 1].$$

Therefore,

$$K_N(x, y) = e^{\frac{i\pi}{T}n(y-x)} \sum_{k=0}^{N-1} \overline{\Pi_k \left( e^{\frac{i\pi}{T}y} \right)} \Pi_k \left( e^{\frac{i\pi}{T}x} \right) = \mathcal{O}(N^2).$$

By the Cauchy-Schwarz inequality and the fact that  $|I_1| \leq 3N^{-1}$ , we have

$$\int_{I_1} |K_N(x, y)| dy \leq \left( \int_{I_1} |K_N(x, y)|^2 dy \right)^{\frac{1}{2}} \left( \int_{I_1} dy \right)^{\frac{1}{2}} \leq \left( \frac{3}{N} \right)^{\frac{1}{2}} \left( \int_{-1}^1 |K_N(x, y)|^2 dy \right)^{\frac{1}{2}}.$$

By the connection between  $K_N$  and  $P_N$ ,  $\int_{-1}^1 |K_N(x, y)|^2 dy = P_N \left( \overline{K_N(x, \cdot)} \right) (x)$ . Since  $\overline{K_N(x, y)} = K_N(y, x)$  and because  $K_N(\cdot, x) \in \mathcal{H}_N$  for each  $x \in [-1, 1]$ , we have,

$$\int_{-1}^1 |K_N(x, y)|^2 dy = K_N(x, x).$$

Therefore,

$$\int_{I_1} |K_N(x, y)| dy = \mathcal{O}(N^{-\frac{1}{2}}) (\mathcal{O}(N^2))^{\frac{1}{2}} = \mathcal{O}(N^{\frac{1}{2}}).$$

Just as in the proof of part (i) of the theorem, but this time using the estimate  $\Pi_N \left( e^{\frac{i\pi}{T}x} \right) = \mathcal{O}(N^{\frac{1}{2}})$ , we have for all  $x, y \in [-1, 1]$ ,

$$|K_N(x, y)| \leq \mathcal{O}(N^{\frac{1}{2}}) \frac{1}{|x - y|} \left| \Pi_N \left( e^{\frac{i\pi}{T}y} \right) \right|.$$

Therefore, for  $y \in I_3$ ,

$$|K_N(x, y)| \leq \mathcal{O}(N^{\frac{1}{2}}) \left| \Pi_N \left( e^{\frac{i\pi}{T}y} \right) \right|,$$

because  $|x - y| > \delta$  for  $y \in I_3$ . Hence,

$$\int_{I_3} |K_N(x, y)| dy \leq \mathcal{O}(N^{\frac{1}{2}}) \int_{I_3} \left| \Pi_N \left( e^{\frac{i\pi}{T}y} \right) \right| dy \leq \mathcal{O}(N^{\frac{1}{2}}) \left( \int_{-1}^1 \left| \Pi_N \left( e^{\frac{i\pi}{T}y} \right) \right|^2 dy \right)^{\frac{1}{2}} = \mathcal{O}(N^{\frac{1}{2}}).$$

All that remains is to show that  $\int_{I_2} |K_N(x, y)| dy = \mathcal{O}(N^{\frac{1}{2}})$  uniformly for  $x \in [1 - \delta, 1]$ . For  $x \in [1 - \delta, 1]$  and  $y \in I_2$ , we have  $y \in [1 - 2\delta, 1]$  so that the asymptotic expression in Theorem 4.5 holds. Define the variables

$$\eta = \cos^{-1} \left( \frac{\sin \left( x \frac{\pi}{2T} \right)}{\sin \left( \frac{\pi}{2T} \right)} \right), \quad \lambda = \cos^{-1} \left( \frac{\sin \left( y \frac{\pi}{2T} \right)}{\sin \left( \frac{\pi}{2T} \right)} \right).$$

Take the asymptotic expressions for  $\Pi_N$  in Theorem 4.5 for the  $x$  and  $y$  currently in question, and consider the numerator in the formula for the kernel  $K_N(x, y)$  (Lemma 4.3). An asymptotic formula is as follows:

$$\text{Imag} \left( \overline{e^{-\frac{i\pi}{T} \frac{N}{2} y} \cdot \Pi_N \left( e^{\frac{i\pi}{T} y} \right)} \cdot e^{-\frac{i\pi}{T} \frac{N}{2} x} \cdot \Pi_N \left( e^{\frac{i\pi}{T} x} \right) \right) \quad (12)$$

$$\begin{aligned} &= \frac{1}{2T} \left( \frac{\pi}{2} N \eta \right)^{\frac{1}{2}} \left( \frac{\pi}{2} N \lambda \right)^{\frac{1}{2}} \\ &\quad \cdot \left( \left( \frac{\sin \left( (1+x) \frac{\pi}{2T} \right)}{\sin \left( (1-x) \frac{\pi}{2T} \right)} \right)^{\frac{1}{4}} J_0(N\eta) \left( \frac{\sin \left( (1-y) \frac{\pi}{2T} \right)}{\sin \left( (1+y) \frac{\pi}{2T} \right)} \right)^{\frac{1}{4}} J_1(N\lambda) \right. \\ &\quad \left. - \left( \frac{\sin \left( (1-x) \frac{\pi}{2T} \right)}{\sin \left( (1+x) \frac{\pi}{2T} \right)} \right)^{\frac{1}{4}} J_1(N\eta) \left( \frac{\sin \left( (1+y) \frac{\pi}{2T} \right)}{\sin \left( (1-y) \frac{\pi}{2T} \right)} \right)^{\frac{1}{4}} J_0(N\lambda) \right) + \mathcal{O}(1) \end{aligned} \quad (13)$$

$$\begin{aligned} &= N^{\frac{1}{2}} \frac{\pi}{4T} \\ &\quad \cdot \left( \left( \frac{\eta^2 \sin \left( (1+x) \frac{\pi}{2T} \right)}{\sin \left( (1-x) \frac{\pi}{2T} \right)} \right)^{\frac{1}{4}} J_0(N\eta) \left( \frac{\sin \left( (1-y) \frac{\pi}{2T} \right)}{\sin \left( (1+y) \frac{\pi}{2T} \right)} \right)^{\frac{1}{4}} (N\lambda)^{\frac{1}{2}} J_1(N\lambda) \right. \\ &\quad \left. - \left( \frac{\sin \left( (1-x) \frac{\pi}{2T} \right)}{\sin \left( (1+x) \frac{\pi}{2T} \right)} \right)^{\frac{1}{4}} (N\eta)^{\frac{1}{2}} J_1(N\eta) \left( \frac{\lambda^2 \sin \left( (1+y) \frac{\pi}{2T} \right)}{\sin \left( (1-y) \frac{\pi}{2T} \right)} \right)^{\frac{1}{4}} J_0(N\lambda) \right) + \mathcal{O}(1). \end{aligned} \quad (14)$$

This was an important step in the proof, because there was cancellation when we took the imaginary part. This cancellation is essential for the result to hold, and it is the reason for deriving and including a fully explicit description of the leading order asymptotics of the polynomials in Appendix A.

We will now proceed to find upper bounds on the resulting expression. We showed in the proof of Theorem 4.5 that,

$$\left( \frac{\eta^2 \sin \left( (1+x) \frac{\pi}{2T} \right)}{\sin \left( (1-x) \frac{\pi}{2T} \right)} \right)^{\frac{1}{4}} = \mathcal{O}(1),$$

uniformly for all  $x \in [0, 1]$ . The same is true when  $x$  and  $\eta$  are replaced by  $y$  and  $\lambda$ .

It is straightforward to also show that  $(1-y)\frac{\pi}{2T} \leq \sin\left(\frac{\pi}{2T}\right)\lambda^2$  for  $y \in [0, 1]$  and  $\lambda \in [0, \frac{\pi}{2}]$ . From this, we have that for  $y \in I_2$ ,  $\lambda \geq \sqrt{\frac{\pi}{2TN}}$ . Combining this with the fact that for  $t \rightarrow \infty$ ,  $J_\alpha(t) = \mathcal{O}\left(t^{-\frac{1}{2}}\right)$  (see [13, Eq. 10.17.3]) we get that  $J_0(N\lambda) = \mathcal{O}\left(N^{-\frac{1}{4}}\right)$ .

Note also that Bessel functions are uniformly bounded in absolute value by 1 (see [13, Eq. 10.14.1]). Furthermore, as  $t \rightarrow \infty$ ,  $t^{\frac{1}{2}}J_\alpha(t) = \mathcal{O}(1)$  (see [13, Eq. 10.17.3]). Collecting the bounds mentioned in the last three paragraphs, we conclude that for  $y \in I_2$ , we have,

$$\text{Imag}\left(\overline{e^{-\frac{i\pi}{T}\frac{N}{2}y} \cdot \Pi_N\left(e^{\frac{i\pi}{T}y}\right)} \cdot e^{-\frac{i\pi}{T}\frac{N}{2}x} \cdot \Pi_N\left(e^{\frac{i\pi}{T}x}\right)\right) = \mathcal{O}(N^{\frac{1}{2}})J_0(N\eta)(1-y)^{\frac{1}{4}} + \mathcal{O}(N^{\frac{1}{4}}). \quad (15)$$

To conclude, we prove two refinements of equation (15), depending on whether  $x \in [1-\delta, 1-N^{-1}]$  or  $x \in [1-N^{-1}, 1]$ . When  $x \in [1-\delta, 1-N^{-1}]$ , we have  $J_0(N\eta) = \mathcal{O}(N^{-\frac{1}{4}})$  (just like for  $y \in I_2$  discussed above), and so,

$$\text{Imag}\left(\overline{e^{-\frac{i\pi}{T}\frac{N}{2}y} \cdot \Pi_N\left(e^{\frac{i\pi}{T}y}\right)} \cdot e^{-\frac{i\pi}{T}\frac{N}{2}x} \cdot \Pi_N\left(e^{\frac{i\pi}{T}x}\right)\right) = \mathcal{O}(N^{\frac{1}{4}}).$$

