ORTHOGONAL POLYNOMIALS ON THE UNIT CIRCLE: STEKLOV PROBLEMS AND WEIGHT PERTURBATIONS

By

Keith Rush

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

(Mathematics)

at the

UNIVERSITY OF WISCONSIN – MADISON

2016

Date of final oral examination: December 16, 2016

The dissertation is approved by the following members of the Final Oral Committee:

Professor S. Denissov, Professor, Mathematics (Reader)

Professor A. Seeger, Professor, Mathematics (Reader)

Professor B. Street, Associate Professor, Mathematics (Reader)

Professor B. Stovall, Assistant Professor, Mathematics

Professor B. Valkó, Professor, Mathematics

Abstract

We consider problems inspired by the conjecture of Steklov and recent work on sharp estimates in this conjecture. We use various techniques to investigate Steklov problems in the measure dimension as well as that of the Verblunsky parameters. In particular, we consider continuous weights, BMO weights, randomized measures, and the Verblunsky asymptotics of a measure which violates the Steklov conjecture.

Acknowledgements

I owe a tremendous debt of gratitude to my advisor, Serguei Denissov. I thank him once for always knowing what I am saying much better than I do. I thank him again for having the forbearance to avoid mentioning that fact.

I'd like to thank the people that have shaped my life in recent years. In particular, but in no order: Greg Myhr, for showing me that it doesn't have to be fun to be fun; Ed Golding, for showing me that it is OK to silently race the people around you without their knowledge; Mitchell Stein, for showing what dedication to one's friends can really mean; Matthew Jackson, for consistently beating me to the punch, and for living by his own rules; James Sands, for his example; Christopher "Max" Schumitsch for being entirely genuine (and housing me for 3 weeks); Dariush "Tim" Aghai, for teaching me about plateauing ("If you don't want to plateau, then don't") as well as my true advantage ("You live in the well. Drag them down there too"); Vince Bertram, for making my shouting seem less out-of-place at the boathouse; Will Horvat, for providing the perfect counterpoint to my boundless optimism, and sharing my disdain for everything; Chase Tarrier, for always being truly inquiring (and encouraging); Deanna Sell, for smacking sense into me when I've needed it; Jonathan Dobson, for helping me really not care what others think; Chris Clark, for being an incredible coach, the greatest teaching model conceivable, and a constant source of quotes to live by ("You want to be good? You got to be *good*!"); Lowell McNicholas, for always being himself; Beau Hoopman, for showing me that nobody need be impressed with themselves, and that it is certainly not the point; LaMar "Soup" Campbell, for having the grace to treat me as if I were as cool and on-the-ball as him; my students, for breaking the monotony and reminding me that there are, in fact, people out there; Penn heavyweight 2013, for the hours we put in together, and for never having pity; Penn heavyweight 2015, for accepting an old man into their ranks; Wisconsin 2014 for generally being on the constant verge of psychosis; everyone else, for just doing whatever you've done–blood and bandage, sweat and sweetness, laughter and tears, I wouldn't have it any other way.

I thank my sister, Margaret Rush, for helping me keep in mind that nobody's life was at stake.

Finally, I thank my parents, David and Catherine Rush, for teaching me early and often that like integrability of the log of your a.c. part and square summability of your Verblunsky coefficients, blessings and burdens are identical.

Epigraph

"We should consider every day lost on which we have not danced at least once. And we should call every truth false which was not accompanied by at least one laugh."

–Friedrich Nietzsche

"Done saying I'm done playing." –Aubrey Drake Graham

Contents

Chapter 1

Introduction

1.1 Szegő, his recurrence, and the beginning of the subject

Consider a probability measure $d\mu$ defined on the unit circle $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$,

$$\int_{\mathbb{T}} d\mu = 1$$

We will assume infinitely many points of growth so that $L^2(\mathbb{T}, d\mu)$ is infinite-dimensional, although the finite-dimensional case is also interesting (see [41]).

By applying the Gram-Schmidt orthogonalization procedure (without normalization) to the set of monomials $\{1, z, z^2, \dots\}$, we create a sequence of monic orthogonal polynomials $\{\Phi_0, \Phi_1, \Phi_2, \dots\}$. These satisfy

$$\int_{\mathbb{T}} \Phi_j \overline{\Phi_k} d\mu = 0, \quad j \neq k, \quad \operatorname{coeff}(\Phi_n, n) = 1 \tag{1.1}$$

where $\operatorname{coeff}(P, j)$ denotes the j^{th} coefficient of the polynomial P. One may similarly define the orthonormal polynomials by setting

$$\phi_n(z) = \frac{\Phi_n(z)}{\|\Phi_n\|_{L^2(d\mu)}}$$

These are simple and natural objects with remarkable properties. Some of the gateways to their study are (see [65], Section 1.1):

1. *Linear prediction theory.* Given a stationary sequence of complex random variables $\{\omega_j\}_{j=-\infty}^{\infty}$, the Carathéodory-Toeplitz Theorem asserts the existence of a measure $d\mu$ on $\mathbb{T}$ such that

$$\mathbb{E}(\omega_j \overline{\omega_k}) = \int_{\mathbb{T}} e^{i(j-k)\theta} d\mu(\theta)$$

Therefore, representing ω_j by z^j yields $L^2(\mathbb{T}, d\mu)$ as a model for the linear part of the stochastic process. See ([43, 44, 45, 50, 72]) for this point of view.

2. *Toeplitz operators and determinants.* A classical subject in analysis (especially statistical mechanics), the Toeplitz operator T_f with symbol f is defined to be the compression of the multiplication operator M_f onto the Hardy space $H^2(\mathbb{T})$. These operators have constant diagonals in their matrix representations on $\{z^n\}_{n=0}^{\infty}$. Several identities relate the asymptotics of orthogonal polynomials to Toeplitz matrices and determinants, for example:

Theorem 1.1 ([65], Theorem 1.5.11)**.** *Let $\{c_n\}$ be the moments of $d\mu$ and Φ_n its orthogonal polynomials. Let*

$$T^{(n)} = \begin{bmatrix} c_0 & c_1 & \dots & c_{n-1} \\ c_{-1} & c_0 & \dots & c_{n-2} \\ \dots & \dots & \dots & \dots \\ c_{-n+1} & \dots & \dots & c_0 \end{bmatrix}$$

and

$$D_n(d\mu) = \det(T^{(n+1)})$$

Then

- *(Heine's Formula)*

$$\Phi_n(z) = D_{n-1}(d\mu)^{-1} \begin{vmatrix} c_0 & \overline{c_1} & \dots & \overline{c}_n \\ c_1 & c_0 & \dots & \overline{c}_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n-1} & c_{n-2} & \dots & \overline{c}_1 \\ 1 & z & \dots & z^n \end{vmatrix}$$

-

$$\|\Phi_n(z)\|^2_{L^2(d\mu)} = \prod_{j=0}^{n-1}(1-|\alpha_j|^2) = \frac{D_n(d\mu)}{D_{n-1}(d\mu)}$$

-

$$1-|\alpha_n|^2 = \frac{D_{n+1}D_{n-1}}{D_n^2}$$

3. *Spectral theory of one-dimensional Schrödinger operators.* A discrete one-dimensional Schrödinger operator for a single particle in electric field is defined by the equation

$$hu(n) = u(n+1) + u(n-1) + V(n)u(n)$$

where V is the potential. This equation leads naturally to the study of Jacobi matrices and orthogonal polynomials on the real line (OPRL), and the connections between OPRL and OPUC allow in some cases for back-and-forth translations of results in the OPUC literature and those in the spectral theory of operators h. This perspective also relates to the theory of Krein systems. Ideas and results in each area has influenced the development of the others.

4. *Random matrix theory.* Through the fomulae relating certain random matrix models and orthogonal polynomials, asymptotics of OPs have entered random matrix

theory in a big way, often providing the original proofs of many major results in the random matrix literature.

As a field distinct from the classical orthogonal polynomials (e.g., Hermite and Laguerre) their theory began with Gabor Szegő, who pioneered a structural view of the subject. Take for example Simon's simple proof of the Szegő recurrence.

Theorem 1.2 (Szegő). *For a nontrivial probability measure $d\mu$ on $\mathbb{T}$, the orthogonal polynomials defined in* (1.1) *and their $*$-polynomials defined by*

$$\Phi_n^*(z) = z^n \overline{\Phi_n\left(\frac{1}{\bar{z}}\right)}, \quad z \in \mathbb{C}\backslash\{0\}$$

satisfy the following recurrence, for $\alpha_n \in \mathbb{D}$:

$$\begin{cases} \Phi_{n+1}(z) = z\Phi_n(z) - \overline{\alpha_n}\Phi_n^*(z) \\ \Phi_{n+1}^*(z) = \Phi_n^*(z) - z\alpha_n\Phi_n(z) \end{cases} \tag{1.2}$$

Moreover,

$$\|\Phi_n\|^2_{L^2(d\mu)} = \prod_{j=0}^{n-1}(1 - |\alpha_j|)^2 \tag{1.3}$$

Proof. The second equation is obtained by applying the $(n+1)$-st order $*$-operation to the first so it suffices to show the first.

Φ_{n+1} is the unique monic polynomial of degree $n+1$ which is orthogonal to $\{1, z, \ldots, z^n\}$ in $L^2(d\mu)$. Notice that, since $\operatorname{coeff}(\Phi_n(z), n) = 1$, the right hand side of the claimed equation is monic of degree $n+1$. It remains to check the orthogonality condition.

Let $1 \leq j \leq n$. Then

$$(z\Phi_n(z), z^j)_{L^2(d\mu)} = (\Phi_n(z), z^{j-1})_{L^2(d\mu)} = 0$$

Likewise

$$(\Phi_n^*(z), z^j)_{L^2(d\mu)} = (z^n\overline{\Phi_n(z)}, z^j)_{L^2(d\mu)} = (\overline{\Phi_n(z)}, z^{j-n})_{L^2(d\mu)} = \overline{(z^{n-j}, \Phi_n(z))}_{L^2(d\mu)} = 0$$

So if we set

$$\overline{\alpha_n} = \frac{\int_{\mathbb{T}} z\Phi_n(z)d\mu(z)}{\int_{\mathbb{T}} \Phi_n^*(z)d\mu(z)}$$

the claimed equation is satisfied.

Since Φ_n^* is a polynomial of degree n, the Pythagorean Theorem says

$$\|\Phi_{n+1}\|^2_{L^2(d\mu)} + |\alpha_n|^2\|\Phi_n^*\|^2_{L^2(d\mu)} = \|\Phi_n\|^2_{L^2(d\mu)}$$

By definition $\|\Phi_n^*\|^2_{L^2(d\mu)} = \|\Phi_n\|^2_{L^2(d\mu)}$, so

$$\|\Phi_{n+1}\|^2_{L^2(d\mu)} = (1 - |\alpha_n|^2)\|\Phi_n\|^2_{L^2(d\mu)}$$

which proves $\alpha_n \in \mathbb{D}$ and by induction the equation

$$\|\Phi_n\|^2_{L^2(d\mu)} = \prod_{j=0}^{n-1}(1 - |\alpha_j|)^2$$

□

Szegő was able to prove the basic Theorems which touch every aspect of the theory. One particularly fundamental result bears his name.

Theorem 1.3 (Szegő's Theorem)**.** *Let $d\mu$ be a nontrivial probability measure on $\mathbb{T}$, and denote $d\mu = w(\theta)\frac{d\theta}{2\pi} + d\mu_s$. Then*

$$\exp\left(\int_{\mathbb{T}} \log(w(\theta))\frac{d\theta}{2\pi}\right) = \prod_{j=0}^{\infty}(1 - |\alpha_j|^2) \tag{1.4}$$

Remark. Coupled with the equation (1.3), Szegő's Theorem tells us that the monic and orthonormal polynomials are uniformly comparable in n if the weight w satisfies the Szegő condition, that the integral on the left-hand side of (1.4) is finite. All of the weights we work with will satisfy this condition.

The Szegő function is defined

$$\mathcal{D}(z; d\mu) = \exp\left(\frac{1}{4\pi}\int_{\mathbb{T}} \frac{e^{i\theta}+z}{e^{i\theta}-z}\log w(\theta)d\theta\right) \tag{1.5}$$

for $d\mu = w\frac{d\theta}{2\pi} + d\mu_s$ which satisfies the Szegő condition. This function is intimately related to the polynomials (see [65], section 2.4). $\mathcal{D}$ will appear repeatedly in what follows, so we list some of its basic properties.

Theorem 1.4 (Szegő)**.** *Suppose the Szegő condition holds. Then*

1. $\mathcal{D}(z)$ *is analytic and nonvanishing in* $\mathbb{D}$.

2. $\mathcal{D}$ *lies in the Hardy space* $H^2(\mathbb{D})$. *In fact*

$$\sup_{0<r<1} \int |\mathcal{D}(re^{i\theta})|^2 \frac{d\theta}{2\pi} \equiv \|\mathcal{D}\|_{H^2}^2 \leq 1$$

3.

$$\lim_{n\to\infty} \phi_n^*(z) = \mathcal{D}(z)^{-1}$$

uniformly on compact subsets of $\mathbb{D}$

4. $\lim_{r\uparrow 1} \mathcal{D}(re^{i\theta}) = \mathcal{D}(e^{i\theta})$ *exists a.e.* $d\theta$, *and*

$$|\mathcal{D}(e^{i\theta})|^2 = w(\theta)$$

5. *We have the following limit formulae*

$$\lim_{n\to\infty} \int |\overline{\mathcal{D}(e^{i\theta})}\phi_n(e^{i\theta}) - e^{in\theta}|^2 \frac{d\theta}{2\pi} = 0$$

$$\lim_{n\to\infty} \int |\mathcal{D}(e^{i\theta})\phi_n^*(e^{i\theta}) - 1|^2 \frac{d\theta}{2\pi} = 0$$

6.

$$\lim_{n\to\infty} \int |\phi_n^*(e^{i\theta})|^2 d\mu_s(\theta) = 0$$

1.2 Renewed interest via determinantal point processes

Interest in the field of orthogonal polynomials was spurred further by the growth of random matrix theory, begun in 1955 by Eugene Wigner [73]. By 1960, M.L. Mehta in [53] had recognized a fundamental connection between orthogonal polynomials and determination of eigenvalue densities through the Vandermonde determinant:

$$\Delta_n = \prod_{1\le i<j\le n}(x_j - x_i) = \begin{vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^2 & x_2^2 & \dots & x_n^2 \\ \vdots & & & \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{vmatrix}$$

Consider the Gaussian Unitary Ensemble (GUE), the ensemble of $n \times n$ symmetric matrices whose upper triangular elements are complex $\mathcal{N}(0,1)$ random variables and whose diagonal entries are real $\mathcal{N}(0,1)$ random variables. The eigenvalue density of this ensemble can be computed explicitly; it is of the form, for $\lambda = (\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n$ (see [17], (2.1.27))

$$\rho_n(\lambda) = \frac{1}{(2\pi)^{n/2}} \exp\left(\frac{-|\lambda|^2}{2}\right) |\Delta_n(\lambda)|^2$$

Multiplicity of the determinant yields

$$|\Delta_n(\lambda_1, \dots, \lambda_n)|^2 = \det\left(\sum_{k=0}^{n-1} \lambda_i^k \lambda_j^k\right)_{1\le i,j\le n}$$

Applying elementary row operations which leave the determinant unchanged, for any sequence of polynomials $P_0(x), \ldots P_{n-1}(x)$ where P_i is monic of degree i,

$$|\Delta_n(\lambda_1, \ldots, \lambda_n)|^2 = \det \left(\sum_{k=0}^{n-1} P_k(\lambda_i) P_k(\lambda_j) \right)_{1 \le i,j \le n}$$

This formula is valid for any polynomial sequence, but if the polynomials are chosen to be orthogonal with respect to a particular weight (the Hermite polynomials are chosen) then the sum

$$K_n(x,y) = \sum_{k=0}^{n-1} P_k(x) e^{-x^2/4} P_k(y) e^{-y^2/4}$$

becomes the integral kernel of the projection onto $\{x^i e^{-x^2/4}\}_{i=0}^{n-1}$in $L^2(\mathbb{R})$. This algebraic structure yields the trace and reproducing formulas, which allow computation of k-point correlation functions. This is an example of a determinantal point process, an area where orthogonal polynomials in various incarnations have proved useful computational tools; see [8, 9, 10] for examples. This connection has been heavily utilized and may indeed be taken as a basic point of view on random matrices; see the excellent reference [17].

This correspondence has served to motivate definitions of new OPUC and OPRL as well as motivated their asymptotic study. Techniques and results in each area have proved extremely useful in the other.

1.3 Spectral theory and OPUC

Consider the discrete Schrödinger operator, which in one dimension takes the form

$$(hu)_n = u_{n+1} + u_{n-1} + v_n u_n \tag{1.6}$$

and its continuum analogue,

$$(Hu)(x) = -u''(x) + V(x)u(x) \tag{1.7}$$

In the formalism of quantum mechanics, the first physically relevant and interesting question concerns the type of spectrum associated with this operator. Roughly speaking, eigenvalues correspond to bound states of the system, and absolutely continuous spectrum to scattering states, where particles get "lost to infinity".

It is a classical fact that if $|V(x)| \leq C(1+|x|)^{-a}$ for $a > 1$, H has purely a.c. spectrum at positive energies ([23]). This is a perturbative result, with comparison made to the free particle, with $V \equiv 0$.

A natural question is whether $a = 1$ is the true dividing line. In fact, it is not; $a = \frac{1}{2}$ is the critical case. Christ-Kiselev [15] showed there is always a.c. spectrum for $1 > a > 1/2$. In 1999, Simon produced a list of open problems which included the question of mixed singular continuous spectrum with potential $V \in L^2$.

By 2001, the advisor of this author, Serguei Denissov, had solved this problem. In [20] he showed that there were L^2 potentials with essentially arbitrary singular part on $[0, E_0]$ for any $E_0 < \infty$.

Denissov's proof used the systems of Krein:

$$\begin{cases} \frac{\partial}{\partial \tau} P(\tau, \lambda) = i\lambda P(\tau, \lambda) + P_*(\tau, \lambda) a(\tau) \quad 0 \leq \tau \leq T \\ \frac{\partial}{\partial \tau} P_*(\tau, \lambda) = P(\tau, \lambda) \overline{a(\tau)} \end{cases} \tag{1.8}$$

Solutions to this system are continuous analogs of OPUC (many of the results here could be translated immediately into that context). A main technique in Denissov's proof involved the use of a sum rule which gave the L^2 properties of his potential. These sum rules have provided an important approach to Schrödinger operators. Recall Szegő's

Theorem and equation (1.4), which asserts

$$\exp\left(\frac{1}{2\pi}\int_{\mathbb{T}}\log(w(\theta))d\theta\right)=\prod_{j=0}^{\infty}(1-|\alpha_j|^2)$$

Taking the logarithm, we have:

$$\frac{1}{2\pi}\int_{\mathbb{T}}\log(w(\theta))d\theta=\sum_j\log(1-|\alpha_j|^2)$$

From one perspective this is exactly an if and only if statement of spectral theory. The Verblunsky coefficients take the place of the potential, and their ℓ^2 decay is exactly equivalent to the integrability of the logarithm of the absolutely continuous part of the measure. Notice there is *no dependence on the singular part.* The sum rules which have since appeared repeatedly in mathematical physics (see [12], [42] as examples) are gems of this form. Sum rules have become extremely powerful–indeed, indispensable–tools in spectral theory. OPUC is the developing ground for these rules: Szegő's Theorem, in this form proved by Verblunsky [71], is from 1936! Further, see the recent work of Lukic [51], [52].

The analogy between Schrödinger operators and orthogonal polynomials continues to animate the work on OPUC and OPRL. One interesting open question, introduced to the author by his advisor, is:

What conditions on the weight function w will yield a.e. $d\theta$ finiteness of the maximal function

$$M\Phi(z)=\sup_n|\Phi_n(z)|,\quad z\in\mathbb{T}$$

This question appears to be extremely difficult. The orthogonal polynomials Φ_n are analogous to the eigenfunctions of a one-dimensional Schrödinger operator, a statement to be made more precise shortly. For now, following the excellent survey [23] (themselves following [46]) we introduce the Prüfer variables in the continuous case to better explain the issues involved. Let u be a solution to the eigenfunction equation for $k \in \mathbb{R}$

$$-u'' + Vu = k^2 u$$

Define the variables $R(x, k)$ and $\theta(x, k)$ by

$$u'(x, k) = kR(x, k) \cos \theta(x, k)$$

$$u(x, k) = R(x, k) \sin \theta(x, k)$$

These variables satisfy the new system

$$(\log(R(x, k))^2)' = \frac{1}{k} V(x) \sin 2\theta(x, k) \tag{1.9}$$

$$\theta(x, k)' = k - \frac{1}{k} V(x)(\sin \theta(x, k))^2 \tag{1.10}$$

This is useful in analysis of decaying potentials because R has dropped out of (1.10), and so can be obtained by integration once (1.10) is solved. Further, letting $V = 0$ in some region (a, b), then $\theta(x) = kx + \theta(a)$ and R is constant. Since we consider the decaying potential to be a perturbation of $V = 0$, these are hopeful indications.

It should be unsurprising therefore that there are methods of analysis in some particular cases (e.g., random and sparse potentials), and that these share much with classical approaches in Fourier analysis. That is, formally solving (1.9) yields

$$R(x, k) = \exp\left(\frac{1}{2k} \int_0^x V(s) \sin(2\theta(s, k)) ds\right)$$

so that instead of the usual (linear) Fourier integral $\int_{-\infty}^{\infty} V(x) \sin(kx) dx$, we are investigating a nonlinear variant.

Indeed, much work on the behavior of R can be thought of as analogous to classical work on Fourier series. In this context that we may approach the question of the maximal function $M\Phi$.

Recall Lusin's conjecture (Carleson's Theorem, [13]) on a.e. convergence of Fourier integrals.

Theorem 1.5 (Carleson)**.** *If $g \in L^2(\mathbb{R})$, then the associated Fourier integrals*

$$\int_{-N}^{N} e^{2\pi i \xi x} g(x) dx$$

converge for almost every value of the parameter ξ.

Since $V \in L^2$ is also the natural breaking point in our setup, it makes sense to ask the same question here. That is, for $V \in L^2$, does $\sup_{x \in \mathbb{R}} |R(x,k)|$ make sense for a.e. k?

Since Carleson's Theorem (from 1955!) remains a pinnacle of modern analysis, it is reasonable to weaken the statement. What, we may ask, happens if $V \in L^p$ with $p < 2$? In the linear case, this was much simpler, already solved by Zygmund [74] in 1928 (see also Menshov [54] and Paley [59]). And indeed, the $p < 2$ case has been resolved by Christ and Kiselev [14], but their methods have no hope of working when $p = 2$ as was shown by Muscalu-Tao-Thiele [56].

A different perspective can be had when viewing the problem through the OPUC analogue. To make this more concrete, recall the discrete Schrödinger equation (1.6)

$$(hu)_n = u_{n+1} + u_{n-1} + v_n u_n$$

The eigenfunction u then satisfies a three-term recurrence relation

$$(hu)_n = \lambda u_n = u_{n+1} + u_{n-1} + v_n u_n$$

The set of orthonormal polynomials $\{p_k\}$ with respect to a measure μ on the real line satisfies a similar recurrence

$$xp_n(x) = a_{n+1}p_{n+1}(x) + b_{n+1}p_n(x) + a_n p_{n-1}(x)$$

for some sequence of parameters $\{a_n > 0\}, \{b_n \in \mathbb{R}\}$. Clearly $a_n \equiv 1$ corresponds to the discrete Schrödinger operator. These coefficients determine a Jacobi matrix

$$J_\mu = \begin{bmatrix} b_1 & a_1 & 0 & 0 & \dots \\ a_1 & b_2 & a_2 & 0 & \dots \\ 0 & a_2 & b_3 & a_3 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix}$$

(notice our numbering convention follows [65], itself following [42]) which represents multiplication by x in $L^2(\mathbb{R}, \mu)$ with respect to the basis $\{p_n(x)\}$. Study of the asymptotics of eigenfunctions of the discrete Schrödinger operator is equivalent to the study of the polynomials.

Measures of orthogonality with compact support can be mapped onto the unit circle, and this is the analogy with Schrödinger operators we exploit. The polynomials Φ_n take the place of the eigenfunctions, the Verblunsky parameters that of the potential, and the measure $d\mu$ becomes spectral measure. The transfer matrix approach in spectral theory (see [47]) is now analyzing the large$-n$ asymptotics of the polynomials Φ_n. This is made clear by the matrix equation the polynomials satisfy:

$$\begin{bmatrix} \Phi_{n+1}(z) \\ \Phi^*_{n+1}(z) \end{bmatrix} = \begin{bmatrix} z & -\overline{\alpha_n} \\ -\alpha_n z & 1 \end{bmatrix} \begin{bmatrix} \Phi_n(z) \\ \Phi^*_n(z) \end{bmatrix}$$

And so the questions about asymptotics of these orthogonal polynomials can be viewed as nonlinear analogues of questions about Fourier series. In particular, a.e.

boundedness of the maximal function $M\Phi(z)$ is the analogue of a.e. convergence of the Fourier integrals.

Notice now what Szegő's Theorem 1.3 represents: an equivalence between spectral data and potential decay. If we can say something reasonable about behavior of the eigenfunctions in terms of the spectral data, we could perhaps approach the nonlinear analogue of Lusin in a different dimension. This approach is in fact a different perturbative style; instead of taking potential as primary and perturbing some well-understood case in that dimension (e.g., imposing some strict decay on the Verblunsky parameters), we may take measure as primary and specify some regularity.

1.4 Steklov's conjecture and inspired problems

In 1921, V.A. Steklov conjectured in [67] that if the weight function ρ on $[-1,1]$ stays bounded away from 0 ($\rho(x) \geq \delta$ for a.e. x), the orthonormal polynomials $p_n(x)$ would stay bounded for $x \in (-1,1)$;

$$\limsup_n |p_n(x)| \leq C(\delta), \quad x \in (-1,1)$$

For the full history see the survey [68], as well as the work [31, 32, 33, 34, 35]. This conjecture was disproved by Rakhmanov in 1979 [62], who constructed a weight in Steklov's class on the unit circle and translated it to the real line. The conjecture of Steklov, disproved, became a problem of Steklov: how fast can the polynomials grow under some conditions on the weight functions? The first question one must answer is how to measure this "growth"; in the Steklov problems, this is the uniform norm. It has been well-known that maximal growth in this norm is $o(\sqrt{n})$; see for example [32].

Rakhmanov in [62, 63] began considering the following variational problem:

$$M_{n,\delta} = \sup_{\sigma \in S_\delta} \|\phi_n(z,\sigma)\|_{L^\infty(\mathbb{T})}.$$

where we define the Steklov class of measures as

$$S_\delta = \left\{ \sigma : \int_{\mathbb{T}} d\sigma = 1, \quad \sigma' \geqslant \delta/(2\pi), \quad \text{a.e.} \quad \theta \in [0, 2\pi) \right\},$$

for $\delta \in (0,1)$.

In [63], the following estimates were established

$$\left(\frac{n}{\ln^3 n} \right)^{1/2} <_\delta M_{n,\delta} <_\delta n^{1/2}.$$

The sharp result remained open until 2014, when Aptekarev-Denisov-Tulyakov [2] were able to construct optimal examples. They improved the estimates to

$$M_{n,\delta} \sim_\delta n^{1/2}.$$

The paper of Aptekarev, Denisov and Tulyakov completely settled the original problem of Steklov. In addition, it opened new perspectives on analysis of polynomials in terms of conditions on the weight function. [2] made it clear that the Steklov condition alone would not yield anything more than the nearly-trivial $o(\sqrt{n})$ bound on uniform norm but also raised new questions. The Steklov condition is a condition on how small the weight can be; is there a similar condition on how large the weight can be that will change the behavior of the polynomials?

[2] proved that assuming $w \in L^p(d\theta)$ for any $p < \infty$ still allows for weights achieving $o(\sqrt{n})$. Nazarov (see [21]) shows that L^∞ control of the weight allows one to break the $n^{1/2}$ barrier; in particular, he shows

Theorem 1.6 (Nazarov)**.** *Let $t > 2$. Assume that the polynomials Φ_n are monic orthogonal with respect to $d\mu = w(\theta)d\theta$. If the weight w satisfies*

$$1 \le w(\theta) \le \epsilon, \quad \epsilon < 1$$

then

$$\|\Phi_n\|_p \lesssim 1, \quad p = C\epsilon^{-1}$$

If

$$1 \le w(\theta) \le T, \quad T > 2$$

then

$$\|\Phi_n\|_p \lesssim 1, \quad p = 2 + CT^{-1}$$

The Nikolskii inequality (2.1) immediately implies the bound

$$\|\Phi_n\|_\infty \le C(p)n^{1/p}$$

for the choices of p above. Sharpness of this result is not completely understood; in [21] a weight satisfying $1 \le w \le T$ is constructed which achieves $n^{\frac{1}{2}-\frac{C}{\sqrt[4]{T}}}$. In the small-variation regime, up to a constant in the exponent the upper bound saturates.

These results show that taking the perturbative approach in the measure dimension can in fact bear fruit. This approach and related issues are the main topics of this thesis.

The following notations will be used throughout the thesis. Some special notations may be adopted within each section.

$\|\cdot\|_p$ will always denote the $L^p\left(\mathbb{T}; \frac{d\theta}{2\pi}\right)$ norm.

Absolute constants will be denoted C, c or c_j, C_j. These constants may change between expressions. A constant which depends on a parameter λ may be denoted $C(\lambda)$ or C_λ, and these may likewise change between expressions.

For a function f, f^{-1} will denote its multiplicative inverse.

If $f_{1(2)}(x)$ are two positive functions for which $f_1 < Cf_2$ with some absolute constant C, uniformly in the argument, we will write $f_1 \lesssim f_2$. If $f_1 \lesssim f_2$ and $f_2 \lesssim f_1$, then $f_1 \sim f_2$. If

$$\sup_x \frac{f_1(x)}{f_2(x)} < C(\epsilon),$$

where ϵ is a parameter, then we will write $f_1 <_\epsilon f_2$. Relations $f_1 \sim_\epsilon f_2, f_1 >_\epsilon f_2$ are defined similarly.

$\Phi_n(z; d\mu), \phi_n(z; d\mu)$ will denote the monic orthogonal and orthonormal polynomials respectively of degree n with respect to the measure $d\mu$. As we have been doing in the introduction, $d\mu$ will often be suppressed.

For a measure $d\mu$ on $\mathbb{T}$ we will use w to denote its a.c. part with respect to normalized Lebesgue measure.

If α is a positive parameter, we write $\alpha \ll 1$ to indicate the following: there is an absolute constant α_0 (sufficiently small) such that $\alpha < \alpha_0$. Similarly, we write $\alpha \gg 1$ as a substitute for: there is a constant α_0 (sufficiently large) so that $\alpha > \alpha_0$. The symbol $\alpha_1 \ll \alpha_2$ ($\alpha_1 \gg \alpha_2$) will mean $\alpha_1/\alpha_2 \ll 1$ ($\alpha_1/\alpha_2 \gg 1$).

If A is a linear operator from $L^p(\mathbb{T})$ space to $L^p(\mathbb{T})$, then $\|A\|_{p,p}$ denotes its operator norm.

Chapter 2

Asymptotic polynomial behavior, perturbed weight regime

Questions about a.e. convergence and maximal functions are not the first to be asked from the perspective of analogy with Fourier series. We learn introductory Fourier analysis as an $L^2(\mathbb{T}) \leftrightarrow \ell^2(\mathbb{Z})$ Hilbert space isomorphism. Not far behind this is the Calderón-Zygmund theory of singular integrals, establishing the convergence of Fourier series and integrals in L^p for $1 < p < \infty$. Therefore investigating the $L^p(d\theta)$ norms of the orthogonal polynomials is a natural direction of inquiry.

From the point of view of Steklov problems, there is another reason why these quantities should be highly relevant. The polynomial inequality of Nikolskii states,, for P_n a polynomial of degree n,

$$\|P_n\|_q \leq \Lambda_{n,p,q} n^{1/p-1/q} \|P_n\|_p, \quad 1 \leq p \leq q \leq \infty \tag{2.1}$$

where $\limsup_{n\to\infty} \Lambda_{n,p,q} \leq C(p,q)$ (see [11]).

The exact dependence of this constant on n, p and q has been a difficult problem; the paper [48] provides a bound which is asymptotically sharp in n for $q = \infty$, and [49] discusses more general p and q.

The Nikolskii inequality implies that quantitative control of L^p norms for fixed finite p also yields information on L^∞ norms.

How do we investigate these L^p norms? The main observation to be used in this section is an elementary identity, the usefulness of which goes back to Sergei Bernstein ([7], [69]). The reproducing kernel of projection onto the first n polynomials in the Hilbert space $L^2(d\mu)$ is called the Christoffel-Darboux kernel

$$K_n(z,\zeta;d\mu) = \sum_{j=0}^{n} \phi_j(z;d\mu)\overline{\phi_j(\zeta;d\mu)}$$

The Christoffel-Darboux formula (see [65] Section 2.2) allows us to greatly simplify the expression of this kernel and is extremely useful in bounding the integrals involved:

$$K_n(z,\zeta) = \frac{\overline{\phi^*_{n+1}(z)}\phi^*_{n+1}(\zeta) - \overline{\phi_{n+1}(z)}\phi_{n+1}(\zeta)}{1-\overline{\zeta}z} \tag{2.2}$$

Let $d\mu_1$ and $d\mu_2$ be measures on $\mathbb{T}$ with $\{\phi_n(z;d\mu_1)\}, \{\phi_n(z;d\mu_2)\}$ their associated orthonormal polynomials. $\{1, z, \ldots, z^n\}$ is a finite-dimensional vector space; each of $\{\phi_j(z;d\mu_1)\}_{j=0}^n$ and $\{\phi_j(z;d\mu_2)\}_{j=0}^n$ is a basis for this vector space. So we may expand either in the basis of polynomials defined by the other. For example

$$\phi_n(z;d\mu_1) = \sum_{j=0}^{n} \langle \phi_n(z;d\mu_1), \phi_j(z;d\mu_2)\rangle_{L^2(d\mu_2)} \phi_j(z;d\mu_2)$$

$$= a_n\phi_n(z;d\mu_2) + \int_{\mathbb{T}} K_{n-1}(z,\zeta;d\mu_2)\phi_n(\zeta;d\mu_1)d\mu_2$$

If $d\mu_1, d\mu_2$ both satisfy the Szegő condition then $a_n = O(1)$ uniformly in n since each of $\phi_n(z;d\mu_1), \phi_n(z;d\mu_2)$ is uniformly comparable to its monic orthogonal polynomial.

