

Proceedings of Symposia in PURE MATHEMATICS

Volume 76, Part 2

Spectral Theory and Mathematical Physics: A Festschrift in Honor of Barry Simon's 60th Birthday

Ergodic Schrödinger Operators, Singular
Spectrum, Orthogonal Polynomials,
and Inverse Spectral Theory

Fritz Gesztesy (Managing Editor)

Percy Deift

Cherie Galvez

Peter Perry

Wilhelm Schlag

Editors



American Mathematical Society

**Spectral Theory and
Mathematical Physics:
A Festschrift in Honor of
Barry Simon's 60th Birthday**

**Ergodic Schrödinger Operators, Singular
Spectrum, Orthogonal Polynomials,
and Inverse Spectral Theory**

Proceedings of Symposia in PURE MATHEMATICS

Volume 76, Part 2

Spectral Theory and Mathematical Physics: A Festschrift in Honor of Barry Simon's 60th Birthday

Ergodic Schrödinger Operators, Singular
Spectrum, Orthogonal Polynomials,
and Inverse Spectral Theory

A Conference on Spectral Theory
and Mathematical Physics
in Honor of Barry Simon's 60th Birthday
March 27–31, 2006
California Institute of Technology
Pasadena, California

Fritz Gesztesy (Managing Editor)
Percy Deift
Cherie Galvez
Peter Perry
Wilhelm Schlag
Editors



American Mathematical Society
Providence, Rhode Island

2000 *Mathematics Subject Classification*. Primary 34A55, 34F05, 34L05, 34L40, 37H15, 42C05, 47A10, 47B36, 60H25, 81Q10.

Library of Congress Cataloging-in-Publication Data

Spectral theory and mathematical physics : a festschrift in honor of Barry Simon's 60th birthday : Ergodic Schrödinger operators, singular spectrum, orthogonal polynomials, and inverse spectral theory / Fritz Gesztesy . . . [et al.], editors.

p. cm. — (Proceedings of symposia in pure mathematics ; v. 76, pt. 2)

Includes bibliographical references.

ISBN-13: 978-0-8218-4249-2 (alk. paper) (Part 2)

ISBN-13: 978-0-8218-3783-2 (alk. paper) (Set)

1. Spectral theory (Mathematics)—Congresses. I. Simon, Barry, 1946– II. Gesztesy, Fritz, 1953–

QC20.7.S646S64 2006
515'.7222—dc22

2006047073

Copying and reprinting. Material in this book may be reproduced by any means for educational and scientific purposes without fee or permission with the exception of reproduction by services that collect fees for delivery of documents and provided that the customary acknowledgment of the source is given. This consent does not extend to other kinds of copying for general distribution, for advertising or promotional purposes, or for resale. Requests for permission for commercial use of material should be addressed to the Acquisitions Department, American Mathematical Society, 201 Charles Street, Providence, Rhode Island 02904-2294, USA. Requests can also be made by e-mail to reprint-permission@ams.org.

Excluded from these provisions is material in articles for which the author holds copyright. In such cases, requests for permission to use or reprint should be addressed directly to the author(s). (Copyright ownership is indicated in the notice in the lower right-hand corner of the first page of each article.)

© 2007 by the American Mathematical Society. All rights reserved.

The American Mathematical Society retains all rights
except those granted to the United States Government.

Copyright of individual articles may revert to the public domain 28 years
after publication. Contact the AMS for copyright status of individual articles.

Printed in the United States of America.

∞ The paper used in this book is acid-free and falls within the guidelines
established to ensure permanence and durability.

Visit the AMS home page at <http://www.ams.org/>

10 9 8 7 6 5 4 3 2 1 12 11 10 09 08 07

Contents

Part 1. Quantum Field Theory, Statistical Mechanics, and Nonrelativistic Quantum Systems

Preface	ix
A Mathematical Genealogy	xiii
A Selection of Barry Stories	xvii

Quantum Field Theory, Statistical Mechanics

Perspectives in Statistical Mechanics MICHAEL AIZENMAN	3
Coherent Infrared Representations in Non-Relativistic QED THOMAS CHEN and JÜRIG FRÖHLICH	25
Quantum Spin Systems After DLS 1978 BRUNO NACHTERGAELE	47
Barry Simon's Contributions to Quantum Field Theory LON ROSEN	69

Nonrelativistic Two-Body and N -Body Quantum Systems, Resonances

Isoperimetric Inequalities for Eigenvalues of the Laplacian MARK S. ASHBAUGH and RAFAEL D. BENGURIA	105
Non-Self-Adjoint Operators and Pseudospectra E. B. DAVIES	141
Barry Simon's Contributions to Non-Relativistic Quantum Mechanics: Two-Body and N -Body Schrödinger Operators and Resonances RICHARD FROESE	153
N -Body Quantum Scattering and Quantum Resonances: An Overview CHRISTIAN GÉRARD	169
Mathematical Analysis of Born–Oppenheimer Approximations GEORGE A. HAGEDORN and ALAIN JOYE	203

Perturbation Theory and Atomic Resonances Since Schrödinger's Time EVANS M. HARRELL II	227
On a Theorem for Quantum Mechanical Scattering Theory ANDREW LENARD	249
Analytic Criteria in the Qualitative Spectral Analysis of the Schrödinger Operator VLADIMIR MAZ'YA	257
The Spectral Geometry of Geometrically Finite Hyperbolic Manifolds PETER PERRY	289
Topics in the Theory of Positive Solutions of Second-Order Elliptic and Parabolic Partial Differential Equations YEHUDA PINCHOVER	329
Complex Scaling in Atomic Physics: A Staging Ground for Experimental Mathematics and for Extracting Physics from Otherwise Impossible Computations WILLIAM P. REINHARDT	357
Recent Results on the Bethe–Sommerfeld Conjecture ALEXANDER V. SOBOLEV	383

Electric and Magnetic Fields, Semiclassical Limit

Recent Developments in Quantum Mechanics with Magnetic Fields LÁSZLÓ ERDŐS	401
Aspects of the Integer Quantum Hall Effect GIAN MICHELE GRAF	429
Barry Simon's Work on Electric and Magnetic Fields and the Semi-Classical Limit IRA W. HERBST	443
Some Bound State Problems in Quantum Mechanics DIRK HUNDERTMARK	463

Part 2. Ergodic Schrödinger Operators, Singular Spectrum, Orthogonal Polynomials, and Inverse Spectral Theory

Preface	ix
---------	----

Random and Ergodic Schrödinger Operators, Singular Continuous Spectrum

A New Approach to Spectral Gap Problems JEAN BOURGAIN	499
--	-----

Strictly Ergodic Subshifts and Associated Operators DAVID DAMANIK	505
Lyapunov Exponents and Spectral Analysis of Ergodic Schrödinger Operators: A Survey of Kotani Theory and Its Applications DAVID DAMANIK	539
Spectral Properties of Schrödinger Operators with Decaying Potentials SERGEY A. DENISOV and ALEXANDER KISELEV	565
On the Formation of Gaps in the Spectrum of Schrödinger Operators with Quasi-Periodic Potentials MICHAEL GOLDSTEIN and WILHELM SCHLAG	591
Ergodic Schrödinger Operators (on one foot) SVETLANA JITOMIRSKAYA	613
The Integrated Density of States for Random Schrödinger Operators WERNER KIRSCH and BERND METZGER	649
Exotic Spectra: A Review of Barry Simon's Central Contributions YORAM LAST	697
Orthogonal Polynomials, Inverse Spectral Theory	
Riemann–Hilbert Methods in the Theory of Orthogonal Polynomials PERCY DEIFT	715
Inverse Spectral Theory as Influenced by Barry Simon FRITZ GESZTESY	741
Orthogonal Polynomials: From Jacobi to Simon LEONID GOLINSKII and VILMOS TOTIK	821
Orthogonal Polynomials: The First Minutes SERGEY KHRUSHCHEV	875
Spectral Theory via Sum Rules ROWAN KILLIP	907
Barry Simon's List of Publications	931

Preface

This Festschrift had its origins in a conference called SimonFest held at Caltech, March 27–31, 2006, but it is not a Proceedings volume in the usual sense. Barry Simon requested that both his birthday conference and this Festschrift concentrate not so much on what the speaker or writer has done recently, but instead on reviews of the state of the art, with a focus on recent developments and open problems. While the number of speakers at Simonfest was originally limited by the number of hour slots in a full week, the contributions to this Festschrift contain a few additionally selected reviews. In the end, the bulk of the articles in this Festschrift are of this state of the art survey form with a few that instead review Barry's contributions to a particular area.

In Part 1, the focus is on the areas of Quantum Field Theory, Statistical Mechanics, Nonrelativistic Two-Body and N -Body Quantum Systems, Resonances, Electric and Magnetic Fields, and Semiclassical Limit. Here in Part 2, the focus is on the areas of Random and Ergodic Schrödinger Operators, Singular Continuous Spectrum, Orthogonal Polynomials, and Inverse Spectral Theory.

For a detailed preface, including a short biography of Barry Simon, we refer the reader to Part 1 of this two-volume Festschrift.

We are grateful to Sergei Gelfand, Christine Thivierge, and the staff at AMS for their support throughout the preparations of this Festschrift. We also thank all authors for their contributions and the referees for their invaluable assistance.

We sincerely thank the following sponsors of SimonFest for their financial support: Caltech's Center for the Mathematics of Information (CMI); Caltech's Division of Physics, Mathematics, and Astronomy; International Association of Mathematical Physics (IAMP); International Union of Pure and Applied Physics (IUPAP); U.S. National Science Foundation (NSF).

Percy Deift
Cherie Galvez
Fritz Gesztesy
Peter Perry
Wilhelm Schlag

October 2006

Random and Ergodic
Schrödinger Operators,
Singular Continuous Spectrum

A New Approach to Spectral Gap Problems

Jean Bourgain

To Barry

ABSTRACT. Based on purely analytical methods, we exhibit new families of expanders in $SL_2(p)$ (p prime) and $SU(2)$, contributing to conjectures of Lubotzky and Sarnak. This is a report on joint work with Gamburd.

CONTENTS

- 1. Introduction
 - 2. Results
 - 3. Methods
 - 4. Further Remarks
- References

1. Introduction

Given an undirected d -regular graph \mathcal{G} and a subset X of V , the *expansion of X* , $c(X)$, is defined to be the ratio $|\partial(X)|/|X|$, where $\partial(X) = \{y \in \mathcal{G} : \text{distance}(y, X) = 1\}$. The *expansion coefficient* of a graph \mathcal{G} is defined as follows:

$$c(\mathcal{G}) = \inf\{c(X) \mid |X| < \frac{1}{2}|\mathcal{G}|\}.$$

A family of d -regular graphs $\mathcal{G}_{n,d}$ forms a family of C -expanders if there is a fixed positive constant C , such that

$$\liminf_{n \rightarrow \infty} c(\mathcal{G}_{n,d}) \geq C. \tag{1.1}$$

The *adjacency matrix* of \mathcal{G} , $A(\mathcal{G})$, is the $|\mathcal{G}| \times |\mathcal{G}|$ matrix, with rows and columns indexed by vertices of \mathcal{G} , such that the x, y entry is 1 if and only if x and y are adjacent and 0 otherwise. Using the discrete Cheeger–Buser inequality, the condition

2000 *Mathematics Subject Classification.* 22E45, 42Axx, 54H15, 81Q30.

Key words and phrases. Hecke operators, sphere, rotations, spectrum.

The author was supported in part by NSF grant DMS-0401277 and a Sloan Foundation Fellowship.

(1.1) can be rewritten in terms of the second largest eigenvalue of the adjacency matrix $A(\mathcal{G})$ as follows:

$$\limsup_{n \rightarrow \infty} \lambda_1(A_{n,d}) < d. \tag{1.2}$$

Given a finite group G with a symmetric set of generators S , the Cayley graph $\mathcal{G}(G, S)$ is a graph which has elements of G as vertices and which has an edge from x to y if and only if $x = \sigma y$ for some $\sigma \in S$. Let S be a set of elements in $SL(2, \mathbb{Z})$. If $\langle S \rangle$, the group generated by S , is a finite index subgroup of $SL(2, \mathbb{Z})$, Selberg’s theorem [17] implies that $\mathcal{G}(SL_2(\mathbb{F}_p), S_p)$ (where S_p is a natural projection of S modulo p) form a family of expanders as $p \rightarrow \infty$. A basic problem, posed by Lubotzky [13, 14] and Lubotzky and Weiss [15], is whether Cayley graphs of $SL_2(\mathbb{F}_p)$ are expanders with respect to other generating sets.

In [9] it is proved that if S is a set of elements in $SL(2, \mathbb{Z})$ such that $\langle S \rangle$ is a subgroup of $SL(2, \mathbb{Z})$, whose Hausdorff dimension of the limit set is greater than $\frac{5}{6}$, then $\mathcal{G}(SL_2(\mathbb{F}_p), S_p)$ form a family of expanders.

In [2] we prove that Cayley graphs of $SL_2(p)$ are expanders with respect to projection of fixed elements in $SL_2(\mathbb{Z})$ generating a nonelementary subgroup (that is, a subgroup whose limit set consists of more than two points), and with respect to elements chosen at random in $SL_2(p)$.

The question of the spectral gap for finitely generated subgroups of $SU(2)$ is motivated in part by the problem, posed by Ruziewicz in 1921, of whether Lebesgue measure on the n -sphere is the unique *finitely* additive rotation-invariant measure defined on the Lebesgue subsets; it is also of interest in connection with problems in quantum computation [7]. As is well-known, the existence of a finitely generated subgroup with a spectral gap implies the affirmative answer. For $n = 1$ the answer is negative, using essentially the amenability of $SO(2)$. For $n > 3$ the affirmative answer was obtained in 1980–1981 by Margulis and Sullivan, who used Kazhdan’s property (T) . In 1984 Drinfeld established the affirmative answer in the most difficult case of $n = 2$ by providing the *existence* of an element in the group ring of $SU(2)$ which has a spectral gap.

Drinfeld’s method used some sophisticated machinery from the theory of automorphic representations (in particular, Deligne’s solution of the Ramanujan conjectures.) In [10] a new robust method establishing that certain elements z in the group ring of $SU(2)$ have a spectral gap was presented, and consequently the spectral gap property was proven to hold for many subgroups defined via integral Hamilton quaternions.

In [3] we prove the spectral gap property for free subgroups of $SU(2)$ generated by elements satisfying a noncommutative diophantine property, in particular, for free subgroups generated by elements with algebraic entries. Our method, following the approach in [10], first exploits the trace formula to reduce the question of the spectral gap to estimating from above the number of returns to a small neighborhood of identity. In [10] the required upper bound was obtained by reduction to an appropriate arithmetic problem. The novelty of our approach is to derive the required upper bound by utilizing the tools of additive combinatorics.

2. Results

First consider the expander problem in $SL_2(p)$. Our first result resolves the question completely for projections of fixed elements in $SL(2, \mathbb{Z})$.

THEOREM 1. *Let S be a set of elements in $SL(2, \mathbb{Z})$. Then $\mathcal{G}(SL_2(\mathbb{F}_p), S_p)$ form a family of expanders if and only if $\langle S \rangle$ is nonelementary, that is, the limit set of $\langle S \rangle$ consists of more than two points (equivalently, $\langle S \rangle$ does not contain a solvable subgroup of finite index).*

Our second result resolves the question for random Cayley graphs of $SL_2(\mathbb{F}_p)$. (Given a group G , a random $2k$ -regular Cayley graph of G is the Cayley graph $\mathcal{G}(G, \sigma \cup \sigma^{-1})$ where σ is a set of k elements from G , selected independently and uniformly at random.)

THEOREM 2. *For any $k \geq 2$, random Cayley graphs of $SL_2(\mathbb{F}_p)$ on k generators are expanders.*

Theorems 1 and 2 are consequences of the following quantitative result (recall that the girth of a graph is the length of a shortest cycle).

THEOREM 3. *Let S be a symmetric set in $SL_2(\mathbb{F}_p)$ and assume*

$$\text{girth}(\mathcal{G}(SL_2(\mathbb{F}_p), S)) > c \log p, \quad c > 0$$

an arbitrary given constant. Then the expansion coefficient $c(\mathcal{G}(SL_2(\mathbb{F}_p), S)) > c_1(c) > 0$ (for p sufficiently large).

Next, we discuss our results for $SU(2)$. Let us first recall the notion of ‘diophantine elements’ introduced in [10].

DEFINITION. *For $k \geq 2$, we say that $g_1, g_2, \dots, g_k \in G$ are diophantine (or satisfy a noncommutative diophantine condition) if there is $D = D(g_1, \dots, g_k) > 0$ such that for any $m \geq 1$ and a word R_m in g_1, g_2, \dots, g_k of length m with $R_m \neq \pm e$, we have*

$$\|R_m \pm e\| \geq D^{-m}. \tag{2.1}$$

Here

$$\left\| \begin{bmatrix} a & b \\ c & d \end{bmatrix} \right\| = |a|^2 + |b|^2 + |c|^2 + |d|^2.$$

Note that for $g \in G$, we have

$$\|g \pm e\|^2 = 2 |\text{trace}(g) \mp 2|. \tag{2.2}$$

THEOREM 4. *Let $\{g_1, \dots, g_k\}$ be a set of elements in $SU(2)$ generating a free group and satisfying a noncommutative diophantine property. Then*

$$z_{g_1, \dots, g_k} = g_1 + g_1^{-1} + \dots + g_k + g_k^{-1} \tag{2.3}$$

has a spectral gap.

The following corollary is an immediate consequence of Theorem 4 and Proposition 4.3 in [10] establishing that elements with algebraic entries are diophantine.

COROLLARY 5. *If $\{g_1, \dots, g_k\}$ are elements with algebraic entries (that is, $g_1, \dots, g_k \in G \cap M_2(\mathbb{Q})$) generating a free group, then z_{g_1, \dots, g_k} has a spectral gap.*

3. Methods

The proof of Theorem 3 consists of two crucial ingredients. The first one is the fact that nontrivial eigenvalues of $\mathcal{G}(SL_2(\mathbb{F}_p), S)$ must appear with high multiplicity. This follows from a result going back to Frobenius, asserting that the smallest dimension of a nontrivial irreducible representation of $SL_2(\mathbb{F}_p)$ is $\frac{p-1}{2}$, which is large compared to the size of the group (which is of order p^3). The second crucial ingredient is an upper bound on the number of short closed geodesics, or, equivalently, the number of returns to identity for random walks of length of order $\log |G|$.

The idea of obtaining spectral gap results by exploiting high multiplicity together with the upper bound on the number of short closed geodesics is due to Sarnak and Xue [16]; it was subsequently applied in [9]. In these works the upper bound was achieved by reduction to an appropriate diophantine problem. The novelty of our approach is to derive the upper bound by utilizing the tools of additive combinatorics. In particular, we make crucial use of the noncommutative version of the Balog–Szemerédi–Gowers lemma, obtained by Tao [18, 19], and of the result of Helfgott [11] asserting that subsets of $SL_2(\mathbb{F}_p)$ grow rapidly under multiplication. Helfgott’s paper [11], which served as a starting point and an inspiration for our work, builds crucially on sum-product estimates in finite fields due to Bourgain, Glibichuk and Konyagin [4] and Bourgain, Katz and Tao [5]. We state the result

PROPOSITION 6 ([4, 5]). *For all $\varepsilon > 0$, there is $\delta > 0$ such that if $A \subset \mathbb{F}_p$ (prime), $1 < |A| < p^{1-\varepsilon}$, then*

$$|A + A| + |A \cdot A| > c|A|^{1+\delta}.$$

The structure of the proof of Theorem 4 is very similar to Theorem 3. As mentioned, the trace formula is used the same way as in [10] to reduce the gap problem to estimating the number of returns to a δ -neighborhood of the identity. A bound on these returns follows from an estimate on the convolution powers $\|\nu^{(\ell)} * P_\delta\|_2$ where $\nu = \frac{1}{2k} \sum_{s=1}^k (\delta_{g_s} + \delta_{g_s^{-1}})$ and $P_\delta = \frac{\chi_{B(1,\delta)}}{|B(1,\delta)|}$ with $B(1,\delta) = \{x \in SU(2) \mid \|x - 1\| < \delta\}$. That estimate itself results from a ‘product theorem’ for subsets of $SU(2)$ in a similar vein as Helfgott’s [11] (both the statement and the argument). But in the present case, the formulations for a compact group require ‘measure’ and ‘metrical entropy’ rather than ‘cardinality’ in the finite case. The most significant ingredient at this stage of the proof is the metrical counterpart of the finite fields sum-product theorem from [5]. The relevant property is a slight variant of the ‘discretized ring conjecture’ from [12], proven in [1]. The statement is as follows:

PROPOSITION 7. *For all $0 < \sigma < 1$ and $\kappa > 0$, there is $\varepsilon > 0$ such that the following holds. Let $\delta > 0$ be a small number and $A \subset [-1, 1]$ a union of size- δ intervals satisfying $|A| = \delta^{1-\sigma}$ and $\max_a |A \cap B(a, \rho)| < \rho^\kappa |A|$ for all $\delta < \rho < \delta^\varepsilon$. Then*

$$|A + A| + |A \cdot A| > \delta^{1-\sigma-\varepsilon}. \tag{3.1}$$

4. Further Remarks

1. Proofs of the above mentioned results appear in [2, 3].
2. If we fix $S \subset SL_2(\mathbb{Z})$ generating a free group, it is reasonable to expect that $c(\mathcal{G}(SL_2(\mathbb{Z}_q), S)) > \delta(S) > 0$, $\mathbb{Z}_q = \mathbb{Z}/q\mathbb{Z}$, for all positive integers q , coprime to some finite set (depending on S) of integers. This follows from Theorem 3 if q is prime. Without entering into details, the same basic scheme may be applied in general. The sum-product theorem in \mathbb{F}_p needs to be substituted here by the corresponding result in \mathbb{Z}_q (which is necessarily more restrictive for composite q). These results are presently available (and were developed for slightly different purposes). Precise statements and details will appear shortly.
3. Apart from the Ruziewicz problem and the Solovay–Kitaev algorithm, Theorem 4 and Corollary 5 are also relevant to a number of other problems. For instance, an affirmative solution is obtained to the question considered in [8] on the uniform distribution of the orientations in the quaquaversal tilings of \mathbb{R}^3 , introduced in [6].
4. Theorem 3 has extensions to the case when p is not prime, for instance, to square free moduli q . These have further applications to number theory (combined with the Selberg sieve) that won't be discussed here.
5. It is likely similar methods will also apply to SL_n for $n > 2$, but this awaits certain further developments.

References

- [1] J. Bourgain, *On the Erdős–Volkmann and Katz–Tao ring conjectures*, *Geom. Funct. Anal.* **13** (2003), 334–365.
- [2] J. Bourgain and A. Gamburd, *Uniform expansion bounds for Cayley graphs of $SL_2(\mathbb{F}_p)$* , to appear in *Annals of Math.*
- [3] J. Bourgain and A. Gamburd, *On the spectral gap for finitely-generated subgroups of $SU(2)$* , preprint.
- [4] J. Bourgain, A. Glibichuk and S. Konyagin, *Estimates for the number of sums and products and for exponential sums in fields of prime order*, to appear in *J. London Math. Soc.*
- [5] J. Bourgain, N. Katz and T. Tao, *A sum-product estimate in finite fields, and applications*, *Geom. Funct. Anal.* **14** (2004), 27–57.
- [6] J. H. Conway and C. Radin, *Quaquaversal tilings and rotations*, *Invent. Math.* **132** (1998), 179–188.
- [7] C. M. Dawson and M. A. Nielsen, *The Solovay–Kitaev algorithm*, *Quantum Inf. Comput.* **6** (2006), 81–95.
- [8] B. Draco, L. Sadun and D. Van Wieren, *Growth rates in the quaquaversal tiling*, *Discrete Comput. Geom.* **23** (2000), 419–435.
- [9] A. Gamburd, *On the spectral gap for infinite index “congruence” subgroups of $SL_2(\mathbb{Z})$* , *Israel J. Math.* **127**, (2002) 157–200.
- [10] A. Gamburd, D. Jakobson and P. Sarnak, *Spectra of elements in the group ring of $SU(2)$* , *J. Eur. Math. Soc. (JEMS)* **1**, 1999, 51–85.
- [11] H. A. Helfgott, *Growth and generation in $SL_2(\mathbb{Z}/p\mathbb{Z})$* , to appear in *Annals of Math.*
- [12] N. Katz and T. Tao, *Some connections between Falconer’s distance set conjecture and sets of Furstenberg type*, *New York J. Math.* **7** (2001), 149–187 (electronic).
- [13] A. Lubotzky, *Discrete Groups, Expanding Graphs and Invariant Measures*, *Progress in Mathematics*, **125**, Birkhäuser Verlag, Basel, 1994.
- [14] A. Lubotzky, *Cayley graphs: eigenvalues, expanders and random walks*, *Surveys in Combinatorics*, 1995 (Stirling), pp. 155–189, *London Math. Soc. Lecture Note Ser.*, **218**, Cambridge University Press, Cambridge, 1995.

- [15] A. Lubotzky and B. Weiss, *Groups and expanders*, Expanding Graphs (Princeton, NJ, 1992), pp. 95–109, DIMACS Ser. Discrete Math. Theoret. Comput. Sci., **10**, American Mathematical Society, Providence, RI, 1993.
- [16] P. Sarnak and X. X. Xue, *Bounds for multiplicities of automorphic representations*, Duke Math. J. **64** (1991), 207–227.
- [17] A. Selberg, *On the estimation of Fourier coefficients of modular forms*, Proc. Sympos. Pure Math. **VIII**, pp. 1–15, American Mathematical Society, Providence, RI, 1965.
- [18] T. Tao, *Non-commutative sum set estimates*, preprint.
- [19] T. Tao and V. Vu, *Additive Combinatorics*, Cambridge Studies in Advanced Mathematics, **105**, Cambridge University Press, Cambridge, 2006.

INSTITUTE FOR ADVANCED STUDY, EINSTEIN DRIVE, PRINCETON, NJ 08540, U.S.A.
E-mail address: bourgain@math.ias.edu

Strictly Ergodic Subshifts and Associated Operators

David Damanik

Dedicated to Barry Simon on the occasion of his 60th birthday

ABSTRACT. We consider ergodic families of Schrödinger operators over base dynamics given by strictly ergodic subshifts on finite alphabets. It is expected that the majority of these operators have purely singular continuous spectrum supported on a Cantor set of zero Lebesgue measure. These properties have indeed been established for large classes of operators of this type over the course of the last twenty years. We review the mechanisms leading to these results and briefly discuss analogues for CMV matrices.

CONTENTS

1. Introduction
 2. Strictly Ergodic Subshifts
 3. Associated Schrödinger Operators and Basic Results
 4. Absence of Absolutely Continuous Spectrum
 5. Zero-Measure Spectrum
 6. Absence of Point Spectrum
 7. Quantum Dynamics
 8. CMV Matrices Associated with Subshifts
 9. Concluding Remarks
- References

1. Introduction

When I was a student in the mid-1990's at the Johann Wolfgang Goethe Universität in Frankfurt, my advisor Joachim Weidmann and his students and postdocs would meet in his office for coffee every day and discuss mathematics and life. One day we walked in and found a stack of preprints on the coffee table. What now must

2000 *Mathematics Subject Classification.* Primary 82B44, 47B80; Secondary 47B36, 81Q10.

Key words and phrases. subshifts, ergodic Schrödinger operators, Cantor spectrum, singular continuous spectrum.

This work was supported in part by NSF grant DMS-0500910.

seem like an ancient practice was not entirely uncommon in those days: In addition to posting preprints on the archives, people would actually send out hardcopies of them to their peers around the world.

In this particular instance, Barry Simon had sent a series of preprints, all dealing with singular continuous spectrum. At the time I did not know Barry personally but was well aware of his reputation and immense research output. I was intrigued by these preprints. After all, we had learned from various sources (including the Reed–Simon books!) that singular continuous spectrum is sort of a nuisance and something whose absence should be proven in as many cases as possible. Now we were told that singular continuous spectrum is generic?

Soon after reading through the preprint series it became clear to me that my thesis topic should have something to do with this beast: singular continuous spectrum. Coincidentally, only a short while later I came across a beautifully written paper by Sütő [130] that raised my interest in the Fibonacci operator. I had studied papers on the almost Mathieu operator earlier. For that operator, singular continuous spectrum does occur, but only in very special cases, that is, for special choices of the coupling constant, the frequency, or the phase. In the Fibonacci case, however, singular continuous spectrum seemed to be the rule. At least there was no sensitive dependence on the coupling constant or the frequency as I learned from [15, 130, 131].

Another feature, which occurs in the almost Mathieu case, but only at special coupling, seemed to be the rule for the Fibonacci operator: zero-measure spectrum.

So I set out to understand what about the Fibonacci operator was responsible for this persistent occurrence of zero-measure singular continuous spectrum. Now, some ten years later, I still do not really understand it. In fact, as is always the case, the more you understand (or think you understand), the more you realize how much else is out there, still waiting to be understood.¹

Thus, this survey is meant as a snapshot of the current level of understanding of things related to the Fibonacci operator and also as a thank-you to Barry for having had the time and interest to devote a section or two of his OPUC book to subshifts and the Fibonacci CMV matrix. Happy Birthday, Barry, and thank you for being an inspiration to so many generations of mathematical physicists!

2. Strictly Ergodic Subshifts

In this section we define strictly ergodic subshifts over a finite alphabet and discuss several examples that have been studied from many different perspectives in a great number of papers.

2.1. Basic Definitions. We begin with the definitions of the the basic objects:

DEFINITION 2.1 (full shift). *Let A be a finite set, called the alphabet. The two-sided infinite sequences with values in A form the full shift $A^{\mathbb{Z}}$. We endow A with the discrete topology and the full shift with the product topology.*

DEFINITION 2.2 (shift transformation). *The shift transformation T acts on the full shift by $[T\omega]_n = \omega_{n+1}$.*

¹Most recently, I have come to realize that I do not understand why the Lyapunov exponent vanishes on the spectrum, even at large coupling. Who knows what will be next...

DEFINITION 2.3 (subshift). *A subset Ω of the full shift is called a subshift if it is closed and T -invariant.*

Thus, our base dynamical systems will be given by (Ω, T) , where Ω is a subshift and T is the shift transformation. This is a special class of topological dynamical systems that is interesting in its own right. Basic questions regarding them concern the structure of orbits and invariant (probability) measures. The situation is particularly simple when orbit closures and invariant measures are unique:

DEFINITION 2.4 (minimality). *Let Ω be a subshift and $\omega \in \Omega$. The orbit of ω is given by $\mathcal{O}_\omega = \{T^n\omega : n \in \mathbb{Z}\}$. If \mathcal{O}_ω is dense in Ω for every $\omega \in \Omega$, then Ω is called minimal.*

DEFINITION 2.5 (unique ergodicity). *Let Ω be a subshift. A Borel measure μ on Ω is called T -invariant if $\mu(T(A)) = \mu(A)$ for every Borel set $A \subseteq \Omega$. Ω is called uniquely ergodic if there is a unique T -invariant Borel probability measure on Ω .*

By compactness of Ω , the set of T -invariant Borel probability measure on Ω is non-empty. It is also convex and the extreme points are exactly the ergodic measures, that is, probability measures for which $T(A) = A$ implies that either $\mu(A) = 0$ or $\mu(A) = 1$. Thus, a subshift is uniquely ergodic precisely when there is a unique ergodic measure on it.

We will focus our main attention on subshifts having both of these properties. For convenience, one often combines these two notions into one:

DEFINITION 2.6 (strict ergodicity). *A subshift Ω is called strictly ergodic if it is both minimal and uniquely ergodic.*

2.2. Examples of Strictly Ergodic Subshifts. Let us list some classes of strictly ergodic subshifts that have been studied by a variety of authors and from many different perspectives (e.g., symbolic dynamics, number theory, spectral theory, operator algebras, etc.) in the past.

2.2.1. *Subshifts Generated by Sequences.* Here we discuss a convenient way of defining a subshift starting from a sequence $s \in \mathcal{A}^{\mathbb{Z}}$. Since $\mathcal{A}^{\mathbb{Z}}$ is compact, \mathcal{O}_s has a non-empty set of accumulation points, denoted by Ω_s . It is readily seen that Ω_s is closed and T -invariant. Thus, we call Ω_s the *subshift generated by s* . Naturally, we seek conditions on s that imply that Ω_s is minimal or uniquely ergodic.

Every word w (also called block or string) of the form $w = s_m \dots s_{m+n-1}$ with $m \in \mathbb{Z}$ and $n \in \mathbb{Z}^+ = \{1, 2, 3, \dots\}$ is called a *subword of s* (of length n , denoted by $|w|$). We denote the set of all subwords of s of length n by $\mathcal{W}_s(n)$ and let

$$\mathcal{W}_s = \bigcup_{n \geq 1} \mathcal{W}_s(n).$$

If $w \in \mathcal{W}_s(n)$, let $\dots < m_{-1} < 0 \leq m_0 < m_1 < \dots$ be the integers m for which $s_m \dots s_{m+n-1} = w$. The sequence s is called *recurrent* if $m_n \rightarrow \pm\infty$ as $n \rightarrow \pm\infty$ for every $w \in \mathcal{W}_s$. A recurrent sequence s is called *uniformly recurrent* if $(m_{j+1} - m_j)_{j \in \mathbb{Z}}$ is bounded for every $w \in \mathcal{W}_s$. Finally, a uniformly recurrent sequence s is called *linearly recurrent* if there is a constant $C < \infty$ such that for every $w \in \mathcal{W}_s$, the gaps $m_{j+1} - m_j$ are bounded by $C|w|$.

We say that $w \in \mathcal{W}_s$ *occurs in s with a uniform frequency* if there is $d_s(w) \geq 0$ such that, for every $k \in \mathbb{Z}$,

$$d_s(w) = \lim_{n \rightarrow \infty} \frac{1}{n} |\{m_j\}_{j \in \mathbb{Z}} \cap [k, k+n]|,$$

and the convergence is uniform in k .

For results concerning the minimality and unique ergodicity of the subshift Ω_s generated by a sequence s , we recommend the book by Queffélec [120]; see in particular Section IV.2. Let us recall the main findings.

PROPOSITION 2.7. *If s is uniformly recurrent, then Ω_s is minimal. Conversely, if Ω is minimal, then every $\omega \in \Omega$ is uniformly recurrent. Moreover, $\mathcal{W}_{\omega_1} = \mathcal{W}_{\omega_2}$ for every $\omega_1, \omega_2 \in \Omega$.*

The last statement permits us to define a set \mathcal{W}_Ω for any minimal subshift Ω so that $\mathcal{W}_\Omega = \mathcal{W}_\omega$ for every $\omega \in \Omega$.

PROPOSITION 2.8. *Let s be recurrent. Then, Ω_s is uniquely ergodic if and only if each subword of s occurs with a uniform frequency.*

As a consequence, Ω_s is strictly ergodic if and only if each subword w of s occurs with a uniform frequency $d_s(w) > 0$.

An interesting class of strictly ergodic subshifts is given by those subshifts that are generated by linearly recurrent sequences [65, 104]:

PROPOSITION 2.9. *If s is linearly recurrent, then Ω_s is strictly ergodic.*

2.2.2. Sturmian Sequences. Suppose s is a uniformly recurrent sequence. We saw above that Ω_s is a minimal subshift and all elements of Ω_s have the same set of subwords, $\mathcal{W}_{\Omega_s} = \mathcal{W}_s$. Let us denote the cardinality of $\mathcal{W}_s(n)$ by $p_s(n)$. The map $\mathbb{Z}^+ \rightarrow \mathbb{Z}^+$, $n \mapsto p_s(n)$ is called the *complexity function of s* (also called *factor, block or subword complexity function*).

It is clear that a periodic sequence gives rise to a bounded complexity function. It is less straightforward that every non-periodic sequence gives rise to a complexity function that grows at least linearly. This fact is a consequence of the following celebrated theorem due to Hedlund and Morse [116]:

THEOREM 2.10. *If s is recurrent, then the following statements are equivalent:*

- (i) *s is periodic, that is, there exists k such that $s_m = s_{m+k}$ for every $m \in \mathbb{Z}$.*
- (ii) *p_s is bounded, that is, there exists p such that $p_s(n) \leq p$ for every $n \in \mathbb{Z}^+$.*
- (iii) *There exists $n_0 \in \mathbb{Z}^+$ such that $p_s(n_0) \leq n_0$.*

PROOF. The implications (i) \Rightarrow (ii) and (ii) \Rightarrow (iii) are obvious, so we only need to show (iii) \Rightarrow (i).

Let $R_s(n)$ be the directed graph with $p_s(n)$ vertices and $p_s(n+1)$ edges which is defined as follows. Every subword $w \in \mathcal{W}_s(n)$ corresponds to a vertex of $R_s(n)$. Every $\tilde{w} \in \mathcal{W}_s(n+1)$ generates an edge of $R_s(n)$ as follows. Write $\tilde{w} = axb$, where $a, b \in \mathcal{A}$ and x is a (possibly empty) string. Then draw an edge from the vertex ax to the vertex xb .

We may assume that \mathcal{A} has cardinality at least two since otherwise the theorem is trivial. Thus, $p_s(1) \geq 2 > 1$. Obviously, p_s is non-decreasing. Thus, by assumption (iii), there must be $1 \leq n_1 < n_0$ such that $p_s(n_1) = p_s(n_1 + 1)$. Consider the graph $R_s(n_1)$. Since s is recurrent, there must be a directed path from w_1 to w_2 for every pair $w_1, w_2 \in \mathcal{W}_s(n_1)$. On the other hand, $R_s(n_1)$ has the same number of vertices and edges. It follows that $R_s(n_1)$ is a simple cycle and hence s is periodic of period $p_s(n_1)$. □

The graph $R_s(n)$ introduced in the proof above is called the *Rauzy graph* associated with s and n . It is an important tool for studying (so-called) combinatorics

on words. This short proof of the Hedlund–Morse Theorem is just one of its many applications.

COROLLARY 2.11. *If s is recurrent and not periodic, then $p_s(n) \geq n + 1$ for every $n \in \mathbb{Z}^+$.*

This raises the question whether aperiodic sequences of minimal complexity exist.

DEFINITION 2.12. *A sequence s is called Sturmian if it is recurrent and satisfies $p_s(n) = n + 1$ for every $n \in \mathbb{Z}^+$.*

Remarks. (a) There are non-recurrent sequences s with complexity $p_s(n) = n + 1$. For example, $s_n = \delta_{n,0}$. The subshifts generated by such sequences are trivial and we therefore restrict our attention to recurrent sequences.

(b) We have seen that growth strictly between bounded and linear is impossible for a complexity function. It is an interesting open problem to characterize the increasing functions from \mathbb{Z}^+ to \mathbb{Z}^+ that arise as complexity functions.

Note that a Sturmian sequence is necessarily defined on a two-symbol alphabet \mathcal{A} . Without loss of generality, we restrict our attention to $\mathcal{A} = \{0, 1\}$. The following result gives an explicit characterization of all Sturmian sequences with respect to this normalization; compare [113, Theorem 2.1.13].

THEOREM 2.13. *A sequence $s \in \{0, 1\}^{\mathbb{Z}}$ is Sturmian if and only if there are $\theta \in (0, 1)$ irrational and $\phi \in [0, 1)$ such that either*

$$s_n = \chi_{[1-\theta, 1)}(n\theta + \phi) \text{ or } s_n = \chi_{(1-\theta, 1]}(n\theta + \phi) \tag{1}$$

for all $n \in \mathbb{Z}$.

Remark. In (1), we consider the 1-periodic extension of the function $\chi_{[1-\theta, 1)}(\cdot)$ (resp., $\chi_{(1-\theta, 1]}(\cdot)$) on $[0, 1)$.

Each sequence of the form (1) generates a subshift. The following theorem shows that the resulting subshift only depends on θ . A proof of this result may be found, for example, in the appendix of [44].

THEOREM 2.14. *Assume $\theta \in (0, 1)$ is irrational, $\phi \in [0, 1)$, and $s_n = \chi_{[1-\theta, 1)}(n\theta + \phi)$. Then the subshift generated by s is given by*

$$\Omega_s = \left\{ n \mapsto \chi_{[1-\theta, 1)}(n\theta + \tilde{\phi}) : \tilde{\phi} \in [0, 1) \right\} \cup \left\{ n \mapsto \chi_{(1-\theta, 1]}(n\theta + \tilde{\phi}) : \tilde{\phi} \in [0, 1) \right\}.$$

Moreover, Ω_s is strictly ergodic.

Let us call a subshift *Sturmian* if it is generated by a Sturmian sequence. We see from the previous theorem that there is a one-to-one correspondence between irrational numbers θ and Sturmian subshifts. We call θ the *slope* of the subshift.

Example (Fibonacci case). The Sturmian subshift corresponding to the inverse of the golden mean,

$$\theta = \frac{\sqrt{5} - 1}{2},$$

is called the *Fibonacci subshift* and its elements are called *Fibonacci sequences*.

An important property of Sturmian sequences is their hierarchical, or *S-adic*, structure. That is, there is a natural level of hierarchies such that on each level,

there is a unique decomposition of the sequence into blocks of two types. The starting level is just the decomposition into individual symbols. Then, one may pass from one level to the next by a set of rules that is determined by the coefficients in the continued fraction expansion of the slope θ .

Let

$$\theta = \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \dots}}} \tag{2}$$

be the continued fraction expansion of θ with uniquely determined $a_k \in \mathbb{Z}^+$. Truncation of this expansion after k steps yields rational numbers p_k/q_k that obey

$$p_0 = 0, \quad p_1 = 1, \quad p_k = a_k p_{k-1} + p_{k-2}, \tag{3}$$

$$q_0 = 1, \quad q_1 = a_1, \quad q_k = a_k q_{k-1} + q_{k-2}. \tag{4}$$

These rational numbers are known to be best approximants to θ . See Khinchin [94] for background on continued fraction expansions.

We define words $(w_k)_{k \in \mathbb{Z}^+}$ over the alphabet $\{0, 1\}$ as follows:

$$w_0 = 0, \quad w_1 = 0^{a_1-1}1, \quad w_{k+1} = w_k^{a_{k+1}}w_{k-1} \text{ for } k \geq 1. \tag{5}$$

THEOREM 2.15. *Let Ω be a Sturmian subshift with slope θ and let the words w_k be defined by (2) and (5). Then, for every $k \in \mathbb{Z}^+$, each $\omega \in \Omega$ has a unique partition, called the k -partition of ω , into blocks of the form w_k or w_{k-1} . In this partition, blocks of type w_k occur with multiplicity a_{k+1} or $a_{k+1} + 1$ and blocks of type w_{k-1} occur with multiplicity one.*

SKETCH OF PROOF. The first step is to use the fact that p_k/q_k are best approximants to show that the restriction of $\chi_{[1-\theta, 1)}(n\theta)$ to the interval $[1, q_k]$ is given by w_k , $k \in \mathbb{Z}^+$; compare [15]. The recursion (5) therefore yields a k -partition of $\chi_{[1-\theta, 1)}(n\theta)$ on $[1, \infty)$. Since every $\omega \in \Omega$ may be obtained as an accumulation point of shifts of this sequence, it can then be shown that a unique partition of ω is induced; see [38]. The remaining claims follow quickly from the recursion (5) \square

Example (Fibonacci case, continued). In the Fibonacci case, $a_k = 1$ for every k . Thus, both (p_k) and (q_k) are sequences of Fibonacci numbers (i.e., $p_{k+1} = q_k = F_k$, where $F_0 = F_1 = 1$ and $F_{k+1} = F_k + F_{k-1}$ for $k \geq 1$) and the words w_k are obtained by the simple rule

$$w_0 = 0, \quad w_1 = 1, \quad w_k = w_{k-1}w_{k-2} \text{ for } k \geq 2. \tag{6}$$

Thus, the sequence $(w_k)_{k \in \mathbb{Z}^+}$ is given by 1, 10, 101, 10110, 10110101, ..., which may also be obtained by iterating the rule

$$1 \mapsto 10, \quad 0 \mapsto 1, \tag{7}$$

starting with the symbol 1.

For the proofs omitted in this subsection and much more information on Sturmian sequences and subshifts, we refer the reader to [16, 44, 113, 117].

2.2.3. *Codings of Rotations.* Theorem 2.13 shows that Sturmian sequences are obtained by coding an irrational rotation of the torus according to a partition of the circle into two half-open intervals. It is natural to generalize this and consider codings of rotations with respect to a more general partition of the circle. Thus, let $[0, 1) = I_1 \cup \dots \cup I_l$ be a partition into l half-open intervals. Choosing numbers $\lambda_1, \dots, \lambda_l$, we consider the sequences

$$s_n = \sum_{j=1}^l \lambda_j \chi_{I_j}(n\theta + \phi). \tag{8}$$

Subshifts generated by sequences of this form will be said to be associated with codings of rotations.

THEOREM 2.16. *Let $\theta \in (0, 1)$ be irrational and $\phi \in [0, 1)$. If s is of the form (8), then Ω_s is strictly ergodic. Moreover, the complexity function satisfies $p_s(n) = an + b$ for every $n \geq n_0$ and suitable integers a, b, n_0 .*

See [77] for a proof of strict ergodicity and [1] for a proof of the complexity statement. In fact, the integers a, b, n_0 can be described explicitly; see [1, Theorem 10].

2.2.4. *Arnoux–Rauzy and Episturmian Subshifts.* Let us consider a minimal subshift Ω over the alphabet $\mathcal{A}_m = \{0, 1, 2, \dots, m - 1\}$, where $m \geq 2$. A word $w \in \mathcal{W}_\Omega$ is called *right-special* (resp., *left-special*) if there are distinct symbols $a, b \in \mathcal{A}_m$ such that $wa, wb \in \mathcal{W}_\Omega$ (resp., $aw, bw \in \mathcal{W}_\Omega$). A word that is both right-special and left-special is called *bispecial*. Thus, a word is right-special (resp., left-special) if and only if the corresponding vertex in the Rauzy graph has out-degree (resp., in-degree) ≥ 2 .

Note that the complexity function of a Sturmian subshift obeys $p(n+1) - p(n) = 1$ for every n , and hence for every length, there is a unique right-special factor and a unique left-special factor, each having exactly two extensions.

Arnoux–Rauzy subshifts and episturmian subshifts relax this restriction on the possible extensions somewhat, and they are defined as follows: A minimal subshift Ω is called an *Arnoux–Rauzy subshift* if $p_\Omega(1) = m$ and for every $n \in \mathbb{Z}^+$, there is a unique right-special word in $\mathcal{W}_\Omega(n)$ and a unique left-special word in $\mathcal{W}_\Omega(n)$, both having exactly m extensions. This implies in particular that $p_\Omega(n) = (m - 1)n + 1$. Arnoux–Rauzy subshifts over \mathcal{A}_2 are exactly the Sturmian subshifts.

A minimal subshift Ω is called *episturmian* if \mathcal{W}_Ω is closed under reversal (i.e., for every $w = w_1 \dots w_n \in \mathcal{W}_\Omega$, we have $w^R = w_n \dots w_1 \in \mathcal{W}_\Omega$) and for every $n \in \mathbb{Z}^+$, there is exactly one right-special word in $\mathcal{W}_\Omega(n)$.

PROPOSITION 2.17. *Every Arnoux–Rauzy subshift is episturmian and every episturmian subshift is strictly ergodic.*

See [63, 88, 122, 137] for these results and more information on Arnoux–Rauzy and episturmian subshifts.

2.2.5. *Codings of Interval Exchange Transformations.* Interval exchange transformations are defined as follows. Given a probability vector $\lambda = (\lambda_1, \dots, \lambda_m)$ with $\lambda_i > 0$ for $1 \leq i \leq m$, let $\mu_0 = 0$, $\mu_i = \sum_{j=1}^i \lambda_j$, and $I_i = [\mu_{i-1}, \mu_i)$. Let $\tau \in S_m$, the symmetric group. Then $\lambda^\tau = (\lambda_{\tau^{-1}(1)}, \dots, \lambda_{\tau^{-1}(m)})$ is also a probability vector and we can form the corresponding μ_i^τ and I_i^τ . Denote the unit interval $[0, 1)$ by I .

The (λ, τ) interval exchange transformation is then defined by

$$T : I \rightarrow I, \quad T(x) = x - \mu_{i-1} + \mu_{\tau(i)-1}^{\tau} \text{ for } x \in I_i, \quad 1 \leq i \leq m.$$

It exchanges the intervals I_i according to the permutation τ .

The transformation T is invertible and its inverse is given by the $(\lambda^{\tau}, \tau^{-1})$ interval exchange transformation.

The symbolic coding of $x \in I$ is $\omega_n(x) = i$ if $T^n(x) \in I_i$. This induces a subshift over the alphabet $\mathcal{A} = \{1, \dots, m\}$: $\Omega_{\lambda, \tau} = \{\omega(x) : x \in I\}$. Every Sturmian subshift can be described by the exchange of two intervals.

Keane [90] proved that if the orbits of the discontinuities μ_i of T are all infinite and pairwise distinct, then T is minimal. In this case, the coding is one-to-one and the subshift is minimal and aperiodic. This holds in particular if τ is irreducible and λ is irrational. Here, τ is called irreducible if $\tau(\{1, \dots, k\}) \neq \{1, \dots, k\}$ for every $k < m$ and λ is called irrational if the λ_i are rationally independent.

Keane also conjectured that all minimal interval exchange transformations give rise to a uniquely ergodic system. This was disproved by Keynes and Newton [92] using five intervals, and then by Keane [91] using four intervals (the smallest possible number). The conjecture was therefore modified in [91] and then ultimately proven by Masur [114], Veech [135], and Boshernitzan [19]: For every irreducible $\tau \in S_m$ and for Lebesgue almost every λ , the subshift $\Omega_{\lambda, \tau}$ is uniquely ergodic.

2.2.6. *Substitution Sequences.* All the previous examples were generalizations of Sturmian sequences. We now discuss a class of examples that generalize a certain aspect of the Fibonacci sequence

$$s_n = \chi_{[1-\theta, 0)}(n\theta), \quad \theta = \frac{\sqrt{5} - 1}{2}.$$

We saw above (see (6) and its discussion) that this sequence, restricted to the right half line, is obtained by iterating the map (7). That is,

$$1 \mapsto 10 \mapsto 101 \mapsto 10110 \mapsto 10110101 \mapsto \dots$$

has $(s_n)_{n \geq 1}$ as its limit. In other words, $(s_n)_{n \geq 1}$ is invariant under the substitution rule (7).

DEFINITION 2.18 (substitution). *Denote the set of words over the alphabet \mathcal{A} by \mathcal{A}^* . A map $S : \mathcal{A} \rightarrow \mathcal{A}^*$ is called a substitution. The naturally induced maps on \mathcal{A}^* and $\mathcal{A}^{\mathbb{Z}^+}$ are denoted by S as well.*

- Examples.** (a) Fibonacci: $1 \mapsto 10, 0 \mapsto 1$
 (b) Thue–Morse: $1 \mapsto 10, 0 \mapsto 01$
 (c) Period doubling: $1 \mapsto 10, 0 \mapsto 11$
 (d) Rudin–Shapiro: $1 \mapsto 12, 2 \mapsto 13, 3 \mapsto 42, 4 \mapsto 43$

DEFINITION 2.19 (substitution sequence). *Let S be a substitution. A sequence $s \in \mathcal{A}^{\mathbb{Z}^+}$ is called a substitution sequence if it is a fixed point of S .*

If $S(a)$ begins with the symbol a and has length at least two, it follows that $|S^n(a)| \rightarrow \infty$ as $n \rightarrow \infty$ and $S^n(a)$ has $S^{n-1}(a)$ as a prefix. Thus, the limit of $S^n(a)$ as $n \rightarrow \infty$ defines a substitution sequence s . In the examples above, we obtain the following substitution sequences.

- (a) Fibonacci: $s_F = 1011010110110\dots$
 (b) Thue–Morse: $s_{TM}^{(1)} = 100101100110\dots$ and $s_{TM}^{(0)} = 0110100110010110\dots$

(c) Period doubling: $s_{PD} = 101110101011101110\dots$

(d) Rudin–Shapiro: $s_{RS}^{(1)} = 1213124212134313\dots$ and $s_{RS}^{(4)} = 4342431343421242\dots$

We want to associate a subshift Ω_s with a substitution sequence s . Since the iteration of S on a suitable symbol a naturally defines a one-sided sequence s , we have to alter the definition of Ω_s used above slightly. One possible way is to extend s to a two-sided sequence \bar{s} arbitrarily and then define

$$\Omega_s = \{\omega \in \mathcal{A}^{\mathbb{Z}} : \omega = T^{n_j} \bar{s} \text{ for some sequence } n_j \rightarrow \infty\}.$$

A different way is to define Ω_s to be the set of all ω 's with $\mathcal{W}_\omega \subseteq \mathcal{W}_s$. Below we will restrict our attention to so-called primitive substitutions and for them, these two definitions are equivalent.

To ensure that Ω_s is strictly ergodic, we need to impose some conditions on S . A very popular sufficient condition is primitivity.

DEFINITION 2.20. *A substitution S is called primitive if there is $k \in \mathbb{Z}^+$ such that for every pair $a, b \in \mathcal{A}$, $S^k(a)$ contains the symbol b .*

It is easy to check that our four main examples are primitive. Moreover, if S is primitive, then every power of S is primitive. Thus, even if $S(a)$ does not begin with a for any symbol $a \in \mathcal{A}$, we may replace S by a suitable S^m and then find such an a , which in turn yields a substitution sequence associated with S^m by iteration.

THEOREM 2.21. *Suppose S is primitive and s is an associated substitution sequence. Then, s is linearly recurrent. Consequently, Ω_s is strictly ergodic.*

See [55, 66]. Linear recurrence clearly also implies that $p_s(n) = O(n)$. Fixed points of non-primitive substitutions may have quadratic complexity. However, there are non-primitive substitutions that have fixed points which are linearly recurrent and hence define strictly ergodic subshifts; see [45] for a characterization of linearly recurrent substitution generated subshifts.

2.2.7. Subshifts with Positive Topological Entropy. All the examples discussed so far have linearly bounded complexity. One may wonder if strict ergodicity places an upper bound on the growth of the complexity function. Here we want to mention the existence of strictly ergodic subshifts that have a very fast growing complexity function. In fact, it is possible to have growth that is arbitrarily close to the maximum possible one on a logarithmic scale.

Given a sequence s over an alphabet \mathcal{A} , $|\mathcal{A}| \geq 2$, its (topological) entropy is given by

$$h_s = \lim_{n \rightarrow \infty} \frac{1}{n} \log p_s(n).$$

The existence of the limit follows from the fact that $n \mapsto \log p_s(n)$ is subadditive. Moreover,

$$0 \leq h_s \leq \log |\mathcal{A}|,$$

where $|\cdot|$ denotes cardinality.

The following was shown by Hahn and Katznelson [75]:

THEOREM 2.22. (a) *If s is a uniformly recurrent sequence over the alphabet \mathcal{A} , then $h_s < \log |\mathcal{A}|$.*

(b) *For every $\delta \in (0, 1)$, there are alphabets $\mathcal{A}^{(j)}$ and sequences $s^{(j)}$ over $\mathcal{A}^{(j)}$, $j \in \mathbb{Z}^+$, such that $|\mathcal{A}^{(j)}| \rightarrow \infty$ as $j \rightarrow \infty$, $h_{s^{(j)}} \geq \log [|\mathcal{A}^{(j)}|(1 - \delta)]$, and every $\Omega_{s^{(j)}}$ is strictly ergodic.*

3. Associated Schrödinger Operators and Basic Results

In this section we associate Schrödinger operators with a subshift Ω and a sampling function f mapping Ω to the real numbers. In subsequent sections we will study spectral and dynamical properties of these operators.

Let Ω be a strictly ergodic subshift with invariant measure μ and let $f : \Omega \rightarrow \mathbb{R}$ be continuous. Then, for every $\omega \in \Omega$, we define a potential $V_\omega : \mathbb{Z} \rightarrow \mathbb{R}$ by $V_\omega(n) = f(T^n\omega)$ and a bounded operator H_ω acting on $\ell^2(\mathbb{Z})$ by

$$[H_\omega\psi](n) = \psi(n + 1) + \psi(n - 1) + V_\omega(n)\psi(n).$$

Example. The most common choice for f is $f(\omega) = g(\omega_0)$ with some $g : \mathcal{A} \rightarrow \mathbb{R}$. This is a special case of a locally constant function that is completely determined by the values of ω_n for n 's from a finite window around the origin, that is, f is called *locally constant* if it is of the form $f(\omega) = h(\omega_{-M} \dots \omega_N)$ for suitable integers $M, N \geq 0$ and $h : \mathcal{A}^{N+M+1} \rightarrow \mathbb{R}$. Clearly, every locally constant f is continuous.

The family $\{H_\omega\}_{\omega \in \Omega}$ is an ergodic family of discrete one-dimensional Schrödinger operators in the sense of Carmona and Lacroix [24]. By the general theory it follows that the spectrum and the spectral type of H_ω are μ -almost surely ω -independent [24, Sect. V.2]:

THEOREM 3.1. *There exist sets $\Omega_0 \subseteq \Omega, \Sigma, \Sigma_{pp}, \Sigma_{sc}, \Sigma_{ac} \subseteq \mathbb{R}$ such that $\mu(\Omega_0) = 1$ and*

$$\sigma(H_\omega) = \Sigma \tag{9}$$

$$\sigma_{pp}(H_\omega) = \Sigma_{pp} \tag{10}$$

$$\sigma_{sc}(H_\omega) = \Sigma_{sc} \tag{11}$$

$$\sigma_{ac}(H_\omega) = \Sigma_{ac} \tag{12}$$

for every $\omega \in \Omega_0$.

Here, $\sigma(H), \sigma_{pp}(H), \sigma_{sc}(H), \sigma_{ac}(H)$ denote the spectrum, the closure of the set of eigenvalues, the singular continuous spectrum and the absolutely continuous spectrum of the operator H , respectively.

Since Ω is minimal and f is continuous, a simple argument involving strong approximation shows that (9) even holds everywhere, rather than almost everywhere:

THEOREM 3.2. *For every $\omega \in \Omega, \sigma(H_\omega) = \Sigma$.*

PROOF. By symmetry it suffices to show that for every pair $\omega_1, \omega_2 \in \Omega, \sigma(H_{\omega_1}) \subseteq \sigma(H_{\omega_2})$. Due to minimality, there exists a sequence $(n_j)_{j \geq 1}$ such that $T^{n_j}\omega_2 \rightarrow \omega_1$ as $j \rightarrow \infty$. By continuity of $f, H_{T^{n_j}\omega_2}$ converges strongly to H_{ω_1} as $j \rightarrow \infty$. Thus,

$$\sigma(H_{\omega_1}) \subseteq \overline{\bigcup_{j \geq 1} \sigma(H_{T^{n_j}\omega_2})} = \sigma(H_{\omega_2}).$$

Here, the first step follows by strong convergence and the second step is a consequence of the fact that each of the operators $H_{T^{n_j}\omega_2}$ is unitarily equivalent to H_{ω_2} and hence has the same spectrum. □

Far more subtle is the result that (12) also holds everywhere:

THEOREM 3.3. *For every $\omega \in \Omega, \sigma_{ac}(H_\omega) = \Sigma_{ac}$.*

For strictly ergodic models, such as the ones considered here, there are two proofs of Theorem 3.3 in the literature. It was shown, based on unique ergodicity, by Kotani in [102]. A proof based on minimality was given by Last and Simon in [106].

A map $A \in C(\Omega, \text{SL}(2, \mathbb{R}))$ induces an $\text{SL}(2, \mathbb{R})$ -cocycle over T as follows:

$$\tilde{A} : \Omega \times \mathbb{R}^2 \rightarrow \Omega \times \mathbb{R}^2, (\omega, v) \mapsto (T\omega, A(\omega)v).$$

Note that when we iterate this map n times, we get

$$\tilde{A}^n(\omega, v) = (T^n\omega, A_n(\omega)v),$$

where $A_n(\omega) = A(T^{n-1}\omega) \cdots A(\omega)$. We are interested in the asymptotic behavior of the norm of $A_n(\omega)$ as $n \rightarrow \infty$. The multiplicative ergodic theorem ensures the existence of $\gamma_A \geq 0$, called the *Lyapunov exponent*, such that

$$\begin{aligned} \gamma_A &= \lim_{n \rightarrow \infty} \frac{1}{n} \int \log \|A_n(\omega)\| d\mu(\omega) \\ &= \inf_{n \geq 1} \frac{1}{n} \int \log \|A_n(\omega)\| d\mu(\omega) \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \|A_n(\omega)\| \quad \text{for } \mu\text{-a.e. } \omega. \end{aligned}$$

In the study of the operators H_ω the following cocycles are relevant:

$$A^{f,E}(\omega) = \begin{pmatrix} E - f(\omega) & -1 \\ 1 & 0 \end{pmatrix},$$

where f is as above and E is a real number, called the *energy*. We regard f as fixed and write $\gamma(E)$ instead of $\gamma_{A^{f,E}}$ to indicate that our main interest is in the mapping $E \mapsto \gamma(E)$. Let

$$\mathcal{Z} = \{E \in \mathbb{R} : \gamma(E) = 0\}.$$

Note that we leave the dependence on Ω and f implicit.

These cocycles are important in the study of H_ω because $A_n^{f,E}$ is the transfer matrix for the associated difference equation. That is, a sequence u solves

$$u(n+1) + u(n-1) + V_\omega(n)u(n) = Eu(n) \tag{13}$$

if and only if it solves

$$\begin{pmatrix} u(n) \\ u(n-1) \end{pmatrix} = A_n^{f,E} \begin{pmatrix} u(0) \\ u(-1) \end{pmatrix},$$

as is readily verified.

4. Absence of Absolutely Continuous Spectrum

Let Ω be a strictly ergodic subshift and $f : \Omega \rightarrow \mathbb{R}$ locally constant. It follows that the resulting potentials V_ω take on only finitely many values. In this section we study the absolutely continuous spectrum of H_ω , equal to Σ_{ac} for every $\omega \in \Omega$ by Theorem 3.3. In 1982, Kotani made one of the deepest and most celebrated contributions to the theory of ergodic Schrödinger operators by showing that Σ_{ac} is completely determined by the Lyapunov exponent, or rather the set \mathcal{Z} . Namely, his results, together with earlier ones by Ishii and Pastur, show that Σ_{ac} is given by the essential closure of \mathcal{Z} . In 1989, Kotani found surprisingly general consequences

of his theory in the case of potentials taking on finitely many values. We will review these results below.

By assumption, the potentials V_ω take values in a fixed finite subset \mathcal{B} of \mathbb{R} . Thus, they can be regarded as elements of $\mathcal{B}^{\mathbb{Z}}$, equipped with product topology. Let ν be the measure on $\mathcal{B}^{\mathbb{Z}}$ which is the push-forward of μ under the mapping

$$\Omega \rightarrow \mathcal{B}^{\mathbb{Z}}, \quad \omega \rightarrow V_\omega.$$

Recall that the support of ν , denoted by $\text{supp } \nu$, is the complement of the largest open set U with $\nu(U) = 0$. Let

$$\begin{aligned} (\text{supp } \nu)_+ &= \{V|_{\mathbb{Z}_0^+} : V \in \text{supp } \nu\} \\ (\text{supp } \nu)_- &= \{V|_{\mathbb{Z}^-} : V \in \text{supp } \nu\}, \end{aligned}$$

where $\mathbb{Z}_0^+ = \{0, 1, 2, \dots\}$ and $\mathbb{Z}^- = \{\dots, -3, -2, -1\}$.

DEFINITION 4.1. *The measure ν is called deterministic if every $V_+ \in (\text{supp } \nu)_+$ comes from a unique $V \in \text{supp } \nu$ and every $V_- \in (\text{supp } \nu)_-$ comes from a unique $V \in \text{supp } \nu$.*

Consequently, if ν is deterministic, there is a bijection $C : (\text{supp } \nu)_- \rightarrow (\text{supp } \nu)_+$ such that for every $V \in \text{supp } \nu$, $V|_{\mathbb{Z}_0^+} = C(V|_{\mathbb{Z}^-})$ and $V|_{\mathbb{Z}^-} = C^{-1}(V|_{\mathbb{Z}_0^+})$.

DEFINITION 4.2. *The measure ν is called topologically deterministic if it is deterministic and the map C is a homeomorphism.*

Thus, when ν is topologically deterministic, we can continuously recover one half line from the other for elements of $\text{supp } \nu$.

Let us only state the part of Kotani theory that is of immediate interest to us here:

- THEOREM 4.3.** (a) *If \mathcal{Z} has zero Lebesgue measure, then Σ_{ac} is empty.*
 (b) *If \mathcal{Z} has positive Lebesgue measure, then ν is topologically deterministic.*

This theorem holds in greater generality; see [100, 101, 102, 125]. The underlying dynamical system (Ω, T, μ) is only required to be measurable and ergodic and the set \mathcal{B} can be any compact subset of \mathbb{R} . Part (a) is a particular consequence of the Ishii–Kotani–Pastur identity

$$\Sigma_{\text{ac}} = \overline{\mathcal{Z}}^{\text{ess}},$$

where the essential closure of a set $S \subseteq \mathbb{R}$ is given by

$$\overline{S}^{\text{ess}} = \{E \in \mathbb{R} : \text{Leb}((E - \varepsilon, E + \varepsilon) \cap S) > 0 \text{ for every } \varepsilon > 0\}.$$

The following result was proven by Kotani in 1989 [101]. Here it is crucial that the set \mathcal{B} is finite.

THEOREM 4.4. *If ν is topologically deterministic, then $\text{supp } \nu$ is finite. Consequently, all potentials in $\text{supp } \nu$ are periodic.*

Combining Theorems 4.3 and 4.4 we arrive at the following corollary.

COROLLARY 4.5. *If Ω and f are such that $\text{supp } \nu$ contains an aperiodic element, then \mathcal{Z} has zero Lebesgue measure and Σ_{ac} is empty.*

Note that by minimality, the existence of one aperiodic element is equivalent to all elements being aperiodic. This completely settles the issue of existence/purity of absolutely continuous spectrum. In the periodic case, the spectrum of H_ω is purely absolutely continuous for every $\omega \in \Omega$, and in the aperiodic case, the spectrum of H_ω is purely singular for every $\omega \in \Omega$.

5. Zero-Measure Spectrum

Suppose throughout this section that Ω is strictly ergodic, $f : \Omega \rightarrow \mathbb{R}$ is locally constant, and the resulting potentials V_ω are aperiodic.² This section deals with the Lebesgue measure of the set Σ , which is the common spectrum of the operators H_ω , $\omega \in \Omega$. It is widely expected that Σ always has zero Lebesgue measure. This is supported by positive results for large classes of subshifts and functions. We present two approaches to zero-measure spectrum, one based on trace map dynamics and sub-exponential upper bounds for $\|A_n^{f,E}(\omega)\|$ for energies in the spectrum, and another one based on uniform convergence of $\frac{1}{n} \log \|A_n^{f,E}(\omega)\|$ to $\gamma(E)$ for all energies. Both approaches have in common that they establish the identity

$$\Sigma = \mathcal{Z}. \tag{14}$$

Zero-measure spectrum then follows immediately from Corollary 4.5.

5.1. Trace Map Dynamics. Zero-measure spectrum follows once one proves

$$\Sigma \subseteq \mathcal{Z}. \tag{15}$$

Note, however, that the Lyapunov exponent is always positive away from the spectrum. Thus, $\mathcal{Z} \subseteq \Sigma$, and (15) is in fact equivalent to (14).

A trace map is a dynamical system that may be associated with a family $\{H_\omega\}$ under suitable circumstances. It is given by the iteration of a map $T : \mathbb{R}^k \rightarrow \mathbb{R}^k$. Iteration of this map on some energy-dependent initial vector, v_E , will then describe the evolution of a certain sequence of transfer matrix traces. Typically, these iterates will diverge rather quickly. The stable set, B_∞ , is defined to be the set of energies for which $T^n v_E$ does not diverge quickly. The inclusion (15) is then established in a two-step procedure:

$$\Sigma \subseteq B_\infty \subseteq \mathcal{Z}. \tag{16}$$

Again, by the remark above, this establishes equality and hence

$$\Sigma = B_\infty = \mathcal{Z}.$$

For the sake of clarity of the main ideas, we first discuss the trace-map approach for the Fibonacci subshift Ω_F and $f : \Omega_F \rightarrow \mathbb{R}$ given by $f(\omega) = g(\omega_0)$, $g(0) = 0$, $g(1) = \lambda > 0$. See [25, 98, 118, 130, 131] for the original literature concerning this special case.

Given the partition result, Theorem 2.15, and the recursion (6), it is natural to decompose transfer matrix products into factors of the form $M_k(E)$, where

$$M_{-1}(E) = \begin{pmatrix} 1 & -\lambda \\ 0 & 1 \end{pmatrix}, \quad M_0(E) = \begin{pmatrix} E & -1 \\ 1 & 0 \end{pmatrix} \tag{17}$$

and

$$M_{k+1}(E) = M_{k-1}(E)M_k(E), \quad \text{for } k \geq 0. \tag{18}$$

²Even when we make explicit assumptions on Ω and f , aperiodicity of the potentials will always be assumed implicitly; for example, in Theorem 5.5 and Corollary 5.8.

PROPOSITION 5.1. *Let $x_k = x_k(E) = \frac{1}{2} \text{Tr} M_k(E)$. Then,*

$$x_{k+2} = 2x_{k+1}x_k - x_{k-1} \quad \text{for } k \in \mathbb{Z}_0^+ \tag{19}$$

and

$$x_{k+1}^2 + x_k^2 + x_{k-1}^2 - 2x_{k+1}x_kx_{k-1} = 1 + \frac{\lambda^2}{4} \quad \text{for } k \in \mathbb{Z}_0^+. \tag{20}$$

PROOF. The recursion (19) follows readily from (18). Using (19), one checks that the left-hand side of (20) is independent of k . Evaluation for $k = 0$ then yields the right-hand side. See [98, 118, 130] for more details. \square

The recursion (19) is called the *Fibonacci trace map*. The x_k 's may be obtained by the iteration of the map $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, $(x, y, z) \mapsto (xy - z, x, y)$ on the initial vector $((E - \lambda)/2, E/2, 1)$.

PROPOSITION 5.2. *The sequence $(x_k)_{k \geq -1}$ is unbounded if and only if*

$$|x_{k_0-1}| \leq 1, \quad |x_{k_0}| > 1, \quad |x_{k_0+1}| > 1 \tag{21}$$

for some $k_0 \geq 0$. In this case, the k_0 is unique, and we have

$$|x_{k+2}| > |x_{k+1}x_k| > 1 \quad \text{for } k \geq k_0 \tag{22}$$

and

$$|x_k| > C^{F_k - k_0} \quad \text{for } k \geq k_0 \tag{23}$$

and some $C > 1$. If $(x_k)_{k \geq -1}$ is bounded, then

$$|x_k| \leq 1 + \frac{\lambda}{2} \quad \text{for every } k. \tag{24}$$

PROOF. Suppose first that (21) holds for some $k_0 \geq 0$. Then, by (19),

$$|x_{k_0+2}| \geq |x_{k_0+1}x_{k_0}| + (|x_{k_0+1}x_{k_0}| - |x_{k_0-1}|) > |x_{k_0+1}x_{k_0}| > 1.$$

By induction, we get (22), and also that the k_0 is unique. Taking log's, we see that $\log |x_k|$ grows faster than a Fibonacci sequence for $k \geq k_0$, which gives (23).

Conversely, suppose that (21) fails for every $k_0 \geq 0$. Consider a value of k for which $|x_k| > 1$. Since $x_{-1} = 1$, it follows that $|x_{k-1}| \leq 1$ and $|x_{k+1}| \leq 1$. Thus, the invariant (20) shows that

$$\begin{aligned} |x_k| &\leq |x_{k+1}x_{k-1}| + \left(|x_{k+1}x_{k-1}|^2 - x_{k+1}^2 - x_{k-1}^2 + 1 + \frac{\lambda^2}{4} \right)^{1/2} \\ &= |x_{k+1}x_{k-1}| + \left((1 - x_{k+1}^2)(1 - x_{k-1}^2) + \frac{\lambda^2}{4} \right)^{1/2}, \end{aligned}$$

which implies that the sequence $(x_k)_{k \geq -1}$ is bounded and obeys (24). \square

The dichotomy described in Proposition 5.2 motivates the following definition:

$$B_\infty = \left\{ E \in \mathbb{R} : |x_k| \leq 1 + \frac{\lambda}{2} \text{ for every } k \right\}. \tag{25}$$

This set provides the link between the spectrum and the set of energies for which the Lyapunov exponent vanishes.

THEOREM 5.3. *Let $\Omega = \Omega_F$ be the Fibonacci subshift and let $f : \Omega \rightarrow \mathbb{R}$ be given by $f(\omega) = g(\omega_0)$, $g(0) = 0$, $g(1) = \lambda > 0$. Then, $\Sigma = B_\infty = \mathcal{Z}$ and Σ has zero Lebesgue measure.*

PROOF. We show the two inclusions in (16). Let $\sigma_k = \{E : |x_k| \leq 1\}$. On the one hand, σ_k is the spectrum of an F_k -periodic Schrödinger operator H_k . It is not hard to see that $H_k \rightarrow H$ strongly, where H is the Schrödinger operator with potential $V(n) = \lambda\chi_{[1-\theta,1)}(n\theta)$ and hence Σ is contained in the closure of $\bigcup_{\bar{k} \geq k} \sigma_{\bar{k}}$ for every k . On the other hand, Proposition 5.2 shows that $\sigma_{k+1} \cup \sigma_{k+2} \subseteq \sigma_k \cup \sigma_{k+1}$ and $B_\infty = \bigcap_k \sigma_k \cup \sigma_{k+1}$. Thus,

$$\Sigma \subseteq \bigcap_k \overline{\bigcup_{\bar{k} \geq k} \sigma_{\bar{k}}} = \bigcap_k \sigma_k \cup \sigma_{k+1} = B_\infty.$$

This is the first inclusion in (16). The second inclusion follows once we can show that for every $E \in B_\infty$, we have that $\log \|M_n\| \lesssim n$, where the implicit constant depends only on λ . From the matrix recursion (18) and the Cayley–Hamilton Theorem, we obtain

$$M_{k+1} = M_{k-1}M_k^2M_k^{-1} = M_{k-1}(2x_kM_k - \text{Id})M_k^{-1} = 2x_kM_{k-1} - M_{k-2}^{-1}.$$

If $E \in B_\infty$, then $2|x_k| \leq 2 + \lambda$, and hence we obtain by induction that $\|M_k\| \leq C^k$. Combined with the partition result, Theorem 2.15, this yields the claim since the F_k grow exponentially. \square

The same strategy works in the Sturmian case, as shown by Bellissard et al. [15], although the analysis is technically more involved. Because of (5), we now consider instead of (18) the matrices defined by the recursion

$$M_{k+1}(E) = M_{k-1}(E)M_k(E)^{a_{k+1}},$$

where the a_k 's are the coefficients in the continued fraction expansion (2) of θ . This recursion again gives rise to a trace map for $x_k = \frac{1}{2}\text{Tr}M_k(E)$ which involves Chebyshev polynomials. These traces obey the invariant (20) and the exact analogue of Proposition 5.2 holds. After these properties are established, the proof may be completed as above. Namely, B_∞ is again defined by (25) and the same line of reasoning yields the two inclusions in (16). We can therefore state the following result:

THEOREM 5.4. *Let Ω be a Sturmian subshift with irrational slope $\theta \in (0, 1)$ and let $f : \Omega \rightarrow \mathbb{R}$ be given by $f(\omega) = g(\omega_0)$, $g(0) = 0$, $g(1) = \lambda > 0$. Then, $\Sigma = B_\infty = \mathcal{Z}$ and Σ has zero Lebesgue measure.*

We see that every operator family associated with a Sturmian subshift admits a trace map and an analysis of this dynamical system allows one to prove the zero-measure property.

Another class of operators for which a trace map always exists and may be used to prove zero-measure spectrum is given by those that are generated by a primitive substitution. The existence of a trace map is even more natural in this case and not hard to verify; see, for example, [3, 4, 5, 99, 119] for general results and [8, 10, 124] for trace maps with an invariant. However, its analysis is more involved and has been completed only in 2002 by Liu et al. [111], leading to Theorem 5.5 below, after a number of earlier works had established partial results [13, 14, 22]. Bellissard et al., on the other hand, had proved their Sturmian result already in 1989—shortly after Kotani made his crucial observation leading to Corollary 4.5.

THEOREM 5.5. *Let Ω be a subshift generated by a primitive substitution $S : \mathcal{A} \rightarrow \mathcal{A}^*$ and let $f : \Omega \rightarrow \mathbb{R}$ be given by $f(\omega) = g(\omega_0)$ for some function $g : \mathcal{A} \rightarrow \mathbb{R}$. Then,*

the associated trace map admits a stable set, B_∞ , for which we have $\Sigma = B_\infty = \mathcal{Z}$. Consequently, Σ has zero Lebesgue measure.

5.2. Uniform Hyperbolicity. Recall that the Lyapunov exponent associated with the Schrödinger cocycle $A^{f,E}$ obeys

$$\gamma(E) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \|A^{f,E}(T^{n-1}\omega) \cdots A^{f,E}(\omega)\| \tag{26}$$

for μ -almost every $\omega \in \Omega$.

DEFINITION 5.6 (uniformity). *The cocycle $A^{f,E}$ is called uniform if the convergence in (26) holds for every $\omega \in \Omega$ and is uniform in ω . It is called uniformly hyperbolic if it is uniform and $\gamma(E) > 0$. Define*

$$\mathcal{U} = \{E \in \mathbb{R} : A^{f,E} \text{ is uniformly hyperbolic} \}.$$

Uniform hyperbolicity of $A^{f,E}$ is equivalent to E belonging to the resolvent set as shown by Lenz [108] (see also Johnson [86]):

THEOREM 5.7. $\mathbb{R} \setminus \Sigma = \mathcal{U}$.

Recall that we assumed at the beginning of this section that the potentials V_ω are aperiodic. Thus, combining Corollary 4.5 and Theorem 5.7, we arrive at the following corollary.