This implies that  $K_N(x, y) = \mathcal{O}\left(\frac{N^{\frac{1}{4}}}{|x-y|}\right)$  for  $x \in [1-\delta, 1-N^{-1}]$  and  $y \in I_2$ . Therefore,

$$\int_{I_2} |K_N(x, y)| dy \leq \mathcal{O}(N^{\frac{1}{4}}) \int_{I_2} \frac{1}{|x-y|} dy \leq \mathcal{O}(N^{\frac{1}{4}}) \int_{N^{-1}}^{\delta} \frac{1}{s} ds = \mathcal{O}(N^{\frac{1}{4}} \log(N)).$$

Finally, when  $x \in [1-N^{-1}, 1]$  and  $y \in I_2$ , we have  $1-y = x-y+1-x \leq x-y+N^{-1}$  (since  $x \geq y$ ). By concavity of the function  $t \mapsto |t|^{\frac{1}{4}}$  at  $t = x-y > 0$ , we have

$$(x-y+N^{-1})^{\frac{1}{4}} \leq (x-y)^{\frac{1}{4}} + \frac{1}{4}N^{-1}(x-y)^{-\frac{3}{4}}.$$

Substituting this bound into equation (15), we get,

$$\begin{aligned} K_N(x, y) &= \mathcal{O}\left(\frac{N^{\frac{1}{2}}(1-y)^{\frac{1}{4}}}{|x-y|}\right) \\ &= \mathcal{O}\left(N^{\frac{1}{2}}|x-y|^{-\frac{3}{4}}\right) + \mathcal{O}\left(N^{-\frac{1}{2}}|x-y|^{-\frac{7}{4}}\right) \end{aligned}$$

The integral is bounded in the predictable manner as follows:

$$\begin{aligned} \int_{I_2} |K_N(x, y)| dy &= \mathcal{O}(N^{\frac{1}{2}}) \int_{I_2} |x-y|^{-\frac{3}{4}} dy + \mathcal{O}(N^{-\frac{1}{2}}) \int_{I_2} |x-y|^{-\frac{7}{4}} dy \\ &\leq \mathcal{O}(N^{\frac{1}{2}}) + \mathcal{O}(N^{-\frac{1}{2}}) \int_{N^{-1}}^1 s^{-\frac{7}{4}} ds \\ &= \mathcal{O}(N^{\frac{1}{2}}) + \mathcal{O}(N^{-\frac{1}{2}} \cdot N^{\frac{3}{4}}) \\ &= \mathcal{O}(N^{\frac{1}{2}}). \end{aligned}$$

Since this covers all  $x \in [-1, 1]$  with finitely many uniform  $\mathcal{O}(N^{\frac{1}{2}})$  upper bounds, we have the final result uniformly for all  $x \in [-1, 1]$ .  $\square$

## 5 Best uniform approximation by Fourier extensions

We will compare best uniform approximation in three spaces. For odd positive integer  $N = 2n + 1$ , define,

$$\begin{aligned}\mathcal{H}_N &= \text{span} \left\{ e^{\frac{i\pi}{T} kx} \right\}_{k=-n}^n \subset C([-1, 1]), \\ \mathcal{T}_N &= \text{span} \left\{ e^{\frac{i\pi}{T} kx} \right\}_{k=-n}^n \subset C_{\text{per}}([-T, T]), \\ \mathcal{P}_N &= \text{span} \{ x^k \}_{k=0}^{2n} \subset C([-1, 1]).\end{aligned}$$

We will see that best uniform approximation by Fourier extensions is more similar to that of algebraic polynomials. The best approximation error functionals for these spaces are defined by

$$\begin{aligned}E(f; \mathcal{H}_N) &= \inf_{r_N \in \mathcal{H}_N} \|f - r_N\|_{L^\infty(-1, 1)} \quad \text{for all } f \in C([-1, 1]), \\ E(g; \mathcal{T}_N) &= \inf_{t_N \in \mathcal{T}_N} \|g - t_N\|_{L^\infty(-T, T)} \quad \text{for all } g \in C_{\text{per}}([-T, T]), \\ E(h; \mathcal{P}_N) &= \inf_{p_N \in \mathcal{P}_N} \|h - p_N\|_{L^\infty(-1, 1)} \quad \text{for all } h \in C([-1, 1]).\end{aligned}$$

We wish to find bounds in terms of  $N$  and the regularity of the functions to be approximated.

For  $f \in C([-1, 1])$  the *modulus of continuity* is defined by [11, 29],

$$\omega(f; \delta) = \sup_{\substack{x, y \in [-1, 1] \\ |x - y| < \delta}} |f(x) - f(y)|. \quad (16)$$

For  $g \in C_{\text{per}}([-T, T])$  we define the periodic modulus of continuity to be,

$$\omega_{\text{per}}(g; \delta) = \sup_{\substack{x, y \in [-T, T] \\ d_T(x, y) < \delta}} |g(x) - g(y)|,$$

where  $d_T(x, y)$  is the distance between  $x, y$  as elements of the periodic interval  $[-T, T]$ . The following results are immediate.

**Lemma 5.1.** *If  $f$  is in the Holder space  $C^\alpha([-1, 1])$  for  $\alpha \in [0, 1]$ , then  $\omega(f; \delta) \leq \delta^\alpha |f|_{C^\alpha([-1, 1])}$  for all  $\delta > 0$ .*

**Lemma 5.2.** *If  $f \in C([-1, 1])$  is uniformly Dini-Lipschitz [42], i.e.*

$$\lim_{\delta \searrow 0} \sup_{\substack{x, y \in [-1, 1] \\ |x - y| < \delta}} |(f(x) - f(y)) \log \delta| = 0,$$

then  $\omega(f; \delta) = o(1/|\log \delta|)$ .

### 5.1 A Jackson-type theorem

The original Jackson Theorem for classical Fourier series asserts that for all  $k = 0, 1, 2, \dots$  and all functions  $g \in C_{\text{per}}^k([-T, T])$ , we have

$$E(g; \mathcal{T}_N) = \mathcal{O}(N^{-k}) \omega_{\text{per}} \left( g^{(k)}; \frac{1}{N} \right),$$

where the constant in the big  $\mathcal{O}$  depends on  $k$  and  $T$  [20, Thm. 1.IV].

There is also a polynomial version of Jackson's Theorem, which states that for all  $k = 0, 1, 2, \dots$  and all functions  $h \in C^k([-1, 1])$ , we have

$$E(h; \mathcal{P}_N) = \mathcal{O}(N^{-k}) \omega \left( h^{(k)}; \frac{1}{N} \right), \quad (17)$$

where the constant in the big  $\mathcal{O}$  depends only on  $k$  [20, Thm. 1.VIII]. We prove a version of Jackson's Theorem for Fourier extensions.

**Theorem 5.3** (Jackson-type). *For all  $k = 0, 1, 2, \dots$  and all functions  $f \in C^k([-1, 1])$ ,*

$$E(f; \mathcal{H}_N) = \mathcal{O}(N^{-k}) \omega\left(f^{(k)}; \frac{1}{N}\right),$$

where the constant in the big  $\mathcal{O}$  depends only on  $k$  and  $T$ .

**Lemma 5.4** (Periodic extension). *Let  $f \in C^k([-1, 1])$ . Then  $f$  can be extended to a function  $g \in C_{\text{per}}^k([-T, T])$  such that*

$$\omega_{\text{per}}(g^{(k)}; \delta) \leq \frac{T}{T-1} \omega(f^{(k)}; \delta).$$

*Proof.* First let  $k = 0$ . Define the function  $g \in C_{\text{per}}([-T, T])$  such that for  $x \in [-1, 1]$ ,  $g(x) = f(x)$  and for  $x \in [-T, T] \setminus [-1, 1]$ ,  $g(x)$  is the linear function which interpolates  $f$  at  $\{-1, 1\}$ . We distinguish between 4 different cases for points  $x, y \in [-T, T]$  such that  $d_T(x, y) \leq \delta$ : (i) if  $x, y \in [-1, 1]$ , then

$$|g(x) - g(y)| = |f(x) - f(y)| \leq \omega(f; \delta);$$

(ii) if  $x, y \in [-T, T] \setminus [-1, 1]$ , then since  $g$  is linear in this region,

$$|g(x) - g(y)| \leq \frac{|f(1) - f(-1)|}{2(T-1)} \delta;$$

(iii) if  $x \in [-1, 1], y \in [-T, T] \setminus [-1, 1]$  then

$$|g(x) - g(y)| \leq |f(\xi) - f(x)| + |g(y) - g(\xi)| \leq \omega(f; \delta) + \frac{|f(1) - f(-1)|}{2(T-1)} \delta,$$

where  $\xi$  is the closest of the endpoints to  $x$ ; and (iv) if  $x \in [-T, T] \setminus [-1, 1], y \in [-T, T]$  the bound is similar to the previous one. Now it remains to bound  $|f(1) - f(-1)|$  in terms of  $\omega(f; \delta)$ . For any positive integer  $m$ , we can use a telescoping sum,

$$|f(1) - f(-1)| \leq \sum_{k=0}^{2m-1} \left| f\left(-1 + \frac{k}{m}\right) - f\left(-1 + \frac{k+1}{m}\right) \right| \leq 2m\omega\left(f; \frac{1}{m}\right)$$

It suffices to take  $m > 1/\delta$  to show that  $\frac{|f(1) - f(-1)|}{2(T-1)} \delta \leq \frac{1}{T-1} \omega(f; \delta)$ . Combining all four cases demonstrates  $\omega_{\text{per}}(g, \delta) \leq \frac{T}{T-1} \omega(f, \delta)$ .