Notice that $K_{n-1}(z,\zeta;d\mu_2)$ is a polynomial of degree $n-1$. Since $\phi_n(\zeta;d\mu_1)$ is orthogonal to all such polynomials in $L^2(\mathbb{T}, d\mu_1)$ we have the trivial identity

$$\int_{\mathbb{T}} K_{n-1}(z,\zeta;d\mu_2)\phi_n(\zeta;d\mu_1)d\mu_1 = 0$$

and therefore

$$\phi_n(z;d\mu_1) = a_n\phi_n(z;d\mu_2) + \int_{\mathbb{T}} K_{n-1}(z,\zeta;d\mu_2)\phi_n(\zeta;d\mu_1)(d\mu_2 - d\mu_1) \tag{2.3}$$

The equation (2.3) is the basis of the weight-pertubative approach which yields the upper bounds to follow.

2.1 Asymptotic behavior with respect to a continuous weight, upper and lower bounds

2.1.1 History and results

The question of possible growth of polynomials orthogonal with respect to a continuous weight on the real line was first investigated by S. Bernstein in [7], where he originated (2.3). This work was taken up by Szegő in his book [69] where he translated Bernstein's results to the circle, proving

Theorem 2.1 (Szegő)**.** *Let $f(\theta)$ be a strictly positive weight function on the unit circle which satisfies the Lipschitz-Dini condition*

$$|f(\theta+\delta) - f(\theta)| < L|\log\delta|^{-1-\lambda}$$

where L and λ are fixed positive numbers. Let $\mathcal{D}$ the Szegő function defined in (1.5)*. Then we have, for $|z| = 1$,*

$$\phi_n^*(z) = z^n\mathcal{D}(z)^{-1} + \epsilon_n(z) \tag{2.4}$$

where

$$|\epsilon_n(x)| < C(\log n)^{-\lambda}$$

$C > 0$ *depends on* L, λ, $\|f\|_\infty$ *and* $\|f^{-1}\|_\infty$.

In the early 1990s Ambroladze also investigated orthogonal polynomials with respect to continuous weights, proving a negative result. [1] showed that, given sufficiently slow decay of the modulus of continuity to 0, there would be continuous weight functions whose associated polynomial sequence was unbounded at the point $z = 1$. More precisely, he proved:

Theorem 2.2 (Ambroladze)**.** *For any* $L > 0$ *and* $-1 < \lambda < 0$ *there exists a weight* $h(\theta) \in B_{L,\lambda}$ *such that*

$$\limsup_{n\to\infty} |\phi_n(1; hd\theta)| = \infty$$

for

$$B_{L,\lambda} := \{f \in C(\mathbb{T}) : f(\theta) > 0, |f(\theta+\delta) - f(\theta)| < L|\ln\delta|^{-(1+\lambda)}, 0 < \delta < 1\}$$

Given Szegő's asymptotic formula (2.4), the result of Ambroladze identified the roughly minimal continuous weight functions whose polynomials could grow, with a gap at the point $\lambda = 0$.

We consider the rate of polynomial growth possible for a given continuous function. Define

$$h_f(\delta) := \sup_{|x-y|<\delta} |f(x) - f(y)|$$

as the modulus of continuity of the function f. We assume $f \in C(\mathbb{T})$. Since trigonometric polynomials are dense in $C(\mathbb{T})$, we should be able to apply classical perturbative techniques in this setup. Indeed, they yield interesting results sharp in some regimes.

Theorem 2.3. *Let* f *be a real-valued function in* $C(\mathbb{T})$ *such that* $A \geq f \geq \delta$ *and* $\int_{\mathbb{T}} f(\theta)\frac{d\theta}{2\pi} = 1$, ϕ_n *the* n^{th} *orthonormal polynomial with respect to the measure* $d\mu := f(\theta)\frac{d\theta}{2\pi}$. *For* $n \geq N_0(f)$, *the polynomials admit the following bound:*

$$||\phi_n||_{p_0} \leq C(A,\delta)$$

for $p_0 = O\left(C(A,\delta)h_f\left(\frac{1}{n}\right)^{-1}\right)$ *and*

$$\|\phi_n\|_{L^\infty(\mathbb{T})} \leq C(A,\delta)n^{C(A,\delta)h_f\left(\frac{1}{n}\right)}$$

Remark. The condition that f is bounded below by δ is restrictive, but necessary. In particular, if we allow smooth f to go down and touch zero we may obtain growth of the associated polynomials. For example (see [65], example 1.6.4), for $d\mu = (1-\cos(\theta))\frac{d\theta}{2\pi}$, we have

$$\phi_n(z;d\mu) = \sqrt{\frac{2}{(n+1)(n+2)}}\sum_{j=0}^{n}(j+1)z^j$$

So

$$|\phi_n(1,d\mu)| = n\sqrt{\frac{n+2}{2(n+1)}}$$

If the modulus of continuity is sufficiently irregular, we will construct a weight which yields a subsequence of polynomials saturating the upper bound above.

Theorem 2.4. *Let* h_0 *be a modulus of continuity on* $\mathbb{T}$*; in particular,* h_0 *is subadditive, uniformly continuous, and* $h_0(0) = 0$*. If there is* $x_0 > 0$ *so that for* $x \leq x_0$*,*

$$|h_0(x)| \geq C\sqrt{\frac{|\ln|\ln(x)||}{|\ln x|}} \tag{2.5}$$

then there is a weight $w(\theta)$ *and a subsequence* $\{n_k\}$*,* $n_k \to \infty$ *such that*

$$h_w(\delta) \leq h_0(\delta)$$

and

$$|\phi_{n_k}(1,w)| \geq c_1 n_k^{C(A,\delta)h_0\left(\frac{1}{n_k}\right)}$$

Remark. This condition can be weakened to a similar condition on the limsup of the ratio in (2.5). We state the result in this way for simplicity.

Remark. This $C(A, \delta)$ is not necessarily the same as that in Theorem 2.3.

Notice that the dependence on the modulus of continuity is the same in both Theorems 2.3 and 2.4. This is the sense in which these results are sharp.

One application of the upper bounds above is L^p-asymptotics of the orthogonal polynomials for *all* $p < \infty$. Further, due to decay of the exponent in n, some standard harmonic analysis allows us to provide explicit rates of convergence for asymptotic quantities such as the polynomial entropies. These have attracted serious interest in recent years (see [2, 3, 4, 6, 24]). We show

Corollary 2.5. *If $f \in C(\mathbb{T})$ with $\delta \leq f \leq A$ and $\int_{\mathbb{T}} f \frac{d\theta}{2\pi} = 1$, for ϕ_n the orthonormal polynomial of degree n with respect to f and $\mathcal{D}$ the Szegő function of the weight f, then*

$$\phi_n^* \to \mathcal{D}^{-1}$$

in $L^p(\mathbb{T})$ for all finite p. The rate of this convergence can be quantified

$$\phi_n^* = \mathcal{D}^{-1} + e(n)$$

where

$$\|e(n)\|_p \leq cph_f\left(\frac{1}{n}\right)(\|\phi_n\|_p + 1)$$

Since $\|\phi_n\|_p \leq C(A, \delta, p)$ for all finite p, this says

$$\|e(n)\|_p \leq C(A, \delta, p)h_f\left(\frac{1}{n}\right)$$

This allows us to immediately show

Corollary 2.6. *Under the same assumptions above, for the polynomial entropy defined as*

$$E(n,f) := \int_{\mathbb{T}} |\phi_n|^2 \log|\phi_n| f \frac{d\theta}{2\pi}$$

we have the limiting statement

$$E(n,f) = -\frac{1}{4\pi}\int_{\mathbb{T}} \log(f) d\theta + C(A,\delta) h_f\left(\frac{1}{n}\right)$$

Notice this implies that $E(n,f)$ obey a uniform bound in n. This was a new result when the upper bounds in this section were proved, but the more general results on BMO weights in the next section published in [25] also imply this fact.

2.1.2 Upper bounds

Before we prove the main results, we introduce some auxiliary Lemmas for Theorem 2.3.

Lemma 2.7. *For $f \in C(\mathbb{T})$ such that $\delta \leq f \leq A$ and $\int_{\mathbb{T}} f(\theta)\frac{d\theta}{2\pi} = 1$, there is a trigonometric polynomial g of degree $n > n_0(f)$ such that*

$$|g^{-1} - f| \leq C(A,\delta) h_f\left(\frac{1}{n}\right), \quad \frac{\delta}{2} \leq g^{-1} \leq 2A$$

Moreover, we may take $\int_{\mathbb{T}} g^{-1}\frac{d\theta}{2\pi} = 1$.

Proof. This is a classical result of Jackson and can be found in [38]. We prove it for completeness.

We claim that it suffices to prove this result for Lipschitz functions. That is, assume

Proposition 2.1.1. *If f is a Lipschitz function with period 2π and Lipschitz constant M, then for each n there exists a trigonometric polynomial T of degree at most n such that*

$$\|f - T\|_\infty \leq \frac{cM}{n}$$

$$\int_{\mathbb{T}} T(\theta)\frac{d\theta}{2\pi} = 1$$

where c is an absolute constant.

Consider the function $f^{-1}(\theta) \in C(\mathbb{T})$. Let γ be the Lipschitz continuous function on the circle which takes on the value $f(\theta_n)^{-1}$ for $\theta_n \in \{0, \frac{2\pi}{n}, \dots, \frac{2\pi(n-1)}{n}\}$ and linearly interpolates between these values. Since f is bounded above and below, the modulus of continuity of f^{-1} is comparable to that of f with constants in terms of A and δ. So there is a constant $C(A,\delta)$ so that the slopes of the function γ never exceed $nC(A,\delta)h_f\left(\frac{1}{n}\right)$. Applying Proposition 2.1.1 to this function yields a trigonometric polynomial T so that

$$\|T-\gamma\|_\infty \le C(A,\delta)h_f\left(\frac{1}{n}\right) \tag{2.6}$$

and we claim $\|\gamma - f^{-1}\|_\infty \lesssim h_f\left(\frac{1}{n}\right)$. Let $x \in \mathbb{T}$. Let a_j be the closest point in $\{0, \frac{2\pi}{n}, \dots, \frac{2\pi(n-1)}{n}\}$ to x; note $|x-a_j| \le \frac{2\pi}{n}$. Therefore

$$|\gamma(x)-f^{-1}(x)| \le |\gamma(x)-\gamma(a_j)|+|f^{-1}(a_j)-f^{-1}(x)| \le C(A,\delta)h_f\left(\frac{1}{n}\right)+h_f\left(\frac{2\pi}{n}\right) \le C(A,\delta)h_f\left(\frac{1}{n}\right)$$

yields the Lemma.

We proceed to prove the proposition. The strategy is to convolve with the Jackson kernel and estimate the error. Let

$$\mathcal{J}_n(x) = c_n\left(\frac{\sin(nx/4)}{\sin(x/4)}\right)^4, \quad \int_{\mathbb{T}} \mathcal{J}_n(x) = 1$$

be the Jackson kernel, with c_n normalization constant. Claim that $c_n \sim n^{-3}$, and $|\mathcal{J}_n * f - f| \lesssim \frac{M}{n}$. We will show both by a simple integral estimate. For $j = 0,1,2$ we estimate

$$\int_0^{2\pi} t^j \frac{\sin^4(nx/4)}{\sin^4(x/4)}dx$$

Since for $t \in [0, \pi/2]$, $\sin(t) \sim t$, we have

$$\int_0^{2\pi} \left| x^j \frac{\sin^4(nx/4)}{\sin^4(x/4)} \right| dx \sim \int_0^{2\pi} \left| x^{j-4} \sin^4(nx/4) \right| dx \leq n^{3-j} \int_0^{\infty} \left| \xi^{j-4} \sin^4(\xi) \right| d\xi = cn^{3-j}$$

Therefore $c_n \sim n^{-3}$, and

$$|\mathcal{J}_n * f(x) - f(x)| = \left| \int_{\mathbb{T}} (f(x+t) - f(x)) \mathcal{J}_n(t) dt \right| \lesssim M \int_0^{2\pi} t \mathcal{J}_n(t) dt$$
$$\sim \frac{M}{n^3} \int_{\mathbb{T}} \frac{t \sin^4(nt/4)}{\sin^4 t/4} dt \sim \frac{M}{n}$$

Notice $\mathcal{J}_n \geq 0$ so g is real and nonnegative.

Since $|g^{-1} - f| \leq C(A, \delta) h_f \left(\frac{1}{n}\right)$,

$$\left| \int_{\mathbb{T}} g^{-1} \frac{d\theta}{2\pi} - \int_{\mathbb{T}} f \frac{d\theta}{2\pi} \right| \leq C(A, \delta) h_f \left(\frac{1}{n} \right)$$

So replacing g^{-1} with $\psi = \frac{g^{-1}}{\int_{\mathbb{T}} g^{-1} \frac{d\theta}{2\pi}}$ yields

$$|\psi - f| \leq |\psi - g^{-1}| + |g^{-1} - f| \leq |g^{-1}| \left| \frac{1}{1 - C(A, \delta) h_f \left(\frac{1}{n}\right)} - 1 \right| + C(A, \delta) h_f \left(\frac{1}{n} \right) \leq C(A, \delta) h_f \left(\frac{1}{n} \right)$$

□

Lemma 2.8. *Let $g, f, n = n_0(f)$ as in Lemma 2.7, $\mathcal{K}_{n-1}$ be the operator defined by*

$$\mathcal{K}_{n-1}[f](z) = \int_{\mathbb{T}} \left(\sum_{k=0}^{n-1} \widetilde{\phi}_k(z) \overline{\widetilde{\phi}_k(\zeta)} \right) f(\zeta) \frac{d\theta}{2\pi} \tag{2.7}$$

for $\{\widetilde{\phi}_k\}$ the orthogonal polynomials with respect to $g^{-1}(\theta) \frac{d\theta}{2\pi}$.

Then we have the operator bound

$$||\mathcal{K}_{n-1}||_{p,p} \leq C(A, \delta) p$$

Proof. This is an immediate consequence of the Christoffel-Darboux formula. Recall (see [65], Theorem 2.2.7)

$$\sum_{k=0}^{n-1} \widetilde{\phi}_k(e^{i\eta})\overline{\widetilde{\phi}_k(e^{i\theta})} = \frac{\overline{\widetilde{\phi}_n^*(e^{i\theta})}\widetilde{\phi}_n^*(e^{i\eta}) - \overline{\widetilde{\phi}_n(e^{i\theta})}\widetilde{\phi}_n(e^{i\eta})}{1 - e^{i\theta}e^{i\eta}}$$

Therefore

$$\mathcal{K}_{n-1}[f](e^{i\eta}) = \int_{\mathbb{T}} \sum_{k=0}^{n-1} \widetilde{\phi}_k(e^{i\eta})\overline{\widetilde{\phi}_k(e^{i\theta})} f(e^{i\theta}) \frac{d\theta}{2\pi} =$$

$$\int_{\mathbb{T}} \frac{\overline{\widetilde{\phi}_n^*(e^{i\theta})}\widetilde{\phi}_n^*(e^{i\eta}) f(e^{i\theta})}{1 - e^{i(\eta-\theta)}} \frac{d\theta}{2\pi} - \int_{\mathbb{T}} \frac{\overline{\widetilde{\phi}_n(e^{i\theta})}\widetilde{\phi}_n(e^{i\eta}) f(e^{i\theta})}{1 - e^{i(\eta-\theta)}} \frac{d\theta}{2\pi}$$

Notice $\frac{1}{1^{i\theta}}$ is a Calderón-Zygmund kernel, so

$$||\mathcal{K}_{n-1} f||_p \lesssim p||\widetilde{\phi}_n||_\infty^2 ||f||_p$$

We chose g to be a degree n trigonometric polynomial, bounded away from 0 by $\frac{1}{2A}$ and from ∞ by $2\delta^{-1}$ uniformly for $n > n_0(f)$. By Fejér-Riesz there is a polynomial h of degree n (that is, having *only* positive powers of z) with all its zeroes in $\mathbb{D}$ so that

$$g(\theta) = |h(\theta)|^2$$

And we have $c_1\sqrt{A^{-1}} \leq |h(\theta)| \leq c_2\sqrt{\delta^{-1}}$.

Since $g^{-1}(\theta)\frac{d\theta}{2\pi} = |h(\theta)|^{-2}\frac{d\theta}{2\pi}$ and h is a polynomial of degree n with all its zeroes in $\mathbb{D}$, $\widetilde{\phi}_n(z) = h(z)$. This is a consequence of the Bernstein-Szegő approximation, Theorem 1.7.8 in [65], and our choice of h as permissible OPUC via Fejér-Riesz.

So we have

$$||\mathcal{K}_{n-1}[f]||_p \leq pC(A, \delta)||f||_p$$

as desired. □

Given Lemmas 2.7 and 2.8, we may prove Theorem 2.3.

Theorem 2.3. Take f as in the statement of the Theorem, g as in Lemma 2.7. Use $\{\widetilde{\phi}_k\}$ to denote the polynomials orthogonal with respect to $d\nu := g^{-1}\frac{d\theta}{2\pi}$, and $\{\phi_k\}$ to denote the polynomials orthogonal with respect to $d\mu := f\frac{d\theta}{2\pi}$. Using $d\mu$ as $d\mu_1$ and $d\nu$ as $d\mu_2$ in our equation (2.3) yields

$$\phi_n(e^{i\eta}) = a_n\widetilde{\phi}_n(e^{i\eta}) + \mathcal{K}_{n-1}[\phi_n](e^{i\eta}) = a_n\widetilde{\phi}_n(e^{i\eta}) + \int_{\mathbb{T}} \phi_n(e^{i\theta})K_{n-1}(e^{i\eta}, e^{i\theta}, d\nu)g^{-1}(e^{i\theta})\frac{d\theta}{2\pi}$$

$$= a_n\widetilde{\phi}_n(e^{i\eta}) + \int_{\mathbb{T}} (g^{-1}(e^{i\theta}) - f(e^{i\theta}))K_{n-1}(e^{i\eta}, e^{i\theta}, d\nu)\phi_n(e^{i\theta})\frac{d\theta}{2\pi} \tag{2.8}$$

So

$$||\phi_n||_p \leq |a_n|||\widetilde{\phi}_n||_p + ||g^{-1} - f||_\infty||\mathcal{K}_{n-1}||_{p,p}||\phi_n||_p$$

By Lemma 2.7, $||g^{-1} - f||_\infty \leq C(A, \delta)h_f\left(\frac{1}{n}\right)$. By Lemma 2.8, $||\mathcal{K}_{n-1}||_{p,p} \leq pC(A, \delta)$. So

$$||\phi_n||_p \leq |a_n|||\widetilde{\phi}_n||_p + pC(A, \delta)h_f\left(\frac{1}{n}\right)||\phi_n||_p$$

$\widetilde{\phi}_n = h$ implies $||\widetilde{\phi}_n||_p \leq C(A, \delta)$, so we have

$$||\phi_n||_p \leq C(A, \delta) + pC(A, \delta)h_f\left(\frac{1}{n}\right)||\phi_n||_p$$

Take p so that

$$h_f\left(\frac{1}{n}\right) = \frac{1}{Cp}C(A, \delta)$$

for appropriate universal constant C. Since f is uniformly continuous, $p = p(n) \to \infty$ as $n \to \infty$. We have then, for this choice of p,

$$||\phi_n||_p \leq C(A, \delta)$$

Applying Nikolskii's inequality (2.1), we have control of the L^∞ norm as well:

$$||\phi_n||_\infty \leq C(A,\delta) n^{C(A,\delta) h_f\left(\frac{1}{n}\right)}$$

as desired.

□

2.1.3 Lower bounds

The lower bound in Theorem 2.4 is a consequence of

Lemma 2.9. *For $\epsilon > 0$ sufficiently small and any $n \geq n_0(\epsilon) = \exp\left(\frac{C|\ln \epsilon|}{\epsilon^2}\right)$, there is $g \in C^\infty(\mathbb{T})$ and absolute constant c so that*

- $\phi_n(1; g) \gtrsim n^{c\epsilon}$
- $\|1 - g\|_\infty \leq \epsilon$
- $\|g'\|_\infty \leq n\epsilon$

We prove this Lemma in Appendix B. It is simply the quantification of ϵ in terms of n in Theorem 3.1 of [21].

Proof (Theorem 2.4). We will assume Lemma 2.9.

We may assume the modulus of continuity h_0 is concave. Indeed, for any modulus of continuity h, there is a concave modulus $\tilde{h}$ satisfying $\tilde{h} \leq 2h$ (see [27], Proposition 3.15).

We construct a sequence of weights w_i with parameters n_i, ϵ_i so that the moduli of continuity of all weights lie below that of f. We will put these weights together into a single weight w, using the localization principle B.4 to argue that this does not change the size of the polynomials too much, and estimate the modulus of continuity h_w. First we choose n_i so that

1.
$$\sum_i \frac{3}{n_i} \leq 1 \tag{2.9}$$

2.
$$c_1 i^{-2} \geq n_i^{-c_2 h_0\left(\frac{1}{n_i}\right)} \tag{2.10}$$

for sufficiently small constants c_1, c_2.

The first condition will be used in the construction to estimate h_w. The second we use to argue that localization will not change the value of the polynomials too much.

Now let

$$\epsilon_i = \frac{1}{2} h_0\left(\frac{1}{n_i}\right) \tag{2.11}$$

We choose this value for ϵ_i so that the point $\left(\frac{1}{n_i}, 2\epsilon_i\right)$ lies on the curve $h_0(x)$. By definition of ϵ_i and since h_0 is concave, $2\epsilon_i n_i \delta \leq h_0(\delta)$ for $\delta \leq \frac{1}{n_i}$. To see this, note $y = 2\epsilon_i n_i x$ is the equation of the line passing through $(0,0)$ and $\left(\frac{1}{n_i}, 2\epsilon_i\right)$, two points on the curve $h_0(x)$. Since h_0 is monotonic increasing, $h_0(\delta) \geq 2\epsilon_i$ for $\delta \geq \frac{1}{n_i}$. Putting these estimates together gives

$$2\min\{\epsilon_i n_i \delta, \epsilon_i\} \leq h_0(\delta) \tag{2.12}$$

We also see

$$n_i^{c\epsilon_i} = n_i^{c h_0\left(\frac{1}{n_i}\right)} \tag{2.13}$$

directly by definition.

By the assumption (2.5), we claim

$$n_i \geq \exp\left(\frac{C|\ln \epsilon_i|}{\epsilon_i^2}\right) \tag{2.14}$$

This can be seen as follows. Recall (2.5), which tells us

$$h_0(x) \geq C\sqrt{\frac{|\ln|\ln(x)||}{|\ln x|}}$$

Since $\epsilon_i = ch_0\left(\frac{1}{n_i}\right)$, we have

$$\epsilon_i \geq C\sqrt{\frac{|\ln|\ln\left(\frac{1}{n_i}\right)||}{|\ln\left(\frac{1}{n_i}\right)|}} = C\sqrt{\frac{|\ln|\ln(n_i)||}{|\ln(n_i)|}}$$

Let $y = \ln n_i$. Then

$$\epsilon_i \geq C\sqrt{\frac{\ln(y)}{y}}$$

Thus

$$\frac{y}{\ln y} \geq \frac{C}{\epsilon_i^2} \tag{2.15}$$

For sufficiently small $\epsilon_i > 0$, (2.15) shows we may restrict to $y \leq \epsilon_i^{-3}$ and still solve the inequality. Therefore

$$y \geq \frac{C\ln y}{\epsilon_i^2} \geq \frac{C|\ln \epsilon_i|}{\epsilon_i^2} \tag{2.16}$$

which, when exponentiated, yields (2.14).

Therefore by Lemma 2.9 for each i we may construct a weight $w_i \in C^\infty(\mathbb{T})$ which satisfies

1.

$$\phi_{n_i}(1; w_i) \gtrsim n_i^{c\epsilon_i} = n_i^{ch_0\left(\frac{1}{n_i}\right)}, \quad c \text{ universal}$$

2.

$$\|1 - w_i\|_\infty \leq \epsilon_i$$

3.

$$\|w'\|_\infty \leq \epsilon_i n_i$$

Items (2) and (3) tell us about the modulus of continuity of the w_i; in particular,

$$h_{w_i}(x) \leq \min\{\epsilon_i n_i x, \epsilon_i\} \tag{2.17}$$

(2.13) and (2.17) say that

$$h_{w_i}(x) \leq \min\{\epsilon_i n_i x, \epsilon_i\} \leq \frac{1}{2} h_0(x) \tag{2.18}$$

for all x and i.

We now put these weights w_i together into one weight w whose polynomials grow along subsequence $\{n_i\}$ at rate which meets our upper bound.

Divide the circle $\mathbb{T}$ into intervals $\{I_i\}$, each of size $|I_i| = ci^{-2}$ for small constant c. Choose c sufficiently small so that $\sum_i ci^{-2} + \sum_i 3n_i^{-1} \leq 2\pi$. Notice it is here we use our condition (2.9). Spread the I_i around the circle so that I_i is separated from I_{i-1} by an interval of length at least

$$\frac{2}{n_{i-1}} + \frac{1}{n_i} \tag{2.19}$$

Denote by θ_i the center of I_i. Let

$$w(\theta) = w_i(\theta - \theta_i) : \theta \in I_i$$

Let b_1, b_2 denote the endpoints of I_i for fixed i. Linearly interpolate at slope $\pm\epsilon_i n_i$ between $(b_j, w_i(b_j))$ for $j \in \{1, 2\}$ and points $(a_j, 1)$ for $a_j \in \mathbb{T}$ (where the positive slope is chosen if $w_i(b_j) < 1$ and the negative if $w_i(b_j) < 1$). This yields an expanded interval L_i for which the w on these L_i is still at worst Lipschitz with Lipschitz constant $\epsilon_i n_i$ and bounded above by ϵ_i. Since $\|1 - w_i\|_\infty \leq \epsilon_i$, $|L_i| \leq \delta_i + \frac{2}{n_i}$.

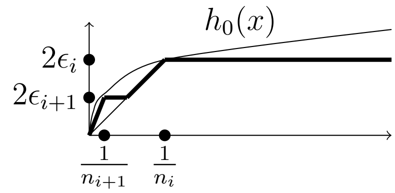

Figure 1: Bounding modulus of continuity of w

Connect the rest of the circle with the constant value 1 and denote by M_{j-1} the interval (on which w is constant 1) between L_{j-1} and L_j. Notice by (2.19) that

$$|M_{j-1}| \geq \frac{1}{n_{j-1}} \tag{2.20}$$

and $\mathbb{T} = \cup_j (M_j \cup L_j)$. Let the intervals L_j be closed and M_j be open.

Then we make two claims. First

$$\|\phi_{n_i}(z; w)\|_\infty \gtrsim n_i^{ch_0\left(\frac{1}{n_i}\right)}$$

and second

$$h_w(x) \leq 2 \sup_i \min\{\epsilon_i n_i x, \epsilon_i\} \leq h_0(x)$$

The first claim follows from the localization principle; B.4 says

$$|\phi_{n_i}(\theta_i; w)| \geq C|I_i||\phi_{n_i}(1; w_i)| \geq C i^{-2} n_i^{c\epsilon_i} \gtrsim n_i^{ch_0\left(\frac{1}{n_i}\right)}$$

by the construction of the parameters ϵ_i and n_i, in particular equations (2.10) and (2.13). We investigate the second claim (see Figure 1). Note the inequality

$$2 \sup_i \min\{\epsilon_i n_i x, \epsilon_i\} \leq h_0(x)$$

is implied by (2.18), so we focus on the first inequality. We wish to estimate $|w(x) - w(y)|$. There are three cases:

- If $x, y \in L_j$ for some j, then since w is Lipschitz with constant $\epsilon_j n_j$ and bounded by ϵ_j we are done.

- If neither is in any L_j, then they are in some M_r and some M_s, on which $w \equiv 1$, so $|w(x) - w(y)| = 0$.

- If one is in L_j and the other is elsewhere, we make use of our separation condition (2.20) on L_j. We will investigate the possibilities separately. Without loss of generality we assume $x \in L_j, y \in M_r \cup L_r$.

  If $r < j - 1$ or $y \in L_{j-1}$,

  $$|w(x) - w(y)| \leq 2\epsilon_r = 2 \min \{\epsilon_r n_r |x - y|, \epsilon_r\}$$

  since by our separation condition $|x - y| > \frac{1}{n_r}$.

  If $y \in M_{j-1} \cup M_j$, then $w(y) = 1$ so

  $$|w(x) - w(y)| = |w(x) - 1| \leq \epsilon_j$$

  If $r \geq j + 1$ then uniformly in r,

  $$|w(x) - w(y)| \leq 2 \min \{\epsilon_j n_j |x - y|, \epsilon_j\}$$

  since $|x - y| \geq \frac{1}{n_j}$.

So we have

$$h_w(x) \leq 2 \sup_i \min\{\epsilon_i n_i x, \epsilon_i\} \leq h_0(x)$$

as well as

$$\|\phi_{n_i}(z; w)\|_\infty \gtrsim n_i^{c h_0\left(\frac{1}{n_i}\right)}$$

and hence the Theorem is proved.

□

Notice that although we have only shown an upper bound for the modulus of continuity h_w, Theorem 2.3 shows that $h_w(x)$ achieves this bound as $x \to 0$ up to a multiplicative factor.

2.1.4 Applications

Our upper bounds in this section utilize methods reminiscent of Szegő's book [69] (see especially chapter XI) originally due to Bernstein. With this in mind, we may hope to prove similar convergence results to the previously mentioned

Theorem 2.10 (Szegő). *Let $f(\theta)$ be a strictly positive weight function on the unit circle which satisfies the Lipschitz-Dini condition*

$$|f(\theta+\delta) - f(\theta)| < L|\log \delta|^{-1-\lambda}$$

where L and λ are fixed positive numbers, and $\mathcal{D}$ the Szegő function defined in (1.5). *Then we have, for $|z| = 1$,*

$$\phi_n^*(z) = z^n \mathcal{D}(z)^{-1} + \epsilon_n(z)$$

where

$$|\epsilon_n(x)| < C(\log n)^{-\lambda}$$

$C > 0$ depends on L, λ, and the minimum and maximum of f.

This is Theorem 12.1.3 in [69]. Of course if we relax the modulus of continuity condition to allow for something worse than Dini-Lipschitz we will not have such a strong asymptotic; the polynomials may be unbounded at a point. But because our regime is so classically perturbative there is hope to adapt the methods of Szegő and similarly extract rates of convergence in a weaker sense. This is the content of Corollaries 2.5 and 2.6.

Proof (Corollary 2.5) . In the proof of Theorem 2.3 we have seen the expansion (2.8)

$$\phi_n(e^{i\eta}) = a_n\widetilde{\phi}_n(e^{i\eta}) + \int_{\mathbb{T}} K_{n-1}(e^{i\eta}, e^{i\theta}, d\nu)g^{-1}\frac{d\theta}{2\pi} = a_n h + \int_{\mathbb{T}} K_{n-1}(e^{i\eta}, e^{i\theta}, d\nu)(g^{-1} - f)\phi_n\frac{d\theta}{2\pi}$$

where h is the Fejér-Riesz factorization of $g \sim \mathcal{J}_n * f^{-1}$ and $\widetilde{\phi}_n$ is orthogonal with respect to the measure $|h|^{-2}$.

We wish to estimate $\phi_n^*(z) - \mathcal{D}^{-1}$. We write

$$\|\phi_n^* - \mathcal{D}^{-1}\|_p \leq \|a_n h^* - h^*\|_p + \|h^* - \mathcal{D}^{-1}\|_p + \left\| \int_{\mathbb{T}} K_{n-1}(z,\zeta,d\nu)(g^{-1} - f)\phi_n^* \right\|_p := I_1 + I_2 + I_3$$

I_3 we have already estimated; we have

$$I_3 = \left\| \int_{\mathbb{T}} K_{n-1}(z,\zeta,d\nu)(g^{-1} - f)\phi_n^* d\theta(\zeta) \right\|_p \lesssim pC(A,\delta)h_f\left(\frac{1}{n}\right)\|\phi_n\|_p$$

which is sufficient for our purposes.

We estimate the first term.