COROLLARY 5.8. *If $A^{f,E}$ is uniform for every $E \in \mathbb{R}$, then $\Sigma = \mathcal{Z}$ and Σ has zero Lebesgue measure.*

We thus seek a sufficient condition on Ω and f such that $A^{f,E}$ is uniform for every $E \in \mathbb{R}$, which holds for as many cases of interest as possible. Such a condition was recently found by Damanik and Lenz in [46]. In [48] it was then shown by the same authors that this condition holds for the majority of the models discussed in Section 2.

DEFINITION 5.9 (condition (B)). *Let Ω be a strictly ergodic subshift with unique T -invariant measure μ . It satisfies the Boshernitzan condition (B) if*

$$\limsup_{n \rightarrow \infty} \left(\min_{w \in \mathcal{W}_\Omega(n)} n \cdot \mu([w]) \right) > 0. \tag{27}$$

Remarks. (a) $[w]$ denotes the cylinder set

$$[w] = \{\omega \in \Omega : \omega_1 \dots \omega_{|w|} = w\}.$$

(b) It suffices to assume that Ω is minimal and there exists some T -invariant measure μ with (27). Then, Ω is necessarily uniquely ergodic.

(c) The condition (27) was introduced by Boshernitzan in [20]. His main purpose was to exhibit a useful sufficient condition for unique ergodicity. The criterion proved to be particularly useful in the context of interval exchange transformations [19, 136], where unique ergodicity holds almost always, but not always [91, 114, 135].

It was shown by Damanik and Lenz that condition (B) implies uniformity for all energies and hence zero-measure spectrum [46].

THEOREM 5.10. *If Ω satisfies (B), then $A^{f,E}$ is uniform for every $E \in \mathbb{R}$.*

Remarks. (a) The Boshernitzan condition holds for almost all of the subshifts discussed in Section 2. For example, it holds for every Sturmian subshift, almost every subshift generated by a coding of a rotation with respect to a two-interval partition, a dense set of subshifts associated with general codings of rotations, almost every subshift associated with an interval exchange transformation, almost every episturmian subshift, and every linearly recurrent subshift; see [48].³

(b) There was earlier work by Lenz who proved uniformity for all energies assuming a stronger condition, called (PW) for positive weights [107]. Essentially, (PW) requires (27) with \limsup replaced by \liminf . The condition (PW) holds for all linearly recurrent subshifts but it fails, for example, for almost every Sturmian subshift.

(c) Lenz in turn was preceded and inspired by Hof [77] who proved uniform existence of the Lyapunov exponents for Schrödinger operators associated with primitive substitutions. Extensions of [77] to linearly recurrent systems, including higher-dimensional ones, were found by Damanik and Lenz [40].

In particular, all zero-measure spectrum results obtained by the trace map approach also follow from Theorem 5.10. Moreover, the applications of Theorem 5.10 cover operator families that are unlikely to be amenable to the trace map approach. However, as we will see later, the trace map approach yields additional information that is crucial in a study of detailed spectral and dynamical properties. Thus, it is worthwhile to carry out an analysis of the trace map whenever possible.

5.3. The Hausdorff Dimension of the Spectrum. Once one knows that the spectrum has zero Lebesgue measure it is a natural question if anything can be said about its Hausdorff dimension. There is very important unpublished work by Raymond [121] who proved in the Fibonacci setting of Theorem 5.3 that the Hausdorff dimension is strictly smaller than one for λ large enough ($\lambda \geq 5$ is sufficient) and it converges to zero as $\lambda \rightarrow \infty$. Strictly positive lower bounds for the Hausdorff dimension of the spectrum at all couplings λ follow from the Hausdorff continuity results of [29, 83], to be discussed in Section 7. Several aspects of Raymond's work were used and extended in a number of papers [35, 51, 53, 95, 112]. In particular, Liu and Wen carried out a detailed analysis of the Hausdorff dimension of the spectrum in the general Sturmian case in the spirit of Raymond's approach; see [112].

6. Absence of Point Spectrum

We have seen that two of the three properties that are expected to hold in great generality for the operators discussed in this paper hold either always (absence of absolutely continuous spectrum) or almost always (zero-measure spectrum). In this section we turn to the third property that is expected to be the rule—the absence of point spectrum. As with zero-measure spectrum, no counterexamples are known and there are many positive results that have been obtained by essentially two different methods. Both methods rely on local symmetries of the potential. The existence of square-summable eigenfunctions is excluded by showing that these local symmetries are reflected in the solutions of the difference equation (13). Since

³Here, notions like “dense” or “almost all” are with respect to the natural parameters of the class of models in question. We refer the reader to [48] for detailed statements of these applications of Theorem 5.10.

there are exactly two types of symmetries in one dimension, the effective criteria for absence of eigenvalues that implement this general idea therefore rely on translation and reflection symmetries, respectively. In the following, we explain these two methods and their range of applicability.

6.1. Local Repetitions. In 1976, Gordon showed how to use the Cayley–Hamilton theorem to derive quantitative solution estimates from local repetitions in the potential [71]. The first major application of this observation was in the context of the almost Mathieu operator: For super-critical coupling and Liouville frequencies, there is purely singular continuous spectrum for all phases, as shown by Avron and Simon in 1982 [6].⁴ The first application of direct relevance to this survey was obtained by Delyon and Petritis [58] in 1986 who proved absence of eigenvalues for certain codings of rotations, including most Sturmian models. Further applications will be mentioned below.

Gordon’s Lemma is a deterministic criterion and may be applied to a fixed potential $V : \mathbb{Z} \rightarrow \mathbb{R}$. Analogous to the discussion in Section 3, we define transfer matrices $A_n^E = T_{n-1} \cdots T_0$, where

$$T_j = \begin{pmatrix} E - V(j) & -1 \\ 1 & 0 \end{pmatrix}.$$

Then, a sequence u solves

$$u(n+1) + u(n-1) + V(n)u(n) = Eu(n) \tag{28}$$

if and only if it solves $U(n) = A_n^E U(0)$, where

$$U(j) = \begin{pmatrix} u(j) \\ u(j-1) \end{pmatrix}.$$

LEMMA 6.1. *Suppose the potential V obeys $V(m+p) = V(m)$, $0 \leq m \leq p-1$. Then,*

$$\max \{ \|U(2p)\|, \|U(p)\| \} \geq \frac{1}{2 \max\{|\operatorname{Tr} A_p^E|, 1\}} \|U(0)\|.$$

PROOF. This is immediate from the Cayley–Hamilton Theorem, applied to the matrix A_p^E and the vector $(u(0), u(-1))^T$. \square

For obvious reasons, we call this criterion the *two-block* Gordon Lemma. A slight variation of the argument gives the following (*three-block*) version of Gordon’s Lemma.

LEMMA 6.2. *Suppose the potential V obeys $V(m+p) = V(m)$, $-p \leq m \leq p-1$. Then,*

$$\max \{ \|U(2p)\|, \|U(p)\|, \|U(-p)\| \} \geq \frac{1}{2}.$$

Remark. The original criterion from [71] uses four blocks. For the application to the almost Mathieu operator, this is sufficient; but for Sturmian models, for example, the improvements above are indeed needed, as we will see below. The two-block version can be found in Sütő’s paper [130] and the three-block version was proved in [58] by Delyon and Petritis.

The two-block version is especially useful in situations where a trace map exists and we have bounds on trace map orbits for energies in the spectrum. Note that

⁴As a consequence, positive Lyapunov exponents do not in general imply spectral localization.

the two-block version gives a stronger conclusion. This will be crucial in the next section when we discuss continuity properties of spectral measures with respect to Hausdorff measures in the context of quantum dynamics.

6.2. Palindromes. Gordon-type criteria give quantitative estimates for solutions of (28) in the sense that repetitions in the potential are reflected in solutions, albeit in a weak sense. One would hope that local reflection symmetries in the potential give similar information. Unfortunately, such a result has not been found yet. However, it is possible to exclude square-summable solutions in this way by an indirect argument. Put slightly simplified, *if* a solution is square-summable, *then* local reflection symmetries are mirrored by solutions and these solution symmetries in turn prevent the solution from being square-summable.

The original criterion for absence of eigenvalues in this context is due to Jitomirskaya and Simon [85] and it was developed in the context of the almost Mathieu operator to prove, just as the result by Avron and Simon did, an unexpected occurrence of singular continuous spectrum. An adaptation of the Jitomirskaya–Simon method to the subshift context can be found in a paper by Hof et al. [78]. Let us state their result:⁵

LEMMA 6.3. *Let $V : \mathbb{Z} \rightarrow \mathbb{R}$ be given. There is a constant B , depending only on $\|V\|_\infty$, with the following property: if there are $n_j \rightarrow \infty$ and l_j with $B^{n_j}/l_j \rightarrow 0$ as $j \rightarrow \infty$ such that V is symmetric about n_j on an interval of length l_j centered at n_j for every j , then the Schrödinger operator H with potential V has empty point spectrum.*

SKETCH OF PROOF. Suppose that V satisfies the assumptions of the lemma. Assume that u is a square-summable solution of (28), normalized so that $\|u\|_2 = 1$. Fix some j and reflect u about n_j . Call the reflected sequence $u^{(j)}$. Since the potential is reflection-symmetric on an interval of length l_j about n_j , the Wronskian of u and $u^{(j)}$ is constant on this interval. By $\|u\|_2 = 1$, it is pointwise bounded in this interval by $2/l_j$. From this, it follows that u and $u^{(j)}$ are close (up to a sign) near n_j . Now apply transfer matrices and compare u and $u^{(j)}$ near zero. The assumption $B^{n_j}/l_j \rightarrow 0$ then implies that, for j large, u and $u^{(j)}$ are very close near zero. In other words, u is bounded away from zero near $2n_j$ for all large j . This contradicts $u \in \ell^2(\mathbb{Z})$. □

Thus, eigenvalues can be excluded if the potential contains infinitely many suitably located palindromes. Here, a *palindrome* is a word that is the same when read backwards. Sequences obeying the assumption of Lemma 6.3 are called *strongly palindromic* in [78].

Hof et al. also prove the following general result for subshifts:

PROPOSITION 6.4. *Suppose Ω is an aperiodic minimal subshift. If \mathcal{W}_Ω contains infinitely many palindromes, then the set of strongly palindromic ω 's in Ω is uncountably infinite.*

In any event, since the set $\mathcal{C}_\Omega = \{\omega \in \Omega : \sigma_{\text{pp}}(H_\omega) = \emptyset\}$ is a G_δ set as shown by Simon [126] (see also Choksi and Nadkarni [27] and Lenz and Stollmann [109]), it is a dense G_δ set as soon as it is non-empty by minimality of Ω and unitary equivalence of H_ω and $H_{T\omega}$.

⁵There is also a half-line version, which is the palindrome analogue of Lemma 6.1; see [36].

Thus, when excluding eigenvalues we are interested in three kinds of results. We say that eigenvalues are *generically absent* if \mathcal{C}_Ω is a dense G_δ set. To prove generic absence of eigenvalues it suffices to treat one $\omega \in \Omega$, as explained in the previous paragraph. Absence of eigenvalues holds *almost surely* if $\mu(\mathcal{C}_\Omega) = 1$. To prove almost sure absence of eigenvalues one only has to show $\mu(\mathcal{C}_\Omega) > 0$ by ergodicity and T -invariance of \mathcal{C}_Ω . Finally, absence of eigenvalues is said to hold *uniformly* if $\mathcal{C}_\Omega = \Omega$.

6.3. Applications. Let us now turn to applications of the two methods just described. We emphasize that absence of eigenvalues is expected to hold in great generality and no counterexamples are known.

As in Section 5, things are completely understood in the Sturmian case and absence of eigenvalues holds uniformly.

THEOREM 6.5. *Let Ω be a Sturmian subshift with irrational slope $\theta \in (0, 1)$ and let $f : \mathbb{R} \rightarrow \mathbb{R}$ be given by $f(\omega) = g(\omega_0)$, $g(0) = 0$, $g(1) = \lambda > 0$. Then, H_ω has empty point spectrum for every $\omega \in \Omega$.*

SKETCH OF PROOF. Given any $\lambda > 0$ and $\omega \in \Omega$, absence of point spectrum follows if Lemma 6.1 can be applied to V_ω for infinitely many values of p . Considering only p 's of the form q_k , where the q_k 's are associated with θ via (4), the trace bounds established in Theorem 5.4 show that we can focus our attention on the existence of infinitely many two-block structures aligned at the origin. Using Theorem 2.15, a case-by-case analysis through the various levels of the hierarchy detects these structures and completes the proof. \square

Remarks. (a) For details, see [37, 38]. In fact, the argument above has to be extended slightly for θ 's with $\limsup a_k = 2$. To deal with these exceptional cases, one also has to consider p 's of the form $q_k + q_{k-1}$.

(b) Here is a list of earlier partial results for Sturmian models: Delyon and Petritis proved absence of eigenvalues almost surely for every $\lambda > 0$ and Lebesgue almost every θ [58]. Their proof employs Lemma 6.2. Using Lemma 6.1, Sütő proved absence of eigenvalues for $\lambda > 0$, $\theta = (\sqrt{5} - 1)/2$, and $\phi = 0$ [130], and hence generic absence of eigenvalues in the Fibonacci case. His proof and result were extended to all irrational θ 's by Bellissard et al. [15].⁶ Hof et al. proved generic absence of eigenvalues for every $\lambda > 0$ and every θ using Lemma 6.3 [78]. Kaminaga then showed an almost sure result for every $\lambda > 0$ and every θ [89]. His proof is based on Lemma 6.2 and refines the arguments of Delyon and Petritis.

(c) If most of the continued fraction coefficients are small, eigenvalues cannot be excluded using a four-block Gordon Lemma. This applies in particular in the Fibonacci case where $a_k \equiv 1$. The reason for this is that there simply are no four-block structures in the potential. See [41, 42, 87, 134] for papers dealing with local repetitions in Sturmian sequences.

(d) The palindrome method is very useful to prove generic results. However, it cannot be used to prove almost sure or uniform results for linearly recurrent subshifts (e.g., subshifts generated by primitive substitutions). Namely, for these subshifts, the strongly palindromic elements form a set of zero μ -measure as shown by Damanik and Zare [55].

⁶They do not state the result explicitly in [15], but given their analysis of the trace map and the structure of the potential, it follows as in [130].

Let us now turn to subshifts generated by codings of rotations. The key papers were mentioned above [58, 78, 89].

THEOREM 6.6. *Suppose Ω is the subshift generated by a sequence of the form (8) with irrational $\theta \in (0, 1)$, some $\phi \in [0, 1)$, and a partition on the circle into l half-open intervals. Let $f : \Omega \rightarrow \mathbb{R}$ be given by $f(\omega) = g(\omega_0)$ with some non-constant function g . Suppose that the continued fraction coefficients of θ satisfy*

$$\limsup_{k \rightarrow \infty} a_k \geq 2l. \tag{29}$$

Then, H_ω has empty point spectrum for μ -almost every $\omega \in \Omega$.

Remarks. (a) For every $l \in \mathbb{Z}^+$, the condition (29) holds for Lebesgue almost every θ . In fact, almost every θ has unbounded continued fraction coefficients; see [94].

(b) The proof of Theorem 6.6, given in [58, 89], is based on Lemma 6.2.

(c) Hof et al. prove a generic result using Lemma 6.3 for every θ provided that the partition of the circle has a certain symmetry property, which is always satisfied in the case $l = 2$ [78].

(d) It is possible to prove a result similar to Theorem 6.6 for a locally constant f . In this case, the number $2l$ in (29) has to be replaced by a larger integer, determined by the size of the window $f(\omega)$ depends upon. Still, this gives almost sure absence of eigenvalues for almost every θ .

A large number of papers deal with the eigenvalue problem for Schrödinger operators generated by primitive substitutions; for example, [7, 14, 22, 30, 31, 32, 34, 59, 78].⁷ We first describe general results that can be obtained using the two general methods we discussed and then turn to some specific examples, where more can be said.

We start with an application of Lemma 6.2. Fix some strictly ergodic subshift Ω and define, for $w \in \mathcal{W}_\Omega$, the *index* of w to be

$$\text{ind}(w) = \sup\{r \in \mathbb{Q} : w^r \in \mathcal{W}_\Omega\}.$$

Here, w^r denotes the word $(xy)^m x$, where $m \in \mathbb{Z}^+$, $w = xy$, and $r = m + |x|/|w|$. The *index* of Ω is given by

$$\text{ind}(\Omega) = \sup\{\text{ind}(w) : w \in \mathcal{W}_\Omega\} \in [1, \infty].$$

Then, the following result was shown in [32] using three-block Gordon.

THEOREM 6.7. *Suppose Ω is generated by a primitive substitution and $\text{ind}(\Omega) > 3$. Let $f : \Omega \rightarrow \mathbb{R}$ be given by $f(\omega) = g(\omega_0)$ with some non-constant function $g : \mathcal{A} \rightarrow \mathbb{R}$. Then, H_ω has empty point spectrum for μ -almost every $\omega \in \Omega$.*

Remarks. (a) See [31] for a weaker result, assuming $\text{ind}(\Omega) \geq 4$.

(b) The result extends to the case of a locally constant f .

(c) Consider the case of the period doubling substitution. Since $s_{\text{PD}} = 101110101011101110\dots$, we see that $\text{ind}(\Omega) \geq \text{ind}(10) \geq 3.5 > 3$. Thus, Theorem 6.7 implies almost sure absence of eigenvalues, recovering the main result of [30].

⁷There are also papers dealing with Schrödinger operators associated with non-primitive substitutions [45, 61, 62, 110]. The subshifts considered in these papers are, however, linearly recurrent and hence strictly ergodic, so that the theory is quite similar.

A substitution belongs to class P if there is a palindrome p and, for every $a \in \mathcal{A}$, a palindrome q_a such that $S(a) = pq_a$. Here, p is allowed to be the empty word and, if p is not empty, q_a may be the empty word. Clearly, if a subshift is generated by a class P substitution, it contains arbitrarily long palindromes. Thus, by Proposition 6.4, it contains uncountably many strongly palindromic elements. The following result from [78] is therefore an immediate consequence.

THEOREM 6.8. *Suppose Ω is generated by a primitive substitution S that belongs to class P . Let $f : \Omega \rightarrow \mathbb{R}$ be given by $f(\omega) = g(\omega_0)$ with some non-constant function $g : \mathcal{A} \rightarrow \mathbb{R}$. Then, eigenvalues are generically absent.*

Notice that the Fibonacci, period doubling, and Thue–Morse subshifts are generated by class P substitutions. See [78] for more examples. The Rudin–Shapiro subshift, on the other hand, is not generated by a class P substitution. In fact, it does not contain arbitrarily long palindromes [2, 7].

We mentioned earlier that the proof of Theorem 6.8 cannot give a stronger result since the set of strongly palindromic sequences is always of zero measure for substitution subshifts [55]. Moreover, it was shown in [32] that the three-block Gordon argument cannot prove more than an almost everywhere statement in the sense that for every minimal aperiodic subshift Ω , there exists an element $\omega \in \Omega$ such that ω does not have the infinitely many three block structures needed for an application of Lemma 6.2. Thus, proofs of uniform results should use Lemma 6.1 in a crucial way. Theorem 6.5 shows that a uniform result is known in the Fibonacci case, for example, and Lemma 6.1 along with trace map bounds was indeed the key to the proof of this theorem.

Another example for which a uniform result is known is given by the period doubling substitution [34]. The trace map bounds are weaker than in the Fibonacci case, but a combination of two-block and three-block arguments was shown to work. Further applications of this idea can be found in [110].

The other two examples from Section 2, the Thue–Morse and Rudin–Shapiro substitutions, are not as well understood as Fibonacci and period doubling. Almost sure or uniform absence of eigenvalues for these cases are open, though expected. Generic results can be found in [59, 103].

The eigenvalue problem in the context of the other examples mentioned in Section 2 has been studied only in a small number of papers. For Arnoux–Rauzy subshifts, see [54]; and for interval exchange transformations, see [60].

7. Quantum Dynamics

In this section we focus on the time-dependent Schrödinger equation

$$i \frac{\partial}{\partial t} \psi = H\psi, \quad \psi(0) = \psi_0, \quad (30)$$

where H is a Schrödinger operator in $\ell^2(\mathbb{Z})$ with a potential $V : \mathbb{Z} \rightarrow \mathbb{R}$, typically from a strictly ergodic subshift, and $\psi_0 \in \ell^2(\mathbb{Z})$. By the spectral theorem, (30) is solved by $\psi(t) = e^{-itH}\psi_0$. Thus, the question we want to study is the following: Given some potential V and some initial state $\psi_0 \in \ell^2(\mathbb{Z})$, what can we say about $e^{-itH}\psi_0$ for large times t ?

7.1. Spreading of Wavepackets. Since ψ_0 is square-summable, it is in some sense localized near the origin. For simplicity, one often considers the special case

$\psi_0 = \delta_0$ —the delta-function at the origin. With time, $\psi(t)$ will in general spread out in space. Our goal is to measure this spreading of the wavepacket and relate spreading rates to properties of the potential. As a general rule of thumb, spreading rates decrease with increased randomness of the potential. We will make this more explicit below.

A popular way of measuring the spreading of wavepackets is the following. For $p > 0$, define

$$\langle |X|_{\psi_0}^p \rangle(T) = \sum_n |n|^p a(n, T), \quad (31)$$

where

$$a(n, T) = \frac{2}{T} \int_0^\infty e^{-2t/T} |\langle e^{-itH} \psi_0, \delta_n \rangle|^2 dt. \quad (32)$$

Clearly, the faster $\langle |X|_{\psi_0}^p \rangle(T)$ grows, the faster $e^{-itH} \psi_0$ spreads out, at least averaged in time.⁸ One typically wants to obtain power-law bounds on $\langle |X|_{\psi_0}^p \rangle(T)$ and hence it is natural to define the following quantities: For $p > 0$, define the *upper* (resp., *lower*) *transport exponent* $\beta_{\psi_0}^\pm(p)$ by

$$\beta_{\psi_0}^\pm(p) = \limsup_{T \rightarrow \infty} \frac{\log \langle |X|_{\psi_0}^p \rangle(T)}{p \log T}$$

Both functions $p \mapsto \beta_{\psi_0}^\pm(p)$ are nondecreasing and obey $0 \leq \beta_{\psi_0}^-(p) \leq \beta_{\psi_0}^+(p) \leq 1$.

For periodic V , $\beta_{\psi_0}^\pm(p) \equiv 1$ (*ballistic transport*); while for random V , $\beta_{\psi_0}^\pm(p) \equiv 0$ (a weak version of *dynamical localization*—stronger results are known). For V 's that are intermediate between periodic and random, and in particular Sturmian V 's, it is expected that the transport exponents take values between 0 and 1.

7.2. Spectral Measures and Subordinacy Theory. By the spectral theorem, $\langle e^{-itH} \psi_0, \psi_0 \rangle = \int e^{-itE} d\mu_{\psi_0}(E)$, where μ_{ψ_0} is the spectral measure associated with H and ψ_0 . Thus, it is natural to investigate quantum dynamical questions by relating them to properties of the spectral measure corresponding to the initial state. This approach is classical and the Riemann–Lebesgue Lemma and Wiener's Theorem may be interpreted as statements in quantum dynamics. The RAGE theorem establishes basic dynamical results in terms of the standard decomposition of the Hilbert space into pure point, singular continuous, and absolutely continuous subspaces. We refer the reader to Last's well-written article [105] for a review of these early results.

The results just mentioned are very satisfactory for initial states whose spectral measure has an absolutely continuous component. This is, to some extent, also true for pure point measures. However, if the measure is purely singular continuous, it is desirable to obtain results that go beyond Wiener's Theorem and the RAGE theorem.

Last also addressed this issue in [105] and proposed a decomposition of spectral measures with respect to Hausdorff measures. This was motivated by earlier results of Guarneri [72] and Combes [28] who proved dynamical lower bounds for

⁸Taking time averages is natural since the operators of interest in this paper have purely singular continuous spectrum; compare Wiener's Theorem. While Wiener's Theorem would suggest taking a Cesàro time average, the Abelian time average we choose is more convenient for technical purposes. The transport exponents are the same for both ways of time averaging.

initial states with uniformly Hölder continuous spectral measures. By approximation with uniformly Hölder continuous measures, Last proved in [105] that these bounds extend to measures that are absolutely continuous with respect to a suitable Hausdorff measure:

THEOREM 7.1. *If μ_{ψ_0} has a non-trivial component that is absolutely continuous with respect to the α -dimensional Hausdorff measure h^α on \mathbb{R} , then*

$$\beta_{\psi_0}^-(p) \geq \alpha \text{ for every } p > 0. \tag{33}$$

Remarks. (a) Here, h^α is defined by

$$h^\alpha(S) = \lim_{\delta \rightarrow 0} \inf_{\delta\text{-covers}} \sum |I_m|^\alpha,$$

where $S \subseteq \mathbb{R}$ is a Borel set and a δ -cover is a countable collection of intervals I_m of length bounded by δ such that the union of these intervals contains the set in question. Note that h^1 coincides with Lebesgue measure and h^0 is the counting measure.

(b) For further developments of quantum dynamical lower bounds in terms of continuity or dimensionality properties of spectral measures, see [11, 12, 73, 74].

(c) The result and its proof have natural analogues in higher dimensions; see [105].

While a bound like (33) is nice, it needs to be complemented by effective methods for verifying the input to Theorem 7.1. In the context of one-dimensional Schrödinger operators, it is always extremely useful to connect a problem at hand to properties of solutions to the difference equation (28). The classical decomposition of spectral measures can be studied via subordinacy theory as shown by Gilbert and Pearson [69]; see also [68, 93]. Subordinacy theory has proved to be one of the major tools in one-dimensional spectral theory and many important results have been obtained with its help. Jitomirskaya and Last were able to refine subordinacy theory to the extent that Hausdorff-dimensional spectral issues can be investigated in terms of solution behavior [81, 82, 83]. The key result is the Jitomirskaya–Last inequality, which explicitly relates the Borel transform of the spectral measure to solutions in the half-line setting [82, Theorem 1.1].

Using the maximum modulus principle together with the Jitomirskaya–Last inequality, Damanik et al. then proved the following result for operators on the line [37]:

THEOREM 7.2. *Suppose $\Sigma \subseteq \mathbb{R}$ is a bounded set and there are constants γ_1, γ_2 such that for each $E \in \Sigma$, every solution u of (28) with $|u(-1)|^2 + |u(0)|^2 = 1$ obeys the estimate*

$$C_1(E)L^{\gamma_1} \leq \left(\sum_{n=1}^L |u(n)|^2 \right)^{1/2} \leq C_2(E)L^{\gamma_2} \tag{34}$$

for $L > 0$ sufficiently large and suitable constants $C_1(E), C_2(E)$. Let $\alpha = 2\gamma_1/(\gamma_1 + \gamma_2)$. Then, for any $\psi_0 \in \ell^2(\mathbb{Z})$, the spectral measure for the pair (H, ψ_0) is absolutely continuous with respect to h^α on Σ . In particular, the bound (33) holds for every non-trivial initial state whose spectral measure is supported in Σ .

This shows that suitable bounds for solutions of (28) imply statements on Hausdorff-dimensional spectral properties, which in turn yield quantum dynamical lower bounds. There is an extension to multi-dimensional Schrödinger operators by Kiselev and Last [96].

Two remarks are in order. First, while there are some important applications of the method just presented, proving the required solution estimates is often quite involved. The number of known applications is therefore still relatively small. Second, dynamical bounds in terms of Hausdorff-dimensional properties are strictly one-sided. It is not possible to prove dynamical upper bounds purely in terms of dimensional properties. There are a number of examples that demonstrate this phenomenon. For example, modifications of the super-critical almost Mathieu operator lead to spectrally localized operators with almost ballistic transport [57, 67]. Another important example that is spectrally, but not dynamically, localized is given by the random dimer model [56, 84].

7.3. Plancherel Theorem. There is another approach to dynamical bounds that is also based on solution (or rather, transfer matrix) estimates, which relates dynamics to integrals over Lebesgue measure, as opposed to integrals over the spectral measure. Compared with the approach discussed above, it has two main advantages: One can prove both upper and lower bounds in this way, and the proof of a lower bound is so soft that it applies to a greater number of models.

The key to this approach is a formula due to Kato, which follows quickly from the Plancherel Theorem:

LEMMA 7.3.

$$2\pi \int_0^\infty e^{-2t/T} |\langle e^{-itH} \psi_0, \delta_n \rangle|^2 dt = \int_{-\infty}^\infty | \langle (H - E - \frac{i}{T})^{-1} \psi_0, \delta_n \rangle|^2 dE. \tag{35}$$

PROOF. Consider the function

$$F(t) = \begin{cases} e^{-t/T} \langle e^{-itH} \psi_0, \delta_n \rangle & t \geq 0, \\ 0 & t < 0. \end{cases}$$

Using the spectral theorem, it is readily verified that the Fourier transform of F obeys $\widehat{F}(-E) = i \langle (H - E - \frac{i}{T})^{-1} \psi_0, \delta_n \rangle$. Thus, (35) follows if we apply the Plancherel theorem to F . \square

For simplicity, let us consider the case $\psi_0 = \delta_0$. Note that

$$u(n) = \langle (H - E - i/T)^{-1} \delta_0, \delta_n \rangle$$

solves the difference equation (28) (with E replaced by $E + i/T$) away from the origin and can therefore be studied by means of transfer matrices! In particular, we may infer a bound from below in terms of $\|A_n^{E+i/T}\|^{-1}$. Thus, upper bounds on transfer matrix norms are of interest.

THEOREM 7.4. *Suppose that the transfer matrices obey the bound $\|A_n^E\| \leq C|n|^\alpha$ for every $n \neq 0$, some fixed energy $E \in \mathbb{R}$ and suitable constants C, α . Then,*

$$\beta_{\delta_0}^-(p) \geq \frac{1}{1 + 2\alpha} - \frac{1 + 8\alpha}{p + 2\alpha}$$

for every $p > 0$.

Remarks. (a) This is the one-energy version of a more general result due to Damanik and Tcheremchantsev [51]. See [50] for extensions of [51] and supplementary material and [67] for related work.

(b) An interesting application of Theorem 7.4 (and its proof) to the random dimer model may be found in the paper [84] by Jitomirskaya et al., which confirms a

prediction of Dunlap et al. [64].

- (c) There is also a version of Theorem 7.4 for more general initial states ψ_0 [49].
- (d) The idea of the proof of Theorem 7.4 is simple. A Gronwall-type perturbation argument derives upper bounds on $\|A_n^{\tilde{E}}\|$ for \tilde{E} close to E and n not too large. The right-hand side of (35) may then be estimated from below by integrating only over a small neighborhood of E , where u is controlled by the upper bound on the transfer matrix. The bound for $\beta_{\delta_0}^-(p)$ then follows by rather straightforward arguments.
- (e) The paper [52] (using some ideas from [133]) shows that a combination of the two approaches may sometimes (e.g., in the Fibonacci case) give better bounds.
- (f) Killip et al. used (35) to prove dynamical upper bounds for the slow part of the wavepacket [95]. Their work inspired the use of (35) in [51].

Since (35) is an identity, rather than an inequality, it can be used to bound $a(n, T)$ from both below and above. Clearly, proving an upper bound is more involved and will require assumptions that are global in the energy. It was shown by Damanik and Tcheremchantsev that the following assumption on transfer matrix growth is sufficient to allow one to infer an upper bound for the transport exponents [53]:

THEOREM 7.5. *Let $K \geq 4$ be such that $\sigma(H) \subseteq [-K + 1, K - 1]$. Suppose that, for some $C \in (0, \infty)$ and $\alpha \in (0, 1)$, we have*

$$\int_{-K}^K \left(\max_{1 \leq n \leq CT^\alpha} \|A_n^{E + \frac{i}{T}}\|^2 \right)^{-1} dE = O(T^{-m}) \tag{36}$$

and

$$\int_{-K}^K \left(\max_{1 \leq -n \leq CT^\alpha} \|A_n^{E + \frac{i}{T}}\|^2 \right)^{-1} dE = O(T^{-m}) \tag{37}$$

for every $m \geq 1$. Then, $\beta_{\delta_0}^+(p) \leq \alpha$ for every $p > 0$.

7.4. Applications. Let us discuss the applications of these general methods to Schrödinger operators with potentials from strictly ergodic subshifts.

We begin with the Fibonacci case. In fact, every approach to quantum dynamical bounds has been tested on this example and there are many papers proving dynamical results for it; for example, [29, 35, 37, 51, 52, 53, 83, 95].

Upper bounds for transfer matrices were established by Iochum and Testard [80] who proved, for zero phase, that the norms of the transfer matrices grow no faster than a power law for every energy in the spectrum. The power can be chosen uniformly on the spectrum and depends only on the sampling function f . Notice that this improves on the statement that the Lyapunov exponent vanishes on the spectrum. An extension to Sturmian subshifts whose slope has (essentially) bounded continued fraction coefficients was obtained by Iochum et al. [79]. Note that upper bounds for transfer matrix norms yield the input to Theorem 7.4 and one half of the input to Theorem 7.2. The other half of the input to Theorem 7.2, lower bounds for solutions, was obtained in [29, 83]. The proof of these bounds uses the bound for the trace map for energies from the spectrum, Gordon’s two-block lemma, and a mass-reproduction technique based on cyclic permutations of repeated blocks.⁹

⁹Using partitions (cf. Theorem 2.15), these solution estimates described in this paragraph can be shown for all elements of the subshift [37, 39].

THEOREM 7.6. *Let $V(n) = \lambda \chi_{[1-\theta,1]}(n\theta)$, where $\theta = (\sqrt{5} - 1)/2$ and $\lambda > 0$. Then,*

$$\beta_{\delta_0}^-(p) \geq \begin{cases} \frac{p+2\kappa}{(p+1)(\alpha+\kappa+1/2)} & p \leq 2\alpha + 1, \\ \frac{1}{\alpha+1} & p > 2\alpha + 1. \end{cases} \tag{38}$$

where κ is an absolute constant ($\kappa \approx 0.0126$) and $\alpha \asymp \log \lambda$.

Remarks. (a) In the form stated, the result is from [52]. The bound (38) is the best known dynamical lower bound for the Fibonacci operator and is a culmination of the sequence of works [29, 37, 51, 83, 95] leading up to [52].

(b) When we write $\alpha \asymp \log \lambda$, we mean that α is a positive λ -dependent quantity that satisfies $C_1 \log \lambda \leq \alpha \leq C_2 \log \lambda$ for positive constants C_1, C_2 and all large λ . See [52] for the explicit dependence of α on λ .

To apply Theorem 7.5 to the Fibonacci operator, one has to prove the estimates (36) and (37). This was done in [53]. Let us describe the main idea. Clearly, to prove the desired lower bounds for transfer matrix norms, it suffices to prove lower bounds for transfer matrix traces. We know a way to establish such lower bounds: Lemma 5.2. Since all relevant energies in (36) and (37) are non-real, we know that the trace map will eventually enter the escape region described in Lemma 5.2. The point is to control the number of iterates it takes for this to occur. To this end, define the complex analogue of the set σ_κ from Section 5 by

$$\sigma_k^{\mathbb{C}} = \{z \in \mathbb{C} : |x_k(z)| \leq 1\}.$$

Notice that the x_k 's are polynomials and hence defined for all complex z . As before, being in the complement of two consecutive $\sigma_k^{\mathbb{C}}$'s is a sufficient condition for escape at an explicit rate; compare Lemma 5.2, whose proof extends to complex energies. It is therefore useful to bound the imaginary width of these sets from above. This will give an upper bound on the number of iterates it takes at a given energy to enter the escape region. For λ sufficiently large, the connected components of $\sigma_k^{\mathbb{C}}$ can be studied with the help of Koebe's Distortion Theorem; see [53] for details. The resulting dynamical upper bound has the same asymptotics for large λ as the lower bound above:

THEOREM 7.7. *Let $V(n) = \lambda \chi_{[1-\theta,1]}(n\theta)$, where $\theta = (\sqrt{5} - 1)/2$ and $\lambda \geq 8$. Then,*

$$\beta_{\delta_0}^+(p) \leq \tilde{\alpha} \quad \text{for every } p > 0,$$

where $\tilde{\alpha} \in (0, 1)$ and $\tilde{\alpha} \asymp (\log \lambda)^{-1}$.

In particular, for the Fibonacci operator with $\lambda \geq 8$, all transport exponents $\{\beta_{\delta_0}^\pm(p)\}_{p>0}$ are strictly between zero and one.

The dynamical lower bounds have been established for more general models; see [29, 37, 43, 49, 50, 51]. On the other hand, Theorem 7.7 is the only explicit result of this kind, but as mentioned in [53], the ideas of [35] should permit one to extend this theorem to more general slopes and all elements of the subshift.

8. CMV Matrices Associated with Subshifts

Given a strictly ergodic subshift Ω and a continuous/locally constant function $f : \Omega \rightarrow \mathbb{D}$, we can define $\alpha_n(\omega) = f(T^n\omega)$ for $n \in \mathbb{Z}$ and $\omega \in \Omega$. Let C_ω be the CMV matrix associated with Verblunsky coefficients $\{\alpha_n(\omega)\}_{n \geq 0}$ and \mathcal{E}_ω the

extended CMV matrix associated with Verblunsky coefficients $\{\alpha_n(\omega)\}_{n \in \mathbb{Z}}$. That is, with $\rho_n(\omega) = (1 - |\alpha_n(\omega)|)^{-1/2}$, \mathcal{C}_ω is given by

$$\begin{pmatrix} \bar{\alpha}_0(\omega) & \bar{\alpha}_1(\omega)\rho_0(\omega) & \rho_1(\omega)\rho_0(\omega) & 0 & 0 & \dots \\ \rho_0(\omega) & -\bar{\alpha}_1(\omega)\alpha_0(\omega) & -\rho_1(\omega)\alpha_0(\omega) & 0 & 0 & \dots \\ 0 & \bar{\alpha}_2(\omega)\rho_1(\omega) & -\bar{\alpha}_2(\omega)\alpha_1(\omega) & \bar{\alpha}_3(\omega)\rho_2(\omega) & \rho_3(\omega)\rho_2(\omega) & \dots \\ 0 & \rho_2(\omega)\rho_1(\omega) & -\rho_2(\omega)\alpha_1(\omega) & -\bar{\alpha}_3(\omega)\alpha_2(\omega) & -\rho_3(\omega)\alpha_2(\omega) & \dots \\ 0 & 0 & 0 & \bar{\alpha}_4(\omega)\rho_3(\omega) & -\bar{\alpha}_4(\omega)\alpha_3(\omega) & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

and \mathcal{E}_ω is the analogous two-sided infinite matrix. See [127, 128] for more information on CMV and extended CMV matrices.

For these unitary operators in ℓ^2 , we can ask questions similar to the ones considered above in the context of Schrödinger operators. That is, is the spectrum of zero Lebesgue measure, are spectral measures purely singular continuous, etc. Since we are dealing with ergodic models, it is more natural to consider the whole-line situation. On the other hand, from the point of view of orthogonal polynomials on the unit circle, the half-line situation is more relevant. The zero-measure property is independent of the setting, whereas the spectral type for half-line models is almost always (i.e., when the “boundary condition” is varied) pure point as soon as zero-measure spectrum is established. The latter statement follows quickly from spectral averaging; compare [128, Theorem 10.2.2]. Thus, the key problem for CMV matrices associated with subshifts is proving zero-measure spectrum. In fact, Simon conjectured the following; see [128, Conjecture 12.8.2].

Simon’s Subshift Conjecture. Suppose \mathcal{A} is a subset of \mathbb{D} , the subshift Ω is minimal and aperiodic and let $f : \Omega \rightarrow \mathbb{D}$, $f(\omega) = \omega(0)$. Then, Σ has zero Lebesgue measure.

Here, Σ is the common spectrum of the operators \mathcal{E}_ω , $\omega \in \Omega$. Equivalently, it is the common essential spectrum of \mathcal{C}_ω , $\omega \in \Omega$.

Simon proved the zero-measure property for the Fibonacci case by means of the trace map approach [128, Section 12.8]. Since the approach based on the Boshernitzan condition has a wider scope in the Schrödinger case, it is natural to try and extend it to the CMV case. This was done by Damanik and Lenz in [47] where the following result was shown.

THEOREM 8.1. *Suppose the subshift Ω is aperiodic and satisfies the Boshernitzan condition. Let $f : \Omega \rightarrow \mathbb{D}$ be locally constant. Then, Σ has zero Lebesgue measure.*

This proves Simon’s Subshift Conjecture for a large number of models since we saw above that many of the prominent aperiodic subshifts satisfy the Boshernitzan condition.

Regarding the spectral type, it should not be hard to extend the material from Section 6 to the CMV case. This will imply purely singular continuous spectrum for \mathcal{E}_ω for many subshifts Ω and many (generic, almost all, all) $\omega \in \Omega$. However, as was noted above, the Aleksandrov measures associated with \mathcal{C}_ω will almost surely be pure point whenever Theorem 8.1 applies.

Quantum dynamics, on the other hand, is less natural in the CMV case than in the Schrödinger case, and has not really been studied.¹⁰ Most of the ideas leading to the results presented in Section 7 should have CMV counterparts. In particular, it should be possible to prove absolute continuity of spectral measures with respect to suitable Hausdorff measures for extended CMV matrices over Fibonacci-like subshifts.

9. Concluding Remarks

The material presented in this survey is motivated by and closely related to the theory of quasicrystals; compare, for example, [9, 115]. More specifically, the surveys [33, 132] deal with the Fibonacci operator and its generalizations and the interested reader may find references to the original physics literature in those papers.

Regarding future research in this field, it would be interesting to see how far one can take the philosophy that potentials taking finitely many values preclude localization phenomena. Since the Bernoulli Anderson model is localized [23], this cannot hold in full generality. On the other hand, Gordon potentials are much more prevalent in the subshift case than in the uniformly almost periodic case. Moreover, for smooth quasi-periodic potentials, it is expected that the Lyapunov exponent is positive at all energies if the coupling is large enough. This is known for trigonometric potentials [76], analytic potentials [21, 70, 129], and Gevrey potentials [97]. See also [17, 26] for recent results in the C^r category. These potentials should be contrasted with those coming from quasi-periodic subshifts satisfying the Boshernitzan condition. The Boshernitzan condition is independent of the coupling constant and yields vanishing Lyapunov exponent throughout the spectrum. Since it is satisfied on a dense set of sampling (step-)functions, upper-semicontinuity arguments allow one to derive surprising phenomena that hold generically in the C^0 category [18].

To shed some light on this, it should be helpful to analyze more examples. That is, take one of the popular base transformations of the torus (e.g., shifts, skew-shifts, or expanding maps) and define an ergodic family of potentials by choosing a sampling function on the torus that takes finitely many values. These models, with the exception of rotations of the circle, are not well understood! There is a serious issue about the competition between the flat pieces of the sampling function and the randomness properties of the base transformation (expressed, e.g., in terms of mixing properties). For example, take a 1-periodic step function f and consider $V_\omega(n) = \lambda f(2^n \omega)$, $\lambda > 0$, $\omega \in [0, 1)$. Is it true that the Lyapunov exponent is positive? For all λ 's or all large λ 's? For all energies or all but finitely many?

References

- [1] P. Alessandri and V. Berthé, Three distance theorems and combinatorics on words, *Enseign. Math.* **44** (1998), 103–132
- [2] J.-P. Allouche, Schrödinger operators with Rudin–Shapiro potentials are not palindromic, *J. Math. Phys.* **38** (1997), 1843–1848
- [3] J.-P. Allouche and J. Peyrière, Sur une formule de récurrence sur les traces de produits de matrices associés à certaines substitutions, *C. R. Acad. Sci. Paris* **302** (1986), 1135–1136

¹⁰Simon did extend the Jitomirskaya–Last theory to OPUC in [128]. This theory has its roots in quantum dynamics; compare [57, 72, 82, 83, 105].

- [4] Y. Avishai and D. Berend, Trace maps for arbitrary substitution sequences, *J. Phys. A* **26** (1993), 2437–2443
- [5] Y. Avishai, D. Berend, and D. Glaubman, Minimum-dimension trace maps for substitution sequences, *Phys. Rev. Lett.* **72** (1994), 1842–1845
- [6] J. Avron and B. Simon, Singular continuous spectrum for a class of almost periodic Jacobi matrices, *Bull. Amer. Math. Soc.* **6** (1982), 81–85
- [7] M. Baake, A note on palindromicity, *Lett. Math. Phys.* **49** (1999), 217–227
- [8] M. Baake, U. Grimm, and D. Joseph, Trace maps, invariants, and some of their applications, *Int. J. Mod. Phys. B* **7** (1993), 1527–1550
- [9] M. Baake and R. Moody (Editors), *Directions in Mathematical Quasicrystals*, American Mathematical Society, Providence (2000)
- [10] M. Baake and J. Roberts, Reversing symmetry group of $GL(2, \mathbb{Z})$ and $PGL(2, \mathbb{Z})$ matrices with connections to cat maps and trace maps, *J. Phys. A* **30** (1997), 1549–1573
- [11] J.-M. Barbaroux, F. Germinet, and S. Tcheremchantsev, Fractal dimensions and the phenomenon of intermittency in quantum dynamics, *Duke Math. J.* **110** (2001), 161–193
- [12] J.-M. Barbaroux and S. Tcheremchantsev, Universal lower bounds for quantum diffusion, *J. Funct. Anal.* **168** (1999), 327–354
- [13] J. Bellissard, Spectral properties of Schrödinger’s operator with a Thue–Morse potential, in *Number Theory and Physics (Les Houches, 1989)*, Springer, Berlin (1990), 140–150
- [14] J. Bellissard, A. Bovier, and J.-M. Ghez, Spectral properties of a tight binding Hamiltonian with period doubling potential, *Commun. Math. Phys.* **135** (1991), 379–399
- [15] J. Bellissard, B. Iochum, E. Scoppola, and D. Testard, Spectral properties of one-dimensional quasicrystals, *Commun. Math. Phys.* **125** (1989), 527–543
- [16] J. Berstel, Recent results in Sturmian words, in *Developments in Language Theory*, World Scientific, Singapore (1996), 13–24
- [17] K. Bjerklöv, Positive Lyapunov exponent and minimality for a class of 1-d quasi-periodic Schrödinger equations, to appear in *Ergod. Th. & Dynam. Sys.*
- [18] K. Bjerklöv, D. Damanik, and R. Johnson, Lyapunov exponents of continuous Schrödinger cocycles over irrational rotations, to appear in *Ann. Mat. Pura Appl.*
- [19] M. Boshernitzan, A condition for minimal interval exchange maps to be uniquely ergodic, *Duke Math. J.* **52** (1985), 723–752
- [20] M. Boshernitzan, A condition for unique ergodicity of minimal symbolic flows, *Ergod. Th. & Dynam. Sys.* **12** (1992), 425–428
- [21] J. Bourgain and M. Goldstein, On nonperturbative localization with quasi-periodic potential, *Ann. of Math.* **152** (2000), 835–879
- [22] A. Bovier and J.-M. Ghez, Spectral properties of one-dimensional Schrödinger operators with potentials generated by substitutions, *Commun. Math. Phys.* **158** (1993), 45–66; Erratum: *Commun. Math. Phys.* **166** (1994), 431–432
- [23] R. Carmona, A. Klein, and F. Martinelli, Anderson localization for Bernoulli and other singular potentials, *Commun. Math. Phys.* **108** (1987), 41–66
- [24] R. Carmona and J. Lacroix, *Spectral Theory of Random Schrödinger Operators*, Birkhäuser, Boston (1990)
- [25] M. Casdagli, Symbolic dynamics for the renormalization group of a quasiperiodic Schrödinger equation, *Commun. Math. Phys.* **107** (1986), 295–318
- [26] J. Chan, Method of variations of potential of quasi-periodic Schrödinger equation, to appear in *Geom. Funct. Anal.*
- [27] J. Choksi and M. Nadkarni, Genericity of certain classes of unitary and self-adjoint operators, *Canad. Math. Bull.* **41** (1998), 137–139
- [28] J.-M. Combes, Connections between quantum dynamics and spectral properties of time-evolution operators, in *Differential Equations with Applications to Mathematical Physics*, Academic Press, Boston (1993), 59–68
- [29] D. Damanik, α -continuity properties of one-dimensional quasicrystals, *Commun. Math. Phys.* **192** (1998), 169–182
- [30] D. Damanik, Singular continuous spectrum for the period doubling Hamiltonian on a set of full measure, *Commun. Math. Phys.* **196** (1998), 477–483
- [31] D. Damanik, Singular continuous spectrum for a class of substitution Hamiltonians, *Lett. Math. Phys.* **46** (1998), 303–311

- [32] D. Damanik, Singular continuous spectrum for a class of substitution Hamiltonians II., *Lett. Math. Phys.* **54** (2000), 25–31
- [33] D. Damanik, Gordon-type arguments in the spectral theory of one-dimensional quasicrystals, in *Directions in Mathematical Quasicrystals*, American Mathematical Society, Providence (2000), 277–305
- [34] D. Damanik, Uniform singular continuous spectrum for the period doubling Hamiltonian, *Ann. Henri Poincaré* **2** (2001), 101–108
- [35] D. Damanik, Dynamical upper bounds for one-dimensional quasicrystals, *J. Math. Anal. Appl.* **303** (2005), 327–341
- [36] D. Damanik, J.-M. Ghez, and L. Raymond, A palindromic half-line criterion for absence of eigenvalues and applications to substitution Hamiltonians, *Ann. Henri Poincaré* **2** (2001), 927–939
- [37] D. Damanik, R. Killip, and D. Lenz, Uniform spectral properties of one-dimensional quasicrystals. III. α -continuity, *Commun. Math. Phys.* **212** (2000), 191–204
- [38] D. Damanik and D. Lenz, Uniform spectral properties of one-dimensional quasicrystals, I. Absence of eigenvalues, *Commun. Math. Phys.* **207** (1999), 687–696
- [39] D. Damanik and D. Lenz, Uniform spectral properties of one-dimensional quasicrystals, II. The Lyapunov exponent, *Lett. Math. Phys.* **50** (1999), 245–257
- [40] D. Damanik and D. Lenz, Linear repetitivity, I. Uniform subadditive ergodic theorems and applications, *Discrete Comput. Geom.* **26** (2001), 411–428
- [41] D. Damanik and D. Lenz, The index of Sturmian sequences, *European J. Combin.* **23** (2002), 23–29
- [42] D. Damanik and D. Lenz, Powers in Sturmian sequences, *European J. Combin.* **24** (2003), 377–390
- [43] D. Damanik and D. Lenz, Uniform spectral properties of one-dimensional quasicrystals, IV. Quasi-Sturmian potentials, *J. Anal. Math.* **90** (2003), 115–139
- [44] D. Damanik and D. Lenz, Half-line eigenfunction estimates and purely singular continuous spectrum of zero Lebesgue measure, *Forum Math.* **16** (2004), 109–128
- [45] D. Damanik and D. Lenz, Substitution dynamical systems: Characterization of linear repetitivity and applications, to appear in *J. Math. Anal. Appl.*
- [46] D. Damanik and D. Lenz, A condition of Boshernitzan and uniform convergence in the Multiplicative Ergodic Theorem, *Duke Math. J.* **133** (2006), 95–123
- [47] D. Damanik and D. Lenz, Uniform Szegő cocycles over strictly ergodic subshifts, to appear in *J. Approx. Theory*
- [48] D. Damanik and D. Lenz, Zero-measure Cantor spectrum for Schrödinger operators with low-complexity potentials, *J. Math. Pures Appl.* **85** (2006), 671–686
- [49] D. Damanik, D. Lenz, and G. Stolz, Lower transport bounds for one-dimensional continuum Schrödinger operators, to appear in *Math. Ann.*
- [50] D. Damanik, A. Sütő, and S. Tcheremchantsev, Power-law bounds on transfer matrices and quantum dynamics in one dimension II, *J. Funct. Anal.* **216** (2004), 362–387
- [51] D. Damanik and S. Tcheremchantsev, Power-law bounds on transfer matrices and quantum dynamics in one dimension, *Commun. Math. Phys.* **236** (2003), 513–534
- [52] D. Damanik and S. Tcheremchantsev, Scaling estimates for solutions and dynamical lower bounds on wavepacket spreading, *J. Anal. Math.* **97** (2005), 103–131
- [53] D. Damanik and S. Tcheremchantsev, Upper bounds in quantum dynamics, Preprint (2005), arXiv/math-ph/0502044
- [54] D. Damanik and L. Q. Zamboni, Combinatorial properties of Arnoux–Rauzy subshifts and applications to Schrödinger operators, *Rev. Math. Phys.* **15** (2003), 745–763
- [55] D. Damanik and D. Zare, Palindrome complexity bounds for primitive substitution sequences, *Discrete Math.* **222** (2000), 259–267
- [56] S. De Bièvre and F. Germinet, Dynamical localization for the random dimer Schrödinger operator, *J. Stat. Phys.* **98** (2000), 1135–1148
- [57] R. del Rio, S. Jitomirskaya, Y. Last, and B. Simon, Operators with singular continuous spectrum. IV. Hausdorff dimensions, rank one perturbations, and localization, *J. Anal. Math.* **69** (1996), 153–200
- [58] F. Delyon and D. Petritis, Absence of localization in a class of Schrödinger operators with quasiperiodic potential, *Commun. Math. Phys.* **103** (1986), 441–444

- [59] F. Delyon and J. Peyrière, Recurrence of the eigenstates of a Schrödinger operator with automatic potential, *J. Stat. Phys.* **64** (1991), 363–368
- [60] C. de Oliveira and C. Gutierrez, Almost periodic Schrödinger operators along interval exchange transformations, *J. Math. Anal. Appl.* **283** (2003), 570–581
- [61] C. de Oliveira and M. Lima, A nonprimitive substitution Schrödinger operator with generic singular continuous spectrum, *Rep. Math. Phys.* **45** (2000), 431–436
- [62] C. de Oliveira and M. Lima, Singular continuous spectrum for a class of nonprimitive substitution Schrödinger operators, *Proc. Amer. Math. Soc.* **130** (2002), 145–156
- [63] X. Droubay, J. Justin, and G. Pirillo, Epi-Sturmian words and some constructions of de Luca and Rauzy, *Theoret. Comput. Sci.* **255** (2001), 539–553
- [64] D. Dunlap, H.-L. Wu, and P. Phillips, Absence of localization in a random-dimer model, *Phys. Rev. Lett.* **65** (1990), 88–91
- [65] F. Durand, Linearly recurrent subshifts have a finite number of non-periodic subshift factors, *Ergod. Th. & Dynam. Sys.* **20** (2000), 1061–1078
- [66] F. Durand, B. Host, and C. Skau, Substitutional dynamical systems, Bratteli diagrams and dimension groups, *Ergod. Th. & Dynam. Sys.* **19** (1999), 953–993
- [67] F. Germinet, A. Kiselev, and S. Tcheremchantsev, Transfer matrices and transport for Schrödinger operators, *Ann. Inst. Fourier* **54** (2004), 787–830
- [68] D. Gilbert, On subordinacy and analysis of the spectrum of Schrödinger operators with two singular endpoints, *Proc. Roy. Soc. Edinburgh A* **112** (1989), 213–229
- [69] D. Gilbert and D. Pearson, On subordinacy and analysis of the spectrum of one-dimensional Schrödinger operators, *J. Math. Anal. Appl.* **128** (1987), 30–56
- [70] M. Goldstein and W. Schlag, Hölder continuity of the integrated density of states for quasi-periodic Schrödinger equations and averages of shifts of subharmonic functions, *Ann. of Math.* **154** (2001), 155–203
- [71] A. Gordon, On the point spectrum of the one-dimensional Schrödinger operator, *Usp. Math. Nauk.* **31** (1976), 257–258
- [72] I. Guarneri, Spectral properties of quantum diffusion on discrete lattices, *Europhys. Lett.* **10** (1989), 95–100
- [73] I. Guarneri and H. Schulz-Baldes, Lower bounds on wave packet propagation by packing dimensions of spectral measures, *Math. Phys. Electron. J.* **5** (1999), paper 1
- [74] I. Guarneri and H. Schulz-Baldes, Intermittent lower bound on quantum diffusion, *Lett. Math. Phys.* **49** (1999), 317–324
- [75] F. Hahn and Y. Katznelson, On the entropy of uniquely ergodic transformations, *Trans. Amer. Math. Soc.* **126** (1967), 335–360
- [76] M. Herman, Une méthode pour minorer les exposants de Lyapunov et quelques exemples montrant le caractère local d'un théorème d'Arnold et de Moser sur le tore de dimension 2, *Comment. Math. Helv.* **58** (1983), 4453–502
- [77] A. Hof, Some remarks on discrete aperiodic Schrödinger operators, *J. Stat. Phys.* **72** (1993), 1353–1374
- [78] A. Hof, O. Knill, and B. Simon, Singular continuous spectrum for palindromic Schrödinger operators, *Commun. Math. Phys.* **174** (1995), 149–159
- [79] B. Iochum, L. Raymond, and D. Testard, Resistance of one-dimensional quasicrystals, *Physica A* **187** (1992), 353–368
- [80] B. Iochum and D. Testard, Power law growth for the resistance in the Fibonacci model, *J. Stat. Phys.* **65** (1991), 715–723
- [81] S. Jitomirskaya and Y. Last, Dimensional Hausdorff properties of singular continuous spectra, *Phys. Rev. Lett.* **76** (1996), 1765–1769
- [82] S. Jitomirskaya and Y. Last, Power-law subordinacy and singular spectra. I. Half-line operators, *Acta Math.* **183** (1999), 171–189
- [83] S. Jitomirskaya and Y. Last, Power-law subordinacy and singular spectra. II. Line operators, *Commun. Math. Phys.* **211** (2000), 643–658
- [84] S. Jitomirskaya, H. Schulz-Baldes, and G. Stolz, Delocalization in random polymer models, *Commun. Math. Phys.* **233** (2003), 27–48
- [85] S. Jitomirskaya and B. Simon, Operators with singular continuous spectrum: III. Almost periodic Schrödinger operators, *Commun. Math. Phys.* **165** (1994), 201–205
- [86] R. Johnson, Exponential dichotomy, rotation number, and linear differential operators with bounded coefficients, *J. Differential Equations* **61** (1986), 54–78

- [87] J. Justin and G. Pirillo, Fractional powers in Sturmian words, *Theoret. Comput. Sci.* **255** (2001), 363–376
- [88] J. Justin and G. Pirillo, Episturmian words and episturmian morphisms, *Theoret. Comput. Sci.* **276** (2002), 281–313
- [89] M. Kaminaga, Absence of point spectrum for a class of discrete Schrödinger operators with quasiperiodic potential, *Forum Math.* **8** (1996), 63–69
- [90] M. Keane, Interval exchange transformations, *Math. Z.* **141** (1975), 25–31
- [91] M. Keane, Non-ergodic interval exchange transformations, *Israel J. Math.* **26** (1977), 188–196
- [92] H. B. Keynes and D. Newton, A minimal, non-uniquely ergodic interval exchange transformation, *Math. Z.* **148** (1976), 101–105
- [93] S. Khan and D. Pearson, Subordinacy and spectral theory for infinite matrices, *Helv. Phys. Acta* **65** (1992), 505–527
- [94] A. Khintchine, *Continued Fractions*, Dover, Mineola (1997)
- [95] R. Killip, A. Kiselev, and Y. Last, Dynamical upper bounds on wavepacket spreading, *Amer. J. Math.* **125** (2003), 1165–1198
- [96] A. Kiselev and Y. Last, Solutions, spectrum, and dynamics for Schrödinger operators on infinite domains, *Duke Math. J.* **102** (2000), 125–150
- [97] S. Klein, Anderson localization for the discrete one-dimensional quasi-periodic Schrödinger operator with potential defined by a Gevrey-class function, *J. Funct. Anal.* **218** (2005), 255–292
- [98] M. Kohmoto, L. Kadanoff, and C. Tang, Localization problem in one dimension: Mapping and escape, *Phys. Rev. Lett.* **50** (1983), 1870–1872
- [99] M. Kolár and F. Nori, Trace maps of general substitutional sequences, *Phys. Rev. B* **42** (1990), 1062–1065
- [100] S. Kotani, Ljapunov indices determine absolutely continuous spectra of stationary random one-dimensional Schrödinger operators, in *Stochastic Analysis (Katata/Kyoto, 1982)*, North Holland, Amsterdam (1984), 225–247
- [101] S. Kotani, Jacobi matrices with random potentials taking finitely many values, *Rev. Math. Phys.* **1** (1989), 129–133
- [102] S. Kotani, Generalized Floquet theory for stationary Schrödinger operators in one dimension, *Chaos Solitons Fractals* **8** (1997), 1817–1854
- [103] L. Kroon and R. Riklund, Absence of localization in a model with correlation measure as a random lattice, *Phys. Rev. B* **69** (2004), paper 094204 (5 pages)
- [104] J. Lagarias and P. Pleasants, Repetitive Delone sets and quasicrystals, *Ergod. Th. & Dynam. Sys.* **23** (2003), 831–867
- [105] Y. Last, Quantum dynamics and decompositions of singular continuous spectra, *J. Funct. Anal.* **142** (1996), 406–445
- [106] Y. Last and B. Simon, Eigenfunctions, transfer matrices, and absolutely continuous spectrum of one-dimensional Schrödinger operators, *Invent. Math.* **135** (1999), 329–367
- [107] D. Lenz, Uniform ergodic theorems on subshifts over a finite alphabet, *Ergod. Th. & Dynam. Sys.* **22** (2002), 245–255
- [108] D. Lenz, Singular continuous spectrum of Lebesgue measure zero for one-dimensional quasicrystals, *Commun. Math. Phys.* **227** (2002), 119–130
- [109] D. Lenz and P. Stollmann, Generic sets in spaces of measures and generic singular continuous spectrum for Delone Hamiltonians, *Duke Math. J.* **131** (2006), 203–217
- [110] M. Lima and C. de Oliveira, Uniform Cantor singular continuous spectrum for nonprimitive Schrödinger operators, *J. Statist. Phys.* **112** (2003), 357–374
- [111] Q.-H. Liu, B. Tan, Z.-X. Wen, and J. Wu, Measure zero spectrum of a class of Schrödinger operators, *J. Statist. Phys.* **106** (2002), 681–691
- [112] Q.-H. Liu and Z.-Y. Wen, Hausdorff dimension of spectrum of one-dimensional Schrödinger operator with Sturmian potentials, *Potential Anal.* **20** (2004), 33–59
- [113] M. Lothaire, *Algebraic combinatorics on words*, Cambridge University Press, Cambridge (2002)
- [114] H. Masur, Interval exchange transformations and measured foliations, *Ann. of Math.* **115** (1982), 168–200
- [115] R. Moody (Editor), *The Mathematics of Long-Range Aperiodic Order*, Kluwer, Dordrecht (1997)

- [116] M. Morse and G. Hedlund, Symbolic dynamics, *Amer. J. Math.* **60** (1938), 815–866
- [117] M. Morse and G. Hedlund, Symbolic dynamics, II. Sturmian trajectories, *Amer. J. Math.* **62** (1940), 1–42
- [118] S. Ostlund, R. Pandit, D. Rand, H. Schellnhuber, and E. Siggia, One-dimensional Schrödinger equation with an almost periodic potential, *Phys. Rev. Lett.* **50** (1983), 1873–1877
- [119] J. Peyrière, Z.-Y. Wen and Z.-X. Wen, Polynomes associés aux endomorphismes de groupes libres, *Enseign. Math.* **39** (1993), 153–175
- [120] M. Queffélec, *Substitution Dynamical Systems – Spectral Analysis*, Springer, Berlin (1987)
- [121] L. Raymond, A constructive gap labelling for the discrete Schrödinger operator on a quasisperiodic chain, Preprint (1997)
- [122] R. Risley and L. Q. Zamboni, A generalization of Sturmian sequences: combinatorial structure and transcendence, *Acta Arith.* **95** (2000), 167–184
- [123] J. Roberts, Escaping orbits in trace maps, *Physica A* **228** (1996), 295–325
- [124] J. Roberts and M. Baake, Trace maps as 3D reversible dynamical systems with an invariant, *J. Stat. Phys.* **74** (1994), 829–888
- [125] B. Simon, Kotani theory for one dimensional stochastic Jacobi matrices, *Commun. Math. Phys.* **89** (1983), 227–234
- [126] B. Simon, Operators with singular continuous spectrum. I. General operators, *Ann. of Math.* **141** (1995), 131–145
- [127] B. Simon, *Orthogonal Polynomials on the Unit Circle. Part 1. Classical theory*, American Mathematical Society, Providence (2005)
- [128] B. Simon, *Orthogonal Polynomials on the Unit Circle. Part 2. Spectral theory*, American Mathematical Society, Providence (2005)
- [129] E. Sorets and T. Spencer, Positive Lyapunov exponents for Schrödinger operators with quasi-periodic potentials, *Commun. Math. Phys.* **142** (1991), 543–566
- [130] A. Sütő, The spectrum of a quasisperiodic Schrödinger operator, *Commun. Math. Phys.* **111** (1987), 409–415
- [131] A. Sütő, Singular continuous spectrum on a Cantor set of zero Lebesgue measure for the Fibonacci Hamiltonian, *J. Stat. Phys.* **56** (1989), 525–531
- [132] A. Sütő, Schrödinger difference equation with deterministic ergodic potentials, in *Beyond Quasicrystals (Les Houches, 1994)*, Springer, Berlin (1995), 481–549
- [133] S. Tcheremchantsev, Dynamical analysis of Schrödinger operators with growing sparse potentials, *Commun. Math. Phys.* **253** (2005), 221–252
- [134] D. Vandeth, Sturmian words and words with a critical exponent, *Theoret. Comput. Sci.* **242** (2000), 283–300
- [135] W. A. Veech, Gauss measures for transformations on the space of interval exchange maps, *Ann. of Math.* **115** (1982), 201–242
- [136] W. A. Veech, Boshernitzan’s criterion for unique ergodicity of an interval exchange transformation, *Ergod. Th. & Dynam. Sys.* **7** (1987), 149–153
- [137] N. Wozny and L. Q. Zamboni, Frequencies of factors in Arnoux–Rauzy sequences, *Acta Arith.* **96** (2001), 261–278

MATHEMATICS 253–37, CALIFORNIA INSTITUTE OF TECHNOLOGY, PASADENA, CA 91125, U.S.A.

E-mail address: damanik@caltech.edu

URL: <http://www.math.caltech.edu/people/damanik.html>

Lyapunov Exponents and Spectral Analysis of Ergodic Schrödinger Operators: A Survey of Kotani Theory and Its Applications

David Damanik

Dedicated to Barry Simon on the occasion of his 60th birthday

ABSTRACT. The absolutely continuous spectrum of an ergodic family of one-dimensional Schrödinger operators is completely determined by the Lyapunov exponent as shown by Ishii, Kotani, and Pastur.

Moreover, the part of the theory developed by Kotani gives powerful tools for proving the absence of absolutely continuous spectrum, the presence of absolutely continuous spectrum, and even the presence of purely absolutely continuous spectrum.

We review these results and their recent applications to a number of problems: the absence of absolutely continuous spectrum for rough potentials, the absence of absolutely continuous spectrum for potentials defined by the doubling map on the circle, and the absence of singular spectrum for the subcritical almost Mathieu operator.

CONTENTS

1. Introduction
 2. The Description of the AC Spectrum
 3. The Induced Measure and Its Topological Support
 4. Potentials Generated by the Doubling Map
 5. Absence of AC Spectrum for Rough Potentials
 6. Uniform Lyapunov Exponents and Zero-Measure Spectrum
 7. Purely AC Spectrum for the Subcritical AMO
- References

2000 *Mathematics Subject Classification.* Primary 82B44, 47B80; Secondary 47B36, 81Q10.
Key words and phrases. ergodic Schrödinger operators, Lyapunov exponents.
The author was supported in part by NSF grant DMS-0500910.

1. Introduction

Schrödinger operators with ergodic potentials have enjoyed quite some popularity for several decades now. This is in no small part due to Barry Simon's contributions to the field, through research articles on the one hand, but also through survey articles and his way of putting his personal stamp on results and conjectures on the other hand. Ergodic Schrödinger operators continue to be dear to him as seven of the fifteen Schrödinger operator problems he singles out in [60] for further investigation in the 21st century deal with them. Moreover, the immense activity in the area of ergodic Schrödinger operators is reflected by the fact that the ratio 7/15 improves to 3/4 when it comes to the problems from that list that have been solved so far. One may say that this is due to the uneven distribution of difficulty among these fifteen problems, but this is balanced by the fact that of the remaining four ergodic problems at least three are very hard and that further progress should be expected on some of the remaining non-ergodic problems.

In the area of ergodic Schrödinger operators there are several powerful methods (e.g., fractional moment) and analyses (multi-scale) but few theories (Kotani). What appears to be wordplay wants to express the fact that Kotani theory is distinguished from the other greats by its immensely general scope. It really is a theory that applies to the class of all ergodic operators and it is central in many ways. In addition, Kotani theory has played a crucial role in the solution of two of the four recently solved 21st century problems.

Our goal here is to present the core parts of Kotani theory with more or less complete proofs and to discuss several recent applications of the theory to a number of concrete classes of models for which, whenever possible, we at least outline the main ideas that go into the proofs of the results we mention.

Suppose (Ω, μ) is a probability measure space, $T : \Omega \rightarrow \Omega$ is an invertible ergodic transformation, and $f : \Omega \rightarrow \mathbb{R}$ is bounded and measurable. We define potentials,

$$V_\omega(n) = f(T^n \omega), \quad \omega \in \Omega, \quad n \in \mathbb{Z},$$

and the corresponding discrete Schrödinger operators in $\ell^2(\mathbb{Z})$,

$$[H_\omega \psi](n) = \psi(n+1) + \psi(n-1) + V_\omega(n)\psi(n). \quad (1)$$

We will call $\{H_\omega\}_{\omega \in \Omega}$ an ergodic family of Schrödinger operators.

Examples. (a) Quasi-periodic potentials: $\Omega = \mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$, μ is the normalized Lebesgue measure on \mathbb{T}^d , and $T\omega = \omega + \alpha$ is some ergodic shift (i.e., $1, \alpha_1, \dots, \alpha_d$ are rationally independent).¹

(b) Potentials defined by the skew shift: $\Omega = \mathbb{T}^2$, μ is the normalized Lebesgue measure on \mathbb{T}^2 , and $T(\omega_1, \omega_2) = (\omega_1 + \omega_2, \omega_2 + \alpha)$ for some irrational α .

¹What we call quasi-periodic here is more general than the notion of quasi-periodicity as defined in [7], for example, where a quasi-periodic potential is almost periodic with a finitely generated frequency module. In particular, a quasi-periodic potential as defined here is not necessarily almost periodic, that is, the translates of a given quasi-periodic potential are not necessarily precompact in $\ell^\infty(\mathbb{Z})$. We allow discontinuous sampling functions f here because we want to include Fibonacci-type potentials.

(c) Potentials defined by the doubling map: $\Omega = \mathbb{T}$, μ is the normalized Lebesgue measure on \mathbb{T} , and $T\omega = 2\omega$.²

(d) Potentials defined by the left shift: $\Omega = I^{\mathbb{Z}}$, where I is a compact subset of \mathbb{R} , $\mu = \mathbb{P}^{\mathbb{Z}}$, where \mathbb{P} is a Borel probability measure on I , and $[T\omega](n) = \omega(n + 1)$.

The following pair of results, proven in [50, 56], shows that for ergodic families of Schrödinger operators, the spectrum and the spectral type are deterministic in the sense that they are constant μ -almost surely.

THEOREM 1 (Pastur 1980). *There exists a set $\Sigma \subset \mathbb{R}$ such that for μ -almost every ω , $\sigma(H_\omega) = \Sigma$.*

THEOREM 2 (Kunz–Souillard 1980). *There are sets $\Sigma_{ac}, \Sigma_{sc}, \Sigma_{pp} \subset \mathbb{R}$ such that for μ -almost every ω , $\sigma_\bullet(H_\omega) = \Sigma_\bullet$, $\bullet \in \{ac, sc, pp\}$.*

Thus, in the spectral analysis of a given ergodic family of Schrödinger operators, a fundamental problem is the identification of the sets $\Sigma, \Sigma_{ac}, \Sigma_{sc}$, and Σ_{pp} .

The almost sure spectrum, Σ , is completely described by the integrated density of states as shown by Avron and Simon [9]. Denote the restriction of H_ω to $[0, N - 1]$ with Dirichlet boundary conditions by $H_\omega^{(N)}$. For $\omega \in \Omega$ and $N \geq 1$, define measures $dk_{\omega,N}$ by placing uniformly distributed point masses at the eigenvalues $E_\omega^{(N)}(1) < \dots < E_\omega^{(N)}(N)$ of $H_\omega^{(N)}$, that is,

$$\int f(E) dk_{\omega,N}(E) = \frac{1}{N} \sum_{n=1}^N f(E_\omega^{(N)}(n)).$$

Then, it can be shown that for μ -almost every $\omega \in \Omega$, the measures $\tilde{dk}_{\omega,N}$ converge weakly to a non-random measure dk , called the *density of states measure*, as $N \rightarrow \infty$. The function k defined by

$$k(E) = \int \chi_{(-\infty, E]}(E') dk(E')$$

is called the *integrated density of states*. It is not hard to show that

$$\int f(E) dk(E) = \mathbb{E}(\langle \delta_0, f(H_\omega)\delta_0 \rangle) \tag{2}$$

for bounded measurable f . Here, $\mathbb{E}(\cdot)$ denotes integration with respect to the measure μ , that is, $\mathbb{E}(g) = \int g(\omega) d\mu(\omega)$. Thus, the density of states measure is given by an average of the spectral measures associated with H_ω and δ_0 . The T -invariance of μ then implies the following result:

THEOREM 3 (Avron–Simon 1983). *The almost sure spectrum is given by the points of increase of k , that is, $\Sigma = \text{supp}(dk)$.*

There is a similarly general description of the set Σ_{ac} in terms of the Lyapunov exponent. Let $E \in \mathbb{C}$ and

$$A^E(\omega) = \begin{pmatrix} E - f(\omega) & -1 \\ 1 & 0 \end{pmatrix}. \tag{3}$$

²Strictly speaking, this example does not fall within our general framework as T is not invertible, but potentials of this kind have been studied in several works and it is possible to tweak the model a little to fit it in the framework above.

Define the *Lyapunov exponent* $\gamma(E)$ by

$$\gamma(E) = \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E} (\log \|A_n^E(\omega)\|),$$

where $A_n^E(\omega) = A^E(T^{n-1}\omega) \cdots A^E(\omega)$.

The integrated density of states and the Lyapunov exponent are related by the Thouless formula (see, e.g., [23, Theorem 9.20]), which reads

$$\gamma(E) = \int \log |E - E'| dk(E'). \tag{4}$$

The significance of the transfer matrices $A_n^E(\omega)$ is that a sequence u solves the difference equation

$$u(n + 1) + u(n - 1) + V_\omega(n)u(n) = Eu(n) \tag{5}$$

if and only if

$$\begin{pmatrix} u_n \\ u_{n-1} \end{pmatrix} = A_n^E(\omega) \begin{pmatrix} u_0 \\ u_{-1} \end{pmatrix}$$

for every n . Since $\det A^E(\omega) = 1$, we always have $\gamma(E) \geq 0$. Let us define

$$\mathcal{Z} = \{E : \gamma(E) = 0\}.$$

By general principles, $\mathcal{Z} \subseteq \Sigma$.

For a subset S of \mathbb{R} , the essential closure of S is given by

$$\overline{S}^{\text{ess}} = \{E \in \mathbb{R} : \text{Leb}(S \cap (E - \varepsilon, E + \varepsilon)) > 0 \text{ for every } \varepsilon > 0\}$$

Then, the following theorem combines results from [37, 43, 56].³

THEOREM 4 (Ishii 1973, Pastur 1980, Kotani 1984). *The almost sure absolutely continuous spectrum is given by the essential closure of the set of energies for which the Lyapunov exponent vanishes, that is, $\Sigma_{\text{ac}} = \overline{\mathcal{Z}}^{\text{ess}}$.*

While there is an analog of Theorem 3 for higher-dimensional ergodic Schrödinger operators, Theorem 4 is, by its very nature, a strictly one-dimensional result. It is one of the great challenges for researchers in the area of ergodic Schrödinger operators to develop effective tools for the study of the absolutely continuous spectrum in higher dimensions. That said, Theorem 4 holds in virtually all one-dimensional and quasi-one-dimensional situations: see Kotani [43] for continuous one-dimensional Schrödinger operators (see also Kirsch [41] for a useful extension), Minami [54] for generalized Sturm–Liouville operators, Kotani–Simon [49] for discrete and continuous Schrödinger operators with matrix-valued potentials, and Geronimo [31] and Geronimo–Teplyaev [32] for orthogonal polynomials on the unit circle.

³To be more precise, the discrete version of the Kotani half of this result here can be found in the paper [59] by Simon and the work of Ishii and Pastur was preceded by closely related work by Casher and Lebowitz [19].

2. The Description of the AC Spectrum

In this section we discuss the main ideas that go into the proof of Theorem 4. The proof naturally breaks up into the proof of two inclusions.

The inclusion “ \subseteq ” was proved by Ishii and Pastur. The other inclusion was proved by Kotani and is a much deeper result. In fact, the proof of the Ishii–Pastur half of the result we give below is based on more modern techniques and shows that this half is really an immediate consequence of the general theory of one-dimensional Schrödinger operators.

There are at least three different proofs of the Ishii–Pastur half of Theorem 4 in the literature. One of them uses the existence of generalized eigenfunctions; compare Cycon et al. [23]. The second one, due to Deift and Simon [30], is close in spirit to, and uses techniques from, Kotani’s proof of the other half of the result. Finally, there are two somewhat related proofs given by Buschmann [18] and Last and Simon [51], which are both either directly or indirectly based on a result of Gilbert and Pearson that describes a support of the singular spectrum of a Schrödinger operator with fixed (non-random) potential in terms of subordinate solutions. We will follow the argument from Buschmann’s paper below.

