ELECTROSTATIC INTERPRETATION OF ZEROS OF ORTHOGONAL POLYNOMIALS

STEFAN STEINERBERGER

(Communicated by Mourad Ismail)

ABSTRACT. We study the differential equation $-(p(x)y')'+q(x)y' = \lambda y$, where $p(x)$ is a polynomial of degree at most 2 and $q(x)$ is a polynomial of degree at most 1. This includes the classical Jacobi polynomials, Hermite polynomials, Legendre polynomials, Chebychev polynomials, and Laguerre polynomials. We provide a general electrostatic interpretation of zeros of such polynomials: a set of distinct, real numbers $\{x_1, \ldots, x_n\}$ satisfies

$$p(x_i) \sum_{\substack{k=1 \\ k \neq i}}^{n} \frac{2}{x_k - x_i} = q(x_i) - p'(x_i) \qquad \text{for all } 1 \leq i \leq n$$

if and only if they are zeros of a polynomial solving the differential equation. We also derive a system of ODEs depending on $p(x), q(x)$ whose solutions converge to the zeros of the orthogonal polynomial at an exponential rate.

1. INTRODUCTION

1.1. **Introduction.** We start by describing an 1885 result of Stieltjes for Jacobi polynomials [18]. Jacobi polynomials $P_n^{\alpha,\beta}(x)$, for real $\alpha, \beta > -1$, are the unique (up to a constant factor) solutions of the equation

$$(1 - x^2)y''(x) - (\beta - \alpha - (\alpha + \beta + 2)x) \, y'(x) = n(n + \alpha + \beta + 1)y(x)$$

where $n \in \mathbb{N}$. The solution is a polynomial of degree n having all its zeros in $(-1, 1)$. Stieltjes defined a notion of energy for any set $\{x_1, \ldots, x_n\} \subset (-1, 1)$ as

$$E = - \sum_{\substack{i,j=1 \\ i \neq j}}^{n} \log |x_i - x_j| - \sum_{i=1}^{n} \left(\frac{\alpha + 1}{2} \log |x - 1| + \frac{\beta + 1}{2} \log |x + 1| \right)$$

and showed that the minimal energy is exactly given by the zeros of the Jacobi polynomial. Differentiating this expression in all the n variables leads to a system of equations describing an electrostatic equilibrium

$$\sum_{\substack{k=1 \\ k \neq i}}^{n} \frac{1}{x_k - x_i} = \frac{1}{2} \frac{\alpha + 1}{x_i - 1} + \frac{1}{2} \frac{\beta + 1}{x_i + 1} \qquad \text{for all } 1 \leq i \leq n.$$

A traditional application of the formula is to establish monotonicity of zeros with respect to the parameters α, β. This argument can be found in Szegő's book [19], where he also discusses Laguerre and Hermite polynomials. These types of questions

have been studied by a large number of people; we refer to [2, 7, 10, 11], the survey [16], the 1978 survey and 2001 book by F. Calogero [4, 5], the more recent papers of F. A. Grünbaum [8, 9] and M. E. H. Ismail [11, 12], and references therein.

1.2. **Equilibrium.** The purpose of this paper is to prove a simple general result that characterizes zeros of orthogonal polynomials. The statement itself is a bit more general but most interesting when applied to the classical orthogonal polynomials. We first discuss second order equations, then describe a system of ODEs converging to zeros, and conclude with a short remark on higher order equations.