A priori, $|a_n|$ may depend on δ. It is in fact uniformly close to 1. Recall $\widetilde{\phi}_n = \frac{\widetilde{\Phi}_n}{\|\widetilde{\Phi}_n\|_{L^2(g^{-1})}}$ and $\phi_n = \frac{\Phi_n}{\|\Phi_n\|_{L^2(f)}}$. Since the only z^n term in an expansion with respect to the basis $\{\widetilde{\phi}_0, \ldots, \widetilde{\phi}_n\}$ is in the polynomial $\widetilde{\phi}_n$,

$$a_n^{-1} = \frac{\|\Phi_n\|_{L^2(f)}}{\|\widetilde{\Phi}_n\|_{L^2(g^{-1})}} = \frac{\left\|\widetilde{\Phi}_n + \int_{\mathbb{T}} K_{n-1}\Phi_n(f - g^{-1})d\theta\right\|_{L^2(f)}}{\|\widetilde{\Phi}_n\|_{L^2(g^{-1})}} = \frac{\|\widetilde{\Phi}_n\|_{L^2(f)}}{\|\widetilde{\Phi}_n\|_{L^2(g^{-1})}} + O\left(C(A,\delta)h_f\left(\frac{1}{n}\right)\right)$$

$$= 1 + O\left(C(A,\delta)h_f\left(\frac{1}{n}\right)\right) \tag{2.21}$$

Therefore

$$I_1 = O\left(C(A,\delta)h_f\left(\frac{1}{n}\right)\right)$$

We turn to estimating the second term. Notice that $(h^*)^{-1}$ is the Szegő function with respect to the weight g^{-1}, and $\mathcal{D}$ the Szegő function with respect to f. Therefore (see [65] Section 2.4) we have the integral representations

$$h^*(z) = \exp\left(\frac{-1}{4\pi}\int_{\mathbb{T}}\frac{e^{i\theta}+z}{e^{i\theta}-z}\log\left(g^{-1}(\theta)\right)\frac{d\theta}{2\pi}\right)$$

$$\mathcal{D}^{-1}(z) = \exp\left(\frac{-1}{4\pi}\int_{\mathbb{T}}\frac{e^{i\theta}+z}{e^{i\theta}-z}\log\left(f(\theta)\right)\frac{d\theta}{2\pi}\right)$$

and we are investigating

$$I_2 = \left\|\exp\left(\frac{-1}{4\pi}\int_{\mathbb{T}}\frac{e^{i\theta}+z}{e^{i\theta}-z}\log\left(g^{-1}(\theta)\right)\frac{d\theta}{2\pi}\right) - \exp\left(\frac{-1}{4\pi}\int_{\mathbb{T}}\frac{e^{i\theta}+z}{e^{i\theta}-z}\log\left(f(\theta)\right)\frac{d\theta}{2\pi}\right)\right\|_p$$

$$\lesssim \left\|\exp\left(\frac{-1}{4\pi}\int_{\mathbb{T}}\frac{e^{i\theta}+z}{e^{i\theta}-z}\left(\log\left(g^{-1}(\theta)\right)-\log\left(f(\theta)\right)\right)\frac{d\theta}{2\pi}\right)-1\right\|_p$$

We expand the exponential in its series and see

$$I_1 \lesssim \sum_{j=1}^{\infty}\frac{\left\|\left(\frac{-1}{4\pi}\int_{\mathbb{T}}\frac{e^{i\theta}+z}{e^{i\theta}-z}\left(\log\left(g^{-1}(\theta)\right)-\log\left(f(\theta)\right)\right)\frac{d\theta}{2\pi}\right)^j\right\|_p}{j!}$$

Since g^{-1}, f, g and f^{-1} are bounded uniformly in θ, we may apply the Mean Value Theorem to say

$$\left|\log\left(g^{-1}(\theta)\right)-\log\left(f(\theta)\right)\right| \le C(A,\delta)\left|g^{-1}(\theta)-f(\theta)\right| \le C(A,\delta)h_f\left(\frac{1}{n}\right)$$

The Cauchy kernel is a Calderón-Zygmund kernel so L^p bounded for all finite p with norm growing like p at infinity. So

$$I_1 \lesssim \sum_{j=1}^{\infty}\frac{\left\|\frac{1}{4\pi}\int_{\mathbb{T}}\frac{e^{i\theta}+z}{e^{i\theta}-z}\left(\log\left(g^{-1}(\theta)\right)-\log\left(f(\theta)\right)\right)\frac{d\theta}{2\pi}\right\|_{jp}^j}{j!} \le \sum_{j\ge 1}\frac{\left(jpC(A,\delta)h_f\left(\frac{1}{n}\right)\right)^j}{j!}$$

$j! \lesssim (cj)^j$ by Stirling's formula so

$$I_1 \lesssim \sum_{j\geq 1}\left(pC(A,\delta)h_f\left(\frac{1}{n}\right)\right)^j = pC(A,\delta)h_f\left(\frac{1}{n}\right)\frac{1}{1-pC(A,\delta)h_f\left(\frac{1}{n}\right)} = pC(A,\delta)h_f\left(\frac{1}{n}\right), \quad n \gg 1$$

Therefore

$$\|\phi_n^* - \mathcal{D}^{-1}\|_p \leq C(A,\delta)ph_f\left(\frac{1}{n}\right)(\|\phi_n\|_p + 1)$$

□

The second corollary is now immediate.

Proof (Corollary 2.6). Following [25], we have the following inequality from the mean value formula

$$|x^2\log x - y^2\log y| \leq C(1 + x|\log x| + y|\log y|)|x-y|$$

Therefore

$$\int_{\mathbb{T}} ||\phi_n|^2\log|\phi_n| - |\mathcal{D}|^{-2}\log|\mathcal{D}|^{-1}|f\frac{d\theta}{2\pi} \lesssim \int_{\mathbb{T}}(1+|\phi_n\log\phi_n|+|\mathcal{D}^{-1}\log\mathcal{D}^{-1}|)||\phi_n^*|-|\mathcal{D}^{-1}||f\frac{d\theta}{2\pi}$$

Since $x|\log x| \leq C(\delta)(1+x^{1+\delta})$, applying the generalized Hölder inequality to $|\phi_n|^{1+\delta}$, $|\mathcal{D}^{-1-\delta}|$, $|\phi_n^*| - |\mathcal{D}^{-1}|$, and f and using $f \in L^\infty$ gives us

$$\int_{\mathbb{T}} ||\phi_n|^2\log|\phi_n| - |\mathcal{D}|^{-2}\log|\mathcal{D}|^{-1}|f\frac{d\theta}{2\pi} \leq C(A)h_f\left(\frac{1}{n}\right)$$

To conclude the proof it suffices to note

$$\int_{\mathbb{T}}|\mathcal{D}|^{-2}\log|\mathcal{D}|^{-1}f\,d\theta = -\frac{1}{2}\int_{\mathbb{T}}\log(f)\frac{d\theta}{2\pi}$$

since $|\mathcal{D}|^2 = f$. □

It would be nice to provide rates of convergence to these asymptotic quantities in other situations. There is a clear breaking point in this approach: we are in a truly classical perturbative regime. That is, our weight function f is uniformly well-approximated by trigonometric polynomials. We expand with respect to these polynomials, show the "tail" is small and investigate the convergence of this "other" set of polynomials to the Szegő function of the weight f in n. If we cannot answer this secondary convergence question (for example, in the next section), this classical approach is out of luck. For this reason it seems that the method above is fairly limited when "weight perturbation" is interpreted in a different sense.

2.2 Asymptotic behavior with respect to a BMO($\mathbb{T}$) weight, upper bounds

The question of what can be said about the polynomials orthogonal with respect to a BMO weight was inspired by the gap which remained after the works [2] and [21]. These showed that L^∞ control over the weight w was sufficient to improve the $\frac{1}{2}$ exponent on n in growth of uniform norm of the polynomials, but L^p control for finite p was not. A natural question, then, was whether any class which sits between L^∞ and L^p on the circle would be sufficient to break $n^{1/2}$ growth.

BMO in particular is a natural choice considering its appearance in place of L^∞ in a range of contexts. Two in particular are the H^1-BMO duality proved by Fefferman and Stein [29] and the estimates on commutators of BMO functions with singular integrals proved by Coifman, Rochberg and Weiss [16]. See the Appendix for a brief introduction to BMO functions following [55].

In this context the author and his advisor began investigating the case of $w \in \mathrm{BMO}(\mathbb{T}), w \geq \delta$. Interestingly, perturbative results were quickly proved which did not rely on $w \geq \delta$ but rather on the weaker condition $w^{-1} \in \mathrm{BMO}(\mathbb{T})$. This arose due to a natural symmetry in the algebraic structure of the problem. The results on large-BMO weights proved more involved.

In the paper [25] we proved the following:

Theorem 2.11 (Denisov-R. [25]). *If w is a real-valued weight satisfying $w : \|w^{-1}\|_{BMO} \leqslant s, \|w\|_{BMO} \leqslant t$ for $st \gg 1$, then there is $\Pi \in L^{p_0}[-\pi, \pi], p_0 > 2$ such that*

$$\lim_{n\to\infty} \|\Phi_n^* - \Pi\|_{p_0} = 0$$

and we have for p_0:

$$p_0 = 2 + \frac{C_1}{(st)\log^2(st)} \tag{2.22}$$

If w is a weight with $w : \|w^{-1}\|_{BMO} \leqslant s, \|w\|_{BMO} \leqslant t$ for $st \ll 1$, then there is $\Pi \in L^{p_0}[-\pi, \pi], p_0 > 2$ such that

$$\lim_{n\to\infty} \|\Phi_n^* - \Pi\|_{p_0} = 0$$

and here we have for p_0:

$$\frac{C_2}{(st)^{1/4}} \tag{2.23}$$

We also have the bound for the uniform norm

$$\|\Phi_n^*\|_\infty \leqslant C_{(st)} n^{1/p_0} \tag{2.24}$$

where $C_{(u)}$ denotes a function of u.

In the case when an additional information is given, e.g., $w \in L^\infty$ or $w^{-1} \in L^\infty$, this result can be improved.

Theorem 2.12 (Denisov-R. [25]). *Under the conditions of the previous Theorem, we have*

- *If $w \geqslant 1$, then p_0 can be taken as*

$$p_0 = \begin{cases} 2 + \dfrac{C_1}{t \log t}, & \text{if } t \gg 1 \\ \dfrac{C_2}{\sqrt{t}}, & \text{if } 0 < t \ll 1 \end{cases}$$

- *If $w \leqslant 1$, then we have*

$$p_0 = \begin{cases} 2 + \dfrac{C_1}{s \log s}, & \text{if } s \gg 1 \\ \dfrac{C_2}{\sqrt{s}}, & \text{if } 0 < s \ll 1 \end{cases}$$

We also have the bound for the uniform norm

$$\|\Phi_n^*\|_\infty \leqslant C_{(t,s)} n^{1/p_0} \tag{2.25}$$

where $C_{(t,s)}$ depends on t or s.

Remark. It is clear that the allowed exponent p_0 is decaying in s and t so it can be chosen larger than 2 for all values of s and t.

Remark. The following scaling invariance holds: $\Phi_n(z, \sigma) = \Phi_n(z, \alpha\sigma)$, $\alpha > 0$. The BMO norm is 1-homogeneous, e.g., $\|\alpha w\|_{BMO} = \alpha \|w\|_{BMO}$, so the estimates in the Theorems are invariant under scaling $w \to \alpha w$, as they should be.

Remark. The weight above could be taken to be complex-valued with some restriction on the argument, and our analysis below will show existence of the polynomials for all n as well as the bounds in the Theorem above. This work is ongoing.

In the case when $w = C$, we get $\|w\|_{BMO} = \|w^{-1}\|_{BMO} = 0$ and, although $\Phi_n^*(z,w) = 1$, we can not say anything about the size of $\phi_n^*(z,w)$. The next Lemma explains how our results can be extended to $\{\phi_n^*\}$.

Lemma 2.13. *In the Theorem 2.11, if one makes an additional assumption that $\|w\|_1 = 1$, then $\|\phi_n^* - \mathcal{D}^{-1}\|_{p_0} \to 0$ with p_0 as above.*

Proof. Indeed, Lemma A.6 from Appendix A shows that

$$\int_{-\pi}^{\pi} \log w \frac{d\theta}{2\pi} > -\infty$$

and thus the sequence $\{\|\Phi_n\|_{2,w}\}$ has a finite positive limit [65]. Therefore, $\{\phi_n^*\} = \left\{\frac{\Phi_n^*}{\|\Phi_n\|_{2,w}}\right\}$ has an L^{p_0} limit by Theorem 1.3. By Theorem 1.4, $\{\phi_n^*\}$ converges weakly to $\mathcal{D}^{-1}$ and therefore we have the statement of the Lemma. ☐

Recall the polynomial entropy as defined in the previous section,

$$E(n,\sigma) = \int_{\mathbb{T}} |\phi_n|^2 \log|\phi_n| d\sigma$$

where $\{\phi_n\}$ are orthonormal with respect to σ. In [24], the sharp lower and upper bounds were obtained for σ in the Szegő class. In [2], it was shown that $E(n,w)$ can not exceed $C \log n$ if $w \geqslant 1$ and $w \in L^p[-\pi,\pi], p < \infty$, and that this bound saturates. This leaves us with the very natural question: what are regularity assumptions on w that guarantee boundedness of $E(n,w)$? The following corollary of Lemma 2.13 gives the partial answer.

Corollary 2.14 (Denisov-R. [25])**.** *If $w : w, w^{-1} \in \mathrm{BMO}(\mathbb{T})$ and $\|w\|_1 = 1$, then*

$$\lim_{n\to\infty} E(n,w) = -\frac{1}{4\pi}\int_{-\pi}^{\pi} \log(w) d\theta$$

So far, the only classes in which the $E(n,w)$ was known to be bounded were the Baxter's class [65]: $d\sigma = w\frac{d\theta}{2\pi}$, $w \in W(\mathbb{T}), w > 0$ (for $W(\mathbb{T})$ the Wiener algebra) or the class given by positive weights with a certain modulus of continuity [69]. The previous section of this thesis showed that in fact all polynomials orthogonal with respect to continuous weight bounded away from 0 obeyed uniform bounds on their entropies. In this section our conditions are obviously much weaker and, in a sense, sharp.

We use the following notation in this section:

$P_{[i,j]}$ denotes the $L^2\left(\frac{d\theta}{2\pi}\right)$ projection to the $(i,\ldots,j)$ Fourier modes.

Given two operators A and B, we write $[A,B] = AB - BA$ for their commutator. If w is a function, then in the expression like $[w,A]$, the symbol w is identified with the operator of multiplication by w. The Hunt-Muckenhoupt-Wheeden characteristic of the weight $w \in A_p$ will be denoted by $[w]_{A_p}$. For the basic facts about the BMO class, A_p and their relationship, we refer the reader to Appendix A or the classical text [66].

2.2.1 Proofs of main results

Before proving the main result, Theorem 2.11, we need some auxiliary Lemmas. We start with the following observation which goes back to S. Bernstein [7, 69]. Notice this is the same fact used in the last section, here with $d\mu_2 = \frac{d\theta}{2\pi}$. It is so useful in this section we restate it in an alternate form.

Lemma 2.15. *For a monic polynomial Q of degree n, we have:*

$$Q(z) = \Phi_n(z,w) \quad \textit{if and only if} \quad P_{[0,n-1]}(wQ) = 0. \tag{2.26}$$

Proof. It is sufficient to notice that (2.26) is equivalent to

$$\int_{-\pi}^{\pi} Q(e^{i\theta})e^{-ij\theta}w(\theta)\frac{d\theta}{2\pi} = 0, \quad j = 0, \dots, n-1$$

which is the orthogonality condition. □

Lemma 2.16. *If $f \in L^2(\mathbb{T})$ is real-valued function, $Q \in L^\infty(\mathbb{T})$, then*

$$z^n \overline{P_{[0,n-1]}(fz^n\overline{Q})} = P_{[1,n]}(fQ)$$

In particular, for a polynomial P of degree at most n with $P(0) = 1$, we have:

$$P(z) = \Phi_n^*(z, w) \quad \textit{if and only if} \quad P_{[1,n]}(wP) = 0.$$

Proof. The first statement is immediate. The second one follows from the Lemma above and the formula $\Phi_n = z^n\overline{\Phi_n^*}$, $z \in \mathbb{T}$. □

We have the following three identities for $\Phi_n^*(z, w)$; the first one was used in [21] recently. They are immediately implied by the Lemma above.

$$\Phi_n^* = 1 + P_{[1,n]}\Big((1 - \alpha w)\Phi_n^*\Big), \quad \alpha \in \mathbb{R} \tag{2.27}$$

$$\Phi_n^* = 1 + w^{-1}[w, P_{[1,n]}]\Phi_n^* \tag{2.28}$$

$$\Phi_n^* = 1 - [w^{-1}, P_{[1,n]}](w\Phi_n^*) \tag{2.29}$$

Denote the higher order commutators recursively:

$$\mathbf{C}_0 = P_{[1,n]}, \quad \mathbf{C}_1 = [w, P_{[1,n]}], \quad \mathbf{C}_l = [w, \mathbf{C}_{l-1}], \quad l = 2, 3, \dots$$

Define the multiple commutators of w^{-1} and $P_{[1,n]}$ (in that order!) by $\widetilde{\mathbf{C}}_j$.

Lemma 2.17. *The following representations hold*

$$w^j P_{[1,n]}\Phi_n^* = \sum_{l=1}^{j} \binom{j-1}{l-1} \mathbf{C}_l w^{j-l}\Phi_n^* \tag{2.30}$$

and

$$w^{-j} P_{[1,n]}\Phi_n^* = -\sum_{l=0}^{j} \binom{j}{l} \widetilde{\mathbf{C}}_{l+1} w^{-(j-l)}(w\Phi_n^*) \tag{2.31}$$

where $j = 1, 2, \ldots$.

Moreover, if f is a polynomial of degree at most n which satisfies (2.30), (2.31) *with constant term 1, then* $f = \Phi_n^*$

Proof. We will prove (2.30), the other formula can be obtained in the similar way. The case $j = 1$ of this expression is our formula $wP_{[1,n]}\Phi_n^* = [w, P_{[1,n]}]\Phi_n^*$, which follows directly from Lemma 2.16. Now the proof proceeds by induction. Suppose we have

$$w^{k-1} P_{[1,n]}\Phi_n^* = \sum_{l=1}^{k-1} \binom{k-2}{l-1} \mathbf{C}_l w^{k-1-l}\Phi_n^*$$

Multiply both sides by w and write

$$w^k P_{[1,n]}\Phi_n^* = \sum_{l=1}^{k-1} \binom{k-2}{l-1} w\mathbf{C}_l w^{k-1-l}\Phi_n^* = \sum_{l=1}^{k-1} \binom{k-2}{l-1} \left(\mathbf{C}_{l+1} w^{k-1-l}\Phi_n^* + \mathbf{C}_l w^{k-l}\Phi_n^*\right) =$$

$$\sum_{l=1}^{k-1} \binom{k-2}{l-1} \mathbf{C}_l w^{k-l}\Phi_n^* + \sum_{l=2}^{k} \binom{k-2}{l-2} \mathbf{C}_l w^{k-l}\Phi_n^* = \sum_{l=1}^{k} \binom{k-1}{l-1} \mathbf{C}_l w^{k-l}\Phi_n^*$$

because

$$\binom{k-1}{l-1} = \binom{k-2}{l-2} + \binom{k-2}{l-1}$$

.

The uniqueness statement is proved by the same induction. Lemma 2.26 shows that uniqueness holds for $j = 1$, and the inductive step above completes the proof. □

Motivated by the previous Lemma, we introduce certain operators. Given $f \in L^p$, define $\{y_j\}$ recursively by

$$y_0 = f,\, y_j = w^j + \sum_{l=0}^{j-1} \binom{j-1}{l} \mathbf{C}_{l+1} y_{j-1-l}$$

Then, we let

$$z_j = w^{-j} - \sum_{l=0}^{j} \binom{j}{l} \widetilde{\mathbf{C}}_{l+1} z_{j-l-1}$$

where $z_{-1} = y_1, z_0 = y_0 = f$. Notice that for fixed j both y_j and z_j are affine linear transformations in f. We can write

$$y_j = y_j' + y_j''$$

where

$$y_0' = f, \quad y_0'' = 0$$

and, recursively,

$$y_j' = \sum_{l=0}^{j-1} \binom{j-1}{l} \mathbf{C}_{l+1} y_{j-1-l}', \quad y_j'' = w^j + \sum_{l=0}^{j-1} \binom{j-1}{l} \mathbf{C}_{l+1} y_{j-1-l}''$$

Similarly, we write $z_j = z_j' + z_j''$ where

$$z_{-1}' = y_1', \quad z_{-1}'' = y_1'', \quad z_0' = f, \quad z_0'' = 0$$

and

$$z_j' = -\sum_{l=0}^{j} \binom{j}{l} \widetilde{\mathbf{C}}_{l+1} z_{j-l-1}', \quad z_j'' = w^{-j} - \sum_{l=0}^{j} \binom{j}{l} \widetilde{\mathbf{C}}_{l+1} z_{j-l-1}'',$$

Let us introduce linear operators: $B_j f = y_j'$, $D_j f = z_j'$. We need an important Lemma.

Lemma 2.18.

$$w^j \Phi_n^* = y_j'' + B_j \Phi_n^*, \quad w^{-j} \Phi_n^* = z_j'' + D_j \Phi_n^*$$

Proof. This follows from

$$w^j\Phi_n^* = w^j + w^j P_{[1,n]}\Phi_n^*, \quad w^{-j}\Phi_n^* = w^{-j} + w^{-j}P_{[1,n]}\Phi_n^*$$

and the previous Lemma. □

The next Lemma, in particular, provides the bounds for B_j and D_j.

Lemma 2.19. *Assume* $w \geqslant 0, \|w\|_{BMO} = t, \|w^{-1}\|_{BMO} = s,\ \|w\|_1 = 1,$ *and* $p \in [2,3]$. *Then,*

$$\|B_j\|_{p,p} \leqslant (Ctj)^j, \quad \|D_j\|_{p,p} \leqslant (1+st)(Csj)^j$$

Moreover,

$$\|y_j''\|_p \leqslant (C\widetilde{t}j)^j, \quad \|z_j''\|_p \leqslant \widetilde{s}\widetilde{t}(C\widetilde{s}j)^j$$

with

$$\widetilde{t} = \max\{t,1\}, \quad \widetilde{s} = \max\{s,1\}$$

Proof. We will prove the estimates for $\|B_j\|_{p,p}$ and $\|y_j''\|_p$ only, the bounds for $\|D_j\|_{p,p}, \|z_j''\|_p$ are shown similarly. By John-Nirenberg inequality (Theorem A.2, see also [66], p.144), we get

$$\int_{-\pi}^{\pi} |w-(2\pi)^{-1}|^{jp}d\theta \lesssim j\int_0^\infty x^{jp-1}\exp(-Cx/t)dx = j(Ct)^{jp}\Gamma(jp) \leqslant (C_1tj)^{pj}$$

where Stirling's formula was used for the gamma function Γ.

Since

$$|w|^{jp} \leqslant (|w-(2\pi)^{-1}| + (2\pi)^{-1})^{jp} \leqslant C^{jp}(|w-(2\pi)^{-1}|^{jp} + 1)$$

we have

$$\int_{-\pi}^{\pi} |w|^{jp}d\theta \leqslant C^{jp}(1+(tj)^{jp}) \leqslant (C_1\widetilde{t}j)^{jp}$$

Lemma A.5 yields

$$\|y_j'\|_p \leqslant \sum_{l=0}^{j-1} \frac{(j-1)!}{l!(j-1-l)!}(\widetilde{C}(l+1)t)^{l+1}\|y_{j-1-l}'\|_p \leqslant (Ct)^j j! \sum_{k=0}^{j-1} \frac{(Ct)^{-k}}{k!}\|y_k'\|_p$$

Divide both sides by $(Ct)^j j!$ and denote $\beta_j = \dfrac{\|y_j'\|_p}{(Ct)^j j!}$. Then,

$$\beta_j \leqslant \sum_{l=0}^{j-1} \beta_l$$

Since $\beta_0 = \|f\|_p$, we have $\beta_j \leqslant 3^j\|f\|_p$ by induction and thus $\|y_j'\|_p \leqslant (Ctj)^j\|f\|_p$. The estimates for $\|y_j''\|_p, \|z_j'\|_p, \|z_j''\|_p$ can be obtained similarly. □

Lemma 2.20. *If* $\|w\|_1 = 1, \|w\|_{BMO} = t, \|w^{-1}\|_{BMO} = s$, *and* $p \in [2,3]$, *then*

$$\min_{l\in\mathbb{N}}\Big(\Lambda^{-l}\|B_l\|_{p,p}\Big) \leqslant \exp\left(-\frac{C\Lambda}{t}\right)$$

and

$$\min_{j\in\mathbb{N}}\Big(\epsilon^j\|D_j\|_{p,p}\Big) \leq (1+st)\exp\left(-\frac{C}{\epsilon s}\right)$$

provided that $\Lambda \gg t$ *and* $\epsilon \ll s^{-1}$.

Proof. By the previous Lemma, we have

$$\Big(\Lambda^{-l}\|B_l\|_{p,p}\Big) \leqslant \left(\frac{Ctl}{\Lambda}\right)^l$$

Optimizing in l we get $l^* \sim C\Lambda/(te)$ and it gives the first estimate. The proof for the second one is identical. □

Now we are ready to prove the main results of the section.

Proof. (Theorem 2.11). Notice first that (2.24) follows from the Nikolskii inequality (2.1) as long as the L^{p_0} norms are estimated.

By scaling invariance, we can assume that $\|w\|_1 = 1$. We consider two cases separately: $st \gg 1$ and $st \ll 1$. The proofs will be different.

1. The case $st \gg 1$.

Let $p = 2 + \delta$ with $\delta < 1$. Take two n-independent parameters ϵ and Λ such that $\epsilon s \ll 1$ and $\Lambda t^{-1} \gg 1$. Consider the following sets $\Omega_1 = \{x : w \leqslant \epsilon\}$, $\Omega_2 = \{x : \epsilon < w < \Lambda\}, \Omega_3 = \{x : w \geqslant \Lambda\}$. Notice that

$$\epsilon s \ll 1, t\Lambda^{-1} \ll 1 \implies (\epsilon s)(t\Lambda^{-1}) \ll 1 \implies \epsilon\Lambda^{-1} \ll (st)^{-1} \ll 1 \implies \epsilon \ll \Lambda$$

From (2.27), we have

$$\Phi_n^* = 1 + P_{[1,n]}(1 - w/\Lambda)\Phi_n^*$$

The idea of our proof is to rewrite this identity in the form

$$\Phi_n^* = f_n + \mathcal{O}(n)\Phi_n^*$$

where $\|f_n\|_p < C(s,t)$ and $\mathcal{O}(n)$ is a contraction in L^p for the suitable choice of p.

Towards showing this contraction, we consider operators

$$\mathcal{O}_1(n)f = \epsilon^j P_{[1,n]}(1 - w/\Lambda)\chi_{\Omega_1}\left(\frac{w}{\epsilon}\right)^j D_j f \tag{2.32}$$

$$\mathcal{O}_2(n)f = P_{[1,n]}(1 - w/\Lambda)\chi_{\Omega_2} f \tag{2.33}$$

$$\mathcal{O}_3(n)f = \Lambda^{-l} P_{[1,n]}(1 - w/\Lambda)(\Lambda/w)^l \chi_{\Omega_3} B_l f \tag{2.34}$$

where j and l will be fixed later. They will be n-independent. Let us estimate the (L^p, L^p) norms of these operators. Since $\|P_{[1,n]}\|_{p,p} \leqslant 1 + C\delta$ (Lemma A.4), we choose j

and l as in Lemma 2.20 to ensure

$$\|\mathcal{O}_1(n)\|_{p,p} \leqslant st\exp\left(-\frac{\widehat{C}}{\epsilon s}\right)$$
$$\|\mathcal{O}_2(n)\|_{p,p} \leqslant (1+C\delta)(1-\epsilon\Lambda^{-1})$$
$$\|\mathcal{O}_3(n)\|_{p,p} \leqslant \exp\left(-\frac{\widehat{C}\Lambda}{t}\right)$$

Lemma 2.18 now yields

$$\Phi_n^* = 1 + f_1(n) + f_3(n) + (\mathcal{O}_1(n) + \mathcal{O}_2(n) + \mathcal{O}_3(n))\Phi_n^*$$

where

$$f_1(n) = \epsilon^j P_{[1,n]}(1-w/\Lambda)\chi_{\Omega_1}\left(\frac{w}{\epsilon}\right)^j z_j'', \quad f_3(n) = \Lambda^{-l}P_{[1,n]}(1-w/\Lambda)(\Lambda/w)^l\chi_{\Omega_3}y_l''$$

$$\text{Let} \quad f(n) = 1 + f_1(n) + f_3(n)$$

Then Lemma 2.19 provides the bound

$$\|f(n)\|_p \leqslant C(s,t) \tag{2.35}$$

uniform in n. Denote $\mathcal{O}(n) = \mathcal{O}_1(n) + \mathcal{O}_2(n) + \mathcal{O}_3(n)$ and select parameters $\epsilon, \Lambda, \delta$ such that $\|\mathcal{O}(n)\|_{p,p} < 1 - C\delta$. To do so, we first let $\delta = c\epsilon\Lambda^{-1}$ with small positive absolute constant c. Then, we consider

$$st\exp\left(-\frac{\widehat{C}}{\epsilon s}\right) + \exp\left(-\frac{\widehat{C}\Lambda}{t}\right) = \frac{c_1\epsilon}{\Lambda}$$

with c_1 again being a small constant. Now, solving equations

$$st\exp(-\widehat{C}/(\epsilon s)) = \exp(-\widehat{C}\Lambda/t), \quad c_1\epsilon/\Lambda = 2\exp(-\widehat{C}\Lambda/t)$$

we get the statement of the Theorem. Indeed, we have two equations:

$$\epsilon = \frac{\widehat{C}t}{s(\widehat{C}\Lambda + t\log(st))}$$

and

$$\frac{\Lambda}{t} = \frac{1}{\widehat{C}}\left(C + \log(s\Lambda) + \log\Big(\frac{\Lambda}{t} + \frac{\log(st)}{\widehat{C}}\Big)\right)$$

Denote

$$u = \widehat{C}\Lambda/t$$

and then

$$u = C + \log(st) + 2\log u + \log\Big(1 + \frac{\log(st)}{u}\Big)$$

To find the required root, we restrict the range of u to $c_1 \log(st) < u < c_2 \log(st)$ for $c_1 \ll 1$, $c_2 \gg 1$. Rewrite the equation above as

$$u - 2\log u - \log\left(1 + \frac{\log(st)}{u}\right) = \log(st) + C$$

Differentiating the left hand side in u, we see that $\text{l.h.s.}' \sim 1$ within the given range. Therefore, there is exactly one solution u and $u \sim \log(st)$. Then, since

$$\log\left(1 + \frac{\log(st)}{u}\right)$$

is $O(1)$, we get

$$u = \log(st) + 2\log u + O(1) = \log(st) + 2\log\log(st) + O(1)$$

by iteration. Thus,

$$\frac{\epsilon}{\Lambda} = Ce^{-u} \sim \frac{1}{st\log^2(st)}$$

and $\delta \sim \frac{1}{st\log^2(st)}$. Now that we proved $\|\mathcal{O}(n)\|_{p,p} \leqslant 1 - C\delta < 1$, we can rewrite

$$\Phi_n^* = f(n) + \sum_{j=1}^{\infty} \mathcal{O}^j(n) f(n)$$

and the series converges geometrically in L^p with tail being uniformly small in n due to (2.35).

To show that Φ_n^* converges in L^p as $n \to \infty$, it is sufficient to prove that $\mathcal{O}^j(n) f(n)$ converges for each j. This, however, is immediate. Indeed,

$$P_{[1,n]} f \to P_{[1,\infty]} f, \quad \text{as} \quad n \to \infty$$

in L^q for all $f \in L^q, 1 < q < \infty$. Since $w, w^{-1} \in \text{BMO} \subset \cap_{p \geqslant 1} L^p$ ([66], this again follows from the John-Nirenberg estimate), we see that multiplication by $w^{\pm j}$ maps L^{p_1} to L^{p_2} continuously by Hölder's inequality provided that $p_2 < p_1$ and $j \in \mathbb{Z}$. Therefore, if $\mu_j \in L^\infty, j = 1, \ldots, k$, then

$$\mu_1 w^{\pm j_1} P_{[1,n]} \mu_2 w^{\pm j_2} \ldots \mu_{k-1} w^{\pm j_{k-1}} P_{[1,n]} \mu_k w^{\pm j_k} \tag{2.36}$$

has a limit in each $L^p, p < \infty$ when $n \to \infty$. Notice that each $f(n)$ and $\mathcal{O}^j(n) f(n)$ can be written as a linear combination of expressions of type (2.36) ($\{\mu_j\}$ taken as the characteristic functions). Now that δ is chosen, we define p_0 in the statement of the Theorem as $p_0 = 2 + \delta$.

2. **The case $st \ll 1$.**

The proof in this case is much easier. Let us start with two identities

$$\Phi_n^* = 1 + w^{-1}[w, P_{[1,n]}]\Phi_n^*, \quad \Phi_n^* = 1 + [P_{[1,n]}, w^{-1}] w \Phi_n^*$$

which can be recast as

$$w\Phi_n^* = w + [w, P_{[1,n]}]\Phi_n^*, \quad \Phi_n^* = 1 + [P_{[1,n]}, w^{-1}] w \Phi_n^*$$

Substitution of the first formula into the second one gives

$$\Phi_n^* = 1 + [P_{[1,n]}, w^{-1}] w + G_n \Phi_n^*$$

where

$$G_n = [P_{[1,n]}, w^{-1}][w, P_{[1,n]}]$$

We have

$$\|1 + [P_{[1,n]}, w^{-1}]w\|_p \leqslant C(s,t,p)$$

and

$$\|G_n\|_{p,p} \lesssim stp^4$$

by Lemma A.7. Taking $p < p_0 \sim (st)^{-1/4}$ we have that G_n is a contraction. Convergence of all terms in the geometric series can be proved as before. $\square$

Let us give a sketch of how the arguments can be modified to prove Theorem 2.12.

Proof. (Theorem 2.12). Consider the case $w \geqslant 1$ first.

1. The case $t \gg 1$.

The proof is identical except that we can chose $\epsilon = 1/2$ so that $\Omega_1 = \emptyset$. We get an equation for Λ

$$\frac{C}{\Lambda} = \exp\left(-\frac{\widehat{C}\Lambda}{t}\right), \quad \Lambda = \widehat{C}^{-1} t(\log\Lambda - \log C)$$

Denote $\widehat{C}\Lambda/t = u$, then

$$u = \log t + \log u + O(1),\, u = \log t + \log\log t + O(1)$$

and $\delta \sim (t\log t)^{-1}$.

2. The case $t \ll 1$.

We have

$$\Phi_n^* = 1 + L_n\Phi_n^*, \quad L_n f = w^{-1}[w, P_{[1,n]}]f$$

and Lemma A.7 yields

$$\|L_n\|_{p,p} \lesssim p^2 t < 0.5$$

for $p < p_0 = O(t^{-1/2})$.