We first recall a central result from Gilbert and Pearson’s subordinacy theory [33, 34]. Consider the discrete Schrödinger operator H in $\ell^2(\mathbb{Z})$ with potential V and solutions of the difference equation

$$u(n + 1) + u(n - 1) + V(n)u(n) = Eu(n). \tag{6}$$

A non-zero solution u of (6) is called *subordinate* at $\pm\infty$ if for every linearly independent solution \tilde{u} of (6), we have

$$\frac{\sum_{n=0}^{N-1} |u(\pm n)|^2}{\sum_{n=0}^{N-1} |\tilde{u}(\pm n)|^2} \rightarrow 0 \text{ as } N \rightarrow \infty.$$

Let

$$S = \{E \in \mathbb{R} : (6) \text{ has solutions } u_+ \text{ and } u_- \text{ such that } u_{\pm} \text{ is subordinate at } \pm\infty\}.$$

Then, S has zero weight with respect to the absolutely continuous part of any spectral measure, that is,

$$\mathcal{P}^{(\text{ac})}(S) = 0. \tag{7}$$

PROOF OF THE ISHII–PASTUR HALF OF THEOREM 4. Note that when $\gamma(E) > 0$, Oseledec’s theorem [55] says that for almost every ω , there are solutions $u_+(E, \omega)$ and $u_-(E, \omega)$ of (5) such that $u_{\pm}(E, \omega)$ is exponentially decaying, and hence subordinate, at $\pm\infty$. Applying Fubini’s theorem, we see that for μ -almost every ω , the set of $E \in \mathbb{R} \setminus \mathcal{Z}$ for which the property just described fails, has zero Lebesgue measure. In other words, for these ω ’s, $\mathbb{R} \setminus \mathcal{Z} \subseteq S_{\omega}$ up to a set of zero Lebesgue measure. Since sets of zero Lebesgue measure have zero weight with respect to the absolutely continuous part of any spectral measure, we obtain from (7) that for μ -almost every ω ,

$$\mathcal{P}_{\omega}^{(\text{ac})}(\mathbb{R} \setminus \mathcal{Z}) = 0.$$

This shows that for μ -almost every ω , $\sigma_{\text{ac}}(H_{\omega}) \subseteq \overline{\mathcal{Z}}^{\text{ess}}$. □

Let us now turn to the Kotani half of Theorem 4. Kotani worked in the continuum setting. Carrying his results over to the discrete case is not entirely straightforward and it was worked out by Simon [59] whose proof we give below.

Given $z \in \mathbb{C}_+ = \{z \in \mathbb{C} : \Im z > 0\}$ and $\omega \in \Omega$, there are (up to a multiplicative constant) unique solutions $u_{\pm}(n, \omega)$ of

$$u(n + 1) + u(n - 1) + V_{\omega}(n)u(n) = zu(n) \tag{8}$$

such that u_{\pm} is square-summable at $\pm\infty$. (Take $u_{\pm}(n, \omega) = \langle \delta_n, (H_{\omega} - z)^{-1} \delta_1 \rangle$ near $\pm\infty$ to show existence; uniqueness follows from constancy of the Wronskian.) Note that $u_{\pm}(0, \omega) \neq 0$ for otherwise z would be a non-real eigenvalue of a self-adjoint half-line operator. Thus, we can define

$$m_{\pm}(z, \omega) = -\frac{u_{\pm}(\pm 1, \omega)}{u_{\pm}(0, \omega)}. \tag{9}$$

Clearly,

$$m_{\pm}(z, T^n \omega) = -\frac{u_{\pm}(n \pm 1, \omega)}{u_{\pm}(n, \omega)}. \tag{10}$$

By Oseledec’s theorem, we have for μ -almost every ω ,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \left| \frac{u_{\pm}(n, \omega)}{u_{\pm}(0, \omega)} \right| = -\gamma(z).$$

By (10),

$$\log \left| \frac{u_{\pm}(n, \omega)}{u_{\pm}(0, \omega)} \right| = \sum_{m=0}^{n-1} \log |m_{\pm}(z, T^{\pm m} \omega)|$$

and hence Birkhoff’s ergodic theorem implies

$$\mathbb{E}(\log |m_{\pm}(z, \omega)|) = -\gamma(z). \tag{11}$$

PROPOSITION 2.1. *We have that*

$$\mathbb{E} \left(\log \left(1 + \frac{\Im z}{\Im m_{\pm}(z, \omega)} \right) \right) = 2\gamma(z).$$

PROOF. By the difference equation (8) that u_{\pm} obeys,

$$m_{\pm}(z, T^n \omega) = V_{\omega}(n) - z - [m_{\pm}(z, T^{n \mp 1} \omega)]^{-1}. \tag{12}$$

Taking imaginary parts,

$$\Im m_{\pm}(z, \omega) = -\Im z - \Im ([m_{\pm}(z, T^{\mp 1} \omega)]^{-1}).$$

Dividing by $\Im m_{+}(z, \omega)$,

$$1 = -\frac{\Im z}{\Im m_{\pm}(z, \omega)} - \frac{\Im ([m_{\pm}(z, T^{\mp 1} \omega)]^{-1})}{\Im m_{\pm}(z, \omega)}.$$

Taking the logarithm,

$$\begin{aligned} \log \left(1 + \frac{\Im z}{\Im m_{\pm}(z, \omega)} \right) &= \log (-\Im ([m_{\pm}(z, T^{\mp 1} \omega)]^{-1})) - \log (\Im m_{\pm}(z, \omega)) \\ &= \log \left(\frac{\Im m_{\pm}(z, T^{\mp 1} \omega)}{|m_{\pm}(z, T^{\mp 1} \omega)|^2} \right) - \log (\Im m_{\pm}(z, \omega)) \end{aligned}$$

Taking expectations and using invariance,

$$\begin{aligned} \mathbb{E} \left(\log \left(1 + \frac{\Im z}{\Im m_{\pm}(z, \omega)} \right) \right) &= \mathbb{E} \left(\log \left(\frac{\Im m_{\pm}(z, T^{\mp 1} \omega)}{|m_{\pm}(z, T^{\mp 1} \omega)|^2} \right) - \log (\Im m_{\pm}(z, \omega)) \right) \\ &= -2 \mathbb{E} (\log |m_{\pm}(z, \omega)|) \\ &= 2\gamma(z), \end{aligned}$$

where we used (11) in the last step. □

Denote

$$b(z, \omega) = m_+(z, \omega) + m_-(z, \omega) + z - V_\omega(0).$$

PROPOSITION 2.2. *We have*

$$\mathbb{E} \left(\Im \left(\frac{1}{b(z, \omega)} \right) \right) = -\frac{\partial \gamma(z)}{\partial(\Im z)}.$$

PROOF. It follows from (12) that

$$\frac{u_-(1, \omega)}{u_-(0, \omega)} = m_-(z, \omega) + z - V_\omega(0). \tag{13}$$

It is not hard to check that for $n \leq m$,

$$G_\omega(n, m; z) := \langle \delta_n, (H_\omega - z)^{-1} \delta_m \rangle = \frac{u_-(n, \omega)u_+(m, \omega)}{u_+(1, \omega)u_-(0, \omega) - u_-(1, \omega)u_+(0, \omega)}. \tag{14}$$

From (9), (13), and (14), we get

$$-G_\omega(0, 0; z)^{-1} = m_+(z, \omega) + m_-(z, \omega) + z - V_\omega(0) = b(z, \omega). \tag{15}$$

The definition of $G_\omega(n, m; z)$ gives

$$\mathbb{E}(G_\omega(0, 0; z)) = \int \frac{1}{E' - z} dk(E'). \tag{16}$$

Thus,

$$\begin{aligned} \mathbb{E} \left(\Im \left(\frac{1}{b(z, \omega)} \right) \right) &= -\Im \mathbb{E}(G_\omega(0, 0; z)) \\ &= -\Im \int \frac{1}{E' - z} dk(E') \\ &= -\frac{\partial}{\partial(\Im z)} \int \log |z - E'| dk(E') \\ &= -\frac{\partial \gamma(z)}{\partial(\Im z)}, \end{aligned}$$

where we used (15), (16), and the Thouless formula (4). □

Denote

$$n_\pm(z, \omega) = \Im m_\pm(z, \omega) + \frac{1}{2} \Im z.$$

PROPOSITION 2.3. *We have that*

$$\mathbb{E} \left(\frac{1}{n_\pm(z, \omega)} \right) \leq \frac{2\gamma(z)}{\Im z} \tag{17}$$

and

$$\mathbb{E} \left(\frac{\left[\frac{1}{n_+} + \frac{1}{n_-} \right] \cdot [(n_+ - n_-)^2 + (\Re b)^2]}{|b|^2} \right) \leq 4 \left[\frac{\gamma(z)}{\Im z} - \frac{\partial \gamma(z)}{\partial(\Im z)} \right]. \tag{18}$$

PROOF. For $x \geq 0$, consider the function $A(x) = \log(1+x) - \frac{x}{1+\frac{x}{2}}$. Clearly, $A(0) = 0$ and $A'(x) = \frac{1}{1+x} - \frac{1}{1+x+\frac{x^2}{4}} \geq 0$. Therefore,

$$\log(1+x) \geq \frac{x}{1+\frac{x}{2}} \quad \text{for all } x \geq 0. \tag{19}$$

Thus,

$$\begin{aligned} \mathbb{E} \left(\frac{1}{n_{\pm}(z, \omega)} \right) &= \mathbb{E} \left(\frac{1}{\Im m_{\pm}(z, \omega) + \frac{1}{2} \Im z} \right) \\ &= \frac{1}{\Im z} \mathbb{E} \left(\frac{\frac{\Im z}{\Im m_{\pm}(z, \omega)}}{1 + \frac{\frac{\Im z}{\Im m_{\pm}(z, \omega)}}{2}} \right) \\ &\leq \frac{1}{\Im z} \mathbb{E} \left(\log \left(1 + \frac{\Im z}{\Im m_{\pm}(z, \omega)} \right) \right) \\ &= \frac{2\gamma(z)}{\Im z}, \end{aligned}$$

which is (17). We used (19) in the third step and Proposition 2.1 in the last step.

Notice that $n_+(z, \omega) + n_-(z, \omega) = \Im b(z, \omega)$. Thus, the integrand on the left-hand side of (18) is equal to

$$\begin{aligned} \frac{\left[\frac{1}{n_+} + \frac{1}{n_-} \right] \cdot [(n_+ + n_-)^2 - 4n_+n_- + (\Re b)^2]}{|b|^2} &= \frac{\left[\frac{1}{n_+} + \frac{1}{n_-} \right] \cdot [|b|^2 - 4n_+n_-]}{|b|^2} \\ &= \frac{1}{n_+} + \frac{1}{n_-} - 4 \frac{n_+ + n_-}{|b|^2} \\ &= \frac{1}{n_+} + \frac{1}{n_-} + 4 \Im \left(\frac{1}{b} \right). \end{aligned}$$

The bound (18) now follows from (17) and Proposition 2.2. □

PROOF OF THE KOTANI HALF OF THEOREM 4. The Thouless formula (4) says that

$$\gamma(z) = \int \log |z - E'| dk(E') = \Re \int \log(z - E') dk(E')$$

and hence $-\gamma(z)$ is the real part of a function whose derivative is a Borel transform (namely, of the measure dk). By general properties of the Borel transform, it follows that the limit $\gamma'(E+i0)$ exists for Lebesgue almost every $E \in \mathbb{R}$ and, in particular, for almost every $E \in \mathcal{Z}$. For these E , we have that

$$\lim_{\varepsilon \downarrow 0} \frac{\gamma(E+i\varepsilon)}{\varepsilon} = \lim_{\varepsilon \downarrow 0} \frac{\gamma(E+i\varepsilon) - \gamma(E)}{\varepsilon - 0} = \lim_{\varepsilon \downarrow 0} \frac{\partial \gamma}{\partial (\Im z)}(E+i\varepsilon), \tag{20}$$

and, in particular, the limit is finite. Thus, by (17),

$$\limsup_{\varepsilon \downarrow 0} \mathbb{E} \left(\frac{1}{\Im m_{\pm}(E+i\varepsilon, \omega)} \right) < \infty$$

for almost every $E \in \mathcal{Z}$. Since m_{\pm} are Borel transforms as well (of the spectral measures associated with half-line restrictions of H_{ω}), we also have that, for every $\omega \in \Omega$, $m_{\pm}(E+i0, \omega)$ exists for Lebesgue almost every $E \in \mathbb{R}$, and hence, for

almost every E , $m_{\pm}(E + i0, \omega)$ exists for almost every ω . Combining the last two observations with Fatou's lemma, we find that

$$\mathbb{E} \left(\frac{1}{\Im m_{\pm}(E + i0, \omega)} \right) < \infty \tag{21}$$

for almost every E in \mathcal{Z} . So, for almost every $\omega \in \Omega$ and $E \in \mathcal{Z}$, $\Im m_{\pm}(E + i0, \omega) > 0$.

On the other hand, $m_{+}(E + i\varepsilon, \omega) + m_{-}(E + i\varepsilon, \omega) + E + i\varepsilon - V_{\omega}(0)$ has a finite limit for almost every $\omega \in \Omega$ and $E \in \mathcal{Z}$.

Hence, (15) shows that $0 < \Im G_{\omega}(0, 0; E + i0) < \infty$ for almost every $\omega \in \Omega$ and $E \in \mathcal{Z}$, which implies the result. \square

Denote the measure associated with the Herglotz function $G_{\omega}(0, 0; z)$ by ν_{ω} , that is,

$$G_{\omega}(0, 0; z) = \int \frac{d\nu_{\omega}(E)}{E - z}.$$

The results above imply the following for μ -almost every ω :

$$\begin{aligned} \nu_{\omega}^{(\text{ac})}(E) &= 0 \text{ for Lebesgue almost every } E \in \mathbb{R} \setminus \mathcal{Z}, \\ \nu_{\omega}^{(\text{ac})}(E) &> 0 \text{ for Lebesgue almost every } E \in \mathcal{Z}. \end{aligned}$$

Here, $\nu_{\omega}^{(\text{ac})}(E)$ denotes the density of the absolutely continuous part of ν_{ω} . Write $k^{(\text{ac})}(E)$ for the density of the absolutely continuous part of the density of states measure.

There is a direct relation between these densities [47]:

THEOREM 5 (Kotani 1997). *For almost every $E \in \mathcal{Z}$,*

$$k^{(\text{ac})}(E) = \mathbb{E} \left(\nu_{\omega}^{(\text{ac})}(E) \right). \tag{22}$$

PROOF. The inequality “ \geq ” in (22) follows from (2) (i.e., the density of states measure is the average of the measures ν_{ω}) and the fact that the average of absolutely continuous measures is absolutely continuous.

To prove the opposite inequality, we first note that for almost every $E \in \mathcal{Z}$, (16), (20), and Cauchy–Riemann imply

$$k^{(\text{ac})}(E) = \frac{1}{\pi} \lim_{\varepsilon \downarrow 0} \frac{\gamma(E + i\varepsilon)}{\varepsilon}. \tag{23}$$

Because of (20), (21), and Fatou's lemma, (18) implies that for almost every pair $(E, \omega) \in \mathcal{Z} \times \Omega$,

$$\Im m_{+}(E + i0, \omega) = \Im m_{-}(E + i0, \omega) \tag{24}$$

and

$$\Re m_{+}(E + i0, \omega) + \Re m_{-}(E + i0, \omega) + E - V_{\omega}(0) = 0. \tag{25}$$

Thus, for almost every $(E, \omega) \in \mathcal{Z} \times \Omega$,

$$\begin{aligned}
 \nu_\omega^{(\text{ac})}(E) &= \frac{1}{\pi} \Im G_\omega(0, 0; E + i0) \\
 &= \frac{1}{\pi} \Im \frac{-1}{m_+(E + i0, \omega) + m_-(E + i0, \omega) + E - V_\omega(0)} \\
 &= \frac{1}{\pi} \Im \frac{-1}{2i \Im m_+(E + i0, \omega)} \\
 &= \frac{1}{2\pi} \frac{1}{\Im m_+(E + i0, \omega)}.
 \end{aligned}
 \tag{26}$$

Let P_ϵ be the Poisson kernel for the upper half-plane, that is,

$$P_\epsilon(E) = \frac{1}{\pi} \frac{\epsilon}{E^2 + \epsilon^2}.$$

Write

$$C_\epsilon(E) = \int_{\mathcal{Z}} P_\epsilon(E - E') dE'$$

and

$$\tilde{P}_\epsilon(E, E') = P_\epsilon(E - E') C_\epsilon(E)^{-1}.$$

Then, by (26) and Jensen’s inequality, we obtain for almost every $(E, \omega) \in \mathcal{Z} \times \Omega$,

$$\begin{aligned}
 \int_{\mathbb{R}} \nu_\omega^{(\text{ac})}(E') P_\epsilon(E - E') dE' &\geq \int_{\mathcal{Z}} \nu_\omega^{(\text{ac})}(E') P_\epsilon(E - E') dE' \\
 &= \int_{\mathcal{Z}} \left(\frac{1}{2\pi} \frac{1}{\Im m_+(E' + i0, \omega)} \right) P_\epsilon(E - E') dE' \\
 &= C_\epsilon(E) \int_{\mathcal{Z}} \left(\frac{1}{2\pi} \frac{1}{\Im m_+(E' + i0, \omega)} \right) \tilde{P}_\epsilon(E, E') dE' \\
 &\geq C_\epsilon(E) \left(\int_{\mathcal{Z}} \left(\frac{1}{2\pi} \frac{1}{\Im m_+(E' + i0, \omega)} \right)^{-1} \tilde{P}_\epsilon(E, E') dE' \right)^{-1} \\
 &\geq C_\epsilon(E)^2 \left(\int_{\mathbb{R}} \left(\frac{1}{2\pi} \frac{1}{\Im m_+(E' + i0, \omega)} \right)^{-1} P_\epsilon(E - E') dE' \right)^{-1} \\
 &= \frac{C_\epsilon(E)^2}{2\pi} \frac{1}{\Im m_+(E + i\epsilon, \omega)}.
 \end{aligned}$$

Thus, for almost every $E \in \mathcal{Z}$,

$$\int_{\mathbb{R}} \mathbb{E} \left(\nu_\omega^{(\text{ac})}(E') \right) P_\epsilon(E - E') dE' \geq \frac{C_\epsilon(E)^2}{2\pi} \mathbb{E} \left(\frac{1}{\Im m_+(E + i\epsilon, \omega)} \right),$$

and hence

$$\mathbb{E} \left(\nu_\omega^{(\text{ac})}(E) \right) \geq \frac{1}{2\pi} \limsup_{\epsilon \downarrow 0} \mathbb{E} \left(\frac{1}{\Im m_+(E + i\epsilon, \omega)} \right) \tag{27}$$

since $C_\epsilon(E) < 1$ and $C_\epsilon(E) \rightarrow 1$ as $\epsilon \downarrow 0$.

Using (23), Proposition 2.1, the inequality $\log(1 + x) \leq x$ for $x \geq 0$, and then (27), we find that

$$\begin{aligned} k^{(\text{ac})}(E) &= \frac{1}{\pi} \lim_{\varepsilon \downarrow 0} \frac{\gamma(E + i\varepsilon)}{\varepsilon} \\ &= \frac{1}{\pi} \lim_{\varepsilon \downarrow 0} \frac{1}{2\varepsilon} \mathbb{E} \left(\log \left(1 + \frac{\varepsilon}{\Im m_{\pm}(E + i\varepsilon, \omega)} \right) \right) \\ &\leq \frac{1}{2\pi} \limsup_{\varepsilon \downarrow 0} \mathbb{E} \left(\frac{1}{\Im m_{\pm}(E + i\varepsilon, \omega)} \right) \\ &\leq \mathbb{E} \left(\nu_{\omega}^{(\text{ac})}(E) \right), \end{aligned}$$

concluding the proof of “ \leq ” in (22). □

COROLLARY 1. *The spectrum is almost surely purely absolutely continuous if and only if the integrated density of states is absolutely continuous and the Lyapunov exponent vanishes almost everywhere with respect to the density of states measure.*

There is a different approach to purely absolutely continuous spectrum as pointed out by Yoram Last (unpublished):⁴ Using a result of Deift and Simon [30, Theorem 7.1], one can show that there is a set $\mathcal{R} \subseteq \mathcal{Z}$ such that $\text{Leb}(\mathcal{Z} \setminus \mathcal{R}) = 0$ and \mathcal{R} has zero singular spectral measure for every $\omega \in \Omega$ due to the absence of subordinate solutions.

The spectrum naturally breaks up into the two components \mathcal{Z} and $\Sigma \setminus \mathcal{Z}$. It is known that either set can support singular continuous spectrum as demonstrated by the almost Mathieu operator at critical and super-critical coupling. However, as we have seen, $\Sigma \setminus \mathcal{Z}$ does not support any absolutely continuous part of the spectral measures. Trivially, Anderson localization is impossible in \mathcal{Z} . However, it is tempting to expect even more:

PROBLEM 1. *Prove or disprove that for all ergodic families, the operators H_{ω} have no eigenvalues in \mathcal{Z} .*

If the answer is affirmative, this will in particular imply the absence of eigenvalues in a number of special cases, such as the operators considered in Section 6 and the critical almost Mathieu operator.⁵

3. The Induced Measure and Its Topological Support

Fix a compact subset R of \mathbb{R} . Endow $R^{\mathbb{Z}}$ with product topology, which makes it a compact metric space. If $V \in R^{\mathbb{Z}}$, we define the functions m_{\pm} by

$$m_{\pm}(z) = \mp \frac{u_{\pm}(1)}{u_{\pm}(0)},$$

where u_{\pm} solves

$$u(n + 1) + u(n - 1) + V(n)u(n) = zu(n) \tag{28}$$

and is ℓ^2 at $\pm\infty$.

Denote $\mathbb{Z}_+ = \{1, 2, 3, \dots\}$ and $\mathbb{Z}_- = \{0, -1, -2, \dots\}$. It is well known that the maps $\mathcal{M}_{\pm} : V_{\pm} = V|_{\mathbb{Z}_{\pm}} \mapsto m_{\pm}$ are one-one and continuous with respect to

⁴The author is grateful to Barry Simon for bringing this to his attention.

⁵For this particular model, this would provide an alternative to the proof based on self-duality and zero-measure spectrum.

uniform convergence on compacta on the m -function side. The m -functions m_{\pm} are Herglotz functions, they have boundary values almost everywhere on the real axis, and they are completely determined by their boundary values on any set of positive Lebesgue measure.

We will be interested in those V for which the functions m_+, m_- obey identities like (24) and (25), that is,

$$m_-(E + i0) = -\overline{m_+(E + i0)} \tag{29}$$

for a rich set of energies. Thus, for a set $\mathcal{Z} \subseteq \mathbb{R}$, we let

$$\mathcal{D}(\mathcal{Z}) = \{V \in \mathbb{R}^{\mathbb{Z}} : m_{\pm} \text{ associated with } V \text{ obey (29) for a.e. } E \in \mathcal{Z}\}.$$

On $\mathbb{R}^{\mathbb{Z}}$, define the shift transformation $[S(V)](n) = V(n + 1)$.

LEMMA 3.1. *Suppose that $\mathcal{Z} \subset \mathbb{R}$ has positive Lebesgue measure. Then:*

- (a) $\mathcal{D}(\mathcal{Z})$ is S -invariant and closed in $\mathbb{R}^{\mathbb{Z}}$.
- (b) For $V \in \mathcal{D}(\mathcal{Z})$, denote the restrictions to \mathbb{Z}_{\pm} by V_{\pm} . Then V_- determines V_+ uniquely among elements of $\mathcal{D}(\mathcal{Z})$ and vice versa.
- (c) If there exist $V^{(m)}, V \in \mathcal{D}(\mathcal{Z})$ such that $V_-^{(m)} \rightarrow V_-$ pointwise, then $V_+^{(m)} \rightarrow V_+$ pointwise.

PROOF. (a) If u_1, u_2 denote the solutions of (28) that obey $u_1(0) = u_2(1) = 1$ and $u_1(1) = u_2(0) = 0$, then we can write (note that we may normalize u_{\pm} by $u_{\pm}(0) = 1$)

$$u_{\pm}(n) = u_1(n) \mp m_{\pm}(z)u_2(n).$$

Let us denote the m -functions associated with $S(V)$ by \tilde{m}_{\pm} . Clearly,

$$\tilde{m}_{\pm}(z) = \mp \frac{u_1(2) \mp m_{\pm}(z)u_2(2)}{u_1(1) \mp m_{\pm}(z)u_2(1)}.$$

Since the $u_j(m)$ are polynomials in z with real coefficients, this shows that

$$\tilde{m}_-(E + i0) = -\overline{\tilde{m}_+(E + i0)}$$

for almost every $E \in \mathcal{Z}$ and hence $\mathcal{D}(\mathcal{Z})$ is S -invariant. It follows from the continuity of the maps \mathcal{M}_{\pm} that $\mathcal{D}(\mathcal{Z})$ is closed. (For a proof that the identity between the boundary values of the associated m -functions is preserved after taking limits, see [44, Lemma 5] and [45, Lemma 7.4].)

(b) V_- determines m_- and then (29) determines the boundary values of m_+ on a set of positive Lebesgue measure. By general properties of Herglotz functions, this determines m_+ (and hence V_+) completely. By the same argument, V_+ determines V_- .

(c) By compactness, there is a subsequence of $\{V^{(m)}\}$ that converges pointwise, that is, there is \tilde{V} such that $V^{(m_k)} \rightarrow \tilde{V}$ as $k \rightarrow \infty$. By part (a), $\tilde{V} \in \mathcal{D}(\mathcal{Z})$. By assumption, $V_- = \tilde{V}_-$. Thus, by part (b), $V_+ = \tilde{V}_+$, and hence $V = \tilde{V}$. Consequently, $V_+^{(m_k)} \rightarrow V_+$ pointwise. In fact, we claim that $V_+^{(m)} \rightarrow V_+$ pointwise. Otherwise, we could reverse the argument (i.e., go from right to left) and show that $V_-^{(m_k)} \not\rightarrow V_-$ for some other subsequence. \square

We will now derive two important consequences of Lemma 3.1: the absence of absolutely continuous spectrum for topologically non-deterministic families and the support theorem. To do so, we will consider the push-forward ν of μ on the sequence space and its topological support.

More precisely, given an ergodic dynamical system (Ω, μ, T) and a measurable bounded sampling function $f : \Omega \rightarrow \mathbb{R}$ defining potentials $V_\omega(n) = f(T^n\omega)$ as before, we associate the following dynamical system $(R^{\mathbb{Z}}, \nu, S)$: R is a compact set that contains the range of f , ν is the Borel measure on $R^{\mathbb{Z}}$ induced by μ via $\Phi(\omega) = V_\omega$ (i.e., $\nu(A) = \mu(\Phi^{-1}(A))$), and S is the shift transformation on $R^{\mathbb{Z}}$ introduced above. Recall that the *topological support* of ν , $\text{supp } \nu$, is given by the intersection of all compact sets $B \subseteq R^{\mathbb{Z}}$ with $\nu(B) = 1$. Clearly, $\text{supp } \nu$ is closed and S -invariant.

THEOREM 6 (Kotani 1989⁶). *Let (Ω, μ, T, f) and $(R^{\mathbb{Z}}, d\nu, S)$ be as just described. Assume that the set*

$$\mathcal{Z} = \{E \in \mathbb{R} : \gamma(E) = 0\}$$

has positive Lebesgue measure. Then,

(a) *Each $V \in \text{supp } \nu$ is determined completely by V_- (resp., V_+).*

(b) *If we let*

$$(\text{supp } \nu)_\pm = \{V_\pm : V \in \text{supp } \nu\},$$

then the mappings

$$E_\pm : (\text{supp } \nu)_\pm \ni V_\pm \mapsto V_\mp \in (\text{supp } \nu)_\mp \tag{30}$$

are continuous with respect to pointwise convergence.

PROOF. (a) By our earlier results, we know that $\mathcal{D}(\mathcal{Z})$ is compact and has full ν -measure. Thus, $\text{supp } \nu \subseteq \mathcal{D}(\mathcal{Z})$ and the assertion follows from Lemma 3.1(b).

(b) This follows from Lemma 3.1(c). □

We say that (Ω, μ, T, f) is *topologically deterministic* if there exist continuous mappings $E_\pm : (\text{supp } \nu)_\pm \rightarrow (\text{supp } \nu)_\mp$ that are formal inverses of one another and obey $V_-^\# \in \text{supp } \nu$ for every $V_- \in (\text{supp } \nu)_-$, where

$$V_-^\#(n) = \begin{cases} V_-(n) & n \leq 0, \\ E_-(V_-)(n) & n \geq 1. \end{cases}$$

This also implies $V_+^\# \in \text{supp } \nu$ for every $V_+ \in (\text{supp } \nu)_+$, where

$$V_+^\#(n) = \begin{cases} V_+(n) & n \geq 1, \\ E_+(V_+)(n) & n \leq 0. \end{cases}$$

Otherwise, (Ω, μ, T, f) is *topologically non-deterministic*.

COROLLARY 2. *If (Ω, μ, T, f) is topologically non-deterministic, then $\Sigma_{\text{ac}} = \emptyset$.*

Our next application of Lemma 3.1 is the so-called support theorem; compare [44]. For a Borel measure ν on $R^{\mathbb{Z}}$, let $\Sigma_{\text{ac}}(\nu) \subseteq \mathbb{R}$ denote the almost sure absolutely continuous spectrum, that is, $\sigma_{\text{ac}}(\Delta + V) = \Sigma_{\text{ac}}(\nu)$ for ν -almost every V . If ν comes from (Ω, μ, T, f) , then $\Sigma_{\text{ac}}(\nu)$ coincides with the set Σ_{ac} introduced earlier. The support theorem says that $\Sigma_{\text{ac}}(\nu)$ is monotonically decreasing in the support of ν .

THEOREM 7 (Kotani 1985). *For every $V \in \text{supp } \nu$, we have $\sigma_{\text{ac}}(\Delta + V) \supseteq \Sigma_{\text{ac}}(\nu)$. In particular, $\text{supp } \nu_1 \subseteq \text{supp } \nu_2$ implies that $\Sigma_{\text{ac}}(\nu_1) \supseteq \Sigma_{\text{ac}}(\nu_2)$.*

⁶The result appears explicitly in the 1989 paper [46]. The main ingredients of the proof, however, were found earlier [44, 45].

PROOF. We know that $\text{supp } \nu \subseteq \mathcal{D}(\mathcal{Z}) = \{V \in \mathbb{R}^{\mathbb{Z}} : m_{\pm} \text{ associated with } V \text{ obey (29) for a.e. } E \in \mathcal{Z}\}$. Bearing in mind the Riccati equation (13), a calculation like the one in (26) therefore shows that for every $V \in \text{supp } \nu$, the Green function associated with the operator $\Delta + V$ obeys $\Im G(0, 0; E + i0) > 0$ for almost every $E \in \mathcal{Z}$. This implies $\overline{\mathcal{Z}}^{\text{ess}} \subseteq \sigma_{\text{ac}}(\Delta + V)$ and hence the result by Theorem 4. \square

A different proof can be found in Last–Simon [51, Sect. 6]. Here is a typical application of the support theorem:

COROLLARY 3. *Let Per_{ν} be the set of $V \in \text{supp } \nu$ that are periodic, that is, $S^p V = V$ for some $p \in \mathbb{Z}_+$. Then,*

$$\Sigma_{\text{ac}}(\nu) \subseteq \bigcap_{V \in \text{Per}_{\nu}} \sigma(\Delta + V).$$

If there are sufficiently many gaps in the spectra of these periodic operators, one can show in this way that $\Sigma_{\text{ac}}(\nu)$ is empty.

Corollaries 2 and 3 have been used in a variety of scenarios to prove the absence of absolutely continuous spectrum. In fact, while the Kotani half of Theorem 4 concerns the *presence* of absolutely continuous spectrum on \mathcal{Z} , it could be argued that the criteria for the *absence* of absolutely continuous spectrum that are by-products of the theory have been applied more often. To a certain extent, this is due to the fact that the majority of the ergodic families of Schrödinger operators are expected to have no absolutely continuous spectrum. The following “conjecture” is tempting because it is supported by a plethora of results, both on the positive side and on the negative side. It has been verbally suggested by Yoram Last and it has appeared explicitly in print in several places, including [39, 48].

PROBLEM 2. *Show that $\text{Leb}(\mathcal{Z}) > 0$ implies almost periodicity, that is, the closure in $\ell^{\infty}(\mathbb{Z})$ of the set of translates of V_{ω} is compact.*

Namely, the presence of (purely) absolutely continuous spectrum is known for all periodic potentials, many limit-periodic potentials,⁷ and some quasi-periodic potentials (that are all uniformly almost periodic). On the other hand, the absence of absolutely continuous spectrum is known for large classes of non-almost periodic ergodic potentials. We will see some instances of the latter statement below. Nevertheless, proving this conjecture is presumably very hard and it would already be interesting to find further specific results that support the conjecture. For example, it is an open (and seemingly hard) problem to prove the absence of absolutely continuous spectrum for potentials defined by the skew shift.

We close this section with a problem concerning strips. As was mentioned at the end of the Introduction, Kotani and Simon developed the analog of Kotani theory for discrete and continuous Schrödinger operators with matrix-valued potentials in their 1988 paper [49]. This framework includes in particular discrete Schrödinger operators on strips. That is, operators of the form $\Delta + V_{\omega}$ on $\ell^2(\mathbb{Z} \times \{1, \dots, L\})$, where Δ is again given by the summation over nearest neighbors. Transfer matrices are now $2L \times 2L$ and, modulo symmetry, there are L Lyapunov exponents, $\gamma_L(E) \geq \gamma_{L-1}(E) \geq \dots \geq \gamma_1(E) \geq 0$. For the general matrix-valued situation, they proved that the largest Lyapunov exponent is positive for almost every energy

⁷A sequence V is limit-periodic if there are periodic sequences $V^{(m)}$ such that $\|V - V^{(m)}\|_{\infty} \rightarrow 0$ as $m \rightarrow \infty$. See, for example, Avron–Simon [7], Chulaevsky [20], Chulaevsky–Molchanov [21], and Pastur–Tkachenko [58].

if the potential is non-deterministic. This result cannot be improved in this general setting. However, it is reasonable to expect that for strips, the following result should hold.

PROBLEM 3. *Prove that for non-deterministic Schrödinger operators on a strip, all Lyapunov exponents are non-zero for Lebesgue almost all energies.*

4. Potentials Generated by the Doubling Map

In this section we discuss potentials defined over the doubling map, that is, Example (c) from the Introduction. The underlying dynamical system is strongly mixing, and one would hope that the spectral theory of the associated operators is akin to that of the Anderson model, where the potentials are generated by independent, identically distributed random variables. Alas, by dropping independence, one loses the availability of most tools that have proven useful in the study of the Anderson model.

While localization is expected for Schrödinger operators with potentials over the doubling map, it has not been shown to hold in reasonable generality. There are only two localization results in the literature, and each of them is to some extent unsatisfactory. The first result was found by Bourgain and Schlag [17], who proved localization at small coupling and away from small intervals about the energies ± 2 and 0. Both assumptions seem unnatural. The other result is due to Damanik and Killip [26], who proved localization for essentially all f but only for Lebesgue almost every boundary condition at the origin (recall that we are dealing with operators in $\ell^2(\mathbb{Z}_+)$). A result holding for fixed boundary condition would of course be more desirable. To this end, Damanik and Killip were at least able to show the absence of absolutely continuous spectrum for fixed boundary condition in complete generality. These results are indeed immediate consequences of Kotani theory and spectral averaging, and we give the short proofs below for the reader's convenience.

The first step in a localization proof for a one-dimensional Schrödinger operator is typically a proof of positive Lyapunov exponents for many energies. For the Anderson model, this can be done for all energies using Fürstenberg's theorem or for Lebesgue almost all energies using Kotani theory. At small coupling there is also a perturbative approach due to Pastur and Figotin [57]. The extension of the approach based on Fürstenberg's theorem to potentials generated by the doubling map is not obvious; see, however, [4]. The perturbative approach extends quite nicely as shown by Chulaevsky and Spencer [22]. Their results form the basis for the proof of the partial localization result in [17]. Finally, the approach based on Kotani theory also extends, as we will now explain.

THEOREM 8 (Damanik–Killip 2005). *Suppose that $f \in L^\infty(\mathbb{T})$ is non-constant and $V_\omega(n) = f(2^n\omega)$ for $n \geq 1$. Then, the Lyapunov exponent $\gamma(E)$ is positive for Lebesgue almost every $E \in \mathbb{R}$ and the absolutely continuous spectrum of the operator H_ω in $\ell^2(\mathbb{Z}_+)$ is empty for Lebesgue almost every $\omega \in \mathbb{T}$.*

PROOF. Since the proof of this result is so short, we reproduce it here in its entirety. The first step is to conjugate the doubling map T to a symbolic shift via the binary expansion. Let $\tilde{\Omega}_+ = \{0, 1\}^{\mathbb{Z}_+}$ and define $D : \tilde{\Omega}_+ \rightarrow \mathbb{T}$ by $D(\omega) = \sum_{n=1}^{\infty} \omega_n 2^{-n}$. The shift transformation, $S : \tilde{\Omega}_+ \rightarrow \tilde{\Omega}_+$, is given by $(S\tilde{\omega})_n = \tilde{\omega}_{n+1}$. Clearly, $D \circ S = T \circ D$.

Next we introduce a family of whole-line operators as follows. Let $\tilde{\Omega} = \{0, 1\}^{\mathbb{Z}}$ and define, for $\tilde{\omega} \in \tilde{\Omega}$, the operator

$$[H_{\tilde{\omega}}\phi](n) = \phi(n + 1) + \phi(n - 1) + V_{\tilde{\omega}}(n)\phi(n)$$

in $\ell^2(\mathbb{Z})$, where

$$V_{\tilde{\omega}}(n) = f[D(\{\tilde{\omega}_n, \tilde{\omega}_{n+1}, \tilde{\omega}_{n+2}, \dots\})].$$

The family $\{H_{\tilde{\omega}}\}_{\tilde{\omega} \in \tilde{\Omega}}$ is non-deterministic since $V_{\tilde{\omega}}$ restricted to \mathbb{Z}_+ only depends on $\{\tilde{\omega}_n\}_{n \geq 1}$ and hence, by non-constancy of f , we cannot determine the values of $V_{\tilde{\omega}}(n)$ for $n \leq 0$ uniquely from the knowledge of $V_{\tilde{\omega}}(n)$ for $n \geq 1$. It follows from Corollary 2 that the Lyapunov exponent for $\{H_{\tilde{\omega}}\}_{\tilde{\omega} \in \tilde{\Omega}}$ is almost everywhere positive and $\sigma_{ac}(H_{\tilde{\omega}})$ is empty for almost every $\tilde{\omega} \in \tilde{\Omega}$ with respect to the $(\frac{1}{2}, \frac{1}{2})$ -Bernoulli measure on $\tilde{\Omega}$.

Finally, let us consider the restrictions of $H_{\tilde{\omega}}$ to $\ell^2(\mathbb{Z}_+)$, that is, let $H_{\tilde{\omega}}^+ = E^*H_{\tilde{\omega}}E$, where $E : \ell^2(\mathbb{Z}_+) \rightarrow \ell^2(\mathbb{Z})$ is the natural embedding. Observe that $H_{\tilde{\omega}}^+ = H_{\omega}$, where $\omega = D(\{\tilde{\omega}_1, \tilde{\omega}_2, \tilde{\omega}_2, \dots\})$. This immediately implies the statement on the positivity of the Lyapunov exponent for the family $\{H_{\omega}\}_{\omega \in \mathbb{T}}$. As finite-rank perturbations preserve absolutely continuous spectrum, $\sigma_{ac}(H_{\tilde{\omega}}^+) \subseteq \sigma_{ac}(H_{\tilde{\omega}})$ for every $\tilde{\omega} \in \tilde{\Omega}$. This proves that $\sigma_{ac}(H_{\tilde{\omega}}^+) = \emptyset$ for almost every $\tilde{\omega} \in \tilde{\Omega}$. \square

Given $\phi \in (-\pi/2, \pi/2)$, let $H_{\omega}^{(\phi)}$ denote the operator which acts on $\ell^2(\mathbb{Z}_+)$ as in (1), but with $\psi(0)$ given by $\cos(\phi)\psi(0) + \sin(\phi)\psi(1) = 0$. Thus, the original operator (with a Dirichlet boundary condition) corresponds to $\phi = 0$. Theorem 8 implies the following result for this family of operators:

COROLLARY 4. *Suppose that f is measurable, bounded, and non-constant. Then, for almost every $\phi \in (-\pi/2, \pi/2)$ and almost every $\omega \in \mathbb{T}$, the operator $H_{\omega}^{(\phi)}$ in $\ell^2(\mathbb{Z}_+)$ with potential $V_{\omega}(n) = f(2^n\omega)$ has pure point spectrum and all eigenfunctions decay exponentially at infinity.*

PROOF. This is standard and follows quickly from spectral averaging; see, for example, [57, Theorem 13.4] or [61, Section 12.3]. \square

For a localization proof without the need for spectral averaging, it will be necessary to prove the positivity of the Lyapunov exponent for a larger set of energies. Sufficient, for example, is positivity away from a discrete set of exceptional energies. For moderately small coupling, such a result will be contained in [4]. The problem for other values of the coupling constant is still open.

PROBLEM 4. *Find a class of functions $f \in L^\infty(\mathbb{T})$ such that for every $\lambda \neq 0$, the Lyapunov exponent associated with the potentials $V_{\omega}(n) = \lambda f(2^n\omega)$ is positive away from a (λ -dependent) discrete set of energies.*

In connection with this problem, important obstructions have been found by Bochi [10] and Bochi and Viana [11]. Namely, positivity of $\gamma(E)$ away from a discrete set will fail generically in $C(\mathbb{T})$ (this result holds for rather general underlying dynamics) and hence the Hölder continuity assumptions made in [4] and [22] are natural.

In some sense a large value of λ alone should ensure the positivity of the Lyapunov exponent. This is indeed the basis of several results for quasi-periodic potentials or potentials generated by the skew shift. For hyperbolic base transformations such as the doubling map, however, there is a competition between two different

kinds of hyperbolic behavior that presents problems that have not been solved yet. Moreover, in the large coupling regime, it would be especially interesting to prove uniform (in energy) lower bounds on the Lyapunov exponents along with the natural log λ asymptotics.

PROBLEM 5. *Find a class of functions $f \in L^\infty(\mathbb{T})$ such that for every $\lambda \geq \lambda_0(g)$, the Lyapunov exponent associated with the potentials $V_\omega(n) = \lambda f(2^n \omega)$ obeys $\inf_E \gamma(E) \geq c \log \lambda$ for some suitable positive constant c .*

In this context it should be noted that Herman’s subharmonicity proof for trigonometric polynomials over ergodic shifts on the torus [36], works in the case of the doubling map.⁸ It seems much less clear, however, how to carry over the Sorets–Spencer proof for real-analytic f [62] from the case of irrational rotations of the circle to the case of the doubling map, let alone the proof of Bourgain for real-analytic functions over ergodic shifts on higher-dimensional tori [14].

For $SL(2, \mathbb{R})$ cocycles over the doubling map that are not of Schrödinger form (i.e., with a general $SL(2, \mathbb{R})$ matrix replacing (3)), Young has developed a method for proving positive Lyapunov exponents “at large coupling” that works in the C^1 category [64]. Her method does not immediately apply to Schrödinger cocycles, but it would be interesting to find a suitable extension.

PROBLEM 6. *Modify Young’s method and apply it to Schrödinger cocycles.*

5. Absence of AC Spectrum for Rough Potentials

It is in some way surprising that a rough sampling function f can make the resulting potentials non-deterministic. Traditionally, non-determinism had been thought of as a feature induced by the underlying dynamics. In particular, quasi-periodic potentials (in the generalized sense considered in this paper, which allows discontinuous f ’s) had for a long time been considered deterministic. The situation changed with an important observation by Kotani in his short 1989 paper [46].

He proved the following very general result:

THEOREM 9 (Kotani 1989). *Suppose that (Ω, T, μ) is ergodic, $f : \Omega \rightarrow \mathbb{R}$ takes finitely many values, and the resulting potentials V_ω are μ -almost surely not periodic. Then, $\text{Leb}(\mathcal{Z}) = 0$ and therefore $\Sigma_{ac} = \emptyset$.*

In particular, operators with quasi-periodic potentials of the form

$$V_\omega(n) = \lambda \sum_{m=1}^N \gamma_m \chi_{[a_{m-1}, a_m)}(n\alpha + \omega),$$

where $0 = a_0 < a_1 < \dots < a_{N-1} < a_N = 1$, $\gamma_1, \dots, \gamma_N \in \mathbb{R}$ (taking at least two values) and $\lambda \neq 0$, have no absolutely continuous spectrum. Note that this result holds for all non-zero couplings and hence it is particularly surprising for small values of λ . We will have more to say about these potentials in the next section.

The family $\{V_\omega\}_{\omega \in \mathbb{T}}$ does not seem to be non-deterministic in an intuitive sense as ω is uniquely determined by the sequence $V_\omega|_{\mathbb{Z}_-}$. However, the family becomes non-deterministic when we pass to the closed topological support of the induced measure on $\{\lambda\gamma_1, \dots, \lambda\gamma_N\}^{\mathbb{Z}}$. Then we will indeed find two distinct sequences

⁸The author is grateful to Kristian Bjerklöv for pointing this out.

that belong to the support of the induced measure, whose restrictions to \mathbb{Z}_- coincide. This fact will become more transparent when we discuss Theorem 10 below. Kotani’s proof proceeded in a slightly different way: he proved that there can be no *continuous* mapping from the left half-line to the right half-line.

PROOF OF THEOREM 9. As above we denote the push-forward of μ under the map $\omega \mapsto V_\omega$ by ν . It suffices to show that $\text{Leb}(\mathcal{Z}) > 0$ implies that $\text{supp } \nu$ is finite since then all elements of $\text{supp } \nu$ are periodic.

By the continuity of the maps E_\pm from (30) and the fact that $\text{Ran } f$ is finite, there is a finite M such that the knowledge of $V(-M), \dots, V(-1)$ determines $V(0)$ uniquely for $V \in \text{supp } \nu$. Now shift and iterate! It follows that $V(-M), \dots, V(-1)$ completely determine $\{V(n) : n \geq 0\}$ and hence $\text{supp } \nu$ has cardinality at most $(\# \text{Ran } f)^M$. □

Consider the case where Ω is a compact metric space, T is a homeomorphism, and μ is an ergodic Borel probability measure. This covers most, if not all, applications of interest. In this scenario, Damanik and Killip realized in [27] that the finite range of f is not essential. What is important, however, is that f is discontinuous at some point $\omega_0 \in \Omega$. One can then use this point of discontinuity to actually “construct” two elements of $\text{supp } \nu$ that coincide on a half-line.

We say that $l \in \mathbb{R}$ is an *essential limit* of f at ω_0 if there exists a sequence $\{\Omega_k\}$ of sets each of positive measure such that for any sequence $\{\omega_k\}$ with $\omega_k \in \Omega_k$, both $\omega_k \rightarrow \omega_0$ and $f(\omega_k) \rightarrow l$. If f has more than one essential limit at ω_0 , we say that f is *essentially discontinuous* at this point.

THEOREM 10 (Damanik–Killip 2005). *Suppose Ω is a compact metric space, $T : \Omega \rightarrow \Omega$ a homeomorphism, and μ an ergodic Borel probability measure. If there is an $\omega_0 \in \Omega$ such that f is essentially discontinuous at ω_0 but continuous at all points $T^n\omega_0$, $n < 0$, then $\Sigma_{\text{ac}} = \emptyset$.*

PROOF. We again denote the induced measure by ν . For each essential limit l of f at ω_0 , we will find $V_l \in \text{supp } \nu$ with $V_l(0) = l$. By assumption and construction, $V_l(n)$ is independent of l for every $n < 0$. This shows $\text{Leb}(\mathcal{Z}) = 0$ and hence $\Sigma_{\text{ac}} = \emptyset$.

Let l be an essential limit of f at ω_0 and let $\{\Omega_k\}$ be a sequence of sets which exhibits the fact that l is an essential limit of f . Since each has positive μ -measure, we can find points $\omega_k \in \Omega_k$ so that V_{ω_k} is in $\text{supp } \nu$; indeed, this is the case for almost every point in Ω_k .

As $\omega_k \rightarrow \omega_0$ and f is continuous at each of the points $T^n\omega_0$, $n < 0$, it follows that $V_{\omega_k}(n) \rightarrow V_{\omega_0}(n)$ for each $n < 0$. Moreover, since $f(\omega_k)$ converges to l , we also have $V_{\omega_k}(0) \rightarrow l$. We can guarantee convergence of $V_{\omega_k}(n)$ for $n > 0$ by passing to a subsequence because $\mathbb{R}^{\mathbb{Z}}$ is compact. Let us denote this limit potential by V_l . As each V_{ω_k} lies in $\text{supp } \nu$, so does V_l ; moreover, $V(0) = l$ and $V_l(n) = V_{\omega_0}(n)$ for each $n < 0$. □

Here is an illustration of this result and a strengthening of the derived consequence:

COROLLARY 5. *Suppose $\Omega = \mathbb{T}$, μ is normalized Lebesgue measure, and $T\omega = \omega + \alpha$ for some irrational α . If f has a single (non-removable) discontinuity at ω_0 , then for all $\omega \in [0, 1)$, the operator H_ω has no absolutely continuous spectrum.*

PROOF. Let us say that l is a limiting value of f at ω_0 if there is a sequence $\{\omega_k\}$ in $\mathbb{T} \setminus \{\omega_0\}$ such that $\omega_k \rightarrow \omega_0$ and $f(\omega_k) \rightarrow l$. As f has a non-removable discontinuity at ω_0 , it has more than one limiting value at this point. Moreover, since f is continuous away from ω_0 , any limiting value is also an essential limit since we can choose each Ω_k to be a suitably small interval around ω_k .

This shows that f has an essential discontinuity at ω_0 . As the orbit of ω_0 never returns to this point, f is continuous at each point $T^n\omega_0$, $n \neq 0$. Therefore, Theorem 10 is applicable and shows that H_ω has no absolutely continuous spectrum for Lebesgue-almost every $\omega \in [0, 1)$.

It remains to show that the absolutely continuous spectrum of H_ω is empty for all $\omega \in \mathbb{R}/\mathbb{Z}$. We begin by fixing ω_1 such that H_{ω_1} has no absolutely continuous spectrum and such that the orbit of ω_1 does not meet ω_0 ; almost all ω_1 have these properties.

Given an arbitrary $\omega \in \mathbb{R}/\mathbb{Z}$, we may choose a sequence of integers $\{n_i\}$ so that $T^{n_i}(\omega) \rightarrow \omega_1$. As f is continuous on the orbit of ω_1 , the potentials associated to $T^{n_i}(\omega)$ converge pointwise to V_{ω_1} . By a result of Last and Simon [51], the absolutely continuous spectrum cannot shrink under pointwise approximation using translates of a single potential. Thus, the operator with potential V_ω cannot have absolutely continuous spectrum. This concludes the proof. \square

These results are particularly interesting in connection with Problem 2. Quasi-periodic potentials (as defined in this paper) are almost periodic if and only if f is continuous. Thus, proving the absence of absolutely continuous spectrum for quasi-periodic potentials with discontinuous f 's is a way of providing further support for the conjecture that $\Sigma_{ac} \neq \emptyset$ implies almost periodicity.

Let us now turn to the case of continuous sampling functions f . The proof of Theorem 10 certainly breaks down and it is not clear where some sort of non-determinism should come from in the quasi-periodic case, for example. Of course, absence of absolutely continuous spectrum does not hold for a general continuous f . Thus, the following result from [2] is somewhat surprising:

THEOREM 11 (Avila–Damanik 2005). *Suppose Ω is a compact metric space, $T : \Omega \rightarrow \Omega$ a homeomorphism, and μ a non-atomic ergodic Borel probability measure. Then, there is a residual set of functions f in $C(\Omega)$ such that $\Sigma_{ac}(f) = \emptyset$.*

Recall that a subset of $C(\Omega)$ is called residual if it contains a countable intersection of dense open sets. A residual set is locally uncountable.

One would expect some absolutely continuous spectrum for weak perturbations with sufficiently nice potentials; especially in the one-frequency quasi-periodic case. More precisely, if f is nice enough, then $\Delta + \lambda f(n\alpha + \omega)$ should have some/purely absolutely continuous spectrum for $|\lambda|$ sufficiently small. It is known that real-analyticity is sufficiently “nice enough” [15] (when α is Diophantine), but it was expected that this assumption is much too strong and could possibly be replaced by mere continuity. The proof of Theorem 11 can easily be adapted to yield the following result, also contained in [2], which shows that continuity of the sampling function is not sufficient to ensure the existence of absolutely continuous spectrum for weakly coupled quasi-periodic potentials.

THEOREM 12 (Avila–Damanik 2005). *Suppose Ω is a compact metric space, $T : \Omega \rightarrow \Omega$ a homeomorphism, and μ a non-atomic ergodic Borel probability measure.*

Then, there is a residual set of functions f in $C(\Omega)$ such that $\Sigma_{ac}(\lambda f) = \emptyset$ for almost every $\lambda > 0$.

PROOF. We only sketch the proofs of Theorems 11 and 12. The key technical issue is to establish that the maps

$$(L^1(\Omega) \cap B_r(L^\infty(\Omega)), \|\cdot\|_1) \rightarrow \mathbb{R}, \quad f \mapsto \text{Leb}(\mathcal{Z}(f)) \tag{31}$$

and

$$(L^1(\Omega) \cap B_r(L^\infty(\Omega)), \|\cdot\|_1) \rightarrow \mathbb{R}, \quad f \mapsto \int_0^\Lambda \text{Leb}(\mathcal{Z}(\lambda f)) \, d\lambda \tag{32}$$

are upper semi-continuous. Here, $\Lambda > 0$, $B_r(L^\infty(\Omega)) = \{f \in L^\infty(\Omega) : \|f\|_\infty < r\}$, and $\mathcal{Z}(f)$ denotes the set of energies for which the Lyapunov exponent associated with (Ω, T, μ, f) vanishes.

Upper semi-continuity of the map (31) can be shown using the fact that γ is harmonic in the upper half-plane and subharmonic on the real line; see [2] for details. Fatou’s Lemma then implies upper semi-continuity of (32).

For $\delta > 0$, define

$$M_\delta = \{f \in C(\Omega) : \text{Leb}(\mathcal{Z}(f)) < \delta\}.$$

By the upper semi-continuity statement above, M_δ is open. By approximation with discontinuous functions and upper semi-continuity again, we see that M_δ is also dense.

It follows that

$$\{f \in C(\Omega) : \Sigma_{ac}(f) = \emptyset\} = \{f \in C(\Omega) : \text{Leb}(\mathcal{Z}(f)) = 0\} = \bigcap_{\delta > 0} M_\delta$$

is residual and Theorem 11 follows. Given upper semi-continuity of (32), the proof of Theorem 12 is analogous. □

While continuous functions can be approximated in the C^0 norm by discontinuous functions, this does not work in the C^ϵ norm for any $\epsilon > 0$. Thus, the proof just given does not extend to Hölder classes. It would be interesting to explore possible extensions of the results themselves; thus motivating the following problem.

PROBLEM 7. Prove or disprove statements like the ones in Theorems 11 and 12 for Hölder classes $C^\epsilon(\Omega)$, $\epsilon > 0$.

6. Uniform Lyapunov Exponents and Zero-Measure Spectrum

The Kotani result for potentials taking finitely many values, Theorem 9, is central to the study of one-dimensional quasi-crystal models. The main results in this area have been reviewed in [24, 25, 63]. In this section we will therefore focus on recent progress and discuss why zero-measure spectrum is a consequence of Kotani theory when there is uniform convergence to the Lyapunov exponent.

One-dimensional quasi-crystals are typically modelled by sequences over a finite alphabet which are aperiodic but which have very strong long-range order properties. An important class of examples is given by one-frequency quasi-periodic potentials with step functions as sampling functions. That is, the potentials are of the form

$$V_\omega(n) = \lambda \sum_{m=1}^N \gamma_m \chi_{[a_{m-1}, a_m)}(n\alpha + \omega), \tag{33}$$

where $0 = a_0 < a_1 < \dots < a_N = 1$ is a partition of the unit circle, $\lambda, \gamma_1, \dots, \gamma_N$ are real numbers, α is irrational, and $\omega \in \mathbb{T}$. We obtain aperiodic potentials if $\lambda \neq 0$ and $\{\gamma_1, \dots, \gamma_N\}$ has cardinality at least two. We will assume these conditions throughout this section.

One of the properties that has been established for many quasi-crystal models is zero-measure spectrum. By general principles, this implies that the spectrum is a Cantor set because it cannot contain isolated points. Given the Kotani result, $\text{Leb}(\mathcal{Z}) = 0$, the natural way of proving this is via the identity $\Sigma = \mathcal{Z}$.

THEOREM 13 (Damanik–Lenz 2006). *Suppose the potentials are of the form (33) and in addition all discontinuity points $\{a_m\} \subset \mathbb{T}$ are rational. Then, the Lyapunov exponent vanishes identically on the spectrum, that is, $\Sigma = \mathcal{Z}$. As a consequence, the spectrum is a Cantor set of zero Lebesgue measure.*

PROOF. We only sketch the main ideas. More details can be found in [28, 29]. Denote

$$\mathcal{UH} = \{E : \frac{1}{n} \log \|A_n^E(\omega)\| \rightarrow \gamma(E) > 0 \text{ uniformly in } \omega\}.$$

Then, by Lenz [52] (see also Johnson [40]), $\mathcal{UH} = \mathbb{C} \setminus \Sigma$. In particular,

$$\Sigma = \mathcal{Z} \cup \mathcal{NUH},$$

where

$$\mathcal{NUH} = \{E : \gamma(E) > 0 \text{ and } E \notin \mathcal{UH}\}.$$

The result follows once $\mathcal{NUH} = \emptyset$ is established. Thus, given E with $\gamma(E) > 0$, we need to show that $\frac{1}{n} \log \|A_n^E(\omega)\| \rightarrow \gamma(E)$ uniformly in ω .

Uniform convergence along a special subsequence, $n_k \rightarrow \infty$, can be shown using results from Boshernitzan [13] and Lenz [53]. Namely, the assumption that all a_m are rational implies that there is a sequence of integers $n_k \rightarrow \infty$ such that for each k , all words of length n_k that occur in the potentials V_ω do so with comparable frequencies. That is, there is a uniform $C > 0$ such that for every k ,

$$\min_{|\omega|=n_k, \omega \text{ occurs}} \liminf_{J \rightarrow \infty} \frac{1}{J} \#\{j : 1 \leq j \leq J, V_\omega(j) \dots V_\omega(j + n_k - 1) = w\} \geq \frac{C}{n_k} \quad (34)$$

uniformly in ω [13]. Using this result, one can then use ideas from [53] to show that $\frac{1}{n_k} \log \|A_{n_k}^E(\omega)\| \rightarrow \gamma(E)$ as $k \rightarrow \infty$, uniformly in ω .

Finally, the avalanche principle of Goldstein and Schlag [35] allows one to interpolate and prove the desired uniform convergence of $\frac{1}{n} \log \|A_n^E(\omega)\|$ to $\gamma(E)$ as $n \rightarrow \infty$. □

The rationality assumption in Theorem 13 holds on a dense set of parameters, which makes it suitable for an approximation of a continuous sampling function by a sequence of step functions. Some consequences that may be drawn from this can be found in Bjerklöv et al. [12]. On the other hand, the assumption is certainly not necessary and more general results than the one presented here can be found in [29]. It would be nice if the assumption could be removed altogether:

PROBLEM 8. *Prove zero-measure spectrum for all finite partitions of the circle, that is, remove the rationality assumption from Theorem 13.*

It should be mentioned, however, that the proof sketched above will not work in this generality. The approach is based on the Boshernitzan condition (34), and it was shown in [29] that this condition fails for certain parameter values.

From a mathematical point of view, the following problem is natural:

PROBLEM 9. *Study multi-frequency analogs. That is, for a finite partition $\mathbb{T}^d = J_1 \cup \dots \cup J_N$ and the operators with potential*

$$V_\omega(n) = \lambda \sum_{m=1}^N \gamma_m \chi_{J_m}(n\alpha + \omega),$$

what is the measure of the spectrum and what is the spectral type?

These potentials are not directly motivated by quasi-crystal theory but they form an interesting class that may again prove useful in the understanding of the phenomena that arise for continuous sampling functions. Moreover, very little is understood about the associated operators apart from the general Kotani result, which says that there is never any absolutely continuous spectrum. It is unclear, however, whether there can be any point spectrum, for example.

7. Purely AC Spectrum for the Subcritical AMO

In this final section, we describe the (to the best of our knowledge) first application of Corollary 1. In his 1997 paper, Kotani writes that at the time Kotani theory was developed, “it was not clear whether we could know the pure absolute continuity ... only from the IDS, but this corollary has answered this question affirmatively.”

The application we will present involves the almost Mathieu operator

$$[H_\omega \psi](n) = \psi(n + 1) + \psi(n - 1) + 2\lambda \cos(2\pi(n\alpha + \omega))\psi(n),$$

that is, the Schrödinger operator with one-frequency quasi-periodic potential associated with the sampling function $f(\omega) = 2\lambda \cos(2\pi\omega)$. This operator is known to exhibit a metal-insulator transition at $|\lambda| = 1$, that is, for almost every α , the almost sure spectral type is purely absolutely continuous for $|\lambda| < 1$, purely singular continuous for $|\lambda| = 1$, and pure point (with exponentially decaying eigenfunctions) for $|\lambda| > 1$. See Jitomirskaya [38] for this result and its history. For $|\lambda| > 1$, one indeed has to exclude a zero-measure set of frequencies α since it was shown by Avron and Simon, using Gordon’s lemma, that for Liouville α , there are no eigenvalues [8, 9]. For $|\lambda| \leq 1$, it has long been expected that the results above extend to all irrational frequencies. For $|\lambda| = 1$, the issue was resolved by Avila and Krikorian [6]. The question of what happens for $|\lambda| < 1$ was addressed by Problem 6 in [60].

Since Bourgain and Jitomirskaya showed in [16] that the Lyapunov exponent associated with the almost Mathieu operator obeys $\gamma(E) = \max\{0, \log|\lambda|\}$ for every $E \in \Sigma$ for all irrational frequencies α , the problem reduces to a study of the integrated density of states, that is, to a proof of its absolute continuity. The following result was shown in [3] and it completely settles this regularity issue for the integrated density of states.

THEOREM 14 (Avila–Damanik 2006). *The integrated density of states of the almost Mathieu operator is absolutely continuous if and only if $|\lambda| \neq 1$.*

Combining this result with the one from [16] just quoted, along with Corollary 1, we obtain almost surely purely absolutely continuous spectrum for the subcritical (i.e., $|\lambda| < 1$) almost Mathieu operator:

COROLLARY 6. *If $|\lambda| < 1$, then $\Sigma_{\text{sing}} = \emptyset$.*

Subsequently, Avila even extended this result and proved purely absolutely continuous spectrum for $|\lambda| < 1$, α irrational, and every ω [1].

PROBLEM 10. *Extend the results of this section to more general $f \in L^\infty(\mathbb{T})$; for example, real-analytic f 's.*

It was shown by Bourgain and Jitomirskaya that for Diophantine frequency α and analytic f , $\Sigma_{\text{sing}}(\lambda f) = \emptyset$ for λ sufficiently small [15]. An extension of this result to all ω is contained in [5]. The proofs probably break down for Liouville frequencies (for this, one needs to extend the method based on Gordon's lemma to matrices that are not banded, but which do have exponential off-diagonal decay). Thus, it seems natural to attack the problem for Liouville frequencies in the same way as above.

References

- [1] A. Avila, Absolutely continuous spectrum for the almost Mathieu operator with subcritical coupling, in preparation
- [2] A. Avila and D. Damanik, Generic singular spectrum for ergodic Schrödinger operators, *Duke Math. J.* **130** (2005), 393–400
- [3] A. Avila and D. Damanik, Absolute continuity of the integrated density of states for the almost Mathieu operator with non-critical coupling, preprint (2006)
- [4] A. Avila and D. Damanik, in preparation
- [5] A. Avila and S. Jitomirskaya, Almost localization and almost reducibility, in preparation
- [6] A. Avila and R. Krikorian, Reducibility or non-uniform hyperbolicity for quasiperiodic Schrödinger cocycles, to appear in *Ann. of Math.*
- [7] J. Avron and B. Simon, Almost periodic Schrödinger operators. I. Limit periodic potentials, *Commun. Math. Phys.* **82** (1981), 101–120
- [8] J. Avron and B. Simon, Singular continuous spectrum for a class of almost periodic Jacobi matrices, *Bull. Amer. Math. Soc.* **6** (1982), 81–85
- [9] J. Avron and B. Simon, Almost periodic Schrödinger operators. II. The integrated density of states, *Duke Math. J.* **50** (1983), 369–391
- [10] J. Bochi, Genericity of zero Lyapunov exponents, *Ergod. Th. & Dynam. Sys.* **22** (2002), 1667–1696
- [11] J. Bochi and M. Viana, The Lyapunov exponents of generic volume-preserving and symplectic maps, *Ann. of Math.* **161** (2005), 1–63
- [12] K. Bjerklov, D. Damanik, and R. Johnson, Lyapunov exponents of continuous Schrödinger cocycles over irrational rotations, to appear in *Ann. Mat. Pura Appl.*
- [13] M. Boshernitzan, Rank two interval exchange transformations, *Ergod. Th. & Dynam. Sys.* **8** (1988), 379–394
- [14] J. Bourgain, Positivity and continuity of the Lyapunov exponent for shifts on \mathbb{T}^d with arbitrary frequency vector and real analytic potential, *J. Anal. Math.* **96** (2005), 313–355
- [15] J. Bourgain and S. Jitomirskaya, Absolutely continuous spectrum for 1D quasiperiodic operators, *Invent. Math.* **148** (2002), 453–463
- [16] J. Bourgain and S. Jitomirskaya, Continuity of the Lyapunov exponent for quasiperiodic operators with analytic potential, *J. Statist. Phys.* **108** (2002), 1203–1218
- [17] J. Bourgain and W. Schlag, Anderson localization for Schrödinger operators on \mathbb{Z} with strongly mixing potentials, *Commun. Math. Phys.* **215** (2000), 143–175
- [18] D. Buschmann, A proof of the Ishii–Pastur theorem by the method of subordinacy, *Univ. Japel. Acta Math.* **34** (1997), 29–34
- [19] A. Casher and J. Lebowitz, Heat flow in regular and disordered harmonic chains, *J. Math. Phys.* **12** (1971), 1701–1711
- [20] V. Chulaevsky, An inverse spectral problem for limit-periodic Schrödinger operators, *Functional Anal. Appl.* **18** (1984), 220–233
- [21] V. Chulaevsky and S. Molchanov, The structure of a spectrum of the lacunary-limit-periodic Schrödinger operator, *Functional Anal. Appl.* **18** (1984), 343–344

- [22] V. Chulaevsky and T. Spencer, Positive Lyapunov exponents for a class of deterministic potentials, *Commun. Math. Phys.* **168** (1995), 455–466
- [23] H. Cycon, R. Froese, W. Kirsch, and B. Simon, *Schrödinger Operators with Application to Quantum Mechanics and Global Geometry*, Texts and Monographs in Physics, Springer-Verlag, Berlin, 1987
- [24] D. Damanik, Gordon-type arguments in the spectral theory of one-dimensional quasicrystals, *Directions in Mathematical Quasicrystals*, pp. 277–305, CRM Monograph Series **13**, American Mathematical Society, Providence, RI, 2000
- [25] D. Damanik, Strictly ergodic subshifts and associated operators, in this Festschrift
- [26] D. Damanik and R. Killip, Almost everywhere positivity of the Lyapunov exponent for the doubling map, *Commun. Math. Phys.* **257** (2005), 287–290
- [27] D. Damanik and R. Killip, Ergodic potentials with a discontinuous sampling function are non-deterministic, *Math. Res. Lett.* **12** (2005), 187–192
- [28] D. Damanik and D. Lenz, A criterion of Boshernitzan and uniform convergence in the multiplicative ergodic theorem, *Duke Math. J.* **133** (2006), 95–123
- [29] D. Damanik and D. Lenz, Zero-measure Cantor spectrum for Schrödinger operators with low-complexity potentials, *J. Math. Pures Appl.* **85** (2006), 671–686
- [30] P. Deift and B. Simon, Almost periodic Schrödinger operators. III. The absolutely continuous spectrum in one dimension, *Commun. Math. Phys.* **90** (1983), 389–411
- [31] J. Geronimo, Polynomials orthogonal on the unit circle with random recurrence coefficients, *Methods of Approximation Theory in Complex Analysis and Mathematical Physics* (Leningrad, 1991), pp. 43–61, Lecture Notes in Math., **1550**, Springer, Berlin, 1993
- [32] J. Geronimo and A. Teplyaev, A difference equation arising from the trigonometric moment problem having random reflection coefficients—an operator-theoretic approach, *J. Funct. Anal.* **123** (1994), 12–45
- [33] D. Gilbert, On subordinacy and analysis of the spectrum of Schrödinger operators with two singular endpoints, *Proc. Roy. Soc. Edinburgh A* **112** (1989), 213–229
- [34] D. Gilbert and D. Pearson, On subordinacy and analysis of the spectrum of one-dimensional Schrödinger operators, *J. Math. Anal. Appl.* **128** (1987), 30–56
- [35] M. Goldstein and W. Schlag, Hölder continuity of the integrated density of states for quasi-periodic Schrödinger equations and averages of shifts of subharmonic functions, *Ann. of Math.* **154** (2001), 155–203
- [36] M. Herman, Une méthode pour minorer les exposants de Lyapunov et quelques exemples montrant caractère local d’un théorème d’Arnold et de Moser sur le tore de dimension 2, *Comment. Math. Helv* **58** (1983), 453–502
- [37] K. Ishii, Localization of eigenstates and transport phenomena in the one dimensional disordered system, *Supp. Theor. Phys.* **53** (1973), 77–138
- [38] S. Jitomirskaya, Metal-insulator transition for the almost Mathieu operator, *Ann. of Math.* **150** (1999), 1159–1175
- [39] S. Jitomirskaya, Ergodic Schrödinger operators (on one foot), in this Festschrift
- [40] R. Johnson, Exponential dichotomy, rotation number, and linear differential operators with bounded coefficients, *J. Differential Equations* **61** (1986), 54–78
- [41] W. Kirsch, On a class of random Schrödinger operators, *Adv. in Appl. Math.* **6** (1985), 177–187
- [42] W. Kirsch, S. Kotani, and B. Simon, Absence of absolutely continuous spectrum for some one-dimensional random but deterministic Schrödinger operators, *Ann. Inst. H. Poincaré Phys. Théor.* **42** (1985), 383–406
- [43] S. Kotani, Ljapunov indices determine absolutely continuous spectra of stationary random one-dimensional Schrödinger operators, *Stochastic Analysis* (Katata/Kyoto, 1982), pp. 225–247, North-Holland Math. Library **32**, North-Holland, Amsterdam, 1984
- [44] S. Kotani, Support theorems for random Schrödinger operators, *Commun. Math. Phys.* **97** (1985), 443–452
- [45] S. Kotani, One-dimensional random Schrödinger operators and Herglotz functions, *Probabilistic methods in mathematical physics* (Katata/Kyoto, 1985), pp. 219–250, Academic Press, Boston, 1987
- [46] S. Kotani, Jacobi matrices with random potentials taking finitely many values, *Rev. Math. Phys.* **1** (1989), 129–133

- [47] S. Kotani, Generalized Floquet theory for stationary Schrödinger operators in one dimension, *Chaos Solitons Fractals* **8** (1997), 1817–1854
- [48] S. Kotani and M. Krishna, Almost periodicity of some random potentials, *J. Funct. Anal.* **78** (1988), 390–405
- [49] S. Kotani and B. Simon, Stochastic Schrödinger operators and Jacobi matrices on the strip, *Commun. Math. Phys.* **119** (1988), 403–429
- [50] H. Kunz and B. Souillard, Sur le spectre des opérateurs aux différences finies aléatoires, *Commun. Math. Phys.* **78** (1980/81), 201–246
- [51] Y. Last and B. Simon, Eigenfunctions, transfer matrices, and absolutely continuous spectrum of one-dimensional Schrödinger operators, *Invent. Math.* **135** (1999), 329–367
- [52] D. Lenz, Singular continuous spectrum of Lebesgue measure zero for one-dimensional quasicrystals, *Commun. Math. Phys.* **227** (2002), 119–130
- [53] D. Lenz, Uniform ergodic theorems on subshifts over a finite alphabet, *Ergod. Th. & Dynam. Sys.* **22** (2002), 245–255
- [54] N. Minami, An extension of Kotani’s theorem to random generalized Sturm–Liouville operators, *Commun. Math. Phys.* **103** (1986), 387–402
- [55] V. Oseledec, A multiplicative ergodic theorem. Characteristic Ljapunov exponents of dynamical systems, (Russian) *Trudy Moskov. Mat. Obsč.* **19** (1968), 179–210; translation in *Trans. Moscow Math. Soc.* **19** (1968), 197–231
- [56] L. Pastur, Spectral properties of disordered systems in the one-body approximation, *Commun. Math. Phys.* **75** (1980), 179–196
- [57] L. Pastur and A. Figotin, *Spectra of Random and Almost-Periodic Operators*, Grundlehren der Mathematischen Wissenschaften **297**, Springer-Verlag, Berlin, 1992
- [58] L. Pastur and V. Tkachenko, Spectral theory of a class of one-dimensional Schrödinger operators with limit-periodic potentials, (Russian) *Trudy Moskov. Mat. Obsč.* **51** (1988), 114–168, 258; translation in *Trans. Moscow Math. Soc.* (1989), 115–166
- [59] B. Simon, Kotani theory for one-dimensional stochastic Jacobi matrices, *Commun. Math. Phys.* **89** (1983), 227–234
- [60] B. Simon, Schrödinger operators in the twenty-first century, *Mathematical Physics 2000*, 283–288, Imperial College Press, London, 2000
- [61] B. Simon, *Trace Ideals and Their Applications*, second edition, Mathematical Surveys and Monographs **120**, American Mathematical Society, Providence, RI, 2005
- [62] E. Sorets and T. Spencer, Positive Lyapunov exponents for Schrödinger operators with quasi-periodic potentials, *Commun. Math. Phys.* **142** (1991), 543–566
- [63] A. Sütő, Schrödinger difference equation with deterministic ergodic potentials, *Beyond Quasicrystals* (Les Houches, 1994), pp. 481–549, Springer, Berlin, 1995
- [64] L.-S. Young, Some open sets of nonuniformly hyperbolic cocycles, *Ergod. Th. & Dynam. Sys.* **13** (1993), 409–415

MATHEMATICS 253–37, CALIFORNIA INSTITUTE OF TECHNOLOGY, PASADENA, CA 91125, U.S.A.

E-mail address: damanik@caltech.edu

URL: <http://www.math.caltech.edu/people/damanik.html>

Spectral Properties of Schrödinger Operators with Decaying Potentials

Sergey A. Denisov and Alexander Kiselev

Dedicated to Barry Simon on the occasion of his 60th birthday

ABSTRACT. We review recent advances in the spectral theory of Schrödinger operators with decaying potentials. The area has seen spectacular progress in the past few years, stimulated by several conjectures stated by Barry Simon starting at the 1994 International Congress on Mathematical Physics in Paris. The one-dimensional picture is now fairly complete, and provides many striking spectral examples. The multidimensional picture is still far from clear and may require deep original ideas for further progress. It might hold the keys for better understanding of a wide range of spectral and dynamical phenomena for Schrödinger operators in higher dimensions.

CONTENTS

1. Introduction
 2. The One-Dimensional Case
 3. Striking Examples
 4. Dirac Operators, Krein Systems, Jacobi Matrices, and OPUC
 5. The Multidimensional Case
 6. The Bethe Lattice
- References

1. Introduction

Schrödinger operators with decaying potentials are used to study the behavior of a charged particle in a local electric field. The operator is defined by

$$H_V = -\Delta + V(x) \tag{1.1}$$

on $L^2(\mathbb{R}^d)$; in one dimension it is common to consider the operator on a half-axis with some self-adjoint boundary condition at zero. The spectral and dynamical

2000 *Mathematics Subject Classification.* 35J10, 34L40.

Key words and phrases. Schrödinger operator, decaying potential, spectral analysis.

effects that we are interested in are those depending on the rate of decay of the potential rather than its singularities, so we will often freely assume that V is bounded. If the decay of the potential is sufficiently fast (short range), one expects scattering motion. The corresponding results have been rigorously proved by Weidmann [76] in one dimension (where short-range case means that the potential is in $L^1(\mathbb{R})$) and by Agmon [1] in higher dimensions (where the natural short range class is defined by $|V(x)| \leq C(1 + |x|)^{-1-\epsilon}$). These results established pure absolute continuity of the spectrum on the positive semi-axis and asymptotic completeness of the wave operators. There has been a significant amount of work on longer range potentials with additional symbol-like conditions (see e.g. [60] for further references), or oscillating potentials of specific Wigner–von Neumann type structure (see e.g. [3, 24, 78] for further references). However, until the 1990s, there had been very limited progress on understanding slowly decaying potentials with no additional assumptions on behavior of derivatives. The short range or classical WKB methods did not seem to apply in this case, and the possible spectral properties remained a mystery. The celebrated Wigner–von Neumann example [77] provides a potential $V(x)$ behaving like $8 \sin(2x)/x + O(x^{-2})$ as $x \rightarrow \infty$ and leading to an imbedded eigenvalue $E = 1$, thus showing that surprising things can happen once the potential is not short range. On the other hand, the work of Kotani and Ushiroya [42] implied that for potentials decaying at power rate slower than $x^{-\alpha}$, $\alpha \leq 1/2$, the spectrum may become purely singular, and thus the scattering picture may be completely destroyed. There was a clear gap in the decay rates where very little information on the possible spectral properties was available. In recent years, there has been significant progress in the area, largely stimulated by Barry Simon’s research and ideas. At the ICMP in Paris in 1994, Simon posed a problem of understanding the spectral properties of Schrödinger operators with potentials satisfying $|V(x)| \leq C(1 + |x|)^{-\alpha}$, $1 > \alpha > 1/2$. Later, at the 2000 ICMP in London [71], he compiled a list of fifteen problems in Schrödinger operators “for the twenty-first century”. Two of the problems on the list concern long range potentials.