Theorem 1. *Let $p(x), q(x)$ be polynomials of degree at most 2 and 1, respectively. Then the set $\{x_1, \ldots, x_n\}$, assumed to be in the domain of definition, satisfies*

$$p(x_i) \sum_{\substack{k=1 \\ k \neq i}}^{n} \frac{2}{x_k - x_i} = q(x_i) - p'(x_i) \qquad for\ all \quad 1 \leq i \leq n$$

if and only if

$$y(x) = \prod_{k=1}^{n} (x - x_k) \quad solves \quad -(p(x)y')' + q(x)y' = \lambda y \quad for\ some\ \lambda \in \mathbb{R}.$$

In classical applications, there is a unique polynomial solution of fixed degree corresponding to a fixed value of λ. This removes λ as a variable and leads to a complete characterization. One direction of the statement (zeros of a polynomial solutions satisfy the system of equations) is a fairly straightforward computation, and the outline of the argument can be found, for example, in [8] or, as a remark, in [11, §3]. Moreover, results in this direction can be obtained for much more general differential equations of higher order (see, for example, [4–6]). Arguments in the other direction seem to exist only in special cases, as in the work of Stieltjes, and are based on interpreting the system of equations as the critical point of an associated energy functional (see e.g. [11, 16]). Our argument proceeds by a different route and bypasses considerations of an underlying energy.

We remark that there is no assumption of orthogonality, nor is there any restriction on the domain which may be either bounded or unbounded, but we do assume that the n points are contained in the domain of definition. Let us quickly consider two special cases: we start with the Hermite differential equation

$$-y'' + xy' = \lambda y.$$

This corresponds to $p(x) \equiv 1$ and $q(x) \equiv x$. We deduce from Theorem 1 that the zeros of the n-th solution satisfy a relationship also going back to Stieltjes (see [1]),

$$\sum_{\substack{k=1 \\ k \neq i}}^{n} \frac{2}{x_k - x_i} = x_i \qquad for\ all \quad 1 \leq i \leq n.$$

Returning to a special case of Jacobi polynomials $\alpha = 0 = \beta$ (for simplicity of exposition), we obtain $p(x) = x^2 - 1$, $q(x) = 0$, and thus

$$(x_i^2 - 1) \sum_{\substack{k=1 \\ k \neq i}}^{n} \frac{2}{x_k - x_i} = 2x_i \qquad for\ all \quad 1 \leq i \leq n,$$

which is easily seen to be equivalent to Stieltjes' electrostatic equilibrium.

1.3. **A system of ODEs.** One interesting byproduct of our approach is a proce-
dure that allows for the computation of zeros without ever computing the poly-
nomial. More precisely, for any given set $x_1(0) < x_2(0) < \cdots < x_n(0)$ (assumed
to be in the domain of definition), we consider the system of ordinary differential
equations

$$(\diamond) \qquad \frac{d}{dt} x_i(t) = -p(x_i) \sum_{\substack{k=1 \\ k \neq i}}^{n} \frac{2}{x_k(t) - x_i(t)} + p'(x_i(t)) - q(x_i(t)).$$

The result above implies that the unique stationary point of this system of ODEs
is given by the n zeros of a polynomial solution of degree n. We will now show that
under standard assumptions on the equation $-(p(x)y')' + q(x)y' = \lambda y$, the system
converges to this fixed point at an exponential rate. We require somewhat stronger
assumptions than we do for Theorem 1 and ask that

(1) the equation $-(p(x)y'(x))' + q(x)y'(x) = \lambda y(x)$ be a Sturm-Liouville prob-
lem with a discrete set of solutions indexed by $0 = \lambda_0 < \lambda_1 < \lambda_2 < \cdots$
and

(2) the solution corresponding to y_n be a polynomial of degree n for all $n \in \mathbb{N}$.

These assumptions cover the classical polynomials but are fairly strong. We refer
to Bochner's theorem and its nice exposition in [13, §20].

Theorem 2. *The system $(\diamond)$ converges for all initial values $x_1(0) < \cdots < x_n(0)$ to
the zeros $x_1 < \cdots < x_n$ of the degree n polynomial solving the equation. Moreover,*

$$\max_{1 \leq i \leq n} |x_i(t) - x_i| \leq c e^{-\sigma_n t},$$

where $c > 0$ depends on everything and $\sigma_n \geq \lambda_n - \lambda_{n-1}$.