Now let  $k > 0$  and choose as extension of  $f$  the  $2(k+1)$ th degree Hermite interpolant in the points  $x = 1$ , and  $x = -1$ ; then  $g^{(k)}(x)$  is the linear interpolation between  $f^{(k)}(1)$  and  $f^{(k)}(-1)$  for  $x \in [-T, T] \setminus [-1, 1]$ . By the case  $k = 0$  proved above,  $\omega_{\text{per}}(g^{(k)}; \delta) \leq \frac{T}{T-1} \omega(f^{(k)}; \delta)$ .  $\square$

*Proof of Theorem 5.3.* Let  $f \in C^k([-1, 1])$ . By Lemma 5.4, this function can be extended to a function  $g \in C_{\text{per}}^k([-T, T])$  such that  $\omega_{\text{per}}(g^{(k)}; \delta)$  is bounded by  $\frac{T}{T-1} \omega(f^{(k)}; \delta)$ . Let  $t_N \in \mathcal{T}_N$  be the best uniform approximation to  $g$ , then (trivially) there exists a function  $r_N \in \mathcal{H}_N$  such that  $r_N(x) = t_N(x)$  for all  $x \in [-1, 1]$ . Hence,

$$E(f; \mathcal{H}_N) \leq \|f - r_N\|_{L^\infty(-1,1)} \leq \|g - t_N\|_{L^\infty(-T,T)} = E(g; \mathcal{T}_N).$$

The original Jackson Theorem can now be used to bound  $E(g; \mathcal{T}_N)$ :

$$E(g; \mathcal{T}_N) = \mathcal{O}(N^{-k}) \omega_{\text{per}}(g^{(k)}; \delta) \leq \mathcal{O}(N^{-k}) \omega(f^{(k)}; \delta).$$

This proves the result.  $\square$

The combination of Lemma 5.1 and Theorem 5.3 yields the following useful fact. If  $f \in C^{k,\alpha}([-1, 1])$  for  $k \geq 0$  and  $\alpha \in [0, 1]$ , then

$$E(f; \mathcal{H}_N) = \mathcal{O}(N^{-k-\alpha}) |f^{(k)}|_{C^\alpha([-1,1])}.$$

In the sequel we will see that this is not actually tight, in the sense that functions in  $C^{k,\alpha}([-1, 1])$  can see a decay of best approximation error with a rate faster than  $N^{-k-\alpha}$ . This is in contrast to the situation for classical Fourier series in which it is indeed tight (see Theorem 5.5).

## 5.2 A Bernstein-type theorem

While Jackson-type theorems bound the best approximation error functional by powers of  $N$  and moduli of continuity of derivatives, Bernstein-type theorems attempt to do the opposite.

Bernstein-type theorems follow from Bernstein inequalities. For classical Fourier series, the Bernstein inequality is,

$$\|t'_N\|_\infty \leq \frac{\pi}{T} n \|t_N\|_\infty, \quad (18)$$

for all  $t_N \in \mathcal{T}_N$ , where  $N = 2n + 1$  [11, Ch. 4, Th. 2.4]. Equality holds when  $t_N(x) \propto e^{\pm \frac{i\pi}{T} nx}$ . From Bernstein's inequality it is possible to show that there exists  $C_T > 0$  such that [11, Ch. 7, Thm. 3.1],

$$\omega_{\text{per}}\left(g; \frac{1}{N}\right) \leq \frac{C_T}{n} \sum_{\substack{k=3 \\ k \text{ odd}}}^N E(g; \mathcal{T}_k).$$

Now, this is not precisely a converse to Jackson's Theorem, but it implies the following tightness property.

**Theorem 5.5** (Jackson–Bernstein [11, Ch. 7, Thm. 3.3]). *Let  $g \in C_{\text{per}}([-T, T])$  and  $\alpha \in (0, 1)$ . It holds that*

$$E(g; \mathcal{T}_N) = \mathcal{O}(N^{-\alpha}) \iff \omega_{\text{per}}(g; \delta) = \mathcal{O}(\delta^\alpha).$$

The direct analogue of Theorem 5.5 for best uniform approximation by algebraic polynomials in  $C([-1, 1])$  is not true. Indeed, consider the function  $h(x) = (1 - x^2)^\alpha$ , whose modulus of continuity satisfies  $\omega(h; \delta) = \mathcal{O}(\delta^\alpha)$  by Lemma 5.1. Define the function  $g(\theta) = h(\cos(\theta)) = |\sin(\theta)|^{2\alpha}$  for  $\theta \in [-\pi, \pi]$ . If  $\alpha < \frac{1}{2}$  then  $g \in C^{2\alpha}([-\pi, \pi])$ , so  $E(g; \mathcal{T}_N) = \mathcal{O}(N^{-2\alpha})$  by Theorem 5.5. Furthermore, the best approximations will be even since  $g$  is even, so the approximants are in fact polynomials in  $\cos(\theta)$ . This implies that the best approximations to  $h$  are polynomials in  $x$ , showing that  $E(h; \mathcal{P}_N) = \mathcal{O}(N^{-2\alpha})$ , twice as good as would be expected from Jackson's Theorem for algebraic polynomials (equation (17)).

It was only in the late twentieth century that characterizations of functions  $h \in C([-1, 1])$  for which  $E(h; \mathcal{P}_N) = \mathcal{O}(N^{-\alpha})$  were developed [11, Ch. 8]. The key insight is to use *weighted* moduli of continuity. The weighted modulus of continuity with weight  $\phi : [-1, 1] \rightarrow [0, \infty)$  for a function  $f \in C([-1, 1])$  is defined as

$$\omega_\phi(f; \delta) = \sup_{\substack{x \pm h \in [-1, 1] \\ 0 \leq h < \phi(x)\delta}} |f(x+h) - f(x-h)|.$$

Taking the weight  $\phi(x) = \frac{1}{2}$  returns the standard modulus of continuity in equation (16).

It turns out that if this weighted modulus of continuity is used with  $\phi(x) = \sqrt{1 - x^2}$ , then there is a direct analogue of Theorem 5.5 for best uniform approximation by algebraic polynomials.

**Theorem 5.6** (Ditzian–Totik [12, Cor. 7.2.5]). *Let  $h \in C([-1, 1])$  and  $\alpha \in (0, 1)$ . It holds that*

$$E(h; \mathcal{P}_N) = \mathcal{O}(N^{-\alpha}) \iff \omega_\phi(h; \delta) = \mathcal{O}(\delta^\alpha),$$

where  $\phi(x) = \sqrt{1 - x^2}$ .

The proof of Theorem 5.6 depends on the Bernstein inequality for algebraic polynomials, which states that for all  $p_N \in \mathcal{P}_N$ ,

$$\|\phi \cdot p'_N\|_{L^\infty(-1, 1)} \leq N \|p_N\|_{L^\infty(-1, 1)}, \quad (19)$$

where  $\phi(x) = \sqrt{1 - x^2}$  [11, Ch. 4, Cor. 1.2]. Compare this with the Bernstein inequality for classical Fourier series (equation (18)). If we wish to remove the factor of  $\phi$  in the left-hand-side of equation (19) then we must change  $N$  to  $N^2$  on the right-hand-side; this is then Markov's inequality [11, Ch. 4, Thm. 1.4].



A Bernstein inequality was proved for Fourier extensions by Videnskii [38], see also [6, p. 242] and [30, Sec. 2]. It states that for all  $r_N \in \mathcal{H}_N$ ,

$$\|\phi \cdot r'_N\|_{L^\infty(-1,1)} \leq \frac{\pi}{T} n \|r_N\|_{L^\infty(-1,1)}, \quad (20)$$

where the weight function is

$$\phi(x) = \frac{\sqrt{\sin\left(\frac{\pi}{2T}(1-x)\right) \sin\left(\frac{\pi}{2T}(1+x)\right)}}{\cos\left(\frac{\pi}{2T}x\right)}. \quad (21)$$

Note that since the sine function is concave in  $[0, \pi]$ , we have  $|\sin\left(\frac{\pi}{2T}(1 \pm x)\right)| \geq \sin\left(\frac{\pi}{2T}\right) |1 \pm x|$  for  $x \in [-1, 1]$ . Also,  $|\sin\left(\frac{\pi}{2T}(1 \pm x)\right)| \leq \frac{\pi}{2T} |1 \pm x|$  and  $\cos\left(\frac{\pi}{2T}x\right) \in [\cos\left(\frac{\pi}{2T}\right), 1]$  for  $x \in [-1, 1]$ . Therefore,

$$\sin\left(\frac{\pi}{2T}\right) \sqrt{1-x^2} \leq \phi(x) \leq \frac{\pi}{2T \cos\left(\frac{\pi}{2T}\right)} \sqrt{1-x^2},$$

and we can change equation (20) to

$$\|\phi \cdot r'_N\|_{L^\infty(-1,1)} \leq \frac{\pi}{T \sin\left(\frac{\pi}{2T}\right)} n \|r_N\|_{L^\infty(-1,1)}, \quad (22)$$

where  $\phi(x) = \sqrt{1-x^2}$ . Using the Bernstein inequality in equation (22) we can prove a Bernstein-type theorem for Fourier extensions.

**Theorem 5.7** (Bernstein-type). *There exists a constant  $C_T > 0$  such that for all  $f \in C([-1, 1])$ , the following holds:*

$$\omega_\phi\left(f; \frac{1}{N}\right) \leq \frac{C_T}{n} \sum_{\substack{k=3 \\ k \text{ odd}}}^N E(f; \mathcal{H}_k),$$

where  $\phi(x) = \sqrt{1-x^2}$  and  $N = 2n + 1$ .

*Proof.* This follows directly from [11, Ch. 6, Thm. 6.2] and [11, Ch. 7, Thm. 5.1(b)], with  $r = 1$ ,  $\mu = 1$ ,  $X = L^\infty(-1, 1)$ ,  $\Phi_n = \mathcal{H}_N$ , and  $Y = W_\infty^1(\phi) := \{f \in W^{1,1}(-1, 1) : \phi \cdot f' \in L^\infty(-1, 1)\}$ , where  $W^{1,1}(-1, 1)$  is the Sobolev space of absolutely continuous functions on  $(-1, 1)$ .  $\square$

From this Bernstein-type theorem for Fourier extensions, we get one half of an equivalence theorem between best approximation errors and weighted moduli of continuity. For the full equivalence, one must prove a Conjecture 5.9 below.