The case $w \leqslant 1$ can be handled similarly. When s is large, we take $\Lambda = 1$ in the proof of the previous Theorem and get an equation for ϵ:

$$C\epsilon = s \exp\left(-\frac{\widehat{C}}{\epsilon s}\right)$$

Its solution for large s gives the required asymptotics for ϵ and, correspondingly, for δ and p_0. If s is small, it is enough to consider the equation

$$\Phi_n^* = 1 - [w^{-1}, P_{[1,n]}]w\Phi_n^*$$

where the operator $[w^{-1}, P_{[1,n]}]w$ is contraction in L^{p_0} for the specified p_0. $\square$

Now we are ready to prove Corollary 2.14.

Proof. (of Corollary 2.14). The following inequality follows from the Mean Value Formula

$$|x^2 \log x - y^2 \log y| \leqslant C(1 + x|\log x| + y|\log y|)|x - y|, \quad x, y \geqslant 0$$

Since $w \in \cap_{p<\infty} L^p$, the Theorem 2.11 yields

$$\int_{-\pi}^{\pi} ||\phi_n|^2 \log|\phi_n| - |\mathcal{D}|^{-2} \log\left(|\mathcal{D}|^{-1}\right)| w\frac{d\theta}{2\pi} \lesssim \int_{-\pi}^{\pi} (1 + |\phi_n \log \phi_n| + |\mathcal{D}^{-1} \log \mathcal{D}|)||\phi_n^*| - |\mathcal{D}^{-1}|| w\frac{d\theta}{2\pi} \to 0,$$

$$n \to \infty$$

by applying the trivial bound: $u|\log u| \leqslant C(\delta)(1 + u^{1+\delta}), \delta > 0$ and the generalized Hölder's inequality to $|\phi_n|^{1+\delta}$(or $|\mathcal{D}|^{-1-\delta}$), $||\phi_n^*| - |\mathcal{D}|^{-1}|$, and w. To conclude the proof, it is sufficient to notice that

$$\int_{-\pi}^{\pi} |\mathcal{D}|^{-2} \log |\mathcal{D}^{-1}| w \frac{d\theta}{2\pi} = -\frac{1}{4\pi} \int_{-\pi}^{\pi} \log(w) d\theta$$

because $|\mathcal{D}|^2 = w$. $\square$

Chapter 3

Randomized Verblunsky Parameters–Steklov was almost surely right

At this point we have seen several senses in which Steklov was wrong, including continuous weights staying bounded away from 0 whose polynomials diverge at a point. But the Steklov conjecture remained open for over 50 years and survived attacks by talented mathematicians; empirically, the counterexamples seem to be sparse.

We are led to a probabilistic Steklov question: if we take a random measure, what behavior can we expect from the uniform norm?

The natural way to place a probability distribution on measures of orthogonality is via the Szegő recurrence (1.2); we randomize the Verblunsky parameters. Recall that in the analogy between orthogonal polynomials and Fourier analysis, the Verblunsky parameters $\{\alpha_n\}$ play the role of the Fourier coefficients. Asking about the regularity of polynomials with random Verblunsky parameters can be thought of as a nonlinear analogue of the convergence of Fourier series with random coefficients. So the randomized Steklov question is analogous to those addressed by Salem and Zygmund in their classical paper [64]. We expect the same techniques to come into play and similar results to be provable. The nonlinear case should be more difficult than the linear case, and we might need to enforce more conditions to guarantee the same outcome. In the main result of this section, we find a logarithm more than Salem and Zygmund results in an almost sure bound on the maximum of the polynomials.

Let $\{a_n\}$ be a fixed (positive) sequence, $|a_n| < 1$, with decay to be specified later. Let $\{\omega_n\}$ be an independent sequence of random variables, uniform on the interval $[0, 1)$.

Let $\alpha_n = a_n e^{2\pi i \omega_n}$, so that $\{\alpha_n\}_{n=0}^{\infty}$ is a sequence of independent random variables with the same decay properties as $\{|a_n|\}$. The particular distribution of α_n plays no role as long as it is bounded by $|a_n|$ and its argument is symmetrically distributed. Denote by dS the measure thus defined on $\mathbb{D}^\infty$.

Our filtration on $(\mathbb{D}^\infty, dS)$ is the natural one; let

$$\mathcal{G}_{n-1} = \sigma(\alpha_0, \dots, \alpha_{n-1})$$

where $\sigma(\alpha_0, \dots, \alpha_{n-1})$ denotes the σ-algebra on $\mathbb{D}^\infty$ generated by the random variables $\{\alpha_0, \dots, \alpha_{n-1}\}$. That is, $\sigma(\alpha_0, \dots, \alpha_{n-1})$ is the smallest σ-algebra under which all α_i for $0 \le i \le n-1$ are measurable.

Since the $\{\alpha_j\}$ are independent random variables, we have

$$\alpha_j \perp \mathcal{G}_n \quad j > n$$

The Szegő recursion for the polynomials Φ_n^* says

$$\Phi_{n+1}^*(z) = \Phi_n^*(z) - \alpha_n z \Phi_n(z) \tag{3.1}$$

Therefore

$$\mathbb{E}[\Phi_j^*(z) | \mathcal{G}_n] = \Phi_n^*(z) \qquad j \ge n \tag{3.2}$$

The equation (3.2) endows the polynomial sequence $\{\Phi_n^*\}$ with martingale structure.

Definition 3.1. *A sequence of integrable random variables* $\{M_n\}$ *along with a filtration* $\{\mathcal{F}_n\}$ *on a probability space* Ω *which satisfies*

$$\mathbb{E}\left[M_n|\mathcal{F}_{n-1}\right] = M_{n-1} \tag{3.3}$$

is called a martingale.

Martingales are a classical and extraordinarily useful object in probability and related areas of analysis. An associated concept we will use is that of a so-called submartingale.

Definition 3.2. *A sequence of integrable random variables* $\{M_n\}$ *along with a filtration* $\{\mathcal{F}_n\}$ *on a probability space* Ω *which satisfies*

$$\mathbb{E}\left[M_n|\mathcal{F}_{n-1}\right] \geq M_{n-1}$$

is a submartingale.

This technology allows us to prove the main result of this section.

Theorem 3.3. *Let* $\{a_j\}$ *be fixed. Let*

$$R_n = \sum_{k=n}^{\infty} a_k^2$$

and assume

$$\sum_{n=1}^{\infty} \frac{\sqrt{R_n}}{n} < \infty$$

Let $\alpha_n = a_n e^{2\pi i \omega_n}$ *with* ω_n *i.i.d. uniform on* $[0,1)$. *Then with probability 1 there is a random constant* $C(\omega)$ *such that*

$$\sup_n \|\Phi_n^*\|_\infty \leq C(\omega)$$

Remark. One example of admissible R_n is

$$R_n \leq \log^{-\beta} n, \quad \beta > 2$$

Remark. The Salem-Zygmund paper [64] has slightly weaker conditions on the decay of R_n. Precisely, it proves

Theorem 3.4 (Salem-Zygmund [64])**.** *Let* $R_n = \sum_{m=n+1}^{\infty} r_m^2$. *If*

$$\sum_n \frac{\sqrt{R_n}}{n\sqrt{\log n}} < \infty$$

then the series

$$\sum_{m=1}^{\infty} r_m \phi_m(t) \cos mx$$

represents a continuous function for almost every value of t*, where* $\{\phi_n\}$ *is the Rademacher system.*

Our result is analogous to theirs, but our assumption on R_n misses theirs by a logarithm. The obvious methods of improving the approach in this section (e.g. further sparsification) will not yield this further logarithm, and it is unclear to the author whether this is a consequence of suboptimal technique or the structure of the problem. A reasonable lower bound on decay necessary to achieve a similar result would be of interest.

Remark. This Theorem asserts that if we impose logarithmically stronger than ℓ^2 decay on the Verblunsky parameters, our resulting measure, which was a priori guaranteed to be only Szegő, is almost surely in fact Steklov with polynomials obeying a uniform bound.

3.1 Basic Lemmas

We use the Markovian nature of this martingale to decouple in a sparse manner and control the process. In this section we collect the Lemmas which will be needed to prove Theorem 3.3 above. The polynomial recursion can be restated

$$\Phi_{n+1}^*(z) = \Phi_n^*(z)\left(1 - \frac{\alpha_n z \Phi_n}{\Phi_n^*}\right)$$

Therefore

$$\Phi_{n+1}^*(z) = \Phi_k^*(z)\prod_{j=k}^{n}\left(1 - \frac{\alpha_j z \Phi_j}{\Phi_j^*}\right)$$

We denote

$$M_{j\to k} \coloneqq \prod_{n=j}^{k}\left(1 - \frac{\alpha_n z \Phi_n}{\Phi_n^*}\right)$$

For now we consider $z \in \mathbb{T}$ a fixed parameter.

Notice that these $M_{j\to k}$ inherit the martingale structure of the polynomials. Indeed

$$\mathbb{E}\left[M_{j\to k+1}|\mathcal{G}_k\right] = M_{j\to k} - z\frac{\Phi_k}{\Phi_n^*}M_{j\to k}\mathbb{E}\left[\alpha_{k+1}\right] = M_{j\to k} \tag{3.4}$$

Since $\phi(x) = |x|$ is convex we immediately have that $|M_{j\to k}(z)|$ is a submartingale by the conditional Jensen inequality:

$$\mathbb{E}\left[|M_{j\to k+1}||\mathcal{G}_k\right] \geq |\mathbb{E}\left[M_{j\to k+1}|\mathcal{G}_k\right]| = |M_{j\to k}| \tag{3.5}$$

We will denote

$$M_j \coloneqq \prod_{n=G(j)+1}^{G(j+1)}\left(1 - \frac{\alpha_n z \Phi_n}{\Phi_n^*}\right)$$

for $G(j)$ some sparse sequence; later we will take $G(j) = 2^{2^j}$.

Lemma 3.5. *For $k|a_n| \leq c_0 \ll 1$,*

$$\mathbb{E}\left[\left|1 - \frac{\alpha_n z \Phi_n}{\Phi_n^*}\right|^k \middle| \mathcal{G}_{n-1}\right] \leq 1 + C(k|\alpha_n|)^2$$

Remark. Notice the left-hand side is a random variable but the right-hand side is not; this Lemma says the random variable on the left is $L^\infty(\Omega, \mathbb{P})$ with bound given by the quantity on the right.

Proof. We write

$$\left|1 - \frac{\alpha_n z \Phi_n}{\Phi_n^*}\right|^k = \left(\left|1 - \frac{\alpha_n z \Phi_n}{\Phi_n^*}\right|^2\right)^{k/2} = \left(1 + |\alpha_n|^2 - 2|\alpha_n| \cos\gamma_n\right)^{k/2}$$

where γ_n is uniformly distributed on the circle given $\frac{\Phi_n}{\Phi_n^*}$ and thus $\cos\gamma_n$ has conditional expectation 0 when conditioned on $\mathcal{G}_{n-1}$. Along with boundedness by $|a_n|$, this is the only condition we need out of the particular distribution of α_n. We estimate this quantity when conditioned

$$\mathbb{E}\left[(1 + |\alpha_n|^2 - 2|\alpha_n|\cos\gamma_n)^{k/2} | \mathcal{G}_{n-1}\right] = \mathbb{E}\left[\exp\left(\frac{k}{2}\ln(1 + |\alpha_n|^2 - 2|\alpha_n|\cos\gamma_n)\right) | \mathcal{G}_{n-1}\right]$$
$$\leq \mathbb{E}\left[\exp\left((k/2)(|\alpha_n|^2 - 2|\alpha_n|\cos\gamma_n)\right) | \mathcal{G}_{n-1}\right]$$

since $\log(1+u) \leq u$.

By our assumption that $k|a_n| \leq c_0$ for $c_0 \ll 1$, we may bound the series expansion for the exponential by a convergent geometric series to say

$$\mathbb{E}\left[(1 + |\alpha_n|^2 - 2|\alpha_n|\cos\gamma_n)^{k/2} | \mathcal{G}_{n-1}\right] \leq \mathbb{E}\left[1 + (k/2)|\alpha_n|^2 - 2|\alpha_n|\cos\gamma_n + O(|k\alpha_n|^2) | \mathcal{G}_{n-1}\right]$$

Because $\cos\gamma_n$ has conditional expectation 0,

$$\mathbb{E}\left[(1 + |\alpha_n|^2 - 2|\alpha_n|\cos\gamma_n)^{k/2} | \mathcal{G}_{n-1}\right] \leq \mathbb{E}\left[1 + C(k|\alpha_n|)^2\right] = 1 + C(k|\alpha_n|)^2$$

$\square$

We stitch this estimate together in a sparse manner.

Lemma 3.6. *If $k_j|a_n| \leq c_0$ for $c_0 \ll 1$ for all $n \in \{G(j)+1, \dots, G(j+1)\}$, then*

$$\mathbb{E}\left[|M_j|^{k_j}\right] = \mathbb{E}\left[\prod_{n=G(j)+1}^{G(j+1)} \left|1 - \alpha_n z \frac{\Phi_n}{\Phi_n^*}\right|^{k_j}\right] \leq \prod_{n=G(j)+1}^{G(j+1)} \left(1 + C(k_j|\alpha_n|)^2\right)$$

and

$$\mathbb{P}\left\{\|M_j\|_{L^{k_j}(\mathbb{T})} \geq \Lambda_j\right\} \leq \frac{\prod_{n=G(j)+1}^{G(j+1)} \left(1 + C(k_j|\alpha_n|)^2\right)}{\Lambda_j^{k_j}}$$

Proof. This follows by conditioning

$$\mathbb{E}\left[\prod_{n=G(j)+1}^{G(j+1)} \left|1 - \alpha_n z \frac{\Phi_n}{\Phi_n^*}\right|^k\right] = \mathbb{E}\left[\mathbb{E}\left[\prod_{n=G(j)+1}^{G(j+1)} \left|1 - \alpha_n z \frac{\Phi_n}{\Phi_n^*}\right|^k \middle| \mathcal{G}_{G(j+1)-1}\right]\right]$$

$$\mathbb{E}\left[\prod_{n=G(j)+1}^{G(j+1)-1} \left|1 - \alpha_n z \frac{\Phi_n}{\Phi_n^*}\right|^k \mathbb{E}\left[\left|1 - \alpha_{G(j+1)} z \frac{\Phi_{G(j+1)}}{\Phi_{G(j+1)}^*}\right|^k \middle| \mathcal{G}_{G(j+1)-1}\right]\right]$$

$$\leq \left(1 + c(k|\alpha_{G(j+1)}|)^2\right) \mathbb{E}\left[\prod_{n=G(j)+1}^{G(j+1)-1} \left|1 - \alpha_n z \frac{\Phi_n}{\Phi_n^*}\right|^k\right]$$

and iterating.

Markov's inequality provides the claimed bound on the probability.

□

This Lemma only controls the process at the lattice points $G(j)$. It is the martingale structure which allows us to extend these estimates to uniform bounds over n between these lattice points. In particular, we will use Doob's L^p maximum inequality. This is the first and only time the nature of the setup as martingale and not simply Markovian process contributes heavily.

Lemma 3.7 (Doob)**.** *If X_n is a nonnegative submartingale then for $1 < p < \infty$,*

$$\mathbb{E}\left[\left(\sup_{0 \leq m \leq n} X_m\right)^p\right] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}\left[X_n^p\right]$$

Proof. Our proof will follow that of Durrett, Theorem 5.4.3 [28]. Denote

$$\widetilde{X}_n = \sup_{0 \le m \le n} X_m$$

Let

$$\widetilde{X}_n \wedge M = \min(\widetilde{X}_n, M)$$

We have

$$\mathbb{E}\left[(\widetilde{X}_n \wedge M)^p\right] = \int_0^\infty p\lambda^{p-1}\mathbb{P}\left[\widetilde{X}_n \wedge M \ge \lambda\right] d\lambda$$

Claim

$$\mathbb{P}\left[\widetilde{X}_n \wedge M \ge \lambda\right] \le \lambda^{-1}\mathbb{E}\left[X_n 1_{(\widetilde{X}_n \wedge M \ge \lambda)}\right] \tag{3.6}$$

We may see this by defining a stopping time $N = \inf\{m : X_m \ge \lambda \text{ or } m = n\}$. Let $A = \{\widetilde{X}_n \ge \lambda\}$. Then since N is a stopping time which satisfies $\mathbb{P}[N \le n] = 1$, $X_n - X_{N \wedge n}$ is a submartingale and so

$$\mathbb{E}[X_N] \le \mathbb{E}[X_n] \tag{3.7}$$

Therefore Markov's inequality says

$$\lambda\mathbb{P}[A] \le \mathbb{E}[X_N 1_A]$$

and using (3.7) we have

$$\mathbb{E}[X_N] = \mathbb{E}[X_N 1_A] + \mathbb{E}[X_N 1_{A^c}] = \mathbb{E}[X_N 1_A] + \mathbb{E}[X_n 1_{A^c}] \le \mathbb{E}[X_N] \Rightarrow \mathbb{E}[X_N 1_A] \le \mathbb{E}[X_n 1_A]$$

Since $\{\widetilde{X}_n \wedge M \ge \lambda\}$ is always either $\{\widetilde{X}_n \ge \lambda\}$ or $\emptyset$, the inequality (3.6) is proved. And so we have

$$\mathbb{E}\left[(\widetilde{X}_n \wedge M)^p\right] \le \int_0^\infty p\lambda^{p-1}\left(\lambda^{-1}\int X_n 1_{(\widetilde{X}_n \wedge M \ge \lambda)} dP\right) d\lambda$$

Since everything in sight is nonnegative we may switch the order of integration

$$\mathbb{E}\left[\left(\widetilde{X}_n \wedge M\right)^p\right] \le \int X_n \int_0^{\widetilde{X}_n \wedge M} p\lambda^{p-2} d\lambda dP = \frac{p}{p-1} \int X_n (\widetilde{X}_n \wedge M)^{p-1} dP$$
$$\le \frac{p}{p-1} \left(\mathbb{E}\left[X_n^p\right]\right)^{1/p} \left(\mathbb{E}\left[(\widetilde{X}_n \wedge M)^p\right]\right)^{1/p'}$$

where p, p' are conjugate exponents, by Hölder's inequality.

Dividing both sides by $\left(\mathbb{E}\left[(\widetilde{X}_n \wedge M)^p\right]\right)^{1/p'}$, letting $M \to \infty$ and using monotone convergence gives the statement. □

3.2 Proof of Theorem 3.3

Proof. (Theorem 3.3)

We adopt our notation with $G(j) = 2^{2^j}$.

Let

$$T_j = \sum_{n=2^{2^j}}^{\infty} a_n^2 = R_{2^{2^j}}$$

Notice by the Cauchy condensation test, convergence of $\sum_n \frac{\sqrt{R_n}}{n}$ implies convergence of $\sum_k \sqrt{R_{2^k}}$ which in turn implies that of $\sum_j 2^j \sqrt{T_j}$, the condition we really use.

We first wish to bound, for appropriate k_j,

$$\prod_{n=2^{2^j}}^{2^{2^{j+1}}} \left(1 + C|\alpha_n k_j|^2\right) \le \exp\left(C \sum_{n=2^{2^j}}^{2^{2^{j+1}}} |\alpha_n k_j|^2\right) \le \exp\left(C k_j^2 T_j\right) \tag{3.8}$$

Let $k_j \sim T_j^{-1/2}$. Then $|a_n| k_j \le c_0$ uniformly in $n : 2^{2^j} \le n \le 2^{2^{j+1}}$ because T_j is at least the sum of the squares of $|a_n|$ in this range. Therefore we can apply Lemma 3.6, Fubini's Theorem and (3.8) to say

$$\mathbb{E}\left[\|M_j\|_{k_j}^{k_j}\right] = \int_{\mathbb{T}} \mathbb{E}\left[|M_j(z)|^{k_j}\right] \frac{d\theta}{2\pi} \le C$$

And so

$$\mathbb{P}\left[\|M_j\|_{k_j} \geq \Lambda_j\right] \leq \frac{C}{\Lambda_j^{k_j}}$$

Notice that we can strengthen this statement by Doob's maximal L^p inequality. That is, by Lemma 3.7 we have

$$\mathbb{E}\left[\max_{2^{2^j} \leq n \leq 2^{2^{j+1}}} |M_{2^{2^j} \to n}(z)|^{k_j}\right] \leq \left(\frac{k_j}{k_j - 1}\right)^{k_j} \mathbb{E}\left[|M_j(z)|^{k_j}\right]$$

and so

$$\mathbb{E}\left[\max_{2^{2^j} \leq n \leq 2^{2^{j+1}}} \|M_{2^{2^j} \to n}\|_{L^{k_j}(\mathbb{T})}^{k_j}\right] \leq \int_{\mathbb{T}} \mathbb{E}\left[\max_{2^{2^j} \leq n \leq 2^{2^{j+1}}} |M_{2^{2^j} \to n}(z)|^{k_j}\right] \leq \mathbb{E}\left[\|M_j\|_{L^{k_j}(\mathbb{T})}^{k_j}\right] \leq C$$

where the first inequality is simply the statement that the maximum of integrals is at most the integral of the maximum in conjunction with Fubini's Theorem.

Therefore we also have

$$\mathbb{P}\left[\max_{2^{2^j} \leq n \leq 2^{2^{j+1}}} \|M_{2^{2^j} \to n}\|_{L^{k_j}(\mathbb{T})} \geq \Lambda_j\right] \leq \frac{C}{\Lambda_j^{k_j}}$$

We wish to have $\prod_j \Lambda_j < \infty$ as well as $\sum_j \Lambda_j^{-k_j} < \infty$.

That is, we take $\Lambda_j = 1 + l_j$ and need $l_j \in \ell^1$. We will choose $l_j = j^{-2}$. Examine the series:

$$\sum_j \Lambda_j^{-k_j} = \sum_j (1 + l_j)^{-k_j} \leq \sum_j \exp\left(-c k_j l_j\right) = \sum_j \exp\left(-c\sqrt{\frac{l_j^2}{T_j}}\right)$$

where the inequality is by estimating the Taylor series of $\log(1 + l_j)$, using our choice of l_j.

Since as we have noted

$$\sum_j 2^j \sqrt{T_j} < \infty$$

we have $T_j \lesssim 2^{-2j}$, so

$$\sum_j \exp\left(-c\sqrt{\frac{l_j^2}{T_j}}\right) \leq C\sum_j \exp\left(-2^j l_j\right)$$

So since $l_j = j^{-2}$ we see

$$\sum_j \mathbb{P}\left[\max_{2^{2^j} \leq n \leq 2^{2^{j+1}}} \|M_{2^{2^j} \to n}\|_{L^{k_j}(\mathbb{T})} \geq \Lambda_j\right] < \infty$$

Clearly there is a lot of room in the choice of l_j. The requirement to sum up the probabilities while enforcing

$$\prod_j \Lambda_j < \infty$$

is not our true obstruction.

Since the probabilities of the events sum up, the Borel-Cantelli Lemma implies that $\left\{\max_{2^{2^j} \leq n \leq 2^{2^{j+1}}} \|M_{2^{2^j} \to n}\|_{L^{k_j}(\mathbb{T})} \geq \Lambda_j\right\}$ occurs infinitely often with probability 0.

Independent of the probability we have, for $n \geq 2^{2^j}$,

$$\|\Phi_n\|_{k_j} \leq \|M_{2^{2^j} \to n}\|_{k_j} \|\Phi_{2^{2^j}}\|_\infty \tag{3.9}$$

We use another simple deterministic Lemma.

Lemma 3.8. *For* $2^{2^{j+1}} \geq n \geq 2^{2^j} + 1$, *we have*

$$\|\Phi_n\|_\infty \leq (1 - 2^{-2^{j+1}})^{-1} (2^{2^{j+1}})^{\frac{2}{k_j}} \|\Phi_n\|_{k_j} \tag{3.10}$$

Proof. This follows from the Bernstein estimate, which states that for a polynomial P_n of degree n,

$$\|P_n'\|_\infty \leq n\|P_n\|_\infty \tag{3.11}$$

(3.11) shows that for Φ_n and any $\epsilon_n > 0$, there is a set Ω with $|\Omega| \geq \frac{\epsilon_n}{n}$ for which

$$|\Phi_n(z)| \geq (1 - \epsilon_n)\|\Phi_n(z)\|_\infty, \quad z \in \Omega$$

Therefore

$$\|\Phi_n\|_p^p \geq \int_\Omega |\Phi_n(e^{i\theta})|^p \frac{d\theta}{2\pi} \geq \frac{\epsilon_n}{n}\left((1-\epsilon_n)\|\Phi_n\|_\infty\right)^p$$

Since $n \leq 2^{2^{j+1}}$ for $n \in \{2^{2^j}+1, \ldots, 2^{2^{j+1}}\}$, choosing $\epsilon_n \equiv 2^{-2^{j+1}}$ and $p = k_j$ yields (3.10). □

So by (3.9) and (3.10) we have

$$\max_{2^{2^j} \leq n \leq 2^{2^{j+1}}} \|\Phi_n\|_\infty \leq \max_{2^{2^j} \leq n \leq 2^{2^{j+1}}} (1-2^{-2^{j+1}})^{-1} (2^{2^{j+1}})^{\frac{2}{k_j}} \|\Phi_n\|_{k_j}$$

$$\leq \max_{2^{2^j} \leq n \leq 2^{2^{j+1}}} (1-2^{-2^{j+1}})^{-1} (2^{2^{j+1}})^{\frac{2}{k_j}} \|M_{2^{2^j} \to n}\|_{k_j} \|\Phi_{2^{2^j}}\|_\infty$$

And so with probability 1, if j is sufficiently large we have

$$\max_{2^{2^j} \leq n \leq 2^{2^{j+1}}} \|\Phi_n\|_\infty \leq (1-2^{-2^{j+1}})^{-1} (2^{2^{j+1}})^{\frac{2}{k_j}} \Lambda_j \|\Phi_{2^{2^j}}\|_\infty \tag{3.12}$$

We have seen $\prod_j \Lambda_j < \infty$. For the first term,

$$\prod_j (1-2^{-2^{j+1}})^{-1} = \exp\left(-\sum_j (\log(1-2^{-2^{j+1}}))\right) \leq \exp\left(\sum_j 2^{-2^{j+1}}\right) < \infty$$

So we must investigate the convergence of

$$\prod_j (2^{2^{j+1}})^{\frac{2}{k_j}} < \infty \iff \sum_j \frac{2^j}{k_j} < \infty$$

Recall that $k_j \sim T_j^{-1/2}$, and so since $\sum_j 2^j \sqrt{T_j} < \infty$ we can iterate (3.12) forward with probability 1 to obtain

$$\sup_n \|\Phi_n\|_\infty \leq C(\omega)$$

□

Chapter 4

On Verblunsky Parameters in Steklov's Problem

The Steklov conjecture and its associated problems have a long history. Many measures which violate the conjecture have been constructed. But the Verblunsky parameters associated to these measures have remained unknown. This is surprising because the Verblunsky parameters are an independently interesting aspect of OPUC.

One perspective on the Verblunsky parameters is physical–they play the role of the electric potential. The analogy between Schrödinger operators and OPUC also yields a physical interpretation of violating Steklov's conjecture.

Roughly speaking, a system of OPUC which violates the Steklov conjecture is an example of a resonance. There is a specific frequency at which the system oscillates with high amplitude, but it is regular at infinity. The conjecture of Steklov is the conjecture that if we enforce boundedness away from 0, resonances in the associated quantum systems are disallowed. Since the potential lives in physical space and can be observed or modified directly, the Verblunsky parameters which yield a resonance are interesting from a physical perspective; how can the potential behave in order to create this resonance?

Verblunsky parameters also provide a natural approach to disproving the Steklov conjecture, namely the following strategy:

1. Fix an arbitrarily large number n.

2. Construct a polynomial ϕ_n which is permissible as OPUC whose uniform norm is large in n.

3. Demonstrate that there is a sequence of Verblunsky coefficients $\{\alpha_{n+1}, \ldots, \alpha_N\}$, which take ϕ_n to ϕ_N obeying $\|\phi_N\|_\infty \leq C$.

4. Finally, take $d\mu = |\phi_N|^{-2} \frac{d\theta}{2\pi}$.

Indeed, this is essentially the strategy followed by [2]. The difficulty lies in step 3. For this step [2] develops a powerful new technology, the so-called decoupling Lemma (Lemma 4.4) which decouples the problem at degree n.

The construction of a polynomial which grows in n and then decays back down to $O(1)$ for $N \gg n$ (steps 2 and 3 above) could be accomplished in one fell swoop if a particularly useful set of Verblunsky parameters could be identified and the calculations carried out. Indeed, this is a strategy the author tried, and within which he made very little progress. It does not seem to be quite as simple as it appears.

This conclusion is due in part to the high-powered results the author and his advisor used to produce (asymptotics of) recursion parameters which would create a resonance. These parameters are computed via the paper of Deift, Its and Krasovsky [18] which used a matrix-valued Riemann-Hilbert problem to asymptotically compute all polynomials and related quantities for the case of weights on the circle with a particular class of singularities, known as Fisher-Hartwig weights.

4.1 Verblunsky asymptotics in Steklov's problem

The polynomials of the second kind will be denoted by $\{\Psi_n\}$ (monic) and $\Psi_n^*(z) = z^n\overline{\Psi_n(\bar{z}^{-1})},\ z \in \mathbb{C}$, $\{\psi_n\}$ (orthonormal), and $\{\psi_n^*\}$. Recall that the pair (Φ_n, Φ_n^*) satisfies the Szegő recurrence:

$$\begin{cases} \Phi_{n+1}(z,\sigma) = & z\Phi_n(z,\sigma) - \overline{\alpha}_n\Phi_n^*(z,\sigma) \\ \Phi_{n+1}^*(z,\sigma) = & \Phi_n^*(z,\sigma) - \alpha_n z\Phi_n(z,\sigma) \end{cases} \tag{4.1}$$

The pair (Ψ_j, Ψ_j^*) satisfies the same recurrence except that the parameters are $\{-\alpha_n\}$. The recursion parameters $\{\alpha_n\} \subseteq \mathbb{D}^\infty$ were called Schur parameters due to their relationship with Schur functions and the Schur algorithm in [26], where these results originally appeared, but here we will keep the name Verblunsky for consistency.

The main result of this section is

Theorem 4.1 (Denisov-R. [26])**.** *Fix $\epsilon \in (0, \epsilon_0]$ where ϵ_0 is sufficiently small. Then, there is $n_0(\epsilon)$ such that for every $n > n_0(\epsilon)$, there is a weight $\tilde{w}^{(n)}$ such that*

$$\tilde{w}^{(n)} \textit{ satisfies the uniform Steklov condition}: \quad \left\| \frac{1}{\tilde{w}^{(n)}} \right\|_{L^\infty(\mathbb{T})} \lesssim 1, \tag{4.2}$$

$$\|\phi_{2n+1}(z, \tilde{w}^{(n)})\|_{L^\infty(\mathbb{T})} >_\epsilon \ln n, \tag{4.3}$$

and the asymptotics for $\alpha_j^{(n)}$ is given by

$$\alpha_j^{(n)} = -\begin{cases} \dfrac{i^{j+1}}{j+1}\displaystyle\sum_{s=\pm 1} s^{j+1}(2(j+1))^{\frac{2i\epsilon s}{\pi}}\frac{\Gamma(1-\frac{i\epsilon s}{\pi})}{\Gamma\left(\frac{i\epsilon s}{\pi}\right)} + r_{j,\epsilon}, & 0 \le j \le n-1 \\ -\dfrac{i^{j'+1}}{j'+1}\displaystyle\sum_{s=\pm 1} s^{j'+1}(2(j'+1))^{\frac{2i\epsilon s}{\pi}}\frac{\Gamma(1-\frac{i\epsilon s}{\pi})}{\Gamma\left(\frac{i\epsilon s}{\pi}\right)} + r_{j',\epsilon}, & n \le j \le 2n-1 \\ 0, & j = 2n \\ \dfrac{(-1)^j i^{j-2n}}{j-2n}\displaystyle\sum_{s=\pm 1} s^{j-2n}(2(j-2n))^{\frac{2i\epsilon s}{\pi}}\frac{\Gamma(1-\frac{i\epsilon s}{\pi})}{\Gamma\left(\frac{i\epsilon s}{\pi}\right)} + r_{j-2n,\epsilon}, & 2n+1 \le j \le 3n \\ 0, & j > 3n \end{cases} \tag{4.4}$$

where $j' = 2n - 1 - j$ *and* $|r_{j,\epsilon}| < C_\epsilon (j+1)^{-2}$.

Remark. The existence of the measure satisfying these conditions is not a new result [62] and the logarithmic growth is not optimal [21].The polynomial constructed by our method has a structure similar to the one from [22, 62].

Remark. Careful analysis of (4.4) shows that the main terms are real-valued and converge to 0 as $\epsilon \to 0$ for every fixed j. The results in [18] allow one to control the dependence of C_ϵ on ϵ when it converges to zero and we conjecture that $\lim_{\epsilon \to 0} C_\epsilon = 0$.

Remark. The asymptotics above would be made exact if the recursion parameters corresponding to the weight which is flat away from 2 jumps were computed exactly.

4.2 Proof of Theorem 4.1

The measures considered below will be symmetric with respect to $\mathbb{R}$ and so the related Verblunsky parameters will be real. We will need the following simple Lemma.