We are not aware of any result of this type being known; the connection be-
tween partial differential equations and zeros of polynomials (or poles of rational
functions) is, of course, classical; we refer to the extensive survey of Calogero [4]
and his more recent book [5]. Our underlying dynamical system is the *backward*
heat equation (well-posed for algebraic reasons), and existence of a solution for
all time is a consequence of Sturm-Liouville theory. We assume that the initial
values are all contained in the domain where the solution is defined. For classical
Sturm-Liouville problems, the Weyl law suggests $\lambda_n - \lambda_{n-1} \sim n$. The rapid con-
vergence can be easily observed in examples. The assumptions are satisfied for all
classical polynomials. We consider, as a specific example, the Laguerre polynomials
satisfying

$$-xy'' + (x - 1)y' = ny$$

or, in our notation, $p(x) = x = q(x)$. Theorem 2 then implies that for any choice
of initial values $x(0) < y(0) < z(0)$, the system of ordinary differential equations

$$\dot{x}(t) = \frac{2x(t)}{x(t) - y(t)} + \frac{2x(t)}{x(t) - z(t)} + 1 - x(t),$$

$$\dot{y}(t) = \frac{2y(t)}{y(t) - x(t)} + \frac{2y(t)}{y(t) - z(t)} + 1 - y(t),$$

$$\dot{z}(t) = \frac{2z(t)}{z(t) - x(t)} + \frac{2z(t)}{z(t) - y(t)} + 1 - z(t)$$

has a solution for all $t > 0$. As $t \to \infty$, the solutions converge to the zeros of the third Laguerre polynomial

$$L_3(x) = x^3 - 9x^2 + 18x - 6,$$

and thus

$$x(t) \to 0.4157\ldots, \quad y(t) \to 2.2942\ldots, \quad \text{and} \quad z(t) \to 6.2899\ldots.$$

Moreover, this convergence happens at an exponential rate (roughly with speed e^{-t}). We consider a large-scale example (see Figure 1) next. The equation

$$-\frac{d}{dx}\left((1 - x^2)\frac{d}{dx}y(x)\right) = n(n + 1)y(x)$$

is solved by the Legendre polynomials P_n defined on $(-1, 1)$. In our notation, we have $p(x) = 1 - x^2$ and $q(x) = 0$. We simulate the system of ODEs

$$\frac{d}{dt}x_i(t) = -(1 - x_i(t))^2 \sum_{\substack{k=1 \\ k \neq i}}^{n} \frac{2}{x_k(t) - x_i(t)} - 2x_i(t)$$

for $n = 100$ equations. The initial values $\{x_1(0), \ldots, x_{100}(0)\}$ are chosen as uniform random variables in the interval $(-0.1, 0.1)$. Figure 1 shows the evolution of the system and convergence to the zeros of the Legendre polynomial of degree 100.
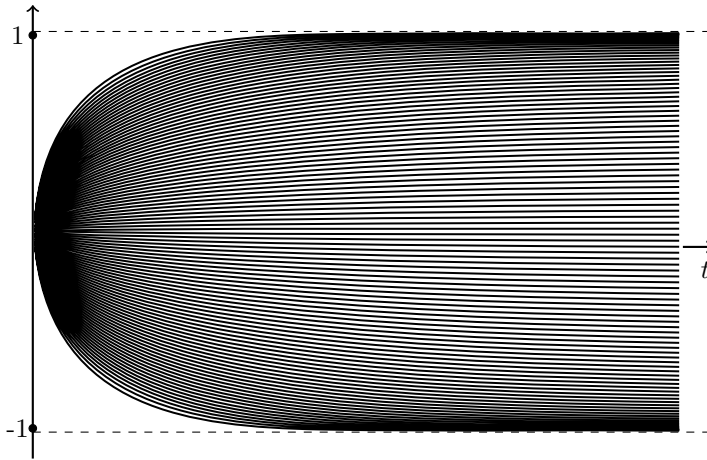

FIGURE 1. Evolution of the system of ODEs for $0 \le t \le 0.01$, which approaches the zeros of the Legendre polynomial P_{100} in $(-1, 1)$.