While there remain many open questions, the recent effort to improve understanding of the long range potentials led to many high quality mathematical works. Fruitful new links between the spectral theory of Schrödinger operators and orthogonal polynomials as well as Fourier analysis have been discovered and exploited. Surprising examples of intricate spectral properties have been produced. Advances have been made towards a better understanding of effects possible in higher dimensions. In this review, we try to survey recent results in this vital area, as well as underline most active current directions and questions of interest. In the second section we discuss the one-dimensional case, where the picture is much more detailed and complete. The third section is devoted to a number of interesting spectral and dynamical examples, typically one-dimensional, but easily extendable to any dimension by spherically symmetric construction. We briefly mention certain relations to Dirac operators, Jacobi matrices and polynomials orthogonal on the unit circle (OPUC) in the fourth section. In the last section, we consider the higher dimensional case, where the main question—known as Barry Simon’s conjecture—is still open and is at the focus of current research.

Most of this work is a compilation and review of known results. There are, however, three nuggets that are new to the best of our knowledge. In Section 3, we

provide a new proof of Theorem 3.1, the construction of an example with a dense set of imbedded eigenvalues. In Section 5, Theorem 5.2 and Theorem 6.2 are new.

2. The One-Dimensional Case

The initial progress in understanding slowly decaying potentials started from particular cases, such as random and sparse. Both of these classes have been treated in a single framework by Simon in a joint paper with Last and Kiselev [40]; random decaying potentials in the discrete setting had been pioneered by Simon in a joint work with Delyon and Souillard [14]. Let H_V be a half-line Schrödinger operator, and fix some boundary condition at the origin. Let us call $V(x)$ a Pearson potential if $V(x) = \sum_n a_n W(x - x_n)$, where $a_n \rightarrow 0$, $x_n/x_{n-1} \rightarrow \infty$ as $n \rightarrow \infty$, and $W(x) \in C_0^\infty(\mathbb{R})$.

THEOREM 2.1. *Let $V(x)$ be a Pearson potential. If $\sum_n a_n^2 < \infty$, the spectrum of H_V on $(0, \infty)$ is purely absolutely continuous. If $\sum_n a_n^2 = \infty$, the spectrum of H_V on $(0, \infty)$ is purely singular continuous.*

This result [40] generalizes the original work of Pearson [58], who essentially proved Theorem 2.1 under the assumption that x_n grow sufficiently fast (with no explicit estimate). See also [61] for related results. In a sense, Pearson's theorem was the first indication of a clear spectral transition at $p = 2$ when the potential is viewed in L^p scale. A similar picture is true for the random potentials. Let $V(x) = n^{-\alpha} a_n(\omega) W(x - n)$, where $W \in C_0^\infty(0, 1)$ and $a_n(\omega)$ are random i.i.d. variables with mean zero and compactly supported probability density function.

THEOREM 2.2. *If $\alpha > 1/2$, then the spectrum of H_V on the positive half-axis is purely absolutely continuous with probability one. If $\alpha < 1/2$, the spectrum on $(0, \infty)$ is pure point with probability one. If $\alpha = 1/2$, the spectrum is a mixture of pure point and singular continuous spectrum with probability one.*

See [40] for more details in the $\alpha = 1/2$ case, as well as for the proof of a more general theorem. The first result of the type of Theorem 2.2 is due to Delyon, Simon and Souillard [14], who handled the discrete case. In the continuous setting, Kotani and Ushiroya [42] proved a version of Theorem 2.2 for a slightly different model.

Theorems 2.1 and 2.2 show a transition which is reminiscent of some classical results on almost everywhere convergence and divergence of Fourier series. Random Fourier series also converge or diverge at almost every point with probability one depending on whether the coefficients are square summable (see, e.g. [32]). A similar result holds for the lacunary Fourier series (see, e.g. [81] for further references). This analogy is not accidental. Indeed, the spectral properties are related to the behavior of solutions of the Schrödinger equation. Although the precise link between solutions and local (in energy) properties of spectral measure is given by the subordinacy condition discovered by Gilbert and Pearson [26], boundedness of the solutions is typically associated with the absolutely continuous spectrum. In particular, it has been shown by Behncke [4], Stolz [74] and Simon [70] that if all solutions of the equation $-u'' + V(x)u = Eu$ are bounded for each E in a set S of positive Lebesgue measure, then the absolutely continuous part of the spectral measure gives positive weight to S , and the singular part of the spectral measure does not give any weight to S . Establishing the boundedness of the solutions, on

the other hand, can be thought of as a nonlinear analog of proving the convergence of Fourier series, at least for the potentials in $L^2(\mathbb{R})$. To clarify this idea, it is convenient to introduce the generalized Prüfer transform, a very useful tool for studying the solutions in one dimension. We will very roughly sketch the idea behind Theorems 2.1 and 2.2 following [40].

Let $u(x, k)$ be a solution of the eigenfunction equation

$$-u'' + V(x)u = k^2u. \tag{2.1}$$

The modified Prüfer variables are introduced by

$$u'(x, k) = kR(x, k) \cos \theta(x, k); \quad u(x, k) = R(x, k) \sin \theta(x, k). \tag{2.2}$$

The variables $R(x, k)$ and $\theta(x, k)$ satisfy

$$(\log R(x, k)^2)' = \frac{1}{k}V(x) \sin 2\theta(x, k) \tag{2.3}$$

$$\theta(x, k)' = k - \frac{1}{k}V(x)(\sin \theta)^2. \tag{2.4}$$

Fix some point x_0 far enough, and set $\theta(x_0) = \theta_0, R(x_0) = R_0$. From (2.4), we see that

$$\theta(x) = \theta_0 + k(x - x_0) - \frac{V(x)}{k}(\sin(k(x - x_0) + \theta_0))^2 + O(V^2) = \theta_0 + k(x - x_0) + \delta\theta + O(V^2). \tag{2.5}$$

Then

$$\sin 2\theta(x, k) = \sin(2\theta_0 + 2k(x - x_0)) + 2 \cos(2\theta_0 + 2k(x - x_0)) \delta\theta + O(V^2).$$

From (2.5) and the equation (2.3) for the amplitude, we find

$$\frac{d}{dx} (\log(R^2(x))) = t_1 + t_2 + O(V^3), \tag{2.6}$$

where

$$t_1 = \frac{V(x)}{k} \sin(2(\theta_0 + k(x - x_0))) - \frac{V(x)}{2k^2} \left(\int_{x_0}^x V(y) dy \right) \cos(2(\theta_0 + k(x - x_0))),$$

and

$$t_2 = \frac{1}{4k^2} \frac{d}{dx} \left[\int_{x_0}^x V(y) \cos(2(\theta_0 + k(y - x_0))) dy \right]^2.$$

In both the random and sparse cases, we obtain the asymptotic behavior of $R(x)$ by summing up contributions from finite intervals. In the random case, these intervals correspond to the independent random parts of the potential, while in the sparse case, $R(x)$ remains unchanged between the neighboring bumps, and we only have to add the contributions of the bumps. In both cases, for different reasons, the contributions of the t_1 terms can be controlled and are finite (with probability one in the random case). For random potentials, one uses the linearity of terms entering t_1 in V and the independence of different contributions; the argument is then similar to the Fourier transform case and gives convergence as far as $V \in L^2$ by the Kolmogorov three series theorem. We note that for the second term in t_1 , one actually has to use a bit more subtle reasoning, also taking into account the oscillations in energy. In the sparse case, one uses the fact that contributions from different steps are oscillating in k with very different frequency due to the large

distance between x_n and x_{n+1} . Again, the argument is related to the techniques used to study the lacunary Fourier series. On the other hand, the sum of t_2 terms is finite if $V \in L^2$, leading to the boundedness of the solutions and absolutely continuous spectrum. If V is not L^2 , the sum of t_2 terms diverges and can be shown to dominate the other terms due to lack of sign cancellations.

The question of whether the absolutely continuous spectrum is preserved for general L^2 potentials remained open longer. The initial progress in this direction focused on proving boundedness of solutions for almost every energy. There are many examples, starting from the celebrated Wigner and von Neumann [77] construction of an imbedded eigenvalue, which show that spectrum does not have to be purely absolutely continuous if $V \notin L^1$, and imbedded singular spectrum may occur. We will discuss some of these examples in Section 3. Thus there can be exceptional energies with decaying and growing solutions. Again, one can think of a parallel with the Fourier transform, where the integral $\int_{-N}^N e^{ikx}g(x) dx$ may diverge for some energies if $g \in L^2$. It was conjectured by Luzin early in the twentieth century that nevertheless the integral converges for a.e. k . The question turned out to be difficult, and required an extremely subtle analysis by Carleson to be solved positively in 1955 in a famous paper [7]. If $g \in L^p$ with $p < 2$, the problem is significantly simpler, and had been solved by Zygmund in 1928 [80] (see also Menshov [51] and Paley [57] for the discrete case). As the equation (2.3) suggests, the problem of boundedness of solutions to the Schrödinger equation may be viewed as a question about a.e. convergence of a nonlinear Fourier transform. Research in this direction started from work of Christ, Kiselev, Molchanov and Remling on power decaying potentials [36, 37, 52, 8, 62]. An elegant and simple paper by Deift and Killip [15] used a completely different idea, sum rules, to prove the sharp result, the preservation of the absolutely continuous spectrum for L^2 potentials. The two approaches can be regarded as complementary: the study of solutions gives more precise information about the operator and dynamics, but has so far been unable to handle the borderline case $p = 2$. The sum-rule methods give the sharp result on the nature of the spectrum, but less information about the nature of the eigenfunctions and dynamical properties. We will briefly sketch the most current results in both areas, starting with the solutions approach.

Let H_V be the whole-line Schrödinger operator. Recall that the modified wave operators are defined by

$$\Omega_{m\pm}^\pm g = L^2 - \lim_{t \rightarrow \mp\infty} e^{iH_V t} e^{-iW(-i\partial_x, t)} g, \quad (2.7)$$

where the operator $e^{-iW(-i\partial_x, t)}$ acts as a multiplier on the Fourier transform of g . Let

$$W(k, t) = k^2 t + \frac{1}{2k} \int_0^{2kt} V(s) ds. \quad (2.8)$$

The following theorem was proved in [9].

THEOREM 2.3. *Assume that $V \in L^p$, $p < 2$. Then for a.e. k , there exist solutions $u_\pm(x, k)$ of the eigenfunction equation (2.1) such that*

$$u_\pm(x, k) = e^{ikx - \frac{i}{2k} \int_0^x V(y) dy} (1 + o(1)) \quad (2.9)$$

as $x \rightarrow \pm\infty$. Moreover, the modified wave operators (2.7) exist.

Assume, in addition, that $V(x)|x|^\gamma \in L^p$ for some $p < 2$ and $\gamma > 0$. Then the Hausdorff dimension of the set of k where (2.9) fails cannot exceed $1 - \gamma p'$ (where p' is the Hölder conjugate exponent to p)

The asymptotic behavior in (2.9) as well as the phase in (2.7) coincide with the WKB asymptotic behavior, which has been known for a long time for potentials satisfying additional conditions on the derivatives. The main novelty of (2.7) is that no such condition is imposed. Note that if the integral $\int_0^N V(s) ds$ converges, the asymptotic behavior of u_\pm becomes identical to the solutions of the unperturbed equation, and modified wave operators can be replaced by the usual Möller wave operators. The proof of Theorem 2.3 proceeds by deriving an explicit series representation for the solutions u_\pm via an iterative procedure. The terms in the series may diverge for some values of k , but converge almost everywhere. The first term in the series is a generalization of the Fourier transform, $\int_0^N \exp(ikx - \frac{i}{k} \int_0^x V(y) dy) V(x) dx$. The main difficulty in the proof comes from proving the estimates for the multilinear higher order terms such that the series can be summed up for a.e. k . The estimate (2.9) implies that all solutions of (2.1) are bounded for a.e. k if $V \in L^p$, $p < 2$, and can be thought of as a nonlinear version of Zygmund's result for the Fourier transform. The question of whether (2.9) holds and whether the modified operators exist for $V \in L^2$ is still open, and appears to be very hard, especially the a.e. boundedness of the eigenfunctions. Indeed, proving (2.9) would be the nonlinear analog of the Carleson theorem. Moreover, Muscalu, Tao and Thiele showed [55] that the method of [9] has no chance of succeeding when $p = 2$ (since some terms in the multilinear series expansion may diverge on a set of positive measure). The techniques behind Theorem 2.3 have been used to prove related results on slowly varying potentials (with derivatives in L^p , $p < 2$) and perturbations of Stark operators. See [10, 11] for more details.

The sum rules approach to proving absolute continuity of the spectrum was pioneered by Deift and Killip and led to an explosion of activity in the area and many impressive new results. Assume that $V(x) \in C_0^\infty(\mathbb{R})$. Let us consider the solution $f(x, k)$ of (2.1) such that $f(x, k) = \exp(ikx)$ when x is to the right of the support of V . Then, for x to the left of the support of V , we have $f(x, k) = a(k) \exp(ikx) + b(k) \exp(-ikx)$. The solution $f(x, k)$ is called the Jost solution, and $f(0, k)$ the Jost function. The coefficient $t(k) = a^{-1}(k)$ is the transmission coefficient in classical scattering theory. Denote by E_j the eigenvalues of the operator H_V . The following identity is well known (see e.g. [25]):

$$\int_{-\infty}^{\infty} (\log |a(k)|) k^2 dk + \frac{2\pi}{3} \sum_j |E_j|^{3/2} = \frac{\pi}{8} \int_{-\infty}^{\infty} V^2(x) dx. \tag{2.10}$$

The identity can be proved, for example, by a contour integration in the complex upper half-plane of an asymptotic expansion in k^{-1} of the integral equation one can write for $f(x, k)$. There is a whole hierarchy of formulas (sum rules) similar to (2.10). This fact is related to the role that the inverse scattering transform for Schrödinger operators plays in understanding the KdV dynamics. The expressions involving $V(x)$ which appear on the right hand side in such sum rules are the KdV invariants. The inequalities of type (2.10) have been applied in the past to derive bounds on the moments of the eigenvalues of H_V (Lieb–Thirring inequalities). Deift and Killip realized that the coefficient $a(k)$ is directly linked to the spectral measure

of H_V . Building a sequence of compact support approximations to $V(x)$, and then passing to the limit, one essentially derives a lower bound on the entropy of the absolutely continuous part of the spectral measure:

$$\int_I \log \mu'(\lambda) d\lambda > -\infty,$$

where $d\mu(\lambda)$ is the spectral measure and I is an arbitrary bounded subinterval of \mathbb{R}^+ . This proves

THEOREM 2.4. *For any $V \in L^2$, the essential support of the absolutely continuous spectrum of the operator H_V coincides with the half-axis $(0, \infty)$. That is, the absolutely continuous part of the spectral measure μ_{ac} gives positive weight to any set $S \subset (0, \infty)$ of the positive Lebesgue measure.*

Killip [33] later proved a strengthened version of Theorem 2.4, also applicable to potentials from L^3 given additional assumptions on the Fourier transform, and to Stark operators. The key advance in [33] is a local in energy version of (2.10), which is more flexible and useful in different situations. The important fact exploited in [33] is that the Jost function is actually the perturbation determinant of the Schrödinger operator. That yields a natural path to obtaining estimates necessary to control the boundary behavior of the Jost function. The square of the inverse of the Jost function, on the other hand, is proportional to the density of the spectral measure (see (5.5) for a similar higher dimensional relation). Therefore, the estimates on the Jost function have deep spectral consequences.

The results of [15] have been extended to slowly varying potentials with higher order derivative in L^2 by Molchanov, Novitski and Vainberg [53], using the higher order KdV invariants. Some improvements were made in [17] where the asymptotical methods for ODE were used.

In the discrete setting, the application of sum rules led Killip and Simon [34] to a beautiful result giving a complete description of the spectral measures of Jacobi matrices which are Hilbert–Schmidt perturbations of a free Jacobi matrix. Further extensions to slower decaying perturbations of Jacobi matrices and Schrödinger operators have been obtained in different works by Laptev, Naboko, Rybkin and Safronov, [45, 46, 66, 67]. Recently, Killip and Simon [35] proved a continuous version of their Jacobi matrix theorem, giving a precise description of spectral measures that can occur for Schrödinger operators with L^2 potentials. We will further discuss their result in the following section.

Certain extensions of the sum rules method have also been applied to higher dimensional problems, and will be discussed in Section 5.

We complete this section with a somewhat philosophical remark. The technique of the Deift–Killip proof (and its developments) has a certain air of magic about it. After all, it is based on an identity, sum rule (2.10), a rarity in analysis. Recall the classical von Neumann–Kuroda theorem, which says that given an arbitrary self-adjoint operator A , one can find an operator Y with arbitrary small Hilbert–Schmidt norm (or any Schatten–von Neumann class norm weaker than trace class) such that $A + Y$ has pure point spectrum. Theorem 2.4 says that the situation is very different if one restricts perturbations to potentials in the case of $A = H_0$. The result is so clear cut that one has to wonder if there is a general, operator theory type of result which, for a given A with absolutely continuous spectrum, describes

classes of perturbations which would be less efficient in diagonalizing it. Such a more general understanding could prove useful in other situations, but currently is completely missing.

3. Striking Examples

Apart from the general results described in the previous section, there are fairly explicit descriptions of decaying potentials leading to quite amazing spectral properties. The examples we discuss here deal with imbedded singular spectrum. Although all the examples we mention are constructed in one dimension, in most cases it is not difficult to extend them to an arbitrary dimension using spherically symmetric potentials. The grandfather of all such examples is a Wigner–von Neumann example of a potential which has oscillatory asymptotic behavior at infinity, $V(x) = 8 \sin 2x/x + O(x^{-2})$, and leads to an imbedded eigenvalue at $E = 1$. The imbedded singular spectrum for decaying potentials is the resonance phenomenon, and requires oscillation in the potential, similarly to the divergence of Fourier series or integrals. It is also inherently unstable—for example, for a.e. boundary condition in the half-line case there is no imbedded singular spectrum. The first examples we are going to discuss are due to Naboko [56] and Simon [69], who provided different constructions for potentials leading to a similar phenomenon.

THEOREM 3.1. *For any positive monotone increasing function $h(x) \rightarrow \infty$, there exist potentials satisfying $|V(x)| \leq \frac{h(x)}{1+|x|}$ such that the half-line operator H_V (with, say, Dirichlet boundary condition) has dense point spectrum in $(0, \infty)$.*

If $|V(x)| \leq \frac{C}{1+|x|}$, the eigenvalues E_1, \dots, E_n, \dots of H_V lying in $(0, \infty)$ must satisfy $\sum_n E_n < \infty$.

The last statement of Theorem 3.1 was proved in [40].

The construction of Naboko used the first order system representation of the Schrödinger equation, and had a restriction that the square roots of eigenvalues in $(0, \infty)$ had to be rationally independent. Simon's construction can be used to obtain any dense countable set of eigenvalues in $(0, \infty)$. The idea of the latter construction is, roughly, given a set of momenta k_1, \dots, k_n, \dots , take

$$V(x) = W(x) + \sum_n \chi_{(x_n, \infty)}(x) B_n \frac{\sin(2k_n x + \beta_n)}{x}.$$

Here x_n , B_n and β_n have to be chosen appropriately, and $W(x)$ is a compactly supported potential whose job is to make sure that the L^2 eigenfunctions at E_n satisfy the right boundary condition at zero. Thus, basically, the potential is a sum of resonant pieces on all frequencies where the eigenvalues are planned. To explain the argument better, we will outline a third construction of such an example, which in our view is technically the simplest one to implement. We will only sketch the proof; the details are left to the interested reader.

PROOF OF THEOREM 3.1. Without loss of generality, we assume that $h(x)$ does not grow too fast, say $|h(x)| \leq x^{1/4}$. Recall the Prüfer variables $R(x, k)$, $\theta(x, k)$ and equations (2.4), (2.3) they satisfy. For $x \leq x_1$, x_1 to be determined later, let

$$V(x) = -\frac{h(x)}{2(1+|x|)} \sin 2\theta(x, k_1). \quad (3.1)$$

Here k_1^2 is the first eigenvalue from the list we are trying to arrange. Note that the seeming conflict between defining V in terms of θ and θ in (2.4) in terms of V is resolved by plugging (3.1) into (2.4), solving the resulting nonlinear equation for $\theta(x, k_1)$, and defining $V(x)$ as in (3.1). Now if V is defined according to (3.1) on the whole half-axis, one can see, using (2.3), (2.4) and integration by parts, that

$$\log(R(x, k_1)^2) = - \int_0^x \frac{h(y)}{2k_1(1+|y|)} dy + O(1).$$

Because of our assumptions on $h(x)$, $R(x, k_1)$ is going to be square integrable.

Now we define our potential $V(x)$ by

$$V(x) = - \sum_{j=1}^{\infty} \frac{h(x)}{2^j(1+|x|)} \chi_{(x_n, \infty)}(x) \sin 2\theta(x, k_j). \quad (3.2)$$

Each $x_n > x_{n-1}$ is chosen inductively, so that the following condition is satisfied: for any $j < n$,

$$\sup_{x > x_n} \left| \int_{x_n}^x \frac{h(y)}{(1+|y|)} \sin 2\theta(y, k_n) \sin 2\theta(y, k_j) dy \right| \leq 1. \quad (3.3)$$

Using (2.4) and integration by parts, it is easy to see that on each step, the condition (3.3) will be satisfied for all sufficiently large x_n . A calculation using (2.4), (2.3), and integration by parts then shows that $R(x, k_n)$ is square integrable for each n . From (3.2) it also follows that $V(x) \leq h(x)/(1+|x|)$. \square

Examples with imbedded singular continuous spectrum are significantly harder to construct. The main difficulty is that while to establish point spectrum one just needs make sure that the L^2 norm of the solution is finite, it is not quite clear what one needs to control to prove the existence of the singular continuous component of the spectral measure. At the ICMP Congress in London, Simon [71] posed the problem of finding a decaying potential leading to imbedded singular continuous spectrum. The first progress in this direction was due to Remling and Kriecherbauer [63, 44]. In particular, they constructed fairly explicit examples of potentials satisfying $|V(x)| \leq C(1+|x|)^{-\alpha}$, $\alpha > 1/2$, such that the Hausdorff dimension of the set of singular energies where the WKB asymptotic behavior (2.9) fails is equal to $2(1-\alpha)$. This is sharp according to Theorem 2.3 (and earlier work of Remling [64] on power decaying potentials). The set of singular energies is the natural candidate to support the singular continuous part of the measure, but the actual presence of the singular continuous part of the measure remained open.

The first breakthrough came in a work of Denisov [18] where the following was proved

THEOREM 3.2. *There exist potentials $V \in L^2$ such that the operator H_V has imbedded singular continuous spectrum in $(0, \infty)$.*

The method was inspired by some ideas in approximation theory (see the next section) and by inverse spectral theory. The classical inverse spectral theory results (see e.g. [49, 50]) imply that one can find potentials corresponding to spectral measures with an imbedded singular continuous component. The standard procedure, however, does not guarantee a decaying potential. In the meantime, one can develop an inverse spectral theory type of construction where one also controls the L^2 norm of the potentials corresponding to certain approximations of the desired

spectral measure, in the limit obtaining the L^2 potential. The key control of the L^2 norm appears essentially from the sum rule used by Deift and Killip. The construction in [18] employed some estimates for the Krein systems, a certain system of first order differential equations. The amazing aspect of the construction is the great flexibility on how the singular part of the spectral measure may look. Later, in [23], it was proved that the imbedded singular continuous spectrum can occur for faster decaying potentials, namely, if

$$\int_x^\infty q^2(t) dt \leq C(1+x)^{-1+D+\epsilon},$$

then the spectral measure can have a singular continuous component of exact dimension D .

Killip and Simon [34] realized that the idea of [18] is not tied to the Krein systems. They proved a comprehensive theorem, providing a complete characterization of the spectral measures of Jacobi matrices which are Hilbert–Schmidt perturbations of the free matrix. This theorem should be regarded as an analog of the celebrated Szegő theorem for polynomials orthogonal on the unit circle [75, 73]. Recently, they also extended their result to the continuous case, where it reads as follows [35]. Denote $d\rho(E)$ the spectral measure of H_V , set $d\rho_0(E) = \pi^{-1}\chi_{[0,\infty)}(E)\sqrt{E}dE$, and define a signed measure $\nu(k)$ on $(1, \infty)$ by

$$\frac{2}{\pi} \int f(k^2)k d\nu(k) = \int f(E)[d\rho(E) - d\rho_0(E)].$$

Given a (signed) Borel measure ν , define

$$M_s\nu(k) = \sup_{0 < L \leq 1} \frac{1}{2L} |\nu|([k - L, k + L]).$$

Denote by $d\mu/d\sigma$ the Radon–Nikodym derivative of μ with respect to σ .

THEOREM 3.3. *A positive measure ρ on \mathbb{R} is the spectral measure associated to a $V \in L^2(\mathbb{R}^+)$ if and only if*

(i) $\text{supp}(d\rho) = [0, \infty) \cup \{E_j\}_{j=1}^N$ with $E_1 < E_2 < \dots < 0$ and $E_j \rightarrow 0$ if $N = \infty$.

(ii)

$$\int \log \left[1 + \left(\frac{M_s\nu(k)}{k} \right)^2 \right] k^2 dk < \infty \tag{3.4}$$

(iii)

$$\sum_j |E_j|^{3/2} < \infty$$

(iv)

$$\int_0^\infty \log \left[\frac{1}{4} \frac{d\rho}{d\rho_0} + \frac{1}{2} + \frac{1}{4} \frac{d\rho_0}{d\rho} \right] \sqrt{E} dE < \infty.$$

The theorem shows explicitly that the singular part of the spectral measure corresponding to an L^2 potential can be pretty much anything on the positive half-axis, as far as a certain normalization condition (3.4) is satisfied.

The last example that we would like to mention provides the sharp rate of decay for which the imbedded singular continuous spectrum may appear [38].

THEOREM 3.4. *For any positive monotone increasing function $h(x) \rightarrow \infty$, there exist potentials satisfying $|V(x)| \leq \frac{h(x)}{1+|x|}$ such that the half-line operator H_V (with, say, Dirichlet boundary condition) has imbedded singular continuous spectrum. The potential $V(x)$ can be chosen so that the Möller wave operators exist, but are not asymptotically complete due to the presence of the singular continuous spectrum.*

On the other hand, if $|V(x)| \leq \frac{C}{1+|x|}$, the singular continuous spectrum of H_V is empty.

The proof is based on building a sequence of approximating potentials V_n which have, respectively, 2^n imbedded eigenvalues E_j^n , approaching a Cantor set. The key is to obtain uniform control of the norms of the corresponding eigenfunctions, $\|u(x, E_j^n)\|_2^2 \leq C2^{-n}$. Such an estimate allows one to control the weights the spectral measure assigns to each eigenvalue, and to pass to the limit obtaining a nontrivial singular continuous component. The estimate of the norms of the eigenfunctions is difficult and is proved using the Prüfer transform, and a Splitting Lemma allowing one to obtain two imbedded eigenvalues from one. This lemma is based on a model nonlinear dynamical system providing an elementary block of construction.

4. Dirac Operators, Krein Systems, Jacobi Matrices, and OPUC

It was understood a long time ago that the spectral theory of one-dimensional differential operators (Schrödinger, Dirac, canonical systems) has a lot in common with the classical theory of polynomials orthogonal on the real line. These polynomials are eigenfunctions of the Jacobi matrix, also quite a classical object in analysis. So, naturally, to understand the problems for differential operators one might first study analogous problems for the discrete version. Unfortunately, Jacobi matrices are not so easy to study either. That difficulty was encountered by many famous analysts (such as Szegő) and the answer was found in the theory of polynomials orthogonal on the unit circle. It turns out that for many questions (especially in scattering theory) that is a more natural and basic object to study. Then many results and ideas can be implemented for Jacobi matrices. For differential operators, the situation is similar. In many cases, instead of a Schrödinger operator, it makes sense to consider a Dirac operator and for good reason. Already in 1955, Krein [43] gave an outline of the construction that led to the theory of continuous analogs of polynomials orthogonal on the circle. Instead of complex polynomials, one has the functions of exponential type that satisfy the corresponding system of differential equations (the Krein system)

$$\begin{cases} P'(r, \lambda) = i\lambda P(r, \lambda) - \bar{A}(r)P_*(r, \lambda), & P(0, \lambda) = 1, \\ P'_*(r, \lambda) = -A(r)P(r, \lambda), & P_*(0, \lambda) = 1. \end{cases} \quad (4.1)$$

Although more complicated than the OPUC case, the corresponding theory can be developed. It turns out that the Krein systems happen to be in one-to-one correspondence with the canonical Dirac operators:

$$D \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} -b & d/dr - a \\ -d/dr - a & b \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}, f_2(0) = 0. \quad (4.2)$$

In fact, $a(r) = 2\Re A(2r)$, $b(r) = 2\Im A(2r)$.

In his pioneering paper [43], Krein states the following

THEOREM 4.1. *If $a(r), b(r) \in L^2(\mathbb{R}^+)$, then $\sigma_{ac}(D) = \mathbb{R}$. Also,*

$$\int_{-\infty}^{\infty} \frac{\ln \sigma'(\lambda)}{1 + \lambda^2} > -\infty$$

where σ is the spectral measure for Dirac operator.

This result is actually the corollary of the analogous statement for Krein systems. Just as OPUC are related to Jacobi matrices, Krein systems and Dirac operators generate Schrödinger operators. Assume that $b(r) = 0$ and $a(r)$ is absolutely continuous. Taking the square of the operator D , we obtain the 2×2 diagonal matrix with Schrödinger operators on the diagonal. They have potentials $q_{1(2)} = a^2 \pm a'$ and certain boundary conditions. That clearly can be used to reduce the study of Schrödinger operators to the problems in the Krein systems theory. Indeed, assume that the potential q in a Schrödinger operator decays at infinity, say, $q \in L^p(\mathbb{R}^+), p < \infty$. Add the large constant γ to the operator and consider one of the corresponding Riccati equations, say:

$$q + \gamma = a^2 + a'.$$

We try to find a solution in the form $a = \sqrt{\gamma} + \mu$. Then,

$$\mu' + 2\sqrt{\gamma}\mu = q - \mu^2$$

and a decaying solution μ can be found by the contraction principle for the corresponding integral equation as long as γ is large enough. It will also be in $L^p(\mathbb{R}^+)$. Since adding a constant γ does not change the spectral types, the study of a Schrödinger operator with a decaying potential is essentially equivalent to the study of a Dirac operator (and so a Krein system) with $b = 0, a = \sqrt{\gamma} + \mu$ where γ is a constant and μ decays at infinity in essentially the same way as the Schrödinger potential does [16, 17].

We already mentioned the problem of proving the existence of wave operators for L^2 potentials in the second section. In [21], it was proved that for the Dirac operator, wave operators do exist if $a, b \in L^2$. The proof bypasses the question about a.e. in energy behavior of the eigenfunctions, and employs instead integral estimates. The key difference between the Dirac and Schrödinger cases is different free evolution. One manifestation of this difference is the fact that no WKB correction is needed in the definition of wave operators; the usual Möller wave operators exist for L^2 perturbations of the Dirac operator. Nevertheless, the result may indicate that the L^2 wave operator question for Schrödinger operators is easier to resolve than the question of the asymptotic behavior (2.9).

The study of Dirac operators is often more streamlined than that of Schrödinger operators in both one-dimensional and multidimensional cases (see the next section), but it already poses significant technical difficulties whose resolution is far from obvious and has proved to be very fruitful for the subject in general.

5. The Multidimensional Case

As opposed to the one-dimensional theory, the spectral properties of Schrödinger operators with slowly decaying potentials in higher dimensions are much less understood. Early efforts focused on the short range case, $|V(x)| \leq C(1 + |x|)^{-1-\epsilon}$, culminating in the proof by Agmon [1] of the existence and asymptotic completeness of wave operators in this case. In the long range case, Hörmander [29] considered

a class of symbol-like potentials, proving existence and completeness of wave operators. For a review of these results and other early literature, see [60, 79]. For potentials with less regular derivatives, the conjecture by Simon [71] states that the absolutely continuous spectrum of the operator H_V should fill the entire positive half-axis if

$$\int_{\mathbb{R}^d} V^2(x)(1 + |x|)^{-d+1} dx < \infty. \quad (5.1)$$

To avoid problems with the definition of the corresponding self-adjoint operator (that might appear for dimension high enough because of the local singularity of potential), we also assume that V belongs to, say, a Kato class: $V \in K_d(\mathbb{R}^d)$ [13].

Recall that there exist potentials $W(r)$ in one dimension which satisfy $|W(r)| \leq Cr^{-1/2}$ and lead to purely singular spectrum [42, 40]. By taking a spherically symmetric potential $V(x) = W(|x|)$, one can obtain multidimensional examples showing that (5.1) is sharp in many natural scales of spaces. Notice also that the potential satisfying (5.1) does not have to decay at infinity pointwise in all directions: it can even grow along some of them. Nevertheless, it does decay in the average and that makes the conjecture plausible.

Motivated by Simon's conjecture, much of the recent research focused on long range potentials with either no additional conditions on the derivatives, or with weaker conditions than in the classical Hörmander work. The solutions method so far had little success in higher dimensions. There are results linking the behavior of solutions and spectrum which work in higher dimensions, such as, for example, the following theorem proved in [39]. In higher dimensional problems, there is no canonical spectral measure, and the spectral multiplicity can be infinite. Given any function $\phi \in L^2(\mathbb{R}^d)$, denote μ^ϕ the spectral measure of H_V corresponding to ϕ , that is, a unique finite Borel measure such that $\langle f(H_V)\phi, \phi \rangle = \int f(E)d\mu^\phi(E)$ for all continuous f with compact support.

THEOREM 5.1. *Assume that the potential V is bounded from below. Suppose that there exists a solution $u(x, E)$ of the generalized eigenfunction equation $(H_V - E)u(x, E) = 0$ such that*

$$\liminf_{R \rightarrow \infty} R^{-1} \int_{|x| \leq R} |u(x, E)|^2 dx < \infty. \quad (5.2)$$

Fix some $\phi(x) \in C_0^\infty(\mathbb{R}^d)$ such that

$$\int_{\mathbb{R}^d} \phi(x)u(x, E) dx \neq 0.$$

Then we have

$$\limsup_{\delta \rightarrow 0} \frac{\mu^\phi(E - \delta, E + \delta)}{2\delta} > 0. \quad (5.3)$$

Notice that if (5.3) holds on some set S of positive Lebesgue measure, this implies that the usual Lebesgue derivative of μ^ϕ is positive a.e. on S , and so the presence of the absolutely continuous spectrum. The condition (5.2) corresponds to the power decay $|x|^{(1-d)/2}$, just as spherical wave solutions decay for the free Laplacian. One may ask how precise this condition is: perhaps the existence of just bounded solutions on a set S of positive Lebesgue measure is sufficient for the

presence of the absolutely continuous spectrum? It turns out that, in general, the condition (5.2) cannot be relaxed.

THEOREM 5.2. *There exists a potential V such that for any $\sigma > 0$, there exists an energy interval I_σ with the following properties:*

- *For a.e. $E \in I_\sigma$, there exists a solution $u(x, E)$ of the generalized eigenfunction equation satisfying $|u(x, E)| \leq C(E)(1 + |x|)^{\sigma+(1-d)/2}$.*
- *The spectrum on I_σ is purely singular.*

One way to prove Theorem 5.2 is to use one-dimensional random decaying potentials with $|x|^{-1/2}$ rate of decay. The results of [42] or [40] show that the spectrum is singular almost surely, and the eigenfunctions decay at a power rate. Taking spherically symmetric potentials of this type in higher dimensions, it is not difficult to see that one gets examples proving Theorem 5.2.

The link between the behavior of solutions and spectral measures has been made even more general, sharp and abstract in [12]. However, the difficulty is that obtaining enough information about solutions in problems of interest is hard: there seems to be no good PDE analog for the ODE perturbation techniques which can be used to understand the solutions in one dimension. On the other hand, the sum rules in higher dimensions typically involve spaces of potentials which are far from the conjectured class (5.1). Some important progress, however, has been made using the one-dimensional ideas [20, 19, 22, 47, 48, 59, 68]. After reviewing these results, we will discuss random decaying potentials in higher dimensions [5, 6, 65] as well as quickly mention some interesting recent progress on a new class of short range potentials [31] and imbedded eigenvalues [30].

The OPUC, the Krein systems, and the Dirac operators with matrix-valued and even operator-valued coefficients have been studied relatively well. The matrix-valued case can give some clues to the understanding of partial differential equations. Indeed, writing up, say, the Schrödinger operator in the spherical coordinates, one obtains the one-dimensional Schrödinger operator with an operator-valued potential. The difficulty is that this potential \tilde{V} is not bounded since it involves a Laplace–Beltrami operator on the unit sphere. Also $\tilde{V}(r_1)$ and $\tilde{V}(r_2)$ do not commute for different values of r .

Consider the three-dimensional Dirac operator with the following type of interactions

$$D = -i\alpha \cdot \nabla + V(x)\beta. \tag{5.4}$$

Here

$$\alpha_j = \begin{pmatrix} 0 & \sigma_j \\ \sigma_j & 0 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Matrices σ_j are called the Pauli matrices and

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

Then the multidimensional analog of Theorem 4.1 says [20] that if $V(x) \in L^\infty$ and the estimate (5.1) holds then $\sigma_{ac}(D) = \mathbb{R}$. Thus the multidimensional result for the Dirac operator is quite satisfactory, and the Dirac analog of Simon's conjecture (5.1) holds. We will sketch the ideas behind this result later. For the Schrödinger

operator, we first state the following interesting result by Safronov [68]. Let us denote by B_r the ball of radius r centered at the origin.

THEOREM 5.3. *Let $d \geq 3$ and suppose $V \in L^\infty(\mathbb{R}^d)$ is such that $V(x) \rightarrow 0$ as $x \rightarrow \infty$. Assume also that $V(x) \in L^{d+1}(\mathbb{R})$ and for some positive $\delta > 0$, the Fourier transform of V satisfies $\hat{V}(\xi) \in L^2(B_\delta)$. Then, $\sigma_{ac}(H) = \mathbb{R}^+$.*

Note that the L^{d+1} condition corresponds to the $|x|^{-\frac{d}{d+1}}$ power decay. There are several methods to obtain this kind of result. The first one, developed by Laptev, Naboko and Safronov [45], is based on writing the operator in the spherical coordinates. Then one takes the Feshbach projection corresponding to the first harmonic and studies the corresponding one-dimensional Schrödinger operator with nonlocal operator-valued potential. Instead of trace equality one can get an inequality only which still is enough to conclude the presence of a.c. spectrum. On the other hand, instead of dealing with Feshbach projections, one can carefully study the matrix-valued Dirac or Schrödinger operator and obtain the estimates on the entropy of the spectral measure independent of the size of the matrix [20]. Then, an analogous estimate can be obtained for the corresponding PDE.

Another approach allows one to work directly with PDE [19]. It consists of the following observation. Consider, for example, the three-dimensional Schrödinger operator with compactly supported potential. Taking $\phi \in L^\infty$ with compact support, we introduce $u(x, k) = (H - k^2)^{-1}\phi$ with $k \in \mathbb{C}^+$, $\Re k > 0$. Then, clearly, $u(x, k)$ has the following asymptotic behavior at infinity:

$$u(x, k) = \frac{\exp(ikr)}{r} (A_\phi(k, \theta) + \bar{o}(1)), \quad r = |x|, \theta = x/r.$$

The amplitude $A_\phi(k, \theta)$ can be regarded as an analytic operator on $L^2(\Sigma)$, where Σ is the unit sphere. For the potential V with compact support, it is continuous up to each boundary point $k > 0$ and the following factorization identity holds [79]

$$\mu'_\phi(E) = k\pi^{-1} \|A_\phi(k, \theta)\|_{L^2(\Sigma)}^2, \quad E = k^2 > 0. \quad (5.5)$$

Loosely speaking, the density of the spectral measure for any vector ϕ can be factorized on the positive interval via the boundary value of some analytic operator-valued function defined in an adjacent domain in \mathbb{C}^+ . Therefore, one can try to consider the general potential V , establish existence of $A_\phi(k, \theta)$ for all $k \in \mathbb{C}^+$, $\Re k > 0$ with some (probably rather crude) bounds on the boundary behavior near the real line. Then the analyticity will be enough to conclude the necessary estimate on the entropy of the spectral measure, similarly to the one-dimensional considerations. Here is the general result.

THEOREM 5.4. *Consider a potential $V \in L^\infty(\mathbb{R}^d)$. Let $V_n(x) = V(x)\chi_{|x|<n}$ be its truncation and $A_{\phi,n}(k, \theta)$, the corresponding amplitude. Consider an interval $0 < a < k < b$. Assume that for $k_0 = (a + b)/2 + i\sigma$, $\sigma > 0$, we have an estimate $\|A_{\phi,n}(k_0, \theta)\|_{L^2(\Sigma)} > \delta > 0$ uniformly in n and $\|A_{\phi,n}(\tau + i\varepsilon, \theta)\|_{L^2(\Sigma)} < C \exp(\varepsilon^{-\gamma})$, $\gamma > 0$ uniformly in n , $\tau \in [a, b]$, $0 < \varepsilon < 2\sigma$. Then the spectral measure of the function ϕ has an a.c. component whose support contains an interval $[a^2, b^2]$.*

The situation is reminiscent of one in the Nevanlinna theory in the classical analysis when the analyticity and rough bounds close to the boundary are enough to say a lot about the function. Thus the whole difficulty here is to obtain the

necessary bounds for the particular PDE. That turns out to be a tricky task but doable in some cases. For example, the following theorem was proved in [19].

THEOREM 5.5. *Let $Q(x)$ be a $C^1(\mathbb{R}^3)$ vector-field in \mathbb{R}^3 and*

$$|Q(x)| < \frac{C}{1 + |x|^{0.5+\epsilon}}, \quad |\operatorname{div} Q(x)| < \frac{C}{1 + |x|^{0.5+\epsilon}}, \quad |V_1(x)| < \frac{C}{1 + |x|^{1+\epsilon}}, \quad \epsilon > 0.$$

Then, $H = -\Delta + \operatorname{div} Q + V_1$ has an a.c. spectrum that fills \mathbb{R}^+ .

This theorem is a multidimensional analog of the following result in dimension one [16]: if $V = a' + V_1$ where $V_1 \in L^1(\mathbb{R}^+)$, $a \in W_2^1(\mathbb{R}^+)$, then the a.c. spectrum of a one-dimensional Schrödinger operator with potential V covers the positive half-line. The proof involves Theorem 5.4, and is based on the uniform estimates for the Green’s function on complex energies. The PDE approach used to prove Theorem 5.5 succeeds because it allows one to not deal with negative eigenvalues and the corresponding Lieb–Thirring inequalities often arising in the sum rule approach (see also [33]).

Although rather elaborate, the conditions on the potential from Theorem 5.5 are not very difficult to check. Essentially, they mean that in addition to decay, one has to have certain oscillation of potential. This condition is related, although not identical, to the condition on the Fourier transform in Theorem 5.3. For example, an application to random potentials is possible. Consider the following model. Take a smooth function $f(x)$ with the support inside the unit ball. Consider

$$V_0(x) = \sum_{j \in \mathbb{Z}^+} a_j f(x - x_j)$$

where the points x_j are scattered in \mathbb{R}^3 such that $|x_k - x_l| > 2, k \neq l$, and $a_j \rightarrow 0$ in a way that $|V_0(x)| < C/(1 + |x|^{0.5+\epsilon})$. Let us now “randomize” V_0 as follows:

$$V(x) = \sum_{j \in \mathbb{Z}^+} a_j \xi_j f(x - x_j), \tag{5.6}$$

where ξ_j are real-valued, bounded, independent random variables with even distribution.

THEOREM 5.6. *For V given by (5.6), we have $\sigma_{ac}(-\Delta + V) = \mathbb{R}^+$ almost surely.*

It turns out that for dimension high enough, these slowly decaying random potentials fall into the class considered in Theorem 5.5. We note that a similar model has been considered by Bourgain [5, 6] in the discrete setting, and will be discussed below. Theorem 5.6 and Bourgain’s results suggest very strongly that Simon’s conjecture (5.1) is true at least in a certain “almost sure” sense (however, the assumption that the random variables are mean zero is crucial for the proofs).

The method used in [19] was also applied by Perelman in the following situation [59]:

THEOREM 5.7. *For $d = 3$, $\sigma_{ac}(-\Delta + V) = \mathbb{R}^+$ as long as*

$$|V(x)| + |x| |\nabla' V(x)| < C/(1 + |x|^{0.5+\epsilon}), \quad \epsilon > 0$$

where ∇' means the angular component of the gradient.

Here the oscillation is arbitrary in the radial variable and slow in the angular variable. In this case, the Green function has a WKB-type correction which can be explicitly computed (and has essentially one-dimensional, integration along a ray, form). The notion of the amplitude $A_\phi(k, \theta)$ can be modified accordingly and the needed estimates on the boundary behavior can be obtained.

We now return to random decaying potentials and discuss recent developments in more detail. It is natural to tackle slowly decaying random potentials if one tries to approach one of the most important open problems in mathematical quantum mechanics: the existence of extended states in the Anderson model in higher dimensions. Important progress in understanding random slowly decaying potentials is due to Bourgain [5, 6]. Consider a random lattice Schrödinger operator on \mathbb{Z}^2 : $H_\omega = \Delta + V_\omega$, where Δ is the usual discrete Laplacian and V_ω is a random potential

$$V_\omega = \omega_n v_n$$

with $|v_n| < C|n|^{-\rho}$, $\rho > 1/2$. The random variables ω_n are Bernoulli or normalized Gaussian (and, in particular, are mean zero). Then,

THEOREM 5.8 ([5]). *Fix $\tau > 0$ and denote $I = \{E \mid \tau < |E| < 4 - \tau\}$. Assume that $\rho > 1/2$, and $\sup_n |v_n| |n|^{-\rho} < \kappa$. Then for $\kappa \leq \kappa(\rho, \tau)$ and for ω outside a set of small measure (which tends to zero as $\kappa \rightarrow 0$) we have*

1. H_ω has purely absolutely continuous spectrum on I
2. Denoting $E_0(I)$, the spectral projections for Δ , the wave operators $W_\pm(H, \Delta)E_0(I)$ exist and are complete.

Using the fact that the absolutely continuous spectrum and the existence of the wave operators are stable under finitely supported perturbations, one readily obtains absolute continuity and existence of wave operators almost surely for potentials satisfying $|v_n| \leq C|n|^{-\rho}$. The method can also be extended effortlessly to dimensions $d > 2$.

Bourgain's approach is based on a careful analysis of the Born approximation series for the resolvent. In summation, each of the terms $[R_0(z)V]^s R_0(z)$ is considered. Then a dyadic decomposition of V is introduced: $V = \sum_j V \chi_{2^j < |x| < 2^{j+1}}$. In the end, the analysis is reduced to getting multilinear bounds for the resulting terms. An interesting (and novel in this context) ingredient of the proof is the smart use of a certain entropy bound (the so-called "dual to Sudakov" inequality). Later [6] this approach was further developed to deal with different situations, such as L^p and slower power decaying potentials. The main result of [6] for the slower power decay is the almost sure existence of a bounded, not tending to zero solution at a single energy. This, however, is not yet sufficient for any spectral conclusions. The problem of handling the random decay with the coefficient ρ even a little less than $1/2$ remains an interesting open question. So far, all attempts to deal with this case were not successful.

In another paper on random decaying potentials [65], Rodnianski and Schlag showed existence of modified wave operators for the model with the slow random decay and additional assumptions ensuring slow variation of the derivatives. The standard techniques of scattering theory, but also with averaging over the randomness, are employed. This allows them to prove scattering with weaker assumptions than in the standard Hörmander's case.

Another case for which scattering can be established is the Schrödinger operator on the strip [22]. One can show the presence of the a.c. part of the spectrum using

the following general result. Assume that we are given two operators H_1 and H_2 that both act in the same Hilbert space \mathcal{H} . Take H in the following form:

$$H = \begin{bmatrix} H_1 & V \\ V^* & H_2 \end{bmatrix}. \tag{5.7}$$

THEOREM 5.9. *Let H_1, H_2 be two bounded self-adjoint operators in the Hilbert space \mathcal{H} . Assume that $\sigma_{\text{ess}}(H_2) \subseteq [b, +\infty]$ and $[a, b] \subseteq \sigma_{\text{ac}}(H_1)$, ($a < b$). Then, for any Hilbert–Schmidt V , ($V \in \mathcal{J}^2$), we have that $[a, b] \subseteq \sigma_{\text{ac}}(H)$, with H given by (5.7).*

This theorem can be effectively applied to study the Schrödinger operator on the strip. Indeed, let

$$L = -\Delta + Q(x, y),$$

considered on the strip $\Pi = \{x > 0, 0 < y < \pi\}$, and impose Dirichlet conditions on the boundary of Π . Consider the matrix representation of L . For $f(x, y) \in L^2(\Pi)$,

$$f(x, y) = \sqrt{\frac{2}{\pi}} \sum_{n=1}^{\infty} \sin(ny) f_n(x), \quad f_n(x) = \sqrt{\frac{2}{\pi}} \int_0^{\pi} f(x, y) \sin(ny) dy$$

and L can be written as follows

$$L = \begin{bmatrix} -\frac{d^2}{dx^2} + Q_{11}(x) + 1 & Q_{12}(x) & \dots \\ Q_{21}(x) & -\frac{d^2}{dx^2} + Q_{22}(x) + 4 & \dots \\ \dots & \dots & \dots \end{bmatrix} \tag{5.8}$$

$$Q_{lj}(x) = \frac{2}{\pi} \int_0^{\pi} Q(x, y) \sin(ly) \sin(jy) dy$$

Assume $\sup_{0 \leq y < \pi} |Q(x, y)| \in L^2(\mathbb{R}^+) \cap L^\infty(\mathbb{R}^+)$. Since $Q_{11}(x) \in L^2(\mathbb{R}^+)$, we can use [15, 54] and Theorem 5.9 to show the presence of a.c. spectrum.

We now turn to the results on the (generalized) short range potentials and imbedded eigenvalues in \mathbb{R}^d . Recently, new short range type results have been obtained for the multidimensional Schrödinger operator with potential from L^p and more general classes [27, 31]. The main goal was to establish limiting absorption estimates for the resolvent acting in certain Banach spaces, which are more detailed and precise than Agmon’s classical results. The standard techniques developed in the works of Agmon can be improved if one uses the Stein–Tomas restriction theorem. Here is one of the results in that direction [31]:

THEOREM 5.10. *Assume that V is such that*

$$M_q(V)(x) \in L^{(d+1)/2}(\mathbb{R}^d)$$

where

$$M_q f = \left[\int_{|y| < 1/2} |f(x+y)|^q dy \right]^{1/q}$$

and $q = d/2$ if $d \geq 3$, $q > 1$ if $d = 2$. Then, the following is true for the operator $H = -\Delta + V$:

- The set of nonzero eigenvalues is discrete with the only possible accumulation point at zero. Each nonzero eigenvalue has finite multiplicity.
- $\sigma_{\text{sc}}(H) = \emptyset$, and $\sigma_{\text{ac}}(H) = \mathbb{R}^+$
- The wave operators $\Omega^\pm(H, H_0)$ exist and are complete.

The actual result is a bit stronger. The authors of [31] present the whole class of “admissible” perturbations for which their method works, including some first order differential operators.

Another direction in which there has been significant recent progress concerns imbedded eigenvalues. The following result, in particular, has been proved by Ionescu and Jerison [30].

THEOREM 5.11. *Let $V(x) \in L^{d/2}(\mathbb{R}^d)$ and for $d = 2$ we also assume $V(x) \in L^r_{\text{loc}}(\mathbb{R}^2)$, $r > 1$. Then, $H = -\Delta + V$ does not have positive eigenvalues.*

The paper [30] actually contains a more general result, which allows for slower decay of the potential if its singularities are weaker. The method relies on the Carleman inequality of special type.

Surprisingly, [30] also provides an example of potential V satisfying

$$|V(x)| < C(|x_1| + x_2^2 + \dots + x_d^2)^{-1} \quad (5.9)$$

for which a positive eigenvalue appears (the multidimensional analog of Wigner–von Neumann potential). From the point of view of physical intuition, the existence of an imbedded eigenvalue for such a potential may seem strange. Indeed, one would expect that tunneling in the directions x_2, \dots, x_d of fast decay should make the bound state impossible. Yet, Wigner–von Neumann type oscillation and slow Coulomb decay in just one direction turn out to be sufficient. The corresponding eigenfunction decays rather slowly but enough to be from L^2 . We note that the potential satisfying (5.9) just misses $L^{(d+1)/2}(\mathbb{R}^d)$. Quite recently, Koch and Tataru [41] improved Theorem 5.11 and showed the absence of imbedded eigenvalues for the optimal $L^{(d+1)/2}(\mathbb{R}^d)$ case. They also considered various long range potentials and more general elliptic operators.

6. The Bethe Lattice

A great example of how the one-dimensional technique works for the multidimensional problem is the case of scattering on the Bethe lattice [22]. The step-by-step sum rules used by Simon to study Jacobi matrices [72] can be adjusted to that case. Let us consider this model. Take the Cayley tree (Bethe lattice) \mathbb{B} and the discrete Laplacian on it

$$(H_0 u)_n = \sum_{|i-n|=1} u_i.$$

Assume, for simplicity, that the degree at each point (the number of neighbors) is equal to 3. It is well known that $\sigma(H_0) = [-2\sqrt{2}, 2\sqrt{2}]$ and the spectrum is purely absolutely continuous. Let $H = H_0 + V$, where V is a potential. Consider any vertex O . It is connected to its neighbors by three edges. Delete one edge together with the corresponding part of the tree stemming from it. What is left will be called \mathbb{B}_O . The degree of O within \mathbb{B}_O is equal to 2. The solution to Simon’s conjecture in this case is given by the following theorem. We denote by the symbol $c_0(\mathbb{B})$ the functional space of sequences decaying at infinity on \mathbb{B} .

THEOREM 6.1. *If $V \in \ell^\infty(\mathbb{B}) \cap c_0(\mathbb{B}_O)$ and*

$$\sum_{n=0}^{\infty} \frac{1}{2^n} \sum_{x \in \mathbb{B}_O, |x-O|=n} V^2(x) < +\infty \tag{6.1}$$

then

$$[-2\sqrt{2}, 2\sqrt{2}] = \sigma_{ac}(H_{|\mathbb{B}_O}) \subseteq \sigma_{ac}(H).$$

The idea of the proof of Theorem 6.1 is based on [72]. This result is sharp in the following sense. Take $V(x) : |V(x)| < C|x - O|^{-\gamma}$. Then for any $\gamma > 0.5$, the condition of the theorem is satisfied (just like in \mathbb{R}^d). In the meantime, one can find the spherically symmetric V with slower decay $0 < \gamma < 0.5$, such that there will be no absolutely continuous spectrum at all.

An important role in the proof is played by the well-known recursive relation for $\langle (H_{|\mathbb{B}_O} - z)^{-1} \delta_O, \delta_O \rangle$, where δ_O is the discrete delta-function at the point O . In particular, one can derive the following important and physically meaningful identity

$$\frac{1}{\pi} \int_{-2\sqrt{2}}^{2\sqrt{2}} \sqrt{8 - \lambda^2} \ln[\mu'_O(\lambda)] d\lambda \geq \frac{1}{\pi} \int_{-2\sqrt{2}}^{2\sqrt{2}} \sqrt{8 - \lambda^2} \ln \left(\frac{\mu'_{O_1}(\lambda) + \mu'_{O_2}(\lambda)}{2} \right) d\lambda - V^2(O) \tag{6.2}$$

where $\mu'_{O_{1(2)}}$ correspond to the densities of the spectral measures at points $O_{1(2)}$ on the trees obtained from \mathbb{B}_O by throwing away the point O along with the corresponding two edges. Using inequality $(x + y)/2 \geq \sqrt{xy}$ and iterating (6.2), one proves Theorem 6.1. Formula (6.2) says, in particular, that no matter what happens along one branch of the tree, the scattering is possible through the other branch. It is also clear that the presence of some “bad” points in the tree (say, points where we have no control over the potential) should not destroy the scattering as long as these points are rather “sparse”. What is an accurate measurement of this sparseness? We suggest the following improvement of Theorem 6.1. Consider the tree \mathbb{B}_O with potential V having finite support, that is, $V(x) = 0$ for $|x - O| > R$. Consider all paths running from the origin O to infinity without self-intersections. Using dyadic decomposition of the real numbers on the interval $[0, 1]$, we can assign to each path the real number in the natural way. That is one way of coding the points at infinity. In principle, this map F is not a bijection, e.g. sequences $(1, 0, 0, \dots)$ and $(0, 1, 1, \dots)$ represent the same real number 0.5 but different paths. Fortunately, these numbers have Lebesgue measure zero and will be of no importance for us. Let us define the following functions

$$\phi(t) = \sum_{n=1}^{\infty} V^2(x_n)$$

where x_n are all vertices of the path representing the point $t \in [0, 1]$. Since the support of V is within the ball of radius R , function $\phi(t)$ is constant on dyadic intervals $[j2^{-R}, (j+1)2^{-R}]$, $j = 0, 1, \dots, 2^R - 1$. Notice that F , not being a bijection, causes no trouble in defining $\phi(t)$.

Define the probability measure with density $w(\lambda) = (4\pi)^{-1}(8 - \lambda^2)^{1/2}$ on $[-2\sqrt{2}, 2\sqrt{2}]$, and $\rho_O = \sigma'_O(\lambda)[w(\lambda)]^{-1}$, a relative density of the spectral measure

at the point O . Define

$$s_O = \int_{-2\sqrt{2}}^{2\sqrt{2}} \ln \rho_O(\lambda) w(\lambda) d\lambda.$$

Consider the probability space obtained by assigning to each path the same weight (i.e., as we go from O to infinity, we toss the coin at any vertex and move to one of the neighbors farther from O depending on the result). Our goal is to prove the following

THEOREM 6.2. *For any V bounded on \mathbb{B}_O , the following inequality is true*

$$\exp s_O \geq \mathbb{E} \left\{ \exp \left[-\frac{1}{4} \sum_{n=1}^{\infty} V^2(x_n) \right] \right\} = \int_0^1 \exp \left(-\frac{1}{4} \phi(t) \right) dt \quad (6.3)$$

where the expectation is taken with respect to all paths $\{x_n\}$ going from O to infinity without self-intersections. In particular, if the r.h.s. of (6.3) is positive, then $[-2\sqrt{2}, 2\sqrt{2}] \subseteq \sigma_{ac}(H)$.

PROOF. Assume that V has finite support. The estimate (6.2) can be rewritten as

$$\int_{-2\sqrt{2}}^{2\sqrt{2}} \ln \rho_O(\lambda) w(\lambda) d\lambda \geq \int_{-2\sqrt{2}}^{2\sqrt{2}} \ln \left(\frac{\rho_{O_1}(\lambda) + \rho_{O_2}(\lambda)}{2} \right) w(\lambda) d\lambda - V^2(O)/4. \quad (6.4)$$

Now, we will use Young's inequality

$$\frac{x^p}{p} + \frac{y^q}{q} \geq xy; \quad x, y \geq 0, \quad 1 \leq p \leq \infty, \quad p^{-1} + q^{-1} = 1$$

in (6.4) to obtain

$$s_O \geq p^{-1} s_{O_1} + q^{-1} s_{O_2} + p^{-1} \ln p + q^{-1} \ln q - \ln 2 - V^2(O)/4. \quad (6.5)$$

Considering s_{O_1} and s_{O_2} to be fixed parameters and maximizing the r.h.s. over $p \in [1, \infty]$, we get the following inequality

$$s_O \geq \ln \frac{\exp s_{O_1} + \exp s_{O_2}}{2} - V^2(O)/4 \quad (6.6)$$

with optimal $p^* = 1 + \exp(s_{O_2} - s_{O_1})$. Iterate (6.6) until we leave the support of V . Thus, we get (6.3). Consider now the general case of bounded V . Define the truncation of V to the ball of radius n : $V_n(x) = V(x)\chi_{\{|x-O| \leq n\}}$. For the corresponding $s_O^{(n)}$, we use (6.3), take $n \rightarrow \infty$ and apply to the l.h.s. a standard by now argument on the semicontinuity of the entropy ([34], p. 293). Notice that the functions $\phi^{(n)}(t)$ are nonnegative and increasing in n for each t (this is why $\phi(t)$ is always well defined). Therefore, we get (6.3) from the theorem on monotone convergence. \square

Using Jensen's inequality, we obtain

COROLLARY 6.3. *Assume that A is any subset of $[0, 1]$ of positive Lebesgue measure and $\phi(t) \in L^1(A)$, then*

$$s_0 \geq -\frac{1}{4|A|} \int_A \phi(t) dt + \ln |A|.$$

In particular, $[-2\sqrt{2}, 2\sqrt{2}] \subseteq \sigma_{ac}(H)$.

It is interesting that the set A does not have to have some special topological structure, say, to be an interval like in the standard scattering theory [2].

There remain many interesting and important open problems regarding the multidimensional slowly decaying potentials. Simon's conjecture (5.1) remains open, and new ideas are clearly needed to make progress. Improving our understanding of random slowly decaying potentials is another quite challenging direction. Other difficult and intriguing open questions involve multidimensional sparse potentials, the appearance of imbedded singular continuous spectrum, and decaying potentials with additional structural assumptions. This vital area is bound to challenge and inspire mathematicians for years to come.

Acknowledgements. S. D. is supported in part by an Alfred P. Sloan Research Fellowship and NSF grant DMS-0500177. A. K. is supported in part by an Alfred P. Sloan Research Fellowship and NSF grant DMS-0314129.

References

- [1] S. Agmon, *Spectral properties of Schrödinger operators and scattering theory*, Ann Scuola Norm. Sup. Pisa II, **2**(1975), 151–218
- [2] W. O. Amrein and D. B. Pearson, *Flux and scattering into cones for long range and singular potentials*, J. Phys. A, **30** (1997), no. 15, 5361–5379
- [3] M. Ben-Artzi, *On the absolute continuity of Schrödinger operators with spherically symmetric, long-range potentials. I, II*. J. Differential Equations **38** (1980), 41–50, 51–60
- [4] H. Behncke, *Absolute continuity of Hamiltonians with von Neumann Wigner potentials. II*, Manuscripta Math. **71** (1991), 163–181
- [5] J. Bourgain, *On random Schrödinger operators on \mathbb{Z}^2* , Discrete Contin. Dyn. Syst. **8** (2002), 1–15
- [6] J. Bourgain, *Random lattice Schrödinger operators with decaying potential: some higher dimensional phenomena*, Geometric Aspects of Functional Analysis, Lecture Notes in Math. **1807** (2003), Springer, Berlin, 70–98
- [7] L. Carleson, *On convergence and growth of partial sums of Fourier series*, Acta Math. **116** (1966), 135–157
- [8] 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** (1998), 771–797
- [9] M. Christ and A. Kiselev, *Scattering and wave operators for one-dimensional Schrödinger operators with slowly decaying nonsmooth potentials*, Geom. Funct. Anal. **12** (2002), 1174–1234
- [10] M. Christ and A. Kiselev, *WKB and spectral analysis of one-dimensional Schrödinger operators with slowly varying potentials*, Comm. Math. Phys. **218** (2001), 245–262
- [11] M. Christ and A. Kiselev, *Absolutely continuous spectrum of Stark operators*, Ark. Mat. **41** (2003), 1–33
- [12] M. Christ, A. Kiselev and Y. Last, *Approximate eigenvectors and spectral theory*, in Differential Equations and Mathematical Physics, (R. Weikard and G. Weinstein, eds.), American Mathematical Society/International Press, Cambridge, 2000
- [13] H. L. Cycon, R. G. Froese, W. Kirsch and B. Simon, *Schrödinger Operators with Application to Quantum Mechanics and Global Geometry*. Texts and Monographs in Physics. Springer Study Edition. Springer-Verlag, Berlin, 1987

- [14] F. Delyon, B. Simon and B. Souillard, *From pure point to continuous spectrum in disordered systems*, Ann. Inst. Henri Poincaré **42** (1985), 283–309
- [15] P. Deift and R. Killip, *On the absolutely continuous spectrum of one-dimensional Schrödinger operators with square summable potentials*, Comm. Math. Phys. **203** (1999), 341–347
- [16] S. Denisov, *On the application of some of M. G. Krein's results to the spectral analysis of Sturm–Liouville operators*, J. Math. Anal. Appl. **261** (2001), 177–191
- [17] S. Denisov, *On the existence of the absolutely continuous component for the measure associated with some orthogonal systems*, Comm. Math. Phys. **226**, (2002), 205–220
- [18] 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** (2003), 90–104
- [19] S. Denisov, *Absolutely continuous spectrum of multidimensional Schrödinger operator*, Int. Math. Res. Not. **74** (2004), 3963–3982
- [20] S. Denisov, *On the absolutely continuous spectrum of Dirac operator*, Comm. Partial Diff. Eq. **29** (2004), 1403–1428
- [21] S. Denisov, *On the existence of wave operators for some Dirac operators with square summable potentials*, Geom. Funct. Anal. **14** (2004), 529–534
- [22] S. Denisov, *On the preservation of absolutely continuous spectrum for Schrödinger operators*, J. Funct. Anal. **231** (2006), 143–156
- [23] S. Denisov and S. Kupin, *On the singular spectrum of Schrödinger operators with decaying potential*, Trans. Amer. Math. Soc. **357** (2005), 1525–1544
- [24] A. Devinatz, *The existence of wave operators for oscillating potentials*, J. Math. Phys. **21** (1980), 2406–2411
- [25] L. Faddeev and L. Takhtajan, *Hamiltonian Methods in the theory of solitons*, (translated from the Russian by A. G. Reiman), Springer, Berlin, 1987
- [26] D. J. Gilbert and D. B. Pearson, *On subordinacy and analysis of the spectrum of one-dimensional Schrödinger operators*, J. Math. Anal. Appl. **128** (1987), 30–56
- [27] M. Goldberg and W. Schlag, *A limiting absorption principle for the three-dimensional Schrödinger equation with L^p potentials*, Int. Math. Res. Not. (2004), no. 75, 4049–4071
- [28] L. Grafakos, *Classical and Modern Fourier Analysis*, Pearson Education, Prentice Hall, Upper Saddle River, NJ, 2004
- [29] L. Hörmander, *The existence of wave operators in scattering theory*, Math. Z. **146** (1976), 69–91
- [30] A. Ionescu and D. Jerison, *On the absence of positive eigenvalues of Schrödinger operators with rough potentials*, Geom. Funct. Anal. **13** (2003), 1029–1081
- [31] A. Ionescu and W. Schlag, *Agmon–Kato–Kuroda theorems for a large class of perturbations*, Duke Math. J. **131**, (2006), 397–440
- [32] J. P. Kahane, *Some Random Series of Functions*, Cambridge University Press, Cambridge, 1985
- [33] R. Killip, *Perturbations of one-dimensional Schrödinger operators preserving the absolutely continuous spectrum*, Int. Math. Res. Not. **2002**, 2029–2061
- [34] R. Killip and B. Simon, *Sum rules for Jacobi matrices and their applications to spectral theory*, Ann. of Math. **158** (2003), 253–321
- [35] R. Killip and B. Simon, *Sum rules and spectral measures of Schrödinger operators with L^2 potentials*, to appear in Ann. of Math.
- [36] A. Kiselev *Absolutely continuous spectrum of one-dimensional Schrödinger operators and Jacobi matrices with slowly decreasing potentials*, Comm. Math. Phys. **179** (1996), 377–400
- [37] A. Kiselev, *Stability of the absolutely continuous spectrum of the Schrödinger equation under slowly decaying perturbations and a.e. convergence of integral operators*, Duke Math. J. **94** (1998), 619–646
- [38] A. Kiselev *Imbedded singular continuous spectrum for Schrödinger operators*, J. Amer. Math. Soc. **18** (2005), 571–603
- [39] A. Kiselev and Y. Last, *Solutions and spectrum of Schrödinger operators on infinite domains*, Duke Math. J. **102** (2000), 125–150
- [40] A. Kiselev, Y. Last and B. Simon, *Modified Prüfer and EFGP transforms and the spectral analysis of one-dimensional Schrödinger operators*, Comm. Math. Phys. **194** (1998), 1–45
- [41] H. Koch and D. Tataru, *Carleman estimates and absence of embedded eigenvalues*, preprint

- [42] S. Kotani and N. Ushiroya, *One-dimensional Schrödinger operators with random decaying potentials*, Comm. Math. Phys. **115** (1988), 247–266
- [43] M. Krein, *Continuous analogues of propositions on polynomials orthogonal on the unit circle*, Dokl. Akad. Nauk SSSR **105** (1955), 637–640
- [44] T. Kriecherbauer and C. Remling, *Finite gap potentials and WKB asymptotics for one-dimensional Schrödinger operators*, Comm. Math. Phys. **223** (2001), 409–435
- [45] A. Laptev, S. Naboko and O. Safronov, *On new relations between spectral properties of Jacobi matrices and their coefficients*, Comm. Math. Phys. **241** (2003), 91–110
- [46] A. Laptev, S. Naboko and O. Safronov, *Absolutely continuous spectrum of Jacobi matrices*, Mathematical Results in Quantum Mechanics (Taxco, 2001), 215–223, Contemp. Math. **307**, Amer. Math. Soc., Providence, RI, 2002
- [47] A. Laptev, S. Naboko and O. Safronov, *Absolutely continuous spectrum of Schrödinger operators with slowly decaying and oscillating potentials*, Comm. Math. Phys. **253** (2005), 611–631
- [48] A. Laptev, S. Naboko and O. Safronov, *A Szegő condition for a multidimensional Schrödinger operator*, J. Funct. Anal. **219** (2005), 285–305
- [49] B. Levitan, *Inverse Sturm–Liouville Problems*, VNU Science Press, Utrecht, 1987
- [50] V. Marchenko, *Sturm–Liouville Operators and Applications*, Birkhauser, Basel, 1986
- [51] D. Moshov, *Sur les series de fonctions orthogonales*, Fund. Math. **10**, (1927) 375–420
- [52] S. Molchanov, unpublished
- [53] S. Molchanov, M. Novitskii and B. Vainberg, *First KdV integrals and absolutely continuous spectrum for 1-D Schrödinger operator*, Comm. Math. Phys. **216** (2001), 195–213
- [54] S. Molchanov and B. Vainberg, *Schrödinger operator with matrix potentials. Transition from the absolutely continuous to the singular spectrum*, J. Funct. Anal. **215** (2004), 111–129
- [55] C. Muscalu, T. Tao and C. Thiele, *A counterexample to a multilinear endpoint question of Christ and Kiselev*, Math. Res. Lett. **10** (2003), 237–246
- [56] S. N. Naboko, *Dense point spectra of Schrödinger and Dirac operators*, Theor.-math. **68** (1986), 18–28
- [57] R. E. A. C. Paley, *Some theorems on orthonormal functions*, Studia Math. **3** (1931) 226–245
- [58] D. Pearson, *Singular continuous measures in scattering theory*, Comm. Math. Phys. **60** (1978), 13–36
- [59] G. Perelman, *On the absolutely continuous spectrum of multi-dimensional Schrödinger operators*, Int. Math. Res. Not. 2005, no. 37, 2289–2313
- [60] M. Reed and B. Simon, *Methods of Modern Mathematical Physics, III. Scattering Theory*, Academic Press, London–San Diego, 1979
- [61] C. Remling, *A probabilistic approach to one-dimensional Schrödinger operators with sparse potentials*, Comm. Math. Phys. **185** (1997), 313–323
- [62] C. Remling, *The absolutely continuous spectrum of one-dimensional Schrödinger operators with decaying potentials*, Comm. Math. Phys. **193** (1998), 151–170
- [63] C. Remling, *Schrödinger operators with decaying potentials: some counterexamples*, Duke Math. J. **105** (2000), 463–496
- [64] C. Remling, *Bounds on embedded singular spectrum for one-dimensional Schrödinger operators*, Proc. Amer. Math. Soc. **128** (2000), 161–171
- [65] I. Rodnianski and W. Schlag, *Classical and quantum scattering for a class of long range random potentials*, Int. Math. Res. Not. **5** (2003), 243–300
- [66] A. Rybkin, *On the spectral L^2 conjecture, 3/2-Lieb–Thirring inequality and distributional potentials*, J. Math. Phys. **46** (2005), 123505
- [67] O. Safronov, *The spectral measure of a Jacobi matrix in terms of the Fourier transform of the perturbation*, Ark. Mat. **42** (2004), 363–377
- [68] O. Safronov, *On the absolutely continuous spectrum of multi-dimensional Schrödinger operators with slowly decaying potentials*, Comm. Math. Phys. **254** (2005), 361–366
- [69] B. Simon, *Some Schrödinger operators with dense point spectrum*, Proc. Amer. Math. Soc. **125** (1997), 203–208
- [70] B. Simon, *Bounded eigenfunctions and absolutely continuous spectra for one-dimensional Schrödinger operators*, Proc. Amer. Math. Soc. **124** (1996), 3361–3369
- [71] B. Simon, *Schrödinger operators in the twenty-first century*, Mathematical Physics 2000 (A. Fokas, A. Grigoryan, T. Kibble and B. Zegarlinski, eds.), pp. 283–288, Imperial College Press, London, 2000

- [72] B. Simon, *A canonical factorization for meromorphic Herglotz functions on the unit disk and sum rules for Jacobi matrices*, J. Funct. Anal. **214** (2004), 396–409
- [73] B. Simon, *Orthogonal Polynomials on the Unit Circle, Part 1: Classical Theory*, AMS Colloquium Publications Series, Vol. **54**, Providence, RI, 2005
- [74] G. Stolz, *Bounded solutions and absolute continuity of Sturm–Liouville operators*, J. Math. Anal. Appl. **169** (1992), 210–228
- [75] G. Szegő, *Orthogonal polynomials*, AMS, Providence, RI, 1975
- [76] J. Weidmann, *Zur Spektraltheorie von Sturm–Liouville Operatoren*, Math. Z. **98** (1967), 268–302
- [77] J. von Neumann and E. P. Wigner, *Über merkwürdige diskrete eigenwerte*, Z. Phys. **30** (1929), 465–467
- [78] D. A. W. White, *Schrödinger operators with rapidly oscillating central potentials*, Trans. Amer. Math. Soc. **275** (1983), 641–677
- [79] D. Yafaev, *Scattering Theory: Some Old and New Results*, Lecture Notes in Mathematics **1735**, Springer, 2000
- [80] A. Zygmund, *A remark on Fourier transforms*, Proc. Camb. Phil. Soc. **32** (1936), 321–327
- [81] A. Zygmund, *Trigonometric Series*, 3rd edition, Cambridge University Press, Cambridge, 2002

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WISCONSIN, MADISON, WI 53706, U.S.A.

E-mail address: denissov@math.wisc.edu; kiselev@math.wisc.edu

On the Formation of Gaps in the Spectrum of Schrödinger Operators with Quasi-Periodic Potentials

Michael Goldstein and Wilhelm Schlag

Dedicated to Barry Simon on the occasion of his 60th birthday

ABSTRACT. In this article we review some recent developments in the theory of Schrödinger operators with quasi-periodic potentials on the discrete line. We focus in particular on the work by the authors on the formation of a dense set of gaps in the spectrum of such operators with general analytic potentials, provided the Lyapunov exponent is positive.

CONTENTS

1. Introduction
 2. The IDS and Cantor Spectrum for General Potentials
 3. A Finite-Volume Mechanism for Anderson Localization
 4. Elimination of Bad Phases and Energies
- References

1. Introduction

In this brief review we consider the class of operators defined as

$$(H_\theta\psi)(n) = -\psi_{n+1} - \psi_{n-1} + \lambda V(\theta + n\omega)\psi_n \tag{1}$$

with $n \in \mathbb{Z}$ and $\omega \in \mathbb{T}$ Diophantine or, in the continuum,

$$(H_\theta\psi)(x) = -\psi''(x) + \lambda V(\theta + x\omega)\psi(x) \tag{2}$$

with $x \in \mathbb{R}$ and $\omega \in \mathbb{T}^\nu$, $\nu \geq 2$, Diophantine. In both cases we shall assume that V is a real-analytic function on a suitable torus. An important special case is the almost Mathieu operator (also known as Harper's operator) for which $V(\theta) = \cos(2\pi\theta)$. Although these operators H_θ depend on a parameter, it is a basic fact that their

2000 *Mathematics Subject Classification.* 82B44 (60H25 81Q10).

Key words and phrases. quasi-periodic potentials, Lyapunov exponents, spectral gaps.

The first author was supported by an NSERC grant. The second author was partially supported by NSF grant DMS-0617854.

spectrum $\sigma(H_\theta)$ and its spectral parts, i.e., the absolutely continuous (a.c.), singular continuous (s.c.), and pure point (p.p.) parts, are *deterministic*. This means that there are fixed compact sets which equal these respective parts of the spectrum for a.e. θ . In the case of (1) this follows from the ergodic theorem and the identity $H_{T\theta} = U^{-1} \circ H_\theta \circ U$, where $T\theta = \theta + \omega \pmod{\mathbb{Z}}$ and U is the left shift.

We now recall some of the advances in this field:

- *Dinaburg–Sinai* (1975): For (2) and a large set of large energies, there are Bloch–Floquet waves $\psi_{\theta,E}(x) = a_{\theta,E}(x)e^{ik(\theta,E)x}$ where $a_{\theta,E}$ is quasi-periodic, $H_\theta\psi_{\theta,E} = E\psi_{\theta,E}$. Their argument is perturbative off the periodic case and they use KAM and Floquet theory.
- *Avron–Simon, Craig–Simon* (1981–1984): Discovered some of the most basic general properties of Schrödinger operators with almost periodic potentials. For example, they established purely s.c. spectrum for Liouville rotation numbers and the existence of Cantor spectrum for limit-periodic potentials. Furthermore, they gave a rigorous derivation of the Thouless formula which provides a relation between the Lyapunov exponent $L(E)$ and the integrated density of states $k(E)$, viz.,

$$L(E) = \int \log |E - E'| dk(E').$$

Here the Lyapunov exponent is defined to be

$$L(E) = \lim_{N \rightarrow \infty} N^{-1} \int_0^1 \log \|M_N(\theta, \omega, E)\| d\theta$$

where

$$M_N(\theta, \omega, E) = \prod_{j=N}^1 \begin{bmatrix} \lambda V(T^j\theta) - E & -1 \\ 1 & 0 \end{bmatrix} \tag{3}$$

and the IDS is defined as the limiting distribution of the eigenvalues, viz.,

$$k(E) = \lim_{N \rightarrow \infty} \frac{1}{2N + 1} \#\{j : E_j^{(N)}(x, \omega) \leq E\}$$

where $E_j^{(N)}(x, \omega)$ are the eigenvalues of $H_{[-N,N]}(x, \omega)$. They observed that the IDS is log-Hölder continuous as a consequence of the Thouless formula. They also gave an exact argument for Aubry duality for the almost Mathieu case.