**Theorem 5.8.** *Let  $f \in C([-1, 1])$  and  $\alpha \in (0, 1)$ . It holds that*

$$E(f; \mathcal{H}_N) = \mathcal{O}(N^{-\alpha}) \implies \omega_\phi(f; \delta) = \mathcal{O}(\delta^\alpha).$$

*If Conjecture 5.9 is true, then the reverse implication holds too.*

*Proof.* The forward implication follows immediately from Theorem 5.7. Suppose now that Conjecture 5.9 is true. Then we would have

$$E(f; \mathcal{H}_N) \leq \frac{C_T}{n} \|\phi \cdot f'\|_{L^\infty(-1,1)} \quad (23)$$

for all  $f \in C^1([-1, 1])$  by setting  $f(x) = F\left(e^{\frac{i\pi}{T}x}\right)$ , because  $f \in C^1([-1, 1])$  if and only if  $F \in C^1(A)$ ,  $x \mapsto q_n(e^{\frac{i\pi}{T}x}) \in \mathcal{H}$ , and  $|\mu(e^{\frac{i\pi}{T}x})| \leq \frac{\pi}{T} \phi(x)$ . We wish to extend this to all  $f \in W^{1,1}(-1, 1)$  such that  $\phi \cdot f' \in L^\infty(-1, 1)$  by a density argument, where  $W^{1,p}(-1, 1)$  is the Sobolev space of  $L^p(-1, 1)$  functions whose weak derivatives lie in  $L^p(-1, 1)$ . For such a function  $f$ , one can verify that the functions  $f_\rho(x) = f(\rho x)$  for  $\rho \in (0, 1)$  satisfy:  $f_\rho \in W^{1,\infty}(-1, 1)$ ,  $f_\rho \rightarrow f$  in

$L^\infty$ , and  $\|\phi \cdot f'_\rho\|_\infty \leq \|\phi \cdot f'\|_\infty$ . For each  $\rho$  and  $\varepsilon > 0$  there exists  $f_{\rho,\varepsilon} \in C^1([-1,1])$  such that  $\|f_{\rho,\varepsilon} - f_\rho\|_{W^{1,\infty}} < \varepsilon$  by density of  $C^1([-1,1])$  in  $W^{1,\infty}(-1,1)$ . Therefore there exists  $f_\varepsilon \in C^1([-1,1])$  such that  $\|f - f_\varepsilon\|_{L^\infty(-1,1)} < \varepsilon$  and  $\|\phi \cdot f'_\varepsilon\|_\infty \leq \|\phi \cdot f'\|_\infty + \varepsilon$ . Hence  $E(f; \mathcal{H}_N) \leq \|f - f_\varepsilon\|_{L^\infty(-1,1)} + E(f_\varepsilon; \mathcal{H}_N) \leq (1 + \frac{C_T}{n})\varepsilon + \frac{C_T}{n}\|\phi \cdot f'\|_\infty$ . Since  $\varepsilon$  is arbitrary, we have the desired inequality. A similar argument may be found in [11, pp. 280].

From the above it would follow that there exists a constant  $C_T > 0$  such that

$$E(f; \mathcal{H}_N) \leq C_T \omega_\phi \left( f; \frac{1}{N} \right), \quad (24)$$

from equation (23), [11, Ch. 6, Thm. 6.2] and [11, Ch. 7, Thm. 5.1(a)], with  $r = 1$ ,  $\mu = 1$ ,  $X = L^\infty(-1,1)$ ,  $\Phi_n = \mathcal{H}_N$ , and  $Y = W_\infty^1(\phi) := \{f \in W^{1,1}(-1,1) : \phi \cdot f' \in L^\infty(-1,1)\}$ . Equation (24) would imply that if  $\omega_\phi(f; \delta) = \mathcal{O}(\delta^\alpha)$  then  $E(f; \mathcal{H}_N) = \mathcal{O}(N^{-\alpha})$ , as required.  $\square$

**Conjecture 5.9** (Jackson inequality for polynomials on a circular arc). *For any  $T > 1$ , define the arc on the complex unit circle,*

$$A = \left\{ e^{i\theta} : \theta \in \left( -\frac{\pi}{T}, \frac{\pi}{T} \right) \right\}.$$

*There exists a constant  $C_T > 0$  such that for all  $F \in C^1(A)$  and all  $n \in \mathbb{N}$ , there exists a polynomial  $q_n$  of degree  $n$  such that*

$$\sup_{z \in A} |F(z) - q_n(z)| \leq \frac{C_T}{n} \sup_{z \in A} |\mu(z) F'(z)|,$$

where  $\mu(z) = \sqrt{(z - e^{\frac{i\pi}{T}})(z - e^{-\frac{i\pi}{T}})}$ .

Notice that to approximate  $f$  we conjecture that we only need to use positive powers of  $z$ , which means we do not need to utilize all of the functions in  $\mathcal{H}_N$ . This is because by Mergelyan's Theorem [33, Thm. 20.5] polynomials are dense in the space  $C(A)$ . It is not surprising because of the redundant nature of approximation by Fourier extensions.

## 6 A localization theorem for Fourier extensions

The theorem proved in this section is a modification of a theorem of Freud ([16, Thm. IV.5.4]), which is a localization theorem for orthogonal polynomials on an interval. We, however, are working with the orthonormal basis given in Lemma 4.2, and there are some clear differences between the two situations. We show that these differences do not change the statement of the result.

**Theorem 6.1** (Localization theorem). *Let  $f \in L^2(-1,1)$  be such that  $f(x) = 0$  for all  $x \in [a,b] \subseteq [-1,1]$ . Then  $P_N(f) \rightarrow 0$  uniformly in all subintervals  $[c,d] \subset (a,b)$ .*

*Proof.* First note that the pointwise error can be written in terms of the prolate kernel discussed in Section 4 as,

$$P_N(f)(x) - f(x) = \int_{-1}^1 (f(y) - f(x)) K_N(x, y) dy.$$

Let  $x \in [c,d] \subset (a,b)$ , so that  $f(x) = 0$ . By the formula for the prolate kernel (Lemma 4.3),

$$P_N(f)(x) - f(x) = \int_{-1}^1 \frac{f(y)}{\sin\left(\frac{\pi}{2T}(x-y)\right)} \text{Imag} \left( \overline{e^{-\frac{i\pi}{2T} \frac{N}{2} y} \cdot \Pi_N\left(e^{\frac{i\pi}{2T} y}\right)} \cdot e^{-\frac{i\pi}{2T} \frac{N}{2} x} \cdot \Pi_N\left(e^{\frac{i\pi}{2T} x}\right) \right) dy.$$

By expressing the imaginary part as  $\frac{1}{2i}$  times the difference of the complex conjugates, it is easy to see that for this expression to tend to zero as  $N \rightarrow \infty$ , it is sufficient that for any  $f$  as in the statement of the theorem, we have

$$\lim_{N \rightarrow \infty} \int_{-1}^1 \frac{f(y)}{\sin\left(\frac{\pi}{2T}(x-y)\right)} e^{\frac{i\pi}{2T} \frac{N}{2} y} \cdot \overline{\Pi_N\left(e^{\frac{i\pi}{2T} y}\right)} dy = 0.$$

To prove this we consider the functions

$$g_\xi(y) = \frac{f(y)e^{\frac{i\pi}{2T}y}}{\sin\left(\frac{\pi}{2T}(\xi - y)\right)},$$

for  $\xi \in [c, d]$ . It holds that  $g_\xi \in L^2(-1, 1)$ , because  $g_\xi$  is equal to 0 inside  $[a, b]$  and equal to  $f$  (an  $L^2(-1, 1)$  function) multiplied by a bounded function ( $y \mapsto e^{\frac{i\pi}{2T}y}/\sin\left(\frac{\pi}{2T}(\xi - y)\right)$ ) outside of  $[a, b]$ .

Let  $\varepsilon > 0$ . By Proposition 2.4, for any  $\xi \in [c, d]$ , there exists  $K_\xi \in \mathbb{N}$  and a function  $h_{K_\xi} \in \mathcal{H}_{K_\xi}$  such that  $\|g_\xi - h_{K_\xi}\|_{L^2(-1,1)} < \varepsilon$ . A key property of the function  $h_{K_\xi}$  is that for  $N \geq K_\xi$ ,

$$\int_{-1}^1 h_{K_\xi}(y) e^{\frac{i\pi}{T} \frac{N-1}{2} y} \cdot \overline{\Pi_N\left(e^{\frac{i\pi}{T}y}\right)} dy = 0, \quad (25)$$

because  $h_{K_\xi}(y) e^{\frac{i\pi}{T} \frac{N-1}{2} y}$  is a polynomial of degree  $\frac{K_\xi-1}{2} + \frac{N-1}{2} \leq N-1$  in the variable  $z = \exp\left(\frac{i\pi}{T}y\right)$ . Now, because the map  $x \mapsto g_x$  is a continuous mapping from  $[c, d] \rightarrow L^2(-1, 1)$ , there exists an interval  $I(\xi)$  such that for all  $x \in I(\xi)$ ,  $\|g_x - h_{K_\xi}\|_{L^2(-1,1)} < \varepsilon$  is still valid. In consequence of the Heine-Borel Compactness Theorem [33], the interval  $[c, d]$  will be covered by finitely many of these intervals  $I(\xi)$ , which we denote,  $I(\xi_1), I(\xi_2), \dots, I(\xi_s)$ .

Let  $K_\varepsilon$  be an odd integer such that  $h_{K_{\xi_i}} \in \mathcal{H}_{K_\varepsilon}$  for  $i = 1, \dots, s$ . For an arbitrary  $x \in [c, d]$  there is an interval  $I(\xi_r)$  such that  $x \in I(\xi_r)$  and for  $N > K_\varepsilon$ , we have (using the expression (25)),

$$\begin{aligned} \left| \int_{-1}^1 \frac{f(y)}{\sin\left(\frac{\pi}{2T}(x - y)\right)} e^{\frac{i\pi}{T} \frac{N}{2} y} \cdot \overline{\Pi_N\left(e^{\frac{i\pi}{T}y}\right)} dy \right| &= \left| \int_{-1}^1 (g_x(y) - h_{K_{\xi_r}}(y)) e^{\frac{i\pi}{T} \frac{N-1}{2} y} \overline{\Pi_N\left(e^{\frac{i\pi}{T}y}\right)} dy \right| \\ &\leq \left( \int_{-1}^1 |g_x(y) - h_{K_{\xi_r}}(y)|^2 dy \right)^{\frac{1}{2}} \\ &\quad \cdot \left( \int_{-1}^1 \left| e^{\frac{i\pi}{T} \frac{N-1}{2} y} \overline{\Pi_N\left(e^{\frac{i\pi}{T}y}\right)} \right|^2 dy \right)^{\frac{1}{2}} \\ &< \varepsilon. \end{aligned}$$

This last line used the normality of the basis for  $\mathcal{H}_N$  discussed in Lemma 4.2.