As indicated in Theorem 2, the constant determining the speed of exponential convergence grows linearly in n: this is reflected in the rather short time-scale $(0 \le t \le 0.01)$ in the picture.

1.4. Higher order equations.

Theorem 1 as well as its proof immediately generalizes to polynomial solutions of the equation

$$a_n(x)y^{(n)}(x) + a_{n-1}(x)y^{(n-1)}(x) + a_1(x)y'(x) = \lambda y(x),$$

where a_n is a polynomial of degree at most n and the solution y is assumed to have only single zeros: one half of the statement can be found in [6, Proposition 1]; the

other direction follows from our approach. We are not aware of an extension of Theorem 2 since our proof relies on Sturm-Liouville theory.

2. Proofs

2.1. **Proof of Theorem 1.**

Proof. One direction of the argument, showing that the zeros of any polynomial solution of the equation are in the desired electrostatic equilibrium, is a simple computation and known in greater generality [1,2]. However, it also follows immediately from our argument and is, for completeness' sake, sketched at the end of the proof. We assume that the system of equations is satisfied and will prove that the associated polynomial solves the differential equation. Fix $x_1, \ldots, x_n$ and consider the candidate polynomial

$$f(x) = \prod_{k=1}^{n} (x - x_k).$$

We want to show that this polynomial satisfies the differential equation

$$-(p(x)f')' + q(x)f' = \lambda f \qquad \text{for some } \lambda \in \mathbb{R}.$$

We introduce the function

$$\frac{\partial}{\partial t} u(t,x) = \frac{\partial}{\partial x} \left(p(x) \frac{\partial}{\partial x} u(t,x) \right) - q(x) \frac{\partial}{\partial x} u(t,x),$$

$$u(0,x) = f(x).$$

This is a parabolic partial differential equation. Generally, unless we specify the sign of p, it might be an inverse heat equation, and the solution of such an equation need not even exist for a short amount of time. Here, however, since p is at most a polynomial of degree 2 and q is a polynomial of degree at most 1, we see that the right-hand side is always a polynomial of degree at most n. In particular, we can rewrite the partial differential equation as a linear system of $n + 1$ ordinary differential equations, and this guarantees existence for all $t > 0$. Suppose $f(x)$ is not a solution of the differential equation. Then

$$u_t(0,x) = \frac{\partial}{\partial x} \left(p(x) \frac{\partial}{\partial x} f(x) \right) - q(x) \frac{\partial}{\partial x} f(x)$$

is a polynomial of degree at most n and not identically 0 (since otherwise f would solve the differential equation for $\lambda = 0$). We observe that if this polynomial vanishes in exactly $\{x_1, \ldots, x_n\}$, then it has to be a multiple of f, and we have obtained the desired result for some $\lambda \neq 0$. If this is not the case, then $u_t(x_i, 0) \neq 0$ for at least one $1 \leq i \leq n$. We fix this value of i. Moreover, we note that

$$f'(x_i) = \prod_{\substack{k=1 \\ k \neq i}}^{n} (x_i - x_k) \neq 0.$$

The implicit function theorem implies that there is a neighborhood of 0 for which there is a function $x_i(t)$ such that

$$u(t, x_i(t)) = 0 \qquad \text{and} \qquad \frac{\partial}{\partial t} x_i(t) \big|_{t=0} \neq 0.$$

We now compute the expression. Differentiation implies that

$$0 = \frac{\partial}{\partial t} u(t, x_i(t))\big|_{t=0} = u_x(0, x_i) \left(\frac{\partial}{\partial t} x_i(t)\big|_{t=0} \right) + u_t(0, x_i).$$