- *Johnson–Moser* (1982) and *Herman* (1983): Introduced a rotation number $\alpha(\lambda) := \lim_{x \rightarrow \infty} x^{-1} \arg(\phi + i\phi')(x, \lambda)$ for solutions of $H_\theta\phi(\cdot, \lambda) = \lambda\phi(\cdot, \lambda)$ with H_θ as in (2) (or more generally, with almost periodic potential). They also proved continuity of α and related the intervals of constancy of α to gaps in the spectrum; more precisely, the value of $2\alpha(\lambda)$ belongs to the frequency module of the almost periodic potential. Moreover, α is the same as the IDS $k(E)$ up to a factor of π .
- *Avron–van Mouche–Simon* (1990): For the almost Mathieu case and in the limit over periodic approximants, they established that $|\sigma(H_\theta)| = |4 - 2|\lambda||$ if $|\lambda| \neq 2$. *Last* then obtained the full theorem without the limit over periodic approximants and also for $|\lambda| = 2$. Thus, for that case he showed that the spectrum has measure zero. In particular, the spectrum is nowhere dense.

- *Eliasson* (1992): Again for the case of (2), Floquet–Bloch solutions exist for a.e. large E and $\sigma(H_\theta)$ is purely a.c.. More precisely, he established reducibility: the Schrödinger cocycle is conjugate to a constant one for a.e. energy. Moreover, for generic potential the spectrum is a Cantor set.
- *Jitomirskaya, Last, Simon* (1994, 1997): Showed that the p.p. and s.c. parts of the spectrum are not constant; rather, only a.e. constant, see above. Moreover, the a.c. part is constant in the phase θ (this also due to Kotani; see the following bullet). There is an open conjecture due to Simon that the union of the s.c. and p.p. spectrum is constant.
- *Kotani* (1984–1997): Showed that the closure of the set of positivity of the Lyapunov exponent equals the essential support of the a.c. part of the spectral measure of an ergodic operator (2) (*Simon* extended these results to the discrete case (1))—thus proving the converse of the *Ishii–Pastur* theorem. Furthermore, Kotani introduced a natural notion of “determinacy” of an ergodic potential V_ω and showed that the Lyapunov exponent is always positive (and thus there is no a.c. spectrum) unless V_ω is deterministic in his sense. This body of techniques, which is based on the Weyl–Titchmarsh m -function, has become known as “Kotani theory.”
- *Gordon–Jitomirskaya–Last–Simon* (1997): Refined the $|\lambda| = 2$ analysis for the almost Mathieu operator and established purely s.c. spectrum.
- *Krikorian and Avila–Krikorian* (2000–2004): Recent work on reducibility; they established a dichotomy between *nonuniform hyperbolicity* (i.e., continuous Oseledec splitting) and *reducibility* for quasi-periodic Schrödinger cocycles with analytic potentials; their results hold for a.e. ω and a.e. energy. They obtain further that for the almost Mathieu operator with $|\lambda| = 2$ and Diophantine ω , the spectrum is measure zero and purely s.c. for a.e. phase. This completes the analysis of Last (see above).

An important question about the operator class (1) and (2) concerns **Anderson localization (AL)**. This means that the spectrum is pure point with exponentially decreasing eigenfunctions.

- *Fröhlich–Spencer–Wittwer, Sinai* (late 1980s): Established (AL) for cosine-like potentials and large disorder. Their arguments are perturbative, treating the operators as perturbations of the multiplication operator given by the potential.
- *Jitomirskaya* (1999): For almost Mathieu, $|\lambda| > 2$, as well as Diophantine ω , proved (AL). By a lower bound on the Lyapunov exponent due to *Avron–Simon*, as well as *Herman*, $|\lambda| > 2$ implies that $\inf_E L(E) > 0$. Moreover, by Aubry duality and the *Ishii–Pastur* theorem, this fails when $|\lambda| \leq 2$. Hence, this result shows that (AL) holds precisely when $L(E) > 0$ for all energies.
- *Bourgain, Goldstein* (2000): Established (AL) for analytic potentials, positive Lyapunov exponents, and almost all ω . They introduced large deviation theorems and the use of semi-algebraic sets (for the elimination of double resonances) into this field.

Finally, we turn to another important question, namely: **When is the spectrum a Cantor set?** This simply means that it is nowhere dense (a more refined version of this statement is whether all gaps allowed by the gap labeling theorem are open). So far, this question had apparently only been considered for the *almost*

Mathieu case. We now list a series of results that answer this question affirmatively in the almost *Mathieu* case.

- *Bellissard–Simon* (1982): For a dense G_δ set of λ, ω .
- *Choi–Elliott–Yui* (1990): Liouville case, $|\lambda| = 2$ using methods of C^* algebras; their proof also applies to all $\lambda \neq 0$.
- *Puig* (2003): $|\lambda| \neq 2$, ω Diophantine. His argument proceeds by contradiction. He shows that eigenvalues of H with phase 0 are necessarily the endpoints of open gaps.
- *Avila–Jitomirskaya* (2005): For $|\lambda| \neq 0$, bridge the gap between the Liouville and Diophantine cases. Thus, they show that whenever the potential is nonperiodic (i.e., for irrational ω) there is a dense set of gaps. Thus, they provide a solution of the so-called Ten Martini problem. As in Puig’s argument, the proof again proceeds by contradiction.
- The case $|\lambda| = 2$ is covered by the previous reference, but had also been settled earlier by *Last* and *Avila–Krikorian* (see above).

2. The IDS and Cantor Spectrum for General Potentials

We now discuss some recent work by the authors. We will in general emphasize ideas over technical correctness. For a review which is much more technical by design, we refer to reader to [29]. However, this reference does not contain any material about gaps. Rather, it reviews the long paper [27] which develops the machinery needed for the formation of gaps in [28]. This paper is intended as an exposition of the main ideas needed to pass from [27] to [28].

We remark that unless indicated otherwise, Diophantine henceforth means that, for all $n \geq 1$,

$$\|n\omega\| \geq \frac{c}{n(\log n)^a}$$

with $c > 0$ and $a > 1$. The following is proved in [26, 27, 28].

THEOREM 1. *Consider (1) with $\lambda = 1$. Let $V : \mathbb{T} \rightarrow \mathbb{R}$ be analytic and suppose that*

$$\inf_{E, \omega} L(E, \omega) > 0.$$

Then the following properties hold:

- (1) *If ω is Diophantine, then the IDS is Hölder continuous. If V is in a small L^∞ neighborhood of a trigonometric polynomial of degree k , then the IDS is Hölder $(2k + \varepsilon)^{-1}$ -continuous for all $\varepsilon > 0$.*
- (2) *For a.e. ω , the following holds: Off a set of Hausdorff dimension zero the IDS is Lipschitz continuous.*
- (3) *For a.e. ω , the IDS is absolutely continuous.*
- (4) *For a.e. ω , the spectrum is a Cantor set.*

Further results include:

- estimates on the distribution of the zeros of

$$\det(H_{[-N, N]}(z, \omega) - E)$$

in the z -plane

- a quantitative separation property of the eigenvalues of $H_{[-N, N]}(z, \omega)$
- a constructive, finite-volume mechanism for the development of Anderson localization

- a constructive finite-volume mechanism for the formation of gaps based on resonances

We now give an example of some finite-volume statements which imply, and are finer than, the corresponding statements concerning the IDS in the previous theorem; they are obtained in [27] and address the problem of bounding the expected number of eigenvalues at a finite scale falling into a small interval.

THEOREM 2. *Let $E_j^{(N)}(x, \omega)$ be the eigenvalues of $H_{[-N, N]}(x, \omega)$. Then*

- *For any $\eta > N^{-1+\delta}$, $\delta > 0$, and ω Diophantine, there is the following bound on the expected number of eigenvalues falling into small intervals: for arbitrary E , $\kappa > 0$ and large N ,*

$$\int_0^1 \#\{j : |E_j^{(N)}(x, \omega) - E| < \eta\} dx \leq N\eta^{\frac{1}{2k} - \kappa}$$

where k is the degree of the underlying trigonometric polynomial.

- *Let $\varepsilon > 0$ be arbitrary but fixed. If $\omega \notin \Omega(\varepsilon)$, $E \notin \mathcal{E}_\omega(\varepsilon)$, then*

$$\int_0^1 \#\{j : |E_j^{(N)}(x, \omega) - E| < \eta\} dx < \exp((\log \varepsilon^{-1})^A) N\eta$$

where $|\Omega(\varepsilon)| + \sup_{\omega \notin \Omega(\varepsilon)} |\mathcal{E}_\omega(\varepsilon)| < \varepsilon$.

According to Yakov Sinai, “Anderson localization is a game of resonances”.

A resonance here means the following: Let $\Lambda, \Lambda' \subset \mathbb{Z}$ be intervals, typically of comparable length and separated by more than their length. Then the Hamiltonians $H_\Lambda(x, \omega)$ and $H_{\Lambda'}(x, \omega)$ are said to be in resonance provided their spectra are very close.

In order for this concept to be useful, it needs to be quantified. Note that for any Λ , one can of course find $\Lambda' = \Lambda + k$ (a translate) for which the distance between the spectra of $H_\Lambda(x, \omega)$ and $H_{\Lambda'}(x, \omega)$ is arbitrarily small. This follows simply from recurrence of the rotation map. Therefore, the point will be to apply this definition inside a fixed box Λ_0 of a given size and then look for (much) smaller $\Lambda, \Lambda' \subset \Lambda_0$ which produce resonances (where the notion of distance between the spectra needs to be adjusted to the length of Λ_0 —typically $\ll |\Lambda_0|^{-C}$ with C large). Of particular importance is to keep track of how many smaller intervals Λ' there are for a given $\Lambda \subset \Lambda_0$ which produce resonances with Λ and also satisfy $\text{dist}(\Lambda', \Lambda) \gg |\Lambda|$. If there is exactly one such Λ' , then one speaks of a *double resonance*, otherwise of *triple or higher order resonance*. It is particularly important to eliminate the occurrence of such higher order resonances for Sinai [47], Fröhlich–Spencer–Wittwer [23], Bourgain [7], as well as the authors’ argument for gap formation [28]. In the former three references this is accomplished perturbatively, by assuming that the potential has no more than two monotonicity intervals; in [28] we proceed nonperturbatively and no assumption other than analyticity is made on the potential—in order to prevent triple resonances it is necessary to eliminate some small set of ω and E (this process of elimination is essentially the one of Chan [12]).

We remark that it is also customary to speak of eigenfunctions of $H_\Lambda(x, \omega)$ and $H_{\Lambda'}(x, \omega)$ to be in resonance. This simply means that these eigenfunctions

correspond to close eigenvalues. At the larger scale, such eigenfunctions “merge” to form a two-humped eigenfunction of $H_{\Lambda_0}(x, \omega)$, cf. Figure 9.

To illustrate this concept, let us recall the KAM-type scheme used in Fröhlich–Spencer–Wittwer’s as well as Sinai’s perturbative approach to localization. In it, one starts off with singletons Λ, Λ' : consider the 2×2 matrix

$$A(x) = \begin{bmatrix} V(x) & \lambda^{-1} \\ \lambda^{-1} & V(x + \omega) \end{bmatrix}$$

where $V(x_0) = V(x_0 + \omega)$ and $V'(x_0) < 0, V'(x_0 + \omega) > 0$. Its eigenvalues $E^\pm(x)$ for x close to x_0 form two branches with

$$E^\pm(x_0) = V(x_0) \pm \lambda^{-1}.$$

Figure 1 shows how the two separate branches of the potential (which are the diagonal entries of $A(x)$) separate when they yield the eigenvalues of $A(x)$. The dashed line is supposed to have length exactly equal to ω . The projections of the intersection points of this dashed line with the graphs are x_0 and $x_0 + \omega$, respectively.

If V has two monotonicity intervals and ω is Diophantine, then it is easy to see that this 2×2 block determines the invertibility of the $N \times N$ Hamiltonian close to x_0 provided λ is large (however, this largeness depends on N). More precisely, it follows from the Feshbach formula that

$$\det(H_N(x, \omega) - E)$$

$$= \begin{vmatrix} \lambda V(x) - E & -1 & 0 & \cdots & \cdots & 0 \\ -1 & \lambda V(x + \omega) - E & -1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \\ 0 & \dots\dots\dots & 0 & -1 & \lambda V(x + (N - 1)\omega) - E & -1 \end{vmatrix}$$

$$= F(x, \lambda, E) \det A(x)$$

where $F(x, \lambda, E) \neq 0$ for all x close to x_0 . It turns out that this fact can be considered as the 0-order step in a KAM scheme. To pass to bigger scales one again uses the Feshbach formula using the previous scale as the information needed to invert the larger blocks.

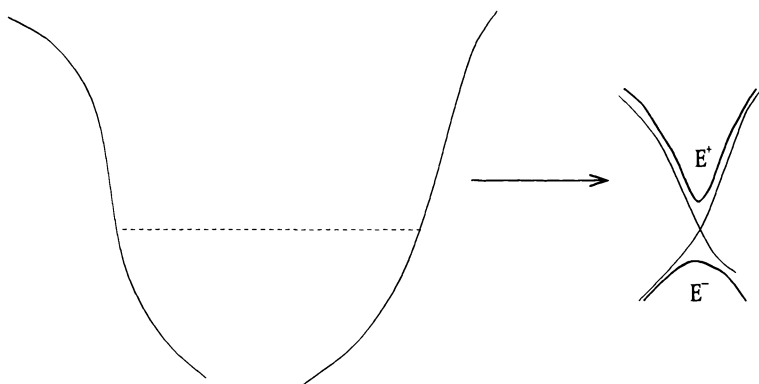


FIGURE 1. A double resonance

Sinai (1987) and Fröhlich–Spencer–Wittwer (1990) carried out versions of this scheme. Later, Bourgain (2001) obtained in this way Hölder $\frac{1}{2} - \varepsilon$ continuity of the IDS for cosine-like potentials and large disorder. The point of the exponent $\frac{1}{2}$ here is the following: the obstruction to invertibility of a finite volume Hamiltonian always arises in the form of a second order polynomial. In the proof of Theorems 1 and 2, which are of course nonperturbative, it will also be of crucial importance to obtain such fixed-degree control. Note that the degree of these polynomial will vary with the degree k of the potential (for $k = 1$, it will again be quadratic and more generally, be of degree $2k$; hence the Hölder exponent $(2k)^{-1}$ of the IDS). To extract these polynomials of fixed degree we use a Weierstrass preparation theorem. Finally, we note that in the free case the IDS is no better than Hölder $\frac{1}{2}$ -continuous at the edges of the spectrum.

3. A Finite-Volume Mechanism for Anderson Localization

Because of the self-adjointness of H and the analyticity of V , the solutions of

$$0 = \det (H_{[1,N]}(x, \omega) - E).$$

are real-analytic functions (which we refer to as *Rellich functions*)

$$E_1(x, \omega) < E_2(x, \omega) < \dots < E_N(x, \omega)$$

One cannot have equality here since the eigenvalues of the Dirichlet problem are simple. Figure 2 displays the graphs of such Rellich functions for the periodic Mathieu operator with periodic boundary conditions. The interval is $[1, N = 12]$ where the latter number arises as denominator of a convergent of $\omega = \sqrt{2}$. In Figure 3 the same is plotted for the almost Mathieu operator with $\omega = \sqrt{2}$ and Dirichlet conditions at the boundary. Although the graphs appear to intersect at many points, they actually separate at these points, as can be seen in a much finer resolution. Finally, Figure 4 shows the Rellich functions for a potential given by a third degree polynomial again with the same ω but the next larger denominator in the sequence of convergents.

At this point it seems natural to ask some **basic quantitative questions**:

- What is the width of the strip around \mathbb{R} to which $E_j(z, \omega)$ can be analytically continued?
- What is the size of the separation between the $E_j(x, \omega)$?
- What is a reasonable lower bound on the absolute values of the slopes of the $E_j(x, \omega)$ away from critical points?

The answer turns out to be e^{-N^δ} provided $\omega \in \mathbb{T} \setminus \Omega_N$, $E_j(x, \omega) \notin \mathcal{E}_N(\omega)$. Here Ω_N , $\mathcal{E}_N(\omega)$ have small measure and complexity (the latter refers to the number of connected components a set needs to have to cover $\mathcal{E}_N(\omega)$ without increasing the measure significantly—by a multiplicative constant, say). More precisely, we have the bounds

$$|\Omega_N| < \exp(-(\log N)^{2A}), \quad \text{compl}(\Omega_N) < \exp((\log N)^A)$$

and similarly for $\mathcal{E}_N(\omega)$. Note that these bounds reflect that the “bad sets” have Hausdorff dimension zero.

Central to the separation of the $E_j(x, \omega)$ is a *finite-volume understanding of (AL)*, which we now describe: Let $H_{[-N,N]}(x, \omega)\psi = E\psi$ with $\|\psi\|_2 = 1$. We seek a window $\Lambda_0 \subset [-N, N]$ so that $\|\psi\|_{\ell^2(\Lambda_0)} = 0.999$, say, and $|\Lambda_0| \ll N^\varepsilon$. Consider

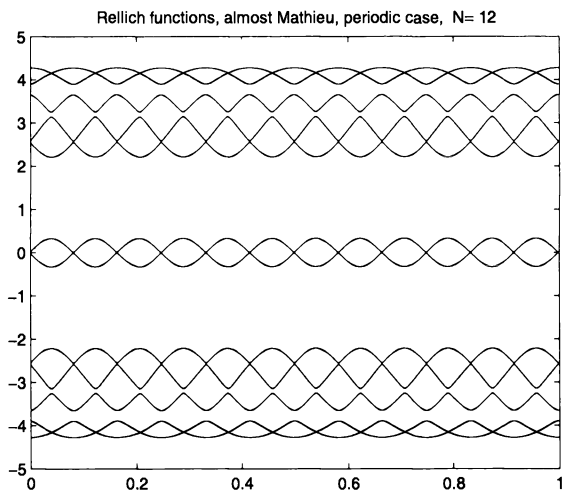


FIGURE 2. Rellich functions in the periodic case

all intervals $\Lambda \subset [-N, N]$ with $|\Lambda| = n \sim (\log N)^C$: there exists such an interval Λ so that with $f_\Lambda(x, \omega, E) = \det(H_\Lambda(x, \omega) - E)$,

$$\log |f_\Lambda(x, \omega, E)| < |\Lambda|L(\omega, E) - |\Lambda|^{\frac{1}{2}}.$$

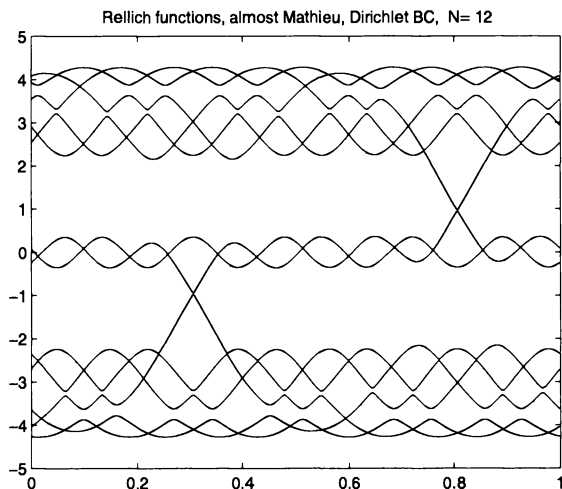


FIGURE 3. Rellich functions in the aperiodic case I

Otherwise each Green's function $G_\Lambda(x, \omega, E)$ would exhibit exponential off-diagonal decay. Indeed, by Cramer's rule,

$$\left| (H_{[-N,N]}(x, \omega) - E)^{-1}(k, m) \right| = \frac{|f_{[-N,k]}(e(x), \omega, E)| |f_{[m+1,N]}(e(x), \omega, E)|}{|f_{[-N,N]}(e(x), \omega, E)|}.$$

In [27] the following uniform upper bounds are proved:

$\log |f_{[-N,k]}(e(x), \omega, E)| + \log |f_{[m+1,N]}(e(x), \omega, E)| \leq 2NL(\omega, E) + (\log N)^{A_1}$
 for any $x \in \mathbb{T}$. Therefore,

$$\| (H_{[-N,N]}(x, \omega) - E)^{-1} \| \leq N^2 \frac{\exp(2NL(\omega, E) + (\log N)^A)}{|f_{[-N,N]}(e(x), \omega, E)|}$$

for any $x \in \mathbb{T}$. Moreover, one obtains the aforementioned off-diagonal decay in this way. This latter property, however, would force ψ to be very small everywhere in contradiction to the fact that $\|\psi\|_2 = 1$.

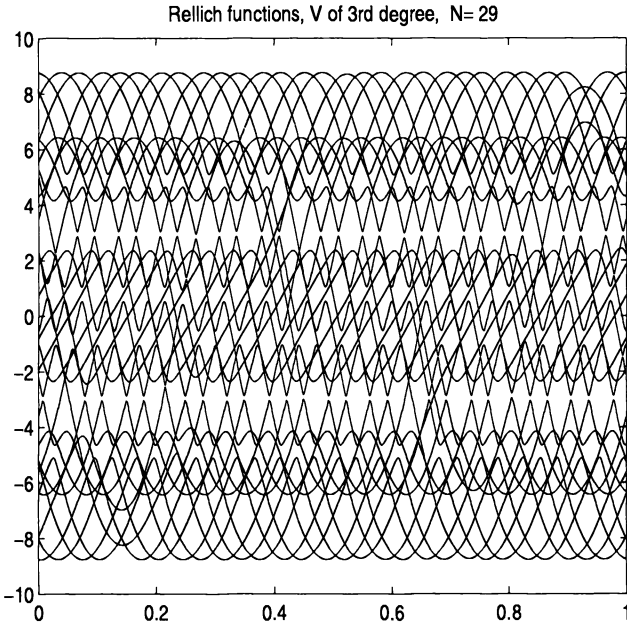


FIGURE 4. Rellich functions in the aperiodic case II

We shall now outline three main points related to the phenomenon of Anderson localization and the property of separation of the eigenvalues:

- **Point 1:** By eliminating bad ω and E (i.e., $\omega \notin \Omega_N, E \notin \mathcal{E}_N(\omega)$) we can ensure that

$$\log |f_{\Lambda'}(x, \omega, E)| \geq |\Lambda'|L(\omega, E) - |\Lambda'|^{\frac{1}{2}} \tag{4}$$

for all $|\Lambda'| \sim |\Lambda|, \text{dist}(\Lambda', \Lambda) > N^\epsilon$. This will be obtained by means of a reduction to a statement about close zeros of two determinants; see below.

- **Point 2:** From the avalanche principle (see Lemma 1 below), conclude that (with $L(E, \omega) > \gamma > 0$)

$$|\psi(n)| \leq \exp(-\gamma \text{dist}(n, \Lambda_0) + N^\epsilon).$$

This step can be considered to be in the spirit of the multi-scale analysis of Fröhlich–Spencer, albeit with the avalanche principle instead of the resolvent identity.

- **Point 3:** Suppose

$$H_{[-N, N]}(x, \omega)\psi = E\psi, \quad H_{[-N, N]}(x, \omega)\tilde{\psi} = \tilde{E}\tilde{\psi}$$

with $\|\psi\|_2 = \|\tilde{\psi}\|_2 = 1$ and $0 < |E - \tilde{E}| < e^{-N^\delta}$, and $\omega \notin \Omega_N$, $E, \tilde{E} \notin \mathcal{E}_N(\omega)$. Then provided $\epsilon \ll \delta$,

$$\|\psi - \tilde{\psi}\|_2 < |E - \tilde{E}|e^{N^\epsilon} \ll 1.$$

Clearly, this is a contradiction to $\psi \perp \tilde{\psi}$ (note that this step requires self-adjointness of the Hamiltonian). So

$$|E - \tilde{E}| > e^{-N^\delta}.$$

This step requires knowing that (4) holds, not just localization; in fact, we need to use that $\psi(n) = \mu(x, \omega, E) \cdot f_{[-N, n-1]}(x, \omega, E)$ for all $-N \leq n \leq N$. It is not surprising that ψ and $\tilde{\psi}$ are close if E and \tilde{E} are close; indeed, this can be thought of as continuity in E or differentiability in E . Of course, the differentiation has to be carried out on determinants of a smaller scale and not the large-scale determinants. This is achieved by means of the avalanche principle and the fact that the small-scale Hamiltonians are not in resonance with those inside the window of localization.

We now recall the Avalanche Principle (AP) from [26] and [27]. It is a purely deterministic statement.

LEMMA 1 (Avalanche Principle). *Let A_1, \dots, A_n be 2×2 matrices whose determinants satisfy*

$$\max_{1 \leq j \leq n} |\det A_j| \leq 1.$$

Suppose that

$$\begin{aligned} \min_{1 \leq j \leq n} \|A_j\| &\geq \mu > n \\ \max_{1 \leq j < n} [\log \|A_{j+1}\| + \log \|A_j\| - \log \|A_{j+1}A_j\|] &< \frac{1}{2} \log \mu. \end{aligned}$$

Then

$$\left| \log \|A_n \cdots A_1\| + \sum_{j=2}^{n-1} \log \|A_j\| - \sum_{j=1}^{n-1} \log \|A_{j+1}A_j\| \right| < C \frac{n}{\mu}$$

with some absolute constant C .

4. Elimination of Bad Phases and Energies

We now return to Point 1 from above in an attempt to explain some of the underlying issues, especially the need for elimination of bad ω and E . In order to obtain our window of localization, we need to exclude resonances. The latter here

refers to the situation where for some energy E there exists a phase x_0 for which we have two small determinants:

$$\begin{aligned} \log |f_\Lambda(x_0, \omega, E)| &< |\Lambda|L(\omega, E) - |\Lambda|^{\frac{1}{2}} \\ \log |f_{\Lambda'}(x_0, \omega, E)| &< |\Lambda'|L(\omega, E) - |\Lambda'|^{\frac{1}{2}} \end{aligned} \tag{5}$$

where $\Lambda, \Lambda' \subset [-N, N]$, $\text{dist}(\Lambda, \Lambda') \gg |\Lambda| = |\Lambda'| = n$. This can of course happen; indeed, if we formally “eliminate” the phase x from these inequalities, then we obtain a condition on (ω, E) that describes a set of bad parameters that need to be removed. Using large deviation theorems for the determinants (see (9) below), as well as Jensen’s formula (7) from complex analysis, one can show that (5) implies the following: there exist z_1, z_2 with $|z_1 - z_2| + \sum_j |\text{Im}z_j| < e^{-n^{\frac{1}{4}}}$ and

$$f_{[1,n]}(z_1, \omega, E) = f_{[1,n]}(z_2 + t\omega, \omega, E) = 0. \tag{6}$$

Hence, we are required to exclude close zeros of two such determinants which we do by means of the method of resultants (we will return to this issue below). We now recall Jensen’s formula from complex analysis¹: if f is analytic on $|z - z_0| \leq r$, then

$$\int_0^1 \log |f(z_0 + re^{2\pi i\theta})| d\theta = \log |f(z_0)| + \sum_{\zeta: f(\zeta)=0} \log \frac{r}{|\zeta - z_0|} \tag{7}$$

In order to use this to pass from (5) to (6) we simply need to show that the sum on the right-hand side cannot vanish for $r = e^{-n^{\frac{1}{4}}}$; or, in other words, that the difference

$$\int_0^1 \log |f_{[1,n]}(x_0 + re^{2\pi i\theta}, \omega, E)| d\theta - \log |f_{[1,n]}(x_0, \omega, E)| \neq 0.$$

In view of (5) this is accomplished by showing that with $e^{-n} < r \leq n^{-1}$, the integral satisfies

$$\int_0^1 \log |f_{[1,n]}(x_0 + re^{2\pi i\theta}, \omega, E)| d\theta = nL(E, \omega) + O((\log n)^C) \log \frac{1}{r}.$$

This in turn follows from the following estimate, which we call *large deviation theorem (LDT) for the determinants*; see [27]:

$$|\{x \in \mathbb{T} : \log |f_{[1,n]}(x + iy, \omega, E)| < nL(\omega, E) - h\}| < e^{-ch/(\log n)^C} \tag{8}$$

uniformly in $|E| \leq C$ and $|y| \leq n^{-1}$. Figure 5 illustrates this bound for the case of $\log \|M_N(x, \omega, E)\|$ instead of the determinant, with M_N as in (3), $N = 100$ and $V = \cos$. The picture displays the self-similar nature of this function together with its subharmonic features: there are large deviations in the direction of small values, but for large values the function looks relatively “flat.” We remark that the (LDT) for M_N goes back to [9] and [26], whereas the case of the determinant was established in [27].

¹We use this device in [27] repeatedly. However, since it is hard to work with a fixed z_0 we are forced to average over this point as well; this is the origin of the double Jensen averages in that paper. Here, however, it suffices to freeze z_0 . See also the review [29] for these matters.

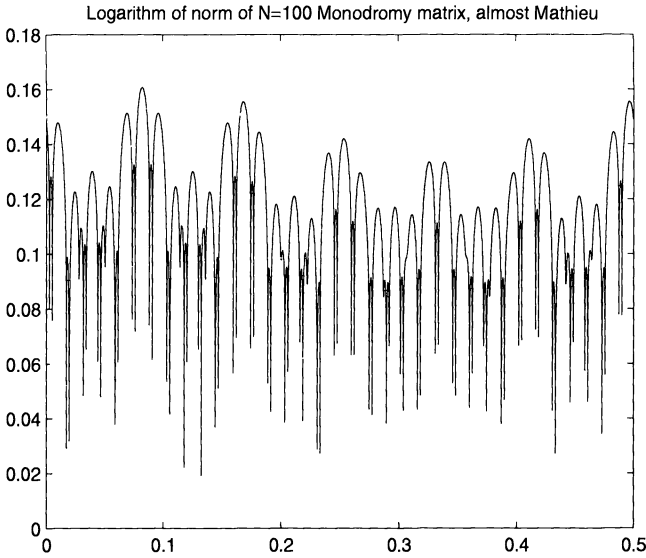


FIGURE 5. The logarithm of a monodromy matrix

By the John–Nirenberg inequality (see Garnett [25]), the bound (8) is equivalent with the statements that

$$\| \log |f_{[1,n]}(\cdot + iy, \omega, E)| \|_{\text{BMO}} \leq (\log n)^C \tag{9}$$

and

$$\int_0^1 \log |f_{[1,n]}(x + iy, \omega, E)| dx = nL(E, \omega) + (\log n)^C \tag{10}$$

where (10) is uniform in $|y| \leq n^{-1}$. We remark that these estimates imply that

$$\sup_{x \in \mathbb{T}} \log |f_{\Lambda'}(x, \omega, E)| < nL(\omega, E) + (\log n)^C \tag{11}$$

via the sub-mean property of subharmonic functions. Another immediate consequence via Jensen’s formula is the following bound

$$\#\{\zeta \in D(x_0, n^{-1}) : f_{[1,n]}(\zeta, \omega, E) = 0\} \leq (\log n)^C \tag{12}$$

for all x_0, E and Diophantine ω . These four facts (9)–(12) are basic to our entire analysis; see [27] and [29]. We remark that the large deviation estimate, and thus the BMO bound (9), are a reflection of that fact that the zeros of $f_{[1,n]}(\cdot, \omega, E)$ are uniformly distributed.

To see this, consider the following classical result of *Erdős–Turan*: Let $\{\zeta_j\}_{j=1}^N \subset \mathbb{T}$ be a collection of N points on the circle. Consider the polynomial

$$P(z) = \prod_{j=1}^N (z - \zeta_j).$$

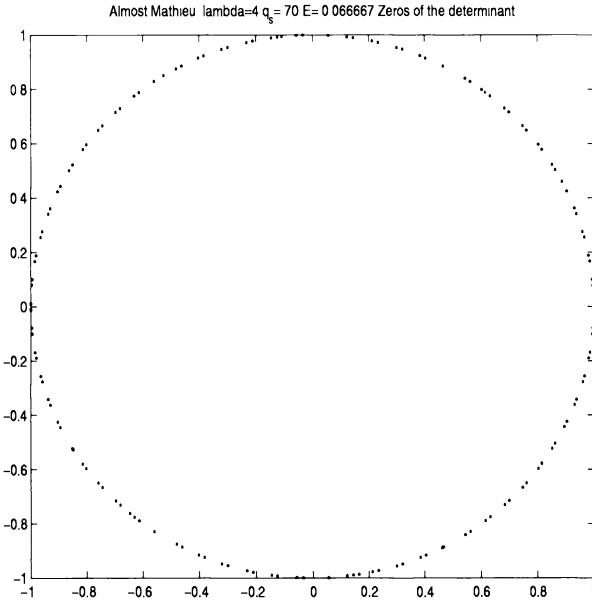


FIGURE 6. Zeros I

Then

$$\sup_{z \in \mathbb{T}} |P(z)| < e^\tau \Rightarrow \|\log |P(e(\cdot))|\|_{\text{BMO}} \sim D_N(\{\zeta_j\}) < \sqrt{N\tau}$$

where

$$D_N(\{\zeta_j\}) = \sup_{I \subset \mathbb{T}} |\#\{j : \zeta_j \in I \pmod{1}\} - N|I||$$

is the usual discrepancy. There is the following analogue for *subharmonic functions*: Let $u : \mathcal{A} \rightarrow \mathbb{R}$ be subharmonic with $\mathbb{T} \subset \mathcal{A} \subset \mathbb{C}$ an annulus of width $\frac{1}{2}$, say. Then the Riesz representation theorem (see Levin [43]) yields that

$$u(z) = \int \log |z - \zeta| d\mu(\zeta) + h(z)$$

with $\mu \geq 0$ and h harmonic. Suppose $\mu(\mathbb{C}) \leq N$. Then the analogue of the result of Erdős–Turan is the following:

$$\|u(e(\cdot))\|_{\text{BMO}} \lesssim \sqrt{N[\sup_{x \in \mathbb{T}} u(e(x)) - \langle u \rangle]}.$$

However, this is insufficient for our purposes (it is inconsistent in the sense that the supremum bound (11) does not imply the BMO bound (9)). Luckily, it can be improved (see Bourgain–Goldstein–Schlag [10] and [8, 29]): Write $u - \langle u \rangle = u_0 + u_1$ on \mathbb{T} . Then

$$\|u(e(\cdot))\|_{\text{BMO}} \lesssim \|u_0\|_\infty + \sqrt{N\|u_1\|_1}$$

It is easy to check that this bound is consistent with our estimates.

In order to obtain the estimates for the determinants which we just described, we need to reveal their “almost-invariance” under the shift. This can be done by

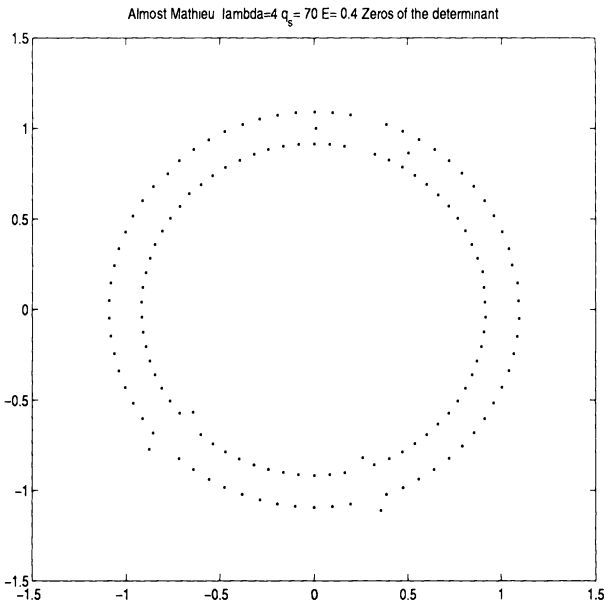


FIGURE 7. Zeros II

means of a “factorization of the determinant” via the avalanche principle. Indeed, write

$$\begin{aligned} \begin{bmatrix} f_N(x, \omega, E) & 0 \\ 0 & 0 \end{bmatrix} &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \prod_{j=N}^1 \begin{bmatrix} V(T^j x) - E & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \\ &= \prod_{k=n}^1 A_k(x, \omega, E) \end{aligned}$$

where each A_k is the product of about $(\log N)^C$ many factors. One needs to use the large deviation theorems for the determinants and monodromies on the small scale to conclude that the conditions of the avalanche principle hold. This requires the removal of a set of phases $x \in \mathbb{T}$ of measure $< \exp(-(\log N)^B)$. For the remaining good phases we conclude that:

$$\log |f_N(x, \omega, E)| = - \sum_{j=2}^{n-1} \log \|A_j\| + \sum_{j=1}^{n-1} \log \|A_{j+1}A_j\| + O(N^{-1000}).$$

Note that this resembles an ergodic average since most of the A_j can be chosen to be shifts in the phase of a fixed one. This is what we mean by “self-similar structure” of the determinants f_N . In Figures 6 and 7, we display two sets of zeros of the determinants in the almost Mathieu case. The first one is for an energy in the spectrum, whereas the second is for an energy outside of the spectrum. Observe that the zeros look approximately evenly distributed with the exception of a few “errant” ones. Loosely speaking, these are related to nonlocalized states in the same way that the “errant” segments of Rellich graphs crossing what appears to be a gap in Figures 3 and 4 correspond to nonlocalized states (see the gap containing energies $[-2, -1] \cup [1, 2]$ in Figure 3 as well as that around energy $E = 3$ in Figure 4).

Recall that we have only discussed so far how to pass from (5) to (6) but not how to exclude the latter. As indicated above, for this we need to eliminate close zeros. The following is proved in [27]:

LEMMA 2. *There exists $\Omega_n \subset \mathbb{T}$, $|\Omega_n| < e^{-n^{\frac{1}{10}}}$, so that for all $\omega \notin \Omega_n$, $t > e^{(\log n)^C}$ there exists $\mathcal{E}_{n,t,\omega} \subset \mathbb{C}$, $|\mathcal{E}_{n,t,\omega}| < e^{-n^{\frac{1}{10}}}$ so that*

$$f_{[1,n]}(z_1, \omega, E) = f_{[1,n]}(z_2 + t\omega, \omega, E) = 0 \Rightarrow |z_1 - z_2| > e^{-n^{\frac{1}{4}}}$$

$$\inf_{x \in \mathbb{T}} \text{dist}(\text{spec}(H_{[1,n]}(x, \omega)) \setminus \mathcal{E}_{n,t,\omega}, \text{spec}(H_{[1,n]}(x + t\omega, \omega))) > e^{-n^{\frac{3}{4}}}$$

This is only one of several statements one can prove in this direction; more precisely, the powers of n can be lowered to n^ϵ with $\epsilon > 0$ arbitrary. Further, we remark that one can also remove x instead of E by a Wegner estimate; see [27] and [29]. Finally, and crucially, we are also able to bound the complexity of the bad sets Ω_n and $\mathcal{E}_{n,t,\omega}$.

For the proof, we need to use the resultant of two polynomials $f(z), g(z)$. It is defined as

$$\text{Res}(f, g) = \prod_{j,k} (z_j - \zeta_k)$$

where $f(z_j) = 0, g(\zeta_k) = 0$. One can show that it is a *polynomial* in the coefficients of f, g (see [39]). Suppose

$$f(z, \omega, E) = z^\nu + a_{\nu-1}(\omega, E)z^{\nu-1} + \dots + a_1(\omega, E)z + a_0(\omega, E)$$

$$g(z, \omega, E) = z^\mu + b_{\mu-1}(\omega, E)z^{\mu-1} + \dots + b_1(\omega, E)z + b_0(\omega, E)$$

with a_j, b_k analytic in ω, E . It follows that Res is also analytic in ω, E . This allows one to use analytical methods to estimate Res from below, at least for most ω, E . Clearly, if we are able to bound the resultant from below, then we are also able to give a quantitative estimate on the separation of the zeros. Figure 8 depicts the algebraic curve $\text{Res}(\omega, E) = 0$ in the (ω, E) -plane. For most values of ω there will be only finitely many E -values on this curve, but we need to remove those ω for which there is a “flank”; this refers to the vertical or near vertical segments of the curve that would lead to a large set in energy for which the resultant is too small.

On a more technical, albeit crucial, point we remark that we cannot use resultants on the full determinants because of their large degree. Rather, we apply the Weierstrass preparation theorem on $f_{[1,n]}(z, \omega, E)$ before applying the resultants locally in ω : the factors which we pull out have very small degree $= (\log n)^C$ because of the control over zeros provided by (12). For more on this topic, see [27] or [29].

We now list some important consequences of the finite-volume Anderson localization and the quantitative separation of the eigenvalues:

- Use a Sard-type argument to conclude that the slopes of the Rellich functions $E_j^{(N)}(x, \omega)$ off a bad energy set $\mathcal{E}_N(\omega)$ are bounded below by e^{-N^δ} in absolute value. The Sard theorem is needed to remove the critical values of the Rellich functions. Particular care needs to be taken concerning the complexity of the resulting set of energies. Of course, it is essentially used in [27] that the Rellich functions are solutions of algebraic equations of controlled degrees.

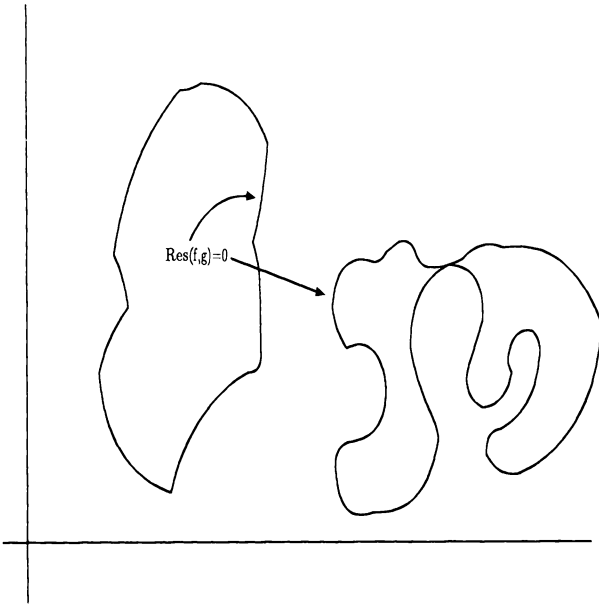


FIGURE 8. Zero set of the resultant

- Use (AL) to obtain almost shift-invariance of the Rellich graphs. Provided $E_j^{(N)}(x, \omega) \notin \mathcal{E}_N(\omega)$ and for most choices of $m \in [-N, N]$, there exists ℓ such that

$$|E_j^{(N)}(x, \omega) - E_\ell^{(N)}(x + m\omega, \omega)| < e^{-N^{\frac{1}{2}}}.$$

This follows from the following fact: if $H_{[-N, N]}(x, \omega)\psi = E\psi$ with ψ localized, then $\psi(n+m)$ is almost an eigenfunction of $H_{[-N, N]}(x+m\omega, \omega)$. This self-similarity of the Rellich graphs is basic to the formation of gaps.

- Gaps (in some finite volume) are formed by the interaction of two localized eigenfunctions with separate supports but close eigenvalues; moreover, their respective Rellich graph segments should have opposite slopes. The following figure describes this schematically. The separate bumps on the left (which depict eigenfunctions on a smaller scale) combine to form an eigenfunction on a larger scale. The curves below depict the Rellich graphs of the two small-scale eigenfunctions which then produce the separated arcs on the right as Rellich graphs of the larger scale (cf. Figure 1) in that regard. An exact way of formulating this requires the notion of a *double resonance*. Roughly speaking, this means that there are exactly two windows of localization in our finite-volume scheme of Anderson localization. To avoid a third or more windows requires elimination of triple resonances as in Chan [12].

We say that (x_0, E_0) is a *point of double resonance* for $H_{[-N, N]}(x_0, \omega)$ provided there exist $\Lambda_1 = [N'_1, N''_1], \Lambda_2 = [N'_2, N''_2]$ as shown in Figure 10 so that for $j = 1, 2$,

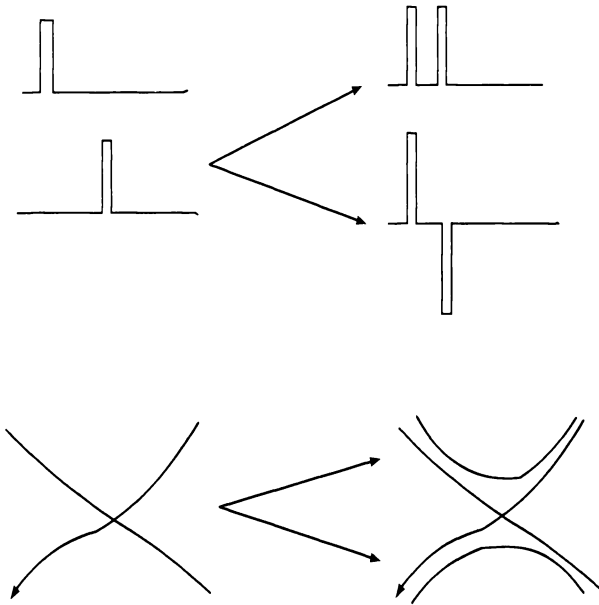


FIGURE 9. A schematic description of a double resonance

and with $\rho = e^{-N^\epsilon}$,

$$\begin{aligned} \text{spec } H_\Lambda(x_0, \omega) \cap (E_0 - \rho, E_0 + \rho) &= \{E_0\} \\ \#\{z \in D(e(x_0), \rho) : f_\Lambda(z, \omega, E_0) = 0\} &= 1 \end{aligned}$$

where $|\Lambda_1| \sim |\Lambda_2| \sim N^{2\epsilon}$ and which are separated from the boundaries of $[-N, N]$. Finally, we need to avoid triple or higher order resonances: for all $\Lambda \subset [-N, N]$ separated from $\Lambda_1 \cup \Lambda_2$ by an amount $\gg N^{2\epsilon}$, we have

$$\text{spec } H_\Lambda(x_0, \omega) \cap (E_0 - \rho, E_0 + \rho) = \emptyset.$$

The importance of this notion of a point of double resonance lies with the fact that it captures the nonperturbative essence of Figure 1. Recall that there we were able to extract a second degree polynomial from the characteristic polynomial of a finite-volume Hamiltonian, i.e., from the determinant $f_N(\cdot, \omega, E)$. This was done perturbatively by means of the Feshbach formula. Here we have to proceed differently—in fact, this extraction of a quadratic factor is accomplished by means of (a quantitative version of) the Weierstrass preparation theorem.

More precisely, if (x_0, E_0) is a point of double resonance, then one can show by means of the (AP) (see “factorization of a determinant” from above) that for all $x \in I := (x_0 - \rho, x_0 + \rho)$ there are exactly two zeros in both the z and E variables locally around the points we are considering:

$$\begin{aligned} \#\left[\text{spec } H_{[-N, N]}(x, \omega) \cap (E_0 - \rho, E_0 + \rho)\right] &= 2 \\ \#\{z \in D(e(x_0), \rho) : f_N(z, \omega, E_0) = 0\} &= 2 \end{aligned} \tag{13}$$

and the corresponding Rellich functions are separated (without any E removal!):

$$E^+(x, \omega) - E^-(x, \omega) > e^{-N^\delta} \quad \forall x \in I.$$

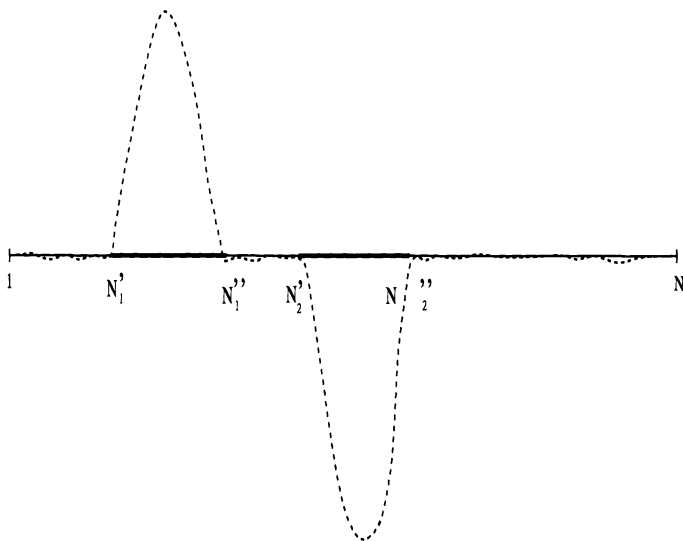


FIGURE 10. The eigenfunction corresponding to a double resonance

It is obviously essential here that we are not forced to eliminate the energy E , as otherwise we would be eliminating the gap which we are trying to construct. The fact that we can obtain separation of the eigenvalues here without eliminating an interval around E relies on the properties of a double resonance; the features needed to obtain the separation property have been included in the definition, see the discussion of Point 3 above. To obtain the zero count (13), we use the avalanche principle (see the “factorization of a determinant” from above) and the Jensen formula. Heuristically, this is a variant of the obvious fact that the number of zeros of any polynomial is the sum of the numbers of the zeros of all factors in a factorization of the polynomial. Here we of course do not have an exact factorization, but only an additive one for the logarithm that holds for most phases and up to a small error. But since the Jensen formula is based on averages and the number of zeros is integer valued, we can afford to make small errors—they produce small errors after averaging and do not affect the zero count.

Finally, in view of Figures 1 and 9, we require that the Rellich functions of the two small-scale windows Λ_1 and Λ_2 which attain the energy E have slopes of opposite signs. The appearance of such slopes is a consequence of the 1-periodicity of the continuous Rellich graphs and the fact that we are working a priori on intervals of energies on which the slopes of the Rellich functions do not vanish. Hence, it is important to realize that our construction does involve the elimination of energies and can thus, in its present form, not capture something that occurs on the entire spectrum (in particular, the argument in [28] shows that the gaps are dense but not much more beyond that).

If we do have graph segments of opposite slopes as in Figure 9, then we do get much more, namely, the desired gap between the branches E^\pm as in Figure 1 and Figure 9:

$$\min_I E^+(x) > \max_I E^-(x) + e^{-N^{2\delta}}. \quad (14)$$

To see this, we need to be able to show that locally in z the determinant of the large scale at energy E behaves like a *quadratic polynomial*; the same is needed also in the E variable. As mentioned before, we use the *Weierstrass preparation theorem* to extract these quadratic polynomials based on the zero count (13).

As mentioned above, the definition of a double resonance point is tailored to our needs in the sense that it produces the exact zero count of two, and thus allows for the Weierstrass preparation argument. The question arises how to find points of double resonance—obviously, they are essential for our gap construction. The fact that there have to be at least two small-scale Hamiltonians which are in resonance is relatively simple and can basically be deduced from the Rellich graphs at the small scale. Much trickier is to ensure that there cannot be more than two resonances—this requires the elimination of ω via an implicit function type argument. For this we crucially need to have some information on the nondegeneracy of certain derivatives, which itself is a consequence of the quantitative nonvanishing of the slopes of the Rellich graphs; recall that we remove energies for this purpose. The elimination of triple resonances employed in [28] is a variant of that introduced by Chan [12].

Returning to (14), we have finally arrived at a gap between two large-scale Rellich functions locally around x_0 ; we call this construct a pre-gap. Using (AL) we can now move this pre-gap around in x by shifting the phase. This is due to the aforementioned “self-similarity” or “almost shift-invariance” of the Rellich graphs.

It remains to show that this pre-gap is not destroyed when passing to larger scales. Among other things, we need to insure that the energies which we remove at the next scale \bar{N} are much smaller in measure than the size of any pre-gap at a previous scale N ; however, the latter is at least $e^{-N^{2\delta}}$, whereas the former is at most $\exp(-(\log \bar{N})^C)$. Hence, we define scales $N_s = \exp(N_{s-1}^{\delta_1})$ with $\delta_1 > 0$ sufficiently small. This ensures that we can safely remove “bad” energies of scale N_{s+1} inside the pre-gap at scale N_s . We can therefore repeat the pre-gap construction at scale N_{s+1} inside the pre-gap of the previous scale provided, of course, Rellich graphs of scale N_{s+1} enter that pre-gap (the case where they do not is easier).

It remains to show that this process has to terminate, i.e., one needs to find a mechanism that will ensure that after some number of steps, no more pre-gaps can form inside a pre-gap of the previous scale. At that point the pre-gap will become a gap of the infinite-volume operator. It turns out that a pre-gap cannot be filled in more than k times if the underlying potential function V has degree k . This involves a counting argument involving complex zeros. The point there is that every pre-gap at scale N_s and locally around (x_0, E_0) corresponds to a pair of complex zeros of

$$\det(H_{[-N_s, N_s]}(z, \omega) - E_0)$$

in the z variable. This pair of zeros lies off the unit circle and close to the point $e(x_0)$. Moreover, their separation from the circle is basically proportional to the size of the gap. By the aforementioned shifting procedure of pre-gaps (which is based on the almost shift-invariance of the Rellich graphs and finite-volume (AL)), each such pair generates almost N_s further pairs. If there was a sequence of consecutive scales producing pre-gaps, then one can show that this would lead to an accumulation of zeros which ultimately violates some degree considerations. This process is somewhat involved and we refer the reader to Lemma 2.24 and Section 9 of [28].

References

- [1] Avila, A., Jitomirskaya, S. *The Ten Martini problem*, to appear in Ann. of Math.
- [2] Avila, A., Krikorian, R. *Reducibility or non-uniform hyperbolicity for quasiperiodic Schrödinger cocycles*, to appear in Ann. of Math.
- [3] Avron, J., Simon, B. *Almost periodic Schrödinger operators. I. Limit periodic potentials*. Comm. Math. Phys. **82** (1981/82), no. 1, 101–120.
- [4] Avron, J., Simon, B. *Singular continuous spectrum for a class of almost periodic Jacobi matrices*. Bull. Amer. Math. Soc. (N.S.) **6** (1982), no. 1, 81–85.
- [5] Avron, J., Simon, B. *Almost periodic Schrödinger operators. II. The integrated density of states*. Duke Math. J. **50** (1983), no. 1, 369–391.
- [6] Bellissard, J., Simon, B. *Cantor spectrum for the almost Mathieu equation*. J. Funct. Anal. **48** (1982), no. 3, 408–419.
- [7] Bourgain, J. *Hölder regularity of integrated density of states for the almost Mathieu operator in a perturbative regime*. Lett. Math. Phys. **51** (2000), no. 2, 83–118.
- [8] Bourgain, J. *Green's function estimates for lattice Schrödinger operators and applications*. Annals of Mathematics Studies, **158**. Princeton University Press, Princeton, NJ, 2005.
- [9] Bourgain, J., Goldstein, M. *On nonperturbative localization with quasi-periodic potential*. Ann. of Math. (2) **152** (2000), no. 3, 835–879.
- [10] Bourgain, J., Goldstein, M., Schlag, W. *Anderson localization for Schrödinger operators on \mathbb{Z} with potentials given by the skew-shift*. Comm. Math. Phys. **220** (2001), no. 3, 583–621.
- [11] Bourgain, J., Jitomirskaya, S. *Continuity of the Lyapunov exponent for quasiperiodic operators with analytic potential*. J. Statist. Phys. **108** (2002), no. 5–6, 1203–1218.
- [12] Chan, J. *Method of variations of potential of quasi-periodic Schrödinger equation*, to appear in Geom. Funct. Anal.
- [13] Choi, M.D., Elliott, G.A., Yui, N. *Gauss polynomials and the rotation algebra*. Invent. Math. **99** (1990), no. 2, 225–246.
- [14] Craig, W., Simon, B. *Subharmonicity of the Lyapunov index*. Duke Math. J. **50** (1983), no. 2, 551–560.
- [15] Craig, W., Simon, B. *Log Hölder continuity of the integrated density of states for stochastic Jacobi matrices*. Comm. Math. Phys. **90** (1983), no. 2, 207–218.
- [16] Cycon, H. L., Froese, R. G., Kirsch, W., Simon, B. *Schrödinger Operators*. Springer Verlag, Berlin, Heidelberg, New York, 1987.
- [17] Deift, P., Simon, B. *Almost periodic Schrödinger operators. III. The absolutely continuous spectrum in one dimension*. Comm. Math. Phys. **90** (1983), no. 3, 389–411.
- [18] Dinaburg, E. I., Sinai, Y. G. *The one dimensional Schrödinger equation with quasiperiodic potential*. (Russian) Funkt. Anal. i. Priloz. **9** (1975), 8–21; English translation: Functional Anal. Appl. **9** (1975), no. 4, 279–289 (1976).
- [19] Eliasson, H. *Floquet solutions for the 1-dimensional quasiperiodic Schrödinger equation*. Comm. Math. Phys. **146** (1992), 447–482.
- [20] Figotin, A., Pastur, L. *Spectra of random and almost-periodic operators*. Grundlehren der mathematischen Wissenschaften **297**, Springer, Berlin, 1992.
- [21] Fröhlich, J., Spencer, T. *Absence of diffusion in the Anderson tight binding model for large disorder or low energy*. Comm. Math. Phys. **88** (1983), 151–189.
- [22] Fröhlich, J., Spencer, T. *A rigorous approach to Anderson localization*. Phys. Rep. **103** (1984), no. 1–4, 9–25.
- [23] Fröhlich, J., Spencer, T., Wittwer, P. *Localization for a class of one dimensional quasi-periodic Schrödinger operators*. Comm. Math. Phys. **132** (1990), 5–25.
- [24] Fürstenberg, H., Kesten, H. *Products of random matrices*. Ann. Math. Statist. **31** (1960), 457–469.
- [25] Garnett, J. *Bounded Analytic Functions*, Academic Press, New York, 1981.
- [26] Goldstein, M., Schlag, W. *Hölder continuity of the integrated density of states for quasiperiodic Schrödinger equations and averages of shifts of subharmonic functions*. Ann. of Math. (2) **154** (2001), no. 1, 155–203.
- [27] Goldstein, M., Schlag, W. *Fine properties of the integrated density of states and a quantitative separation property of the Dirichlet eigenvalues*, preprint 2005.

- [28] Goldstein, M., Schlag, W. *On resonances and the formation of gaps in the spectrum of quasi-periodic Schrödinger equations*, preprint 2005.
- [29] Goldstein, M., Schlag, W. *On Schrödinger operators with dynamically defined potentials*, Moscow Math. J. **5** (2005), no. 3, 577–612.
- [30] Gordon, A., Jitomirskaya, S., Last, Y., Simon, B. *Duality and singular continuous spectrum in the almost Mathieu equation*. Acta Math. **178** (1997), 169–183.
- [31] Harper, P. G. *Single band motion of conduction electrons in a uniform magnetic field*. Proc. Phys. Soc. London **A68** (1955), 874–892.
- [32] Herman, M. *Une méthode pour minorer les exposants de Lyapounov et quelques exemples montrant le caractère local d'un théorème d'Arnold et de Moser sur le tore de dimension 2*. Comment. Math. Helv. **58** (1983), no. 3, 453–502.
- [33] Jitomirskaya, S. Ya. *Metal-insulator transition for the almost Mathieu operator*. Ann. of Math. (2) **150** (1999), no. 3, 1159–1175.
- [34] Jitomirskaya, S., Simon, B. *Operators with singular continuous spectrum. III. Almost periodic Schrödinger operators*. Comm. Math. Phys. **165** (1994), no. 1, 201–205.
- [35] Johnson, R., Moser, J. *The rotation number for almost periodic potentials*. Comm. Math. Phys. **84** (1982), no. 3, 403–438.
- [36] Kotani, S. *Ljapunov indices determine absolutely continuous spectra of stationary random one-dimensional Schrödinger operators*. Stochastic Analysis (Katata/Kyoto, 1982), pp. 225–247, North-Holland Math. Library, **32**. North-Holland, Amsterdam, 1984.
- [37] Kotani, S. *One-dimensional random Schrödinger operators and Herglotz functions*. Probabilistic Methods in Mathematical Physics (Katata/Kyoto, 1985), pp. 219–250, Academic Press, Boston, MA, 1987.
- [38] Kotani, S. *Generalized Floquet theory for stationary Schrödinger operators in one dimension*. Chaos Solitons Fractals **8** (1997), no. 11, 1817–1854.
- [39] Lang, S. *Algebra*, third edition, Addison-Wesley, Reading, MA, 1993.
- [40] Last, Y. *A relation between a.c. spectrum of ergodic Jacobi matrices and the spectra of periodic approximants*. Comm. Math. Phys. **151** (1993), 183–192.
- [41] Last, Y. *Zero measure spectrum for the almost Mathieu operator*. Comm. Math. Phys. **164** (1994), 421–432.
- [42] Last, Y., Simon, B. *Eigenfunctions, transfer matrices, and absolutely continuous spectrum of one-dimensional Schrödinger operators*. Invent. Math. **135** (1999), no. 2, 329–367.
- [43] Levin, B. Ya. *Lectures on Entire Functions*. Transl. of Math. Monographs, **150**. American Mathematical Society, Providence, RI, 1996.
- [44] Puig, J. *Cantor spectrum for the almost Mathieu operator*. Comm. Math. Phys. **244** (2004), no. 2, 297–309.
- [45] Puig, J. *Reducibility of Quasi-Periodic Skew-Products and the Spectrum of Schrödinger Operators*. Ph.D. thesis, Universitat de Barcelona, 2004.
- [46] Simon, B. *Kotani theory for one-dimensional stochastic Jacobi matrices*. Comm. Math. Phys. **89** (1983), no. 2, 227–234.
- [47] Sinai, Y. G. *Anderson localization for one-dimensional difference Schrödinger operator with quasi-periodic potential*. J. Statist. Phys. **46** (1987), 861–909.
- [48] Thouless, D. *Scaling for the discrete Mathieu equation*. Comm. Math. Phys. **127** (1990), 187–193.
- [49] Wegner, F. *Bounds on the density of states in disordered systems*. Z. Phys. **B44** (1981), 9–15.

(Michael Goldstein) DEPARTMENT OF MATHEMATICS, UNIVERSITY OF TORONTO, TORONTO, ONTARIO, CANADA M5S 1A1

E-mail address: gold@math.toronto.edu

(Wilhelm Schlag) DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF CHICAGO, 5734 SOUTH UNIVERSITY AVENUE, CHICAGO, IL 60637, U.S.A.

E-mail address: schlag@math.uchicago.edu

Ergodic Schrödinger Operators (on one foot)

Svetlana Jitomirskaya

Dedicated to Barry Simon on the occasion of his 60th birthday

ABSTRACT. We review several topics related to two most popular ergodic families: the Anderson model and quasiperiodic Schrödinger operators.

CONTENTS

1. Introduction
 2. Random Schrödinger Operators a.k.a. Anderson-type Models
 3. Quasiperiodic Operators
- References

1. Introduction

The title refers to a chapter (volume?) in [175] in the same way that [174] refers to [173]. (See [174] for the story behind this.) The original ambition was for this relation/analogy to be reflected in the content, but it had to be significantly downscaled. So I will concentrate almost entirely on only two models—still with no attempt at being comprehensive—and will focus on selected advances in the past ten years or so.

Writing this review has made it particularly obvious to me how much Barry Simon has influenced this subject, both through his fundamental contributions and review articles that have defined the area and then served as a standard reference for 20+ years, but even more importantly, by throwing his weight behind the area, and particularly through bravely formulating and popularizing various problems and (sometimes wrong) conjectures. As a result, while Barry has not had many

2000 *Mathematics Subject Classification*. 81q10, 82b44, 47b80, 47b36, 47b39, 47b36, 47n50.

Key words and phrases. Anderson model, quasiperiodic, localization, absolutely continuous spectrum.

This work was partially supported by NSF grant DMS-0300974 and BSF grant 2002068.

students of his own working in this field, it's fair to say that he has been largely responsible for attracting a lot of new talent to this area.

Let (Ω, μ) be a probability measure space, and $T_j, j \in \mathbb{Z}^d$, be an ergodic measure-preserving action of \mathbb{Z}^d on Ω . Discrete Schrödinger operators with ergodic potentials are operators acting on $\ell^2(\mathbb{Z}^d)$ and defined by

$$H_\lambda = \Delta + \lambda V \tag{1}$$

where Δ is the lattice tight-binding Laplacian

$$\Delta(n, m) = \begin{cases} 1, & \text{dist}(n, m) = 1 \\ 0, & \text{otherwise} \end{cases}$$

and $V(n, m) = V_n \delta(n, m)$ is a potential given by $V_n = f(T_n \theta), \theta \in \Omega$, where $f : \Omega \rightarrow \mathbb{R}$. In certain cases Δ may also be replaced by a long-range Laplacian $L(n, m) = L(n-m)$ with $L(n) \rightarrow 0$ sufficiently fast. We will also define a continuum analogue later.

The most intensively studied cases in both the physics and mathematical physics literature are those of random V (V_n are i.i.d.r.v.'s), called the *Anderson model* and quasiperiodic V ($T_n : \mathbb{T}^b \rightarrow \mathbb{T}^b$ defined by $T_n \theta = \theta + \sum_{i=1}^d n_i \omega_i, n = (n_1, \dots, n_d) \in \mathbb{Z}^d$, where $\omega_i \in \mathbb{T}^b$ are incommensurate and where we will assume f to have some regularity.) This article is devoted to some recent advances in these two models. The questions of interest are the nature and structure of the spectrum, behavior of the eigenfunctions, and the quantum dynamics: properties of the time evolution $\Psi_t = e^{itH} \Psi_0$ of an initially localized wave packet Ψ_0 .

There are many open problems related to these two models, and we mention some in this paper. We do not claim that the problems we list with numbers are the most important ones. In fact, we discuss extensively but do not list separately the two most important outstanding problems: localization for the Anderson model on \mathbb{Z}^2 and extended states for the Anderson model with $d > 2$, as these problems, being already so famous, do not need further popularizing. Nor do we even mention (other than in this sentence) another outstanding question: continuity of the integrated density of states in the continuum (see [171]). There are also other major problems not touched upon. The problems we do list have a common feature that, given recent advances, they all seem to be more within reach than they were in the past—yet they all remain significant challenges.

Finally, our title may be particularly misleading as we leave out almost the entire discussion of the general theory of ergodic operators, most notably Kotani theory ([140, 143, 167, 69, 141, 144]) and also the regularity of the integrated density of states issues ([61, 74]). The classical results mentioned above can mostly be found in several books on the subject (e.g., [62] where they play a prominent role) but there also have been some interesting applications of Kotani theory lately ([64, 16, 17]). Nor do we discuss at all other ergodic potentials that have been studied successfully and where there are many remaining challenges, e.g., limit-periodic operators (e.g., [23]), skew shifts (e.g., [87, 44]), the Maryland model (e.g., [163, 168]), hyperbolic dynamics (e.g., [57, 50]) or subshifts (see [63] in this Festschrift for a review).

While this is a review article, there are some statements here that have not previously appeared in the literature, for instance, an example presented as Theorem 2.

Operators with ergodic potentials always have spectra (and pure point spectra, understood as closures of the set of eigenvalues) constant for a.e. realization of the potential. The same is true for singular continuous spectra understood as closures of corresponding supports. The individual eigenvalues (singular measure supports) however depend very sensitively on the phase. Moreover, the pure point spectrum of operators with ergodic potentials never contains isolated eigenvalues, so pure point spectrum in such models is dense in a certain perfect set. (See, e.g., [62] for all those classical facts that go back to [160, 146, 127, 24].) An easy example of an operator with dense pure point spectrum is H_∞ which is operator (1) with $\lambda^{-1} = 0$, or pure diagonal. It has a complete set of eigenfunctions, characteristic functions of lattice points, with eigenvalues V_j . H_λ may be viewed as a perturbation of H_∞ for small λ^{-1} , and from this point of view it is natural to expect that the localization will persist for small values of λ^{-1} . One explanation for this intuition is that if, in general, in the regime of localization we consider quantum dynamics of a wave packet with a finite number of excited eigenmodes, then the quantum motion will be quasiperiodic on the torus determined by these modes. The general KAM philosophy implies that under small perturbations the tori should persist, only getting modified slightly, which in turn should imply localization for the perturbed model. However, this, of course, cannot be true in such generality, even if (M, μ, T) is an ergodic system with M sufficiently rich (e.g., [29] or [63] in this Festschrift). Localization can be destroyed for arbitrarily small λ^{-1} by the resonances that are too strong. Even when the result is correct, since V_j are dense, small denominators $(V_i - V_j)^{-1}$ make any perturbation theory difficult, traditionally requiring intricate KAM-type schemes.

One might expect that H_λ with λ small can be treated as a perturbation of $H_0 = \Delta$, and therefore have absolutely continuous spectrum. It is not the case though for random potentials in $d = 1$, where Anderson localization holds for all λ . The same is expected for random potentials in $d = 2$ (but not higher). Moreover, in the one-dimensional case there is strong evidence (numerical, analytical, as well as rigorous) that even models with very mild stochasticity in the underlying dynamics (and sufficiently nice sampling functions) have point spectrum for all values of λ , like in the random case (e.g., $V_n = \lambda f(n^\sigma \alpha + \theta)$, for any $\sigma > 1$; see [87]). While for nonintegral σ , such potentials are not generally ergodic, this still illustrates the point. At the same time, for quasiperiodic potentials one can in many cases show absolutely continuous spectrum for λ small as well as pure point spectrum for λ large (see below), and therefore there is a metal-insulator transition in the coupling constant. It is an interesting question whether quasiperiodic potentials are the only ones with metal-insulator transition in 1D. Absolutely continuous spectrum in 1D requires a potential to be nondeterministic [143]; moreover, in all known examples (e.g., [23]), potentials are almost periodic.

PROBLEM 1. *Does absolutely continuous spectrum for a 1D ergodic potential imply almost periodicity of the potential?*

This was established under some additional assumptions (see e.g., [12]) but the problem as stated does not seem unreasonable.

2. Random Schrödinger Operators a.k.a. Anderson-type Models

In 1958 Anderson [11] proposed to explain the observed transport properties of electrons in solids and, as a corollary, the observed finite conductivity of metals at room temperatures in a single-electron model by the presence of random impurities in a crystal. According to [11], “at sufficiently low densities transport does not take place; the exact wave functions are localized in a small region of space”—the phenomenon that was dubbed Anderson localization (dynamical or spectral, depending on what part of the above statement was emphasized).

The main model here is that of an electron-gas of Fermions with a one-particle Hamiltonian. The one-electron model of a crystal with impurities (a.k.a. the Anderson model) is operator (1) on \mathbb{Z}^d with V_n being the i.i.d.r.v.’s. The following picture represents the state of current beliefs about this model: for large couplings the large fluctuations of the potential dominate leading to Anderson localization. The same happens for all energies near the edges of the spectrum (or edges of the bands), with the corresponding states localized in the regions of such fluctuations. The same also happens in one and two dimensions for all energies. However, in three or more dimensions at small couplings and away from the edges of the spectrum the kinetic term dominates the fluctuations of the potential leading to so-called extended states. Thus one expects a metal-insulator transition from the extended states to the localized states regime. The spectral interpretation of this picture is that one expects pure point spectrum with exponentially decaying eigenfunctions in the localization (insulator) regime, and absolutely continuous spectrum in the extended states (metallic) regime. Dynamically, one expects absence of transport in the insulator range and diffusive transport in the metallic regime.

A similar situation is expected for the continuum analogue where

$$H = -\Delta + V_\omega \tag{2}$$

on $L^2(\mathbb{R}^d)$, and a typical example of the potential is $V_\omega(x) = V_{\text{per}} + \sum_{i \in \mathbb{Z}^d} \omega_i u(x - i)$ where ω_i are i.i.d.r.v.’s and u is a certain bump function. Here one expects localization near the bottom of the spectrum/interior band edges, and extended states for high energies in three or more dimensions.

More precisely, the expected picture is that in all regimes there is a ballistic transport up to times of order λ^{-2} . This is expected in the metallic (small coupling for $d \geq 3$) regime, with the diffusive transport appearing at larger times, but also even in the localization regime, where there is no transport at all in the infinite time-scale. Moreover, for $d = 2$, it is expected that this ballistic regime is followed by the diffusive transport up to times of order $e^{\lambda^{-2}}$, and only after that should the localization effects become pronounced. This picture has recently been partially confirmed in the work of Erdős, Salmhofer, and Yau [83, 84] who established ballistic transport up to times of order λ^{-2} , and diffusive transport up to the times of order $\lambda^{-2-\delta}$, $\delta > 0$, for a certain Anderson-type model in \mathbb{R}^3 . Their approach to prove diffusive transport seems to be potentially extendable up to times of order λ^{-4} , but there are some essential difficulties for going beyond that limit.

We will review the recent rigorous results related to the metal-insulator transition for the Anderson model and corresponding methods. We do not review here other related topics such as, e.g., the regularity of the integrated density of states (see [176, 129] for $d = 1$ results and, e.g., [60] for $d > 1$) where there has been a lot of progress, as well as many other important issues.

2.1. Localization. The one-dimensional case is very well understood, with first proofs of localization going back to the pioneering work of Goldscheid, Molchanov, and Pastur [102]. However, even here some unexpected aspects keep getting discovered. The final result is the following

THEOREM 1. *If $d = 1$, then for any nontrivial distribution of V_n 's there is dynamical localization for all λ .*

This theorem was first proved in [52], with a dynamical enhancement given in [68] (see Section 2.1.5 for more history/discussion). For the multidimensional case, where localization is established for various models under the assumption of either a high disorder or energies near the boundaries of the spectrum of the unperturbed operator, there are now two competing methods: multiscale analysis and fractional moments.

2.1.1. *Multiscale analysis.* Multiscale analysis was the first one historically, preceding fractional moments by some ten years. It was originally developed by Fröhlich and Spencer [91] as a probabilistic KAM-type scheme that involved a definition and detailed study of the geometry and combinatorics of singular (resonant) sets. In fact, what is proved in [91] is what is now called *the single energy estimate* for the Green's function. With probability one, the Green's function $G_\omega(x, y; z) = (H_\omega - z)^{-1}(x, y)$ decays exponentially in the distance $\|x - y\|$, for $x, y \in \mathbb{Z}^d$. That is, there exist finite constants $C_0 > 0$, and $\gamma(E) > 0$, so that

$$\sup_{\epsilon > 0} |G_\omega(x, y; E + i\epsilon)| \leq C_0 e^{-\gamma(E)\|x-y\|}, \tag{3}$$

with probability one.

The argument of [91] required absolute continuity of the distribution; see below. Estimate (3) is not sufficient to obtain localization because the zero measure exceptional set can depend on energy. Several methods have been developed to obtain localization out of such single energy estimates, and they can be divided into two categories:

- (1) Spectral averaging [177, 73] (building on some ideas of [51, 142]; see also [117] where those ideas seem to have appeared for the first time, and [58] for the continuum);
- (2) Energy interval multiscale analysis a.k.a. energy elimination [89, 78].

Spectral averaging, in any form, is based on rank one perturbations and thus requires absolute continuity of the joint distribution.

The multiscale analysis proceeds by inductively passing from smaller to larger finite-volume scales. The technique borrows ideas from statistical mechanics (in fact, it was first developed for the study of the Kosterlitz–Thouless transition [90, 139] and only later adapted to the analysis of localization) and is based on treating the infinite-volume system as the thermodynamic limit of finite-volume ones. It has been significantly simplified, further developed, extended to the continuum, and applied to many situations, [113, 181, 78, 58, 134, 128, 184, 183] being a very incomplete list. The latest and most powerful version, bootstrap multiscale analysis, is developed in [96]. All the proofs of the single energy estimate (or energy interval analysis), however, require two important building blocks:

- (1) the initial length-scale estimate
- (2) the Wegner estimate.

The Hamiltonian for the finite-volume system H_Λ is obtained by restricting H to $\Lambda \subset \mathbb{Z}^d$ with some self-adjoint boundary conditions.

Let $\Lambda_L(k)$ be a cube in \mathbb{Z}^d of side length L centered at $k \in \mathbb{Z}^d$. For $M > 0$ and $E \in \mathcal{R}$, we say a cube $\Lambda_L(k)$ is (m, E) -regular if the Green's function $G_{\Lambda_L(k)}(i, j; E)$ for $H_{\Lambda_L(k)}$ satisfies the bound

$$|G_{\Lambda_L(k)}(k, l; E)| \leq e^{-mL/2} \quad (4)$$

for all $l \in \partial\Lambda_L(k)$. For a fixed realization ω , this bound indicates that the energy E is not too close to the discrete spectrum of $H_{\Lambda_L(k)}$. This estimate has to hold for some initial length scale as an input, which allows one to obtain a sequence of estimates of this type for the increasing sequence of scales. Since singular (i.e., nonregular) blocks will appear with some probability in a larger block, one needs also an a priori estimate on the Green's function of those. It is called a Wegner estimate. It is an estimate on the expectation of the spacing of the eigenvalues of $H_{\Lambda_L(k)}$. A local, random Hamiltonian H_Λ satisfies a Wegner estimate at energy E if there are positive exponents $0 < \alpha \leq 1$ and $1 \leq \beta < \infty$, and a constant $C_W > 0$, so that for $\eta > 0$ small enough,

$$\mathbb{P}\{\text{dist}(\sigma(H_\Lambda), E) \leq \eta\} \leq C_W \eta^\alpha |\Lambda|^\beta. \quad (5)$$

This is one possible form of the Wegner estimate; the proofs of exponential localization work with a much weaker form as well. The original form is in [190], with many later enhancements. For the continuum, it was first proved by Kotani–Simon [145] and improved in [58].

The initial length-scale estimate in the continuum is stated in terms of appropriate bounds on the operator norm of the resolvent of the local Hamiltonian H_Λ . One of the main practical advances of [96] is a clean form of the initial length-scale hypothesis, both for the discrete and continuum cases. Namely, in [96] the resolvent is required to decay only polynomially with respect to the length L with the probability that is independent of the length scale L . This was crucial for some important recent advances described below.

Germinet and Klein [97] also give finite-volume criteria for localization in the continuum. These are explicit conditions, depending on the various parameters of the model, for starting the bootstrap multiscale analysis. These explicit finite-volume criteria yield localization in situations where the crucial quantities of the model that enter the multiscale analysis (the constant in Wegner's estimate, and the constant in the Simon–Lieb-type inequality needed to relate resolvents at different scales in the continuum) depend on the parameters of the model (e.g., the disorder parameter, the energy where localization is to be proven, the strength of the magnetic field).