In conclusion, since  $\varepsilon$  is arbitrary and the inequality above is valid for all  $N > K_\varepsilon$ , the integral must converge to zero as  $N \rightarrow \infty$ , uniformly with respect to  $x \in [c, d]$ , as required.  $\square$

## 7 Numerical experiments

In this section we provide numerically computed examples of pointwise and uniform convergence of Fourier extensions for functions with various regularity properties. It was discussed in the introduction that the condition number of the linear system for computing the Fourier extension is extremely ill-conditioned, making computation of the exact solution to the Fourier extension practically impossible. To deal with this issue, we used sufficiently high precision floating point arithmetic and we did not take  $N$  higher than 129, to ensure that the system could be inverted accurately. The right-hand-side vectors for the computations are computed by quadrature in high precision floating point arithmetic.

In practice, one would compute a fast regularized oversampled interpolation Fourier extension using the algorithm in [27], requiring only  $\mathcal{O}(N \log^2(N))$  floating point operations. However, we are interested in the exact Fourier extension and want to avoid any artefacts that may be caused by the regularization or discretization of the domain.

In some cases, we compare the convergence rate of Fourier extensions to that of Legendre series, because we predict that the qualitative behaviour of Legendre series will be similar (see Section 8). For the Legendre series approximations we computed the Legendre series coefficients one by one using adaptive quadrature in 64 bit floating point precision.

## 7.1 Analytic and entire functions

Theorem 3.1 gives an upper bound on the rate of exponential convergence of Fourier extension approximations to analytic functions. The regions of analyticity in the complex plane which dictate the rate are the mapped Bernstein ellipse  $\mathcal{D}(\rho)$ , where  $\rho > 1$ . The theorem is illustrated in Figure 1, where we approximate an entire function and four analytic functions which each have a pole in a different location in the complex plane. All examples exhibit exponential convergence in the uniform norm at a rate which is predicted by Theorem 3.1. This is also the case for the entire function, where the exponential convergence rate is limited by a  $T$ -dependent upper bound.

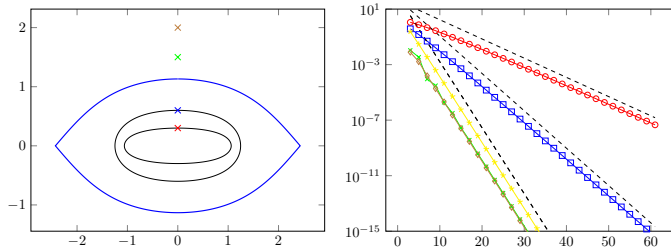


Figure 1: We compute Fourier extension approximations to 5 functions:  $f(x) = e^x$  (yellow stars) and  $f(x) = \frac{1}{x-r}$  for  $r = 0.3i, 0.6i, 1.5i, 2.0i$  (red circles, blue squares, green crosses and brown diamonds, respectively). The  $T$  parameter is 2.43. Left: the mapped Bernstein ellipses  $\mathcal{D}(\rho)$  in the complex plane, for  $\rho = 1.891, 3.454, 8.913$ . The outermost outline (in blue) encloses the maximal mapped Bernstein ellipse; analyticity outside this largest region does not increase the exponential convergence rate. Right: The  $L^\infty(-1, 1)$  error against values of  $N$  for each of the 5 functions. The black dashed lines indicate the convergence rates predicted by Theorem 3.1. Color figures online.

## 7.2 Differentiable functions

We investigate Fourier extension approximation of splines of degree  $d = 3, 9$  and 15 on the interval  $[0, \frac{1}{2}]$ , which lie in the Hölder spaces  $C^{2,1}([0, \frac{1}{2}])$ ,  $C^{8,1}([0, \frac{1}{2}])$  and  $C^{14,1}([0, \frac{1}{2}])$  respectively. By Theorem 3.2, we expect the pointwise errors to be  $\mathcal{O}(N^{-d} \log N)$  in the interior, and  $\mathcal{O}(N^{\frac{1}{2}-d})$  uniformly over the whole interval.

The spline functions and the pointwise approximation errors for Fourier extensions with various values of  $N$  are plotted in Figure 2. The rates of convergence predicted by Theorem 3.2 fit reasonably well, sometimes performing slightly better. For comparison, we include the errors for a Legendre series approximation in a dashed line of the same color. See subsection 8.1 for a full discussion comparing convergence of Legendre series and Fourier extensions.

## 7.3 Non-differentiable functions

We investigate the approximation of functions with algebraic singularities, discontinuities, and Dini-Lipschitz continuity.

Functions with an algebraic singularity at the endpoint are studied in Figure 3. We plot the pointwise errors for Fourier extension and Legendre series approximations to  $f(x) = x^\alpha$  for  $\alpha = \frac{3}{4}, \frac{1}{2}$ , and  $\frac{1}{10}$ . These functions lie in the Hölder spaces  $C^\alpha([0, \frac{1}{2}])$  for their respective values of  $\alpha$ .

While Theorem 3.2 guarantees uniform convergence over  $[-1, 1]$  only for the first function (since for the other two functions  $\alpha \leq \frac{1}{2}$ ), in our experiments we find that the error of the other two functions converges to zero too. We believe that this discrepancy is related to the weighted moduli of continuity of these functions being more favourable than the standard moduli (see Section 5). Overall, the observed convergence rates are sometimes better than the predicted rates, but when Fourier extensions are compared with Legendre series, we see similar rates of pointwise

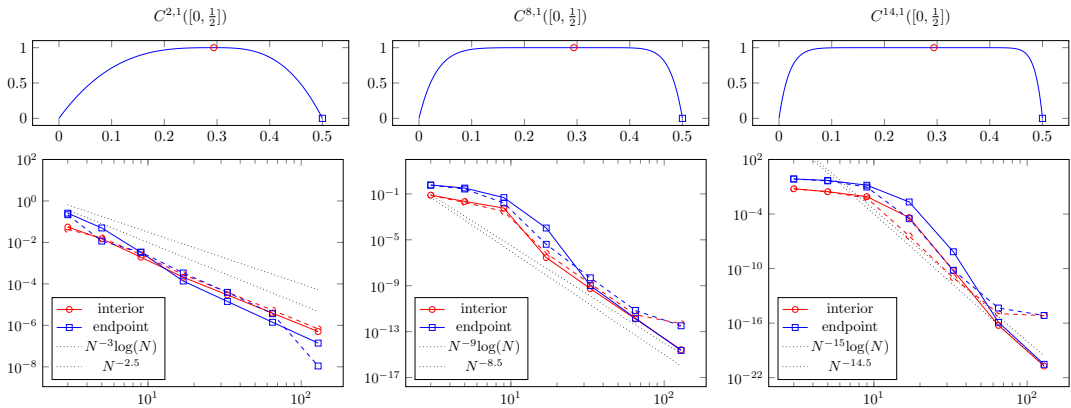


Figure 2: Above : Plots of plines of degree 3, 9, and 15 in  $C^{2,1}$ ,  $C^{8,1}$ , and  $C^{14,1}$ , respectively with an interior point marked using a red circle, and a boundary point marked with a blue square. Below: The pointwise error at an interior point (red circle) and an endpoint (blue square) using Fourier extension with  $T = 2$  (full lines) and using Legendre series (dashed lines) against the number of degrees of freedom,  $N$ . The black dotted lines without markers indicate the upper bounds on the algebraic rates of convergence predicted by Theorem 3.2. Note that while the Fourier extension coefficients were computed in high precision, the Legendre series were not, so some of the data points for Legendre series are affected by numerical error. Color figures online.

convergence, especially at the singular point. See subsection 8.1 for a full discussion comparing convergence of Legendre series and Fourier extensions.

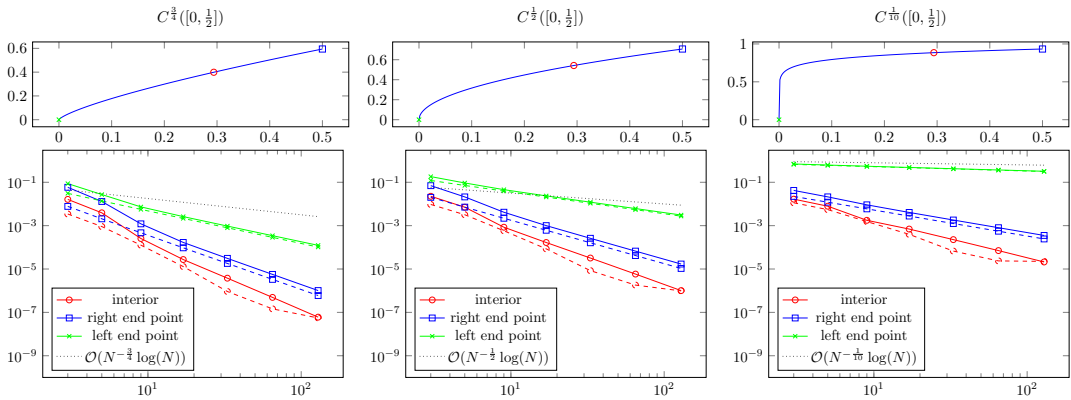


Figure 3: Above left:  $f(x) = x^{3/4}$ . Above middle:  $f(x) = x^{1/2}$ . Above right:  $f(x) = x^{1/10}$ . Below: The pointwise error at an interior point (red circles), singular endpoint (green crosses), and non-singular endpoint (blue square) using Fourier extension with  $T = 2$  (full lines) and Legendre series (dashed lines) against the number of degrees of freedom,  $N$ . The black dotted lines without markers indicate the upper bounds on the algebraic rates of convergence predicted by Theorem 3.2. Color figures online.