Lemma 4.2. *Suppose* $\Phi_k, \Psi_k, \Phi_k^*, \Psi_k^*$ *are the polynomials that correspond to* **real** *Verblunsky parameters* $\{\alpha_j\}_{j=0}^{k-1}$. *Then, the polynomials associated to the sequence of Verblulnsky parameters* $\{\alpha_j\}$ *given by*

$$\alpha_j = \begin{cases} \alpha_j & : 0 \le j \le k-1 \\ -\alpha_{2k-1-j} & : k \le j \le 2k-1 \end{cases}$$

satisfy

$$2\Phi_{2k} = \Phi_k^2 + \Phi_k \Psi_k - z^{-1}(\Phi_k^*)^2 + z^{-1}\Phi_k^* \Psi_k^*,$$

$$2\Phi_{2k}^* = (\Phi_k^*)^2 + \Phi_k^* \Psi_k^* - z\Phi_k^2 + z\Phi_k \Psi_k.$$

Proof. The pair (Ψ_j, Ψ_j^*) satisfies the Szegő recurrence with parameters $\{-\alpha_j\}$. If

$$A = \prod_{j=k-1}^{0} \begin{bmatrix} z & -\alpha_j \\ -z\alpha_j & 1 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \tag{4.5}$$

then

$$\begin{bmatrix} \Phi_k & \Psi_k \\ \Phi_k^* & -\Psi_k^* \end{bmatrix} = A \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Thus,

$$a = \frac{\Phi_k + \Psi_k}{2},\ b = \frac{\Phi_k - \Psi_k}{2},\ c = \frac{\Phi_k^* - \Psi_k^*}{2},\ d = \frac{\Phi_k^* + \Psi_k^*}{2}.$$

First we reverse the dynamics. We are interested in

$$\prod_{j=0}^{k-1} \begin{bmatrix} z & -\alpha_j \\ -z\alpha_j & 1 \end{bmatrix}.$$

We see that

$$A^T = \prod_{j=0}^{k-1} \begin{bmatrix} z & -z\alpha_j \\ -\alpha_j & 1 \end{bmatrix} = \prod_{j=0}^{k-1} \left(\begin{bmatrix} 1 & 0 \\ 0 & z^{-1} \end{bmatrix} \begin{bmatrix} z & -\alpha_j \\ -z\alpha_j & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & z \end{bmatrix} \right)$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & z^{-1} \end{bmatrix} \prod_{j=0}^{k-1} \left(\begin{bmatrix} z & -\alpha_j \\ -z\alpha_j & 1 \end{bmatrix} \right) \begin{bmatrix} 1 & 0 \\ 0 & z \end{bmatrix}.$$

Therefore,

$$\prod_{j=0}^{k-1} \begin{bmatrix} z & -\alpha_j \\ -z\alpha_j & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & z \end{bmatrix} A^T \begin{bmatrix} 1 & 0 \\ 0 & z^{-1} \end{bmatrix}.$$

We have

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} z & -\alpha_j \\ -z\alpha_j & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} z & \alpha_j \\ z\alpha_j & 1 \end{bmatrix}.$$

Therefore, (4.5) implies

$$\prod_{j=k-1}^{0} \begin{bmatrix} z & \alpha_j \\ z\alpha_j & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} A \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Combining the above results, we get

$$\prod_{j=0}^{k-1} \begin{bmatrix} z & \alpha_j \\ z\alpha_j & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -z \end{bmatrix} A^T \begin{bmatrix} 1 & 0 \\ 0 & -z^{-1} \end{bmatrix},$$

and so

$$\begin{bmatrix} \Phi_{2k} & \Psi_{2k} \\ \Phi_{2k}^* & -\Psi_{2k}^* \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -z \end{bmatrix} \begin{bmatrix} a & c \\ b & d \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -z^{-1} \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} a(a+b) - z^{-1}c(c+d) & a(a-b) + z^{-1}c(d-c) \\ d(c+d) - zb(a+b) & d(c-d) + zb(b-a) \end{bmatrix}.$$

Thus,

$$2\Phi_{2k} = \Phi_k^2 + \Phi_k\Psi_k - z^{-1}(\Phi_k^*)^2 + z^{-1}\Phi_k^*\Psi_k^*$$

and

$$2\Phi_{2k}^* = (\Phi_k^*)^2 + \Phi_k^*\Psi_k^* - z\Phi_k^2 + z\Phi_k\Psi_k.$$

□

Notice that this computational approach bears some resemblance to symmetry methods in differential equations; we are using a particular algebraic symmetry of the difference equation to simplify a potentially complex calculation.

The following result is well-known.

Lemma 4.3. *If $\{\alpha_j\}$ are the Verblunsky parameters for the measure $\sigma(\theta)$, then parameters $\{\alpha_j^{(\beta)}\}$ for the translated measure $\sigma^{(\beta)} = \sigma(\theta - \beta)$ are given by*

$$\alpha_j^{(\beta)} = e^{-i(j+1)\beta}\alpha_j, \quad j = 0, 1, \ldots$$

Proof. The proof follows from making the following two observations:

$$\Phi_n(z, \sigma^{(\beta)}) = e^{in\beta}\Phi_n(ze^{-i\beta})$$

(which follows from the definition) and, take $z = 0$ in (1.2), $\Phi_{j+1}(0, \sigma^{(\beta)}) = -\overline{\alpha}_j^{(\beta)}$. □

We will need the decoupling Lemma. Its proof is contained in Appendix B but we include the statement for the reader's convenience.

Lemma 4.4. *Suppose we are given a polynomial ϕ_n of degree n and Carathéodory function $\widetilde{F}$ which satisfy the following properties*

1. *$\phi_n^*(z)$ has no roots in $\overline{\mathbb{D}}$.*

2. *Normalization on the size and "rotation"*

$$\int_{\mathbb{T}} |\phi_n^*(z)|^{-2} d\theta = 2\pi \ , \quad \phi_n^*(0) > 0 \ . \tag{4.6}$$

3. *$\widetilde{F} \in C^\infty(\mathbb{T})$, $\operatorname{Re} \widetilde{F} > 0$ on $\mathbb{T}$, and*

$$\frac{1}{2\pi} \int_{\mathbb{T}} \operatorname{Re} \widetilde{F}(e^{i\theta}) d\theta = 1 \ . \tag{4.7}$$

Denote the Verblunsky parameters given by the probability measures μ_n and $\widetilde{\sigma}$

$$d\mu_n = \frac{d\theta}{2\pi|\phi_n^*(e^{i\theta})|^2}, \quad d\widetilde{\sigma} = \widetilde{\sigma}' d\theta = \frac{\operatorname{Re} \widetilde{F}(e^{i\theta})}{2\pi} d\theta,$$

as $\{\alpha_j\}$ and $\{\widetilde{\alpha}_j\}$, respectively. Then, the probability measure σ, corresponding to Verblunsky coefficients

$$\alpha_0, \ldots, \alpha_{n-1}, \widetilde{\alpha}_0, \widetilde{\alpha}_1, \ldots$$

is purely absolutely continuous with the weight given by

$$\sigma' = \frac{4\widetilde{\sigma}'}{|\phi_n + \phi_n^* + \widetilde{F}(\phi_n^* - \phi_n)|^2} = \frac{2 \operatorname{Re} \widetilde{F}}{\pi|\phi_n + \phi_n^* + \widetilde{F}(\phi_n^* - \phi_n)|^2} \ . \tag{4.8}$$

The polynomial ϕ_n is the n^{th} orthonormal polynomial for σ.

The next result is a consequence of the recent analysis of the asymptotics of the polynomials orthogonal with the respect to the so-called Fisher-Hartwig weights. We will use [18] as the main reference.

Consider the weight on $\mathbb{T}$ given by

$$f(z) = e^{\epsilon} g_{i,-\frac{i\epsilon}{\pi}}(z) g_{-i,\frac{i\epsilon}{\pi}}(z), \tag{4.9}$$

where $z = e^{i\theta}$, $\theta \in [0, 2\pi)$ and $\epsilon \in (0, \epsilon_0]$ with ϵ_0 to be chosen sufficiently small. For $z_j = e^{i\theta_j}$,

$$g_{z_j,\beta_j} = \begin{cases} e^{i\pi\beta_j} & : 0 \leq \arg z < \theta_j \\ e^{-i\pi\beta_j} & : \theta_j \leq \arg z < 2\pi \end{cases}.$$

That is, f is a weight with two jumps, one from e^{ϵ} to $e^{-\epsilon}$ around i (in counterclockwise direction), and one from $e^{-\epsilon}$ back to e^{ϵ} around $-i$ (again in counterclockwise direction). It does not define a probability measure but this will not influence the polynomials much due to invariance of the monic orthogonal polynomials under scaling. Notice that f is symmetric with respect to $\mathbb{R}$ so all its recursion parameters are real, as follows from the proof of the Szegő recurrence (1.2) upon replacing the weight w with $\overline{w}$.

We will need the following result, the proof of which (essentially contained in [18]) will be discussed in the Appendix.

Lemma 4.5. *Let $\epsilon \in (0, \epsilon_0]$ and $n > n_0(\epsilon)$. Then, for the weight f given by (4.9), the n-th associated monic polynomials of the first and second kinds Φ_n and Ψ_n respectively, and their $*$-polynomials Ψ_n^* and Φ_n^* satisfy the following estimates:*

$$|\Phi_n^*(z)| \sim 1, \quad z \in \mathbb{T}, \tag{4.10}$$

$$\|\Phi_n^*\Psi_n^* + z\Phi_n\Psi_n\|_{L^\infty(\mathbb{T})} >_\epsilon \ln n, \quad z \in \mathbb{T}\,, \tag{4.11}$$

$$\left|\frac{\Psi_n^*(z)}{\Phi_n^*(z)} + \frac{\Psi_n^*(-z)}{\Phi_n^*(-z)}\right| \lesssim 1, \quad z \in \mathbb{T}\,. \tag{4.12}$$

Remark. The following general identity is immediate from the Szegő recursion by taking the determinant

$$\det\begin{bmatrix} \Phi_n & \Psi_n \\ \Phi_n^* & -\Psi_n^* \end{bmatrix} = -2z^n \prod_{j=0}^{n-1}(1-|\alpha_j|^2)$$

so

$$\overline{\Phi}_n^*\Psi_n^* + \Phi_n^*\overline{\Psi}_n^* = 2\prod_{j=0}^{n-1}(1-|\alpha_j|^2)\,.$$

Then, (4.10) and (4.12) give

$$\left|\frac{\Psi_n^*}{\Phi_n^*} + \overline{\left(\frac{\Psi_n^*}{\Phi_n^*}\right)}\right| \lesssim 1 \tag{4.13}$$

uniformly over $\mathbb{T}$ if the polynomials correspond to the weight (4.9).

The asymptotics of the polynomials yields the following result as well.

Lemma 4.6. *The Verblunsky parameters $\{\alpha_j\}$ associated to the f above satisfy:*

$$\alpha_j = -(j+1)^{-1} i^{j+1}\left((2(j+1))^{\frac{2i\epsilon}{\pi}}\frac{\Gamma(1-\frac{i\epsilon}{\pi})}{\Gamma\left(\frac{i\epsilon}{\pi}\right)} + (-1)^{j+1}(2(j+1))^{-\frac{2i\epsilon}{\pi}}\frac{\Gamma\left(1+\frac{i\epsilon}{\pi}\right)}{\Gamma\left(\frac{-i\epsilon}{\pi}\right)}\right) + r_{j,\epsilon},$$

$$|r_{j,\epsilon}| < C_\epsilon (j+1)^{-2}\,.$$

We give the proof of Lemmas 4.5 and 4.6 in the Appendix.

Now we are in position to give a proof of Theorem 4.1.

Proof. (of Theorem 4.1). Given $\epsilon > 0$ and large n, we consider f defined by (4.9). If the corresponding coefficients are denoted by $\{\alpha_j\}$, we consider the weight $w^{(n)}$ given by Verblunsky parameters $\{\alpha_j^{(n)}\}$ defined as

$$\alpha_j^{(n)} = \begin{cases} \alpha_j, & j \le n-1 \\ -\alpha_{j'}, & j' = 2n-1-j,\ n \le j \le 2n-1 \\ 0, & j = 2n \\ (-1)^{j-2n}\alpha_{j-2n-1}, & 2n+1 \le j \le 3n \end{cases}.$$

Now, (4.4) follows immediately from Lemma 4.6 and we need to show (4.2) and (4.3).

Let us prove (4.3). Notice first that $|\phi_j| \sim |\Phi_j|$ by (1.3) so it is sufficient to consider Φ_{2n+1}. We apply Lemma 4.2 to say

$$2\Phi_{2n} = \Phi_n^2 + \Phi_n\Psi_n - z^{-1}\Phi_n^{*\,2} + z^{-1}\Phi_n^*\Psi_n^*,$$

$$2\Phi_{2n}^* = \Phi_n^{*\,2} + \Phi_n^*\Psi_n^* - z\Phi_n^2 + z\Phi_n\Psi_n\,.$$

Since $\alpha_{2n}^{(n)} = 0$, we get $\Phi_{2n+1} = z\Phi_{2n}, \Psi_{2n+1} = z\Psi_{2n}, \Phi_{2n+1}^* = \Phi_{2n}^*$, and $\Psi_{2n+1}^* = \Psi_{2n}^*$ so

$$2\Phi_{2n+1} = z\Phi_n^2 + z\Phi_n\Psi_n - \Phi_n^{*\,2} + \Phi_n^*\Psi_n^*, \quad 2\Phi_{2n+1}^* = \Phi_n^{*\,2} + \Phi_n^*\Psi_n^* - z\Phi_n^2 + z\Phi_n\Psi_n. \tag{4.14}$$

We have

$$2\|\Phi_{2n+1}\|_\infty \ge \|\Phi_n^*\Psi_n^* + z\Phi_n\Psi_n\|_\infty - 2\|\Phi_n\|_\infty^2 >_\epsilon \ln n$$

by Lemma 4.5.

To show (4.2), we will use Lemma 4.4. We choose Carathéodory function for the decoupled problem as

$$\widetilde{F}(z) = \frac{\psi_n^*(-z)}{\phi_n^*(-z)} = \frac{\Psi_n^*(-z)}{\Phi_n^*(-z)} = -\frac{\Psi_n^*(z)}{\Phi_n^*(z)} + O(1) \quad \text{by (4.12).} \tag{4.15}$$

We have (see [65], Theorem 3.2.4)

$$\operatorname{Re} \widetilde{F}(e^{i\theta}) = |\phi_n^*(e^{i(\theta+\pi)})|^{-2}$$

and $\widetilde{F}$ is Carathéodory function of the Bernstein-Szegő weight $(2\pi)^{-1}|\phi_n^*(e^{i(\theta+\pi)})|^{-2}$ having the Verblunsky parameters

$$\begin{cases} (-1)^{j+1}\alpha_j, & j < n \\ \quad 0, & j \geq n \end{cases}$$

as follows from Lemma 4.3. Since our Verblunsky parameters $\alpha_j^{(n)} = 0, j > 3n$, we have

$$|\widetilde{\Pi}(-z)| = |\phi_n^*(-z, f)| \sim 1, \quad \text{by} \quad (4.10).$$

So by the identity (4.8) we only need to show

$$|\Phi_{2n+1} + \Phi_{2n+1}^* + \widetilde{F}(\Phi_{2n+1}^* - \Phi_{2n+1})| \lesssim 1$$

uniformly on $\mathbb{T}$. Recall that $\widetilde{F} = -\dfrac{\Psi_n^*}{\Phi_n^*} + O(1)$ by (4.12). We introduce auxiliary

$$D = \frac{\Phi_n^* \Psi_n^*}{2}, \quad A = -\frac{{\Phi_n^*}^2}{2}.$$

Notice that

$$|A| \sim 1, \quad \frac{D}{A} = -\frac{\Psi_n^*}{\Phi_n^*} = -\overline{\left(\frac{D}{A}\right)} + O(1) \tag{4.16}$$

by (4.10) and (4.13). We can now rewrite (4.14) as

$$\Phi_{2n+1} = -A^* + D^* + A + D, \quad \Phi_{2n+1}^* = -A + D + A^* + D^*.$$

(where the $(*)$-operations in these identities are of order $2n+1$).

Then, by (4.15),

$$\Phi_{2n+1} + \Phi_{2n+1}^* + \widetilde{F}(\Phi_{2n+1}^* - \Phi_{2n+1}) = 2\left((D + D^*) + \frac{D}{A}(A^* - A)\right) + O(1) =$$

$$2\left(D^* + \frac{D}{A}A^*\right) + O(1) = \frac{2z^{2n}}{\overline{A}}\left(\frac{D}{A} + \overline{\left(\frac{D}{A}\right)}\right) = O(1)$$

by (4.16).

□

Appendix A

Some standard harmonic analysis

A.1 BMO($\mathbb{T}$) functions

BMO was defined as a class by John and Nirenberg in [39], the functions of bounded mean oscillation. BMO behavior can be captured by the so-called sharp maximal function

$$f^{\#}(x) = \sup_{Q \supset x} \frac{1}{|Q|} \int_Q |f(y) - f_Q| dy \tag{A.1}$$

where Q ranges over subintervals of $\mathbb{T}$ containing x and

$$f_Q = \frac{1}{|Q|} \int_Q f(y) dy$$

With this definition,

$$f^{\#} \in L^\infty(\mathbb{T}) \iff f \in \mathrm{BMO}(\mathbb{T})$$

and the L^∞ norm of $f^{\#}$ is called the "BMO norm" of f, denoted $\|f\|_{BMO}$. The BMO norm is in reality only a quasinorm, satisfying the triangle inequality and linearity in multiplication, but constants have BMO norm 0. After factoring out constants it becomes a legitimate norm.

Clearly an L^∞ function is also BMO. But BMO is strictly larger than L^∞; a logarithmic singularity like $|\log(\theta)|$ will be in BMO by a simple calculation. BMO can be rather tricky; $\mathrm{sgn}(\theta)|\log\theta|$ is *not* BMO.

We collect a few standard facts about BMO functions, and refer to [55] Section 7.4 or [66] Chapter 4 for their proofs.

BMO is intended as a useful substitute for L^∞ so we should be able to interpolate between BMO and L^p. Indeed this is true:

Theorem A.1. *Let T be a linear operator bounded on L^{p_0} for some $1 \le p_0 < \infty$ and bounded from $L^\infty \to BMO$. Then T is bounded on L^p for any $p_0 < p < \infty$.*

A fundamental property of BMO is its geometric/arithmetic scaling. That is, if $Q_2 \supset Q_1$ with $|Q_2| = 2|Q_1|$, we have

$$|f_{Q_1} - f_{Q_2}| \le \frac{|Q_1|}{|Q_2|} \frac{1}{|Q_1|} \int_{Q_1} |f(y) - f_{Q_1}| dy \le 2\|f\|_{\mathrm{BMO}}$$

Iterating this and using the triangle inequality shows, for $\{Q_n\}$ a nested sequence of dyadic cubes,

$$|f_{Q_n} - f_{Q_0}| \le 2n\|f\|_{\mathrm{BMO}}$$

This relation says exactly that averages over cubes of BMO functions increase only on an arithmetic scale while their supports decrease on a geometric scale; this principle is the basis of the John-Nirenberg Theorem, first proved in [39].

Theorem A.2 (John-Nirenberg)**.** *Let $f \in BMO(\mathbb{T})$. Then, for any cube Q, we have*

$$|\{x \in Q | |f(x) - f_Q| > \lambda\}| \le C|Q| \exp\left(-c\frac{\lambda}{\|f\|_{BMO}}\right)$$

for any $\lambda > 0$ with absolute c and C.

Integrating this inequality shows that f is exponentially integrable, and in particular $f \in L^p(\mathbb{T})$ for all finite p.

Another use of the arithmetic/geometric scaling properties of BMO was found by Coifman, Rochberg and Weiss in their singular integral commutator estimate [16] which is central to our section on BMO weights.

Theorem A.3 (Coifman-Rochberg-Weiss)**.** *For T a so-called "strong" Calderón-Zygmund operator, $b \in BMO(\mathbb{T})$, we have*

$$\|[T,b]\|_{p,p} \leq C(d,p)A||b||_{BMO}$$

where the constant A depends on the operator T.

In fact BMO is characterized by this property; if the inequality holds for a single Riesz transform then $b \in BMO$. [16] also proved the iterated commutator bounds we use in Section 3.

Finally, there is a close relationship between BMO functions and A_p weights of Muckenhoupt which will be used in a Lemma to come. A weight w is A_p for $1 < p < \infty$ if there exists $A < \infty$ so that for any cube $Q \subseteq \mathbb{T}$,

$$\left(\frac{1}{|Q|}\int_Q w(x)dx\right)\left(\frac{1}{|Q|}\int_Q w(x)^{-q/p}dx\right)^{p/q} \leq A < \infty$$

where q is such that $\frac{1}{p} + \frac{1}{q} = 1$. For the basic theory of such weights see [66] Chapter 5. The relationship between BMO and A_p will be seen in the proof of Lemma A.5.

A.2 Lemmas of Section 2.2

Lemma A.4. *For every $p \in [2,\infty)$,*

$$\|P_{[1,n]}\|_{p,p} \leqslant 1 + C(p-2) \tag{A.2}$$

Proof. If $\mathcal{P}^+$ is the projection of $L^2(\mathbb{T})$ onto $H^2(\mathbb{T})$ (analytic Hardy space), then

$$P_{[1,n]} = z\mathcal{P}^+z^{-1} - z^{n+1}\mathcal{P}^+z^{-(n+1)} = 0.5(zHz^{-1} - z^{n+1}Hz^{-(n+1)}) + zP_0z^{-1} - z^{n+1}P_0z^{-(n+1)} \tag{A.3}$$

where H is the Hilbert transform. Recall that $\|H\|_{p,p} = \cot(\pi/(2p))$ [61] and $\|P_0\|_{p,p} = 1$ by monotonicity of L^p-norms on probability space. So it suffices to note that

$$\|P_{[1,n]}\|_{2,2} = 1$$

and interpolate between this and e.g. $\|P_{[1,n]}\|_{3,3} = C$. □

The proof of the following Lemma uses some standard results of harmonic analysis.

Lemma A.5. *If $\|w\|_{BMO} = t$ and $p \in [2,3]$, then we have*

$$\|\mathbf{C}_j\|_{p,p} \leqslant (Cjt)^j$$

Proof. Consider the following operator-valued function

$$F(z) = e^{zw}P_{[1,n]}e^{-zw}$$

If we can prove that $F(z)$ is weakly analytic around the origin (i.e., analyticity of the scalar function $\langle F(z)f_1, f_2\rangle$ with fixed $f_{1(2)} \in C^\infty$), then

$$F(z) = \frac{1}{2\pi i}\int_{|\xi|=\epsilon} \frac{F(\xi)}{\xi - z}d\xi, \quad z \in B_\epsilon(0)$$

understood in a weak sense. By induction, one can then easily show the well-known formula

$$\mathbf{C}_j = \partial^j F(0) = \frac{j!}{2\pi i}\int_{|\xi|=\epsilon} \frac{F(\xi)}{\xi^{j+1}}d\xi$$

which explains that we can control $\|\mathbf{C}_j\|_{p,p}$ by the size of $\|F(\xi)\|_{p,p}$ on the circle of radius ϵ. Indeed,

$$\|\mathbf{C}_j\|_{p,p} = \sup_{f_{1(2)}\in C^\infty, \|f_1\|_p \leqslant 1, \|f_2\|_{p'}\leqslant 1} |\langle \mathbf{C}_j f_1, f_2\rangle| \leq$$

$$\frac{j!}{2\pi} \sup_{f_{1(2)}\in C^\infty, \|f_1\|_p \leqslant 1, \|f_2\|_{p'}\leqslant 1} \left| \int_{|\xi|=\epsilon} \frac{\langle F(\xi) f_1, f_2\rangle}{\xi^{j+1}} d\xi \right| \leqslant \frac{j!}{\epsilon^j} \max_{|\xi|=\epsilon} \|F(\xi)\|_{p,p}$$

The weak analyticity of $F(z)$ around the origin follows immediately from, e.g., the John-Nirenberg estimate. To bound $\|F\|_{p,p}$, we use the following well-known result (which is again an immediate corollary from John-Nirenberg inequality, see, e.g., [66], p.218).

There is ϵ_0 such that

$$\|\widetilde{w}\|_{BMO} < \epsilon_0 \implies [e^{\widetilde{w}}]_{A_p} \leqslant [e^{\widetilde{w}}]_{A_2} < C, \quad p > 2$$

The Hunt-Muckenhoupt-Wheeden Theorem ([66], p.205), asserts that

$$\sup_{[\widehat{w}]_{A_p}\leqslant C} \|H\|_{(L^p_{\widehat{w}}(\mathbb{T}), L^p_{\widehat{w}}(\mathbb{T}))} = \sup_{[\widehat{w}]_{A_p}\leqslant C} \|\widehat{w}^{1/p} H \widehat{w}^{-1/p}\|_{p,p} = C(p) < \infty, \quad p\in[2,\infty) \tag{A.4}$$

We must also bound the $\|P_0\|_{L^p(\hat{w}), L^p(\hat{w})}$ terms. But this is immediate by definition of the A_p characteristic:

$$\int_{\mathbb{T}} |\hat{w}^{1/p} P_0 \hat{w}^{-1/p} f|^p d\theta \leq \|f\|_p \int \hat{w} \|\hat{w}^{-1/p}\|_{p'} \leq \|f\|_p \, [\hat{w}]_{A_p}$$

where $\frac{1}{p} + \frac{1}{p'} = 1$. Taking $\epsilon \ll t^{-1}$, we get the statement. □

The following Lemma provides an estimate which is not optimal but it is good enough for our purposes.

Lemma A.6. *Suppose* $w \geqslant 0, \|w\|_{BMO} = t, \|w^{-1}\|_{BMO} = s$, *and* $\|w\|_1 = 1$. *Then,*

$$(2\pi)^2 \leqslant \|w^{-1}\|_1 \lesssim 1 + (1+t)s$$

Proof. Denote $\|w^{-1}\|_1 = M$. Then, by Cauchy-Schwarz inequality,

$$2\pi \leqslant \|w\|_1^{1/2} \|w^{-1}\|_1^{1/2} = M^{1/2}$$

On the other hand, by John-Nirenberg estimate for w^{-1},

$$|\{\theta : |w^{-1} - (2\pi)^{-1}M| > \lambda\}| \lesssim \exp\left(-\frac{C\lambda}{s}\right)$$

Choosing $\lambda = (4\pi)^{-1}M$, we get

$$|\Omega^c| \lesssim \exp\left(-\frac{CM}{s}\right) \lesssim \left(\frac{s}{M}\right)^2, \quad \text{where} \quad \Omega = \left\{\theta : \frac{4\pi}{3M} \leqslant w \leqslant \frac{4\pi}{M}\right\} \tag{A.5}$$

Then, $\|w\|_1 = 1$ and therefore

$$1 = \int_{w \leqslant (4\pi)/M} w d\theta + \int_{w > (4\pi)/M} w d\theta$$

$$\int_{w > (4\pi)/M} w d\theta \geqslant 1 - 8\pi^2 M^{-1} \tag{A.6}$$

By John-Nirenberg inequality, we have

$$\|w - (2\pi)^{-1}\|_p < Ctp, \quad p < \infty \tag{A.7}$$

We choose $p = 2$ in the last estimate and use Cauchy-Schwarz inequality in (A.6) to get

$$1 - 8\pi^2 M^{-1} \leqslant \int_{w > (4\pi)/M} w d\theta \leqslant \|w\|_2 \cdot |\{\theta : w > 4\pi/M\}|^{1/2} \leqslant \|w\|_2 \cdot |\Omega_c|^{1/2} \lesssim \frac{(1+t)s}{M}$$

where we used (A.5) and (A.7) for the last bound. So, $M \lesssim (1+t)s + 1$. □

Lemma A.7. *For* $p \in [2, \infty)$, *we have*

$$\|[w, P_{[1,n]}]\|_{p,p} \lesssim p^2 \|w\|_{BMO}$$

Proof. The proof is standard but we give it here for completeness. Assume $\|w\|_{BMO} = 1$. By duality, formula (A.3) and the John-Nirenberg estimate, it is sufficient to show that

$$\|[w, H]\|_{p,p} \leqslant C(p-1)^{-2}, \quad p \in (1, 2] \tag{A.8}$$

We will interpolate between two bounds: the standard Coifman-Rochberg-Weiss Theorem for $p = 2$ ([16],[66])

$$\|[H, w]\|_{2,2} \leqslant C \tag{A.9}$$

and the following estimate

$$|\{x : |([H, w]f)(x)| > \alpha\}| \leqslant C \int_{\mathbb{T}} \frac{|f(t)|}{\alpha} \left(1 + \log^+ \left(\frac{|f(t)|}{\alpha}\right)\right) dt \tag{A.10}$$

(See [60], the estimate was obtained on $\mathbb{R}$ for smooth f with compact support. The proof, however, is valid for $\mathbb{T}$ as well and, e.g., piece-wise smooth continuous f). Assume a smooth f is given and denote $\lambda_f(t) = |\{x : |f(x)| > t\}|, t \geqslant 0$. Take $A > 0$ and consider $f_A = f \cdot \chi_{|f| \leqslant A} + A \cdot \text{sgn} f \cdot \chi_{|f|>A}$, $g_A = f - f_A$. Let $T = [H, w]$. Then,

$$\|Tf\|_p^p = p \int_0^\infty t^{p-1} \lambda_{Tf}(t) dt \leqslant p \int_0^\infty t^{p-1} \lambda_{Tf_A}(t/2) dt + p \int_0^\infty t^{p-1} \lambda_{Tg_A}(t/2) dt = I_1 + I_2$$

Let $A = t$. From Chebyshev inequality and (A.9), we get

$$I_1 \lesssim \int_0^\infty t^{p-3} \|f_A\|_2^2 dt = 2 \int_0^\infty t^{p-3} \int_0^A \xi \lambda_f(\xi) d\xi dt \lesssim (2-p)^{-1} \int_0^\infty \xi^{p-1} \lambda_f(\xi) d\xi \lesssim (2-p)^{-1} \|f\|_p^p$$

For I_2, we use (A.10) (notice that g_A is continuous and piece-wise smooth)

$$I_2 \lesssim -\int_0^\infty t^{p-1} \int_0^\infty \frac{\xi}{t} \left(1 + \log^+ \frac{\xi}{t}\right) d\lambda_{g_A}(\xi) \lesssim$$

$$\|f\|_p^p + \int_0^\infty t^{p-1} \int_{2t}^\infty t^{-1} \Big(1 + \log^+((\tau - t)/t)\Big) \lambda_f(\tau) d\tau \lesssim \|f\|_p^p \int_0^1 \xi^{p-2} \left(1 + \log^+ \frac{1-\xi}{\xi}\right) d\xi$$

We have

$$\int_0^{1/2} \xi^{p-2}\left(1+\log^+\frac{1-\xi}{\xi}\right)d\xi \lesssim \int_2^\infty u^{-p}\log u du \lesssim \int_0^\infty e^{-\delta t} t dt \lesssim \delta^{-2}$$

with $\delta = p-1$.

□

Appendix B

Some OPUC results

B.1 Classical constructions and results

Here we collect some standard constructions and definitions for the reader's convenience.

We will call a function F to be Carathéodory if it is analytic in $\mathbb{D}$ and its real part is positive. Given a measure σ, we will denote the Carathéodory function given by the Schwarz transform of σ as

$$F(z) = \mathcal{S}(\sigma) = \int_{\mathbb{T}} \frac{e^{i\theta}+z}{e^{i\theta}-z} d\sigma(\theta). \tag{B.1}$$

This Carathéodory function satisfies $\operatorname{Re} F(e^{i\theta}) = \sigma'(\theta)$.

Verblunsky parameters of recursion have a special place in the theory far preceding the parameter-to-potential correspondence. It is clear that any measure will yield a set of Verblunsky parameters, but Verblunsky himself proved

Theorem B.1 (Verblunsky's Theorem)**.** *Let* $\{\alpha_j\}_{j=0}^{\infty}$ *be a sequence of numbers in* $\mathbb{D}$*. Then there is a unique measure* $d\mu$ *with* $\alpha_j(d\mu) = \alpha_j$

The proof of this Theorem made use of a construction called the Bernstein-Szegő approximation.

Theorem B.2 (Bernstein-Szegő Approximation)**.** *Let* $d\mu$ *be a nontrivial probability measure on* $\mathbb{T}$ *and let* $\{\alpha_j(d\mu)\}_{j=0}^{\infty}$ *and* $\{\phi_n(z; d\mu)\}_{j=0}^{\infty}$ *be its Verblunsky parameters and* *orthonormal polynomials. Then for each n,*

$$\int_{\mathbb{T}} \frac{1}{|\phi_n(e^{i\theta})|^2} \frac{d\theta}{2\pi} = 1$$

and the measure

$$d\mu_n = \frac{1}{|\phi_n(e^{i\theta})|^2} \frac{d\theta}{2\pi}$$

obeys

$$\alpha_j(d\mu_n) = \alpha_j(d\mu) \qquad j = 0, 1, \ldots, n-1$$

$$= 0 \qquad j \geq n$$

Moreover, $d\mu_n \to d\mu$ weakly as $n \to \infty$.

Notice that this final statement can also be viewed as a consequence of Theorem 1.4.

F is also approximated by the polynomials in a specific sense. We introduce an important class of measures, those of Aleksandrov.

Let $\lambda \in \mathbb{T}$. For a measure $d\mu$ with Verblunsky parameters $\{\alpha_k\}$, denote by $d\mu_\lambda$ the measure associated to the parameters $\{\lambda\alpha_k\}$. Let $F^{(\lambda)}$ be the Carathéodory function associated to $d\mu_\lambda$. Then

$$F^{(\lambda)}(z) = \frac{(1-\lambda) + (1+\lambda)F}{(1+\lambda) + (1-\lambda)F}$$

In particular,

$$F^{(-1)} = \frac{1}{F}$$

This hints that $\lambda = -1$ will play a special role. The polynomials associated to $d\mu_{-1}$ are the second-kind polynomials we have already seen in Chapter 4, and are denoted Ψ_n for monic and ψ_n for orthonormal. They appear in an important limit formula for $z \in \mathbb{D}$

$$F(z; d\mu) = \lim_{n\to\infty} \frac{\Psi_n^*(z; d\mu)}{\Phi_n^*(z; d\mu)}, \quad z \in \mathbb{D}$$

For an interesting exposition of this and similar limiting formulae in orthogonal polynomials, related to continued fractions, convergents and approximation theory, see [40].

B.2 Proof of the Decoupling Lemma

Proof (Lemma 4.4). First, notice that $\{\widetilde{\alpha}_j\} \in \ell^1$ by Baxter's Theorem (see, e.g., [65], Chapter 5). Therefore, σ is purely absolutely continuous by the same Baxter's criterion. Denote the orthonormal polynomials of the first/second kind corresponding to measure $\widetilde{\sigma}$ by $\{\widetilde{\phi}_j\}, \{\widetilde{\psi}_j\}$. Similarly, let $\{\phi_j\}, \{\psi_j\}$ be orthonormal polynomials for σ. Since by construction μ_n and σ have identical first n Verblunsky parameters, ϕ_n is n-th orthonormal polynomial for σ.