2.1.2. Method of fractional moments. An alternative method for proving localization was found by Aizenman and Molchanov [7]. It is much simpler than the multiscale analysis in the lattice case. The key feature is that it allows one to obtain various statements on localization from single length-scale estimates (or a sequence of constructive finite-volume criteria [8]). Also, it gives a natural meaning to the notion of localization length scale often used in physics: it is the minimum scale for which the estimate on the fractional power of the Green's function holds. The method uses the subharmonicity properties of the Green's function and proceeds to estimate the expectation of the fractional ($s < 1$) moments of the Green's function directly. The fractional moment is needed to compensate for the singularity of the

distribution function of the Green’s function. The main result of the method is that for any $0 < s < 1$, there exists a constant $C_s > 0$ so that for certain energies there is a finite constant $\gamma(E) > 0$ so that

$$\mathbb{E}\{|G_\omega(i, j; E)|^s\} \leq C_s e^{-s\gamma(E)\|i-j\|}. \tag{6}$$

The difference equation for the Green’s function can be rewritten as

$$(V_\omega - E)G_\omega(i, j; E) = \delta_{ij} + \sum_{k:|k-i|=1} G_\omega(i, j; E). \tag{7}$$

A simple important fact is that the dependency of G on the coupling constant at one site can be isolated by writing G using the resolvent formula as

$$G_\omega(i, j; E) = A/(\lambda_k + B). \tag{8}$$

Taking the s^{th} power of (7) and the expectation with respect to the random variable $\lambda_k(\omega)$, one obtains

$$\mathbb{E}\{|\lambda_k(\omega) + \alpha_E|^s / |\lambda_k(\omega) + B|^s\}. \tag{9}$$

For absolutely continuous probability measures, one can prove that there exists a constant $D_s > 0$ so that

$$\mathbb{E}\{|\lambda_k(\omega) + \alpha_E|^s / |\lambda_k(\omega) + B|^s\} \geq D_s \mathbb{E}\{1 / |\lambda_k(\omega) + B|^s\}. \tag{10}$$

The resulting formula is the expression of the subharmonicity of the Green’s function and can be used to derive the boundedness of the expectation of the s^{th} moment. As seen from the above discussion, absolute continuity of the probability distribution is quite crucial here.

As discussed in [6], estimate (6) has several interesting corollaries. First, the dynamical localization in the range of energies for which (6) holds is obtained essentially automatically. The form of dynamical localization implied by (6) is as follows. Suppose that (6) holds on an interval $I \subset \mathbb{R}$. Let ϕ_I be a real-valued function with support in I . There exist finite constants $A > 0$ and $\mu > 0$ so that

$$\mathbb{E}\{\sup_{t \in \mathcal{R}} \langle x | e^{-iH_\omega t} \phi_I(H_\omega) | y \rangle\} \leq A e^{-\mu\|x-y\|} \tag{11}$$

for all $x, y \in \mathbb{Z}^d$. Note that even the most sophisticated multiscale analysis only leads to subexponential bounds here.

Another interesting consequence of the fractional moment bound is the result of Minami [157], extending earlier work of Molchanov [158] on the 1D case, on the energy-level statistics for the Anderson model on a lattice. Let $\{E_j(\Lambda)\}$ be the eigenvalues of the random Hamiltonian H_ω restricted to the box Λ . Consider the rescaled eigenvalues $\xi_j(\Lambda; E) = |\Lambda|(E_j(\Lambda) - E)$ near an energy E . Minami proved that if the fractional moment of the Green’s function at energies near E is exponentially bounded, as in (6), and if the density of states $\rho(E)$ exists at energy E , then the random point process on \mathbb{R} given by the sequence $\{\xi_j(\Lambda; E)\}$ converges weakly to the stationary Poisson process on \mathbb{R} with intensity $\rho(E)$ as $\Lambda \rightarrow \mathbb{R}^d$. This is interpreted as energy-level repulsion in that the energy levels do not exhibit any correlation as the volume of the system increases.

Aizenman et al. [5] have extended the fractional moment method to Schrödinger operators on \mathbb{R}^d with a certain class of Anderson-type potentials. This was a very important development as the method originally seemed to be intrinsically discrete. The main difficulty was that while in the discrete case changing a term

at one site is only a rank one perturbation, in the continuum a change in each of the random variables leads to a perturbation of infinite rank. In [5] the finite-rank perturbation theory was replaced by the use of the Birman–Schwinger kernel. While in the discrete case the fact that changing one parameter leads to only a rank one perturbation is roughly manifested in the property that the number of eigenvalues of a finite box problem that are smaller than a given energy E , the “spectral shift” changes by no more than one upon any change of one parameter, in the continuum case the same quantity is not uniformly bounded. This is overcome in [5] by showing that for random potentials the “spectral shift” is necessarily of finite mean. However, there are many technical difficulties involved, and as a result the continuum version is hardly simpler than a multiscale analysis proof. Also, Minami’s result does not seem to easily extend, although this seems to be less of a problem with the continuum version of fractional moments than with the extension of other estimates used by Minami.

PROBLEM 2. *Prove Poisson statistics for continuum models in $d \geq 2$ in the localization regime.*

2.1.3. Initial length-scale estimates. While the Wegner estimate, given some regularity of the probability distribution, usually holds for all energies, the initial length-scale estimate of multiscale analysis is the one that is energy/model dependent. As should be clear from the previous section, the fractional moments method also requires a bound that holds at a finite scale as a starting point. In the one-dimensional and quasi-one-dimensional (strip) cases, bounds of such form follow from random matrix theory [52, 131] for any coupling. For large couplings in any dimension, it is a very simple bound, but the resulting estimate is probably rather crude and far from being optimal. The case of small couplings (weak disorder) has been studied in [4, 136, 189]. The first quantitative results on the size of the localization region inside the unperturbed band are due to Wang [189] who used the supersymmetric formalism. The best result in this direction belongs to Klopp [136] who used the Lifshitz tails bound on the integrated density of states to obtain the initial scale estimate. Let $H_\omega = H_0 + \lambda V_\omega$ be a random Schrödinger operator with an Anderson-type potential. Let Σ denote the deterministic spectrum. Let us assume that $\inf \Sigma = 0$. If $\lambda \geq 0$ is a measure of the disorder, then Klopp proved that the region of localization extends at least as far as $\lambda^{1+\eta} \mathbb{E}\{V_\omega\}$ into the deterministic spectrum Σ .

The Lifshitz tail estimate (e.g., [138]) allows one to estimate the probability that a finite-volume Hamiltonian contains no spectrum at all in a certain energy interval, through a statement of the form

$$\mathbb{P}\{\sigma(H_\Lambda) \cap [0, E] \neq \emptyset\} \leq C|\Lambda|N(E), \quad (12)$$

showing that the probability of small eigenvalues for the finite-volume operator is exponentially small in the energy. This gives an initial length-scale estimate (with the same probability) by the Combes–Thomas argument that then implies exponential decay of the Green’s function for corresponding realizations.

Once the initial scale estimate is obtained, one can use either the finite-volume criteria for localization [8] in order to obtain strong estimates on the decay of the Green’s function, or the multiscale analysis.

In dimension two, the proof of localization by either method can, as discussed above, be fully reduced to an initial length-scale estimate—that is where the challenge lies, unless a completely new method emerges. The current methods do not capture the specificity of $d = 2$.

2.1.4. *Dynamical localization.* Dynamical localization, i.e., a nonspread of initially localized wave packet, is usually defined mathematically as a suitable decay of the matrix elements of the time evolution operator e^{-iHt} , or boundedness in time of the moments of the position operator, e.g., as in (13) (many different definitions of varying degrees of strength are now available). Pure point spectrum is implied by the dynamical localization; this is the subject of the celebrated RAGE (Ruelle–Amrein–Georgescu–Enss) theorem (see, e.g., [62]), so the latter is a stronger notion. In [70], del Rio, Jitomirskaya, Last, and Simon constructed a potential such that operator H on $\ell^2(\mathbb{Z})$ has a complete set of exponentially decaying eigenfunctions (with same rate of decay) However, for any $\delta > 0$, $\|xe^{-itH}\delta_0\|^2/t^{2-\delta}$ is unbounded as $t \rightarrow \pm\infty$, where $\|xe^{-itH}\delta_0\|^2 = (e^{-itH}\delta_0, x^2e^{-itH}\delta_0)$. Note that by a classical result of Simon [169], $\|xe^{-itH}\delta_0\|^2$ (that can grow at most ballistically, i.e., as t^2) grows necessarily subballistically in case of pure point spectrum.

Other examples with the same property were later constructed in [26, 95]. This example has shown that the result of [169] is optimal and that mere “exponential localization” of eigenfunctions need not have any consequences for the dynamics. This as well as some other theorems in [70] have shown the importance of proving the dynamical localization, which, since then, has become a common point of view.

For the Anderson model in the regime of localization, dynamical localization is easiest to obtain directly by the method of the fractional moments [4]. For the one-dimensional case, dynamical localization was also obtained in [72] and, in a restricted form, in [155].

De Bièvre and Germinet [68] proved that the von Dreifus–Klein version of multiscale analysis implies certain uniformity in the decay of the eigenfunctions that, by [70], leads to dynamical localization. As a consequence, they proved dynamical localization for random Schrödinger operators on the lattice and the continuum where the statement takes the following form. For an energy interval $I \subset \mathbb{R}$ in the localization region, and for any $q > 0$ and wave function ψ with compact support, we have

$$\sup_t \| |x|^q E_I(H_\omega) e^{-itH_\omega} \psi \| < \infty, \tag{13}$$

with probability one. Damanik and Stollmann [66] extended the analysis in [68] to prove partial strong dynamical localization (i.e., they prove (13) for the expectation of the LHS for all $q < q_0$ for some $q_0 < \infty$). Germinet and Klein [96] used a generalized eigenfunction expansion to exploit the bootstrap multiscale analysis, instead of resorting to centers of localization as in [68, 66], and obtained the subexponential decay of any order of (11) and strong dynamical localization in the Hilbert–Schmidt norm. For the deterministic ergodic potentials (see Section 3.3), dynamical upgrades of existing localization proofs have also been made. It was shown in [94] that strong dynamical localization (13) holds for the almost Mathieu operator throughout the regime of localization as in [118] (the earliest result in this direction was in [121]). While proved in [94] with a slightly more restrictive condition on the frequency, this holds as stated by a result of [48]. Bourgain–Jitomirskaya [46] showed that dynamical localization can be obtained by an upgrade

of the Bourgain–Goldstein method [43], and therefore holds wherever this method applies.

All the results above, plus the artificial character of examples in [70, 26, 95], may leave an impression that pure point spectrum cannot coexist with nontrivial transport in physically relevant models. It is not the case though as demonstrated by the example of random polymer models. A random polymer model is a 1D Jacobi matrix randomly composed of two finite building blocks. The best known example is a dimer model, where the random potential takes only two values, but these potential values always come in neighboring pairs (dimers). The dimer model was studied by Dunlap, Wu, and Phillips [79] who argued and showed numerically that the second moment of the position operator X on the lattice grows superdiffusively under the dynamics: $\langle \psi | X^2(t) | \psi \rangle \approx C t^{3/2}$ for a localized initial state ψ and any typical dimer configuration.

This phenomenon, however, was considered controversial by physicists (see, e.g., [154]), as the common perception was that randomness in 1D always leads to strong localization. Indeed, random polymer models always have exponential localization. (See [68] for the dimer model; for other models, the method of [65] applies. Both essentially stem from [52].) Critical energies (which exist for the dimer model and many other, but not all, polymers) are defined as energies at which transfer-matrices over polymer blocks are elliptic and commute (leading to perfect transmission, but only at these discrete energies). For spectral projections on compact sets outside the critical energies, strong dynamical localization holds [68, 65]. In [123] it was shown that for random polymer models with at least one critical energy there is nontrivial transport:

$$\frac{1}{T} \int_0^T \langle \psi | X^q(t) | \psi \rangle \geq C_\epsilon T^{q-1/2-\epsilon},$$

any $\epsilon > 0$, for a.e. configuration. The result, in particular, applies to the dimer model, for which the last inequality with $q = 2$ gives precisely the bound in [79], therefore confirming their findings. This shows that the distinction between spectral and dynamical localization should indeed be made.

2.1.5. *Anderson models with singular probability distributions.* As should be clear from the above discussion, both known methods for the proof of localization have absolute continuity of the underlying probability distribution as a major requirement, as both are based in one way or another, on rank one perturbations. Thus singular distributions, e.g., Bernoulli, when the random variables are allowed to take only two values, represent an important challenge. In the one-dimensional case localization was proved for such models by Carmona, Klein, and Martinelli [52] using the Furstenberg–Lepage theory of products of random matrices. A different proof, using supersymmetry, was given later by [166] (see also [65] for the continuum model with Bernoulli distribution). For the multidimensional case, [52] contains a Wegner estimate for probability measures that are Hölder continuous, in the lattice, and Stollmann [183] extended this result to the continuum. However, none of this is close to being applicable to the Bernoulli distribution, and proving localization for multidimensional Bernoulli–Anderson models is widely regarded as a very significant problem in the area.

A major breakthrough came recently with a work of Bourgain–Kenig [49] where they study a continuous analogue of the Bernoulli–Anderson model: the operator

(2) on \mathbb{R}^d with potential

$$V_\omega(n) = \lambda \sum_{m \in \mathbb{Z}^d} \phi(n - m)\omega_m \tag{14}$$

where ϕ is a smooth compactly supported function.

The result of [49] is localization near the bottom of the spectrum for a.e. realization. The general framework of the method of [49] is multiscale analysis. However, it required many completely new ingredients. Most importantly, the Wegner lemma, which has been a necessary starting point for any multiscale analysis, is not a priori available here and had to be proved inductively as a part of the multiscale scheme. To obtain the usual Wegner lemma for models with absolutely continuous (or Hölder) distributions, it suffices to consider variations of only one parameter. For the Bernoulli case, it is clear that variations of many parameters are needed. The corresponding measure estimates were obtained in [49] by introducing the concept of “free sites”: these are scale-dependent sets such that for each element, the value of the corresponding random variable can vary from 0 to 1 without affecting the estimates on the Green’s function on that scale.

For the Wegner-type estimate, one then needs a lower bound on the variation of an eigenvalue produced by the variation of the parameter at each site to quantitatively ensure that one can move the eigenvalues by such variations. This is resolved in [49] for the continuum model by a refined version of the unique continuation principle: a quantitative formulation of the fact that eigenfunctions cannot vanish locally. The probabilistic estimates then come through the use of Sperner’s lemma for boolean functions.

Finally, while these new ideas allow one to obtain a single energy estimate on the Green’s function, this is not yet sufficient to obtain localization. The reason is that the spectral averaging method (e.g., Simon–Wolff) is not available here as it again requires continuity of the distribution, and the energy-interval multiscale analysis cannot work here either because the probabilistic estimate for the single-energy bound is too weak for that. In fact, the weakness of this estimate is due to the form of the unique continuation principle, and it is argued in [49] that this form is essentially optimal, in general, because of Meshkov’s example. Thus, this does not seem to be improvable within the same circle of ideas. As a result, a completely new method is designed in [49] to obtain localization out of a single energy estimate, without the use of spectral averaging, adding therefore a third method to the two described in Section 2.1.1. The method, while very delicate in the Bernoulli case because of the weak probabilistic bounds in the single energy estimate, becomes relatively simple if combined with, for example, the original Fröhlich–Spencer bound.

A precursor to [49] was in a somewhat earlier work by Bourgain [39] where he studied a correlated multidimensional Anderson–Bernoulli model: a potential of the form (14) with exponentially decaying (but also exponentially bounded away from zero at infinity) ϕ . While many important ingredients of [49] already appear in [39], the nonvanishing of ϕ was crucially important there in order to obtain lower bounds on eigenfunctions in a relatively simple way, without using the unique continuation principle, and leading to better probabilistic bounds, which in turn allowed for a conventional proof of localization at the end.

Most of this method can be adapted to the discrete Bernoulli–Anderson model, except for the unique continuation principle type result. It turns out that on the lattice eigenfunctions can vanish locally. Thus no reasonable analogue of the unique continuation holds. Let $\Omega = \{(x, y) \in \mathbb{Z}^2 : |x - y| > 1\} \subset \mathbb{Z}^2$. Then we have the following simple

THEOREM 2. *For any $v \in \ell^\infty(\Omega)$, there exist $w \in \ell^\infty(\mathbb{Z}^d)$ with $w|_\Omega = v$ such that $\Delta\phi = w\phi$ has a solution with $\phi(0) = 1$ and $\phi(x) = 0$ for $x \in \Omega$.*

PROOF. For $a > 0$ we set $w(x, x) = 0$ and $w(2n, 2n \pm 1) = a$, $w(2n \pm 1, 2n) = -a$. Then ϕ with $\phi(x, x) = 1$, $\phi(2n, 2n \pm 1) = -\frac{2}{a}$, $\phi(2n \pm 1, 2n) = \frac{2}{a}$, $\phi(m, n) = 0$, $(m, n) \in \Omega$, is a solution. \square

A similar statement obviously holds for \mathbb{Z}^d . This takes away a crucial piece of the argument of [49] if applied to \mathbb{Z}^d , thus proving localization for the Bernoulli–Anderson model remains more than a merely technical challenge.

PROBLEM 3. *Prove localization for the Bernoulli–Anderson model on \mathbb{Z}^d for $d > 1$ and either at high disorder or near the edges of the spectrum of the free Laplacian.*

Poisson model. The Poisson model is a model of random placement (rather than strength) of impurities. For a fixed countable set $X \subset \mathbb{R}^d$, we consider an operator

$$H_X := -\Delta + V_X \quad \text{on } L^2(\mathbb{R}^d), \quad \text{where} \quad (15)$$

$$V_X(x) := \sum_{\zeta \in X} u(x - \zeta). \quad (16)$$

Here u is some function (e.g., smooth, nonnegative, with compact support) and $u(x - \zeta)$ is the potential created by the impurity placed at ζ . The Poisson model is the family of operators H_X where X is modelled by a Poisson point process on \mathbb{R}^d with density $\varrho > 0$. This means that the number of points in each Borel subset of Lebesgue measure m has Poisson distribution with mean ϱm and the numbers of points of X in A_i are independent random variables for any finite disjoint collection $A_i \subset \mathbb{R}^d$. V_X is an ergodic potential with respect to translations in \mathbb{R}^d . This is another example of a model where random variables with densities are not available to perform the spectral averaging or obtain an a priori Wegner-type lemma, thus presenting a serious challenge. The Poisson Hamiltonian has Lifshitz tails—an indication of localization at the bottom of the spectrum [77, 53, 161, 135, 137, 182], but this problem has been open for many years other than in $d = 1$ where localization was established by Stolz [185], using a specially designed spectral averaging technique that, however, did not extend to higher dimensions. It turns out that the ideas and techniques of [49] are adaptable (albeit with some nontrivial technical difficulties) to the multidimensional Poisson model as well. This was shown in a recent work by Germinet, Klein, and Hislop [93] where they prove a.e. dynamical localization for H_X at the bottom of the spectrum: for $E < E_0(\varrho)$, or alternatively, at high disorder: for $\varrho < \varrho_0(E_0)$ and all $0 < E < E_0$.

2.2. Multidimensional Case: Extended States. It is strongly believed by physicists that in dimensions three and higher, the Anderson model has some interval of energy (high energy in the continuous case, near the spectrum of the

free Laplacian at small coupling in the lattice case) Σ_T characterized by the fact that states with energy localized in Σ_T propagate out of any compact set in \mathbb{R}^d (or \mathbb{Z}^d) almost surely. The phenomenon is loosely defined as extended states.

The physicists' explanation of this for $d \geq 3$ is based on their formulation of the problem as a nonlinear sigma model [80]. For the two-dimensional case, the corresponding β -function is believed to be nonnegative and monotone increasing. This indicates the absence of a phase transition as a function of the parameter that is a measure of the disorder. In dimension three, a mechanism of spontaneous symmetry breaking should lead to the occurrence of extended states.

One of the most important problems in the area is to prove a spectral or transport statement that would indicate those "extended states" for the Anderson model. (The problem is so widely known and acknowledged that I don't even list it as a separate problem.) While a number of conjectures have been made, it is not even entirely clear what mathematical statement one should expect.

The general belief is that there should be a region of purely absolutely continuous spectrum, with diffusive transport. So far, there are no existence results, neither on the spectral nor on the dynamical side. The results related to this region can be divided into two categories: the ones that prove extended states type effects for other related models, and the ones that are of conditional nature, which prove some properties of this elusive extended states region, without proving that it exists.

2.2.1. *Properties of the extended states region.* There are three remarkable results in this direction, on both spectral (due to Jaksic–Last) and dynamical (Germinet–Klein) sides:

THEOREM 3. (1) *The absolutely continuous spectrum of an Anderson model must be pure [114].*

(2) *The singular spectrum must be simple [116].*

Both statements actually do not require ergodicity, only *some* independence, thus they apply as well to decaying models with some randomness. At present, they do require a discrete setting. The second statement extends a theme (and, in some sense, methods) started by Simon in his famous(ly) short note [170] where he proved simplicity of the point spectrum in the regime of Anderson localization in the discrete case. Simplicity of singular spectrum is also known for general 1D operators [126, 101]; see also Simon [172] for a very simple proof.

The second statement is relevant to the topic "extended states region" in two ways. First, it gives a property of the possible singular continuous spectrum there. Second, it gives an extremely tempting idea for a proof of absolutely continuous spectrum: it reduces the problem to establishing noncyclicity of a certain vector.

On the dynamical side, it was shown in [98] that if there is a region (consisting of at least one point) where localization in a certain strong sense does not hold, then necessarily there is some transport. Namely, Germinet and Klein introduce the local transport exponent roughly defined as the power of growth of expectations of high moments of the position operator with respect to time, and show that in absence of strong localization (where it is equal to zero) it is necessarily at least $1/2d$. This result is obtained by showing that transport slower than that is sufficient to start the bootstrap multiscale analysis [96], which then implies strong dynamical localization in the Hilbert–Schmidt norm. Note that existence of an α -continuous component in the spectrum would imply a bound of the form α/d [150] but this

only goes in one direction, so a lower bound on the transport exponent can in principle coexist with pure point spectrum.

2.3. Other Models. The only model with i.i.d.r.v. potential where extended states were established is the Bethe lattice, \mathbb{B} . It is an infinite connected graph with no closed loops and a fixed degree, k , at each vertex. The Anderson Hamiltonian on the Bethe lattice is given by (2) with \mathbb{Z} replaced by \mathbb{B} . This is still an ergodic operator [3] and the spectrum is easily computable to be $\sigma(H_\lambda) = \sigma(\Delta) + \lambda \text{supp } \mu$ with probability one, where the spectrum of the free Laplacian $\sigma(\Delta)$ is equal to $[-\sqrt{k}, \sqrt{k}]$. This model was proposed in the physics literature [1, 2] as a model where an approximation for the study of localization becomes exact [1]. It was also argued in [1] that localization breaks down at an energy that converges to $\frac{k+1}{2}$ in the limit $\lambda \rightarrow 0$. A perturbative argument [156] showed that outside $[-\sqrt{k}, \sqrt{k}]$, the density of states and conductance vanish to all orders of perturbation theory. The existing rigorous results [7, 4, 130, 9, 88] seem to support the fact that there may be two transitions for this model.

THEOREM 4. *For sufficiently small λ , there are $0 < e(\lambda) < \sqrt{k} < \frac{k+1}{2} < \ell(\lambda)$ with $e(\lambda) \rightarrow \sqrt{k}$ and $\ell(\lambda) \rightarrow \frac{k+1}{2}$ as $\lambda \rightarrow 0$ such that*

- (1) H_λ has pure point spectrum and dynamical localization in $\pm[\ell(\lambda), \infty)$ [7, 4].
- (2) H_λ has purely absolutely continuous spectrum with ballistic behavior in $[-e(\lambda), e(\lambda)]$ [130, 9, 88].

The requirement for the localization part is the absolute continuity of the common distribution of the random variables, while for the extended states part, only finiteness of the second moment is needed.

The absolutely continuous spectrum was originally established by Klein by writing the expected value of the square of the modulus of the diagonal element of the Green's function as the fixed point of a certain nonlinear equation in supersymmetric variables. The result then follows by the implicit function theorem in the appropriate Banach space from the properties of the free Laplacian. However, writing of this nonlinear equation requires the loopless character of the graph. Thus the fact that the Bethe lattice, while in a sense infinitely dimensional, is also in some sense quasi-one-dimensional. This method does not seem to be transposable to the \mathbb{Z}^d —at least no progress has been made in this direction in over ten years since [130] appeared. Two alternative, somewhat simpler, proofs appeared recently in [9, 88]. The result in [9] is also made more general, allowing periodic background, and weakly correlated randomness instead of independence (see also [10] for a result on a related model with a different type of randomness). While the methods of [9, 88] are completely different from that of [130] as well as from each other, they both exploit the quasi-one-dimensionality and thus do not seem to be transposable to \mathbb{Z}^d either.

The bound $\frac{k+1}{2}$ appears in [4] as the edge of the ℓ^1 spectrum of the free Laplacian on the Bethe lattice. It is an extremely intriguing question: What happens between $e(\lambda)$ and $\ell(\lambda)$? It seems natural to conjecture singular continuous spectrum, but it is also natural in a loose analogy with random surfaces [115] or decaying potentials (see [75] in this Festschrift) to conjecture that ℓ^1 phenomena must prevail up to the ℓ^2 threshold, so there must be localization up to near the edge of the ℓ^2 spectrum.

PROBLEM 4. Describe the nature of the spectrum of H_λ in $[\sqrt{k}, \frac{k+1}{2}]$ for small λ .

Random Landau Hamiltonians. A random Landau Hamiltonian is an operator

$$H_{B,\lambda,\omega} = H_B + \lambda V_\omega \tag{17}$$

on $L^2(\mathbb{R}^2)$ where H_B is the free Landau Hamiltonian

$$H_B = (-i\nabla - \mathbf{A})^2 \tag{18}$$

where the vector potential \mathbf{A} is given by $\mathbf{A} = \frac{B}{2}(\mathbf{x}_2, -\mathbf{x}_1)$, $B > 0$ is the strength of the magnetic field, and $V_\omega(x) = \sum_{i \in \mathbb{Z}^2} \omega_i u(x - i)$ is a random potential. Here ω_i are i.i.d.r.v.'s with an absolutely continuous common distribution, and u is a certain bump function. This is a \mathbb{Z}^2 ergodic family of operators. The spectrum of the free Landau Hamiltonian consists of the set of so-called Landau levels: $B_n = (2n - 1)B$, that are all infinitely degenerate eigenvalues. Random Landau Hamiltonians describe a two-dimensional electron in a conductor with impurities and in a constant magnetic field perpendicular to the plane of the conductor. It is a model that has been linked with the quantum Hall effect. Physics results [153, 108] expect bands of extended states near the Landau levels separated by regions of localization and/or gaps.

At the same time, Thouless [186] argued that this extended states region may actually consist of singular continuous spectrum or even just one energy with an extended state. Localization at the edges of the bands for this model has been well understood [59, 188, 97]. Using the idea that the Hall conductance is constant in the gaps and jumps from one integer value to another across the Landau levels, Bellissard, van Elst, and Schulz-Baldes [28] (see also [6] for another derivation) showed (for the discrete version of this model) the existence of energies where a certain localization length diverges—a manifestation of extended states. More recently, Germinet, Klein, and Schenker, well equipped with [98], took this idea further to show that under certain conditions on the bump function and distribution, there is a nontrivial transport (i.e., dynamical delocalization) in each Landau band [99]—a strong manifestation of “extended states.” However, the exact nature of these extended states, even as simple a question as whether the effect is due to a band or a single energy as in [123], remains elusive.

PROBLEM 5. What is the nature of the metal-insulator transition for random Landau Hamiltonians? Is there any nonpoint/absolutely continuous spectrum?

Extended states were also established for several other Anderson-type models, such as random decaying potentials (see [75] for a review), sparse random potentials (e.g., [159]) or random surfaces (e.g., [115]), the models that are quite remarkable by the wealth of results in their own right. However, the nature of the “extended states” in all cases where it was established and understood seems to be different from what is expected for the Anderson model.

3. Quasiperiodic Operators

As discussed at the beginning, quasiperiodic potentials provide the only known model where the metal-insulator transition is well pronounced. Because of limited space, we will only discuss the discrete setting here (which is the one coming from physics), although there were some remarkable advances in the continuous setting

as well, e.g., a series of papers by Fedotov and Klopp (see [85, 86]). Neither the multiscale analysis (at least in the form devised for the Anderson model) nor the fractional moments method, works for quasiperiodic potentials as, among other reasons, quasiperiodicity does not allow for nice perturbations. The situation here is more difficult and the theory is far less developed than for the random case. With a few exceptions, the results are confined to the one-dimensional setting, and also the case of one frequency ($b = 1$) has been much better understood than that of higher frequencies.

The exotic phenomena, such as metal-insulator transitions, thick point spectrum, singular continuous, and Cantor spectrum abound in this family. However, there are several remarkable general results that make the theory of quasiperiodic operators a bit nicer than that of general ergodic operators.

THEOREM 5. *For any quasiperiodic potential:*

- (1) *The spectrum is the same for all phases θ [24].*
- (2) *For $d = 1$ the absolutely continuous spectrum is the same for all phases θ [152, 144].*

For general ergodic operators, both statements are only true for a.e. θ . The point and singular continuous parts (that are constant a.e.) can depend sensitively on θ [124]. However, there is no evidence of such sensitive dependence if they are combined together

PROBLEM 6. *Is singular spectrum the same for all phases θ ?*

All the existing evidence points to a positive answer, which is also conjectured by Simon. There is, however, a remarkable analogy between phases for quasiperiodic operators and rank one perturbations, which may be suggestive of a negative answer to Problem 6. But it is possible that this analogy is not relevant here. It is supported though by another general fact:

THEOREM 6. *The supports of singular parts of spectral measure are mutually disjoint for a.e. θ .*

This theorem appears in this form in [107], but it is essentially contained in Deift–Simon [69]. Another intriguing question in the same direction is whether Theorem 6 and the second statement of Theorem 5 hold for $d > 1$.

It is probably fair to say that much of the “hard analysis” of quasiperiodic operators has first been developed around the almost Mathieu operator, which is

$$H_{\lambda,\omega,\theta} = \Delta + \lambda v(\theta + n\omega) \tag{19}$$

acting on $\ell^2(\mathbb{Z})$, with $v : \mathbb{T} \rightarrow \mathbb{T}$; $v(\theta) = \cos(2\pi\theta)$, a model coming from physics. Several KAM-type approaches, starting with the pioneering work of Dinaburg–Sinai [76], were developed in the eighties and nineties for this or similar models in both large and small coupling regimes. Of those, the most robust and detailed is the reducibility result of Eliasson [81] that settled the (perturbative) case of small couplings for sufficiently regular potentials.

The common feature of these perturbative *results* in the quasiperiodic setting is that they provide no explicit estimates on how large (or small) the parameter λ should be and, more importantly, λ clearly depends on ω at least through the constants in the Diophantine characterization of ω . Thus for any given λ one cannot obtain a result for a.e. ω .

In contrast, the recently developed nonperturbative methods allow effective (in many cases even optimal) and, most importantly, independent of ω , estimates on λ . The latter property (uniform in ω estimates on λ) has often been taken as a definition of a nonperturbative result despite the fact that many such results require a certain degree of smallness.

The nonperturbative methods for localization are also quite different from the perturbative ones in that they do not employ multiscale schemes: usually only a few (from one to three) sufficiently large scales are involved, do not use the eigenvalue parametrization, and rely instead on direct estimates of the Green’s function. They are also significantly less involved, technically. One may think that in these latter respects they resemble the Aizenman–Molchanov method for random localization. It is, however, a superficial similarity as, on the technical side, they are still closer to and do borrow certain ideas from the multiscale analysis proofs of localization. On the other hand, all the existing nonperturbative methods (and therefore results) critically use (almost) analyticity while perturbative results (particularly for the large coupling) often require less regularity.

3.1. Cocycles and Lyapunov Exponents. In the one-dimensional case, many spectral questions can be recast in the dynamical language through the use of transfer matrices. For simplicity we will give the definitions only for the quasiperiodic case here, although they work with obvious adjustments for the general ergodic case as well. A solution $Hu = Eu$ satisfies $A(\theta + n\omega) \begin{pmatrix} u_n \\ u_{n-1} \end{pmatrix} = \begin{pmatrix} u_{n+1} \\ u_n \end{pmatrix}$ where

$$A(x) = S_{v,E}(x) = \begin{pmatrix} E - v(x) & -1 \\ 1 & 0 \end{pmatrix} \tag{20}$$

is the transfer matrix that can be viewed as a cocycle. A cocycle (ω, A) in general is defined by $\omega \in \mathbb{R}$ and an analytic map $A : \mathbb{R}/\mathbb{Z} \rightarrow SL(2, \mathbb{R})$. It is viewed as a linear skew-product $(x, w) \mapsto (x + \omega, A(x) \cdot w)$, $x \in \mathbb{R}/\mathbb{Z}$, $w \in \mathbb{R}^2$.

We say that two analytic cocycles $(\omega, A^{(i)})$, $i = 1, 2$, are analytically conjugate if there exists an analytic map $B : \mathbb{R}/2\mathbb{Z} \rightarrow SL(2, \mathbb{R})$ such that

$$A^{(2)}(x) = B(x + \omega)A^{(1)}(x)B(x)^{-1}. \tag{21}$$

The dynamical properties of cocycles are preserved by conjugacies.

For cocycles such that $A : \mathbb{R}/\mathbb{Z} \rightarrow SL(2, \mathbb{R})$ is homotopic to the identity, one can define the fibered rotation number by

$$\rho(\omega, A) = \int \psi d\mu \pmod{\mathbb{Z}} \tag{22}$$

where ψ is a lift of A and μ is a measure on $\mathbb{R}/\mathbb{Z} \times \mathbb{R}/\mathbb{Z}$ invariant with respect to $T : (x, y) \mapsto (x + \omega, y + \psi(x, y))$ and projecting over Lebesgue measure on the first coordinate [125, 111]. For Schrödinger cocycles (20), the rotation number admits a determination in $[0, \frac{1}{2}]$. It is linked to the integrated density of states by

$$N_{\lambda,\omega}(E) = 1 - 2\rho_{\lambda,\omega}(E) \in [0, 1]. \tag{23}$$

Here N is the usual integrated density of states for irrational ω and the integral of the density of states over different θ for $\omega \in \mathbb{Q}$; see [24, 125].

The *Lyapunov exponent* is defined by

$$\lim \frac{1}{n} \int \ln \|A_n(x)\| dx, \tag{24}$$

so $L(\omega, A) \geq 0$. It is invariant under conjugacy.

The Lyapunov exponent and the integrated density of states are linked through the Thouless formula (see [24]),

$$L(E) = \int \ln |E - E'| dN(E'). \quad (25)$$

Positivity of the Lyapunov exponent is equivalent to hyperbolicity of the corresponding cocycle. Uniform hyperbolicity corresponds to E in the resolvent set, and nonuniform hyperbolicity to E in the spectrum [125].

3.2. Positivity of Lyapunov Exponents. In the physics literature, positivity of the Lyapunov exponent is often taken as an implicit definition of localization, and the Lyapunov exponent is often called the inverse localization length.

If Lyapunov exponents are positive for all $E \in \mathbb{R}$, there is no absolutely continuous component in the spectrum for all θ (see, e.g., [62] for this result known as the Pastur–Ishii theorem). Positivity of Lyapunov exponents, however, does not imply localization or exponential decay of eigenfunctions (in particular, not for the Liouville ω nor for the resonant $\theta \in \mathbb{T}^b$).

To a large extent, nonperturbative methods, at least in their original form, stem from estimates involving the Lyapunov exponents and exploiting their positivity.

The general theme of the results on positivity of L , as suggested by perturbation arguments, is that the Lyapunov exponents are positive for large λ . This subject has had a rich history. The positivity was given several proofs, and all originally seemed to have a rather limited extendability. In particular, a proof by Herman [111] exploiting the subharmonicity, applied to trigonometric polynomials v . Herman's lower bound was in terms of the highest coefficient of the trigonometric polynomial and therefore this did not easily extend to the real analytic case. All the subsequent proofs, however, were also based on subharmonicity. Sorets–Spencer [179] proved that for nonconstant real analytic potentials v on \mathbb{T} ($b = 1$), one has $L > \frac{1}{2} \ln \lambda$ for $\lambda > \lambda(v)$ and all irrational ω . Another proof was given in [43], where this was also extended to the multifrequency case ($b > 1$) with, however, the estimate on λ dependent on the Diophantine condition on ω . Finally, Bourgain [41] proved that Lyapunov exponents are continuous in ω at every incommensurate ω (for $b > 1$; for $b = 1$ this was previously established in [48]), and that led to the following theorem which to date is the strongest result in this general context [41]:

THEOREM 7. *Let v be a nonconstant real analytic function on \mathbb{T}^b , and H be given by (1). Then, for $\lambda > \lambda(v)$, we have $L > \frac{1}{2} \ln \lambda$ for all E and all incommensurate vectors ω .*

For $b = 1$, this has been extended to potentials belonging to a Gevrey class [132]. In the perturbative setting, i.e., with results holding for $\lambda > \lambda(v, \omega)$ or for ω in a set of measure going to zero as $\lambda \rightarrow \infty$, positive Lyapunov exponents are known for $b = 1$ and $v \in C^2$ with v of cos-type (in a certain sense) [178, 92]. Removing the cos-type condition in the smooth category has been a subject of significant efforts and presents a serious challenge in nonuniformly hyperbolic dynamics. Bjerklöv [32] established positivity of Lyapunov exponents for energies outside a set of small (going to zero as $\lambda \rightarrow \infty$) measure. Chan [54] proved a perturbative result on positivity for all energies for typical (in a certain sense) C^3 potentials. On the other hand, generic continuous potentials have zero Lyapunov exponents generically

on the spectrum (see [33, 16]), and the results on discontinuity of the Lyapunov exponents [34] suggest that the same may be the case in C^1 .

3.3. Corollaries of Positive Lyapunov Exponents. As with the KAM methods, the almost Mathieu operator was the first model where the positivity of Lyapunov exponents was effectively exploited as the localization result of [118] used only positivity of the Lyapunov exponents (and no other properties of the coupling constant) as an ingredient.

The method in [118], while so far the only nonperturbative available allowing precise arithmetic conditions, uses some specific properties of the cosine. It extends to certain other situations, e.g., quasiperiodic operators arising from Bloch electrons in a perpendicular magnetic field, where the lattice is triangular or has next-nearest neighbor interactions [119]. However, it does not extend easily to the multifrequency or even general analytic potentials.

A much more robust method was developed by Bourgain–Goldstein [43] which allowed them to obtain a measure-theoretic version of the localization result for the general real analytic as well as the multifrequency case. Note that essentially no results were previously available for the multifrequency case, even perturbative.

THEOREM 8. *Let v be nonconstant real analytic on \mathbb{T}^b and H given by (12). Suppose $L(E, \omega) > 0$ for all $E \in [E_1, E_2]$ and a.e. $\omega \in \mathbb{T}^b$. Then for any θ , H has Anderson localization in $[E_1, E_2]$ for a.e. ω .*

Combining this with Theorem 7, Bourgain [41] obtained that for $\lambda > \lambda(v)$, H as above satisfies Anderson localization for a.e. ω . Theorem 8 was recently extended by S. Klein to potentials belonging to certain Gevrey classes [132]. One very important ingredient of this method is the theory of semi-algebraic sets that allows one to obtain polynomial algebraic complexity bounds for certain “exceptional” sets. Combined with measure estimates coming from the large deviation analysis of $\frac{1}{n} \ln \|M_n(\theta)\|$ (using subharmonic function theory and involving approximate Lyapunov exponents), this theory provides necessary information on the geometric structure of those exceptional sets. Such algebraic complexity bounds also exist for the almost Mathieu operator and are actually sharp, albeit trivial, in this case due to the specific nature of the cosine.

Further corollaries of positive Lyapunov exponents for analytic sampling functions f and $b = 1$ are summarized in the following theorem:

THEOREM 9. *If v is analytic and $L(E, \omega) > 0$ for all $\omega \in \mathbb{R} \setminus \mathbb{Q}$, E , then*

- (1) *For a.e. ω , the spectrum is a Cantor set [105].*
- (2) *For all ω, θ , the spectral measures have zero Hausdorff dimension ([122] contains this result for trigonometric polynomials).*
- (3) *For all ω, θ , the lower transport exponents are equal to zero ([67] contains this result for trigonometric polynomials).*
- (4) *For Diophantine ω and all θ , the upper transport exponents are equal to zero ([67] contains this result for trigonometric polynomials).*
- (5) *For ω satisfying a strong Diophantine condition, $\|q\omega\| \geq \frac{C}{q \ln^A q}$ for some C, A , the integrated density of states (and the Lyapunov exponent as a function of energy) is Hölder continuous [103] and absolutely continuous [104].*

- (6) For Diophantine ω , spectral gaps are almost Lipschitz (i.e., a logarithmic correction) continuous in frequency [120].
- (7) The measure of the spectrum is continuous in frequency [120].

Here, the upper and lower transport exponents refer to roughly the upper and lower (in time) power-law growth rates for the moments of the position operator. Lower exponent being zero corresponds to a very slow growth along a sequence of scales (which for Liouville ω may coexist with almost ballistic growth along another subsequence [150]) and upper being zero corresponds to “almost” dynamical localization (as far as power laws of growth are concerned), coexisting in some cases with singular continuous spectrum.

All the statements can be made local in energy and frequency in a natural way. The theorems combined in Theorem 9 are of widely varying levels of difficulty/depth. Some weaker statements are available for $b > 1$ [103] or v belonging to certain Gevrey classes [132]. The Hölder exponent can be estimated by $\frac{1}{2k} - \epsilon$ for f being in a small C^∞ neighborhood of a trigonometric polynomial of degree k [104], with an estimate becoming almost precise ($\frac{1}{2} - \epsilon$) for the almost Mathieu case (where a precise result is actually available but through the analysis of the dual regime; see Section 3.7.1). The argument of [105] uses some ideas of Sinai [178] who proved Cantor spectrum for cos-type potentials in the perturbative regime. [105] establishes quantitative estimates on separation of eigenvalues using techniques developed in [104]. Those techniques in turn extend the ones developed in [103]: sharp large deviation estimates for the norms [103] and elements [104] of the transfer matrix and the avalanche principle. The latter is a general fact about $SL(2, \mathbb{R})$ matrices that roughly allows one to propagate norm estimates for the large products, provided no two nearest neighbors “almost cancel” each other. This appeared implicitly in [191] and was made explicit and developed into a powerful method to study regularity properties of the Lyapunov exponents in [103].

A Diophantine-type condition is certainly necessary for Statements 4 [150] and 5 [40]. It is also expected to be necessary for Statement 6 (with only $\frac{1}{2}$ -Hölder regularity holding in general). It is not entirely clear whether it is necessary for the Cantor spectrum.

3.4. Almost Mathieu Operator and Simon’s Problems. The almost Mathieu operator is the central quasiperiodic model, mainly due to being the one coming from physics and attracting continued interest there. It first appeared in the work of Peierls [162] and arises as related, in two different ways, to a two-dimensional electron subject to a perpendicular magnetic field [109, 165]. It plays a central role in the Thouless et al. theory of the integer quantum Hall effect [187]. Also, it seems to represent most of the nontrivial properties expected to be encountered in the more general case. On the other hand, it has a very special feature: the duality (essentially a Fourier) transform (see Section 3.7.1) maps H_λ to $H_{4/\lambda}$, hence $\lambda = 2$ is the self-dual (also called critical) point. The development of the rigorous theory of this model (and of general quasiperiodic operators along with it) was strongly motivated by the numerical study of Hofstadter in 1976, the famous Hofstadter’s butterfly [112], and was guided for a long time by two conjectures formulated by Aubry and Andre in [13] and heavily popularized in several of Simon’s articles in the early eighties. One conjecture was on the metal-insulator transition at $\lambda = 2$ and the other one on the exact equality for the measure of the spectrum.

The related history up to early 2005 is presented in [151]. More recently, the developments were centered around “Simon’s problems”: three almost Mathieu related problems formulated in [171]. Two of these problems at the time of their formulation were unresolved only for zero measure sets of parameters, and including those problems in the list of fifteen highlighted the fact that it is starting to be perceived as more and more a “number-theory”-type problem where recent advances have made it possible to seek very precise information for all values of parameters.

Measure of the spectrum. This problem was the first to be fully resolved.

THEOREM 10. *The measure of the spectrum of $H_{\lambda,\omega,\theta}$ is equal to $|4 - 2|\lambda||$ for all irrational ω and all λ .*

That was exactly the original Aubry–Andre conjecture [13]. Note that for irrational ω , the spectrum as a set is independent of θ [24]. This theorem was proved for a.e. ω in [147, 148]. It was extended to complementary sets of full measure in [120] for noncritical λ and in [21] for the critical coupling (the latter alone was a subject of Problem 5 in [171]). The results for the critical point [148, 21] are certainly the most delicate here.

Metal-insulator transition. One remarkable feature of the almost Mathieu operator is that the Lyapunov exponents on the spectrum can be computed exactly:

THEOREM 11 ([48]). *Let $\lambda > 0$, $\omega \in \mathbb{R} \setminus \mathbb{Q}$. For every $E \in \Sigma_{\lambda,\omega}$, $L_{\lambda,\omega}(E) = \max\{\ln \lambda, 0\}$.*

Therefore, a transition from zero to positive Lyapunov exponents on the spectrum happens at $\lambda = 2$. This statement itself was actually one of the first rigorous facts proven for this model, with different proofs based on either the subharmonicity [111] or the duality [13, 24] (see Section 3.7.1 for more details). The original Aubry–Andre conjecture [13], based on this transition for the Lyapunov exponents and supported by the duality considerations, was that the spectrum is purely absolutely continuous for $\lambda < 2$ and pure point for $\lambda > 2$, without worrying about possible arithmetic issues, nor about the critical value $\lambda = 2$. It turned out to be true in the subcritical regime and true a.e. in the supercritical case.

THEOREM 12. *The almost Mathieu operator has*

- (1) *purely absolutely continuous spectrum for $\lambda < 2$ and all ω, θ .*
- (2) *purely singular continuous spectrum for $\lambda = 2$, all ω and a.e. θ .*
- (3) *pure point spectrum for $\lambda > 2$, Diophantine ω and a.e. θ .*

The first statement with “purely” replaced by “a lot of” was known since [147] (for a.e. θ , and a.e. ω including Liouville numbers; it was subsequently enhanced in [100]), and was quite a surprise at the time. It was then established as stated but for Diophantine ω and a.e. θ in [118] (using [107]). Extending it to all ω, θ has been considered a challenge—so much so that it was formulated as Problem 6 in [171]. It was solved only very recently with partial advances in [20] (Diophantine ω and all θ) and [17] (Liouville ω and a.e. θ) and the final solution by Avila in [15].

The second statement was known for (explicitly defined) a.e. ω [107], combining the duality theorem of [107] and [148]. It also followed by duality from the measure of the spectrum theorem of [21] for a complementary full measure set of ω . The a.e. θ part is probably an artifact of the duality proof. By Theorem 10, there is no absolutely continuous spectrum. There are also no eigenfunctions belonging to ℓ^1 [71]. Still one has an open

PROBLEM 7. *Prove that there are no eigenvalues for $\lambda = 2$.*

The third statement of Theorem 12 is due to [118], where a.e. θ is described through an explicit arithmetic condition, so called *nonresonant* property; in particular, this set contains 0. The theorem cannot be improved from “a.e.” to “every”—neither in frequency nor in phase because of

THEOREM 13. *$H_{\lambda,\omega,\theta}$ has purely singular continuous spectrum for $\lambda > 2$, Liouville ω , and all θ or all ω and a certain (arithmetically defined) dense G_δ of θ .*

This theorem is a combination of [106, 24] (Liouville case) and [124] (all ω case). Thus there is a very delicate dependence on the arithmetics of both ω and θ in the positive Lyapunov exponent regime. The recent advances, however, make it reasonable to expect that it may be possible to fully describe the spectral properties of H depending on ω, θ . For a.e. θ the answer should depend on the relation between λ and ω through the exponential rate of approximation of ω by the rationals. Namely, let p_n/q_n be the approximants of $\omega \in \mathbb{R} \setminus \mathbb{Q}$. Let

$$\beta = \beta(\omega) = \limsup \frac{\ln q_{n+1}}{q_n}. \quad (26)$$

PROBLEM 8. 1. *Show that for a.e. θ the spectrum is pure point with eigenfunctions decaying at a rate $\ln \frac{\lambda}{2} - \beta$ for $\lambda > 2e^\beta$.*

2. *Show that for $2 < \lambda < 2e^\beta$ the spectrum is purely singular continuous for all θ .*

The first statement is proved (without the statement on the rate of decay) in [19] for $\lambda > 2e^{\frac{16}{9}\beta}$, and the second one follows immediately from Gordon’s lemma [106] and the exact formula for the Lyapunov exponent [48] for $2 < \lambda < 2e^{\beta/2}$. It is reasonable, however, to expect the conjecture above to hold. It is harder to predict what should happen for the critical case $\lambda = 2e^\beta$; it is possible that the transition there would go deeper, depending on the finer properties of approximation of ω . It is also possible to make natural conjectures for the remaining zero measure set of θ depending on the exponential rate of growth of resonance strength.

Ten Martini Problem. If $\omega = p/q$ is rational, it is well known that the spectrum consists of the union of q intervals possibly touching at endpoints. In the case of irrational ω , the spectrum (which then does not depend on θ) has been conjectured for a long time to be a Cantor set for all $\lambda \neq 0$ [25]. To prove this conjecture has been dubbed the *Ten Martini Problem* by Barry Simon, after an offer of Kac in 1981; see Problem 4 in [171]. Note that for the critical point $\lambda = 2$, the result follows automatically from the zero measure of the spectrum statement.

In 1984 Bellissard and Simon [30] proved the conjecture for generic pairs of (λ, ω) . In 1987 Sinai [178] proved Cantor spectrum for a.e. ω in the perturbative regime: for $\lambda = \lambda(\omega)$ sufficiently large or small. In 1989 Helffer–Sjöstrand proved Cantor spectrum for the critical value $\lambda = 2$ and an explicitly defined generic set of ω [110]. Most developments in the nineties were related to the following observation. For $\omega = p/q$, the spectrum of $H_{\lambda,\omega,\theta}$ can have at most $q - 1$ gaps. It turns out that all these gaps are open, except for the middle one for even q (see, e.g., [55]). Choi, Elliott, and Yui obtained, in fact, an exponential lower bound on the size of the individual gaps, from which they deduced Cantor spectrum for Liouville (exponentially well approximated by the rationals) ω [55]. In 1994 Last, using certain estimates of Avron, van Mouche, and Simon [22], proved zero measure

Cantor spectrum for a.e. ω (for an explicit set that intersects with but does not contain the set in [110]) and $\lambda = 1$ [148].

A major breakthrough came recently with an influx of ideas coming from dynamical systems. Puig, using Aubry duality [13] and localization for $\theta = 0$ and $\lambda > 2$ [118], proved Cantor spectrum for Diophantine ω and any noncritical λ [164]. At about the same time, Avila and Krikorian proved zero measure Cantor spectrum for $\lambda = 2$ and ω satisfying a certain Diophantine condition, therefore extending the result of Last to all irrational ω [21]. The solution of the Ten Martini Problem as originally stated was finally given in [19]:

THEOREM 14 ([19]). *The spectrum of the almost Mathieu operator is a Cantor set for all irrational ω and for all $\lambda \neq 0$.*

Cantor structure of the spectrum for the Diophantine ω [164] follows from localization for the phase $\theta = 0$, with corresponding eigenvalues being the boundaries of noncollapsed gaps. The key idea here is that for energies dual to eigenvalues of $H_{\omega,0}$, corresponding to localized eigenfunctions, the rotation number of the transfer-matrix cocycle is of the form $k\omega \pmod{\mathbb{Z}}$. Thus they are the ends of the gaps (possibly collapsed). However, a collapsed gap in this case would correspond to reducibility of the system to the identity which can be shown to contradict the simplicity of pure point spectrum for the dual model. Since those energies form a dense subset of the spectrum, the result follows.

The argument of [164] relies on two small denominator problems (one needed for the proof of localization and another for the solution of a cohomological equation), thus it cannot be extended beyond a certain Diophantine condition. The argument of [55] for the Liouville side also has obvious technical limitations. That left open, even after pushing the existing approaches to absolute technical limits, the values of parameters belonging to the arithmetically critical region: $2e^{-2\beta} < \lambda < 2e^{2\beta}$ where localization is not expected to hold at the gap edges yet where ω is not close enough to the rationals to guarantee, by simple approximation, a noncollapse of small gaps (that are known to be present). The key to treating the critical region was an analytic extension of Kotani theory [19], showing that m -functions, that by the Kotani theory can be analytically extended through an interval with zero Lyapunov exponents, are also analytic in phase. That allowed one, under the assumption of no Cantor spectrum, to obtain much stronger, unrealistically good, estimates on the modulus of continuity (in the Hausdorff topology) of the spectrum as a function of the frequency, for the Liouville side. For the Diophantine side, it allowed one to effectively bring the problem down to the question of reducibility of the cocycle of rotations, which is a much simpler small denominator problem. See [18] for an account of ideas in the proof of [19]. This analytic extension of Kotani theory was further developed by Avila in [14], where it was used to show density of positive Lyapunov exponents for analytic quasiperiodic $SL(2, \mathbb{R})$ cocycles over arbitrary irrational rotations of \mathbb{T}^b .

3.5. Multidimensional Case: $d > 1$. As mentioned above, there are very few results in the multidimensional lattice case ($d > 1$). Essentially, the only result that existed before the recent developments was a perturbative theorem—an extension by Chulaevsky–Dinaburg [56] of Sinai’s method [178] to the case of operator (1) on $\ell^2(\mathbb{Z}^d)$ with $V_n = \lambda v(n \cdot \omega)$, $\omega \in \mathbb{R}^d$, where v is a cos-type function on \mathbb{T} . This also holds nonperturbatively for any real-analytic v (see [40]). Note

that since $b = 1$, this avoids most serious difficulties and is therefore significantly simpler than the general multidimensional case. We therefore have:

THEOREM 15. *For any $\epsilon > 0$, there is $\lambda(v, \epsilon)$, and for $\lambda > \lambda(v, \epsilon)$, $\Omega(\lambda, v) \subset \mathbb{T}^d$ with $\text{mes}(\Omega) < \epsilon$, so that for $\omega \notin \Omega$, operator (1) with V_n as above has Anderson localization.*

This should be confronted with the following theorem of Bourgain [38, 40]:

THEOREM 16. *Let $d = 2$ and $v(\theta) = \cos 2\pi\theta$ in $H = H_\omega$ defined as above. Then for any λ the measure of ω s.t. H_ω has some continuous spectrum is positive.*

Therefore, for large λ there will be both ω with complete localization as well as those with at least some continuous spectrum. This shows that nonperturbative results do not hold in general in the multidimensional case! Perturbative results, however, had been obtained; see the next section.

A similar (in fact, dual) situation is observed for one-dimensional multifrequency ($d = 1$; $b > 1$) case at small disorder. One has, by duality:

THEOREM 17. *Let H be given by (19) with $\theta, \omega \in \mathbb{T}^b$ and v real analytic on \mathbb{T}^b . Then for any $\epsilon > 0$, there is $\lambda(v, \epsilon)$ s.t. for $\lambda < \lambda(v, \epsilon)$ there is $\Omega(\lambda, v) \subset \mathbb{T}^b$ with $\text{mes}(\Omega) < \epsilon$ so that for $\omega \notin \Omega$, H has purely absolutely continuous spectrum.*

And also

THEOREM 18. *Let $d = 1$, $b = 2$, and v be a trigonometric polynomial on \mathbb{T}^2 with a nondegenerate maximum. Then for any λ measure of ω s.t. H_ω has some point spectrum, dense in a set of positive measure, is positive.*

Therefore, unlike the $b = 1$ case (see Theorem 21), nonperturbative results do not hold for absolutely continuous spectrum at small disorder.

3.6. Perturbative Localization. The key perturbative localization results are [178, 92, 82], and some extensions. [178] and [92] do not use analytic methods and thus the results establish localization under just the C^2 (plus a cos-type) condition, something unmatched by the nonperturbative methods. Sinai's theorem also establishes Cantor spectrum under the same C^2 cos-type condition. Eliasson's result works for general Gevrey-class potentials.

While the previous section demonstrates the limitations of the nonperturbative results, the nonperturbative methods have been applied to significantly simplify the proofs and obtain new perturbative results that previously had been completely beyond reach.

Many such applications that are outside the scope of this article are described in [40]. In particular, new results on the construction of quasiperiodic solutions in Melnikov problems and nonlinear PDE's, obtained by using certain ideas developed for nonperturbative quasiperiodic localization (e.g., the theory of semi-algebraic sets) are presented there. Other results in this group contain localization for the skew-shift model [44], almost periodicity for the quantum kicked rotor model [36, 40], and localization for potentials in higher Gevrey classes [132].

The main goal in a nonperturbative method is to obtain exponential off-diagonal decay for the matrix elements of the Green's function of box-restricted operators along with subexponential bounds on the distance from the spectrum of

such box-restrictions to a given energy. From that result one can obtain localization through elimination of energy via an argument involving complexity bounds on semi-algebraic sets (see [40]).

A nonperturbative way to achieve the desired Green’s function estimates uses Cramer’s rule to represent the matrix elements of the resolvent. Then, in the one-dimensional (in space) case, it is often possible to obtain the estimates from the positivity of Lyapunov exponents: uniformly for the numerator, and from large deviation bounds for the subharmonic functions for the denominator. This is done in one step for a sufficiently large scale (see Section 3.3.)

A perturbative way consists of establishing the desired estimates in a multiscale scheme: namely, the estimates are proved outside a set of parameters of (subexponentially) decaying (in the size of the box) measure. Moreover, this set should be shown to have a semi-algebraic description, in order to make possible sublinear upper bounds on the number of times a trajectory of a given phase (under the underlying rotation or other ergodic transformation of the torus) hits the “forbidden” set. This, plus certain subharmonic function arguments, allows passage to a larger scale through a repeated use of the resolvent identity.

An application that is most relevant to the current article is localization for a “true” $d > 1$ situation. The best currently available result is the following recent theorem [42]:

THEOREM 19. *Let $d = b$ and let v be real analytic on \mathbb{T}^d such that for all $i = 1, \dots, d$ and $(\theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_d) \in \mathbb{T}^{d-1}$, the map*

$$\theta_i \mapsto v(\theta_1, \dots, \theta_i, \dots, \theta_d)$$

is a nonconstant function of $\theta_i \in \mathbb{T}$. Then for any $\epsilon > 0$, there is $\lambda(v, \epsilon)$ s.t. for $\lambda > \lambda(v, \epsilon)$, there is $\Omega(\lambda, v) \subset \mathbb{T}^d$ with $\text{mes}(\Omega) < \epsilon$ so that for $\omega \notin \Omega$ operator (1) with $V_n = \lambda v(n_1\omega_1, \dots, n_d\omega_d)$ has Anderson (and dynamical) localization.

This result was obtained previously, for $d = 2$ only in [45]. There were some serious purely arithmetic difficulties that prevented an extension of this result to higher dimensions. In the previous results on localization, there were two major steps: estimates on the Green’s function for a fixed energy and elimination of energy. The main difficulty in the multidimensional case lies in establishing the sublinear bound, described above, that enters in the first step. It is for this bound that an arithmetic condition on ω was needed. The condition used was to guarantee that the number of $(n_1, n_2) \in [1, N]^2$ such that $(n_1\omega_1, n_2\omega_2) \pmod{\mathbb{Z}^2} \in S$ is bounded from above by N^ω for some $\omega < 1$, uniformly for all semi-algebraic sets S of degree D , with $\frac{D'}{D} = o(\frac{1}{N})$ and with the measure of all horizontal and vertical sections S_x satisfying $\log \text{mes} S_x = o(\log \frac{1}{N})$. This condition roughly means that too many points close to an algebraic curve of a bounded degree would force it to oscillate more than it should. Such a statement is essentially two-dimensional and not extendable to $d \geq 3$. In Theorem 19, Bourgain circumvents it by using from the beginning the theory of semi-algebraic sets to eliminate energy and the translation variable to get conditions on ω (that depend on the potential) already in the first step.

3.7. Small Coupling. Only the one-dimensional case is well understood. The first application of KAM in this setting is due to Dinaburg–Sinai [69]. The most comprehensive perturbative result belongs to Eliasson [81]

THEOREM 20 ([81]). *For Diophantine ω and $V : \mathbb{T}^d \rightarrow \mathbb{R}$ analytic in the strip of width r , there exist $\epsilon = \epsilon(r, \omega)$ such that for $|V| < \epsilon$ and a.e. E the corresponding cocycle is analytically reducible. The spectrum of H is purely absolutely continuous for all θ .*

The dependence of ϵ on ω is through the constants in the Diophantine class. The set of E is characterized by the rotation number of the cocycle being either Diophantine or rational with respect to ω . For other energies in the spectrum, Eliasson’s theorem gives at most linear growth of the norms of the transfer-matrices (sublinear in the case of the irrational rotational number). While Eliasson’s theorem was proved for the continuous setting, the discrete version as above follows similarly. Also, what comes from Eliasson’s method is a more general statement on arbitrary (not necessarily Schrödinger) analytic quasiperiodic cocycles that are close to constant cocycles.

3.7.1. Nonperturbative results. For Schrödinger operators with analytic potential and $b = d = 1$, a nonperturbative method based on duality (in one way or another) and localization for the dual (generally long range) model is now available. Let $\mathcal{H} = L^2([0, 2\pi], d\theta/2\pi \times \mathbb{Z})$. The duality transform is a unitary operator $U : \mathcal{H} \rightarrow \mathcal{H}$ defined by $(U\phi)(\eta, m) = \hat{\phi}(m, \eta + \pi\omega m)$. We then define $\hat{H} = U^{-1}HU$ where by H we understand the direct integral in θ of $H_{\omega, \theta}$. It is shown in [107] that for H of the form (19) with $v \in \ell^2$, the pure point spectrum for a.e. θ for \hat{H} implies absolutely continuous spectrum for a.e. θ for H . For the almost Mathieu operator $\hat{H}_\lambda = H_{4/\lambda}$, leading to the duality property for the almost Mathieu family. For other potentials, \hat{H} is usually long range, thus the methods of proving localization that are based on the Lyapunov exponents do not work. A non-Lyapunov exponent based method for such one- (and quasi-one-) dimensional models was developed in [47]. This method studies large deviations for the quantities of the form $\frac{1}{n} \ln |\det(H - E)_\Lambda|$ and uses path-determinant expansion for the matrix elements of the resolvent. For the cosine potential, many ideas of [118] apply (see [47]). As a result one obtains

THEOREM 21. *Let H be an operator (19), where v is real analytic on \mathbb{T} and ω is Diophantine. Then, for $\lambda < \lambda(v)$, H has purely absolutely continuous spectrum for a.e. θ .*

We note that an analogue of this theorem does not hold in the multifrequency case (see Section 3.5).

A certain measure-theoretic version of it, allowing nonlocal Laplacians but leading only to continuous spectrum, is also available; see [40]. The same ideas allow extension of quasiperiodic localization results to the strip of arbitrary dimension [46].

An important feature of these nonperturbative results [118, 47] is that the duality theorem of [107] is used as a black box, and all the analysis is performed for the dual model $\hat{H} = \hat{H}_{v, \omega, \theta}$ defined on $\ell^2(\mathbb{Z})$

$$(\hat{H}\hat{u})_n = \sum \hat{v}_k \hat{u}_{n-k} + 2 \cos(2\pi\theta + n\omega) \hat{u}_n, \tag{27}$$

where \hat{v}_k are the Fourier coefficients of $v(x) = \sum \hat{v}_k e^{2\pi i k x}$. The existence of an exponentially decaying eigenfunction with eigenvalue E for some θ for the dual model corresponds to reducibility for the original cocycle at E . While localization (and reducibility) fail generically (even in the perturbative regime), it turns out that for all

energies and nonperturbatively small couplings, one can establish so-called almost localization for the dual model, which means exponential decay away from a sparse sequence of resonances [20]. The main advantage of the cosine potential in (27) is that it leads to the sequence of resonances (finite for a.e. θ but infinite for a dense G_δ of θ) being the same for all energies, while for other potentials resonances may depend on energy which leads to the need for the elimination of energy techniques. These almost localization estimates can be used to extend Eliasson’s result to the nonperturbative setting, with the optimal smallness requirement in the case of the almost Mathieu operator.

We will say that a cocycle is almost reducible if it is analytically conjugate to a cocycle in Eliasson’s perturbative regime. The dynamical properties and thus spectral consequences that hold in this regime should then follow for almost reducible cocycles and associated operators. It is shown in [20] that almost localization estimates for \hat{H} lead to almost reducibility for the cocycle associated with H , and thus such almost reducibility can be established for all energies, which allows for making very precise spectral conclusions. The idea of reducing the nonperturbative regime to the perturbative one was already realized by Avila and Krikorian [21] for a.e. E in the spectrum throughout the regime of zero Lyapunov exponents. The analysis of [21] cannot work, however, for all energies, particularly for the gap edges which are responsible for some important effects. It should be mentioned, however, that the dynamical estimates of [20] allow to obtain the spectral and dynamical consequences directly by nonperturbative analysis, without referring to Eliasson’s theorem. Thus, in particular, a larger range of Diophantine frequencies can be covered.

This nonperturbative analysis leads to the following

THEOREM 22 ([20]). *Under the conditions of Theorem 21 and for the same value of $\lambda(v)$, we have*

1. *The integrated density of states is $\frac{1}{2}$ -Hölder.*
2. *The spectral measures are absolutely continuous for all phases.*

For the almost Mathieu case, $v = \cos$, the value of $\lambda(v)$ is optimal: $\lambda(v) = 2$.

The first statement was recently obtained by Ben Hadj Amor in Eliasson’s perturbative regime [31]. Other results on this topic include [35] where $\frac{1}{2} - \epsilon$ regularity was obtained for the almost Mathieu operator in the perturbative regime, and [104] where $\frac{1}{2} - \epsilon$ regularity was shown throughout the noncritical regime for the almost Mathieu operator under a strong Diophantine condition. Goldstein and Schlag [103, 104] also obtained regularity for analytic v in the regime of positive Lyapunov exponents but at the expense of a worse Hölder constant (see Theorem 9 and the discussion after it).

The first statement of Theorem 22 is optimal in several ways. Under the conditions of the theorem, there are square-root singularities for the integrated density of states at the gap edges, so $\frac{1}{2}$ -Hölder cannot be improved. It is easy to see that for a fixed $|\lambda| > 0$ and generic ω , the integrated density of states is not Hölder, and Bourgain [40] showed that for $v = \cos$ and $|\lambda| = 1$ even a fairly mild Diophantine condition is not enough to guarantee Hölder continuity.

The second statement holds for the almost Mathieu operator without a Diophantine condition on frequency (see Theorem 12, [15]). It therefore is natural to conjecture

PROBLEM 9. *Prove that for $\lambda < \lambda(v)$, the spectral measures are absolutely continuous for all phases and all irrational frequencies.*

It is known [149] that for even $v \in C^{3/2+\epsilon}$ there is a large absolutely continuous component for a set of ω that is of full measure and includes all non-Diophantine ω . For even analytic v , this can be combined with Theorem 22 to obtain the existence statement for all ω . However, nothing is known for non-even v , nor about purity other than in the almost Mathieu case.

The analysis of [20] also provides a strengthening of the Ten Martini Problem: a solution of the so-called Dry Ten Martini Problem for the case of Diophantine frequencies.

The gap labeling theorem (see [27]) states that the integrated density of states takes values of the form $k\omega + m \pmod{\mathbb{Z}}$ on each gap in the spectrum. The Dry Ten Martini Problem is to show that all gaps predicted by the gap labeling theorem are actually open in the case of the almost Mathieu operator. It was obtained for Liouville ω [55], as well as for Diophantine ω in perturbative regime, $\lambda > \lambda(\omega)$ ([164], based on [82]). We have

THEOREM 23. *The dry version of the Ten Martini Problem holds for Diophantine ω , $\lambda \neq -2, 0, 2$.*

This leaves open the question for a regime intermediate between Diophantine and Liouville frequencies (which is arithmetically critical, as argued in [19]) as well as the critical value of $\lambda = 2$ and all frequencies.

PROBLEM 10. *Prove the Dry Ten Martini Problem for all values of parameters.*

This would also have some interesting implications for potentials of the form $\cos(\omega n^\tau)$ with $1 < \tau < 2$ [152].

Acknowledgment. I am grateful to D. Damanik, A. Klein, and B. Simon for their suggestions that led to a significant improvement of this manuscript, and to F. Gesztesy for his encouragement.

References

- [1] R. Abou-Chacra, P. Anderson, and D. J. Thouless, *A selfconsistent theory of localization*, J. Phys. C: Solid State Phys. **6**, 1734–1752 (1973).
- [2] R. Abou-Chacra and D. J. Thouless, *Selfconsistent theory of localization. II. Localization near the band edges*, J. Phys. C: Solid State Phys. **7**, 65–75 (1974).
- [3] V. Acosta and A. Klein, *Analyticity of the density of states in the Anderson model on the Bethe lattice*, J. Statist. Phys. **69**, 277–305 (1992).
- [4] M. Aizenman, *Localization at weak disorder: some elementary bounds*, Rev. Math. Phys. **6**, 1163–1182 (1994).
- [5] M. Aizenman, A. Elgart, S. Naboko, J. Schenker, and G. Stolz, *Moment analysis for localization in random Schrödinger operators*, Invent. Math. **163**, 343–413 (2006).
- [6] M. Aizenman and G.-M. Graf, *Localization bounds for an electron gas*, J. Phys. A: Math. Gen. **31**, 6783–6806 (1998).
- [7] M. Aizenman and S. Molchanov, *Localization at large disorder and at extreme energies*, Comm. Math. Phys. **157**, 245–278 (1993).
- [8] M. Aizenman, J. Schenker, R. Friedrich, and D. Hundertmark, *Finite-volume fractional-moment criteria for Anderson localization*, Dedicated to Joel L. Lebowitz, Comm. Math. Phys. **224**, 219–253 (2001).
- [9] M. Aizenman, R. Sims, and S. Warzel, *Absolutely continuous spectra of quantum tree graphs with weak disorder*, Comm. Math. Phys. **264**, 371–389 (2006).

- [10] M. Aizenman, R. Sims, and S. Warzel, *Stability of the absolutely continuous spectrum of random Schrödinger operators on tree graphs*, to appear in Probab. Theory Related Fields.
- [11] P. W. Anderson, *Absence of diffusion in certain random lattices*, Phys. Rev. **109**, 1492–1501 (1958).
- [12] A. J. Antony and M. Krishna, *Inverse spectral theory for Jacobi matrices and their almost periodicity*. Spectral and Inverse Spectral Theory (Bangalore, 1993). Proc. Indian Acad. Sci. Math. Sci. **104**, 777–818 (1994).
- [13] S. Aubry and G. Andre, *Analyticity breaking and Anderson localization in incommensurate lattices*, Ann. Israel Phys. Soc. **3**, 133–140 (1980).
- [14] A. Avila, *Density of positive Lyapunov exponents for quasiperiodic $SL(2, \mathbb{R})$ cocycles in arbitrary dimension*, preprint, 2005.
- [15] A. Avila, *Absolutely continuous spectrum for the almost Mathieu operator with subcritical coupling*, preprint, 2006.
- [16] A. Avila and D. Damanik, *Generic singular spectrum for ergodic Schrödinger operators*, Duke Math. J. **130** (2005), 393–400.
- [17] A. Avila and D. Damanik, *Absolute continuity of the integrated density of states for the almost Mathieu operator with non-critical coupling*, preprint, 2006.
- [18] A. Avila and S. Jitomirskaya, *Solving the ten martini problem*, Proc. QMath 9 (Giens, 2004), Springer Lecture Notes in Physics **690**, 5–16 (2006).
- [19] A. Avila and S. Jitomirskaya, *The ten martini problem*, to appear in Annals of Math.
- [20] A. Avila, S. Jitomirskaya, *Almost localization and almost reducibility*, preprint, 2006.
- [21] A. Avila and R. Krikorian, *Reducibility or non-uniform hyperbolicity for quasiperiodic Schrödinger cocycles*, to appear in Annals of Math.
- [22] J. Avron, P. van Mouche, and B. Simon, *On the measure of the spectrum for the almost Mathieu operator*, Comm. Math. Phys. **132**, 103–118 (1990).
- [23] J. Avron and B. Simon, *Almost periodic Schrödinger operators. I. Limit periodic potentials*, Comm. Math. Phys. **82**, 101–120 (1981/82).
- [24] J. Avron and B. Simon, *Almost periodic Schrödinger operators. II. The integrated density of states*, Duke Math. J. **50**, 369–391 (1983).
- [25] M. Ya. Azbel, *Energy spectrum of a conduction electron in a magnetic field*, Sov. Phys. JETP **19**, 634–645 (1964).
- [26] J.-M. Barbaroux and S. Tcheremchantsev, *Universal lower bounds for quantum diffusion*, J. Funct. Anal. **168**, 327–354 (1999).
- [27] J. Bellissard, A. Bovier, and J.-M. Ghez, *Gap labelling theorems for one dimensional discrete Schrödinger operators*, Rev. Math. Phys. **4**, 1–37 (1992).
- [28] J. Bellissard, A. van Elst, and H. Schulz-Baldes, *The non commutative geometry of the quantum Hall effect*, J. Math. Phys. **35** 5373–5451 (1994).
- [29] J. Bellissard, B. Iochum, and D. Testard, *Continuity properties of the electronic spectrum of 1D quasicrystals*, Comm. Math. Phys. **141**, 353–380 (1991).
- [30] J. Bellissard and B. Simon, *Cantor spectrum for the almost Mathieu equation*, J. Funct. Anal. **48**, 408–419 (1982).
- [31] S. Ben Hadj Amor, *Opérateurs de Schrödinger quasi-périodiques uni-dimensionnels*, Ph.D. thesis, Université Paris 7, 2006.
- [32] K. Bjerklov, *Positive Lyapunov exponent and minimality for a class of one-dimensional quasi-periodic Schrödinger equations*, Ergodic Theory Dynam. Systems **25**, 1015–1045 (2005).
- [33] J. Bochi, *Discontinuity of the Lyapunov exponent for non-hyperbolic cocycles*, preprint.
- [34] J. Bochi and M. Viana, *The Lyapunov exponents of generic volume preserving and symplectic systems*, Annals of Math. **161**, 1423–1485 (2005).
- [35] J. Bourgain, *Hölder regularity of integrated density of states for the almost Mathieu operator in a perturbative regime*, Lett. Math. Phys. **51**, 83–118 (2000).
- [36] J. Bourgain, *Estimates on Green's functions, localization and the quantum kicked rotor model*, Annals of Math. **156**, 249–294 (2002).
- [37] J. Bourgain, *New results on the spectrum of lattice Schrödinger operators and applications*, Contemp. Math. **307**, 27–38 (2002).
- [38] J. Bourgain, *On the spectrum of lattice Schrödinger operators with deterministic potential, I, II*, J. Anal. Math. **87**, 37–75 (2002); **88**, 221–254 (2002).

- [39] J. Bourgain, *On localization for lattice Schrödinger operators involving Bernoulli variables*, Geometric Aspects of Functional Analysis, pp. 77–99, Lecture Notes in Math., 1850, Springer, Berlin, 2004.
- [40] J. Bourgain, *Green's Function Estimates for Lattice Schrödinger Operators and Applications*, Annals of Mathematics Studies, 158. Princeton University Press, Princeton, NJ, 2004.
- [41] J. Bourgain, *Positivity and continuity of the Lyapunov exponent for shifts on \mathbb{T}^d with arbitrary frequency vector and real analytic potential*, J. Anal. Math. **96**, 313–355 (2005).
- [42] J. Bourgain, *Anderson localization for quasi-periodic lattice Schrödinger operators on \mathbb{Z}^d , d arbitrary*, to appear in Geom. Funct. Anal.
- [43] J. Bourgain and M. Goldstein, *On nonperturbative localization with quasiperiodic potential*, Annals of Math. **152**, 835–879 (2000).
- [44] J. Bourgain, M. Goldstein, and W. Schlag, *Anderson localization for Schrödinger operators on \mathbb{Z} with potentials given by the skew shift*, Comm. Math. Phys. **220**, 583–621 (2001).
- [45] J. Bourgain, M. Goldstein, and W. Schlag, *Anderson localization on \mathbb{Z}^2 with quasiperiodic potential*, Acta Math. **188**, 41–86 (2002).
- [46] J. Bourgain and S. Jitomirskaya, *Anderson localization for the band model*, Geometric Aspects of Functional Analysis, pp. 67–79, Lecture Notes in Math., 1745, Springer, Berlin, 2000.
- [47] J. Bourgain and S. Jitomirskaya, *Absolutely continuous spectrum for 1D quasiperiodic operators*, Invent. Math. **148**, 453–463 (2002).
- [48] J. Bourgain and S. Jitomirskaya, *Continuity of the Lyapunov exponent for quasiperiodic operators with analytic potential*, Dedicated to David Ruelle and Yasha Sinai on the occasion of their 65th birthdays, J. Statist. Phys. **108**, 1203–1218 (2002).
- [49] J. Bourgain and C. Kenig, *On localization in the continuous Anderson–Bernoulli model in higher dimension*, Invent. Math. **161**, 389–426 (2005).
- [50] J. Bourgain and W. Schlag, *Anderson localization for Schrödinger operators on \mathbb{Z} with strongly mixing potentials*, Comm. Math. Phys. **215**, 143–175 (2000).
- [51] R. Carmona, *One-dimensional Schrödinger operators with random or deterministic potentials: new spectral types*, J. Funct. Anal. **51**, 229–258 (1983).
- [52] R. Carmona, A. Klein, and F. Martinelli, *Anderson localization for Bernoulli and other singular potentials*, Comm. Math. Phys. **108**, 41–66 (1987).
- [53] R. Carmona and J. Lacroix, *Spectral Theory of Random Schrödinger operators*, Birkhäuser, Boston, 1990.
- [54] J. Chan, *Method of variation of potential of quasiperiodic Schrödinger equation*, to appear in Geom. Funct. Anal.
- [55] M. D. Choi, G. A. Elliott, and N. Yui, *Gauss polynomials and the rotation algebra*, Invent. Math. **99**, 225–246 (1990).
- [56] V. Chulaevsky and E. Dinaburg, *Methods of KAM theory for long range quasi-periodic operators on \mathbb{Z}^n* , Comm. Math. Phys. **153**, 559–577 (1993).
- [57] V. Chulaevsky and T. Spencer, *Positive Lyapunov exponents for a class of deterministic potentials*, Comm. Math. Phys. **168**, 455–466 (1995).
- [58] J. M. Combes and P. D. Hislop, *Localization for some continuous, random Hamiltonians in d -dimensions*, J. Funct. Anal. **124**, 149–180 (1994).
- [59] J. M. Combes and P. D. Hislop, *Landau Hamiltonians with random potentials: localization and the density of states*, Comm. Math. Phys. **177** 603–629 (1996).
- [60] J. M. Combes, P. D. Hislop, F. Klopp, and G. Raikov, *Global continuity of the integrated density of states for random Landau Hamiltonians*, Comm. Partial Differential Equations **29**, 1187–1213 (2004).
- [61] W. Craig and B. Simon, *Log Hölder continuity of the integrated density of states for stochastic Jacobi matrices*, Comm. Math. Phys. **90**, 207–218 (1983).
- [62] H. L. Cycon, R. G. Froese, W. Kirsch, and B. Simon, *Schrödinger Operators*, Springer, Berlin–Heidelberg–New York, 1987.
- [63] D. Damanik, *Strictly ergodic subshifts and associated operators*, in this Festschrift.
- [64] D. Damanik and R. Killip, *Ergodic potentials with a discontinuous sampling function are non-deterministic*, Math. Res. Lett. **12**, 187–192 (2005).
- [65] D. Damanik, R. Sims, and G. Stolz, *Localization for one dimensional, continuum, Bernoulli–Anderson models*, Duke Math. J. **114**, 59–100 (2002).