Three functions with a singularity at the interior are shown in Figure 4. The first has an algebraic singularity:  $f(x) = |x - r|^{1/4}$  where  $r = 0.29384$  (chosen to avoid any symmetry with respect to the domain). We observe agreement with the expected convergence rate of  $\mathcal{O}(N^{1/4} \log N)$  for the error at interior points. The second function has a jump:

$$f(x) = \begin{cases} x & \text{if } x \in [0, \frac{1}{4}], \\ 1 & \text{if } x \in (\frac{1}{4}, \frac{1}{2}]. \end{cases} \quad (26)$$

Even though the function is highly irregular because of the jump, this does not deny convergence at regular points, corroborating the local nature of Theorem 3.5. The last function is uniformly Dini-Lipschitz continuous in  $[0, \frac{1}{2}]$ :

$$f(x) = \begin{cases} (\log(|x - r|))^{-2} & \text{if } x \in [0, \frac{1}{2}] \setminus \{r\} \\ 0 & \text{if } x = r, \end{cases} \quad (27)$$

where  $r = 0.29384$  (chosen to avoid any symmetry with respect to the domain). In Figure 4, the expected convergence rate of  $\mathcal{O}((\log(N))^{-1})$  of Lemma 3.7 is present.

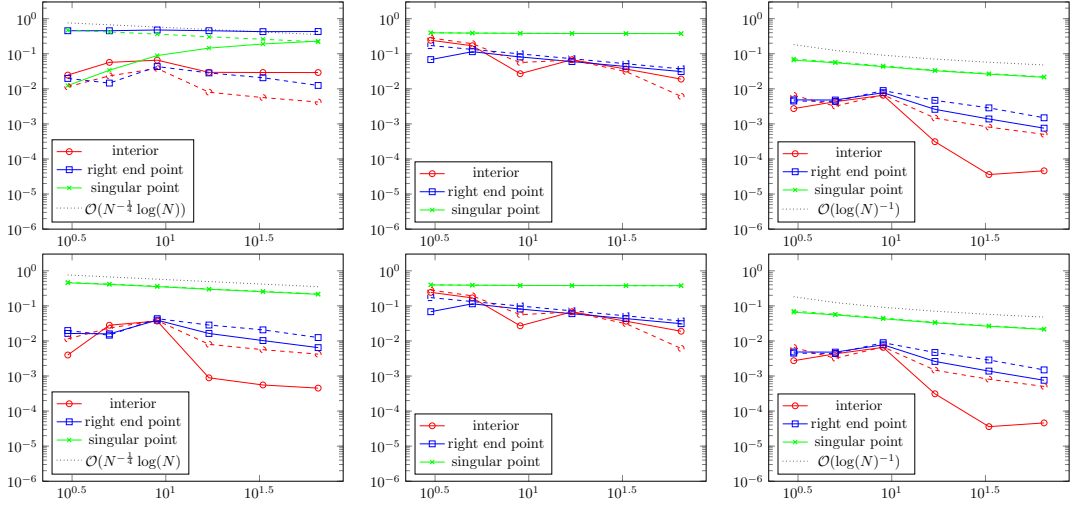


Figure 4: Above left:  $f(x) = |x - r|^{1/4}$  with  $r = 0.29384$ . Above middle: function with a jump, given in equation (26). Above right: Dini-Lipschitz continuous function given in equation (27). It has a strong cusp at  $x = 0.29384$ . Below: The pointwise error at an interior point (red circles), singular interior point (green crosses), and endpoint (blue squares) using Fourier extension with  $T = 2$  (full lines) and Legendre series (dashed lines) against the number of degrees of freedom,  $N$ . The black dotted lines without markers in the bottom left plot indicate the upper bounds on the algebraic rates of convergence predicted by Theorem 3.2. The black dotted line without markers in the bottom right plot indicates the rate of convergence predicted by Lemma 3.7. Color figures online.

In all three cases, we compared the convergence of Fourier extension approximations and Legendre series. While there is sometimes a mismatch between the pessimistic prediction of Theorem 3.2 and Lemma 3.7 for the convergence rates (see Section 5), when we compare Fourier extensions and Legendre series, we observe agreement. See subsection 8.1 for a full discussion comparing convergence of Legendre series and Fourier extensions.

## 8 Discussion

We proved pointwise and uniform convergence results for Fourier extension approximations of functions in Hölder spaces and with local uniform Dini-Lipschitz conditions. This was achieved by proving upper bounds on the associated Lebesgue function and the decay rate of best uniform approximation error for Fourier extensions, then appealing to Lebesgue's Lemma.

## 8.1 Comparison to Legendre series

For a function  $f \in L^2(-1, 1)$ , let us compare the Fourier extension approximant,  $f_N$ , to the Legendre series approximant,

$$f_N^L(x) = \sum_{k=0}^{N-1} a_k p_k^L(x), \quad a_k = \frac{1}{2} \int_{-1}^1 f(x) p_k^L(x) dx,$$

where  $p_k^L$  is the  $k$ th Legendre polynomial normalized so that  $\frac{1}{2} \int_{-1}^1 p_k^L(x)^2 dx = 1$ .

The Lebesgue function of this approximation scheme is  $\mathcal{O}(\log N)$  for  $x \in [a, b] \subset (-1, 1)$  and  $\mathcal{O}(N^{\frac{1}{2}})$  uniformly for  $x \in [-1, 1]$  [20, Ch. 1], [17], which is precisely the same as the Lebesgue function for Fourier extensions (see Theorem 4.1). Best uniform approximation by Fourier extensions was compared to best uniform approximation by algebraic polynomials in Section 5. For any  $f \in C^k([-1, 1])$  for  $k \in \mathbb{Z}_{\geq 0}$ , we have

$$E(f; \mathcal{H}_N) = \mathcal{O}(N^{-k}) \omega\left(f^{(k)}; \frac{1}{N}\right) \text{ and } E(f; \mathcal{P}_N) = \mathcal{O}(N^{-k}) \omega\left(f^{(k)}; \frac{1}{N}\right).$$

It follows that for  $C^{k, \alpha}([-1, 1])$  functions, the statement of Theorem 3.2 also applies to Legendre series approximations. The localized convergence result for Dini–Lipschitz functions, Theorem 3.5 also also applies to Legendre series [16, Thm. IV.5.6]. Some of the experiments in Section 7 demonstrate these similarities.

Theorem 2.1 on exponential convergence differs from the exponential convergence results for Legendre series in two ways. First, the region in the complex plane which determines the rate of exponential convergence is determined by Bernstein ellipses for Legendre series, but by mapped Bernstein ellipses for Fourier extensions. Second, there is an upper limit of  $\cot^2\left(\frac{\pi}{4T}\right)$  for the rate of exponential convergence of Fourier extensions regardless of the region of analyticity, whereas for Legendre series the rate can be arbitrarily fast, and for entire functions the rate of convergence is superexponential [39].

## 8.2 Extensions of this work

It was mentioned in the introduction that our convergence results will be more applicable if we can extend them to regularized and oversampled interpolation versions of Fourier extensions, because those are the kinds of Fourier extensions for which stable and efficient algorithms have been developed.

Regularized Fourier extensions for a given regularization parameter  $\varepsilon > 0$  can be defined as follows. Suppose the matrix  $G \in \mathbb{R}^{N \times N}$ ,

$$G_{k,j} = \text{sinc}\left(\left(k-j\right)\frac{\pi}{T}\right),$$

has eigendecomposition  $G = V S V^*$ . Let  $S_\varepsilon$  be  $S$  but with all entries less than  $\varepsilon$  set to 0. The coefficients  $\mathbf{c}^\varepsilon \in \mathbb{C}^N$  of the regularized Fourier extension of  $f \in L^2(-1, 1)$  are given by

$$\mathbf{c}^\varepsilon = V S_\varepsilon^\dagger V^* \mathbf{b},$$

where  $b_k = \left(\frac{T}{2}\right)^{\frac{1}{2}} \int_{-1}^1 e^{-\frac{i\pi}{T} kx} f(x) dx$  [27]. In other words, the solution is projected onto the eigenvectors of  $G$  whose eigenvalues are greater than or equal to  $\varepsilon$ . These eigenvectors are the Discrete Prolate Spheroidal Sequences (DPSSs), which are the Fourier coefficients of the DPSWFs  $\{\xi_{k,N}\}_{k=1}^N$  discussed in Section 4 [34]. The regularized Fourier extension, therefore, finds the best approximation not in  $\mathcal{H}_N$ , but in the linear space  $\mathcal{H}_{N,\varepsilon} \subset \mathcal{H}_N \subset L^2(-1, 1)$ , where

$$\mathcal{H}_{N,\varepsilon} = \text{span} \{\xi_{k,N} : \lambda_{k,N} \geq \varepsilon\}.$$

Therefore, if the Lebesgue function  $\Lambda(x; P_{N,\varepsilon})$  (where  $P_{N,\varepsilon}$  is the projection operator from  $L^2(-1, 1)$  to  $\mathcal{H}_{N,\varepsilon}$ ) and best approximation error functional  $E(f; \mathcal{H}_{N,\varepsilon})$  can be estimated as in Sections 4

and 5, then we immediately obtain pointwise convergence results for regularized Fourier extensions by Lebesgue’s Lemma. Extensions to the regularized oversampled interpolation version of Fourier extensions can be conducted by considering the analogous quantities for the Periodic Discrete Prolate Spheroidal Sequences (PDPSSs) [41, 27].

Generalization of this work to the multivariate case would be extremely interesting, because the shape of the domain  $\Omega \subset \mathbb{R}^d$  and regularity of its boundary will likely come into play [28].

## Acknowledgements

We benefited from useful discussions with Ben Adcock, Arno Kuijlaars, Walter Van Assche, and Andrew Gibbs. The first author is grateful to FWO Research Foundation Flanders for a postdoctoral fellowship which he enjoyed during the writing of this paper.