Let us compute the polynomials ϕ_j and ψ_j, orthonormal with respect to σ, for the indexes $j > n$. The recursion can be rewritten in the following matrix form

$$\begin{pmatrix} \phi_{n+m} & \psi_{n+m} \\ \phi^*_{n+m} & -\psi^*_{n+m} \end{pmatrix} = \begin{pmatrix} \mathcal{A}_m & \mathcal{B}_m \\ \mathcal{C}_m & \mathcal{D}_m \end{pmatrix} \begin{pmatrix} \phi_n & \psi_n \\ \phi^*_n & -\psi^*_n \end{pmatrix}, \tag{B.2}$$

where $\mathcal{A}_m, \mathcal{B}_m, \mathcal{C}_m, \mathcal{D}_m$ satisfy

$$\begin{pmatrix} \mathcal{A}_0 & \mathcal{B}_0 \\ \mathcal{C}_0 & \mathcal{D}_0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$\begin{pmatrix} \mathcal{A}_m & \mathcal{B}_m \\ \mathcal{C}_m & \mathcal{D}_m \end{pmatrix} = \frac{1}{\widetilde{\rho}_0 \cdot \ldots \cdot \widetilde{\rho}_{m-1}} \begin{pmatrix} z & -\overline{\widetilde{\alpha}}_{m-1} \\ -z\widetilde{\alpha}_{m-1} & 1 \end{pmatrix} \cdot \ldots \cdot \begin{pmatrix} z & -\overline{\widetilde{\alpha}}_0 \\ -z\widetilde{\alpha}_0 & 1 \end{pmatrix}$$

and thus depend only on $\widetilde{\alpha}_0, \ldots, \widetilde{\alpha}_{m-1}$. Moreover, we have

$$\begin{pmatrix} \widetilde{\phi}_m & \widetilde{\psi}_m \\ \widetilde{\phi}^*_m & -\widetilde{\psi}^*_m \end{pmatrix} = \begin{pmatrix} \mathcal{A}_m & \mathcal{B}_m \\ \mathcal{C}_m & \mathcal{D}_m \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Thus, $\mathcal{A}_m = (\widetilde{\phi}_m + \widetilde{\psi}_m)/2$, $\mathcal{B}_m = (\widetilde{\phi}_m - \widetilde{\psi}_m)/2$, $\mathcal{C}_m = (\widetilde{\phi}^*_m - \widetilde{\psi}^*_m)/2$, $\mathcal{D}_m = (\widetilde{\phi}^*_m + \widetilde{\psi}^*_m)/2$ and their substitution into (B.2) yields

$$2\phi^*_{n+m} = \phi_n(\widetilde{\phi}^*_m - \widetilde{\psi}^*_m) + \phi^*_n(\widetilde{\phi}^*_m + \widetilde{\psi}^*_m) = \widetilde{\phi}^*_m \left(\phi_n + \phi^*_n + \widetilde{F}_m(\phi^*_n - \phi_n)\right) \tag{B.3}$$

where

$$\widetilde{F}_m(z) = \frac{\widetilde{\psi}^*_m(z)}{\widetilde{\phi}^*_m(z)} \ .$$

Since $\{\widetilde{\alpha}_n\} \in \ell^1$ and $\{\alpha_n\} \in \ell^1$, we have ([65], p. 225)

$$\widetilde{F}_m \to \widetilde{F} \text{ as } m \to \infty \text{ and } \phi^*_j \to \Pi,\ \widetilde{\phi}^*_j \to \widetilde{\Pi} \text{ as } j \to \infty \ .$$

uniformly on $\overline{\mathbb{D}}$. The functions Π and $\widetilde{\Pi}$ are related to σ and $\widetilde{\sigma}$ as follows: they are the outer functions in $\mathbb{D}$ that satisfy

$$|\Pi|^{-2} = 2\pi\sigma', \quad |\widetilde{\Pi}|^{-2} = 2\pi\widetilde{\sigma}', \quad \Pi(0) > 0, \quad \widetilde{\Pi}(0) > 0 \tag{B.4}$$

on $\mathbb{T}$. In (B.3), send $m \to \infty$ to get

$$2\Pi = \widetilde{\Pi}\left(\phi_n + \phi^*_n + \widetilde{F}(\phi^*_n - \phi_n)\right) \tag{B.5}$$

and we have the Lemma after taking the square of absolute values and using (B.4). □

B.3 Localization principle

For a weight w on $\mathbb{T}$, define

$$\lambda(w) = \exp\left(\frac{1}{2}\int_{\mathbb{T}} \log(2\pi w)\frac{d\theta}{2\pi}\right), \quad \Lambda(w) = \sqrt{\|w\|_1}$$

The following Theorem was proved in [2].

Theorem B.3 ([2]). *Let $w_{1(2)}$ be two weights on $(-\pi, \pi]$ so that*

$$w_1(\theta) = w_2(\theta), \quad \theta \in [-\epsilon, \epsilon]$$

Then

$$\left|\frac{\phi_n(1; w_1)}{\phi_n(1; w_2)}\right| \leq \frac{\Lambda(w_2)}{\lambda(w_1)} + \frac{2\Lambda(w_1)}{\epsilon\lambda(w_1)} \left(\int_{|\theta|>\epsilon} |\phi_n(e^{i\theta}; w_1)\phi_n(e^{i\theta}; w_2)|(w_1 + w_2)d\theta\right)$$

And we get the localization principle as a corollary.

Lemma B.4. *(Localization principle [21]) Assume that*

$$0 < m_1 \leq w_{1(2)}(e^{i\theta}) \leq m_2, \quad \theta \in [-\pi, \pi)$$

and

$$w_1(e^{i\theta}) = w_2(e^{i\theta}), \quad \theta \in [-\epsilon, \epsilon]$$

Then

$$\frac{\epsilon m_1}{m_2} \lesssim \left|\frac{\phi_n(1; w_1)}{\phi_n(1; w_2)}\right| \lesssim \frac{m_2}{\epsilon m_1}$$

The following (quick) proof can be found in [21].

Proof. Notice that

$$\int_{-\pi}^{\pi} |\phi_n(e^{i\theta}, w_1)\phi_n(e^{i\theta}, w_2)|w_1 d\theta \leq \left(\int_{-\pi}^{\pi} |\phi_n(e^{i\theta}, w_2)|^2 w_1 d\theta\right)^{1/2}$$

by Cauchy-Schwarz and normalization. Since

$$1 = \int_{-\pi}^{\pi} |\phi_n(e^{-\theta}, w_2)|^2 w_2 d\theta \geq m_1 \int_{-\pi}^{\pi} |\phi_n(e^{i\theta}, w_2)|^2 d\theta$$

we see

$$\int_{-\pi}^{\pi} |\phi_n(e^{i\theta}, w_1)\phi_n(e^{i\theta}, w_2)|w_1 d\theta \leq \left(\frac{m_2}{m_1}\right)^{1/2}$$

Switching the roles of m_1 and m_2 we get the Lemma.

□

B.4 Denisov's argument in case of small deviation

The argument which comprises this section is due to Denisov in [21]. It is included for the reader's convenience and in order that the proof of Lemma 2.9 be contained within this thesis.

Theorem B.5 (Denisov)**.** *For $\epsilon < 1$,*

$$\sup_{1\leq w\leq 1+\epsilon} \|\Phi_n(e^{i\theta}; w)\|_\infty > C(\epsilon)n^{c\epsilon}$$

for $n > n_0(\epsilon)$.

See [21]. We will follow the strategy used in [2], constructing a large polynomial with particular properties and decoupling the problem. Below, we introduce $\widetilde{F}$, ϕ_n^* and check that they satisfy conditions of Lemma 4.4. Then, we will control the deviation of the weight from the constant. Notice that it is sufficient to consider $\epsilon \in (0, \epsilon_0)$ where ϵ_0 is sufficiently small.

(a) Consider auxiliary function H_n given by $H_n = 2(1-z)^\epsilon * \mathcal{F}_n$. Clearly H_n is a polynomial of degree $n-1$. Since $\mathrm{Re}(1-z)^\epsilon \geq 0, z \in \mathbb{T}$ and $\mathcal{F}_n \geq 0$, its real part is strictly positive over $\mathbb{T}$ so H_n is zero-free in $\overline{\mathbb{D}}$. In Lemmas B.6 and B.9, more detailed information is obtained, in particular,

$$|\arg H_n| \leqslant C\epsilon \tag{B.6}$$

uniformly in $\theta \in [-\pi, \pi)$ and $n > n(\epsilon)$.

Because $\int_{-\pi}^{\pi} \mathcal{F}_n d\theta = 1$, we get $\int_{-\pi}^{\pi} H_n d\theta = 2\pi H(0) = 2(2\pi)$. Since $\mathcal{F}_n$ is even and $\mathrm{Im}(1-z)^\epsilon, z \in [-\pi, \pi]$ is odd, $\mathrm{Im}\, H_n$ is odd on $[-\pi, \pi]$ also.

(b) In Lemma B.9, take

$$\widetilde{F} = 2H_n^{-1} \tag{B.7}$$

From the properties of H_n, we get analyticity of $\widetilde{F}$ in $\mathbb{D}$ and infinite smoothness on the boundary. Since

$$\operatorname{Re}\widetilde{F} = 2\frac{\operatorname{Re} H_n}{|H_n|^2}, \tag{B.8}$$

$\widetilde{F}$ is Carathéodory function. Moreover, since $H_n(0) = 2$, we get the normalization $\operatorname{Re}\widetilde{F}(0) = 1$.

From (B.6) and (B.8), we also have

$$1 \leqslant \operatorname{Re} H_n \cdot \operatorname{Re}\widetilde{F} \leqslant 2 \tag{B.9}$$

uniformly in $\theta \in [-\pi, \pi), n > n(\epsilon)$.

(c) Let Q_n be the analytic polynomial of degree $n-1$ with no zeroes in $\mathbb{D}$, which satisfies

$$|Q_n|^2 = \left(\operatorname{Re}\widetilde{F}\right) * \mathcal{F}_n, \quad z \in \mathbb{T} \tag{B.10}$$

and $Q_n(0) > 0$.

We let $\phi_{2n}^* = \alpha_n(Q_n + Q_n^* + Q_n H_n)$ where α_n is the normalization constant chosen to ensure that $\|\phi_{2n}^{-1}\|^2_{L^2(\mathbb{T})} = 2\pi$. Here Q_n^* is understood as $(*)$–operation of order $2n$ acted on Q_n. Thus $\deg \phi_n^* \leqslant 2n$.

To make sure that ϕ_{2n}^* has no zeroes in $\overline{\mathbb{D}}$, we write

$$Q_n + Q_n^* + Q_n H_n = Q_n\left(1 + H_n + \frac{Q_n^*}{Q_n}\right).$$

Then, Q_n is zero-free and $1 + H_n + Q_n^*/Q_n$ is analytic in $\mathbb{D}$ having the positive real part since

$$\operatorname{Re} H_n > 0, \quad \operatorname{Re}\left(1 + \frac{Q_n^*}{Q_n}\right) \geqslant 0, \quad z \in \mathbb{T}.$$

The last inequality is the consequence of $|Q_n^*| = |Q_n|, z \in \mathbb{T}$.

We will show that

$$\alpha_n \sim 1 \tag{B.11}$$

uniformly in $n > n(\epsilon)$. Indeed,

$$|Q_n + Q_n^* + Q_n H_n|^2 = \left| Q_n \left(1 + H_n + \frac{Q_n^*}{Q_n} \right) \right|^2 \geqslant |Q_n|^2 (\operatorname{Re} H_n)^2 = (\operatorname{Re} H_n)^2 (\operatorname{Re} \widetilde{F} * \mathcal{F}_n)\,.$$

Let us rewrite the last expression using (B.8)

$$(\operatorname{Re} H_n)^2 (\operatorname{Re} \widetilde{F} * \mathcal{F}_n) = (\operatorname{Re} H_n)^2 \operatorname{Re} \widetilde{F} \cdot \frac{(\operatorname{Re} \widetilde{F} * \mathcal{F}_n)}{\operatorname{Re} \widetilde{F}} = 2 \frac{(\operatorname{Re} H_n)^3}{|H_n|^2} \cdot \frac{(\operatorname{Re} \widetilde{F} * \mathcal{F}_n)}{\operatorname{Re} \widetilde{F}}\,.$$

The second factor has absolute value uniformly bounded above and below in $\theta \in [-\pi, \pi)$ and $n > n(\epsilon)$. This is due to Lemma B.8. From (B.6), we get

$$\frac{(\operatorname{Re} H_n)^2}{|H_n|^2} \sim 1\,.$$

Thus, $|Q_n + Q_n^* + Q_n H_n|^2 \geqslant C \operatorname{Re} H_n$ uniformly in $\theta \in [-\pi, \pi), n > n(\epsilon)$. Therefore, Lemmas B.9, B.6, B.8, and (B.9) give

$$|\theta|^\epsilon \lesssim |Q_n + Q_n^* + Q_n H_n|^2 \lesssim |Q_n|^2 = \left(\operatorname{Re} \widetilde{F} \right) * \mathcal{F}_n \lesssim |\theta|^{-\epsilon}$$

for $\theta \in [-\pi, \pi)$, uniformly in $n > n(\epsilon)$. Therefore, we have (B.11).

We can apply the Decoupling Lemma 4.4 now with $2n$ taken instead of n. Consider the value of ϕ_{2n}^* at $z = 1$. Since $|Q_n(e^{i\theta})|$ is even and $Q_n(0) > 0$, we have $Q_n(1) \in \mathbb{R}$. Thus, $Q_n^*(1) = Q_n(1)$ and

$$|\phi_{2n}^*(1, \sigma)|^2 = \alpha_n^2 |Q_n(1)(2 + H_n(0))|^2 \sim |\operatorname{Re} \widetilde{F} * \mathcal{F}_n|_{z=1} \geqslant$$

$$|\operatorname{Re} \widetilde{F}|_{z=1} - |\operatorname{Re} \widetilde{F}|_{z=1} \left| \frac{\operatorname{Re} \widetilde{F} * \mathcal{F}_n}{\operatorname{Re} \widetilde{F}} - 1 \right|_{z=1} \gtrsim \widetilde{F}|_{z=1}$$

by Lemma B.8. Now, (B.7) and Lemma B.9 yield

$$|\phi_n^*(1,\sigma)|^2 \gtrsim n^\epsilon \,.$$

For the weight of orthogonality, we have by the Decoupling Lemma

$$\sigma'^{-1} = \upsilon_n \mathcal{A}\mathcal{B}\mathcal{C}\,,$$

where υ_n is n–dependent parameter, $\upsilon_n \sim 1$, and

$$\mathcal{A} = \frac{|Q_n|^2}{\operatorname{Re}\widetilde{F}}, \quad \mathcal{B} = |2+\overline{H}_n(1-\widetilde{F})|^2, \quad \mathcal{C} = \left|\xi + \frac{2+H_n(1+\widetilde{F})}{2+\overline{H}_n(1-\widetilde{F})}\right|^2, \quad \xi = e^{i(2n\theta-2\Theta)}\,, \tag{B.12}$$

where $\Theta = \arg Q_n$.

Now, in what follows, we will consider a small ϵ–dependent interval I_ϵ around $\theta = 0$ and will control how each of the factors $\mathcal{A}$ and $(\mathcal{B}\mathcal{C})$ deviates from constants on I_ϵ.

By Lemma B.8,

$$|\mathcal{A} - 1| \lesssim \epsilon \tag{B.13}$$

uniformly in $\theta \in [-\pi,\pi)$ and $n > n(\epsilon)$.

For $\mathcal{B}$, we can write

$$\mathcal{B} = |2+\overline{H}_n(1-\widetilde{F})|^2 = \left|2\left(1-\frac{\overline{H}_n}{H_n}\right)+\overline{H}_n\right|^2 \lesssim \epsilon^2 + |H_n|^2 \lesssim \epsilon^2$$

on some I_ϵ (due to estimates on H_n obtained in Lemma B.6), provided that $n > n(\epsilon)$. Consider

$$\mathcal{J} = \frac{2+H_n(1+\widetilde{F})}{2+\overline{H}_n(1-\widetilde{F})}\,.$$

Then, recalling (B.7) and Lemma B.6,

$$\mathcal{B}\mathcal{C} = \mathcal{B}(1+|\mathcal{J}|^2 + 2\operatorname{Re}(\xi\bar{\mathcal{J}})) = \mathcal{B} + \mathcal{B}|\mathcal{J}|^2 + 2\mathcal{B}\operatorname{Re}(\xi\bar{\mathcal{J}}) \tag{B.14}$$

The first term is at most $C\epsilon^2$. For the third one, we have

$$|2\mathcal{B}\,\mathrm{Re}(\xi\overline{\mathcal{J}})| \lesssim \sqrt{\mathcal{B}}|2+H_n(1+\widetilde{F})| \lesssim \epsilon|2+H_n(1+2H_n^{-1})| \lesssim \epsilon\,.$$

By (B.8), the second term in (B.14) can be rewritten as

$$|2+H_n(1+\widetilde{F})|^2 = |4+H_n|^2 = 16+8\,\mathrm{Re}\,H_n+|H_n|^2 = 16+O(\epsilon)$$

over appropriate I_ϵ for $n > n(\epsilon)$. To summarize, we have

$$\sigma' = \omega_n\,(1+O(\epsilon))$$

uniformly over some I_ϵ and $n > n(\epsilon)$, positive ω_n depends on n only and $\omega_n \sim 1$.

To apply Lemma B.4 later, we have to check that σ' satisfies the global lower and upper bounds, i.e., we need to verify that

$$\mathcal{A}\mathcal{B}\mathcal{C} \sim 1$$

uniformly in θ provided that $n > n(\epsilon)$ and ϵ is small. Indeed, for $\mathcal{A}$, we use (B.13). To control $\mathcal{B}\mathcal{C}$ we argue differently:

$$\mathcal{B}\mathcal{C} = \left|4+H_n+\xi\left(\overline{H}_n+2\left(1-\frac{\overline{H}_n}{H_n}\right)\right)\right|^2\,. \tag{B.15}$$

It is clear that $\mathcal{B}\mathcal{C} \lesssim 1$. For the lower bound, we can use

$$\left|1-\frac{\overline{H}_n}{H_n}\right| \lesssim \epsilon$$

to get

$$4+H_n+\xi\left(\overline{H}_n+2\left(1-\frac{\overline{H}_n}{H_n}\right)\right) = 4+H_n+O(\epsilon)+\xi\left(\frac{\overline{H}_n}{H_n}-1\right)H_n+\xi H_n$$

$$= 4 + (\xi + 1)H_n + O(\epsilon)\,.$$

Since $\arg(1+\xi) \in [-\pi/2, \pi/2]$ and $|\arg H_n| < C\epsilon$, we have

$$|4 + (\xi + 1)H_n + O(\epsilon)| > 2$$

for small ϵ and $n > n(\epsilon)$. Therefore, we also have $\sigma' \sim 1$ uniformly in $\theta \in [-\pi, \pi)$ and $n > n(\epsilon)$ provided that $\epsilon < \epsilon_0$.

We constructed a probability measure σ given by weight σ' which is bounded above and below uniformly in n, it deviates from the constant by at most $C\epsilon$ on the interval I_ϵ centered at $\theta = 0$. Moreover,

$$|\phi_{2n}(1, \sigma)| \gtrsim n^{\epsilon/2}\,.$$

Finally, we are to use localization principle to make the deviation of the weight from the constant smaller than $C\epsilon$ over the whole $\mathbb{T}$. To that end, consider w_1:

$$w_1 = \begin{cases} 1, & \theta \notin I_\epsilon \\ \dfrac{\sigma'}{\min\limits_{I_\epsilon} \sigma'}, & \theta \in I_\epsilon \end{cases}\,.$$

We have $1 \leqslant w_1 \leqslant 1 + O(\epsilon)$ uniformly over $\mathbb{T}$. Lemma B.4 gives

$$|\phi_{2n}(1, w_1)| > C(\epsilon)|\phi_{2n}(1, \sigma)| > C(\epsilon)n^{\epsilon/2}, \quad n > n(\epsilon)\,.$$

□

B.5 Quantified properties of auxiliary polynomials in Denisov's argument

Lemma B.6. *Let*

$$H_n = 2(1-z)^{\epsilon} * \mathcal{F}_n$$

for $\mathcal{F}_n$ the Fejér kernel of order n. Then there exist ϵ_0 so that for all $\epsilon \leq \epsilon_0$, there is $n_0(\epsilon) = \exp\left(\frac{C}{\epsilon}\right)$ so that

$$|\arg H_n(e^{i\theta})| \lesssim \epsilon$$

for $\theta \in [-\pi, \pi)$ and $n > n_0$. Further,

$$|H_n(e^{i\theta})| \sim n^{-\epsilon} + |\theta|^{\epsilon}$$

for all $\theta \in [-\pi, \pi)$ and $n > n_0$

Proof. This proof follows the proof of Lemma 6.2 in [21]. Note by Lemma B.9 it suffices to prove this for $|x| < \frac{\pi}{4}$. For such x,

$$H_{2n}(x) = C(\cos(\epsilon\pi/2)H_n^{(1)}(x) - i\sin(\epsilon\pi/2)H_n^{(2)}(x))$$

We deal with the real part first. We have defined

$$H_n^{(1)}(x) = n^{-1}\int_{-\pi}^{\pi} |\theta|^{\epsilon}\frac{\sin^2(n(x-\theta)/2)}{(x-\theta)^2}d\theta + \epsilon_n(x)$$

We will handle $\epsilon_n(x)$ at the end. Let

$$g(t) = \frac{\sin^2(t/2)}{t^2}$$

For the first term, if $|x| < \pi/4$,

$$= n^{-1}\int_{-\pi}^{\pi} |\theta|^{\epsilon}\frac{\sin^2(n(x-\theta))}{(x-\theta)^2}d\theta = n^{-\epsilon}M_n(nx), \quad M_n(\hat{x}) = \int_{-\pi n}^{\pi n} |t|^{\epsilon} g(\hat{\theta} - t)dt$$

For $\hat{x} \in [-\pi n, \pi n]$, estimating the integral yields uniformly

$$M_n(\hat{x}) \sim 1 + |\hat{x}|^\epsilon$$

For the imaginary part,

$$H_n^{(2)}(x) = n^{-1} \int_{-\pi}^{\pi} |\theta|^\epsilon \mathrm{sgn}(\theta) \frac{\sin^2(n(x-\theta))}{(x-\theta)^2} d\theta = n^{-\epsilon} N_n(nx) + \epsilon_n(x)$$

where N_n is odd and satisfies $|N_n(\hat{x})| \leq M_n(\hat{x})$ for all $\hat{x} \in [-\pi n, \pi n]$.

Finally we control ϵ_n. For both real and imaginary parts (see [21]),

$$|\epsilon_n(x)| \leq C(n^{-1} + \epsilon |x|^{1+\epsilon} + \epsilon n^{-1} \log n)$$

Therefore we may take $n \gtrsim \frac{1}{\epsilon}$, and we have the Lemma.

□

We will find quantitative control over $M_n'(t)$ to be useful.

Lemma B.7.

$$|M_n'(t)| \lesssim n^{\epsilon-2} + \epsilon(1+|t|)^{\epsilon-1}$$

Proof.

$$|M_n'(t)| = \left| \int_{-\pi n}^{\pi n} |s|^\epsilon g'(s-t) ds \right|$$

$$\lesssim n^\epsilon \left(\frac{1}{(n\pi - t)^2 + 1} + \frac{1}{(n\pi + t)^2 + 1} \right) + \epsilon \int_{-\pi n}^{\pi n} |s|^{\epsilon-1} \mathrm{sgn} s \cdot g(s-t) ds$$

For the integral above, we write

$$\int_{-\pi n}^{\pi n} |s|^{\epsilon-1} \mathrm{sgn} s \cdot g(s-t) ds = \int_0^1 s^{\epsilon-1}(g(s-t) - g(-s-t)) ds + \int_{1<|s|\leq n\pi} |s|^{\epsilon-1} \mathrm{sgn} s \cdot g(s-t) ds$$

Applying the Mean Value Theorem to the first integrand, we have

$$\left| \int_0^1 s^{\epsilon-1}(g(s-t) - g(-s-t)) ds \right| \lesssim (1+t^2)^{-1}$$

and we have

$$\left|\int_{1<|s|\leq n\pi} |s|^{\epsilon-1}\mathrm{sgn}s \cdot g(s-t)ds\right| \lesssim \int_{\mathbb{R}} (1+|s|)^{\epsilon-1}\frac{1}{(t-s)^2+1}ds \lesssim (1+|t|)^{\epsilon-1}$$

Putting these estimates together yields the Lemma. □

Lemma B.8. *Let*

$$\widetilde{F} = 2H_n^{-1}$$

Then there is ϵ_0 *so that for every* $0 < \epsilon \leq \epsilon_0$, *there is* $n_0(\epsilon)$ *so that*

$$\left|\frac{\mathrm{Re}\,\widetilde{F_n} * \mathcal{F}_n}{\mathrm{Re}\,\widetilde{F_n}} - 1\right| \lesssim \epsilon$$

uniformly for $\theta \in [-\pi, \pi), n > n_0(\epsilon)$. *Moreover we may take*

$$n_0(\epsilon) = \exp\left(\frac{C}{\epsilon}\right)$$

Proof. This proof will follow the proof of Lemma 6.4 of [21]. We merely keep track of the dependencies between n and ϵ.

Fix ϵ_0 small. Notice first that uniformly in this ϵ, $\mathrm{Re}\,\widetilde{F} \to \mathrm{Re}(1-z)^{-\epsilon}$ and $\mathrm{Re}\,\widetilde{F}*\mathcal{F}_n \to \mathrm{Re}(1-z)^{-\epsilon}$ uniformly in $\pi/2 \leq |\theta| \leq \pi$, since the smoothness of $(1-z)^{\epsilon}$ in this region is uniform as $\epsilon \to 0$. Let $x : |x| \leq \pi/4$. Then we have

$$\left|\frac{\mathrm{Re}\,\widetilde{F} * \mathcal{F}_n}{\mathrm{Re}\,\widetilde{F}} - 1\right| \lesssim |H_n(\theta)| \left|\int_{-\pi}^{\pi} g(nx)\left(H_n^{-1}(\theta - x) - H_n^{-1}(\theta)\right)dx\right|$$

$$= \left|\int_{-\pi}^{\pi} g(nx)\frac{H_n(\theta-x) - H_n(\theta)}{H_n(\theta - x)}dx\right|$$

Inserting the asymptotics we know for H_n by Lemma B.6 we may bound

$$\lesssim \int_{-\pi}^{\pi} \frac{\sin^2(nx/2)}{nx^2}\left|\frac{M_n(n(\theta-x)) - M_n(n\theta)}{M_n(n(\theta-x))}\right|dx + \epsilon\int_{-\pi}^{\pi}\frac{\sin^2(nx/2)}{nx^2}\frac{|M_n(n(\theta-x)| + |M_n(n\theta)|}{M_n(n(\theta-x))}dx$$

$$+Cn^{-1}\int_{-\pi}^{\pi}\left|\frac{H_n(\theta-x)-H_n(\theta)}{H_n(\theta-x)}\right|dx$$

By Lemma B.6 the last term is bounded by

$$Cn^{-1}\int_{-\pi}^{\pi}\frac{1}{|\theta-x|^{\epsilon}}dx \lesssim n^{-1}$$

The second term we may bound via the estimate $M_n \sim 1+|x|^{\epsilon}$ by

$$\epsilon\int_{-\pi n}^{\pi n}\frac{\sin^2(\hat{x}/2)}{\hat{x}^2}\frac{1+|\hat{\theta}-\hat{x}|^{\epsilon}+|\hat{\theta}|^{\epsilon}}{1+|\hat{\theta}-\hat{x}|^{\epsilon}}d\hat{x} \lesssim \epsilon$$

For the first term, split the integral in two:

$$\int_{-\pi}^{\pi}\frac{\sin^2(nx/2)}{nx^2}\left|\frac{M_n(n(\theta-x))-M_n(n\theta)}{M_n(n(\theta-x))}\right|dx$$

$$\lesssim \int_{|\hat{x}<\log n}\frac{\sin^2(\hat{x}/2)}{\hat{x}^2}\left|\frac{M_n(\hat{\theta}-\hat{x})-M_n(\hat{\theta}}{M_n(\hat{\theta}-\hat{x})}\right|d\hat{x}+\int_{|\hat{x}|>\log n}\frac{\sin^2(\hat{x}/2)}{\hat{x}^2}\left|\frac{M_n(\hat{\theta}-\hat{x})-M_n(\hat{\theta})}{M_n(\hat{\theta}-\hat{x})}\right|d\hat{x}$$

The last integral can be bounded by

$$\int_{|\hat{x}|>\log n}|\hat{x}|^{-2+\epsilon}d\hat{x}=O(\log^{-1+\epsilon}n)$$

For the first integral, use the estimate for the derivative of M_n in Lemma B.7. This yields

$$|M_n(\hat{\theta}-\hat{x})-M_n(\hat{\theta})| \lesssim n^{\epsilon-2}|\hat{x}|+\left|\int_{\hat{\theta}}^{\hat{\theta}-\hat{x}}\epsilon(1+|\xi|)^{\epsilon-1}d\xi\right| \lesssim n^{\epsilon-2}\log n+\epsilon(1+|\hat{x}|^{0.1})$$

by Hölder's inequality, for ϵ sufficiently small.

Putting together the estimates we have shown yields the Lemma with $n_0(\epsilon) = \exp(\epsilon^{-1})$ as claimed.

□

Lemma B.9. *For $\delta > 0$, we may take $n_0(\delta)$, $\epsilon_0(\delta)$ so that*

$$|H_n(z) - 2(1-z)^\epsilon| \ll 1$$

$$|\widetilde{F_n}(z) - (1-z)^{-\epsilon}| \ll 1$$

for $n > n_0$, $\epsilon < \epsilon_0$, uniformly in $\{z : z = e^{i\theta}, |\theta| \geq \delta\}$

Proof. This is immediate because $\mathcal{F}_n$ is an approximation of the identity.

□

Given these properties we may prove Lemma 2.9.

Proof (Lemma 2.9). We examine the proof of Theorem 3.1 in [21] in the case of small deviation, reproduced in section B.4. The proof of this Theorem constructs a weight w which deviates from 1 by at most ϵ on $\mathbb{T}$ such that the orthonormal polynomials obey

$$|\phi_n(1; w)| \gtrsim C(\epsilon) n^{c\epsilon}$$

for $n = n_0(\epsilon)$.

To prove Lemma 2.9, we need to take n sufficiently large to kill this $C(\epsilon)$; it is here we pay our steep price.

The $C(\epsilon)$ above is a consequence of the localization principle. In the proof in B.4, we estimate 3 quantities which do not vary much from constant on some interval I_ϵ. The price we need to pay to localize is then $|I_\epsilon| \sim C(\epsilon)$ and we must inquire about the size of this constant.

We rely essentially on two estimates. For Q_n the Fejér-Riesz factorization of $\operatorname{Re} \widetilde{F_n} * \mathcal{F}_n$, that is,

$$|Q_n|^2 = \operatorname{Re} \widetilde{F} * \mathcal{F}_n$$

we need

1.

$$\left| \frac{|Q_n|^2}{\operatorname{Re} \widetilde{F}} - 1 \right| \leq C\epsilon$$

2.

$$|H_n| \leq C\epsilon$$

By Lemma B.8, item 1 in fact can be satisfied uniformly in θ if $n = \exp\left(\frac{C}{\epsilon}\right)$ and so it cannot hurt us.

Lemma B.6 tells us we must take the interval I_ϵ sufficiently small and n sufficiently large that

$$n^{-\epsilon} + |\theta|^\epsilon \leq C\epsilon$$

That is,

$$|\theta| \leq \epsilon^{\frac{1}{\epsilon}}, \quad n \geq \exp\left(\frac{|\ln \epsilon|}{\epsilon}\right)$$

So for fixed ϵ, we may take n such that

$$|\phi_n(1;w)| \gtrsim \epsilon^{\frac{1}{\epsilon}} n^{c\epsilon}$$

if we enforce $n \geq \exp\left(\frac{|\ln \epsilon|}{\epsilon}\right)$. We also need n sufficiently large so that

$$\epsilon^{\frac{1}{\epsilon}} n^{c_1 \epsilon} \gtrsim n^{c_2 \epsilon}$$

for some $0 < c_2 < c_1$.

Solving this inequality for n we must take

$$n \geq \exp\left(\frac{C|\log \epsilon|}{\epsilon^2}\right)$$

and this is the quantity which appears in Lemma 2.9.

We have seen that the weight w may be chosen to deviate at most ϵ from constant and have polynomials $|\phi_n(1;w)| \gtrsim n^{c\epsilon}$ for $n \geq \exp\left(\frac{C|\ln \epsilon|}{\epsilon^2}\right) = n_0(\epsilon)$, but we have not yet seen the regularity claimed in properties (2) and (3) of Lemma 2.9. Let

$$\mathcal{VP}_n = 2\mathcal{F}_{2n} - \mathcal{F}_n \tag{B.16}$$

where

$$\mathcal{F}_n = \frac{1}{n}\left(\frac{\sin(n\theta/2)}{\sin(\theta/2)}\right)^2 \tag{B.17}$$

so that $\mathcal{F}_n$ is the Féjer kernel. We call $\mathcal{VP}_n$ the de la Vallée-Poussin kernel of order n. Notice the Fourier transform of $\mathcal{VP}_n$ is constant 1 on modes in $[-n, n]$. Then, for $n \geq n_0(\epsilon)$ fixed, let

$$\tilde{w} = \mathcal{VP}_n * w$$

Then $\tilde{w}$ has the same first n moments as w, and therefore its polynomials agree for indices $j \leq n$; that is

$$\Phi_j(z; w) = \Phi_j(z; \tilde{w}), \quad j \leq n$$

So $\tilde{w}$ satisfies conclusion (1) of Lemma 2.9.