- [66] D. Damanik and P. Stollmann, *Multi-scale analysis implies strong dynamical localization*, Geom. Funct. Anal. **11**, 11–29 (2001).
- [67] D. Damanik and S. Tcheremchantsev, *Upper bounds in quantum dynamics*, preprint, 2005.
- [68] S. de Bièvre and F. Germinet, *Dynamical localization for the random dimer Schrödinger operator*, J. Stat. Phys. **98**, 1135–1148 (2000).
- [69] P. Deift and B. Simon, *Almost periodic Schrödinger operators. III. The absolutely continuous spectrum in one dimension*, Comm. Math. Phys. **90**, 389–411 (1983).
- [70] R. del Rio, S. Jitomirskaya, Y. Last, and B. Simon, *Operators with singular continuous spectrum: IV. Hausdorff dimension, rank one perturbations and localization*, J. d'Analyse Math. **69**, 153–200. (1996).
- [71] F. Delyon, *Absence of localization for the almost Mathieu equation*, J. Phys. A **20**, L21–L23 (1987).
- [72] F. Delyon, H. Kunz, and B. Souillard, *One dimensional wave equations in disordered media*, J. Phys. A **16**, 25–42 (1983).
- [73] F. Delyon, Y. Lévy, and B. Souillard, *Anderson localization for multidimensional systems at large disorder or low energy*, Comm. Math. Phys. **100**, 463–470 (1985).
- [74] F. Delyon and B. Souillard, *Remark on the continuity of the density of states of ergodic finite difference operators*, Comm. Math. Phys. **94**, 289–291 (1984).
- [75] S. Denisov and A. Kiselev, *Spectral properties of Schrödinger operators with decaying potentials*, in this Festschrift.
- [76] E. I. Dinaburg and Ya. G. Sinai, *The one-dimensional Schrödinger equation with a quasi-periodic potential*, Funct. Anal. Appl. **9**, 279–289 (1975).
- [77] M. Donsker and S. R. S. Varadhan, *Asymptotics for the Wiener sausage*, Comm. Pure Appl. Math. **28**, 525–565 (1975).
- [78] H. von Dreifus and A. Klein, *A new proof of localization in the Anderson tight binding model*, Comm. Math. Phys. **124**, 285–299 (1989).
- [79] D. H. Dunlap, H.-L. Wu, and P. W. Phillips, *Absence of localization in random-dimer model*, Phys. Rev. Lett. **65**, 88–91 (1990).
- [80] K. Efetov, *Supersymmetry in Disorder and Chaos*, Cambridge University Press, Cambridge, 1997.
- [81] L. H. Eliasson, *Floquet solutions for the one-dimensional quasi-periodic Schrödinger equation*, Comm. Math. Phys. **146**, 447–482 (1992).
- [82] L. H. Eliasson, *Discrete one-dimensional quasi-periodic Schrödinger operators with pure point spectrum*, Acta Math. **179** 153–196 (1997).
- [83] L. Erdős, M. Salmhofer, and H.-T. Yau, *Quantum diffusion of the random Schrödinger evolution in the scaling limit*, preprint, 2005.
- [84] L. Erdős, M. Salmhofer, and H.-T. Yau, *Quantum diffusion of the random Schrödinger evolution in the scaling limit, II. The recollision diagrams*, to appear in Comm. Math. Phys.
- [85] A. Fedotov and F. Klopp, *Anderson transitions for a family of almost periodic Schrödinger equations in the adiabatic case*, Comm. Math. Phys. **227**, 1–92 (2002).
- [86] A. Fedotov and F. Klopp, *On the singular spectrum for adiabatic quasi-periodic Schrödinger operators on the real line*, Ann. Henri Poincaré **5**, 929–978 (2004).
- [87] S. Fishman and M. Griniasty, *Localization by pseudorandom potentials in one dimension*, Phys. Rev. Lett. **60** 1334–1337 (1987).
- [88] R. Froese, D. Hasler, and W. Spitzer, *Absolutely continuous spectrum for the Anderson model on a tree: a geometric proof of Klein's theorem*, to appear in Comm. Math. Phys.
- [89] J. Fröhlich, F. Martinelli, E. Scoppola, and T. Spencer, *Constructive proof of localization in the Anderson tight-binding model*, Comm. Math. Phys. **101**, 21–46 (1985).
- [90] J. Fröhlich and T. Spencer, *The Kosterlitz–Thouless transition in two-dimensional Abelian spin systems and the Coulomb gas*, Comm. Math. Phys. **81**, 527–602 (1981).
- [91] J. Fröhlich and T. Spencer, *Absence of diffusion in the Anderson tight-binding model for large disorder or low energy*, Comm. Math. Phys. **88**, 151–184 (1983).
- [92] J. Fröhlich, T. Spencer, and P. Wittwer, *Localization for a class of one dimensional quasi-periodic Schrödinger operators*, Comm. Math. Phys. **132**, 5–25 (1990).
- [93] F. Germinet, P. Hislop, and A. Klein, *Localization for the Schrödinger operator with a Poisson random potential*, C. R. Acad. Sci. Paris Ser. I **341**, 525–528 (2005).

- [94] F. Germinet and S. Jitomirskaya, *Strong dynamical localization for the almost Mathieu model*, Rev. Math. Phys. **13**, 755–765 (2001).
- [95] F. Germinet, A. Kiselev, and S. Tcheremchantsev, *Transfer matrices and transport for 1D Schrödinger operator with singular spectrum*, Ann. Inst. Fourier (Grenoble) **54**, 787–830 (2004).
- [96] F. Germinet and A. Klein, *Bootstrap multiscale analysis and localization in random media*, Comm. Math. Phys. **222**, 415–448 (2001).
- [97] F. Germinet and A. Klein, *Explicit finite volume criteria for localization in random media and applications*, Geom. Funct. Anal. **13**, 1201–1238 (2003).
- [98] F. Germinet and A. Klein, *A characterization of the Anderson metal-insulator transport transition*, Duke Math. J. **124**, 309–350 (2004).
- [99] F. Germinet, A. Klein, and J. Schenker, *Dynamical delocalization in random Landau Hamiltonians*, to appear in Annals of Math.
- [100] F. Gesztesy and B. Simon, *The xi function*, Acta Math. **176**, 49–71 (1996).
- [101] D. J. Gilbert, *On subordinacy and spectral multiplicity for a class of singular differential operators*, Proc. Roy. Soc. Edinburgh A **128** 549–584 (1998).
- [102] I. Ya. Goldsheidt, S. Molchanov, and L. Pastur, *A pure-point spectrum of the stochastic one-dimensional Schrödinger equation*, Funct. Anal. Appl. **11**, 1–10 (1977).
- [103] M. Goldstein and W. Schlag, *Hölder continuity of the integrated density of states for quasi-periodic Schrödinger equations and averages of shifts of subharmonic functions*, Annals of Math. **154**, 155–203 (2001).
- [104] M. Goldstein and W. Schlag, *Fine properties of the integrated density of states and a quantitative separation property of the Dirichlet eigenvalues*, preprint, 2005.
- [105] M. Goldstein and W. Schlag, *On resonances and the formation of gaps in the spectrum of quasi-periodic Schrödinger equations*, preprint, 2005.
- [106] A. Gordon, *The point spectrum of the one-dimensional Schrödinger operator*, (Russian) Uspehi Mat. Nauk **31**, 257–258 (1976).
- [107] A. Gordon, S. Jitomirskaya, Y. Last, and B. Simon, *Duality and singular continuous spectrum in the almost Mathieu equation* Acta Math., **178**, 169–183 (1997).
- [108] B. Halperin, *Quantized Hall conductance, current carrying edge states and the existence of extended states in a two dimensional disordered potential*, Phys. Rev. B **25** 2185–2190 (1982).
- [109] P. G. Harper, *Single band motion of conduction electrons in a uniform magnetic field*, Proc. Phys. Soc. London **A 68**, 874–892 (1955).
- [110] B. Helffer and J. Sjöstrand, *Semiclassical analysis for Harper’s equation. III. Cantor structure of the spectrum*, Mem. Soc. Math. France (N.S.) **39**, 1–124 (1989).
- [111] M.-R. Herman, *Une méthode pour minorer les exposants de Lyapounov et quelques exemples montrant le caractère local d’un théorème d’Arnol’d et de Moser s ur le tore de dimension 2*, Comment. Math. Helv. **58**, 453–502 (1983).
- [112] D. R. Hofstadter, *Energy levels and wave functions of Bloch electrons in rational and irrational magnetic fields*, Phys. Rev. B **14**, 2239–2249 (1976).
- [113] H. Holden and F. Martinelli, *On absence of diffusion near the bottom of the spectrum for a random Schrödinger operator*, Comm. Math. Phys. **93**, 197–217 (1984).
- [114] V. Jaksic and Y. Last, *Spectral structure of Anderson type Hamiltonians*, Invent. Math. **141**, 561–577 (2000).
- [115] V. Jaksic and Y. Last, *Surface states and spectra*, Comm. Math. Phys. **218**, 459–477 (2001).
- [116] V. Jaksic and Y. Last, *Simplicity of singular spectrum in Anderson type Hamiltonians*, to appear in Duke Math. J.
- [117] V. A. Javřan, *A certain inverse problem for Sturm-Liouville operators*, (Russian) Izv. Akad. Nauk Armjan. SSR Ser. Mat. **6**, 246–251 (1971).
- [118] S. Jitomirskaya, *Metal-insulator transition for the almost Mathieu operator*, Annals of Math. **150**, 1159–1175 (1999).
- [119] S. Jitomirskaya, D. A. Koslover, and M. S. Schulteis, *Localization for a family of one-dimensional quasiperiodic operators of magnetic origin*, Ann. Henri Poincaré **6**, 103–124 (2005).
- [120] S. Jitomirskaya and I. Krasovsky, *Continuity of the measure of the spectrum for discrete quasiperiodic operators*, Math. Res. Letters **9**, 413–421 (2002).

- [121] S. Jitomirskaya and Y. Last, *Anderson localization for the almost Mathieu equation, III. Uniform localization, continuity of gaps, and measure of the spectrum*, Comm. Math. Phys. **195**, 1–14 (1998).
- [122] S. Jitomirskaya and Y. Last, *Power-law subordinacy and singular spectra, II. Line operators*, Comm. Math. Phys. **211**, 643–658 (2000).
- [123] S. Jitomirskaya, H. Schulz-Baldes, and G. Stolz, *Delocalization in random polymer models*, Comm. Math. Phys. **233**, 27–48 (2003).
- [124] S. Jitomirskaya and B. Simon, *Operators with singular continuous spectrum, III. Almost periodic Schrödinger operators*, Comm. Math. Phys. **165**, 201–206 (1994).
- [125] R. Johnson and J. Moser, *The rotation number for almost periodic potentials*, Comm. Math. Phys. **84**, 403–438 (1982).
- [126] I. S. Kac, *Spectral multiplicity of a second-order differential operator and expansion in eigenfunction*, (Russian) Izv. Akad. Nauk SSSR Ser. Mat. **27**, 1081–1112 (1963).
- [127] W. Kirsch and F. Martinelli, *On the ergodic properties of the spectrum of general random operators*, J. Reine Angew. Math. **334**, 141–156 (1982).
- [128] W. Kirsch, P. Stollmann, and G. Stolz, *Localization for random perturbations of periodic Schrödinger operators*, Random Oper. Stochastic Equations **6**, 241–268 (1998).
- [129] A. Klein, *The supersymmetric replica trick and smoothness of the density of states for random Schrödinger operators*, Operator Theory: Operator Algebras and Applications, Part 1 (Durham, NH, 1988), pp. 315–331, Proc. Sympos. Pure Math., 51, Part 1, Amer. Math. Soc., Providence, RI, 1990.
- [130] A. Klein, *Extended states in the Anderson model on the Bethe lattice*, Adv. Math. **133**, 163–184 (1998).
- [131] A. Klein, J. Lacroix, and A. Speis, *Localization for the Anderson model on a strip with singular potentials*, J. Funct. Anal. **94**, 135–155 (1990).
- [132] S. Klein, *Anderson localization for the discrete one-dimensional quasi-periodic Schrödinger operator with potential defined by a Gevrey-class function*, J. Funct. Anal. **218**, 255–292 (2005).
- [133] F. Klopp, *Localization for semiclassical continuous random Schrödinger operators. II. The random displacement model*, Helv. Phys. Acta **66**, 810–841 (1993).
- [134] F. Klopp, *Localization for continuous random Schrödinger operators*, Comm. Math. Phys. **167**, 553–569 (1995).
- [135] F. Klopp, *A low concentration asymptotic expansion for the density of states of a random Schrödinger operator with Poisson disorder*, J. Funct. Anal. **145** 267–295 (1997).
- [136] F. Klopp, *Weak disorder localization and Lifshitz tails*, Comm. Math. Phys. **232**, 125–155 (2002).
- [137] F. Klopp and L. Pastur, *Lifshitz tails for random Schrödinger operators with negative singular Poisson potential*, Comm. Math. Phys. **206**, 57–103 (1999).
- [138] F. Klopp and T. Wolff, *Lifshitz tails for 2-dimensional random Schrödinger operators*, J. Anal. Math. **88**, 63–147 (2002).
- [139] J. M. Kosterlitz and D. J. Thouless, *Ordering, metastability and phase transitions in two-dimensional systems*, J. Phys. C **6**, 1181–1203 (1973).
- [140] S. Kotani, *Ljapunov indices determine absolutely continuous spectra of stationary random one-dimensional Schrödinger operators*, Stochastic Analysis (Katata/Kyoto, 1982), pp. 225–247, North-Holland Math. Library, 32, North-Holland, Amsterdam, 1984.
- [141] S. Kotani, *Support theorems for random Schrödinger operators*, Comm. Math. Phys. **97**, 443–452 (1985).
- [142] S. Kotani, *Lyapunov exponents and spectra for one-dimensional random Schrödinger operators*, Random Matrices and Their Applications (Brunswick, Maine, 1984), pp. 277–286, Contemp. Math., 50, Amer. Math. Soc., Providence, RI, 1986.
- [143] S. Kotani, *Jacobi matrices with random potentials taking finitely many values*, Rev. Math. Phys. **1**, 129–133 (1989).
- [144] S. Kotani, *Generalized Floquet theory for stationary Schrödinger operators in one dimension*, Chaos Solitons Fractals **8**, 1817–1854 (1997).
- [145] S. Kotani and B. Simon, *Localization in general one dimensional systems. II*, Comm. Math. Phys. **112**, 103–120 (1987).

- [146] H. Kunz and B. Souillard, *Sur le spectre des operateurs aux differences finies alatoires*, [On the spectra of random finite difference operators], *Comm. Math. Phys.* **78**, 201–246 (1980/81).
- [147] Y. Last, *A relation between absolutely continuous spectrum of ergodic Jacobi matrices and the spectra of periodic approximants*, *Comm. Math. Phys.* **151**, 183–192 (1993).
- [148] Y. Last, *Zero measure spectrum for the almost Mathieu operator*, *Comm. Math. Phys.* **164**, 421–432 (1994).
- [149] Y. Last, unpublished (1994).
- [150] Y. Last, *Quantum dynamics and decomposition of singular continuous spectra*, *J. Funct. Anal.* **142**, 402–445 (1996).
- [151] Y. Last, *Spectral theory of Sturm–Liouville operators on infinite intervals: a review of recent developments*, *Sturm–Liouville Theory*, pp. 99–120, Birkhauser, Basel, 2005.
- [152] Y. Last and B. Simon, *Eigenfuctions, transfermatrices, and absolutely continuous spectrum of one-dimensional Schrödinger operators*, *Invent. Math.* **135**, 329–367 (1999).
- [153] R. B. Laughlin, *Quantised Hall conductivity in two dimensions*, *Phys. Rev. B.* **23**, 5632–5633 (1981)
- [154] I. M. Lifshits, A. G. Gredeskul, and L. A. Pastur, *Introduction to the Theory of Disordered Systems*, Wiley–Interscience, New York, 1988.
- [155] F. Martinelli and E. Scoppola, *Introduction to the mathematical theory of Anderson localization*, *Riv. Nuovo Cimento* (3) **10**, 1–90 (1987)
- [156] J. Miller and B. Derrida, *Weak disorder expansion for the Anderson model on a tree*, *J. Stat. Phys.* **75**, 357–389 (1994).
- [157] N. Minami, *Local fluctuation of the spectrum of a multidimensional Anderson tight binding model*, *Comm. Math. Phys.* **177**, 709–725 (1996).
- [158] S. A. Molchanov, *The local structure of the spectrum of the one-dimensional Schrödinger operator*, *Comm. Math. Phys.* **78**, 69–101 (1978).
- [159] S. Molchanov and B. Vainberg, *Multiscattering by sparse scatterers*, *Mathematical and Numerical Aspects of Wave Propagation (Santiago de Compostela, 2000)*, pp. 518–522, SIAM, Philadelphia, PA, 2000.
- [160] L. Pastur, *Spectral properties of disordered systems in the one-body approximation*, *Comm. Math. Phys.* **75**, 179–196 (1980).
- [161] L. Pastur and A. Figotin, *Spectra of Random and Almost-Periodic Operators*, Springer-Verlag, Heidelberg, 1992.
- [162] R. Peierls, *Zur Theorie des Diamagnetismus von Leitungselektronen*, *Z. Phys.* **80**, 763–791 (1933).
- [163] R. Prange, D. Grepfel, and S. Fishman, *A solvable model of quantum incommensurate potential*, *Phys. Rev. B* **29**, 6500–6512 (1984).
- [164] J. Puig, *Cantor spectrum for the almost Mathieu operator*, *Comm. Math. Phys.* **244**, 297–309 (2004).
- [165] A. Rauh, *Degeneracy of Landau levels in crystals*, *Phys. Status Solidi B* **65**, K131–135 (1974).
- [166] C. Shubin, R. Vakilian, and T. Wolff, *Some harmonic analysis questions suggested by Anderson-Bernoulli models*, *Geom. Funct. Anal.* **8**, 932–964 (1998).
- [167] B. Simon, *Kotani theory for one-dimensional stochastic Jacobi matrices*, *Comm. Math. Phys.* **89**, 227–234 (1983).
- [168] B. Simon, *Almost periodic Schrödinger operators, IV. The Maryland model*, *Ann. Phys.* **159**, 157–183 (1985).
- [169] B. Simon, *Absence of ballistic motion*, *Comm. Math. Phys.* **134**, 209–212 (1990).
- [170] B. Simon, *Cyclic vectors in the Anderson model*, *Rev. Math. Phys.* **6**, 1183–1185 (1994).
- [171] B. Simon, *Schrödinger operators in the twenty-first century*, XIIIth International Congress on Mathematical Physics (Imperial College, London, 2000). A. Fokas, A. Grigoryan, T. Kibble and B. Zegarlinski, eds., pp. 283–288, International Press, Boston, 2001.
- [172] B. Simon, *On a theorem of Kac and Gilbert*, *J. Funct. Anal.* **223**, 109–115 (2005).
- [173] B. Simon, *Orthogonal Polynomials on the Unit Circle*, American Mathematical Society Colloquium Publications, 54, American Mathematical Society, Providence, RI, 2005.
- [174] B. Simon, *OPUC on one foot*, *Bull. Amer. Math. Soc. (N.S.)* **42**, 431–460 (2005).
- [175] B. Simon, *A New Book*, (6 volumes); to be written by 2066.

- [176] B. Simon and M. Taylor, *Harmonic analysis on $SL(2, \mathbb{R})$ and smoothness of the density of states in one-dimensional Anderson model*, Comm. Math. Phys. **101**, 1–19 (1985).
- [177] B. Simon and T. Wolff, *Singular continuous spectrum under rank one perturbation and localization for random Hamiltonians*, Comm. Pure Appl. Math. **39**, 75–90 (1986).
- [178] Ya. Sinai, *Anderson localization for one-dimensional difference Schrödinger operator with quasi-periodic potential*, J. Stat. Phys. **46**, 861–909 (1987).
- [179] E. Sorets and T. Spencer, *Positive Lyapunov exponents for Schrödinger operators with quasi-periodic potential*, Comm. Math. Phys. **142**, 543–566 (1991).
- [180] T. Spencer, *Localization for random and quasiperiodic potentials*, J. Stat. Phys. **51**, 1009–1019 (1988).
- [181] T. Spencer, *Lifshitz tails and localization*, preprint.
- [182] P. Stollmann, *Lifshitz asymptotics via linear coupling of disorder*, Math. Phys. Anal. Geom. **2**, 279–289 (1999).
- [183] P. Stollmann, *Wegner estimates and localization for continuum Anderson models with some singular distributions*, Arch. Math. (Basel) **75**, 307–311 (2000).
- [184] P. Stollmann, *Caught by disorder*, Bound States in Random Media, Progress in Mathematical Physics, 20, Birkhäuser, Boston, 2001.
- [185] G. Stolz, *Localization for random Schrödinger operators with Poisson potential*, Ann. Inst. H. Poincaré Phys. Théor. **63**, 297–314 (1995).
- [186] D. J. Thouless, *Localization and the two dimensional Hall effect*, J. Phys. C **14** 3475–3480 (1981).
- [187] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, *Quantised Hall conductance in a two dimensional periodic potential*, Phys. Rev. Lett. **49**, 405–408 (1982).
- [188] W.-M. Wang, *Microlocalization, percolation and Anderson localization for magnetic Schrödinger operator with a random potential*, J. Func. Anal. **146** 1–26 (1997).
- [189] W.-M. Wang, *Localization and universality of Poisson statistics for the multidimensional Anderson model at weak disorder*, Invent. Math. **146**, 365–398 (2001).
- [190] F. Wegner, *Bounds on the density of states in disordered systems*, Z. Phys. **B44**, 9–15 (1981).
- [191] L.-S. Young, *Lyapunov exponents for some quasi-periodic cocycles*, Ergodic Theory Dynam. Systems **17**, 483–504 (1997).

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA AT IRVINE, IRVINE, CA 92697, U.S.A.

E-mail address: szhitomi@math.uci.edu

The Integrated Density of States for Random Schrödinger Operators

Werner Kirsch and Bernd Metzger*

Dedicated to Barry Simon on the occasion of his 60th birthday

ABSTRACT. We survey some aspects of the theory of the integrated density of states (IDS) of random Schrödinger operators. The first part motivates the problem and introduces the relevant models as well as quantities of interest. The proof of the existence of this interesting quantity, the IDS, is discussed in the second section. One central topic of this survey is the asymptotic behavior of the integrated density of states at the boundary of the spectrum. In particular, we are interested in Lifshitz tails and the occurrence of a classical and a quantum regime. In the last section we discuss regularity properties of the IDS. Our emphasis is on the discussion of fundamental problems and central ideas to handle them. Finally, we discuss further developments and problems of current research.

CONTENTS

1. Introduction
 2. The Density of States Measure: Existence
 3. Lifshitz Tails
 4. Regularity of the Integrated Density of States
- References

1. Introduction

1.1. Physical Background: Models and Notation. The time evolution of a quantum mechanical state ψ is obtained from the time dependent Schrödinger equation

$$i \frac{\partial}{\partial t} \psi = H \psi. \tag{1.1}$$

2000 *Mathematics Subject Classification.* 82B44, 35J10, 35P20, 81Q10, 58J35.

Key words and phrases. disordered systems, random Schrödinger operators, integrated density of states, Lifshitz tails.

*Supported by the Deutsche Forschungsgemeinschaft (DFG), SFB-TR 12.

Another important differential equation is the heat or diffusion equation

$$\frac{\partial}{\partial t} \psi = - H \psi. \tag{1.2}$$

In order to investigate either of these equations, it is extremely useful to know as much as possible about the operator H and its spectrum. In general the Schrödinger operator H is of the form

$$H = H_0 + V.$$

The free operator H_0 represents the kinetic energy of the particle. In the absence of magnetic fields, it is given by the Laplacian

$$H_0 = - \Delta = - \sum_{\nu=1}^d \frac{\partial^2}{\partial x_\nu^2}.$$

The potential V encoding the forces $F(x) = -\nabla V(x)$ is acting as a multiplication operator in the Hilbert space $L^2(\mathbb{R}^d)$.

Occasionally, we will replace H_0 with the operator $H_0(B)$ which contains a homogeneous magnetic field. For $d = 2$ $H_0(B)$, $B > 0$, is given by

$$H_0(B) = \left(i \frac{\partial}{\partial x_1} - \frac{1}{2} B x_2 \right)^2 + \left(i \frac{\partial}{\partial x_2} + \frac{1}{2} B x_1 \right)^2.$$

In contrast to $H_0(B = 0)$, the spectrum of $H_0(B)$ for $B > 0$ is a countable set $\sigma(H_0(B)) = \{(2n + 1)B; n \in \mathbb{N}\}$. The energies $E_n = (2n + 1)$ are eigenvalues of $H_0(B)$ of infinite multiplicity, called the Landau levels.

1.2. Random Potentials. First-order physical modeling often assumes an ideal background, e.g., a homogeneous material without any impurities. For example, in ideal crystals the atoms or nuclei are supposed to be distributed on a periodic lattice (say the lattice \mathbb{Z}^d) in a completely regular way. We assume that a particle (electron) at the point $x \in \mathbb{R}^d$ feels a potential of the form $q f(x - i)$ due to an atom (or ion or nucleus) located at the point $i \in \mathbb{Z}^d$. We call the function f the *single site potential*. The coupling constant q represents the charge of the particle at the lattice point i . So, in a regular crystal, the particle is exposed to a total potential

$$V(x) = \sum_{i \in \mathbb{Z}^d} q f(x - i). \tag{1.3}$$

The potential V in (1.3) is periodic with respect to the lattice \mathbb{Z}^d , i.e., $V(x - i) = V(x)$ for all $x \in \mathbb{R}^d$ and $i \in \mathbb{Z}^d$. The mathematical theory of Schrödinger operators with periodic potentials is well developed (see, e.g., [35, 112, 150]). It is based on a thorough analysis of the symmetry properties of periodic operators. The spectrum of periodic Schrödinger operators has a band structure and the spectrum is purely absolutely continuous, i.e.,

$$\sigma(H) = \bigcup_{n=0}^{\infty} [a_n, b_n] \quad \text{with } a_n < b_n \leq a_{n+1}, \tag{1.4}$$

$$\sigma_{\text{sing}}(H) = \emptyset, \quad \sigma(H) = \sigma_{\text{ac}}(H). \tag{1.5}$$

The real world is not ideal. Solids occur in nature in various forms. Sometimes they are (almost) totally ordered, sometimes they are more or less completely disordered. For a mathematical modelling of disordered solids two ingredients are essential: the spatial homogeneity in the mean and the disappearance of long range correlations. In full generality these properties are studied in the theory of ergodic operators. This class of operators contains, e.g., quasicrystals and can also be related to random matrices. In this publication we will be concerned with *random* Schrödinger operators.

The most popular and best understood model of a disordered solid is the alloy-type Schrödinger operator. It models a mixture of different atoms located at lattice positions. The type of atom at the lattice point i is assumed to be random. These particles are represented by randomly distributed coupling constants q_i encoding the different charges. The total potential is given by

$$V_\omega(x) = \sum_{i \in \mathbb{Z}^d} q_i(\omega) f(x - i). \quad (1.6)$$

When talking about the alloy-type model, we will mean (1.6) with the following assumptions:

- (1) The single site potential f is bounded, non-negative and strictly positive on an open set.
- (2) f satisfies $f(x) \leq C(1 + |x|)^{-(d+\varepsilon)}$ for some C and $\varepsilon > 0$.
- (3) The random variables q_i are independent and identically distributed random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.
- (4) The common probability distribution of the q_i is denoted by P_0 . Its support $\text{supp } P_0$ is compact and contains at least two points.

Assumption (1) can be considerably relaxed. For example, for most of the following results one may allow local singularities for f . By assuming (1) we avoid technical difficulties which may obscure the main argument. More details on weaker conditions can be found in the papers cited. Assumption (2) ensures that the sum in (1.6) is convergent. The compactness of $\text{supp } P_0$ is convenient but not always necessary and in some especially marked situations we consider also unbounded single site distributions. However, for many of our results we need that the q_i (and hence $\text{supp } P_0$) are bounded *below*. The physical model suggests that $\text{supp } P_0$ consists of finitely many points only (the possible charges). However, for many mathematical results it is necessary (or at least convenient) to suppose that P_0 has a (bounded) density g , i.e., $P_0 = g(\lambda) d\lambda$.

One might argue that such an assumption is acceptable as a purely technical one. On the other hand one could argue the problem is not understood as long as it is impossible to handle the physically relevant case of finitely many values.

A simplified version of the alloy-type Schrödinger operator above is the (discrete) Anderson model. Here the Hilbert space is the sequence space $\ell^2(\mathbb{Z}^d)$ instead of $L^2(\mathbb{R}^d)$ and the free operator H_0 is replaced by the discrete analogue of the Laplacian. This is the finite-difference operator

$$(h_0 u)(n) = - \sum_{|m-n|=1} (u(m) - u(n)). \quad (1.7)$$

Above we set $|n| = \sum_{i=1}^d |n_i|$ on \mathbb{Z}^d . The potential is a multiplication operator $V = V_\omega$ on $\ell^2(\mathbb{Z}^d)$ with $V_\omega(i)$ independent, identically distributed, and the total

Hamiltonian is defined by $h = h_0 + V_\omega$. We will call this setting the “discrete case” in contrast to Schrödinger operators on $L^2(\mathbb{R}^d)$ which we refer to as the “continuous case.”

The most frequently used and most important approach to model *amorphous material* like glass or rubber defines the random locations of the atoms by a Poisson random measure. This random point measure can be characterized by the number $n_A = \mu_\omega(A)$ of random points in the set A . We assume that the random variables n_A and n_B are independent for disjoint (measurable) sets A, B and $\mathbb{P}(n_A = k) = \frac{|A|^k}{k!} e^{-|A|}$ ($|A|$ denotes the Lebesgue measure of A). With this notation we may write the Poisson potential for an amorphous solid as

$$V_\omega(x) = \int_{\mathbb{R}^d} q f(x - \eta) d\mu_\omega(\eta). \quad (1.8)$$

To model thin disordered layers, we also consider random potentials which are concentrated along a hypersurface in \mathbb{R}^d (or \mathbb{Z}^d). For example, we are going to consider “surface” alloy potentials. To define such a potential let us write $\mathbb{R}^d = \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ then

$$V_\omega(x_1, x_2) = \sum_{i_1 \in \mathbb{Z}^{d_1}} q_{i_1}(\omega) f(x_1 - i_1, x_2)$$

is a random potential which is concentrated along the hypersurface \mathbb{R}^{d_1} in \mathbb{R}^d . In addition, there may be a random or periodic background potential on \mathbb{R}^d .

Most of the theorems we are going to discuss can be proved for rather general single site potentials f and probability distributions P_0 of the q_i . For example, most of the time we can allow some local singularities of f . To simplify the following discussion, we will assume in this paper the conditions defined in the context of the alloy-type potential.

The above random operators are examples of “ergodic operators.” This class of operators includes not only most random operators but also periodic and almost periodic operators. Most of the results of Section 2 and part of Section 4 can be shown for general ergodic operators. We refer to [148, 15, 27, 119, 166] and [68] for a discussion of this general context.

1.3. The Concept of the Integrated Density of States. The (integrated) density of states is a concept of fundamental importance in condensed matter physics. It measures the “number of energy levels per unit volume” near (resp. below) a given energy.

Typical systems arising in solid state physics have periodic or ergodic potentials. Consequently, the spectrum of the corresponding Hamiltonian is not discrete. Therefore, we cannot just count the eigenvalues below E or within an interval $[E_1, E_2]$. On the other hand, the number of electrons in such a system, which extends to infinity, ought to be infinite. For these two reasons, the Pauli exclusion principle does not make immediate sense. (How do we distribute infinitely many electrons on a continuum of spectral energies?)

However, there may be a chance to make sense out of the Pauli principle by first restricting the system to a finite volume Λ . Inside Λ there should be only finitely many electrons, in fact, we may assume that the number of electrons in a given Λ is proportional to the volume $|\Lambda|$ of this set.

If P is a finite-dimensional orthogonal projection, then $\text{tr}(P)$ is the dimension of its range. If $P_{(-\infty, E]}$ is the spectral projection of a random Schrödinger operator (which as a rule has infinite-dimensional range) and if Λ_L is a cube of side length L around the origin, then we may call $\text{tr}(\chi_{\Lambda_L} P_{(-\infty, E]})$ the restriction of $P_{(-\infty, E]}$ to the cube Λ_L . Above χ_A is the characteristic function of the set A . Thus, we may try to define the integrated density of states as

$$N(E) = \lim_{L \rightarrow \infty} \frac{1}{|\Lambda_L|} \text{tr}(\chi_{\Lambda_L} P_{(-\infty, E]}). \quad (1.9)$$

Of course, we have to prove that the limit in (1.9) does exist and is not trivial. We will deal with these questions in Section 2.

There is another way to define the integrated density of states which turns out to be equivalent to (1.9): We restrict the operator H_ω to the Hilbert space $L^2(\Lambda)$. To obtain a self adjoint operator we have to impose boundary conditions at $\partial\Lambda$, e.g., Dirichlet or Neumann boundary conditions. We call the corresponding operators H_Λ^D and H_Λ^N , respectively. These operators have compact resolvents, i.e., their spectra are purely discrete. We denote by

$$E_1(H_\Lambda^D) \leq E_2(H_\Lambda^D) \leq E_3(H_\Lambda^D) \dots \quad (1.10)$$

the eigenvalues of H_Λ^D (and analogously for H_Λ^N) in increasing order, where eigenvalues are repeated according to their multiplicity. The eigenvalue counting function of an operator A with purely discrete spectrum is defined by

$$N(A, E) = \#\{n \mid E_n(A) \leq E\} = \text{tr}(P_{(-\infty, E]}(A)). \quad (1.11)$$

Analogously to (1.9), we can therefore define

$$\begin{aligned} N^D(E) &= \lim_{L \rightarrow \infty} \frac{1}{|\Lambda_L|} N(H_{\Lambda_L}^D, E) \\ &= \lim_{L \rightarrow \infty} \frac{1}{|\Lambda_L|} \text{tr}(P_{(-\infty, E]}(H_{\Lambda_L}^D)) \end{aligned} \quad (1.12)$$

and similarly for Neumann boundary conditions,

$$\begin{aligned} N^N(E) &= \lim_{L \rightarrow \infty} \frac{1}{|\Lambda_L|} N(H_{\Lambda_L}^N, E) \\ &= \lim_{L \rightarrow \infty} \frac{1}{|\Lambda_L|} \text{tr}(P_{(-\infty, E]}(H_{\Lambda_L}^N)). \end{aligned} \quad (1.13)$$

This procedure to define the integrated density of states makes sense only if N , N^D and N^N all exist and agree.

This is, indeed, the case. We will see in the sequel that each of these definitions has its own technical advantage. The integrated density of states N is basic for studying the physical (in particular the thermodynamical) properties of disordered systems. From a mathematical point of view, the properties of N are interesting in their own respect. Moreover, properties of N constitute an essential input to prove localization properties of the system.

It is the aim of this review to discuss some of the problems and results connected with the integrated density of states. In Section 2 we sketch the proof of the existence of the integrated density of states and discuss some fundamental questions concerning the probabilistic and the functional analytic approach. In Section 3 we study the behavior of the integrated density of states at the boundary of the

spectrum. In the last section we discuss some basic ideas concerning the regularity of the integrated density of states.

Acknowledgements. It is a pleasure to thank many colleagues for fruitful collaborations and stimulating discussions on the subject. There are too many to name them here. We also would like to stress the fact that the selection of topics within our subject and the way of presenting them is due to our very personal preferences. We have most certainly left out important topics and works. This is to be blamed on our ignorance and the limitation of time and space.

We would like to thank Jessica Langner and Riccardo Catalano for their skillful typing and careful proofreading of the manuscript.

2. The Density of States Measure: Existence

2.1. Introduction and Historical Remarks. The first existence proofs for the integrated density of states go back at least to Pastur. See [145] for an early review of the subject.

There are a couple of methods to prove the existence of the integrated density of states. One of them, invented and used by Pastur, is based on the Laplace transform of the integrated density of states and of its approximants. For this method, one proves the convergence of the Laplace transform and uses the fact that convergence of the Laplace transform implies the vague convergence of the underlying measures.

To prove the convergence of the Laplace transforms, it is useful to express the Laplace transform of the finite-volume quantities using the Feynman–Kac-representation of the Schrödinger semigroup e^{-tH} . Feynman–Kac and Laplace transform methods are also used to prove the equivalence of the definitions of the integrated density of states (1.9) and (1.12), (1.13) with either Neumann or Dirichlet (or more general) boundary conditions (see, e.g., Pastur [145] or [74, 34, 68]). The definition of the integrated density of states via (1.9) was used by Avron and Simon in the context of almost periodic potentials [3]. They also proved that the spectrum of the operator coincides with the growth points of the integrated density of states. In Section 2.2 we will follow this approach to prove the existence of the integrated density of states.

One of the virtues of the definition of the integrated density of states via boundary conditions (1.12) and (1.13) is the fact that they allow lower and upper bounds of N “for free.” In fact, one way to examine the behavior of N at the bottom (or top = ∞) of the spectrum is based on this approach. We will discuss this approach in Section 2.3 and the estimates based on it in Section 3.

As a rule, quantitative estimates on the effect of introducing boundary conditions are hard to obtain. For example, if one investigates the behavior of N at internal spectral edges it seems extremely difficult to control the perturbation of eigenvalues due to boundary conditions. Klopp [97] proposed an approximation of the random potential by periodic ones with growing period. This way we lose monotonicity which is at the heart of the Neumann–Dirichlet approach. Instead one can prove that the approximation is exponentially fast thus gaining good estimates of the remainder.

Finally, we would like to mention that one can also define the integrated density of states via Krein’s spectral shift function. This reasoning is well known in

scattering theory (see, e.g., [9, 180]). In connection with random Schrödinger operators, the spectral shift function was first used by Simon [159] to investigate spectral averaging. Kostrykin and Schrader [106] applied this technique to prove the existence of the integrated density of states and the density of surface states. This method turns out to be useful also to investigate regularity properties of the integrated density of states [107].

The results of this section are true not only for the specific random potentials discussed in Section 1, but rather for general ergodic operators. In fact, the proofs carry over to this general setting in most cases. We refer to the survey [68] for details.

2.2. The Existence of the Integrated Density of States. In this section we prove the existence of the integrated density of states as defined in (1.9). To do so, we need little more than Birkhoff's ergodic theorem (see, e.g., [113]). Below, as in the rest of this paper, we denote by \mathbb{E} the expectation with respect to the probability measure \mathbb{P} .

PROPOSITION 2.1. *If φ is a bounded measurable function of compact support, then*

$$\lim_{L \rightarrow \infty} \frac{1}{|\Lambda_L|} \operatorname{tr}(\varphi(H_\omega)\chi_{\Lambda_L}) = \mathbb{E}(\operatorname{tr}(\chi_{\Lambda_1}\varphi(H_\omega)\chi_{\Lambda_1})) \quad (2.1)$$

for \mathbb{P} -almost all ω .

PROOF. Define $\xi_i = \operatorname{tr}(\varphi(H_\omega)\chi_{\Lambda_1(i)})$. ξ_i is an ergodic sequence of random variables. Hence, Birkhoff's ergodic theorem implies that

$$\frac{1}{|\Lambda_L|} \operatorname{tr}(\varphi(H_\omega)\chi_{\Lambda_L}) = \frac{1}{|\Lambda_L|} \sum_{i \in \Lambda_L} \xi_i$$

converges to its expectation value. \square

The right-hand side of (2.1) as well as $|\Lambda_L|^{-1} \operatorname{tr}(\varphi(H_\omega)\chi_{\Lambda_L})$ are positive linear functionals on the bounded, continuous functions. They define positive measures ν and ν_L by

$$\int \varphi(\lambda) d\nu(\lambda) = \mathbb{E}(\operatorname{tr}(\chi_{\Lambda_1}\varphi(H_\omega)\chi_{\Lambda_1}))$$

and

$$\int_{\mathbb{R}} \varphi(\lambda) d\nu_L(\lambda) = \frac{1}{|\Lambda_L|} \operatorname{tr}(\varphi(H_\omega)\chi_{\Lambda_L}).$$

Equation (2.1) suggests that the measures ν_L might converge to the limit measure ν as $L \rightarrow \infty$ in the sense of vague convergence of measures. The problem is (2.1) holds only for fixed φ on a set Ω_φ of full probability; respectively (2.1) holds for all φ for $\omega \in \bigcap_\varphi \Omega_\varphi$. However, this is an *uncountable* intersection of sets of probability one. The problem is solved by approximating $C_0(\mathbb{R})$ by a countable, dense subset.

THEOREM 2.2. *The measures ν_L converge vaguely to the measure ν \mathbb{P} -almost surely, i.e., there is a set Ω_0 of probability one, such that*

$$\int \varphi(\lambda) d\nu_L(\lambda) \rightarrow \int \varphi(\lambda) d\nu(\lambda) \quad (2.2)$$

for all $\varphi \in C_0(\mathbb{R})$, the set of continuous functions with compact support, and all $\omega \in \Omega_0$.

DEFINITION 2.3. *The non-random probability measure ν is called the density of states measure. The distribution function N of ν , defined by*

$$N(E) = \nu([\!-\infty, E]),$$

is known as the integrated density of states.

Using Theorem 2.2 it is not hard to see:

PROPOSITION 2.4 ([3]). $\text{supp}(\nu) = \Sigma [= \sigma(H_\omega) \text{ a.s.}]$.

2.3. Existence via Dirichlet–Neumann–Bracketing. Our first approach to define the density of states measure was based on the additivity of $\text{tr}(\varphi(H_\omega)\chi_{\Lambda_L})$ and the ergodic theorem by Birkhoff. This very naturally fits in the concept of self-averaged quantities from physics.

However, for some part of the further analysis, an alternative approach—the Dirichlet–Neumann bracketing—is more suitable. Let $(H_\omega)_\Lambda^N$ and $(H_\omega)_\Lambda^D$ be the restrictions of H_ω to $L^2(\Lambda)$ with Neumann and Dirichlet boundary conditions. See, e.g., [150] for an appropriate definition of these boundary conditions via quadratic forms. Furthermore, we define (for $X = N$ or D , and $E \in \mathbb{R}$)

$$N_\Lambda^X(E) := N((H_\omega)_\Lambda^X, E) = \text{tr}(\chi_{(-\infty, E]}(H_\omega)_\Lambda^X). \tag{2.3}$$

Our aim is to consider the limits

$$N^X(E) = \lim_{L \rightarrow \infty} \frac{1}{|\Lambda_L|} N_{\Lambda_L}^X(E). \tag{2.4}$$

The quantities N_Λ^D and N_Λ^N as defined in (2.3) are distribution functions of point measures ν_Λ^D and ν_Λ^N concentrated in the eigenvalues of H_Λ^D and H_Λ^N , i.e.,

$$N_\Lambda^X(E) = \nu_\Lambda^X((-\infty, E]). \tag{2.5}$$

The convergence in (2.4) is meant as the vague convergence of the corresponding measures or, what is the same, as the pointwise convergence of the distribution function $\frac{1}{|\Lambda|} N_\Lambda^X$ at all continuity points of the limit.

Let us first look at $\frac{1}{|\Lambda|} N_\Lambda^D(E)$. The random field N_Λ^D is *not* additive in Λ , so that we *cannot* use Birkhoff’s ergodic theorem. However, N_Λ^D is *superadditive*, in the sense that $N_\Lambda^D(E) \geq N_{\Lambda_1}^D(E) + N_{\Lambda_2}^D(E)$ whenever $\Lambda = \Lambda_1 \cup \Lambda_2$ with $(\Lambda_1)^\circ \cap (\Lambda_2)^\circ = \emptyset$. (M° denotes the interior of the set M .) Similarly, N_Λ^N is *subadditive*, i.e., $-N^N$ is superadditive.

THEOREM 2.5. N_Λ^D is superadditive and N_Λ^N is subadditive. More precisely, if $\Lambda = \Lambda_1 \cup \Lambda_2$ and $(\Lambda_1)^\circ \cap (\Lambda_2)^\circ = \emptyset$ then

$$N_{\Lambda_1}^D(E) + N_{\Lambda_2}^D(E) \leq N_\Lambda^D(E) \leq N_\Lambda^N(E) \leq N_{\Lambda_1}^N(E) + N_{\Lambda_2}^N(E).$$

Fortunately, there are sub- and superadditive versions of the ergodic theorem, going back at least to Kingman [66]. The situation here is ideal for the superadditive ergodic theorem by Akcoglu and Krengel [2]. Indeed, one can prove that (for fixed E) the processes $N_\Lambda^D(E)$ and $N_\Lambda^N(E)$ are superadditive and subadditive random fields in the sense of [2] respectively (see [74, 111]). This yields the following result.

THEOREM 2.6 ([74]). *The limits*

$$\bar{N}^D(E) = \lim_{L \rightarrow \infty} \frac{1}{|\Lambda_L|} N(H_{\omega, \Lambda}^D, E)$$

and

$$\bar{N}^N(E) = \lim_{L \rightarrow \infty} \frac{1}{|\Lambda_L|} N(H_{\omega, \Lambda}^N, E)$$

exist P -almost surely. Moreover,

$$\bar{N}^D(E) = \sup_L \frac{1}{|\Lambda_L|} \mathbb{E}(N(H_{\omega, \Lambda_L}^D, E))$$

$$\bar{N}^N(E) = \inf_L \frac{1}{|\Lambda_L|} \mathbb{E}(N(H_{\omega, \Lambda_L}^N, E)).$$

The functions \bar{N}^X are increasing functions. However, it is not clear whether they are right continuous. To obtain *distribution functions*, we define N^X by making the \bar{N}^X right continuous

$$N^D(E) = \inf_{E' > E} \bar{N}^D(E') \quad (2.6)$$

$$N^N(E) = \inf_{E' > E} \bar{N}^N(E'). \quad (2.7)$$

Note that \bar{N}^X and N^X disagree at most on a countable set. Since N^D and N^N are distribution functions, they define measures by

$$\nu^D((a, b]) = N^D(b) - N^D(a) \quad (2.8)$$

$$\nu^N((a, b]) = N^N(b) - N^N(a). \quad (2.9)$$

From Theorem 2.6 we obtain the following corollary, which we will use to investigate the asymptotic behavior of the integrated density of states (e.g., for small E).

COROLLARY 2.7. *For any Λ ,*

$$\frac{1}{|\Lambda|} \mathbb{E}(N(H_{\omega, \Lambda}^D, E)) \leq \bar{N}^D(E) \leq \bar{N}^N(E) \leq \frac{1}{|\Lambda|} \mathbb{E}(N(H_{\omega, \Lambda}^N, E)).$$

Our physical intuition would lead to the hope that $N(E) = N^D(E) = N^N(E)$ since, after all, the introduction of boundary conditions was a mathematical artifact that should not play any role for the final physically meaningful quantity. This is, in fact, true under fairly weak conditions (see [145, 74, 34] and references given there).

THEOREM 2.8. *The distribution functions $N(E)$, $N^D(E)$ and $N^N(E)$ agree.*

Theorem 2.8 follows from Theorem 2.10 in the next section. An alternative proof for the Anderson model can be found in the review [70]. Theorem 2.8 implies a fortiori that the quantities $\frac{1}{|\Lambda_L|} N_{\Lambda_L}^D(E)$ and $\frac{1}{|\Lambda_L|} N_{\Lambda_L}^N(E)$ converge to the same limit, except for a countable set of energies E . The exceptional points, if any, are the discontinuity points of N . We will discuss continuity (and, more generally, regularity) properties of N in Section 4.

2.4. A Feynman–Kac Representation for N . In this section we will consider the Laplace transform of the integrated density of states (both $N(E)$ and $N^X(E)$). The Laplace transform of a measure ν with distribution function F is defined by

$$\tilde{\nu}(t) = \tilde{F}(t) := \int e^{-\lambda t} d\nu(\lambda) = \int e^{-\lambda t} dF(\lambda). \tag{2.10}$$

There is a very useful representation of the Laplace transform of N (and of N^X) via the Feynman–Kac formula. Using this representation one can show that N and N^X are, indeed, the same quantities. Moreover, the Feynman–Kac-representation of N is very useful to compute the asymptotic behavior of N for small or large energies.

The key ingredients of the representation formula for $\tilde{N}(t)$ are the Brownian motion, the Brownian bridge and the Feynman–Kac-formula. For material about these concepts, we refer to Reed–Simon [149, 150] and Simon [154].

By $\mathbb{P}_{0,x}^{t,y}$ we denote the measure underlying a Brownian bridge starting in the point $x \in \mathbb{R}^d$ at time 0 and ending at time t in the point y . $\mathbb{E}_{0,x}^{t,y}$ denotes integration over $\mathbb{P}_{0,x}^{t,y}$. A Brownian bridge is a Brownian motion b conditioned on $b(t) = y$. Note that $\mathbb{P}_{0,x}^{t,y}$ is not a probability measure. $\mathbb{P}_{0,x}^{t,y}$ has total mass $p(t, x, y)$ where p denotes the probability kernel of the Brownian motion.

THEOREM 2.9 (Feynman–Kac formula). *If $V \in L_{\text{loc}, \text{unif}}^p(\mathbb{R}^d)$ for $p = 2$ if $d \leq 3$, $p > d/2$ if $d \geq 3$, then e^{-tH} has a jointly continuous integral kernel given by*

$$\begin{aligned} e^{-tH}(x, y) &= \mathbb{E}_{0,x}^{t,y} (e^{-\int_0^t V(b(s)) ds}) \\ &= \int e^{-\int_0^t V(b(s)) ds} d\mathbb{P}_{0,x}^{t,y}(b). \end{aligned} \tag{2.11}$$

The integration here is over paths $b(\cdot) \in C([0, t])$.

We remind the reader that we always assume bounded potentials so that the conditions in Theorem 2.9 are satisfied. In the context of the density of states, we are interested in a Feynman–Kac formula for Hamiltonians on bounded domains. Let us denote by Ω_Λ^t the set of all paths staying inside Λ up to time t , i.e.,

$$\Omega_\Lambda^t = \{b \in C([0, t]) : |b(s)| \in \Lambda \text{ for all } 0 \leq s \leq t\}.$$

Then $e^{-tH_\Lambda^D}$ is simply given by restricting the integration in (2.11) to the set Ω_Λ^t , e.g.,

$$e^{-tH_\Lambda^D}(x, y) = \mathbb{E}_{0,x}^{t,y} (e^{-\int_0^t V(b(s)) ds} \chi_{\Omega_\Lambda^t}).$$

A proof can be found in Simon [154] and Aizenman–Simon [1]. There is also a Feynman–Kac formula for Neumann boundary conditions (see [74] and references given there).

Now we are able to state the probabilistic representation of the density of states measure in terms of Brownian motion.

THEOREM 2.10. *The Laplace transforms of $N(E)$, $N^D(E)$ and $N^N(E)$ agree and are given by*

$$\tilde{N}(t) = \tilde{N}^D(t) = \tilde{N}^N(t) = \mathbb{E} \times \mathbb{E}_{0,0}^{t,0} (e^{-\int_0^t V_\omega(b(s)) ds}). \tag{2.12}$$

For a detailed proof see, e.g., [68] and references there. The first step in the proof is to check that the right-hand side of (2.12) is finite for all $t \geq 0$. Interchanging the expectation values with respect to random potential and the Brownian motion, this follows from Jensen's inequality.

The second step of the proof is to compare (2.12) and the Laplace transforms of the approximating density of states measure ν_L and ν_L^D . To prove the theorem one has to estimate the hitting probability of the boundary of Λ for a Brownian motion starting and ending far away from the boundary. Using standard facts of Brownian motion, this tends to 0 in the limit $|\Lambda| \rightarrow \infty$.

Once we know that the Laplace transforms of N , N^D and N^N agree, it follows from the uniqueness of the Laplace transform that N , N^D and N^N agree themselves (see, e.g., [40]).

2.5. The Density of Surface States. We would like to define a density of states measure for surface potentials as well. Suppose we have a surface potential of the form

$$V_\omega^s(x_1, x_2) = \sum_{i_1 \in \mathbb{Z}^{d_1}} q_{i_1}(\omega) f(x_1 - i_1, x_2)$$

where, as above, $x \in \mathbb{R}^d$ is written as $x = (x_1, x_2)$ with $x_1 \in \mathbb{R}^{d_1}$, $x_2 \in \mathbb{R}^{d_2}$. In addition to the surface potential, there may be a random or periodic potential $V^b(x)$, which we call the “bulk” potential. The bulk potential should be stationary and ergodic with respect to shifts T_j parallel to the surface. Stationarity perpendicular to the surface is not required in the following.

This allows “interfaces” in the following sense: Let $d_1 = d - 1$, so the surface has codimension one. Thus it forms the interface between the upper half space $V_+ = \{x; x_2 > 0\}$ and the lower half space V_- . The bulk potential V^b may then be defined by $V^b(x) = V_1(x)$ for $x_2 \geq 0$ and $= V_2(x)$ for $x_2 < 0$. Here V_1 and V_2 are random or periodic potentials on \mathbb{R}^d . We set $H^b = H_0 + V^b$, which we call the “bulk operator” and $H_\omega = H^b + V_\omega^s$.

We could try to define a density of states measure in the same way as in (2.1), i.e., look at

$$\lim_{L \rightarrow \infty} \frac{1}{L^d} \text{tr}(\varphi(H_\omega) \chi_{\Lambda_L}). \quad (2.13)$$

It is not hard to see that this limit exists and equals

$$\mathbb{E}(\text{tr}(\chi_{\Lambda_1} \varphi(H^b) \chi_{\Lambda_1})). \quad (2.14)$$

In other words, (2.13) gives the density of states measure for the bulk operator. After all, this is not really surprising. The normalization with the volume term L^d is obviously destroying any influence of the surface potential.

So it sounds reasonable to choose a surface term like L^{d_1} as normalization and to consider

$$\lim_{L \rightarrow \infty} \frac{1}{L^{d_1}} \text{tr}(\varphi(H_\omega) \chi_{\Lambda_L}). \quad (2.15)$$

However, Definition (2.15) gives a finite result only when $\text{supp } \varphi \cap \sigma(H^b) = \emptyset$.

To define the density of surface states also inside the spectrum of the bulk operator, we therefore set

$$\nu_s(\varphi) = \lim_{L \rightarrow \infty} \frac{1}{L^{d_1}} \text{tr}((\varphi(H_\omega) - \varphi(H^b)) \chi_{\Lambda_L}). \quad (2.16)$$

Of course, it is not obvious at all that the limit (2.16) exists. In [36, 37] the authors proved that the limit exists for functions $\varphi \in C_0^3(\mathbb{R})$. Hence the density of surface states is defined as a distribution. The order of this distribution is at most 3. Observe that, in contrast to the density of states, the limit in (2.16) is not necessarily positive for positive φ due to the subtraction term. In fact, in the discrete case it is not hard to see that the total integral of the density of states, i.e., $\nu_s(\mathbf{1})$, is zero. Therefore, we cannot conclude that the density of surface states is a (positive) measure.

Kostykin and Schrader [106, 107] proved that the density of surface states distribution is actually the derivative of a measurable locally integrable function. They do not prove that this function is of bounded variation, thus leaving the possibility that ν_s is *not* given by a measure. See also the papers [16, 17] by Chahrouh for regularity properties of the density of surface states on the lattice.

Outside the spectrum of H^b , the distribution ν_s is positive, so that the density of surface states *is* a measure *there*. In [87] it was proven that below the “bulk” spectrum $\sigma(H^b)$ the density of surface states can also be defined by using (Neumann or Dirichlet) boundary conditions. We expect this to be wrong inside the bulk spectrum.

3. Lifshitz Tails

3.1. The Problem. For a periodic potential V the integrated density of states $N(E)$ behaves near the bottom E_0 of the spectrum $\sigma(H_0 + V)$ like

$$N(E) \sim C(E - E_0)^{d/2}. \tag{3.1}$$

This can be shown by explicit calculation for $V \equiv 0$ and was proved for general periodic potentials in [81].

On the basis of physical arguments Lifshitz [117, 118] predicted a completely different behavior for disordered systems, namely,

$$N(E) \sim C_1 e^{-C_2(E - E_0)^{-d/2}} \tag{3.2}$$

as $E \searrow E_0 > -\infty$. This behavior of $N(E)$ is called *Lifshitz behavior* or *Lifshitz tails*. The reason for this peculiar behavior is a collective phenomenon. To simplify the following heuristic argument, let us assume that $V_\omega \geq 0$ and $E_0 = 0$. To find an eigenvalue smaller than E , the potential V_ω has to be small on a rather large region in space. In fact, to have an eigenvalue at small $E > 0$, the uncertainty principle (i.e., the kinetic energy) forces the potential to be smaller than E on a set whose volume is of the order $E^{-d/2}$. That V_ω is small on a large set is a typical “large deviations event” which is very rare—in fact, its probability is exponentially small in terms of the volume of the set, i.e., its probability is of the order

$$e^{-C_2 E^{-d/2}} \tag{3.3}$$

which is precisely the behavior (3.2) predicted by Lifshitz. It is the aim of this section to discuss the Lifshitz behavior (3.2) of the integrated density of states as well as its extensions and limitations.

The first proof of Lifshitz behavior (for the Poisson model (1.8)) was given by Donsker and Varadhan [33]. They estimated the Laplace transform $\tilde{N}(t)$ for $t \rightarrow \infty$ using the Feynman–Kac representation on N (see Section 2.4). Their estimate relied on an investigation of the “Wiener sausage” and the machinery of large deviations for Markov processes developed by these authors. To obtain information about

the behavior of $N(E)$ for $E \searrow 0 = E_0$ from the large t behavior of $\tilde{N}(t)$ one uses Tauberian theorems [8, 11]. This technique was already used by Pastur [145, 6] and developed in [142, 44] and recently in [125].

Donsker and Varadhan [33] needed in their proof of (3.2) that the single site potential f decays faster than $(1 + |x|)^{-(d+2)}$. They asked whether this condition is necessary for the result (Lifshitz tails) or just necessary for *their proof*. It was Pastur ([146]) who observed that, in fact, the Lifshitz asymptotic is qualitatively changed if f has long range tails, i.e., if $f(x) \sim C(1 + |x|)^{-\alpha}$ for $\alpha < d + 2$. Observe that $\alpha > d$ is necessary for the mere existence of V_ω . Pastur proved the behavior

$$N(E) \sim C_1 e^{-C_2(E-E_0)^{-\frac{d}{\alpha-d}}} \quad (3.4)$$

as $E \searrow E_0$ for $d < \alpha < d + 2$. We call this behavior *Pastur tails*. For a disordered system with constant magnetic field in dimension $d = 2$, Pastur tails (3.4) were found for *all* $\alpha > d = 2$ in [10].

These results and more observations of the last several years indicate that the asymptotics of the integrated density of states even at the bottom of the spectrum is more complicated than expected. To be more precise and following the terminology of [148], we can distinguish two qualitatively different behaviors in the low energy asymptotics of the integrated density of states. For short range potentials and “fat” single site distributions, the asymptotics of $N(E)$ is determined by the quantum kinetic energy as predicted by Lifshitz. Hence it is called quantum asymptotics or *quantum regime*. On the other hand, for long range potentials or “thin” single site distributions, the leading asymptotics of the integrated density of states is determined by the potential, i.e., by classical effects. This situation is called the *classical regime*.

We will discuss these phenomena in this section. We start with the short range case (quantum regime). The proof of Lifshitz tails we present here is based on spectral theoretic arguments close to Lifshitz’s original heuristics (see Section 3.2).

We then discuss the long range case (classical regime) (Section 3.3) to some extent, including recent results [86] of single site potentials with anisotropic decay resulting in a mixed classical-quantum regime (Section 3.4).

Classical and quantum behavior of the integrated density of states and the transition between the two regimes is best understood for the Anderson model. The approach of [125] combines spectral theoretic and path integral methods. We will present this in Section 3.5.

Lifshitz predicted the behavior (3.1) and (3.2) not only at the bottom of the spectrum but also for any band edge of the spectrum. To distinguish these two cases, we will speak of *internal Lifshitz tails* in the latter case. Investigating Lifshitz behavior at internal band edges turns out to be much more complicated than at the bottom of the spectrum. In fact, already the investigation of periodic potentials at internal band edges is extremely complicated. We will discuss internal Lifshitz tails (following [97, 104, 105]) in Section 3.6.

Finally, we will look at random Schrödinger operators with magnetic fields in Section 3.8.

3.2. Lifshitz Tails: Quantum Case.

3.2.1. *Statement of the main result.* The aim of this subsection is a proof of Lifshitz behavior close to his original heuristics and without heavy machinery. We will prove the quantum asymptotics in (3.2) for short range single site potentials

and “fat” single site distributions. We will make no attempt to reach high generality but rather emphasize the strategy of the proof.

As before, we consider random alloy-type potentials of the form

$$V_\omega(x) = \sum_{i \in \mathbb{Z}^d} q_i(\omega) f(x - i). \tag{3.5}$$

We assume that the random variables q_i are independent and identically distributed with a common probability distribution P_0 . We suppose that the support of P_0 is compact and contains at least two points.

As always, we also suppose that the single site potential f is non-negative, bounded and decays at infinity as fast as $|x|^{-(d+\varepsilon)}$. The technique we are going to present allows us to treat local singularities of f . (We refer to [80, 86] for details.)

To ensure Lifshitz tails in the sense of (3.2), we need two conditions:

Assumption 1: Define $q_{\min} = \inf \text{supp}(P_0)$. We assume that

$$P_0([q_{\min}, q_{\min} + \varepsilon)) \geq C \varepsilon^N \tag{3.6}$$

for some C, N and all $\varepsilon > 0$ small.

Condition (3.6) means that the distribution P_0 is “fat” at the bottom of its support. Note that this condition is, in particular, satisfied if P_0 has an atom at q_{\min} , i.e., if $P_0(\{q_{\min}\}) > 0$.

The second condition we need is precisely the “short range” condition already encountered by Donsker and Varadhan [33].

Assumption 2:

$$f(x) \leq C (1 + |x|)^{-(d+2)}. \tag{3.7}$$

We are ready to formulate the main result of this subsection.

THEOREM 3.1. *If Assumptions (1) and (2) are satisfied, we have*

$$\lim_{E \searrow E_0} \frac{\ln(-\ln N(E))}{\ln(E - E_0)} = -\frac{d}{2}. \tag{3.8}$$

Observe that equation (3.8) is a weak form of Lifshitz’s original conjecture (3.2). In their work [33], Donsker and Varadhan proved the stronger form for the Poisson potential

$$\lim_{E \searrow 0} \frac{\ln N(E)}{E^{-d/2}} = -C_d \tag{3.9}$$

where C_d is a (computable) positive constant.

Both the short range condition (Assumption 2) and the fatness condition (Assumption 1) turn out to be necessary for the above result, as we will see later. For example, if the single site potential f decays substantially slower than required in the short range condition the integrated density of states decays *faster* than in the (3.8).

We define the *Lifshitz exponent* γ by

$$\gamma = \lim_{E \searrow E_0} \frac{\ln(-\ln N(E))}{\ln(E - E_0)} \tag{3.10}$$

whenever this limit exists. With this notation we may rephrase (3.8) as $\gamma = -d/2$. The Lifshitz exponent for periodic potentials is 0.

3.2.2. *Strategy of the proof.* The proof of Theorem 3.1 consists of an upper and a lower bound. The next subsection will provide us with the tools we need for these bounds.

It will turn out that the bounds are easier and more natural for positive random potentials. Therefore, we will split the random potential in a periodic and a positive random part

$$\begin{aligned} V_\omega(x) &= \sum_{i \in \mathbb{Z}^d} q_{\min} f(x - i) + \sum_{i \in \mathbb{Z}^d} (q_i(\omega) - q_{\min}) f(x - i) \\ &= V_{\text{per}}(x) + \tilde{V}_\omega(x). \end{aligned} \tag{3.11}$$

We will subsume the periodic potential under the kinetic energy and denote the positive random potential \tilde{V}_ω in a slight abuse of notation again by V_ω . Thus we have

$$H_\omega = H_1 + V_\omega \tag{3.12}$$

with $H_1 = H_0 + V_{\text{per}}$ a Hamiltonian with periodic “background” potential V_{per} and

$$V_\omega = \sum_{i \in \mathbb{Z}^d} q_i(\omega) f(x - i) \tag{3.13}$$

where the independent $q_i \geq 0$ have a common probability distribution P_0 with $0 = \inf(\text{supp}P_0)$.

For the upper bound below, we need information about the two lowest eigenvalues of H_1 restricted to a box. If $V_{\text{per}} \equiv 0$, these eigenvalues can be computed explicitly. However, if $V_{\text{per}} \neq 0$, we need a careful analysis of periodic operators. This was done in [81] and [129]. Here we restrict ourselves to the case $V_{\text{per}} \equiv 0$, avoiding some technical complications. Note that this implies $E_0 = \inf(\sigma(H_\omega)) = 0$. We refer the reader to the papers [80] and [129] for the general case. We also remark that [86] contains an extension of the approach presented here that works for Poisson potentials and various other potentials as well.

3.2.3. *The Dirichlet–Neumann bracketing.* The first step in the proof is to bound the integrated density of states from above and from below using the Dirichlet–Neumann bracketing as in Corollary 2.7. We have

$$\frac{1}{|\Lambda_L|} \mathbb{E}(N(H_{\omega\Lambda_L}^D, E)) \leq N(E) \leq \frac{1}{|\Lambda_L|} \mathbb{E}(N(H_{\omega\Lambda_L}^N, E)). \tag{3.14}$$

The side length L of the cube Λ_L will be chosen later in an E -dependent way when we send E to E_0 . We estimate the right-hand side of (3.14) by

$$\begin{aligned} \mathbb{E}(N(H_{\omega\Lambda_L}^N, E)) &= \int N(H_{\omega\Lambda_L}^N, E) d\mathbb{P} \\ &= \int_{E_1(H_{\omega\Lambda_L}^N) \leq E} N(H_{\omega\Lambda_L}^N, E) d\mathbb{P} + \int_{E_1(H_{\omega\Lambda_L}^N) > E} N(H_{\omega\Lambda_L}^N, E) d\mathbb{P} \\ &\leq \mathbb{P}(E_1(H_{\omega\Lambda_L}^N) \leq E) N(H_0_{\Lambda_L}^N, E). \end{aligned}$$

With $N(H_0_{\Lambda_L}^N, E) \leq (C_1 + C_2 E)^{d/2} |\Lambda|$ following from Weyl asymptotics, we get for $0 \leq E \leq 1$ the estimate

$$\frac{1}{|\Lambda_L|} \mathbb{P}(E_1(H_{\omega\Lambda_L}^D) \leq E) \leq N(E) \leq C \mathbb{P}(E_1(H_{\omega\Lambda_L}^N) \leq E). \tag{3.15}$$

The problem now is to find upper and lower bounds such that after taking the double logarithm, the left- and the right-hand side of (3.15) coincide asymptotically. In general the upper bounds are more difficult than the lower bounds. To prove the lower bound we only have to “guess” a good test function, whereas for the upper bounds, one has to prove that all eigenfunctions for energies in $[0, E)$ roughly behave the same way.

It is an astonishing fact that the lower bound from (3.15) in all known cases leads to the *asymptotically correct* behavior of the integrated density of states, a fact emphasized by Pastur.

3.2.4. The lower bound. For simplicity we restrict ourselves to single site potentials f with $\text{supp } f \subset \Lambda_{\frac{1}{2}}$ so that $f(\cdot - i)$ and $f(\cdot - j)$ do not overlap for $i \neq j$. The necessary changes for the general case will become clear when we discuss long range potentials f .

By the Neumann–Dirichlet bracketing in (3.15), we have for arbitrary L and any $\psi \in D(\Delta_{\Lambda_L}^D)$ with $\|\psi\|_{L^2(\Lambda_L)} = 1$,

$$\begin{aligned} N(E) &\geq |\Lambda_L|^{-1} \mathbb{P} (E_1(H_{\omega_{\Lambda_L}}^D) < E) \\ &\geq |\Lambda_L|^{-1} \mathbb{P} (\langle \psi, H_{\omega} \psi \rangle_{L^2(\Lambda_L)} < E). \end{aligned} \tag{3.16}$$

A natural choice of ψ for (3.16) seems to be the ground state ψ_0 of $-\Delta_{\Lambda_L}^N$, $\psi_0(x) \equiv |\Lambda_L|^{-\frac{1}{2}}$. Unfortunately, this function does not obey Dirichlet boundary conditions and is therefore not admissible for (3.16).

This problem can be circumvented by multiplying ψ_0 by a function which is zero at the boundary of Λ_0 . To do so, let us take $\chi \in C^\infty(\mathbb{R}^d)$, $\text{supp } \chi \subset \Lambda_1$, $\chi(x) = 1$ on $\Lambda_{\frac{1}{2}}$ and $0 \leq \chi(x) \leq 1$. We set $\chi_L(x) = \chi(\frac{x}{L})$ and $\psi_L(x) = \chi_L(x)\psi_0(x)$. Then $\psi_L \in D(\Delta_{\Lambda_L}^D)$, $\|\psi_L\| \geq \frac{1}{2}$ and

$$\begin{aligned} \langle \psi_L, H_{\omega} \psi_L \rangle &\leq \langle \psi_0, H_{\omega} \psi_0 \rangle + CL^{-2} \\ &= |\Lambda_L|^{-1} \int_{\Lambda_L} V_{\omega}(x) dx + CL^{-2}. \end{aligned}$$

Note that the “error term” L^{-2} is due to the influence of the kinetic energy (a second-order differential operator). Inserting in (3.16) we get

$$\begin{aligned} N(E) &\geq |\Lambda_L|^{-1} \mathbb{P} \left(|\Lambda_L|^{-1} \int_{\Lambda_L} V_{\omega}(x) < E - CL^{-2} \right) \\ &\geq |\Lambda_L|^{-1} \mathbb{P} \left(|\Lambda_L|^{-1} \|f\|_1 \sum_{i \in \Lambda_L} q_i(\omega) < E - CL^{-2} \right). \end{aligned} \tag{3.17}$$

In principle, we can choose L as we like. However, if $E < CL^{-2}$ estimate (3.17) becomes useless. So it seems reasonable to choose $L = \beta E^{-\frac{1}{2}}$ and we obtain

$$\begin{aligned} (3.17) &\geq |\Lambda_L|^{-1} \mathbb{P} \left(|\Lambda_L|^{-1} \sum_{i \in \Lambda_L} q_i(\omega) < \tilde{C} E \right) \\ &\geq |\Lambda_L|^{-1} \mathbb{P} (q_0 < \tilde{C} E)^{L^d} \\ &\geq C_1 E^{-\frac{d}{2}} (C E^N)^{C_2 E^{-\frac{d}{2}}}. \end{aligned}$$

Thus we conclude

$$\lim_{E \searrow 0} \frac{\ln(-\ln(N(E)))}{\ln E} \geq -\frac{d}{2}. \quad (3.18)$$

3.2.5. The upper bound. The strategy to prove the upper bounds of $\mathbb{P}(E_1(H_{\omega_{\Lambda_L}}^N) < E)$ in (3.15) can be divided in two parts. The first step is to find a *lower* bound for $E_1(H_{\omega_{\Lambda_L}}^D)$ in such a way that it is possible to control the influence of the random potential. This is done by an application of Temple's inequality. The second step is to balance between the size of Λ_L and the probability of a random potential such that most of the potential values are small.

We start by stating Temple's inequality for the reader's convenience. A proof can be found, e.g., in [150].

THEOREM 3.2 (Temple's inequality). *Suppose H is a self-adjoint operator, bounded below which has discrete spectrum and denote by $E_n(H), n = 1, 2, \dots$ its eigenvalues (in increasing order, counting multiplicity). If $\mu \leq E_2(H)$ and $\psi \in D(H)$ with $\|\psi\| = 1$ satisfying $\langle \psi, H\psi \rangle < \mu$, then*

$$E_1(H) \geq \langle \psi, H\psi \rangle - \frac{\langle \psi, H^2\psi \rangle - \langle \psi, H\psi \rangle^2}{\mu - \langle \psi, H\psi \rangle}.$$

To apply Temple's inequality, we set $E_2(-\Delta_{\Lambda_L}^N) := \mu \leq E_2(H_{\omega_{\Lambda_L}}^N)$. Note that by direct computation, $E_1(-\Delta_{\Lambda_L}^N) = 0 = E_0$ and $E_2(-\Delta_{\Lambda_L}^N) \sim L^{-2}$. These facts require a careful analysis if there is a periodic background potential as in (3.11) and (3.12); see [81].

Next we need a good approximation ψ of the ground state of $H_{\omega_{\Lambda_L}}^N$. This is done by choosing ψ to be the ground state of $-\Delta_{\Lambda_L}^N$, which is intuitively close to the correct ground state for small E . The function ψ is given by $\psi(x) = |\Lambda_L|^{-1/2}$.

To apply Temple's inequality we have to ensure that with the above choice, $\langle \psi, H\psi \rangle < \mu \approx cL^{-2}$. We force this to happen by changing the coupling constants q_i to $\tilde{q}_i = \min(q_i(\omega), \alpha L^{-2})$ with a suitable $\alpha > 0$, small enough. If \tilde{H} denotes the corresponding operator, we have $E_1(H) \geq E_1(\tilde{H})$. An application of Temple's inequality to \tilde{H} and an elementary calculation yield the following lemma.

LEMMA 3.3.

$$E_1(H_{\omega_{\Lambda_L}}^N) \geq \frac{1}{2} \frac{1}{|\Lambda_L|} \sum_{i \in \Lambda_L} \tilde{q}_i(\omega). \quad (3.19)$$

A consequence of the lemma above is the intuitively convincing estimate

$$\mathbb{P}(E_1(H_{\omega_{\Lambda_L}}^N) < E) \leq \mathbb{P}\left(\frac{1}{|\Lambda_L|} \sum_{i \in \Lambda_L} \tilde{q}_i(\omega) \leq 2E\right) \quad (3.20)$$

$$\leq \mathbb{P}\left(\frac{1}{L^d} \sum_{|i|_\infty \leq L/2} \tilde{q}_i(\omega) \leq 2E\right). \quad (3.21)$$

The expression (3.21) for E small very much resembles a large deviation probability which would lead to a bound exponentially small in the volume term L^d . At first sight, Cramer's theorem, a result of the theory of large deviation, seems to be applicable (see, e.g., [31, 32, 58]).

However, there is a complication here: To obtain a large deviation event in (3.21) we need that $\mathbb{E}(\tilde{q}_i) < E$. Thus, if we set $L = L(E) = \beta E^{-1/2}$ with $\beta > 0$

small, the event (3.21) is, indeed, a large deviation event and we obtain the following bound.

LEMMA 3.4.

$$\mathbb{P} \left(\frac{1}{|\Lambda_L|} \sum_{i \in \Lambda_L} \tilde{q}_i(\omega) \leq E \right) \leq C_1 e^{-C_2 L^d}$$

for E close enough to zero and $L \leq \beta E^{-1/2}$.

Combining the results above, (3.15) and (3.20), we have proven

$$N(E) \leq C_1^{-C_2 E^{-d/2}}.$$

3.2.6. *Final remarks.* The idea of using Neumann–Dirichlet bracketing to prove Lifshitz tails first appeared in [77]. It was carried over to the discrete Anderson model by Simon [156] who streamlined it at the same time. The proof was extended to more general alloy-type potentials by Kirsch and Simon [80], who still needed reflection symmetry of f . Mezincescu [129] modified the upper bound by introducing other boundary conditions to get rid of this extra assumption. We refer to [86] for a rather general proof using these techniques.

3.3. Long Range Single Site Potentials: A “Classical” Case. In this section we turn to an example of classical behavior of the integrated density of states near $0 = \inf(\sigma(H_\omega))$, in the sense of Section 3.1, namely, to long range single site potential f .

The upper bound on $N(E)$ is easier than for the short range case. While there is a subtle interplay between the kinetic energy and the potential in the short range case ($f(x) \leq |x|^{-(d+2)}$), it is the potential energy alone that determines the leading behavior of $N(E)$ ($E \searrow 0$) in the long range case.