## References

- [1] B. Adcock and D. Huybrechs. On the resolution power of Fourier extensions for oscillatory functions. *Journal of Computational and Applied Mathematics*, 260:312–336, 2014.
- [2] B. Adcock and D. Huybrechs. Frames and numerical approximation II: generalized sampling. *submitted*, 2017.
- [3] B. Adcock and D. Huybrechs. Frames and numerical approximation. *Accepted in SIAM Review*, 2018.
- [4] B. Adcock, D. Huybrechs, and J. Martín-Vaquero. On the numerical stability of Fourier extensions. *Foundations of Computational Mathematics*, 14(4):635–687, 2014.
- [5] B. Adcock and J. Ruan. Parameter selection and numerical approximation properties of Fourier extensions from fixed data. *Journal of Computational Physics*, 273:453–471, 2014.
- [6] P. Borwein and T. Erdélyi. *Polynomials and polynomial inequalities*, volume 161. Springer Science & Business Media, 2012.
- [7] J. P. Boyd. A comparison of numerical algorithms for Fourier extension of the first, second, and third kinds. *Journal of Computational Physics*, 178(1):118–160, 2002.
- [8] J. P. Boyd. Fourier embedded domain methods: extending a function defined on an irregular region to a rectangle so that the extension is spatially periodic and  $C^\infty$ . *Applied mathematics and computation*, 161(2):591–597, 2005.
- [9] O. P. Bruno, Y. Han, and M. M. Pohlman. Accurate, high-order representation of complex three-dimensional surfaces via Fourier continuation analysis. *Journal of Computational Physics*, 227(2):1094–1125, 2007.
- [10] P. Deift. *Orthogonal polynomials and random matrices: a Riemann-Hilbert approach*, volume 3. American Mathematical Soc., 1999.
- [11] R. A. DeVore and G. G. Lorentz. *Constructive approximation*, volume 303. Springer Science & Business Media, 1993.
- [12] Z. Ditzian and V. Totik. *Moduli of smoothness*. Springer-Verlag, Berlin, 1987.
- [13] *NIST Digital Library of Mathematical Functions*. <http://dlmf.nist.gov/>, Release 1.0.20 of 2018-09-15. FWJ Olver, AB Olde Daalhuis, DW Lozier, BI Schneider, RF Boisvert, CW Clark, BR Miller and BV Saunders, eds.



- [14] T. A. Driscoll, N. Hale, and L. N. Trefethen. *Chebfun guide*. Pafnuty Publications, Oxford, 2014.
- [15] L. C. Evans. *Partial differential equations*. American Mathematical Society, 2010.
- [16] G. Freud. *Orthogonal polynomials*. Pergamon Press, 1971.
- [17] T. H. Gronwall. Über die Laplacesche Reihe. *Mathematische Annalen*, 74(2):213–270, 1913.
- [18] D. Huybrechs. On the Fourier extension of nonperiodic functions. *SIAM Journal on Numerical Analysis*, 47(6):4326–4355, 2010.
- [19] D. Huybrechs and R. Matthysen. Computing with functions on domains with arbitrary shapes. In *International Conference Approximation Theory*, pages 105–117. Springer, 2016.
- [20] D. Jackson. *The theory of approximation*. The American Society, 1930.
- [21] S. Karnik, Z. Zhu, M. B. Wakin, J. Romberg, and M. A. Davenport. The fast Slepian transform. *Applied and Computational Harmonic Analysis*, 2017.
- [22] I. V. Krasovsky. Gap probability in the spectrum of random matrices and asymptotics of polynomials orthogonal on an arc of the unit circle. *International Mathematics Research Notices*, 2004(25):1249–1272, 2004.
- [23] A. B. J. Kuijlaars, R. Mclaughlin, W. Van Assche, and M. Vanlessen. The RiemannHilbert approach to strong asymptotics for orthogonal polynomials on  $[-1,1]$ . *Advances in Mathematics*, 188:337–398, 2004.
- [24] M. Lyon. A fast algorithm for Fourier continuation. *SIAM Journal on Scientific Computing*, 33(6):3241–3260, 2011.
- [25] M. Lyon. Approximation error in regularized SVD-based Fourier continuations. *Applied Numerical Mathematics*, 62(12):1790–1803, 2012.
- [26] A. P. Magnus. Freud equations for Legendre polynomials on a circular arc and solution of the Grünbaum–Delsarte–Janssen–Vries problem. *Journal of Approximation Theory*, 139(1-2):75–90, 2006.
- [27] R. Matthysen and D. Huybrechs. Fast algorithms for the computation of Fourier extensions of arbitrary length. *SIAM Journal on Scientific Computing*, 38(2):A899–A922, 2016.
- [28] R. Matthysen and D. Huybrechs. Function approximation on arbitrary domains using Fourier extension frames. *SIAM Journal on Numerical Analysis*, 56(3):1360–1385, 2018.
- [29] H. N. Mhaskar and D. V. Pai. *Fundamentals of approximation theory*. CRC Press, 2000.
- [30] B. Nagy and V. Totik. Bernsteins inequality for algebraic polynomials on circular arcs. *Constructive approximation*, 37(2):223–232, 2013.
- [31] S. Olver and collaborators. ApproxFun v0.10.0. <https://github.com/JuliaApproximation/ApproxFun.jl>, 2018.
- [32] G. M. Phillips. *Interpolation and approximation by polynomials*, volume 14. Springer Science & Business Media, 2003.
- [33] W. Rudin. *Real and Complex Analysis, 3rd Edition*. McGraw-Hill International Editions, 1987.
- [34] D. Slepian. Prolate spheroidal wave functions, Fourier analysis, and uncertainty V: The discrete case. *Bell System Technical Journal*, 57(5):1371–1430, 1978.

- [35] G. Szegő. *Orthogonal polynomials*, volume 23. American Mathematical Soc., 1939.
- [36] L. N. Trefethen. *Approximation theory and approximation practice*, volume 128. SIAM, 2013.
- [37] J. Varah. The prolate matrix. *Linear algebra and its applications*, 187:269–278, 1993.
- [38] V. Videnskii. Extremal estimates for the derivative of a trigonometric polynomial on an interval shorter than its period. In *Soviet Math. Dokl*, volume 1, pages 5–8, 1960.
- [39] H. Wang and S. Xiang. On the convergence rates of Legendre approximation. *Mathematics of Computation*, 81(278):861–877, 2012.
- [40] G. B. Wright, M. Javed, H. Montanelli, and L. N. Trefethen. Extension of Chebfun to periodic functions. *SIAM Journal on Scientific Computing*, 37(5):C554–C573, 2015.
- [41] W. Y. Xu and C. Chamzas. On the periodic discrete prolate spheroidal sequences. *SIAM Journal on Applied Mathematics*, 44(6):1210–1217, 1984.
- [42] A. Zygmund. *Trigonometric series*, volume 1. Cambridge University Press, 2002.

## A Asymptotics of Legendre polynomials on a circular arc

Krasovsky derived the asymptotics of polynomials orthogonal on an arc  $\{e^{i\theta} : \alpha \leq \theta \leq 2\pi - \alpha\}$  with respect to a positive analytic weight  $f_\alpha(\theta)$  by Riemann–Hilbert analysis [22, 23, 10]. We are interested in the case  $f_\alpha(\theta) \equiv 1$ , the Legendre polynomials on an arc of the unit circle. The following lemma follows Krasovsky’s instructions on how to calculate an asymptotic expansion of these polynomials in various regions of the complex plane, where we restrict to the special case of the arc itself.

**Lemma A.1.** *For  $\alpha \in (0, \pi)$ , let  $\{\phi_k(z, \alpha)\}_{k=0}^\infty$  be the polynomials in  $z$  with positive leading coefficient, satisfying*

$$\frac{1}{2\pi} \int_\alpha^{2\pi-\alpha} \phi_n(e^{i\theta}, \alpha) \overline{\phi_m(e^{i\theta}, \alpha)} d\theta = \delta_{n,m},$$

for  $n, m = 0, 1, 2, \dots$ . Then there exists  $\delta > 0$  such that for  $\theta \in [\alpha + \delta, 2\pi - \alpha - \delta]$ ,

$$\begin{aligned} \phi_n(e^{i\theta}, \alpha) &= e^{i\frac{n}{2}\theta} \gamma^{-\frac{1}{2}} \left( e^{i\frac{\alpha-\pi}{4}} \left( \frac{\sin(\frac{1}{2}(\theta - \alpha))}{\sin(\frac{1}{2}(\theta + \alpha))} \right)^{\frac{1}{4}} \cos\left(n\tau(\theta) - \frac{\pi}{4}\right) \right. \\ &\quad \left. - e^{-i\frac{\alpha-\pi}{4}} \left( \frac{\sin(\frac{1}{2}(\theta + \alpha))}{\sin(\frac{1}{2}(\theta - \alpha))} \right)^{\frac{1}{4}} \sin\left(n\tau(\theta) - \frac{\pi}{4}\right) \right) + \mathcal{O}(n^{-1}) \end{aligned} \quad (28)$$

and for  $\theta \in [\alpha, \alpha + \delta]$ ,

$$\begin{aligned} \phi_n(e^{i\theta}, \alpha) &= e^{i\frac{n}{2}\theta} \gamma^{-\frac{1}{2}} \left( \frac{\pi}{2} n\tau(\theta) \right)^{\frac{1}{2}} \left( e^{-i\frac{\alpha+\pi}{4}} \left( \frac{\sin(\frac{1}{2}(\theta + \alpha))}{\sin(\frac{1}{2}(\theta - \alpha))} \right)^{\frac{1}{4}} J_0(n\tau(\theta)) \right. \\ &\quad \left. - e^{i\frac{\alpha+\pi}{4}} \left( \frac{\sin(\frac{1}{2}(\theta - \alpha))}{\sin(\frac{1}{2}(\theta + \alpha))} \right)^{\frac{1}{4}} J_1(n\tau(\theta)) \right) + \mathcal{O}(n^{-\frac{1}{2}}) \end{aligned} \quad (29)$$

where

$$\tau(\theta) = \cos^{-1} \left( \frac{\cos(\theta/2)}{\gamma} \right) \quad \text{and} \quad \gamma = \cos\left(\frac{\alpha}{2}\right).$$

The asymptotics for  $\theta \in [2\pi - \alpha - \delta, 2\pi - \alpha]$  can be determined using  $\phi_n(e^{-i\theta}, \alpha) = \overline{\phi_n(e^{i\theta}, \alpha)}$ .