We are interested in the first term, have already computed the second term $u_x(x_i, 0)$ $= f'(x_i)$, and thus only need to compute the third term. A simple computation shows that

$$u_t(t, x_i)\big|_{t=0} = \frac{\partial}{\partial x} \left(p(x) \frac{\partial}{\partial x} f(x) \right) - q(x) \frac{\partial}{\partial x} f(x)\big|_{x=x_i}$$

$$= p(x) \frac{\partial^2}{\partial x^2} f(x) + (p'(x) - q(x)) \frac{\partial}{\partial x} f(x)\big|_{x=x_i}.$$

The first term simplifies to

$$\frac{\partial^2}{\partial x^2} f(x)\big|_{x=x_i} = 2 \sum_{\substack{k=1 \\ k \neq i}}^{n} \prod_{\substack{j=1 \\ j \notin \{i,k\}}}^{n} (x_i - x_j),$$

and altogether we obtain

$$0 \neq \frac{\partial}{\partial t} x_i(t)\big|_{t=0} = p(x_i) \sum_{\substack{k=1 \\ k \neq i}}^{n} \frac{2}{x_k - x_i} + p'(x_i) - q(x_i),$$

which is a contradiction. Conversely, if f is indeed a solution of the equation, then

$$u(t, x) = e^{\lambda t} f(x) \qquad \text{and thus} \qquad \frac{\partial}{\partial t} x_i(t)\big|_{t=0} = 0$$

for all $1 \leq i \leq n$, which implies that the equations are satisfied. $\qquad \square$
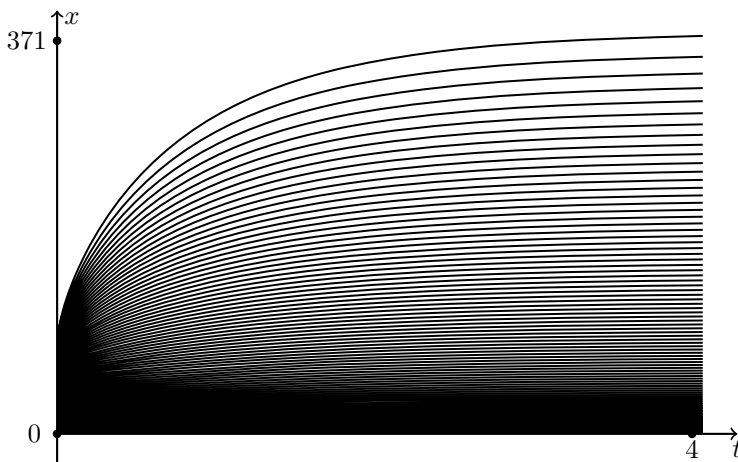

FIGURE 2. Another example: starting with initial values $x_i(0) = i$ for $1 \leq i \leq 100$, the associated system of ODEs approaches the zeros of the Laguerre polynomial L_{100}.

2.2. **Proof of Theorem 2.**

Proof. We assume that the solutions $y_0, y_1, \ldots, y_{n-1}, y_n$ satisfy $\deg(y_i) = i$ for all $0 \leq i \leq n$. We assume that p, q and the domain of definition are such that classical Sturm-Liouville theory applies. Then the polynomial y_j has exactly j zeros all of which are simple. Suppose

$$x_1(0) < x_2(0) < \cdots < x_n(0)$$

is given. We define the function

$$f(x) = \prod_{k=1}^{n} (x - x_k(0)) \quad \text{and write it as} \quad f(x) = \sum_{k=0}^{n} a_k y_k(x)$$

for some coefficients $a_0, \ldots, a_n$. This is possible because of the assumption on the degrees (an upper triangular matrix is invertible). The strong form of the Sturm Oscillation Theorem, which is not very widely known, states that as long as not all coefficients b_i vanish, any function of the form

$$\sum_{k=0}^{n-1} b_k y_k(x) \qquad \text{has at most } n-1 \text{ zeros.}$$

We refer to Bérard and Helffer [3] and Lützen [15] for the history of this remarkable theorem, which seems to have been forgotten ([3] gives rigorous proofs in modern language; [17] gives a quantitative form). Since $f(x)$ has n zeros, the Sturm Oscillation Theorem implies that $a_n \neq 0$. We now define $u(t, x)$ as the solution of the *backward* heat equation