We estimate the smoothness of $\tilde{w}$. Write $w = w - 1 + 1$. Then

$$\mathcal{VP}_n * w = \mathcal{VP}_n * (w - 1) + 1$$

Since $\|\mathcal{VP}_n\|_1 \leq 3$ by (B.16), $\|\mathcal{VP}_n * (w-1)\|_\infty \leq 3\epsilon$. Up to the factor of 3 this is property (2) in Lemma 2.9 for the weight $\tilde{w}$. Further, $\mathcal{VP}_n * (w-1)$ is a trigonometric polynomial so we may apply Bernstein's inequality to say

$$\left\| \frac{d}{d\theta}\left[\mathcal{VP}_n * (w-1)\right]\right\|_\infty \leq 3n\epsilon \tag{B.18}$$

Since $\frac{d}{d\theta}1 = 0$, (B.18) yields property (3) in Lemma 2.9 with a factor of 3. Replacing ϵ with $\epsilon/3$ proves the Lemma. $\square$

Appendix C

Riemann-Hilbert analysis of weight with jump–proof of Lemma 4.5

C.1 Setup

We wish to analyze the asymptotics of the orthogonal polynomials with respect to the weight

$$f(z) = e^{\epsilon} g_{i,-\frac{i\epsilon}{\pi}}(z) g_{-i,\frac{i\epsilon}{\pi}}(z),$$

where $z = |z|e^{i\theta}, \theta \in [0, 2\pi)$. For $z_j = e^{i\theta_j}$,

$$g_{z_j,\beta_j} = \begin{cases} e^{i\pi\beta_j} & : 0 \leq \arg z < \theta_j \\ e^{-i\pi\beta_j} & : \theta_j \leq \arg z < 2\pi \end{cases}$$

as defined in the main text.

This f belongs to the class of Fisher-Hartwig weights considered in [18] (see also [30] for the weight on the real line), so much of this Appendix consists of examining the results of [18]. Recall the three properties of these polynomials we need:

1.

$$|\Phi_n^*(z)| \sim 1,$$

2.

$$\|\Phi_n^*\Psi_n^* + z\Phi_n\Psi_n\|_{L^\infty(\mathbb{T})} >_\epsilon \ln n,$$

3.

$$\frac{\Psi_n^*(z)}{\Phi_n^*(z)} + \frac{\Psi_n^*(-z)}{\Phi_n^*(-z)} = O(1),$$

for $z \in \mathbb{T}, n > n_0(\epsilon)$.

As was noted in [5] and [36], the orthogonal polynomials of the first and second kinds satisfy a particular Riemann-Hilbert problem with contour $C = \mathbb{T}$. If Y is defined by

$$Y(z) = \begin{pmatrix} \Phi_n(z) & \int_{\mathbb{T}} \frac{\Phi_n(\xi)}{\xi - z} \frac{f(\xi)d\xi}{2\pi i \xi^n} \\ -\Phi_{n-1}^*(z) & -\int_{\mathbb{T}} \frac{\overline{\Phi_{n-1}}(\xi)}{\xi - z} \frac{f(\xi)d\xi}{2\pi i \xi} \end{pmatrix} \tag{C.1}$$

then it satisfies the following Riemann-Hilbert problem:

- $Y(z)$ is analytic in $\mathbb{C}\backslash\mathbb{T}$.
- For $z \in \mathbb{T}\backslash\{i, -i\}$, Y has continuous boundary values $Y_+(z)$ as z approaches $\mathbb{T}$ from the inside, and $Y_-(z)$ from the outside, related by the jump condition

$$Y_+(z) = Y_-(z) \begin{pmatrix} 1 & z^{-n}f(z) \\ 0 & 1 \end{pmatrix}.$$

- $Y(z)$ has the following asymptotic behavior at infinity:

$$Y(z) = \left(I + O\left(\frac{1}{z}\right)\right) \begin{pmatrix} z^n & 0 \\ 0 & z^{-n} \end{pmatrix}.$$

- As $z \to \pm i, z \in \mathbb{C}\backslash\mathbb{T}$,

$$Y(z) = \begin{pmatrix} O(1) & O(\ln|z - \pm i|) \\ O(1) & O(\ln|z - \pm i|) \end{pmatrix}.$$

In what follows, we consider n to be a sufficiently large but fixed parameter.

The analysis of Riemann-Hilbert problems of this type proceeds by enclosing the singularity points z_j by small but fixed disks U_{z_j}, enclosing the curve $C = \mathbb{T}$ by lenses meeting in these disks, solving the Riemann-Hilbert problems induced in these regions, and finally stitching them back together. For Fisher-Hartwig singularities this was performed in [18]. Since two of our estimates must be uniform over $z \in \mathbb{T}$, we will be concerned with the formulas that result in the regions enclosed by U_{z_j} as well as the the regions outside U_{z_j} but inside the lenses.

Before we begin the analysis of the Riemann-Hilbert problem, we list a few useful identities and notations. Following [18], our complex logarithm will be cut at the negative real axis unless otherwise noted.

$d\mu = \dfrac{f}{2\pi\cosh\epsilon}d\theta$ is a probability measure. Therefore one has (see [65], equation (3.2.53)):

$$\mathcal{S}(\mu)\Phi^*_{n-1}(z) - \Psi^*_{n-1}(z) = z^{n-1}\int_{\mathbb{T}} \frac{e^{i\theta}+z}{e^{i\theta}-z}\overline{\Phi_{n-1}(e^{i\theta})}d\mu$$
$$= z^{n-1}\int_{\mathbb{T}} \frac{e^{i\theta}-z+2z}{e^{i\theta}-z}\overline{\Phi_{n-1}(e^{i\theta})}d\mu = 0 + 2z^n\int_{\mathbb{T}} \frac{\overline{\Phi_{n-1}}(\xi)}{\xi - z}\frac{f(\xi)d\xi}{i2\pi\xi\cosh\epsilon}$$

so

$$-Y_{22}(z) = \int_{\mathbb{T}} \frac{\overline{\Phi_{n-1}}(\xi)}{\xi - z}\frac{f(\xi)d\xi}{2\pi i\xi} = \frac{1}{2z^n}\left(F(z)\Phi^*_{n-1}(z) - (\cosh\epsilon)\Psi^*_{n-1}(z)\right), \tag{C.2}$$

where F is the Carathéodory function associated to $f(\theta)\frac{d\theta}{2\pi}$ (i.e. $F = \mathcal{S}(f/2\pi)$), and Ψ_n is the second kind polynomial associated to Φ_n.

The exact expression for our F is easy to compute:

$$F(z) = \int_{\mathbb{T}} \frac{e^{i\theta}+z}{e^{i\theta}-z} f(\theta) \frac{d\theta}{2\pi}$$
$$= -\frac{i(e^{\epsilon}-e^{-\epsilon})}{\pi} \ln\left(\frac{i-z}{i+z}\right) + \frac{1}{2}(e^{-\epsilon}+e^{\epsilon}). \tag{C.3}$$

We will use the following notation from [18]:

$$\beta_1 = -i\epsilon/\pi, \quad \alpha_1 = 0, \quad z_1 = i$$

$$\beta_2 = i\epsilon/\pi, \quad \alpha_2 = 0, \quad z_2 = -i$$

Before we discuss the results obtained in [18] we introduce a confluent hypergeometric function $\psi(a,b,\zeta)$ which plays the key role in the analysis of the Riemann-Hilbert problem. We will have to use many facts about ψ. For this purpose we refer the reader to the National Institute of Standards and Technology's Digital Library of Mathematical Functions [57] and the appendix of [37].

The function $\psi(a,b,\zeta)$ is the confluent hypergeometric function of the second kind, often written as $U(a,b,\zeta)$. It is defined as the unique solution to Kummer's equation

$$\zeta\frac{d^2w}{d\zeta^2} + (b-\zeta)\frac{dw}{d\zeta} - aw = 0,$$

satisfying $w(a,b,\zeta) \sim \zeta^{-a}$ as $\zeta \to \infty$. We will be interested in the following choices of the parameters: $b = 1$ and $a = \{\beta_j, 1+\beta_j\}, j = 1,2$. The function ψ is analytic in ζ on the universal cover of $\mathbb{C}\backslash\{0\}$ and can be represented by the series (formula 13.2.9 in [57]):

$$\psi(a,1,\zeta) = -\frac{1}{\Gamma(a)} \sum_{k=0}^{\infty} \frac{(a)_k}{k!^2}\zeta^k \left(\ln\zeta + \frac{\Gamma'(a+k)}{\Gamma(a+k)} - \frac{2\Gamma'(k+1)}{\Gamma(k+1)}\right), \tag{C.4}$$

where $(a)_k = \dfrac{\Gamma(a+k)}{\Gamma(a)}$ is the Pochhammer symbol. This allows us to write ψ as

$$\psi(\zeta) = g(\zeta)\ln\zeta + h(\zeta), \tag{C.5}$$

where g and h are entire and single-valued. In particular, we have

$$\psi(a,1,\zeta) = -\frac{1}{\Gamma(a)}\left(\ln\zeta + \frac{\Gamma'(a)}{\Gamma(a)} - \frac{2\Gamma'(1)}{\Gamma(1)}\right) + O(\zeta\ln\zeta) \tag{C.6}$$

for $\zeta : |\zeta| < 1$. The precise asymptotics of ψ as $\zeta \in \mathbb{C} \to \infty, -3\pi/2 < \arg\zeta < 3\pi/2$ for fixed a is (formula (7.2) in [37])

$$\psi(a,1,\zeta) = \zeta^{-a}[1 - a^2\zeta^{-1} + O(\zeta^{-2})] \tag{C.7}$$

This asymptotics is a consequence of the following integral representation of ψ (formula (7.3), [37]):

$$\psi(a,1,\zeta) = \frac{1}{\Gamma(a)}\int_0^{\infty e^{-i\alpha}} t^{a-1}(1+t)^{-a}e^{-\zeta t}dt, \quad -\pi < \alpha < \pi,\ -\pi/2+\alpha < \arg\zeta < \pi/2+\alpha.$$

That representation, in particular, implies that

$$\sup_{u\in i\mathbb{R},|u|>1} |\psi(\pm i\epsilon, 1, u)| \to 1, \quad \epsilon \to 0 \tag{C.8}$$

and

$$\sup_{u\in i\mathbb{R},|u|>1} |\psi(1 \pm i\epsilon, 1, u) - \psi(1,1,u)| \to 0, \quad \epsilon \to 0. \tag{C.9}$$

The crucial property of ψ which makes it indispensable for the Riemann-Hilbert analysis is the following transformation formula (formula (7.30) in [37])

$$\psi(a,c,e^{-2\pi i}\zeta) = e^{2\pi i a}\psi(a,c,\zeta) - \frac{2\pi i}{\Gamma(a)\Gamma(a-c+1)}e^{i\pi a}e^{\zeta}\psi(c-a,c,e^{-i\pi}\zeta).$$

Following [18], we will use the convention that, unless otherwise mentioned, ζ always satisfies $0 \leqslant \arg\zeta < 2\pi$.

Concerning the logarithmic derivative of the Gamma function (the digamma function) which appears above, we will have occasion to use its reflection formula (equation 5.15.6, [57])

$$\frac{\Gamma'(1-z)}{\Gamma(1-z)} - \frac{\Gamma'(z)}{\Gamma(z)} = \pi\cot(\pi z). \tag{C.10}$$

We compute the Szegő function of the weight under consideration before we begin

$$\mathcal{D}(z) = \exp\left(\frac{1}{2\pi i}\int_{\mathbb{T}} \frac{\ln f(s)}{s-z} ds\right) = \exp\left(\frac{\epsilon}{\pi i}\ln\left(\frac{i-z}{i+z}\right)\right). \tag{C.11}$$

C.2 Asymptotic formulas for solution of the Riemann-Hilbert problem

In the next subsection we recall how asymptotics of Y on $\mathbb{T}$ is obtained through solving the Riemann-Hilbert problem. Then we will apply this asymptotics to prove Lemma 4.5.

In [18], the Riemann-Hilbert problem undergoes various transformations until it is in a form for which explicit solutions can be written. The singularity points: i, $-i$, along with the artificially introduced point $z = 1$ (though the analysis reduces to triviality here and we drop this case) are all enclosed by the small disks U_i, U_{-i} of fixed radius $\delta > 0$. The remainder of the unit circle is enclosed in "lenses" (see Figure 2).

We trace through the various transformations of the RH problem for z away from the points of singularity i and $-i$. These reductions are:

$$Y \to T \to S \to R\,. \tag{C.12}$$

We will explain each of these transformations below. Our analysis will be limited to the case when $\epsilon \in (0, \epsilon_0]$, $n > n_0(\epsilon)$, $|z| \leq 1$ and z belongs to one of the lenses. This choice is

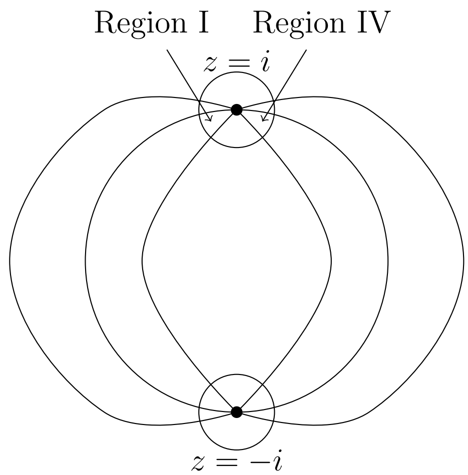

Figure 2: Setup of Riemann-Hilbert problem on $\mathbb{T}$

motivated by our goal to control Y only on the unit circle $\mathbb{T}$ itself so we will only need to take $|z| = 1$ later on. In these domains, some of the transformations in (C.12) are trivial, e.g., $T = Y$ (formula (4.1) in [18]).

Then, (formula (4.3), [18]), S is related to T by

$$S(z) = T(z) \begin{pmatrix} 1 & 0 \\ -f(z)^{-1}z^n & 1 \end{pmatrix}.$$

Now, the original Riemann-Hilbert problem for Y can be written in terms of S and its solution proceeds by first choosing various parametrices (approximate solutions) in each of the domains. The parametrices outside of $\cup_j U_{z_j}$ and inside of each U_{z_j} will be denoted by N and P_{z_j}, respectively. Our final transformation is to R, which satisfies the Riemann-Hilbert problem of very special form. In fact, the correct choices of parametrices N and P_{z_j} makes it possible to say that each jump in the Riemann-Hilbert problem for R is of the form $I + O(n^{-1})$ when $n \to \infty$ on each of the contours involved (see (4.57)–(4.59) of [18]) and an asymptotics of R at infinity is $I + O(1/z)$. Then, the standard argument (see, e.g., [19]) implies that

$$R = I + O(n^{-1}) \tag{C.13}$$

uniformly over $z \in \mathbb{C}$. It is clear now that the main asymptotics of Y is captured by N and P_{z_j}. Below we will discuss these parametrices in detail.

Case 1. Parametrix N, z outside of U_{z_j}

For $z : |z| < 1$, we write (formula (4.7), [18])

$$N(z) = \begin{pmatrix} \mathcal{D}(z) & 0 \\ 0 & \mathcal{D}(z)^{-1} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

$$R(z) = S(z) N^{-1}(z)$$

and

$$R(z) = I + O\left(\frac{1}{n}\right)$$

by equations (4.61) and (4.65)-(4.71) in [18]. Since we are away from the singularities of the weight all terms are uniformly bounded. Collecting $O\left(\frac{1}{n}\right)$ errors, we have

$$S(z) = N(z) + O\left(\frac{1}{n}\right).$$

Reversing these transformations

$$Y(z) = T(z) = S(z) \begin{pmatrix} 1 & 0 \\ -f(z)^{-1} z^n & 1 \end{pmatrix}^{-1} = \left(N(z) + O\left(\frac{1}{n}\right)\right) \begin{pmatrix} 1 & 0 \\ f(z)^{-1} z^n & 1 \end{pmatrix}$$

$$= \left(\begin{pmatrix} \mathcal{D}(z) & 0 \\ 0 & \mathcal{D}(z)^{-1} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} + O\left(\frac{1}{n}\right) \right) \begin{pmatrix} 1 & 0 \\ f(z)^{-1} z^n & 1 \end{pmatrix}.$$

So

$$Y(z) = \begin{pmatrix} z^n f^{-1} \mathcal{D}(z) & \mathcal{D}(z) \\ -\mathcal{D}(z)^{-1} & 0 \end{pmatrix} + O\left(\frac{1}{n}\right), \tag{C.14}$$

when $|z|<1$ and $z\notin U_{z_j}$. Now, to get asymptotical behavior of $Y(z)$ (and thus of $\Phi_n^*(z)$ and $\Psi_n^*(z)$ by formula (C.2)) on $\mathbb{T}$ but outside U_{z_j}, we need to take $|z|\to 1$ in (C.14) and this gives

$$Y(z)=\begin{pmatrix} z^n f^{-1}\mathcal{D}_+(z) & \mathcal{D}_+(z) \\ -\mathcal{D}_+(z)^{-1} & 0 \end{pmatrix}+O\left(\frac{1}{n}\right) \tag{C.15}$$

uniformly over $z\in\mathbb{T}, z\notin U_{z_j}$.

Case 2: Parametrix P_{z_j}, $z\in U_{z_j}$.

The nature of the singularities of the weight f at points z_1 and z_2 is the same so we will discuss only asymptotics in U_{z_1} in detail. By (4.23) in [18], we write

$$P_{z_j}(z)=E(z)\Psi^{(j)}(z)\begin{pmatrix} z^{n/2} & 0 \\ 0 & z^{-n/2} \end{pmatrix},$$

where $E(z)$ is chosen so that P_{z_j} and N approximately match across ∂U_{z_j}. For example, we choose region I on Figure 2 (we can also take the similar domain around z_2). This corresponds to $|z|<1, 0<\arg\frac{z}{z_j}<\pi/2$. By introducing $\zeta=n\ln\frac{z}{z_j}$ we get $\arg\zeta\in(\pi/2,\pi)$ and the choice of E and $\Psi^{(j)}$ are made by (formula (4.32), [18])

$$\Psi_j^{(I)}(\zeta)=\begin{pmatrix} \psi(\beta_j,1,\zeta)e^{2i\pi\beta_j}\left(\frac{z}{z_j}\right)^{-n/2} & -e^{i\pi\beta_j}\psi(1-\beta_j,1,e^{-i\pi}\zeta)\left(\frac{z}{z_j}\right)^{n/2}\frac{\Gamma(1-\beta_j)}{\Gamma(\beta_j)} \\ -e^{i\pi\beta_j}\psi(1+\beta_j,1,\zeta)\left(\frac{z}{z_j}\right)^{-n/2}\frac{\Gamma(1+\beta_j)}{\Gamma(-\beta_j)} & \psi(-\beta_j,1,e^{-i\pi}\zeta)\left(\frac{z}{z_j}\right)^{n/2} \end{pmatrix}$$

$$:=\begin{pmatrix} e^{2i\pi\beta_j}\left(\frac{z}{z_j}\right)^{-n/2}\psi_1^{(j)} & -e^{i\pi\beta_j}\left(\frac{z}{z_j}\right)^{n/2}c_2^{(j)}\psi_3^{(j)} \\ -e^{i\pi\beta_j}\left(\frac{z}{z_j}\right)^{-n/2}c_1^{(j)}\psi_2^{(j)} & \left(\frac{z}{z_j}\right)^{n/2}\psi_4^{(j)} \end{pmatrix}\quad \text{(this defines}\quad \psi_{1,2,3,4}^{(j)}, c_{1,2}^{(j)})$$

and (formula (4.47))

$$E(z)=N(z)\begin{pmatrix} \zeta^{\beta_j} & 0 \\ 0 & \zeta^{-\beta_j} \end{pmatrix}\begin{pmatrix} z_j^{-n/2} & 0 \\ 0 & z_j^{n/2} \end{pmatrix}\begin{pmatrix} e^{-2\pi i\beta_j} & 0 \\ 0 & e^{\pi i\beta_j} \end{pmatrix}.$$

Multiplying matrices, we find

$$P_{z_j}(z) = \begin{pmatrix} 0 & \mathcal{D}(z) \\ -\mathcal{D}(z)^{-1} & 0 \end{pmatrix} \begin{pmatrix} \zeta^{\beta_j}\psi_1^{(j)} & -e^{-\pi i\beta_j}c_2^{(j)}\zeta^{\beta_j}z_j^{-n}\psi_3^{(j)} \\ -e^{2\pi i\beta_j}c_1^{(j)}z_j^n\zeta^{-\beta_j}\psi_2^{(j)} & e^{\pi i\beta_j}\zeta^{-\beta_j}\psi_4^{(j)} \end{pmatrix}.$$

Restricting our attention to $j = 1$, and using the notation $Y^{(1,I)}$ to identify this as the solution Y in region I around point $z_1 = i$:

$$Y^{(1,I)} = T = RP_{z_1} \begin{pmatrix} 1 & 0 \\ f(z)^{-1}z^n & 1 \end{pmatrix}.$$

We have again

$$R(z) = S(z)P_{z_1}^{-1}(z)$$

and

$$R(z) = I + O\left(\frac{1}{n}\right)$$

as follows from (C.13). Since we have singularities in our expressions, we will leave R in this form for now. This yields

$$Y^{(1,I)}(z) = \left(I + O\left(\frac{1}{n}\right)\right) P_{z_1}(z) \begin{pmatrix} 1 & 0 \\ f^{-1}z^n & 1 \end{pmatrix} \tag{C.16}$$

$$= \left(I + O\left(\frac{1}{n}\right)\right) \begin{pmatrix} 0 & \mathcal{D}(z) \\ -\mathcal{D}(z)^{-1} & 0 \end{pmatrix} \times$$

$$\begin{pmatrix} \zeta^{-i\epsilon/\pi}\psi_1^{(1)} & -e^{-\epsilon}c_2^{(1)}\zeta^{-i\epsilon/\pi}i^{-n}\psi_3^{(1)} \\ -e^{2\epsilon}c_1^{(1)}i^n\zeta^{i\epsilon/\pi}\psi_2^{(1)} & e^{\epsilon}\zeta^{i\epsilon/\pi}\psi_4^{(1)} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ f^{-1}z^n & 1 \end{pmatrix}$$

$$= \left(I + O\left(\frac{1}{n}\right)\right) \begin{pmatrix} 0 & \mathcal{D}(z) \\ -\mathcal{D}(z)^{-1} & 0 \end{pmatrix} \times$$

$$\begin{pmatrix} \zeta^{-i\epsilon/\pi}\psi_1^{(1)} - e^{-\epsilon} f^{-1} z^n c_2^{(1)} \zeta^{-i\epsilon/\pi} i^{-n} \psi_3^{(1)} & -e^{-\epsilon} c_2^{(1)} \zeta^{-i\epsilon/\pi} i^{-n} \psi_3^{(1)} \\ -e^{2\epsilon} c_1^{(1)} i^n \zeta^{i\epsilon/\pi} \psi_2^{(1)} + e^{\epsilon} f^{-1} z^n \zeta^{i\epsilon/\pi} \psi_4^{(1)} & e^{\epsilon} \zeta^{i\epsilon/\pi} \psi_4^{(1)} \end{pmatrix}$$

$$= \left(I + O\left(\frac{1}{n}\right)\right) \begin{pmatrix} \mathcal{D}(z)(-e^{2\epsilon} c_1^{(1)} i^n \zeta^{i\epsilon/\pi} \psi_2^{(1)} + e^{\epsilon} f^{-1} z^n \zeta^{i\epsilon/\pi} \psi_4^{(1)}) & \mathcal{D}(z) e^{\epsilon} \zeta^{i\epsilon/\pi} \psi_4^{(1)} \\ -\mathcal{D}(z)^{-1}(\zeta^{-i\epsilon/\pi}\psi_1^{(1)} - e^{-\epsilon} f^{-1} z^n c_2^{(1)} \zeta^{-i\epsilon/\pi} i^{-n} \psi_3^{(1)}) & \mathcal{D}(z)^{-1} e^{-\epsilon} c_2^{(1)} \zeta^{-i\epsilon/\pi} i^{-n} \psi_3^{(1)} \end{pmatrix}.$$

In U_{z_2}, the calculations are exactly the same, but with $z_1 = i$ replaced by $z_2 = -i$ and β_1 replaced by $\beta_2 = -\beta_1$. Therefore, we have

$$Y^{(2,I)}(z) =$$

$$\left(I + O\left(\frac{1}{n}\right)\right) \begin{pmatrix} \mathcal{D}(z)(-e^{-2\epsilon} c_1^{(2)} (-i)^n \zeta^{-i\epsilon/\pi} \psi_2^{(2)} + e^{-\epsilon} f^{-1} z^n \zeta^{i\epsilon/\pi} \psi_4^{(2)}) & \mathcal{D}(z) e^{-\epsilon} \zeta^{-i\epsilon/\pi} \psi_4^{(2)} \\ -\mathcal{D}(z)^{-1}(\zeta^{i\epsilon/\pi}\psi_1^{(2)} - e^{\epsilon} f^{-1} z^n c_2^{(2)} \zeta^{i\epsilon/\pi} (-i)^{-n} \psi_3^{(2)}) & \mathcal{D}(z)^{-1} e^{\epsilon} c_2^{(2)} \zeta^{i\epsilon/\pi} (-i)^{-n} \psi_3^{(2)} \end{pmatrix}.$$

Recalling that

$$f(z) = \begin{cases} e^{\epsilon}, & -\pi/2 \le \arg z < \pi/2 \\ e^{-\epsilon}, & \text{otherwise} \end{cases},$$

we see that $f = e^{-\epsilon}$ in U_{z_1}, region I, and $f = e^{\epsilon}$ in region I of U_{z_2}. Thus, we get

$$Y^{(1,I)}(z) = \left(I + O\left(\frac{1}{n}\right)\right) \times$$

$$\begin{pmatrix} e^{2\epsilon}\mathcal{D}(z)(-c_1^{(1)} \zeta^{i\epsilon/\pi} i^n \psi_2^{(1)} + z^n \zeta^{i\epsilon/\pi} \psi_4^{(1)}) & \mathcal{D}(z) e^{\epsilon} \zeta^{i\epsilon/\pi} \psi_4^{(1)} \\ -\mathcal{D}(z)^{-1}(\zeta^{-i\epsilon/\pi}\psi_1^{(1)} - z^n c_2^{(1)} \zeta^{-i\epsilon/\pi} i^{-n} \psi_3^{(1)}) & \mathcal{D}(z)^{-1} e^{-\epsilon} c_2^{(1)} \zeta^{-i\epsilon/\pi} i^{-n} \psi_3^{(1)} \end{pmatrix} \quad \text{(C.17)}$$

and

$$Y^{(2,I)}(z) = \left(I + O\left(\frac{1}{n}\right)\right) \times$$

$$\begin{pmatrix} e^{-2\epsilon}\mathcal{D}(z)(-c_1^{(2)}\zeta^{-i\epsilon/\pi}(-i)^n\psi_2^{(2)} + z^n\zeta^{-i\epsilon/\pi}\psi_4^{(2)}) & \mathcal{D}(z)e^{-\epsilon}\zeta^{-i\epsilon/\pi}\psi_4^{(2)} \\ -\mathcal{D}(z)^{-1}(\zeta^{i\epsilon/\pi}\psi_1^{(2)} - z^n c_2^{(2)}\zeta^{i\epsilon/\pi}(-i)^{-n}\psi_3^{(2)}) & \mathcal{D}(z)^{-1}e^{\epsilon}c_2^{(2)}\zeta^{i\epsilon/\pi}(-i)^{-n}\psi_3^{(2)} \end{pmatrix}. \tag{C.18}$$

Because of the singularities involved, we must take care in performing this last multiplication. Denote by $\widetilde{Y}_{z_j}$ the right-hand matrix in the above equation. That is

$$\widetilde{Y}_i^{(I)}(z) :=$$

$$\begin{pmatrix} e^{2\epsilon}\mathcal{D}(z)(-c_1^{(1)}\zeta^{i\epsilon/\pi}i^n\psi_2^{(1)} + z^n\zeta^{i\epsilon/\pi}\psi_4^{(1)}) & \mathcal{D}(z)e^{\epsilon}\zeta^{i\epsilon/\pi}\psi_4^{(1)} \\ -\mathcal{D}(z)^{-1}(\zeta^{-i\epsilon/\pi}\psi_1^{(1)} - z^n c_2^{(1)}\zeta^{-i\epsilon/\pi}i^{-n}\psi_3^{(1)}) & \mathcal{D}(z)^{-1}e^{-\epsilon}c_2^{(1)}\zeta^{-i\epsilon/\pi}i^{-n}\psi_3^{(1)} \end{pmatrix}, \tag{C.19}$$

$$\widetilde{Y}_{-i}^{(I)}(z) :=$$

$$\begin{pmatrix} e^{-2\epsilon}\mathcal{D}(z)(-c_1^{(2)}\zeta^{-i\epsilon/\pi}(-i)^n\psi_2^{(2)} + z^n\zeta^{-i\epsilon/\pi}\psi_4^{(2)}) & \mathcal{D}(z)e^{-\epsilon}\zeta^{-i\epsilon/\pi}\psi_4^{(2)} \\ -\mathcal{D}(z)^{-1}(\zeta^{i\epsilon/\pi}\psi_1^{(2)} - z^n c_2^{(2)}\zeta^{i\epsilon/\pi}(-i)^{-n}\psi_3^{(2)}) & \mathcal{D}(z)^{-1}e^{\epsilon}c_2^{(2)}\zeta^{i\epsilon/\pi}(-i)^{-n}\psi_3^{(2)} \end{pmatrix}. \tag{C.20}$$

We begin be proving Lemma 4.6.

Proof. (Lemma 4.6). We only need to use formula (1.23) from [18]. This equation shows, in the notations introduced above,

$$-\overline{\alpha}_{k-1} = \Phi_k(0) = k^{-2\beta_1-1}z_1^k 2^{2\beta_2}\frac{\Gamma(1+\beta_1)}{\Gamma(-\beta_1)} + k^{-2\beta_2-1}z_2^k 2^{2\beta_1}\frac{\Gamma(1+\beta_2)}{\Gamma(-\beta_1)} + r_{k,\epsilon} \tag{C.21}$$

and

$$|r_{k,\epsilon}| < C_\epsilon(k+1)^{-2}. \tag{C.22}$$

□

Remark. The estimates (C.21), (C.22), and (1.3) imply that the recursion coefficients under consideration satisfy

$$\|\{\gamma_k\}\|_{\ell^2} \lesssim \sqrt{\epsilon}, \quad |\gamma_k| <_\epsilon (k+1)^{-1}. \tag{C.23}$$

Recall that Ψ_n and Ψ_n^* satisfy recursion

$$\begin{cases} \Psi_{n+1} = & z\Psi_n + \overline{\gamma}_n \Psi_n^* \\ \Psi_{n+1}^* = & \Psi_n^* + \gamma_n z \Psi_n \end{cases} \tag{C.24}$$

and we have

$$|\Psi_{n+1}| \le |\Psi_n|(1+|\gamma_n|), \quad |\Psi_0| = 1$$

Iterating this formula and using (C.23) we get a rough upper bound

$$\|\Psi_n\|_{L^\infty(\mathbb{T})} <_\epsilon n^{C_\epsilon}. \tag{C.25}$$

This estimate can be substantially improved by Riemann-Hilbert analysis but (C.25) will be good enough for our purposes.

Now we are ready to verify Lemma 4.5.

C.3 First claim of Lemma

In this section we verify

$$|\Phi_n^*(z)| \sim 1, z \in \mathbb{T} \text{ for } \epsilon \in (0, \epsilon_0] \text{ and } n > n_0(\epsilon)$$

We consider two cases.

Case 1: z outside U_{z_j}

By equation (C.14), $\Phi_n^*(z) \to \mathcal{D}_+(z)^{-1}$ uniformly in this region. Since $|\mathcal{D}_+| = f^{1/2}$ a.e. on $\mathbb{T}$, this trivially implies our desired estimate for z outside of $U_{\pm i}$.

Case 2: z inside U_{z_j}

We consider the boundary values as $|z| \to 1$ in the asymptotics for Y. Notice that, since $|\mathcal{D}_+|^2 = f$, we have $|\mathcal{D}_+| = e^{-\epsilon/2}$ in region I around $z = i$. We will focus on region I where $\zeta = iu, u > 0$. In the other regions, analysis is the same. Recall that $\zeta = n \ln \frac{z}{i}$. So, if $z = e^{i(\pi/2+\tau)}, \tau > 0$, we have

$$\zeta^{-i\epsilon/\pi} = \left(n \ln \frac{z}{i}\right)^{-i\epsilon/\pi} = (ni\tau)^{-i\epsilon/\pi} = e^{\epsilon/2} e^{-\frac{i\epsilon}{\pi}\ln(n\tau)}.$$

Therefore, in region I, in which $\arg \frac{z}{i} > 0$, one has

$$\left|\frac{1}{\mathcal{D}(z)\zeta^{i\epsilon/\pi}}\right| = e^{\epsilon}. \tag{C.26}$$

Similarly, in region IV, the other side of i, $|\mathcal{D}(z)| = e^{\epsilon/2}$ and $|\zeta^{-i\epsilon/\pi}| = |e^{-i\epsilon\pi^{-1}\ln\zeta}| = e^{3\epsilon/2}$ (since $\arg\zeta = 3\pi/2$). We again obtain $\left|\frac{1}{\mathcal{D}(z)\zeta^{i\epsilon/\pi}}\right| = e^{\epsilon}$.