*Proof.* In the notation of [22], the arc on which the polynomials are orthogonal is contained within “region 1” of the complex plane. The asymptotics of  $\phi_n(z, \alpha)$  in region 1 (for  $f_\alpha(\theta) = 1$ ) are given by

$$\phi_n(z, \alpha) = z^{\frac{n}{2}} \chi_n \gamma^n \begin{pmatrix} 1 \\ 0 \end{pmatrix}^T R(z) M(z) \begin{pmatrix} (\psi(z)/\sqrt{z})^n \\ (\psi(z)/\sqrt{z})^{-n} \end{pmatrix}, \quad (30)$$

where  $\chi_n$  is the leading coefficient of  $\phi_n(z, \alpha)$ ,  $\gamma = \cos(\alpha/2)$ ,  $R(z)$  is a  $2 \times 2$ -matrix-valued function which is analytic and satisfies  $R(z) = I + \mathcal{O}(n^{-1})$ ,  $M(z)$  is a  $2 \times 2$ -matrix-valued analytic function whose expression changes depending on whether  $z$  is in a neighbourhood of the endpoints of the arc or not (see below), and  $\psi(z)$  is a conformal mapping of the outside of the arc to the outside of the unit circle, given by

$$\psi(z) = \frac{1}{2\gamma} \left( z + 1 + \sqrt{(z - e^{i\alpha})(z - e^{-i\alpha})} \right).$$

The branch of the square root which is positive for positive arguments is taken. This is similar to [22, Eqn. 2.56], which gives the asymptotics of  $\phi_n(z, \alpha)$  in subsets of the complex plane outside a fixed neighbourhood of the arc. The job of this Lemma is to unpack this expression and convert to the variable  $\theta \in [\alpha, 2\pi - \alpha]$  where  $z = e^{i\theta}$ .

The leading coefficient has asymptotic expression  $\chi_n = \gamma^{-n-\frac{1}{2}}(1 + \mathcal{O}(n^{-1}))$  (by [22, Eq. 2.58]), and we have after some algebraic manipulation, for all  $\theta \in [\alpha, 2\pi - \alpha]$ ,

$$\psi(e^{i\theta})/\sqrt{e^{i\theta}} = \frac{1}{\gamma} \left( \cos\left(\frac{\theta}{2}\right) - i\sqrt{\sin\left(\frac{1}{2}(\theta + \alpha)\right)\sin\left(\frac{1}{2}(\theta - \alpha)\right)} \right).$$

Since  $|\psi(e^{i\theta})/\sqrt{e^{i\theta}}| = 1$  (which can be shown directly or inferred from the conformal mapping definition of  $\psi$  above), the function  $\tau(\theta)$  defined in the statement of the lemma maps  $\theta \in [\alpha, 2\pi - \alpha]$  to  $\tau \in [0, \pi]$ , and provides us with the simple identity,  $\psi(e^{i\theta})/\sqrt{e^{i\theta}} = e^{-i\tau(\theta)}$ . Substituting this into equation (30) we write,

$$\begin{aligned} \phi_n(e^{i\theta}, \alpha) &= (1 + \mathcal{O}(n^{-1}))e^{i\frac{n}{2}\theta}\gamma^{-\frac{1}{2}}(M_{11}(e^{i\theta})e^{-in\tau(\theta)} + M_{12}(e^{i\theta})e^{in\tau(\theta)}) \\ &\quad + \mathcal{O}(n^{-1})(M_{21}(e^{i\theta})e^{-in\tau(\theta)} + M_{22}(e^{i\theta})e^{in\tau(\theta)}) \\ &= e^{i\frac{n}{2}\theta}\gamma^{-\frac{1}{2}}(M_{11}(e^{i\theta})e^{-in\tau(\theta)} + M_{12}(e^{i\theta})e^{in\tau(\theta)}) + \mathcal{O}(n^{-1})e_n(\theta), \end{aligned}$$

where  $e_n(\theta) = |M_{11}(e^{i\theta})e^{-in\tau(\theta)} + M_{12}(e^{i\theta})e^{in\tau(\theta)}| + |M_{21}(e^{i\theta})e^{-in\tau(\theta)} + M_{22}(e^{i\theta})e^{in\tau(\theta)}|$ .

According to [22, Eq. 2.23], there exists  $\delta > 0$  so this asymptotic expression is valid for  $\theta \in [\alpha + \delta, 2\pi - \alpha - \delta]$  with  $M$  set as the function,

$$M(e^{i\theta}) = \frac{1}{2} \begin{pmatrix} a + a^{-1} & -i(a - a^{-1}) \\ i(a - a^{-1}) & a + a^{-1} \end{pmatrix}, \quad a(e^{i\theta}) = e^{i\frac{\alpha}{4}} \left( \frac{\sin\left(\frac{1}{2}(\theta - \alpha)\right)}{\sin\left(\frac{1}{2}(\theta + \alpha)\right)} \right)^{\frac{1}{4}},$$

and for  $\theta \in [\alpha, \alpha + \delta]$  with  $M$  set as the function,

$$\begin{aligned} M(e^{i\theta}) &= 2^{-\frac{5}{2}} \begin{pmatrix} a + a^{-1} & -i(a - a^{-1}) \\ i(a - a^{-1}) & a + a^{-1} \end{pmatrix} \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix} \\ &\quad \cdot \begin{pmatrix} (-i\pi n\tau)^{\frac{1}{2}} & 0 \\ 0 & (-i\pi n\tau)^{-\frac{1}{2}} \end{pmatrix} \begin{pmatrix} H_0^{(1)}(-n\tau) & H_0^{(2)}(-n\tau) \\ -i\pi n\tau(H_0^{(1)})'(-n\tau) & -i\pi n\tau(H_0^{(2)})'(-n\tau) \end{pmatrix} \\ &\quad \cdot \begin{pmatrix} e^{in\tau} & 0 \\ 0 & e^{-in\tau} \end{pmatrix}, \end{aligned}$$

where  $H_\nu^{(j)}$  is the  $\nu$ th Hankel function of the  $j$ th kind [13, Sec. 10.2].

For  $\theta \in [\alpha + \delta, 2\pi - \alpha - \delta]$ , we have  $e_n(\theta) = \mathcal{O}(\delta^{-\frac{1}{4}}) = \mathcal{O}(1)$ . Therefore, grouping terms to convert exponentials into trigonometric functions we obtain (28-29).

For  $\theta \in [\alpha, \alpha + \delta]$ , we can simplify the formula for  $M$  and obtain,

$$\begin{aligned} M_{11}(e^{i\theta})e^{-in\tau(\theta)} &= 2^{-\frac{3}{2}}(-i\pi n\tau)^{\frac{1}{2}} \left( a^{-1}H_0^{(1)}(-n\tau) - ia(H_0^{(1)})'(-n\tau) \right) \\ M_{12}(e^{i\theta})e^{in\tau(\theta)} &= 2^{-\frac{3}{2}}(-i\pi n\tau)^{\frac{1}{2}} \left( a^{-1}H_0^{(2)}(-n\tau) - ia(H_0^{(2)})'(-n\tau) \right) \\ M_{21}(e^{i\theta})e^{-in\tau(\theta)} &= 2^{-\frac{3}{2}}(-i\pi n\tau)^{\frac{1}{2}} \left( -ia^{-1}H_0^{(1)}(-n\tau) + a(H_0^{(1)})'(-n\tau) \right) \\ M_{22}(e^{i\theta})e^{in\tau(\theta)} &= 2^{-\frac{3}{2}}(-i\pi n\tau)^{\frac{1}{2}} \left( -ia^{-1}H_0^{(2)}(-n\tau) + a(H_0^{(2)})'(-n\tau) \right) \end{aligned}$$

Using the fact that  $J_\nu = \frac{1}{2}(H_\nu^{(1)} + H_\nu^{(2)})$  [13, Eq. 10.4.4],  $J'_0 = -J_1$ ,  $J_0(-z) = J_0(z)$ , and  $J_1(-z) = -J_1(z)$ , we obtain equations (29-30), sans the remainder term  $\mathcal{O}(n^{-1})e_n(\theta)$ , which we must show to be  $\mathcal{O}(n^{-\frac{1}{2}})$ . Collecting the terms, we have,

$$e_n(\theta) = 2^{-\frac{1}{2}}(-i\pi n\tau)^{\frac{1}{2}} \left( |a(e^{i\theta})^{-1}J_0(n\tau) - ia(e^{i\theta})J_1(n\tau)| + |ia(e^{i\theta})^{-1}J_0(n\tau) - a(e^{i\theta})J_1(n\tau)| \right).$$

For all  $\tau \in [0, \frac{\pi}{2}]$ ,  $\tau^2 \leq \frac{\pi^2}{4}(1 - \cos(\tau))$ , and  $\cos(\tau) = \cos(\theta/2)/\gamma$ , so  $\tau^2 \leq \frac{\pi^2}{4\gamma}(\gamma - \cos(\theta/2)) = \frac{\pi^2}{2\gamma} \sin(\frac{1}{4}(\theta + \alpha)) \sin(\frac{1}{4}(\theta - \alpha))$ . For  $\theta \in [\alpha, \pi]$ , we have  $\sin(\frac{1}{4}(\theta - \alpha)) \leq \sin(\frac{1}{2}(\theta - \alpha))$ , so we can conclude,

$$\left( \frac{\tau(\theta)^2}{\sin(\frac{1}{2}(\theta - \alpha))} \right)^{\frac{1}{4}} = \mathcal{O}(1),$$

uniformly for all  $\theta \in [\alpha, \pi]$ . Note also that Bessel functions are uniformly bounded in absolute value by 1 (see [13, Eq. 10.14.1]). This makes it clear that  $e_n(\theta) = \mathcal{O}(n^{\frac{1}{2}})$ , as required.

The fact that  $\psi_n(e^{-i\theta}, \alpha) = \overline{\psi_n(e^{i\theta}, \alpha)}$  follows from the fact that the weight satisfies  $f(-\theta) = f(\theta)$ , so the coefficients of  $\psi_n(z, \alpha)$  are real (see [35, p. 288]).  $\square$