$$\frac{\partial}{\partial t} u(t, x) = -\frac{\partial}{\partial x}\left(p(x)\frac{\partial}{\partial x}u(t, x)\right) + q(x)\frac{\partial}{\partial x}u(t, x),$$
$$u(0, x) = f(x)$$

and observe that, as explained in the proof of Theorem 1, this equation is well-posed for all $t > 0$ since it can be rewritten as a linear system of $n+1$ ordinary differential equations. Linearity implies that the solution is given by

$$u(t, x) = \sum_{k=0}^{n} a_k e^{\lambda_k t} y_k(x).$$

At the same time, as long as the zeros do not collide, we can write

$$u(t, x) = h(t) \prod_{i=1}^{n} (x - x_i(t)),$$

where $h(t) \neq 0$ for all $t > 0$ and the functions $x_i(t)$ satisfy, following the computation done in the proof of Theorem 1 and reversing the sign, the system of ordinary differential equations

$$\frac{d}{dt} x_i(t) = -p(x_i) \sum_{\substack{k=1 \\ k \neq i}}^{n} \frac{2}{x_k(t) - x_i(t)} + q(x_i(t)) - p'(x_i(t)) \quad \text{for all } 1 \leq i \leq n.$$

It is a property of Sturm-Liouville problems that the number of distinct zeros is nonincreasing in time under the forward heat equation. Since we are dealing with the backward heat equation, we see that the number of distinct zeros is nondecreasing. Moreover, since we start with a polynomial of degree n and the solution

is always a polynomial of degree at most n, this implies that zeros can never collide (and that the solution is always exactly of degree n). This shows that

$$e^{-\lambda_n t} h(t) \prod_{i=1}^{n} (x - x_i(t)) = a_n y_n(t) + \sum_{k=0}^{n-1} a_k e^{(\lambda_k - \lambda_n)t} y_k(x).$$

All zeros of y_n are simple, and the Inverse Function Theorem now implies that the zeros of $u(t, x)$ converge to the zeros of y_n exponentially quickly in t. The speed of convergence depends on the size of $\lambda_{n-1} - \lambda_n$ and the size of y_n' at its zeros; the constant in front will depend on the precise values of the coefficients $\{a_0, a_1, \ldots, a_n\}$. If $a_{n-1} = 0$ or more of the leading terms vanish, then convergence would be even faster. □

References

[1] S. Ahmed, M. Bruschi, and F. Calogero, *On the zeros of combinations of Hermite polynomials*, Lett. Nuovo Cimento (2) **21** (1978), no. 13, 447–452. MR0486691

[2] S. Ahmed, M. Bruschi, F. Calogero, M. A. Olshanetsky, and A. M. Perelomov, *Properties of the zeros of the classical polynomials and of the Bessel functions* (English, with Italian and Russian summaries), Nuovo Cimento B (11) **49** (1979), no. 2, 173–199, DOI 10.1007/BF02737517. MR529047

[3] P. Bérard and B. Helffer, *Sturm's theorem on zeros of linear combinations of eigenfunctions*, (2017) arXiv:1706.08247

[4] F. Calogero, *Motion of poles and zeros of special solutions of nonlinear and linear partial differential equations and related "solvable" many-body problems* (English, with Italian and Russian summaries), Nuovo Cimento B (11) **43** (1978), no. 2, 177–241, DOI 10.1007/BF02721013. MR489287

[5] Francesco Calogero, *Classical many-body problems amenable to exact treatments*, (Solvable and/or integrable and/or linearizable...) in one-, two- and three-dimensional space, Lecture Notes in Physics. New Series m: Monographs, vol. 66, Springer-Verlag, Berlin, 2001. MR1880249

[6] R. G. Campos and L. A. Avila, *Some properties of orthogonal polynomials satisfying fourth order differential equations*, Glasgow Math. J. **37** (1995), no. 1, 105–113, DOI 10.1017/S0017089500030445. MR1316970

[7] D. Dimitrov and B. Shapiro, *Degenerate Lamé operators and electrostatic problems with a polynomial constraint*, preprint, 2018.