Consider the expressions involving the ψ in the first column of $\widetilde{Y}_i^{(I)}(z)$. We focus on $\widetilde{Y}_{i,21}^{(I)}$, the bottom left corner of the $\widetilde{Y}$ matrix. Due to (C.26) and definition of ζ,

$$\left|\widetilde{Y}_{i,21}^{(I)}(z)\right| = e^{\epsilon}\left|\psi_1^{(1)} - \left(\frac{z}{i}\right)^n c_2^{(1)}\psi_3^{(1)}\right| = e^{\epsilon}\left|\psi\left(-\frac{i\epsilon}{\pi}, 1, \zeta\right) - e^{\zeta}\frac{\Gamma(1+\frac{i\epsilon}{\pi})}{\Gamma(-\frac{i\epsilon}{\pi})}\psi\left(1+\frac{i\epsilon}{\pi}, 1, e^{-i\pi}\zeta\right)\right|.$$

Consider

$$\Omega(\zeta,\epsilon) = \psi\left(-\frac{i\epsilon}{\pi}, 1, \zeta\right) - e^{\zeta}\frac{\Gamma(1+\frac{i\epsilon}{\pi})}{\Gamma(-\frac{i\epsilon}{\pi})}\psi\left(1+\frac{i\epsilon}{\pi}, 1, e^{-i\pi}\zeta\right).$$

It is the analysis of this function which concerns us. We want to show that

$$\max_{\zeta=iu, u\in\mathbb{R}} ||\Omega(\zeta,\epsilon)| - 1| \to 0, \quad \epsilon \to 0.$$

We do this in two steps. The estimates (C.7),(C.8), and (C.9) imply that

$$\max_{\zeta=iu, |u|>1} ||\Omega(\zeta,\epsilon)| - 1| \to 0, \quad \epsilon \to 0.$$

For $|\zeta| < 1$, we will use series (C.4) for ψ. We want to show

$$\max_{\zeta=iu, |u|<1} ||\Omega(\zeta,\epsilon)| - 1| \to 0, \quad \epsilon \to 0.$$

Recall that $\Gamma(\zeta)$ has a pole at 0 so $\lim_{\epsilon\to 0}\Gamma^{-1}(\pm i\epsilon)=0$. From (C.5), we get

$$\Omega(\zeta,\epsilon)=$$

$$(\ln\zeta)g(-i\epsilon/\pi,\zeta)+h(-i\epsilon/\pi,\zeta)-e^{\zeta}\frac{\Gamma(1+i\epsilon/\pi)}{\Gamma(-i\epsilon/\pi)}\Big((\ln(e^{-i\pi}\zeta))g(1+i\epsilon/\pi,e^{-i\pi}\zeta)+h(1+i\epsilon/\pi,e^{-i\pi}\zeta)\Big),$$

where

$$g(a,\zeta)=-\frac{1}{\Gamma(a)}\sum_{k=0}^{\infty}\frac{(a)_k}{k!^2}\zeta^k, \tag{C.27}$$

$$h(a,\zeta)=-\frac{1}{\Gamma(a)}\sum_{k=0}^{\infty}\frac{(a)_k}{k!^2}\zeta^k\left(\frac{\Gamma'(a+k)}{\Gamma(a+k)}-\frac{2\Gamma'(k+1)}{\Gamma(k+1)}\right). \tag{C.28}$$

These expansions converge uniformly in $\zeta : |\zeta|<1$ and the coefficients depend on ϵ explicitly. In Ω, the logarithmic singularities cancel each other as follows (recall that $\zeta=iu, u>0$)

$$-\frac{\ln\zeta}{\Gamma(-i\epsilon/\pi)}+\frac{\ln(e^{-i\pi}\zeta)}{\Gamma(-i\epsilon/\pi)}=\frac{-i\pi}{\Gamma(-i\epsilon/\pi)}\to 0,\ \epsilon\to 0,$$

where we accounted for the first terms in the series (C.27) and (C.28) only since for the other terms we can use

$$|\zeta\ln\zeta|\lesssim 1,\quad |\zeta|<1.$$

Therefore, the required asymptotics of Ω will follow from

$$-\lim_{\epsilon\to 0}\frac{\Gamma'(-i\epsilon/\pi)}{\Gamma^2(-i\epsilon/\pi)}=1.$$

Now, recall that (C.1), (C.20), (C.16) yield

$$-\Phi^*_{n-1}(z)=O(n^{-1})\widetilde{Y}^{(I)}_{i,11}(z)+(1+O(n^{-1}))\widetilde{Y}^{(I)}_{i,21}(z).$$

The analysis to show that $\widetilde{Y}^{(I)}_{i,11}(z)=O(1)$ is nearly identical to that showing $|\widetilde{Y}^{(I)}_{i,21}(z)|\sim 1$ except that it may be performed with less care, since only an upper bound is needed. The estimates we obtained prove (4.10) in Lemma 4.5.

C.4 Second statement

Here we verify the statement

$$\|\Phi_n^*\Psi_n^* + z\Phi_n\Psi_n\|_{L^\infty(\mathbb{T})} >_\epsilon \ln n$$

We will investigate $\Phi_n^*\Psi_n^* + z\Phi_n\Psi_n$ for $z = e^{i\theta}, \theta \in (\pi/2 + n^{-0.5}, \pi/2 + 2n^{-0.5})$. Since $\zeta = n \ln \dfrac{z}{i} = iu, u \sim \sqrt{n}$, this puts us in the $\zeta \to \infty$ regime when using the parametrix P_{z_1}. This also allows us to easily perform the final multiplication in (C.16), since all elements in the $\widetilde{Y}$ matrix are $O(1)$ when $|\zeta| > 1$. Recall equation (C.2):

$$2z^n Y_{22}(z) = (\cosh \epsilon)\Psi_{n-1}^*(z) - F(z)\Phi_{n-1}^*(z). \tag{C.29}$$

Performing the multiplication and noting the error, (C.19) gives

$$Y_{22}(z) = \mathcal{D}(z)^{-1} e^{-\epsilon} c_2^{(1)} \zeta^{-i\epsilon/\pi} i^{-n} \psi_3^{(1)} + O\left(\frac{1}{n}\right).$$

By (C.7), $\mathcal{D}(z)^{-1} e^{-\epsilon} c_2^{(1)} \zeta^{-i\epsilon/\pi} i^{-n} \psi_3^{(1)} = O(n^{-1/2})$ and $Y_{22}(z) = O(n^{-1/2})$. Similarly, by equation (C.19),

$$\Phi_n^* = \mathcal{D}^{-1} + O\left(\frac{1}{\sqrt{n}}\right) \tag{C.30}$$

for $\zeta = iu, u \sim \sqrt{n}$. Therefore we have

$$|\Psi_n^*(z)| \sim \epsilon \ln n \tag{C.31}$$

due to (C.3). Recall that $z^n \overline{\Phi_n^*} = \Phi_n$. So, we can write

$$\Phi_n^*\Psi_n^* + z\Phi_n\Psi_n = \Phi_n^*\Psi_n^* + z^{2n+1}\overline{\Phi_n^*\Psi_n^*} = \Phi_n^*\Psi_n^* \left(1 + z^{2n+1} \frac{\overline{\Phi_n^*\Psi_n^*}}{\Phi_n^*\Psi_n^*}\right).$$

Inserting (C.29),(C.30), and using $|\mathcal{D}| \sim 1$, we have

$$\frac{\overline{\Phi_n^*\Psi_n^*}}{\Phi_n^*\Psi_n^*} = \frac{\overline{F(z)(\mathcal{D}(z)^{-2} + O(1/\sqrt{n}))}}{F(z)(\mathcal{D}(z)^{-2} + O(1/\sqrt{n}))} = \frac{\overline{F}}{F} \cdot \left(\frac{\overline{\mathcal{D}}}{\mathcal{D}}\right)^{-2} (1 + O(n^{-0.5})). \tag{C.32}$$

From (C.3) and (C.11) we can compute

$$\frac{\overline{F}}{F} = -1 + o(1), \quad n \to \infty$$

and

$$\frac{\overline{\mathcal{D}}}{\mathcal{D}} = \exp\left(\frac{2i\epsilon}{\pi}\ln(\theta/2 - \pi/4)\right)(1 + o(1)), \quad n \to \infty$$

in our range of $z = e^{i\theta}$. Therefore, substitution into (C.32) shows that there is some $\theta_0 : \theta_0 \in (\pi/2 + n^{-0.5}, \pi/2 + 2n^{-0.5})$ for which

$$\left|1 + z_0^{2n+1}\frac{\overline{\Phi_n^*(z_0)\Psi_n^*(z_0)}}{\Phi_n^*(z_0)\Psi_n^*(z_0)}\right| \sim 1$$

where $z_0 = e^{i\theta_0}$.

Since, by (C.31), we have $|\Psi_n^*(z_0)| = O(\ln n)$, and $|\Phi_n^*| \sim 1$, we get

$$\|\Phi_n^*\Psi_n^* + z\Phi_n\Psi_n\|_{L^\infty(\mathbb{T})} >_\epsilon \ln n$$

as desired.

C.5 Final statement

Here we prove the claim

$$\frac{\Psi_n^*(z)}{\Phi_n^*(z)} + \frac{\Psi_n^*(-z)}{\Phi_n^*(-z)} = O(1), \quad z \in \mathbb{T}, n > n_0(\epsilon)$$

Outside of $U_{z_{1(2)}}$ this statement is trivial by (C.14), so we only consider z inside $U_{z_{1(2)}}$. Further, since the calculations are exactly similar in U_i and U_{-i}, we let $z \in U_i$.

Before we proceed with the analysis let us make two remarks. First, since $|\Phi_n^*| \sim 1, n > n_0(\epsilon)$, we only need to show that U_{2n}, defined by $U_{2n}(z) = \Psi_n^*(z)\Phi_n^*(z) + \Psi_n^*(-z)\Phi_n^*(-z)$, satisfies

$$\|U_{2n}\|_{L^\infty(\mathbb{T})} \lesssim 1,$$

Secondly, U_{2n} is a polynomial of degree at most $2n$ and

$$\|U_{2n}\|_{L^\infty(\mathbb{T})} \lesssim n^{C_\epsilon}$$

by (C.25). The Bernstein inequality gives us

$$\|U'_{2n}\|_{L^\infty(\mathbb{T})} \lesssim n^{1+C_\epsilon}$$

Thus, to prove $\|U_{2n}\|_{L^\infty(\mathbb{T})} = O(1)$, we only need

$$|U_{2n}(e^{i\theta})| \lesssim 1 \tag{C.33}$$

for $\theta : \theta \in (\pi/2 + e^{-\sqrt{n}}, \pi/2 + \delta_1)$ and the parameter δ_1 here is of the same size as the radius of U_i.

Recall that $Y^{(1(2),I)}$ denotes the Y matrix in the region I that corresponds to point $z_{1(2)}$, respectively. By (C.2),

$$2z^n \frac{Y_{22}^{(1,I)}(z)}{Y_{21}^{(1,I)}(z)} + 2(-z)^n \frac{Y_{22}^{(2,I)}(-z)}{Y_{21}^{(2,I)}(-z)} = \Big(F(z) + F(-z)\Big) - (\cosh \epsilon)\left(\frac{\Psi_{n-1}^*(z)}{\Phi_{n-1}^*(z)} + \frac{\Psi_{n-1}^*(-z)}{\Phi_{n-1}^*(-z)}\right),$$

so we want to show

$$\Big(F(z) + F(-z)\Big) - 2\left(z^n \frac{Y_{22}^{(1,I)}(z)}{Y_{21}^{(1,I)}(z)} + (-z)^n \frac{Y_{22}^{(2,I)}(-z)}{Y_{21}^{(2,I)}(-z)}\right) = O(1)$$

uniformly on $\mathbb{T}$ provided that n is large enough.

The formula (C.3) implies

$$F(z) + F(-z) = -\frac{i(e^\epsilon - e^{-\epsilon})}{\pi}\left(\ln\left(\frac{i-z}{i+z}\right) + \ln\left(\frac{i+z}{i-z}\right)\right) + e^{-\epsilon} + e^\epsilon = e^\epsilon + e^{-\epsilon}.$$

Due to this cancellation, we may focus on

$$z^n \frac{Y_{22}^{(1)}(z)}{Y_{21}^{(1)}(z)} + (-z)^n \frac{Y_{22}^{(2)}(-z)}{Y_{21}^{(2)}(-z)}$$

where $z = e^{i\theta}$, $\theta \in (\pi/2 + e^{-\sqrt{n}}, \pi/2 + \delta_1)$. Since we know that $|\Phi_n^*| \sim 1$ uniformly on $\mathbb{T}$, we may multiply out the denominators and only examine

$$z^n Y_{22}^{(1)}(z) Y_{21}^{(2)}(-z) + (-z)^n Y_{22}^{(2)}(-z) Y_{21}^{(1)}(z).$$

We must take care in performing the final multiplication in the Riemann-Hilbert problem. We have previously seen that $Y_{21}(z) = \widetilde{Y}_{21}(z) + O\left(\frac{1}{n}\right) \sim O(1)$ uniformly $z \in \mathbb{T}$. Therefore,

$$\begin{gathered} z^n Y_{22}^{(1)}(z) Y_{21}^{(2)}(-z) + (-z)^n Y_{22}^{(2)}(-z) Y_{21}^{(1)}(z) = \\ z^n \widetilde{Y}_{22}^{(1)}(z) \widetilde{Y}_{21}^{(2)}(-z) + (-z)^n \widetilde{Y}_{22}^{(2)}(-z) \widetilde{Y}_{21}^{(1)}(z) + O\left(\max_{j=1,2} \frac{\left|\widetilde{Y}_{12}^{(j)}(z)\right| + \left|\widetilde{Y}_{22}^{(j)}(z)\right|}{n} \right). \end{gathered} \tag{C.34}$$

Since the final term has a logarithmic singularity at $z = i$, we will handle it away from this point. In the range $z = e^{i\theta}, \pi/2 + e^{-\sqrt{n}} < \theta < \pi/2 + \delta_1$, we have (by (C.19) and estimates on ψ)

$$\max_{j=1,2} \frac{|\widetilde{Y}_{12}^{(j)}(z)| + |\widetilde{Y}_{12}^{(j)}(z)|}{n} = O\left(\frac{|\ln \zeta|}{n}\right) = O(\sqrt{n}/n) = O(n^{-1/2}).$$

So, it suffices to show

$$z^n \widetilde{Y}_{22}^{(1)}(z) \widetilde{Y}_{21}^{(2)}(-z) + (-z)^n \widetilde{Y}_{22}^{(2)}(-z) \widetilde{Y}_{21}^{(1)}(z) = O(1). \tag{C.35}$$

By (C.19) and (C.20), we have

$$z^n \widetilde{Y}_{22}^{(1)}(z) \widetilde{Y}_{21}^{(2)}(-z) + (-z)^n \widetilde{Y}_{22}^{(2)}(-z) \widetilde{Y}_{21}^{(1)}(z) =$$

$$\frac{1}{e^{\epsilon} \mathcal{D}(z) \zeta^{i\epsilon/\pi}} \left(\frac{z}{i}\right)^n c_2^{(1)} \psi_3^{(1)} \left(\frac{-\zeta^{i\epsilon/\pi}}{\mathcal{D}(-z)} \left(\psi_1^{(2)} - \left(\frac{-z}{-i}\right)^n c_2^{(2)} \psi_3^{(2)} \right) \right)$$

$$+\frac{e^{\epsilon}\zeta^{i\epsilon/\pi}}{\mathcal{D}(-z)}\left(\frac{-z}{-i}\right)^{n} c_2^{(2)}\psi_3^{(2)}\left(\frac{-1}{\mathcal{D}(z)\zeta^{i\epsilon/\pi}}\left(\psi_1^{(1)}-\left(\frac{z}{i}\right)^{n} c_2^{(1)}\psi_3^{(1)}\right)\right).$$

Notice it is not ambiguous to leave the arguments of the ψ functions unidentified, as $\zeta = n\ln\frac{z}{z_j}$ and $\frac{z}{i} = \frac{-z}{-i}$. Taking absolute values, we have

$$\left|z^n\widetilde{Y}_{22}^{(1)}(z)\widetilde{Y}_{21}^{(2)}(-z)+(-z)^n\widetilde{Y}_{22}^{(2)}(-z)\widetilde{Y}_{21}^{(1)}(z)\right| =$$

$$|\mathcal{D}(z)\mathcal{D}(-z)|^{-1}\left|e^{-\epsilon}\left(c_2^{(1)}\psi_3^{(1)}\psi_1^{(2)}-\left(\frac{z}{i}\right)^{n}c_2^{(1)}c_2^{(2)}\psi_3^{(2)}\psi_3^{(1)}\right)+e^{\epsilon}\left(c_2^{(2)}\psi_1^{(1)}\psi_3^{(2)}-\left(\frac{z}{i}\right)^{n}c_2^{(1)}c_2^{(2)}\psi_3^{(2)}\psi_3^{(1)}\right)\right|$$

$$\lesssim\left|e^{-\epsilon}\frac{\Gamma(1+\frac{i\epsilon}{\pi})}{\Gamma(-\frac{i\epsilon}{\pi})}\psi\left(1+\frac{i\epsilon}{\pi},1,e^{-i\pi}\zeta\right)\psi\left(\frac{i\epsilon}{\pi},1,\zeta\right)+e^{\epsilon}\frac{\Gamma(1-\frac{i\epsilon}{\pi})}{\Gamma(\frac{i\epsilon}{\pi})}\psi\left(1-\frac{i\epsilon}{\pi},1,e^{-i\pi}\zeta\right)\psi\left(-\frac{i\epsilon}{\pi},1,\zeta\right)\right.$$

$$\left.-(e^{\epsilon}+e^{-\epsilon})\frac{\Gamma(1+\frac{i\epsilon}{\pi})}{\Gamma(-\frac{i\epsilon}{\pi})}\frac{\Gamma(1-\frac{i\epsilon}{\pi})}{\Gamma(\frac{i\epsilon}{\pi})}\left(\frac{z}{i}\right)^{n}\psi\left(1+\frac{i\epsilon}{\pi},1,e^{-i\pi}\zeta\right)\psi\left(1-\frac{i\epsilon}{\pi},1,e^{-i\pi}\zeta\right)\right|.$$

By (C.7),(C.8),(C.9), these expressions are uniformly bounded in $\zeta, |\zeta| > 1, \epsilon < \epsilon_0, n > n_0(\epsilon)$. Thus, we only need to consider the case $\zeta : |\zeta| < 1$. On that interval, $(z/i)^n = e^{\zeta} = 1 + O(\zeta)$. We are concerned with the logarithmic singularities and the constant terms in the series expansions for ψ. We isolate these terms and denote their sum c_0. Using the notation $d(z) = \frac{\Gamma'(z)}{\Gamma(z)}$ for the digamma function, we get

$$c_0 = e^{-\epsilon}\frac{1}{\Gamma(-\frac{i\epsilon}{\pi})\Gamma(\frac{i\epsilon}{\pi})}\left(\ln(e^{-i\pi}\zeta)+d\left(1+\frac{i\epsilon}{\pi}\right)-2d(1)\right)\left(\ln\zeta+d\left(\frac{i\epsilon}{\pi}\right)-2d(1)\right)$$

$$-(e^{\epsilon}+e^{-\epsilon})\frac{1}{\Gamma(\frac{i\epsilon}{\pi})\Gamma(-\frac{i\epsilon}{\pi})}\left(\ln(e^{-i\pi}\zeta)+d\left(1+\frac{i\epsilon}{\pi}\right)-2d(1)\right)\left(\ln(e^{-i\pi}\zeta)+d\left(1-\frac{i\epsilon}{\pi}\right)-2d(1)\right)$$

$$+e^{\epsilon}\frac{1}{\Gamma(\frac{i\epsilon}{\pi})\Gamma(-\frac{i\epsilon}{\pi})}\left(\ln(e^{-i\pi}\zeta)+d\left(1-\frac{i\epsilon}{\pi}\right)-2d(1)\right)\left(\ln\zeta+d\left(\frac{-i\epsilon}{\pi}\right)-2d(1)\right).$$

Performing these multiplications, writing $\ln(iu) = \ln u + i\pi/2$, $\ln(e^{-i\pi}iu) = \ln u - i\pi/2$, and pulling out the common factor yields

$$c_0 = \frac{\ln u}{\Gamma(\frac{i\epsilon}{\pi})\Gamma(-\frac{i\epsilon}{\pi})}\left(e^{-\epsilon}\left(d\left(\frac{i\epsilon}{\pi}\right)-d\left(1-\frac{i\epsilon}{\pi}\right)+i\pi\right)+e^{\epsilon}\left(d\left(\frac{-i\epsilon}{\pi}\right)+d\left(1+\frac{i\epsilon}{\pi}\right)+i\pi\right)\right)+O(1),$$

where $\zeta = iu$. Using the reflection formula (C.10) gives

$$\frac{\ln u}{\Gamma(\frac{i\epsilon}{\pi})\Gamma(-\frac{i\epsilon}{\pi})}(e^{\epsilon}(-\pi\cot(-i\epsilon)) + e^{\epsilon}i\pi + e^{-\epsilon}(-\pi\cot(i\epsilon)) + e^{-\epsilon}i\pi)$$

$$= \frac{\pi \ln u}{\Gamma(\frac{i\epsilon}{\pi})\Gamma(-\frac{i\epsilon}{\pi})}(i(e^{\epsilon} + e^{-\epsilon}) - e^{\epsilon}\cot(-i\epsilon) - e^{-\epsilon}\cot(i\epsilon)) = 0,$$

because $\cot z = \dfrac{i(e^{iz} + e^{-iz})}{e^{iz} - e^{-iz}}$. Therefore,

$$z^n \widetilde{Y}_{22}^{(1)}(z)\widetilde{Y}_{21}^{(2)}(-z) + (-z)^n \widetilde{Y}_{22}^{(2)}(-z)\widetilde{Y}_{21}^{(1)}(z) = O(1)$$

and

$$\frac{\Psi_n^*(z)}{\Phi_n^*(z)} + \frac{\Psi_n^*(-z)}{\Phi_n^*(-z)} = O(1)$$

uniformly in $z \in \mathbb{T}$ for n large enough.

Bibliography

[1] M. U. Ambroladze, *On the possible rate of growth of polynomials that are orthogonal with a continuous positive weight*, (English translation), Math. USSR-Sb. 72.2 (1992): 311–331.

[2] A. Aptekarev, S. Denisov, and D. Tulyakov, *On a problem by Steklov*, JAMS 29.4 (2016): 1117-1165.

[3] A. I. Aptekarev, A. Martinez-Finkelstein and J. S. Dehesa, *Asymptotics of orthogonal polynomial's entropy*, J. Comput. Appl. Math. 233.6 (2010): 1355–1365.

[4] A. I. Aptekarev, V. S. Buyarov and I. S. Dehesa, *Asymptotic behavior of L^p–norms and the entropy for general orthogonal polynomials*, Russian Acad. Sci. Sb. Math. 82.2 (1995): 373–395.

[5] J. Baik, P. Deift and K. Johansson, *On the distribution of the length of the longest increasing subsequence of random permutations*, J. Amer. Math. Soc. 12.4 (1999): 1119–1178.

[6] B. Beckermann, A. Martínez-Finkelshtein, E. A. Rakhmanov, F. Wielonsky, *Asymptotic upper bounds for the entropy of orthogonal polynomials in the Szegő class*, J. Math. Phys. 45.11 (2004): 4239–4254.

[7] S. Bernstein, *Sur les polynomes orthogonaux relatifs a un segment fini*, Journal de Mathematiques 9.9 (1930): 219–286.

[8] A. Borodin and V. Gorin, *Lectures on integrable probability.* arXiv:1212.3351, 2012.

[9] ——, *Moments match between the KPZ equation and the Airy point process.* arXiv:1608.01557v2, 2016.

[10] A. Borodin and G. Olshanski, *The ASEP and determinantal point processes.* arXiv:1608.01564, 2016.

[11] P. Borwein and T. Erdelyi, *Polynomials and Polynomial Inequalities*, Springer, 1995.

[12] J. Breuer, B. Simon and O. Zeitouni, *Large deviations and sum rules for spectral theory - a pedagogical approach.* arXiv: 1608.01467, 2016.

[13] L. Carleson, *On convergence and growth of partial sums of Fourier series*, Acta. Math. 116.1 (1966): 135–157.

[14] M. Christ and A. Kiselev, *Scattering and wave operators for one-dimensional Schrödinger operators with slowly decaying nonsmooth potentials*, Geom. Funct. Anal. 12.6 (2002): 1174–1234.

[15] M. Christ and A. Kiselev, *Absolutely continuous spectrum for one-dimensional Schrödinger operators with slowly decaying potentials: Some optimal results*, J. Amer. Math. Soc. 11.4 (1998): 771–797.

[16] R. Coifman, R. Rochberg and G. Weiss, *Factorization theorems for Hardy spaces in several variables*, Ann. of Math. 103 (1976): 611–635.

[17] P. Deift, *Orthogonal Polynomials and Random Matrices: a Riemann-Hilbert approach*, Courant Lecture Notes in Mathematics, New York, NY, 1999.

[18] P. Deift, A. Its and I. Krasovsky, *Asymptotics of Toeplitz, Hankel, and Toeplitz+Hankel determinants with Fisher-Hartwig singularities*, Ann. Math. 174 (2011): 1243–1299.

[19] P. Deift, T. Kriecherbauer, K. McLaughlin, X. Zhou, *Strong asymptotics of orthogonal polynomials with respect to exponential weight*, Comm. Pure Appl. Math. 52.12, 1999: 1491–1552.

[20] S. Denisov, *On the coexistence of absolutely continuous and singular continuous components of the spectral measure for some Sturm-Liouville operators with square summable potential*, J. Differential Equations 191.1 (2003): 90–104.

[21] ——, *The growth of polynomials orthogonal on the unit circle with respect to a weight w that satisfies $w, w^{-1} \in L^\infty(\mathbb{T})$.* arXiv:1611.00267, 2016

[22] ——, *Remark on the formula by Rakhmanov and Steklov's conjecture*, J. Approx. Theory 205 (2016): 102–113.

[23] S. Denisov and A. Kiselev, *Spectral properties of Schrödinger operators with decaying potentials*, Proceedings of Symposia in Pure Mathematics. Vol. 76. No. 2. Providence, RI; American Mathematical Society, 1998: 565–589.

[24] S. Denisov and S. Kupin, *On the growth of the polynomial entropy integrals for measures in the Szegő class*, Advances in Mathematics 241 (2013): 18–32.

[25] ——, *Orthogonal polynomials on the circle for the weight w satisfying conditions: w,1/w in BMO($\mathbb{T}$).* To appear in Constr. Approx., 2016.

[26] S. Denisov and K. Rush, *On Schur parameters in Steklov's problem.* To appear in J. Approx. Theory, 2016.

[27] R. Dudley and R.M. Norvaïsa, *Concrete Functional Calculus*, Springer-Verlag, New York, NY, 2011.

[28] R. Durrett, *Probability: Theory and Examples. Fourth edition*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge, UK, 2010.

[29] C. Fefferman and E. Stein, *H^p spaces of several variables*, Act. Math. 129.1 (1972): 137–193.

[30] A. Foulquie Moreno, A. Martinez-Finkelstein and V. Sousa, *Asymptotics of orthogonal polynomials for a weight with a jump on* $[-1, 1]$, Constr. Approx. 33.2, 2011: 219–263.

[31] Y. L. Geronimus, *Polynomials orthogonal on the circle and on the interval*, in International Series of Monographs on Pure and Applied Mathematics, Vol. 18, Pergamon Press, New York-Oxford-London-Paris, 1960.

[32] ——, *On a conjecture of V. A. Steklov*, Dokl. Akad. Nauk SSSR 142.3 (1962): 507–509.

[33] ——, *The relation between the order of growth of orthonormal polynomials and their weight function*, Mat. Sb. (N.S.) 103.1 (1963): 65–79.

[34] ——, *Some estimates of orthogonal polynomials and the problem of Steklov*, Dokl. Akad. Nauk SSSR 236.1 (1977): 14–17.

[35] B. L. Golinskii, *The problem of V. A. Steklov in the theory of orthogonal polynomials*, Mat. Zametki 15.1 (1974): 21–32.

[36] A. Its, A.S. Fokas and A. Kitaev, *The isomonodromy approach to matrix models in* $2d$ *quantum gravity*, Comm. Math. Phys. 147.2 (1992): 395–430.

[37] A. Its and I. Krasovsky, *Hankel determinant and orthogonal polynomials for the Gaussian weight with a jump*, Contemp. Math. 458 (2008): 215–248.

[38] D. Jackson, *The Theory of Approximation*, AMS, Providence, RI, 1930.

[39] F. John and L. Nirenberg, *On functions of bounded mean oscillation*, Comm. on Pure and Appl. Math. 14.3 (1961): 415–426.

[40] S. Khrushchev, *Schur's algorithm, orthogonal polynomials, and convergence of Walls continued fractions in* $L^2(\mathbb{T})$, J. Approx. Theory 108.2 (2001): 161–248.

[41] R. Killip and I. Nenciu, *CMV: The unitary analogue of Jacobi matrices*, Comm. Pure Appl. Math. 60.8 (2007): 1148–1188.

[42] R. Killip and B. Simon, *Sum rules for Jacobi matrices and their applications to spectral theory*, Ann. of Math. 158 (2003): 253–321.

[43] A. Kolmogorov, *Stationary sequences in Hilbert space*, Bull. univ. Moscow [Russian] 2 (1941), 40 pp.

[44] M. Krein, *On a generalization of some investigations of G. Szegő, W. M. Smirnov, and A.N. Kolmogorov*, Dokl. Akad. Nauk SSSR 46 (1945): 91–94.

[45] ——, *On a problem of extrapolation by A.N. Kolmogorov*, Dokl. Akad. Nauk SSSR 46 (1945): 306–309.

[46] Y. Last, A. Kiselev and B. Simon, *Modified Prüfer and EFGP transforms and the spectral analysis of one-dimensional Schrödinger operators*, Commun. Math. Phys. 194.1 (1998): 1-45.

[47] Y. Last and B. Simon, *Eigenfunctions, transfer matrices, and absolutely continuous spectrum of one-dimensional Schrödinger operators*, Invent. Math. 135.2 (1999): 329–367.

[48] E. Levin and D.S. Lubinsky, *L_p Christoffel functions, L_p universality, and Paley-Wiener Spaces*, Journal d'Analyse Mathematique 125.1 (2015): 243–283.

[49] E. Levin and D. S. Lubinsky, *Asymptotic behavior of Nikolskii constants for polynomials on the unit circle*, Comp. Methods and Funct. Theory 3.15 (2015): 459–468.

[50] N. Levinson, *The Wiener RMS (root-mean-square) error criterion in filter design and prediction*, J. Math. Phys. Mass. Inst. Tech., 25.1 (1947): 261–278.

[51] M. Lukic, *On a conjecture for higher-order Szegő theorems*, Constr. Approx. 38.1 (2013): 161–169.

[52] ——, *On higher-order Szegő theorems with a single critical point of arbitrary order*, Constr. Approx 44.2 (2016): 283–296.

[53] M. Mehta, *L'emploi del polynomes orthogonaux pour calculer certain determinants*, Rapport S. P. H., Saclay 658 (1960).

[54] D. Menshov, *Sur les series de fonctions orthogonales*, Fund. Math. 4.1 (1923): 82–105.

[55] C. Muscalu and W. Schlag, *Classical and Multilinear Harmonic Analysis, Volume 1*, Cambridge Studies in Advanced Mathematics, 137, Cambridge, UK, 2013.

[56] C. Muscalu, T. Tao and C. Thiele, *A counterexample to a multilinear endpoint question of Christ and Kiselev*, Math. Res. Lett. 10 (2003): 236–246.

[57] NIST, *NIST digital library of mathematical functions.* http://dlmf.nist.gov/, Release 1.0.10 of 2015-08-07. Online companion to [58], 2015.

[58] F.W.J. Olver, D.W. Lozier, R.F. Boisver, and C.W. Clark, Eds. *NIST Handbook of Mathematical Functions.*, Print companion to [57], Cambridge University Press, New York, NY 2010.

[59] R. Paley, *Some theorems on orthogonal functions*, Studia Math. 1.3 (1931): 226–238.

[60] C. Perez, *Endpoint estimates for commutators of singular integral operators*, J. Funct. Anal. 128.1 (1995): 163–185.

[61] S. K. Pichorides, *On the best values of the constants in the theorems of M. Riesz, Zygmund and Kolmogorov*, Studia Mathematica 44.2 (1972): 165-179.

[62] E. A. Rahmanov, *On Steklov's conjecture in the theory of orthogonal polynomials*, (English translation) Math. USSR Sb. 36.4 (1980): 549–575.

[63] ——, *Estimates of the growth of orthogonal polynomials whose weight is bounded away from zero*, (English translation) Math. USSR Sb. 156.2 (1981): 269–298.

[64] R. Salem and A. Zygmund, *Some properties of trigonometric series whose terms have random signs*, Acta Mathematica 91.1 (1954): 245–301.

[65] B. Simon, *Orthogonal Polynomials on the Unit Circle, Volumes 1 and 2*, AMS, Providence, RI, 2005.

[66] E. Stein, *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*, Princeton University Press, Pinceton, NJ, 1993.

[67] V. A. Steklov, *Une methode de la solution du probleme de development des fonctions en series de polynomes de Tchebysheff independante de la theorie de fermeture*, Izv. Rus. Ac. Sci. (1921): 281–302

[68] P.K. Suetin, *V. A. Steklov's problem in the theory of orthogonal polynomials*, (English translation) Journal of Soviet Mathematics 12.6 (1979): 631–682.

[69] G. Szegö, *Orthogonal Polynomials (4th ed.)*, AMS, Providence, RI, 1975.

[70] T. Tao, *Topics in Random Matrix Theory*, Vol. 132 of Graduate Studies in Mathematics, AMS, Providence, RI, 2012.

[71] S. Verblunsky, *On positive harmonic functions (second paper)*, Proc. London Math. Soc. 2.1 (1936): 290–320.

[72] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series. With Engineering Applications.*, The Technology Press of the Massachusetts Institute of Technology, Cambridge, MA, 1949.

[73] E. P. Wigner, *Statistical properties of real symmetric matrices with many dimensions*, in Can. Math. Congr. Proc. Univ. of Toronto Press, Toronto, ON, 1959: 174.

[74] A. Zygmund, *A remark on Fourier transforms*, in Mathematical Proceedings of the Cambridge Philosophical Society. Vol. 32. No. 02. Cambridge University Press, Cambridge, UK, 1936.