[8] F. Alberto Grünbaum, *Variations on a theme of Heine and Stieltjes: an electrostatic interpretation of the zeros of certain polynomials*, Proceedings of the VIIIth Symposium on Orthogonal Polynomials and Their Applications (Seville, 1997), J. Comput. Appl. Math. **99** (1998), no. 1-2, 189–194, DOI 10.1016/S0377-0427(98)00156-3. MR1662694

[9] F. Alberto Grünbaum, *Electrostatic interpretation for the zeros of certain polynomials and the Darboux process*, Proceedings of the Fifth International Symposium on Orthogonal Polynomials, Special Functions and their Applications (Patras, 1999), J. Comput. Appl. Math. **133** (2001), no. 1-2, 397–412, DOI 10.1016/S0377-0427(00)00661-0. MR1858298

[10] E. Hendriksen and H. van Rossum, *Electrostatic interpretation of zeros*, Orthogonal polynomials and their applications (Segovia, 1986), Lecture Notes in Math., vol. 1329, Springer, Berlin, 1988, pp. 241–250, DOI 10.1007/BFb0083363. MR973431

[11] Mourad E. H. Ismail, *An electrostatics model for zeros of general orthogonal polynomials*, Pacific J. Math. **193** (2000), no. 2, 355–369, DOI 10.2140/pjm.2000.193.355. MR1755821

[12] Mourad E. H. Ismail, *More on electrostatic models for zeros of orthogonal polynomials*, Proceedings of the International Conference on Fourier Analysis and Applications (Kuwait, 1998), Numer. Funct. Anal. Optim. **21** (2000), no. 1-2, 191–204, DOI 10.1080/01630560008816948. MR1759996

[13] Mourad E. H. Ismail, *Classical and quantum orthogonal polynomials in one variable*, with two chapters by Walter Van Assche, with a foreword by Richard A. Askey, Encyclopedia of Mathematics and its Applications, vol. 98, Cambridge University Press, Cambridge, 2005. MR2191786

[14] G. López and E. A. Rakhmanov, *Rational approximations, orthogonal polynomials and equilibrium distributions*, Orthogonal polynomials and their applications (Segovia, 1986), Lecture Notes in Math., vol. 1329, Springer, Berlin, 1988, pp. 125–157, DOI 10.1007/BFb0083356. MR973424

[15] Jesper Lützen, *Sturm and Liouville's work on ordinary linear differential equations. The emergence of Sturm-Liouville theory*, Arch. Hist. Exact Sci. **29** (1984), no. 4, 309–376, DOI 10.1007/BF00348405. MR745152

[16] F. Marcellán, A. Martínez-Finkelshtein, and P. Martínez-González, *Electrostatic models for zeros of polynomials: old, new, and some open problems*, J. Comput. Appl. Math. **207** (2007), no. 2, 258–272, DOI 10.1016/j.cam.2006.10.020. MR2345246

[17] S. Steinerberger, Quantitative Projections in the Sturm Oscillation Theorem, (2018) arXiv:1804.05779

[18] T. J. Stieltjes, *Sur certains polynômes* (French), Qui vérifient une équation différentielle linéaire du second ordre et sur la theorie des fonctions de Lamé, Acta Math. **6** (1885), no. 1, 321–326, DOI 10.1007/BF02400421. MR1554669

[19] Gábor Szegő, *Orthogonal polynomials*, 4th ed., American Mathematical Society, Colloquium Publications, Vol. XXIII, American Mathematical Society, Providence, R.I., 1975. MR0372517

DEPARTMENT OF MATHEMATICS, YALE UNIVERSITY, NEW HAVEN, CONNECTICUT 06511
Email address: stefan.steinerberger@yale.edu