Assumption: In this section we suppose that

$$\frac{c}{(1 + |x|)^\alpha} \leq f(x) \leq \frac{C}{(1 + |x|)^\alpha}. \tag{3.22}$$

THEOREM 3.5. *Assume (3.6). If (3.22) holds for an α with $d < \alpha < d + 2$, then*

$$\lim_{E \searrow E_0} \frac{\ln(-\ln N(E))}{\ln E} = -\frac{d}{\alpha - d}. \tag{3.23}$$

In the terminology of (3.10), Theorem 3.5 states that the Lifshitz exponent for the long range case ($\alpha < d + 2$) is $d/(\alpha - d)$.

PROOF. To simplify the argument, we assume as in the short range case, there is no periodic background potential and $q_{\min} = 0$. Consequently, $E_0 = 0$. We start with the upper bound and estimate

$$E_1(H_{\omega \Lambda_1}^N) \geq \inf_{x \in \Lambda_1} \sum_{i \in \mathbb{Z}^d} q_i \frac{c}{(x + |i|)^\alpha}.$$

Hence

$$\begin{aligned}
 N(E) &\leq C_1 \mathbb{P}(E_1(H_{\omega_{\Lambda_1}}^N) < E) \\
 &\leq C_1 \mathbb{P}\left(\sum_{i \in \mathbb{Z}^d} q_i \frac{C_2}{(1+|i|)^\alpha} < E\right) \\
 &\leq C_1 \mathbb{P}\left(\sum_{|i| \leq L} q_i \frac{C'}{L^\alpha} < E\right) \\
 &\leq C_1 \mathbb{P}\left(\frac{1}{L^d} \sum_{|i| \leq L} q_i < C_3 E L^{\alpha-d}\right). \tag{3.24}
 \end{aligned}$$

We choose $L = \frac{\beta}{C_3} E^{-\frac{1}{\alpha-d}}$ (β small). Hence

$$(3.24) \leq C_1 \mathbb{P}\left(\frac{1}{L^d} \sum_{|i| \leq L} q_i \leq \beta\right).$$

If $\beta < \frac{1}{2} \mathbb{E}(q_0)$, standard large deviation theory gives

$$\mathbb{P}\left(\frac{1}{L^d} \sum_{|i| \leq L} q_i \leq \beta\right) < e^{-CL^d} = e^{-\tilde{C}E^{-\frac{d}{\alpha-d}}}.$$

We turn to the lower bound. As in the proof of the lower bound (3.17) in the previous section,

$$\begin{aligned}
 N(E) &\geq \frac{1}{|\Lambda_L|} \mathbb{P}\left(\frac{1}{|\Lambda_L|} \int_{\Lambda_L} V_\omega(x) dx < E - CL^{-2}\right) \\
 &\geq \frac{1}{|\Lambda_L|} \mathbb{P}\left(\sum_{i \in \mathbb{Z}^d} q_i \frac{1}{|\Lambda_L|} \int_{\Lambda_L} f(x-i) dx < E - CL^{-2}\right).
 \end{aligned}$$

Due to the long range tails of f , we cannot ignore the summands with $|i|$ large. Instead, we estimate

$$\begin{aligned}
 &\sum_{i \in \mathbb{Z}^d} q_i \frac{1}{|\Lambda_L|} \int_{\Lambda_L} f(x-i) dx \\
 &\leq \frac{1}{|\Lambda_L|} \sum_{|i|_\infty \leq 2L} q_i \int f(y) dy + \sum_{|i|_\infty > 2L} q_i \frac{1}{|\Lambda_L|} \int_{\Lambda_L} f(x-i) dx \\
 &\leq \frac{C_3}{|\Lambda_{2L}|} \sum_{|i|_\infty \leq 2L} q_i + q_{\max} \sum_{|i|_\infty > 2L} \frac{1}{|\Lambda_L|} \int_{\Lambda_L} f(x-i) dx,
 \end{aligned}$$

where $q_{\max} = \sup(\text{supp} P_0)$, P_0 being the distribution of the q_i . We estimate

$$\begin{aligned}
 \sum_{|i| > 2L} \frac{1}{|\Lambda_L|} \int_{\Lambda_L} f(x-i) dx &\leq C_4 \sum_{|i| > 2L} \frac{1}{|\Lambda_L|} \int_{\Lambda_L} \frac{1}{|x-i|^\alpha} dx \\
 &\leq C_5 \sum_{|i| > L} \frac{1}{|i|^\alpha} \\
 &\leq C_6 \frac{1}{L^{\alpha-d}}.
 \end{aligned}$$

Thus we obtain

$$N(E) \geq \frac{1}{|\Lambda_L|} \mathbb{P} \left(\frac{1}{|\Lambda_{2L}|} \sum_{|i| \leq 2L} q_i \leq C_7 E - C_8 L^{-2} - C_9 L^{(\alpha-d)} \right). \tag{3.25}$$

As the derivation shows, the L^{-2} term comes from manipulating the kinetic energy, while the $L^{-(\alpha-d)}$ term is due to the potential energy. Note that for $\alpha > d + 2$, the term L^{-2} (kinetic energy contribution) dominates in (3.25). In this case we can therefore redo the estimates of Section 3.2.4 and obtain a lower bound as we got there. However, for $\alpha < d + 2$, the term $L^{-(\alpha-d)}$ wins out in (3.25). Remember, this term is due to the potential energy distribution. We obtain

$$N(E) \geq \frac{1}{|\Lambda_L|} \mathbb{P} \left(\frac{1}{|\Lambda_{2L}|} \sum_{|i| \leq 2L} q_i \leq C_7 E - C_{10} L^{-(\alpha-d)} \right).$$

This time, E has to be bigger than $L^{-(\alpha-d)}$; more precisely, $E \geq C_{11} L^{-(\alpha-d)}$. Hence $L = C_{12} E^{-\frac{1}{\alpha-d}}$ so

$$\begin{aligned} N(E) &\geq \frac{1}{|\Lambda_L|} \mathbb{P}(q_i = 0 \text{ for } |i| \leq 2L) \\ &\geq \frac{1}{|\Lambda_L|} e^{-C_{13} L^d} \\ &\geq C_{14} E^{\frac{d}{\alpha-d}} e^{-C_{15} E^{-\frac{d}{\alpha-d}}}. \end{aligned} \quad \square$$

3.4. Anisotropic Single Site Potentials. Recently, Theorems 3.1 and 3.5 were generalized to single site potentials f decaying in an anisotropic way at infinity ([86]). Let us write $x \in \mathbb{R}^d = \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ as $x = (x_1, x_2)$, $x_1 \in \mathbb{R}^{d_1}$, $x_2 \in \mathbb{R}^{d_2}$ and suppose that

$$\frac{a}{|x_1|^{\alpha_1} + |x_2|^{\alpha_2}} \leq f(x) \leq \frac{b}{|x_1|^{\alpha_1} + |x_2|^{\alpha_2}} \tag{3.26}$$

for $|x_1|, |x_2| \geq 1$, and define V_ω in the usual way

$$V_\omega(x) = \sum_{i \in \mathbb{Z}^d} q_i f(x - i). \tag{3.27}$$

Let us define $\gamma_i = \frac{d_i}{\alpha_i}$ and $\gamma = \gamma_1 + \gamma_2$. Then the sum in (3.27) converges (absolutely) if $\gamma < 1$.

In [86] the authors prove that there is Lifshitz behavior of $N(E)$ for potential as in (3.26) and (3.27) in the sense that the Lifshitz exponent η , defined by

$$\eta = \lim_{E \searrow E_0} \frac{\ln |\ln(N(E))|}{\ln E}$$

exists ($E_0 = \inf \sigma(H_\omega)$). η depends on the exponents α_i , of course. If both

$$\frac{\gamma_1}{1 - \gamma} \leq \frac{d_1}{2} \quad \text{and} \quad \frac{\gamma_2}{1 - \gamma} \leq \frac{d_2}{2} \tag{3.28}$$

we obtain the ‘‘quantum’’ exponent:

$$\eta = -\frac{d}{2}.$$

Observe that (3.28) reduce to the condition $\alpha \leq d + 2$ for the isotropic case $\alpha_1 = \alpha_2 = \alpha$. If

$$\frac{\gamma_1}{1-\gamma} > \frac{d_1}{2} \quad \text{and} \quad \frac{\gamma_2}{1-\gamma} > \frac{d_2}{2} \quad (3.29)$$

we are in the “classical” case both in the d_1 - and d_2 -directions. Then

$$\eta = -\frac{\gamma}{1-\gamma}.$$

The third case

$$\frac{\gamma_1}{1-\gamma} \leq \frac{d_1}{2} \quad \text{and} \quad \frac{\gamma_2}{1-\gamma} > \frac{d_2}{2} \quad (3.30)$$

is new compared to the isotropic case. It is, in a sense, a mixed quantum-classical case. The Lifshitz exponent is given by

$$\eta = -\frac{d_1}{2} - \frac{\gamma_2}{1-\gamma}.$$

We note that the d_1 -direction and the d_2 -direction “influence each other” in a rather sophisticated way. In [86] these results are proved for alloy-type potentials as well as for Poisson (and related) models. We summarize:

THEOREM 3.6 ([86]). *Suppose (3.6) and (3.26) hold. Set $\gamma_i = d_i/\alpha_i$ and $\gamma = \gamma_1 + \gamma_2$. Then the Lifshitz exponent η is given by*

$$\eta = -\max \left\{ \frac{d_1}{2}, \frac{\gamma_1}{1-\gamma} \right\} - \max \left\{ \frac{d_2}{2}, \frac{\gamma_2}{1-\gamma} \right\}. \quad (3.31)$$

3.5. Path Integral Methods and the Transition Between Quantum and Classical Regime with Respect to the Single Site Measure. We start this section with two observations which indicate that the asymptotic behavior of the integrated density of states depends qualitatively on the distribution P_0 of the q_i .

Let us first assume that the “fatness” condition (3.6) is satisfied in the strongest form, namely, $\inf \text{supp}(P_0) = 0$ and $P(q_0 = 0) = a > 0$. An inspection of the proofs in Section 3.2 shows that for this case we have actually proven

$$\liminf \frac{\ln N(E)}{(E - E_0)^{-d/2}} \geq -C_1, \quad (3.32)$$

$$\limsup \frac{\ln N(E)}{(E - E_0)^{-d/2}} \leq -C_2 \quad (3.33)$$

with $C_1, C_2 > 0$. If instead $P(q_0 = 0) = 0$ (but still $P(q_0 > \epsilon) \geq B\epsilon^n$) the lower bound requires a logarithmic correction

$$\liminf \frac{\ln N(E)}{(E - E_0)^{-d/2} |\ln(E - E_0)|} \geq -C_1. \quad (3.34)$$

The second observation concerns unbounded single site measures P_0 . In [98] it is proved that both the classical and the quantum regime can occur for the discrete, unbounded Anderson model and more general matrix operators. Depending on the single site measure, collective phenomena may occur similar to those we encountered above. In other situations, the single site measure alone determines the behavior of the integrated density of states.

It seems difficult to understand the mechanisms causing the transition from quantum to classical regime with respect to the single site measure by using the spectral analytic approach close to Lifshitz’s original intuition.

The first approach to prove Lifshitz tails is based on the Donsker–Varadhan technique (see also Section 3.1). This method to compute the Laplace transform of the integrated density of states in the limit $t \rightarrow \infty$ is a far-reaching generalization of the Laplace method known from classical analysis. The starting point is the path integral representation

$$\tilde{N}(t) = \mathbb{E} \times \mathbb{E}_{0,0}^{t,0} \left[e^{-\int_0^t V_\omega(b(s)) ds} \right].$$

The Donsker–Varadhan technique is based on a large deviation principle satisfied by the product probability measure $d\mathbb{P} \times d\mathbb{P}_{0,0}^{t,0}$ combining the random potential and the Brownian motion. In an informal sense, it makes it possible to quantize the asymptotic probability of a Brownian particle to stay most of its lifetime in a pocket with a favorable configuration of potential values. Using the large deviation principle, one can balance between favorable configurations and their small probability by applying Varadhan’s lemma. Last but not least, given the large time asymptotics of the Laplace transform $\tilde{N}(t)$, one can reconstruct the Lifshitz tail behavior using Tauber theory.

The Donsker–Varadhan technique was worked out by Nakao [142] for the Poisson model with $f \geq 0$. He proved

$$\lim_{E \searrow 0} \frac{\ln N(E)}{E^{-d/2}} = -C_d$$

where C_d is a (computable) positive constant. In the 1990’s Lifshitz asymptotics became a starting point for stochastic analysis of diffusion in random media. We mention the work of Sznitman (see, e.g, [167, 168, 169, 170]) in the continuous case, especially for Poisson potentials, and in the discrete context the moment analysis for the so-called parabolic Anderson model (PAM) (see, e.g., [7, 48, 49, 50]). Here Brownian motion has to be replaced by the continuous time Markov chain generated by the discrete Laplacian.

The phenomenology described at the beginning of this subsection was also observed in the moment analysis of the parabolic Anderson model starting in [48]. As we will see, the case of the double exponential distribution discussed in [49] can be interpreted as the borderline between the quantum and the classical regime. The paper [7] clarified the discrepancy between the lower and the upper bounds in (3.32), (3.33) and (3.34). Still, a general principle explaining the transition from quantum to classical regime with respect to the single site measure was not formulated.

We want to systemize the phenomenology discussed above in the following theorem taken from [125]. To combine the bounded and the unbounded case, we assume that the cumulant generating function is finite, e.g.,

$$G(t) := \log \mathbb{E} \left(\exp(-tV_\omega(0)) \right) < \infty \tag{3.35}$$

for all $t \geq 0$. Furthermore, we apply the Legendre transformation to define the rate function

$$I(E) := \sup_{t>0} [Et - G(t)] \tag{3.36}$$

and for $t, \lambda > 0$, we set

$$S(\lambda, t) := (\lambda t)^{-1}G(\lambda t) - t^{-1}G(t). \quad (3.37)$$

Informally the scale function $S(\lambda, t)$ measures the change of the cumulant generating function after rescaling the time.

THEOREM 3.7. *We consider the discrete Anderson tight binding operator $h_\omega = h_0 + V_\omega$ and set $E_0 = \inf \sigma(h_\omega)$. Suppose $G(t) < \infty$ for all $t > 0$. Then we distinguish the following four cases:*

(i) *Let $S(\lambda, t) \sim c(\lambda^\rho - 1)t^\rho$ with $c, \rho > 0$. Then the IDS behaves in the limit $E \rightarrow E_0 = -\infty$ like*

$$\log(N(E)) = -I(E + 2d + o(1))(1 + o(1)).$$

(ii) *In the case $S(\lambda, t) \sim c \log(\lambda)$ with $\rho = 0$ and $c > 0$, we have in the limit $E \rightarrow E_0 = -\infty$,*

$$-KI(E + C_2)(1 + o(1)) \leq \log(N(E)) \leq -I(E + C_1)(1 + o(1))$$

with $K > 0$,

$$C_1 = -2 \sin^2 \left(\frac{\pi}{2} \frac{1}{c^{-1/2} + 1} \right) + \frac{1}{4} c \log(c)$$

and

$$C_2 = K(d) \min \left[-c + c \log(c^{-1}), \max \left[1 - 4 \exp(-Kc), \frac{a(d)}{4} \right] \right].$$

(iii) *In the case $S(\lambda, t) \sim c(1 - \lambda^\rho)t^\rho$ with $-1 < \rho < 0$ and $c > 0$, the IDS behaves in the limit $E \searrow E_0 = 0$ like*

$$\begin{aligned} -K_1 E^{-1/2(d-2\rho^{-1}(\rho+1))}(1 + o(1)) &\leq \log(N(E)) \\ &\leq -K_2 E^{-1/2(d-2\rho^{-1}(\rho+1))}(1 + o(1)). \end{aligned}$$

(iv) *In the case $S(\lambda, t) \sim -c(\lambda t)^{-1} \log(t)$ with $c > 0$, we have in the limit $E \searrow E_0 = 0$,*

$$\begin{aligned} -K_1 E^{-d/2} \log(E)(1 + o(1)) &\leq \log(N(E)) \\ &\leq -K_2 E^{-d/2} \log(E)(1 + o(1)). \end{aligned}$$

The scaling assumption $S(\lambda, t) \sim c(\lambda^\rho - 1)t^\rho$ with $c, \rho > 0$ in the first case corresponds to “fat” unbounded single site distributions; the behavior of the integrated density of states is classical. The second case represents the double exponential case, while in the third situation the single site distribution is bounded, but very thin. The fourth case corresponds to relatively fat single site distributions studied in [74].

Although, by now, there are results covering a lot of possible single site distributions, there seems to be no systematic approach known to explain this phenomenology. Furthermore there exists two relatively different approaches as discussed above. A first step to combine the functional analytic and the path integral approach as well as to systemize the known results with respect to the single site distribution seem to be [125, 126]. In contrast to the direct analysis of the operator H_ω in Section 3.2, but in analogy to the path integral methods, one is interested

in the large time behavior of the semigroup $\exp(-h_\omega t)$. The Lifshitz asymptotics follows by an application of (modified) Tauber theorems [127].

The first step of the argument in [125] is to restrict h_ω to a (time-dependent) box $\Lambda = \Lambda_t(0)$ by introducing discrete Dirichlet boundary conditions and to approximate $\mathbb{E}[\exp(-h_\omega t)(0, 0)]$ in the limit $t \rightarrow \infty$ by

$$\mathbb{E}[\exp(-E_1(h_\omega^\Lambda)t)] = \mathbb{E} \left[\sup_{p \in M_1(\Lambda)} \exp(-t [(\sqrt{p}|h_0^\Lambda \sqrt{p}) + (\sqrt{p}|V(\omega)\sqrt{p})]) \right]. \tag{3.38}$$

Here $E_1(h_\omega^\Lambda) = \inf \sigma(h_\omega^\Lambda)$ is the principal eigenvalue of h_ω^Λ and $M_1(\Lambda)$ is the set of probability measures on Λ . Equation (3.38) is a consequence of the min-max principle and the nonnegativity of the ground state. It is the starting point to find upper and lower bounds of the Laplace transform of the IDS.

To illustrate the central effect explaining the transition from the quantum mechanical to the classical regime, we want to sketch the very elementary proof of the lower bounds starting from (3.38). The first step is to interchange the expectation value and the supremum

$$\begin{aligned} \mathbb{E} [\exp(-E_1(h_\omega^\Lambda)t)] &= \mathbb{E} \left[\sup_{p \in M_1(\Lambda)} \exp(-t [(\sqrt{p}|h_0^\Lambda \sqrt{p}) + (\sqrt{p}|V(\omega)\sqrt{p})]) \right] \\ &\geq \sup_{p \in M_1(\Lambda)} \exp(-t (\sqrt{p}|h_0^\Lambda \sqrt{p})) \mathbb{E} \left[\exp\left(-\sum_{x \in \Lambda} p(x)V_\omega(x)t\right) \right] \\ &= \sup_{p \in M_1(\Lambda)} \exp\left(-t (\sqrt{p}|h_0^\Lambda \sqrt{p}) + \sum_{x \in \Lambda} G(p(x)t)\right). \end{aligned} \tag{3.39}$$

The second step is to define a subset $D \subset M_1(\Lambda)$ of relatively uniform probability distributions concentrated on a subvolume of Λ . With the side length L of Λ , we set $1 \leq l \leq L$ and $\Lambda_l := \{x \in \mathbb{Z}^d : |x|_\infty \leq l\}$. The ground state of the discrete Laplacian $h_0^{\Lambda_l}$ restricted to Λ_l with Dirichlet boundary conditions is given by

$$\phi_l : \Lambda_l \rightarrow [0, \infty), \tag{3.40}$$

$$\phi_l = \prod_{j=1}^d \left(\frac{2}{l+1} \right)^{1/2} \sin\left(\frac{x_j \pi}{l+1}\right) \tag{3.41}$$

and the corresponding principal eigenvalue of $h_0^{\Lambda_l}$ is

$$E_1(h_0^{\Lambda_l}) = 2d \left(1 - \cos\left(\frac{\pi}{l+1}\right) \right). \tag{3.42}$$

The subset $D \subset M_1(\Lambda)$ is then defined by

$$D := \{\phi_l^2 : 1 \leq l \leq L\}. \tag{3.43}$$

Restricting the estimate (3.39) to D , we get

$$\begin{aligned} \mathbb{E} (\exp(-t E_1(h_\omega^\Lambda))) &\geq \exp\left(G(t) + t \sup_{p \in M_1(\Lambda)} \left(-(\sqrt{p}|h_0^\Lambda \sqrt{p}) + \sum_{x \in \Lambda} p(x)S(p(x), t) \right)\right) \\ &\geq \exp\left(G(t) + t \sup_{p \in D} \left(-(\sqrt{p}|h_0^\Lambda \sqrt{p}) + \sum_{x \in \Lambda} p(x)S(p(x), t) \right)\right). \end{aligned}$$

Using the definition of $S(\lambda, t)$, the convexity of the cumulant generating function $G(t)$ and the Jensen inequality, we can estimate for $p \in M_1(\Lambda_l) \subset M_1(\Lambda)$,

$$\begin{aligned} \sum_{x \in \Lambda} p(x) S(p(x), t) &= \sum_{x \in \Lambda_l} p(x) \left(\frac{G(p(x)t)}{p(x)t} - \frac{G(t)}{t} \right) \\ &= t^{-1} l^d \left(l^{-d} \sum_{x \in \Lambda_l} G(p(x)t) \right) - t^{-1} G(t) \\ &\geq t^{-1} l^d G \left(l^{-d} t \sum_{x \in \Lambda_l} p(x) \right) - t^{-1} G(t) \\ &= \frac{G(l^{-d}t)}{l^{-d}t} - \frac{G(t)}{t} \\ &= S(l^{-d}, t). \end{aligned}$$

So the uniform distribution on Λ_l minimizes $\sum_{x \in \Lambda} p(x) S(p(x), t)$ with respect to the variation over $p \in M_1(\Lambda_l)$. We have

$$\mathbb{E} \left(\exp \left(-t E_1 \left(H_\Lambda^D(\omega) \right) \right) \right) \geq \exp \left(G(t) + t \sup_{1 \leq l \leq L} \left(-4d \sin^2 \left(\frac{\pi}{2} \frac{1}{l+1} \right) + S(l^{-d}, t) \right) \right)$$

and the only remaining problem is to maximize

$$\begin{aligned} &-4dt \sin^2 \left(\frac{\pi}{2} \frac{1}{l+1} \right) + t S(l^{-d}, t) \\ &\sim -4dt \sin^2 \left(\frac{\pi}{2} \frac{1}{l+1} \right) + \begin{cases} c(l^{-d\rho} - 1)t^{\rho+1} & \rho > 0 \\ -cdt \log(l) & \rho = 0 \\ c(1 - l^{-d\rho})t^{\rho+1} & -1 < \rho < 0 \\ -cl^d \log(t) & \end{cases} \end{aligned}$$

with respect to l . The exponent of the time t in the scaling expression is responsible for the occurrence of the classical or the quantum regime. In the case $\rho > 0$, the scaling term increases faster in t than the linear time dependence in the diffusion term. Consequently, the maximum will be asymptotically $l = 1$. This corresponds to the classical regime. In the case $-1 < \rho < 0$ as well as in the fourth situation, the scaling term is sublinear and the diffusion term is dominating. Like in Section 3.2, a collective behavior of potential values is necessary and we are in the quantum regime. In the case $\rho = 0$, the diffusion and the scaling term are both linear in time. So the optimal peak size depends strongly on the constants. This is the borderline between the classical and the quantum regime. It corresponds to the double exponential distribution.

The upper bounds are much more complicated. It is not possible to interchange the supremum and the expectation with respect to the random potential. Moreover, we have to estimate (3.38) for all $p \in M_1(\Lambda)$. The first problem is solvable by a variant of the ordinary Laplace method. The second problem is attacked using the convexity of the cumulant generating function $G(t)$ and ideas from spectral geometry. For details, we refer to [125] and [126].

3.6. Internal Band Edges. Lifshitz predicted the ‘‘Lifshitz behavior’’ not only for the bottom of the spectrum but also for other band edges. We refer to this phenomenon as ‘‘internal Lifshitz tails.’’ Internal Lifshitz tails have been proven for

the Anderson model by Mezincescu [128] and Simon [157]. Their proofs apparently cannot be translated to the continuum case. In fact, the band edges of the Anderson model which they can handle are those coming from gaps in $\text{supp } P_0$ together with the boundedness of the kinetic energy. (To be more precise: Since for the Anderson model, $\|h_0\| \leq 4d$ and $h_0 \geq 0$, there are gaps in the spectrum whenever there are gaps in $\text{supp } P_0$ of length exceeding $4d$).

One can also handle the case of a point interaction potential in one dimension, a problem which essentially reduces to a lattice problem. Formally this potential is given by

$$V_\omega = \sum q_i \delta(x - i)$$

where δ is the Dirac-“function.” This potential is also known as the random *Kronig-Penney model*. It turns out in this case that the lower edges do and the upper edges (for $q_i \geq 0$) do not show Lifshitz behavior but polynomial behavior of N as for periodic potentials [79]. This is due to the fact that the upper edges are “stable boundaries” in the sense of [148]. The case of general one-dimensional alloy-type potentials was treated in [130].

The multidimensional case is by far more difficult. The reason is mainly that *periodic* potentials are much less well understood in higher dimensions. For example, it is not true in general that bands are parabolic, as is the case for $d = 1$ and for the ground state band in arbitrary dimension.

The paper [97] marks a breakthrough in this topic. Klopp uses the method of approximation by periodic potentials. Compared to Dirichlet–Neumann bracketing, one loses monotonicity, a property which was very useful above. However, Klopp manages to prove an exponential convergence rate for the periodic approximations.

As mentioned above, not so much is known about the behavior of the band functions (of the periodic operators) at internal band edges. In fact, Klopp has to make assumptions on the behavior of the integrated density of states for the *periodic* operator.

Like Lifshitz tails at the bottom of the spectrum internal Lifshitz tails can be used as an input for a localization proof [175].

We consider an alloy-type potential with a continuous single site potential $f \geq 0$, not identically equal to 0, with decay

$$f(x) \leq C(1 + |x|)^{-(d+2+\epsilon)}. \tag{3.44}$$

The random coupling constants are independent and have a common probability distribution P_0 with $q_{\min} = \inf \text{supp } P_0$. We set $V_{\text{per}} = \sum_{i \in \mathbb{Z}^d} q_{\min} f(x - i)$ and denote the integrated density of states of $H_{\text{per}} = H_0 + V_{\text{per}}$ by $N_{\text{per}}(E)$. Furthermore, we suppose that E_- is a lower band edge of H_{per} , i.e., $E_- \in \sigma(H_{\text{per}})$, but $(E_- - a, E_-) \cap \sigma(H_{\text{per}}) = \emptyset$.

It is reasonable to assume that generically N_{per} behaves like $(E - E_-)^{d/2}$, for $E \searrow E_-$, as it would for a unique parabolic band. In fact, this behavior is known in one dimension and for the bottom of the spectrum in arbitrary dimension. However, it is not clear that this is true in general, even not generically (see, however, [103]).

Thus, we have to *assume* such a behavior of N_{per} :

Assumption: Suppose E_- is a lower band edge of H_{per} . We assume that

$$\lim_{E \searrow E_-} \frac{\ln(N_{\text{per}}(E) - N_{\text{per}}(E_-))}{\ln(E - E_-)} = -\frac{d}{2}. \tag{3.45}$$

Under this assumption, the main result of [97] is:

THEOREM 3.8 (Klopp). *If assumption (3.45) holds, then*

$$\lim_{E \searrow E_-} \frac{\ln \left(-\ln (N(E) - N(E_-)) \right)}{\ln (E - E_-)} = -\frac{d}{2}. \tag{3.46}$$

For the case $d = 2$, one has more information about the periodic operators [105]. In particular, there is always exponential decay of the integrated density of states at band edges. We refer to the review [100] for an introduction and further results.

3.7. Lifshitz Tails for Surface Potentials. In this section we consider surface potentials of the form

$$V_\omega^s(x_1, x_2) = \sum_{i_1 \in \mathbb{Z}^{d_1}} q_{i_1}(\omega) f(x_1 - i_1, x_2) \tag{3.47}$$

and suppose we have some spectrum below 0. This is the case if $q_{\min} = \inf \text{supp } P_0$ is negative enough. Note that for $d_2 \leq 2$, there is negative spectrum as soon as q_{\min} is negative. For $d_2 \geq 3$, there is a threshold $\gamma > 0$ such that the spectrum starts at 0 if $q_{\min} \geq -\gamma$ and there is negative spectrum if $q_{\min} < -\gamma$.

We are going to investigate Lifshitz tails for surface potentials with $E_0 < 0$. Below the bulk spectrum (which starts at 0), the density of surface states is positive, hence a measure. We may therefore define the integrated density of surface states $N_s(E)$ to be the corresponding distribution function.

As before, we decompose the potential into a non-random background potential and a positive random potential

$$\begin{aligned} V_\omega^s(x_1, x_2) &= \sum_{i_1 \in \mathbb{Z}^{d_1}} q_{i_1}(\omega) f(x_1 - i_1, x_2) \\ &= \sum_{i_1 \in \mathbb{Z}^{d_1}} q_{\min} f(x_1 - i_1, x_2) + \sum_{i_1 \in \mathbb{Z}^{d_1}} (q_{i_1} - q_{\min}) f(x_1 - i_1, x_2) \\ &= V_{\text{sp}}^s(x) + \tilde{V}_\omega^s(x). \end{aligned} \tag{3.48}$$

The Neumann–Dirichlet bracketing technique goes through for this case as soon as we have sufficient knowledge about the background operator $H_1 = H_0 + V_{\text{sp}}^s$. Since we want $E_0 < 0$ —which makes the bottom of the spectrum “surface spectrum”—there is no case $V_{\text{sp}}^s = 0$ here. Moreover, the background potential V_{sp}^s is only periodic for the d_1 -directions, but decays perpendicular to them.

The analysis of those partially periodic potentials and the Lifshitz estimates for surface potentials were done in [87] for the continuous case. We assume that

$$P_0([q_{\min}, q_{\min} + \varepsilon]) \geq C \varepsilon^N$$

and

$$0 \leq f(x_1, x_2) \leq f_0 (1 + |x_1|)^{-(d_1+2)}.$$

We also assume that $f(x_1, x_2)$ decays uniformly in x_2 -directions. Then we have:

THEOREM 3.9. *If $E_0 < 0$, then*

$$\lim_{E \searrow E_0} \frac{\ln \left(-\ln(N_s(E)) \right)}{\ln (E - E_0)} = -\frac{d_1}{2}. \tag{3.49}$$

There is also an analogous theorem for long range f . The paper [72] proves Lifshitz tails for surface potentials in the discrete setting by fairly different techniques. This paper also contains an analysis at the energy $E = 0$, i.e., for surface corrections to the bulk Lifshitz tails.

3.8. Lifshitz Tails for Random Landau Hamiltonians. We turn to the density of states for operators of the form

$$H_\omega = H_0(B) + V_\omega$$

with a constant magnetic field $B > 0$ and a non-negative random potential V_ω . (For a careful definition of the density of states and some basics, see [62, 63, 178].)

We discuss the *two-dimensional* case first. The Landau Hamiltonian $H_0(B)$ is given by

$$H_0(B) = \left(i \frac{\partial}{\partial x_1} - \frac{1}{2} B x_2 \right)^2 + \left(i \frac{\partial}{\partial x_2} + \frac{1}{2} B x_1 \right)^2.$$

$H_0(B)$ has a pure point spectrum for $B \neq 0$ and $d = 2$. In fact, the eigenvalues are given by the ‘‘Landau levels’’ $(2n + 1)B$; $n \in \mathbb{N}$ and all Landau levels are infinitely degenerate. One possible choice of the ground state is

$$\psi_0(x) = \frac{B}{\pi} e^{-\frac{B}{2}|x|^2} \tag{3.50}$$

which will play a major role below. For V_ω we take a Poisson potential or an alloy-type potential with $q_{\min} = 0$. In this case, the bottom E_0 of the spectrum of H_ω is given by the lowest Landau level, which is B . In [10] the authors proved for the Poisson model the following result.

THEOREM 3.10. *If $B \neq 0$ and*

$$\frac{C_1}{(1 + |x|)^\alpha} \leq f(x) \leq \frac{C_2}{(1 + |x|)^\alpha}, \tag{3.51}$$

the Lifshitz exponent η for $H_0(B) + V_\omega$ is given by

$$\eta = \frac{2}{2 - \alpha} \quad \left(= \frac{d}{d - \alpha} \right) \tag{3.52}$$

for all $\alpha > d$.

This means that, according to our classification above, we are always in the classical case for $d = 2$ and constant magnetic field.

PROOF. We sketch the lower bound only and restrict ourselves to the alloy-type case. As usual we have to estimate

$$\mathbb{P}(E_1(H_{\omega_{\Lambda_L}}^D) < E_0 + E)$$

from below. This time we have a ground state ψ_0 for $H_0(B)$ which is L^2 . We modify ψ_0 near the boundary of Λ_L to make it satisfy Dirichlet boundary conditions. Due to the (super-)exponential decay of ψ_0 , the error we make is of the order (at most)

$e^{-C_0 L^2}$. Thus

$$\begin{aligned} \mathbb{P}(E_1(H_{\omega \Lambda_L}^D) < E_0 + E) &\geq \mathbb{P}\left(\int V_\omega(x)|\psi_0|^2 dx < E - e^{-C_0 L^2}\right) \\ &\geq \mathbb{P}\left(\int V_\omega(x) dx < C_1 E - e^{-C_2 L^2}\right) \\ &= (*). \end{aligned}$$

At this point we can literally repeat the estimates in the proof of Theorem 3.5 and obtain in analogy to (3.25)

$$(*) \geq \mathbb{P}\left(\frac{1}{|\Lambda_{2L}|} \sum_{|i| \leq 2L} q_i \leq C_1 E - C_2 e^{-cL^2} - C_3 L^{-(\alpha-d)}\right). \quad (3.53)$$

The only difference to the previous case is the error term due to the kinetic energy. In Theorem 3.5 it was of the order L^{-2} causing the different behavior for $\alpha \geq d+2$ and $\alpha < d$. In (3.53) the error term is exponentially small, thus being negligible with respect to the potential term $L^{-(\alpha-d)}$ for all α . Consequently, we may choose $E \sim L^{\frac{1}{\alpha-d}}$. By a large deviation estimate, we obtain

$$\begin{aligned} (3.53) &\geq \mathbb{P}(q_i \leq C' E)^{|\Lambda_{2L}|} \\ &\geq M e^{-\tilde{C} L^d} \\ &= M e^{-\tilde{C} E^{-\frac{d}{\alpha-d}}}. \quad \square \end{aligned}$$

Theorem 3.10 implies that for compactly supported f , the integrated density of states $N(E)$ decays subexponentially.

Erdős [38, 39] proved that it decays, in fact, polynomially. Erdős' proof is based on a careful estimate of the Laplace transform of N . It uses an analog of the Feynman–Kac formula for magnetic Schrödinger operators, the Feynman–Kac–Ito formula (see [154]). There are a couple of complications in the Feynman–Kac expression of \tilde{N} due to the magnetic field. The most serious one is the fact that the integrand is no longer positive but rather oscillating.

Erdős' proof is done for the Poisson model. Recently, Klopp and Raikov [102] found a completely different approach based on the approximation by periodic potentials. Their proof works for alloy-type potentials. Moreover, the latter paper contains results for internal Lifshitz tails as well as for the case of an additional periodic background potential.

We state Erdős' result:

THEOREM 3.11. *If $H_\omega = H_0(B) + V_\omega$ with a Poisson random potential V_ω and if the single site potential $f \geq 0$ has compact support, then*

$$\lim_{E \searrow E_0} \frac{\ln N(E)}{\ln E} = \frac{\pi}{B}. \quad (3.54)$$

The reasoning we gave in the proof of Theorem 3.10 can be used to prove a lower bound for Theorem 3.11 as well. In fact, the bound suggests that the borderline between the two kinds of behavior is given by Gaussian single site potentials. This was actually proved in [64] and [39].

We have seen in this section that random Landau Hamiltonians may have unusual Lifshitz behavior compared to the nonmagnetic case. This originates in

the fact that the lowest Landau band is flat, collapsing into one point. It is well known that this never happens for ($B = 0$ and) periodic *scalar* potential (see [150]).

It is reasonable to expect that the flatness of the ground state Landau band is removed by an additional periodic scalar potential, at least generically. If this is true, we would certainly expect “normal” Lifshitz behavior for such operators, including those with alloy-type potentials and $q_{\min} \neq 0$. There are results in this direction in [102].

We turn to the *three-dimensional* case with homogeneous magnetic field. We consider the Hamiltonian

$$H_0(B) = \left(i \frac{\partial}{\partial x_1} - \frac{1}{2} B x_2 \right)^2 + \left(i \frac{\partial}{\partial x_2} + \frac{1}{2} B x_1 \right)^2 - \frac{\partial^2}{\partial x_3^2}.$$

In $d = 3$, the magnetic field itself introduces an anisotropy. The two space dimensions perpendicular to the magnetic field (x_1, x_2) will be denoted by x_\perp , the direction x_3 parallel to the field by x_\parallel . If we add an anisotropic Anderson (or Poisson) potential with a single site potential f obeying

$$\frac{a}{|x_\perp|^{\alpha_1} + |x_\parallel|^{\alpha_2}} \leq f(x) \leq \frac{b}{|x_\perp|^{\alpha_1} + |x_\parallel|^{\alpha_2}}, \tag{3.55}$$

we will have Lifshitz behavior that resembles the results of Section 3.4, except that in the \perp -direction the behavior is always *classical*.

Following the conventions of Section 3.4, we define

$$\gamma_1 = \frac{2}{\alpha_1}, \quad \gamma_2 = \frac{1}{\alpha_2} \quad \text{and} \quad \gamma = \gamma_1 + \gamma_2. \tag{3.56}$$

If f has compact support in x_\perp -direction, we set $\gamma_1 = 0$. To have the potential well defined, we need $\gamma < 1$.

THEOREM 3.12. *The Lifshitz exponent η for $H_0(B) + V_\omega$ in dimension $d = 3$ is given by*

$$\eta = \frac{\gamma_1}{1 - \gamma} + \max \left\{ \frac{1}{2}, \frac{\gamma_2}{1 - \gamma} \right\}. \tag{3.57}$$

This theorem was proven for f with compact support in [178]. The case of $\gamma_2 > \frac{1}{2}$ is considered in [60] and the mixed classical-quantum case is taken from [73].

We remark that the lower bound given above (in connection with Theorem 3.10) can be used in this case as well. As a test function in this bound, we use

$$\Psi_0(x_\perp, x_\parallel) = \psi_0(x_\perp) \phi_0(x_\parallel) \tag{3.58}$$

where ψ_0 is the (2-d) ground state (3.50) and $\phi_0(x_\parallel) = L^{-1/2}$ is the \parallel -ground state. As before Ψ_0 has to be cut down to zero near the boundary of the cube Λ_L . This gives an error term of the order e^{-L^2} in the \perp -direction and an error term of the order L^{-2} in the \parallel -direction.

For further references about magnetic Lifshitz tails, we refer to [178] and to [116].

There are also results on the integrated density of states and on Lifshitz tails for *random* magnetic fields. We refer to the works of Ueki and Nakamura [172, 173, 174, 138, 139] and the references given there.

3.9. Percolation Models. Recently the integrated density of states for Laplacians on percolation graphs has been investigated. We consider bond percolation on \mathbb{Z}^d with $d \geq 2$. This means that we remove bonds in the graph \mathbb{Z}^d independently with probability $1 - p$ for $0 < p < 1$. The resulting random graph is called the *percolation graph* and is denoted by \mathcal{G}_p or simply by \mathcal{G} .

For p small ($p < p_c$), all connected components (“clusters”) of \mathcal{G}_p are finite almost surely. For $p > p_c$ there is a unique infinite cluster (see, e.g., [54]).

We consider the Laplacian on this graph which is a random operator, due to the randomness of the underlying graph. For a general discussion of percolation Hamiltonians, we refer to [176].

Actually, there are various Laplacians on \mathcal{G} due to different boundary conditions. To define these operators, we start with the adjacency operator $A_{\mathcal{G}}$ defined by the matrix elements with $A_{\mathcal{G}}(i, j) = 1$ if $|i - j| = 1$, $i, j \in \mathcal{G}$ and $A_{\mathcal{G}} = 0$ otherwise. For any $i \in \mathcal{G}$, we let $d_{\mathcal{G}}(i)$ denote the number of sites in \mathcal{G} to which i is connected. We denote by $D_{\mathcal{G}}$ the diagonal matrix with entries $d_{\mathcal{G}}(i)$ on the diagonal. The Neumann Laplacian $L_{\mathcal{G}}^N$ is defined by

$$L_{\mathcal{G}}^N = D_{\mathcal{G}} - A_{\mathcal{G}}. \quad (3.59)$$

The Neumann Laplacian is the intrinsic Laplacian of the graph \mathcal{G} . In a sense it “ignores” the embedding of \mathcal{G} into \mathbb{Z}^d . The Dirichlet Laplacian $L_{\mathcal{G}}^D$ is defined by

$$L_{\mathcal{G}}^D = D_{\mathcal{G}} + 2(2d - D_{\mathcal{G}}) - A_{\mathcal{G}}. \quad (3.60)$$

On its diagonal this operator counts the connections to neighboring sites in \mathcal{G} once and the (lost) connections to sites in $\mathbb{Z}^d \setminus \mathcal{G}$ twice. It was an observation of Simon [156] that the above operator is a good analog of the Dirichlet Laplacian for subgraphs of \mathbb{Z}^d . See also [78] or [70] for a discussion of boundary conditions for discrete Laplacians.

Both for $L_{\mathcal{G}}^N$ and for $L_{\mathcal{G}}^D$, a integrated density of states can be defined in analogy to Theorem 2.2. We call them $N_{\mathcal{G}}^N(E)$ and $N_{\mathcal{G}}^D(E)$, respectively.

Since there are infinitely many clusters containing just one point, the Neumann Laplacian $L_{\mathcal{G}}^N$ has an eigenvalue of infinite multiplicity at the bottom $E = 0$ of its spectrum. This causes the integrated density of states to jump at that energy. Hence

$$\nu_{\mathcal{G}}^N(\{0\}) = N_{\mathcal{G}}^N(0) - N_{\mathcal{G}}^N(0-) = N_{\mathcal{G}}^N(0) \neq 0.$$

THEOREM 3.13. (1) *If $p < p_c$, we have*

$$\lim_{E \searrow 0} \frac{\ln \left(-\ln \left(N_{\mathcal{G}}^N(E) - N_{\mathcal{G}}^N(0) \right) \right)}{\ln(E)} = -\frac{1}{2}. \quad (3.61)$$

(2) *If $p > p_c$, we have*

$$\lim_{E \searrow 0} \frac{\ln \left(N_{\mathcal{G}}^N(E) - N_{\mathcal{G}}^N(0) \right)}{\ln(E)} = \frac{d}{2}. \quad (3.62)$$

(3) *For arbitrary $0 < p < 1$, we have*

$$\lim_{E \searrow 0} \frac{\ln \left(-\ln N_{\mathcal{G}}^D(E) \right)}{\ln(E)} = -\frac{d}{2}. \quad (3.63)$$

Part (1) of this theorem, as well as part (3) for $p < p_c$, was proven in [78]; part (2) and part (3) are taken from [133].

The perhaps surprising behavior of $N_G^N(E)$ for $p < p_c$ is due to the fact that long one-dimensional chains dominate. They lead to small eigenvalues for the Neumann Laplacian since they “don’t know” they are actually in \mathbb{Z}^d . For the Dirichlet Laplacians these one-dimensional chains have rather high eigenvalues due to the additional diagonal term. In fact, they dominate the scene at $E = 4d$, the top of the spectrum. This can be seen by a symmetry argument.

Part (2) of the theorem comes from the fact that for $p > p_c$ the leading behavior of N_G^N comes from the infinite cluster. In a sense, the infinite cluster looks rather d -dimensional. The proof in [133] of this part relies on a celebrated paper by Barlow [5].

4. Regularity of the Integrated Density of States

4.1. Introduction. In this final section we discuss regularity properties of the integrated density of states. So far we have seen that the density of states measure ν is a positive Borel measure with a distribution function $N(E)$.

Of course, the name integrated *density* of states suggests that ν (resp. N) should have a density $n(E)$ in the sense

$$N(E) = \int_{E' \leq E} n(E') dE', \tag{4.1}$$

$$\nu([a, b]) = \int_{[a, b]} n(E') dE'. \tag{4.2}$$

We will, indeed, prove this for the *Anderson model* under certain assumptions on P_0 , the probability distribution of the random potential. The continuous case is more complicated. For this case we will state only some key results and refer the reader to the literature. A good account is the review article by Veselic [177]. However, we remark that there are important developments after this survey; we mention especially the recent preprint [20].

We will see that the integrated density of states does *not* always have a density. We will therefore also look at weaker regularity properties of N . We are especially interested in the question whether the function N is continuous, which is the same as $\nu(\{E\}) = 0$ for all E . It turns out that this is always the case for the Anderson model. However, for the continuous case, no such result is known.

4.2. Continuity of the Integrated Density of States. The results of this section are valid for general ergodic operators on $\ell^2(\mathbb{Z}^d)$. Analogous results for \mathbb{R}^d have not been proven so far. The most general result in this context is the following theorem proved by Craig and Simon [25] in a somewhat stronger form (see Theorem 4.3).

THEOREM 4.1 ([25, 30]). *Let $\{V_\omega(n) : n \in \mathbb{Z}^d\}$ be an ergodic stationary real valued random potential satisfying*

$$\mathbb{E}[\log(1 + |V_\omega(0)|)] < \infty. \tag{4.3}$$

Then the integrated density of states of the Anderson operator $h_\omega = h_0 + V_\omega$ is a continuous function, i.e., $\nu(\{\lambda\}) = 0$ for all $\lambda \in \mathbb{R}$.

This result can be proven with an elementary argument by Delyon and Souillard [30, 27] using a kind of “unique continuation” property of discrete Schrödinger operators. We use this idea in the proof of the following lemma to show that no eigenspace can be sufficiently degenerated to produce a jump of the integrated density of states.

Let us denote by $\mu_H(\cdot)$ the projection-valued spectral measure of H , i.e., $\mu_H(A) = \chi_A(H)$; in particular, $\mu_H(\{\lambda\})$ is the projector onto the eigenspace of H with respect to λ .

LEMMA 4.2. *We have*

$$\dim(\chi_{\Lambda_L} \text{Ran } \mu_{h_\omega}(\{\lambda\})) \leq C L^{d-1}.$$

PROOF. The set

$$\Lambda_L^{(2)} = \left\{ i \in \Lambda_L \mid \max_{\nu=1, \dots, d} |i_\nu| = L \text{ or } \max_{\nu=1, \dots, d} |i_\nu| = L - 1 \right\}$$

consists of the two outermost layers of Λ_L . A solution u of $h_\omega u = \lambda u$ is uniquely determined inside Λ_L by its values on $\Lambda_L^{(2)}$. So, the dimension of $\chi_{\Lambda_L}(\text{Ran } \mu_{h_\omega}(\{\lambda\}))$ is at most the number of points in $\Lambda_L^{(2)}$. \square

With the lemma, we can prove the theorem.

PROOF OF THE THEOREM. By Proposition 2.1 and Theorem 2.2, we have

$$\nu(\{\lambda\}) = \lim_{L \rightarrow \infty} \frac{1}{(2L+1)^d} \text{tr}(\chi_{\Lambda_L} \mu_{h_\omega}(\{\lambda\})). \quad (4.4)$$

If f_i is an orthonormal basis of $\chi_{\Lambda_L}(\text{Ran } \mu_{h_\omega}(\{\lambda\}))$ and g_i an orthonormal basis of $(\chi_{\Lambda_L}(\text{Ran } \mu_{h_\omega}(\{\lambda\})))^\perp$ we have, noting that $\chi_{\Lambda_L}(\text{Ran } \mu_{h_\omega}(\{\lambda\}))$ is finite-dimensional,

$$\begin{aligned} \text{tr}(\chi_{\Lambda_L} \mu_{h_\omega}(\{\lambda\})) &= \sum_{i \in I} \langle f_i, \chi_{\Lambda_L} \mu_{h_\omega}(\{\lambda\}) f_i \rangle \\ &\leq \dim(\chi_{\Lambda_L}(\text{Ran } \mu_{h_\omega}(\{\lambda\}))) \end{aligned} \quad (4.5)$$

$$\leq C L^{d-1} \quad (4.6)$$

hence (4.4) converges to zero and $\nu(\{\lambda\}) = 0$. \square

In [25] Craig and Simon prove a stronger result than just continuity:

THEOREM 4.3 (Craig–Simon). *Under the assumptions of Theorem 4.1, the integrated density of states is locally log-Hölder continuous, e.g., for any positive r there exists a finite constant C_r such that*

$$|N(\lambda) - N(\lambda')| \leq C_r \frac{1}{|\log |\lambda - \lambda'| |} \quad (4.7)$$

for $|\lambda| \leq r$ and $|\lambda - \lambda'| \leq 1$.

The theorem of Craig and Simon is based on the Thouless formula for a strip in \mathbb{Z}^d . For the Thouless formula in the one-dimensional setting, we refer to the next subsection.

4.3. Regularity of the DOS in Dimension One. There are very powerful techniques to study the Schrödinger equation and its discrete analog in dimension *one* which are unfortunately restricted to dimension one exclusively. Such techniques have been successfully applied to random operators as well to investigate the integrated density of states.

At the heart of most of these techniques lies the reformulation of the eigenvalue equation of a second-order equation into an initial value problem for a system of first-order equations. Specifically, let us look at the eigenvalue equation

$$h_\omega = -u(n + 1) - u(n - 1) + (V_\omega(n) - E)u(n) = 0. \tag{4.8}$$

We define

$$U(n) = \begin{pmatrix} u(n + 1) \\ u(n) \end{pmatrix}$$

and

$$A_n(E) = \begin{pmatrix} V_\omega(n) - E & -1 \\ 1 & 0 \end{pmatrix}.$$

The function $u(n)$ is a solution of (4.8) if and only if

$$U(n + 1) = A_{n+1}(E)U(n)$$

for all $n \in \mathbb{Z}$. With the transfer matrix

$$\Phi_n(E) := \prod_{i=1}^n A_i(E)$$

the solution of (4.8) to the right initial condition

$$U(0) = \begin{pmatrix} u(1) \\ u(0) \end{pmatrix}$$

can be expressed by

$$U(n) = \Phi_n(E)U(0).$$

Similarly, it is possible to define the solution to the left. The spectral theory of the operator h_ω is encoded in the matrices $A_n(E)$ (or $\Phi_n(E)$). Note that these matrices belong to the group $SL(2, \mathbb{R})$. Thus it should not come as a surprise that harmonic analysis on $SL(2, \mathbb{R})$, in explicit or implicit form, plays a major role in the analysis of h_ω .

The asymptotic behavior of the eigensolutions of $h = h_0 + V$ is reflected by the Lyapunov exponent given by

$$\gamma(E) := \lim_{N \rightarrow \pm\infty} \frac{1}{N} \ln \|\Phi_N(E)\|. \tag{4.9}$$

This definition of the Lyapunov exponent is well defined by Fürstenberg’s theorem ([46], see also [27]). The link between the Lyapunov exponent and the integrated density of states is expressed in the Thouless formula

$$\gamma(E) = \int \log |E - E'| dN(E'). \tag{4.10}$$

This formula from physics ([56, 171]) was made rigorous in [3]. A simplified proof can be found in [26] or [27].

The Thouless formula and the obvious fact that $\gamma(E)$ is non-negative imply that the integrated density of states N is log-Hölder continuous (in one dimension). This was observed by Craig and Simon in [26]. The proof by the same authors

of the multidimensional analog uses a version of the Thouless formula for strips in higher dimensions [25].

As mentioned in the previous section, these results hold for general ergodic potentials on \mathbb{Z}^d , they are (at least) close to optimal in this generality (see also our discussion below).

However, if we assume that the $V_\omega(n)$ are independent (and identically distributed) much more can be said about the regularity of the integrated density of states. In the next section, we discuss Lifshitz continuity in the multidimensional case under assumptions on the distribution P_0 of the random variable $V_\omega(0)$. In the one-dimensional case we discuss here, we have a fairly complete picture about the regularity of N . For simplicity we assume that $\text{supp } P_0$ is compact.

DEFINITION 4.4. *We call a function f on \mathbb{R} Hölder continuous of order α if*

$$|f(x) - f(y)| \leq C |x - y|^\alpha.$$

THEOREM 4.5. *Let $h = h_0 + V$ be a discrete random Schrödinger operator in dimension one. Assume V is a sequence of independent, identically distributed random variables, such that their distribution P_0 has compact support. Then the integrated density of states N of h is Hölder continuous of some order $\alpha > 0$.*

This theorem is due to [114]. The simple case of the discrete Laplacian h_0 (i.e., $V = 0$) shows that one cannot expect more than Hölder continuity in general. For h_0 the integrated density of states N is merely Hölder continuous of order $\frac{1}{2}$ at the band edges. The derivative of N , which is called the *density of states*, diverges at $E = \pm 2$. This behavior is known as van Hove singularities. The Lifshitz behavior of the integrated density of states seems to suggest that N is smoother at the band edges for a truly random V .

One can expect that more regularity for P_0 implies more regularity for N . In fact, Simon and Taylor [161] proved that already a little regularity of P_0 implies that N is C^∞ . To state their result, we need the following definition of a Sobolev space:

DEFINITION 4.6. *We say that a function $f \in L^p$ belongs to L^p_α if the Fourier transform \hat{f} satisfies: There is a function $g \in L^p$ such that $\hat{g}(k) = (1 + |k|^2)^{\alpha/2} \hat{f}(k)$ is an L^p -function.*

It is not hard to see that the characteristic function χ_I of a finite interval I is in L^1_α for $\alpha < \frac{1}{2}$ (see [161]).

THEOREM 4.7. *Let $h = h_0 + V$ be a discrete random Schrödinger operator in dimension one. Assume V is a sequence of independent, identically distributed random variables. If the distribution P_0 of $V_\omega(0)$ has a density g with compact support and such that $g \in L^1_\alpha$ for some $\alpha > 0$, then $N \in C^\infty$.*

If we do not assume any regularity of P_0 , Theorem 4.5 is the best result we can hope for. To see this, let us look at a Bernoulli distribution for P_0 . We set

$$P_0 = p \delta_a + (1 - p) \delta_b. \tag{4.11}$$

Halperin [55] proved, with some points of rigor clarified by Simon and Taylor [161], that:

THEOREM 4.8. *Assume that the $V_\omega(n)$ are independent with identical distributions P_0 as in (4.11) with $0 < p \leq \frac{1}{2}$. Then the integrated density of states is not Hölder continuous of any order α larger than*

$$\alpha_0 = \frac{2 |\log(1 - p)|}{\operatorname{arccosh}(1 + \frac{1}{2}|b - a|)}.$$

If N is Lifshitz continuous (Hölder continuous of order $\alpha = 1$), then N has a bounded density n , i.e., $N(E) = \int_{-\infty}^E n(\lambda) d\lambda$ and vice versa. If N is Hölder continuous of (strict) order $\alpha < 1$, then N may still have a density which then has to be unbounded. So, Theorem 4.8 does not rule out that the density of states measure ν is absolutely continuous. However, it is true that ν has a singular continuous component if $|b - a|$ is large.

THEOREM 4.9. *If P_0 is Bernoulli (as in (4.11)) and $0 < p < 1$, then the density of states measure ν has a singular continuous component if $|b - a|$ is large.*

This theorem was proved in [14] following ideas from [161]. The paper [14] contains the first proof of Anderson localization for the one-dimensional Bernoulli model. The proof of a singular continuous component of N is based on this knowledge. The article [122] continues these investigations and proves that for $|b - a|$ large, the density of states measure is even purely singular continuous.

For further results concerning the regularity of the DOS of random Schrödinger operators and its discrete analog in dimension one see [12, 28, 13, 88, 90, 91, 92, 93, 94, 120, 153].

As the operators with i.i.d. potentials, *almost periodic* operators belong to the class of ergodic operators. In a sense these two types of operators form the extreme cases within the class of ergodic operators. There are recent results on regularity of the integrated density of states for one-dimensional almost periodic operators by Goldstein and Schlag [51, 53]. We refer to the review [52] for details.

4.4. The Wegner Estimates: Discrete Case. One motivation in physics to study the integrated density of states was the hope to use it as an indicator for different spectral types. The aim was to distinguish pure point spectrum and continuous spectrum at the mobility edge. In some sense the conjectures discussed were inconsistent. Some expected a divergent density of states at the mobility edge, while others assumed a vanishing density of states.

In 1981 Wegner [179] put an end to this discussion by proving upper and lower bounds of the density of states in the discrete setting of the Anderson model. These estimates imply the density of states neither vanishes nor explodes at a mobility edge (or anywhere else).

Wegner’s result (more precisely, his upper bound) soon became a corner piece in the proofs of Anderson localization by the multiscale analysis method and it still is. To formulate Wegner’s result, we introduce boundary conditions on the lattice. The simplest boundary condition on the lattice is defining h_ω^Λ through its matrix elements:

$$h_\omega^\Lambda(i, j) = h_\omega(i, j) \tag{4.12}$$

whenever both i and j belong to Λ . For a discussion of boundary condition on ℓ^2 , see Simon [156] or the review [70].

We define

$$N_\Lambda(E) = \#\{n \mid E_n(h_\omega^\Lambda) \leq E\} = \operatorname{tr} (P_{(-\infty, E]}(h_\omega^\Lambda)). \tag{4.13}$$

THEOREM 4.10 (Wegner-estimate). *Suppose the measure P_0 has a bounded density g , (i.e., $P_0(A) = \int_A g(\lambda)d\lambda$, $\|g\|_\infty < \infty$), then*

$$\mathbb{E}(N_\Lambda(E + \varepsilon) - N_\Lambda(E - \varepsilon)) \leq 2 \|g\|_\infty |\Lambda| \varepsilon. \tag{4.14}$$

REMARK 4.11. *The assumption that P_0 has a density cannot be dropped. We have seen at the end of Section 4.3 that the integrated density of states has a singular continuous component if P_0 is a (certain) Bernoulli distribution.*

Before we discuss this estimate, we note two important consequences. The first concerns the regularity of the density of states.

COROLLARY 4.12. *Under the assumption of Theorem 4.10, the integrated density of states is absolutely continuous with a bounded density $n(E)$.*

PROOF.

$$\begin{aligned} N(E + \varepsilon) - N(E - \varepsilon) &= \lim_{|\Lambda| \rightarrow \infty} \frac{1}{|\Lambda|} \mathbb{E}(N_\Lambda(E + \varepsilon) - N_\Lambda(E - \varepsilon)) \\ &\leq C\varepsilon \quad \text{by Theorem 4.10.} \end{aligned} \quad \square$$

Thus $N(E) = \int_{-\infty}^E n(\lambda) d\lambda$. We call $n(\lambda)$ the *density of states*. Sometimes, N also is called the density of states which, we admit, is an abuse of language. The second consequence of Theorem 4.10 is a key ingredient in proving Anderson localization.

COROLLARY 4.13. *Under the assumptions of Theorem 4.10, we have for any E and Λ ,*

$$\mathbb{P}(\text{dist}(E, \sigma(h_\omega^\Lambda)) < \varepsilon) \leq C\varepsilon|\Lambda|. \tag{4.15}$$

PROOF. By the Chebyshev inequality, we get

$$\begin{aligned} \mathbb{P}(\text{dist}(E, \sigma(h_\omega^\Lambda)) < \varepsilon) &= \mathbb{P}(N_\Lambda(E + \varepsilon) - N_\Lambda(E - \varepsilon) > 1) \\ &\leq \mathbb{E}(N_\Lambda(E + \varepsilon) - N_\Lambda(E - \varepsilon)) \\ &\leq C\varepsilon|\Lambda| \quad \text{by Theorem 4.10.} \end{aligned} \quad \square$$

The first step in the proof of Theorem 4.10 is to average the eigenvalues inside the interval $(E - \varepsilon, E + \varepsilon)$ with respect to the random potential. To do this we consider the eigenvalues $E_n(h_\omega^\Lambda)$ as functions of the arguments $V_i = V_\omega(i)$ with $i \in \Lambda$, i.e., $E_n(h_\omega^\Lambda) = E_n(V_i, i \in \Lambda)$. The resulting estimate is summarized in the following lemma.

LEMMA 4.14. *With $\varepsilon > 0$ let ϱ be a non-decreasing C^∞ -function with $\varrho(\lambda) = 1$ for $\lambda > \varepsilon$, $\varrho(\lambda) = 0$ for $\lambda < -\varepsilon$ and $0 \leq \varrho(\lambda) \leq 1$. Then*

$$N(h_\omega^\Lambda, E + \varepsilon) - N(h_\omega^\Lambda, E - \varepsilon) \leq \sum_{j \in \Lambda} \int_{E-2\varepsilon}^{E+2\varepsilon} \frac{\partial}{\partial V_j} \text{tr} \varrho(h_\omega^\Lambda(\{V_i\}) - \eta) d\eta.$$

Now we are in a position to interchange the expectation of the random potential and the energy integral. Since the random variables $V_\omega(i)$ are independent and have the common distribution $dP_0(V_i)$, the expectation \mathbb{E} is just integration with respect

to the product of these distributions. Hence

$$\begin{aligned}
 & \mathbb{E}(N_\Lambda(E + \varepsilon) - N_\Lambda(E - \varepsilon)) \\
 & \leq \mathbb{E}\left(\sum_{j \in \Lambda} \int_{E-2\varepsilon}^{E+2\varepsilon} \frac{\partial}{\partial V_j} \text{tr} \left(\varrho(h_\omega^\Lambda(\{V_i\}) - \eta) d\eta \right)\right) \\
 & = \int_{E-2\varepsilon}^{E+2\varepsilon} d\eta \sum_{j \in \Lambda} \prod_{i \in \Lambda} \int_{\mathbb{R}} g(V_i) dV_i \frac{\partial}{\partial V_j} \text{tr} \varrho(h_\omega^\Lambda(\{V_i\}) - \eta) \\
 & \leq \|g\|_\infty \int_{E-2\varepsilon}^{E+2\varepsilon} d\eta \sum_{j \in \Lambda} \prod_{i \in \Lambda, i \neq j} \int_{\mathbb{R}} g(V_i) dV_i \int_a^b dV_j \frac{\partial}{\partial V_j} \text{tr} \varrho(h_\omega^\Lambda(\{V_i\}) - \eta) \\
 & \leq \int_{E-2\varepsilon}^{E+2\varepsilon} d\eta \sum_{j \in \Lambda} \prod_{i \in \Lambda, i \neq j} \int g(V_i) dV_i \\
 & \quad \left\{ \text{tr} \varrho(h_\omega^\Lambda(\{V_i\}^{V_j=b}) - \eta) - \text{tr} \varrho(h_\omega^\Lambda(\{V_i\}^{V_j=a}) - \eta) \right\}.
 \end{aligned}$$

Above, a, b are such that $\text{supp} g \subset [a, b]$. The notation $\{V_i\}^{V_j=b}$ means the family $\{V_i\}_{i \in \Lambda}$ with $\tilde{V}_i = V_i$ for $i \neq j$ and $\tilde{V}_j = b$. The problem is now to estimate the trace difference. In the discrete context of the Anderson model, the variation of a potential value at one site is a rank one perturbation.

LEMMA 4.15. *Let A be a self-adjoint operator bounded below with purely discrete spectrum $E_0 \leq E_1 \leq \dots$ (where the eigenvalues are repeated according to multiplicity). If B is a symmetric positive rank one operator, then $\tilde{A} = A + B$ has eigenvalue \tilde{E}_n with $E_n \leq \tilde{E}_n \leq E_{n+1}$.*

Given the lemma, we now continue the proof of the theorem. We set $A = H_\Lambda(\{V_i\}^{V_j=a})$ and $\tilde{A} = H_\Lambda(\{V_i\}^{V_j=b})$. Obviously, their difference is a (positive) rank one operator

$$\begin{aligned}
 \text{tr} \varrho(\tilde{A} - \eta) - \text{tr} \varrho(A - \eta) &= \sum (\varrho(\tilde{E}_n - \eta) - \varrho(E_n - \eta)) \\
 &\leq \sum (\varrho(E_{n+1} - \eta) - \varrho(E_n - \eta)) \\
 &\leq \sup_{\lambda, \mu} \varrho(\lambda) - \varrho(\mu) \\
 &= 1.
 \end{aligned} \tag{4.16}$$

We conclude the proof of Wegner’s estimate by proving Lemma 4.15:

PROOF. $B = c|h\rangle\langle h|$ with $c \geq 0$, i.e., $B\varphi = c\langle h, \varphi \rangle h$ for some h . By the min-max principle (see [150]),

$$\begin{aligned}
 \tilde{E}_n &= \sup_{\psi_1, \dots, \psi_{n-1}} \inf_{\substack{\varphi \perp \psi_1, \dots, \psi_{n-1} \\ \|\varphi\|=1}} \langle \varphi, A\varphi \rangle + c|\langle \varphi, h \rangle|^2 \\
 &\leq \sup_{\psi_1, \dots, \psi_{n-1}} \inf_{\substack{\varphi \perp \psi_1, \dots, \psi_{n-1}, h \\ \|\varphi\|=1}} \langle \varphi, A\varphi \rangle \\
 &\leq \sup_{\psi_1, \dots, \psi_{n-1}, \psi_n} \inf_{\substack{\varphi \perp \psi_1, \dots, \psi_{n-1}, \psi_n \\ \|\varphi\|=1}} \langle \varphi, A\varphi \rangle \\
 &= E_{n+1}.
 \end{aligned} \tag{4.17}$$

□

Wegner's estimate is intimately connected with a method called "spectral averaging." Roughly speaking, spectral averaging says that taking expectation with respect to random parameters will make the spectral measure absolutely continuous. Here is a typical example which comes from the theory of rank one perturbations, see Simon [158] and references given there.

To formulate this result, we take a bounded operator h on $\ell^2(\mathbb{Z}^d)$ and set $h_\alpha = h + \alpha \delta_j$ for any fixed $j \in \mathbb{Z}^d$. Note that the multiplication operator δ_j is a rank one perturbation of h . We denote by m_α the (projection valued) spectral measure of h_α and set $\mu_\alpha(A) = \langle \delta_j, m_\alpha(A) \delta_j \rangle$. We obtain

THEOREM 4.16 (Spectral averaging).

$$\int d\mu_\alpha(E) d\alpha = dE, \quad (4.18)$$

i. e.,

$$\int \left(\int f(E) d\mu_\alpha(E) \right) d\alpha = \int f(E) dE$$

for all integrable f .

Wegner's estimate follows from the spectral averaging result (see [158]).

Spectral averaging was introduced in the theory of random operators by Kotani [109] who used it to prove Anderson localization. He used random boundary conditions in dimension one, but soon Kotani's trick was also used to prove localization in higher-dimensional systems. However, Kotani was not the first to prove a spectral averaging formula. Such a formula was known in the Russian literature earlier, e.g., to Javrjan [65] who proved it for "random" boundary conditions. Spectral averaging also plays a prominent role for continuous Schrödinger operators; see [21] and references given there.

4.5. Regularity in the Continuous Case. To prove a Wegner estimate for the continuous case is considerably harder than for the discrete case. In fact, only the alloy-type model and a few other cases can be treated so far (for these cases see [21] and [41]). For the alloy-type model, one can carry over Wegner's original proof ([69]).

However, the finite rank estimate (Lemma 4.15) cannot be transferred directly to the continuum (see [67]). Thus a direct analog of Wegner's approach only gives

$$\mathbb{E}(N_\Lambda(E + \epsilon) - N_\Lambda(E - \epsilon)) \leq C|\Lambda|^2 \epsilon. \quad (4.19)$$

The estimate (4.19) obviously gives *no* information on the regularity of N . However, it suffices as input to multiscale analysis to prove Anderson localization (this was observed in [121]).

The first to prove a Wegner bound with the "correct" volume dependence were Kotani and Simon in [110]. They required the single site potential f to be the characteristic function of the unit cube. Later Combes and Hislop [18] relaxed this condition.

Meanwhile, there is a large number of results on "generalized" Wegner estimates of the form

$$\mathbb{E}(N_\Lambda(E + \epsilon) - N_\Lambda(E - \epsilon)) \leq C|\Lambda|^k \epsilon^\alpha. \quad (4.20)$$

For $k = 1$, they imply Hölder continuity of the integrated density of states with Hölder exponent α .

We refer to the excellent survey by Veselic [177] on the subject which describes the development until 2004. In addition, we mention the papers [4, 19, 23, 22, 24, 57, 59, 61, 82, 106, 107, 108] for further reading.

Very recently, Combes, Hislop and Klopp [20] published a result which includes (and improves) virtually all previous results.

THEOREM 4.17. *Let q_i be independent random variables with a common distribution P_0 of compact support. Let f be a non-negative single site potential of compact support. Then*

- (1) *If P_0 is Hölder continuous with Hölder exponent α , then N is Hölder continuous with the same exponent.*
- (2) *If P_0 is Lifshitz continuous, then N is Lifshitz continuous as well. In this case N has a bounded density.*

The authors of [20] actually prove a more general theorem allowing the random variables to be dependent and including a magnetic field.

4.6. Beyond the Density of States: Level Statistics. One may look at the energy statistics of a disordered system on a smaller scale than we do for the integrated density of states. This subject is common in the theory of random matrices since the days of Wigner and Dyson (see, e.g., [124] or [29]), but is still in its infancy for random Schrödinger operators.

Suppose we have a random Schrödinger operator H which we restrict to a cube Λ_L by appropriate boundary conditions. We call the resulting operator H_{Λ_L} . For a fixed cube Λ_L , the operator H_{Λ_L} has roughly $|\Lambda_L|$ energy levels around an energy E . So we may say that the averaged level spacing near E is $\frac{1}{|\Lambda_L|}$. We now look at the eigenvalues around E under a microscope zooming the averaged level spacing to 1. The keyword is “unfolding of the spectrum.” By this we mean we look at the measure

$$\mu_L([a, b]) = \#\left\{n; E_n(H_{\Lambda_L}) \in \left[E + \frac{a}{|\Lambda_L|}, E + \frac{b}{|\Lambda_L|}\right]\right\}. \tag{4.21}$$

It is reasonable to ask whether there is a limit of μ_L when L goes to infinity. Moreover, if such a limit exists, what are the properties of the limit measure μ ?

Molchanov [132] proved the existence of this limit for a one-dimensional model in the continuum. Minami [131] investigated the multidimensional discrete case in the regime of Anderson localization. Observe that we have localization for all energies for Molchanov’s model.

We discuss Minami’s case (Molchanov’s case being similar). In the following we give merely a rough sketch of Minami’s result, leaving out many details—even assumptions and precise statements. A complete discussion is far beyond the scope of this paper. We urge the reader to look at the paper [131] to get a complete picture.

Let us suppose we have an integrated density of states N which has a bounded density n , i.e., $N(E) = \int_{-\infty}^E n(\lambda) d\lambda$. For energies near the bottom of the spectrum, Minami shows that μ_L is asymptotically a Poisson measure of intensity $n(E)$. This implies that for $a < b$,

$$\mathbb{P}(\mu_L([a, b] = k)) \approx \frac{n(E)^k (b - a)^k}{k!} e^{-n(E)(b-a)}. \tag{4.22}$$

Moreover, the random variables $\mu_L(A_1), \dots, \mu_L(A_m)$ are approximately independent.

In an informal (but provable) sense, this means that the eigenvalues near E look like independent random variables with a uniform distribution. Especially we have

$$\mathbb{P}(\mu_L([a, b] = 1)) \approx n(E)(b - a) e^{-n(E)(b-a)}, \quad (4.23)$$

$$\mathbb{P}(\mu_L([a, b] = 2)) \approx \frac{n(E)^2(b - a)^2}{2} e^{-n(E)(b-a)} \quad (4.24)$$

and consequently one may prove

$$\mathbb{P}\left(\mu_L([-x/2, x/2]) \leq 2 \mid \mu_L([-x/2, x/2]) \leq 1\right) \approx \frac{(n(E)^2 x^2)/2}{n(E) x} \sim n(E) x.$$

This says in a rough way that the *differences* of energy levels near E have a probability density which is strictly positive near 0. In physics terminology, there is no level repulsion.

One expects that this is not the case in the energy region of extended states. So far nobody has proven the existence of extended states for the Anderson model. A fortiori, nobody has proven level repulsion in this case. However, level repulsion is well established for the classical Wigner–Dyson ensembles of random matrix theory ([124, 29]).

References

- [1] Aizenman, M.; Simon, B.: *Brownian motion and Harnack inequality for Schrödinger operators*. Commun. Pure Appl. Math. **35**, 209–273 (1982).
- [2] Akcoglu, M.A.; Krengel, U.: *Ergodic theorems for superadditive processes*. J. Reine Angew. Math. **323**, 53–67 (1981).
- [3] Avron, J.; Simon, B.: *Almost periodic Schrödinger operators. II: The integrated density of states*. Duke Math. J. **50**, 369–391 (1983).
- [4] Barbaroux, J.M.; Combes, J.M.; Hislop, P.D.: *Localization near band edges for random Schrödinger operators*. Helv. Phys. Acta **70**, 16–43 (1997).
- [5] Barlow, M.T.: *Random walks on supercritical percolation clusters*. Ann. Probab. **32**, 3024–3084 (2004).
- [6] Benderskij, M.M.; Pastur, L.A.: *On the spectrum of the one-dimensional Schrödinger equation with a random potential*. Math. USSR-Sb. **11**, 245–256 (1970).
- [7] Biskup, M.; König, W.: *Long-time tails in the parabolic Anderson model*. Ann. Probab. **29**, 636–682 (2001).
- [8] Bingham, N.H.; Goldie, C.M.; Teugels, J.L.: *Regular variation*. Encyclopedia of Mathematics and Its Applications, Vol. 27. Cambridge: Cambridge University Press (1987).
- [9] Birman, M.Sh.; Yafaev, D.R.: *The spectral shift function. The work of M. G. Krejn and its further development*. St. Petersburg. Math. J. **4**, 833–870 (1993).
- [10] Broderix, K.; Hundertmark, D.; Kirsch, W.; Leschke, H.: *The fate of Lifshits tails in magnetic fields*. J. Stat. Phys. **80**, 1–22 (1995).
- [11] de Bruijn, N.G.: *Asymptotic Methods in Analysis*. 2nd edition. Groningen: Wolters–Noordhoff Publishing (1967).
- [12] Campanino, M.; Klein, A.: *A supersymmetric transfer matrix and differentiability of the density of states in the one-dimensional Anderson model*. Commun. Math. Phys. **104**, 227–241 (1986).
- [13] Campanino, M.; Klein, A.: *Anomalies in the one-dimensional Anderson model at weak disorder*. Commun. Math. Phys. **130**, 441–456 (1990).
- [14] Carmona, R.; Klein, A.; Martinelli, F.: *Anderson localization for Bernoulli and other singular potentials*. Commun. Math. Phys. **108**, 41–66 (1987).

- [15] Carmona, R.; Lacroix, J.: *Spectral theory of random Schrödinger operators*. Probability and Its Applications. Basel: Birkhäuser Verlag (1990).
- [16] Chahrouh, A.: *Densité intégrée d'états surfaciques et fonction généralisée de déplacement spectral pour un opérateur de Schrödinger surfacique ergodique*. (Integrated density of surface states and generalized function of spectral shift for an ergodic surface Schrödinger operator). *Helv. Phys. Acta* **72**, 93–122 (1999).
- [17] Chahrouh, A.: *On the spectrum of the Schrödinger operator with periodic surface potential*. *Lett. Math. Phys.* **52**, 197–209 (2000).
- [18] Combes, J.M.; Hislop, P.D.: *Localization for some continuous, random Hamiltonians in d -dimensions*. *J. Funct. Anal.* **124**, 149–180 (1994).
- [19] Combes, J.M.; Hislop, P.D.; Klopp, F.: *Hölder continuity of the integrated density of states for some random operators at all energies*. *Int. Math. Res. Not.* 2003, No. 4, 179–209 (2003).
- [20] Combes, J.M.; Hislop, P.D.; Klopp, F.: *An optimal Wegner estimate and its application to the global continuity of the integrated density of states for random Schrödinger operators*. Preprint.
- [21] Combes, J.M.; Hislop, P.D.; Mourre, E.: *Spectral averaging, perturbation of singular spectra, and localization*. *Trans. Amer. Math. Soc.* **348**, 4883–4894 (1996).
- [22] Combes, J.M.; Hislop, P.D.; Nakamura, S.: *The L^p -theory of the spectral shift function, the Wegner estimate, and the integrated density of states for some random operators*. *Commun. Math. Phys.* **218**, 113–130 (2001).
- [23] Combes, J.M.; Hislop, P.D.; Klopp, F.; Nakamura, S.: *The Wegner estimate and the integrated density of states for some random operators*. *Proc. Indian Acad. Sci. (Math. Sci.)* **112**, 31–53 (2002).
- [24] Combes, J.M.; Hislop, P.D.; Klopp, F.; Raikov, G.: *Global continuity of the integrated density of states for random Landau Hamiltonians*. *Commun. Partial Differ. Equations* **29**, 1187–1213 (2004).
- [25] Craig, W.; Simon, B.: *Log Hölder continuity of the integrated density of states for stochastic Jacobi matrices*. *Commun. Math. Phys.* **90**, 207–218 (1983).
- [26] Craig, W.; Simon, B.: *Subharmonicity of the Lyapunov index*. *Duke Math. J.* **50**, 551–560 (1983).
- [27] Cycon, H.L.; Froese, R.G.; Kirsch, W.; Simon, B.: *Schrödinger Operators, with Application to Quantum Mechanics and Global Geometry*. (Springer Study ed.) Texts and Monographs in Physics. Berlin: Springer-Verlag (1987).
- [28] Damanik, D.; Sims, R.; Stolz, G.: *Localization for one-dimensional, continuum, Bernoulli–Anderson models*. *Duke Math. J.* **114**, 59–100 (2002).
- [29] Deift, P.: *Orthogonal polynomials and random matrices: a Riemann–Hilbert approach*. Courant Lecture Notes in Mathematics, 3. Providence, RI: American Mathematical Society (1999).
- [30] Delyon, F.; Souillard, B.: *Remark on the continuity of the density of states of ergodic finite difference operators*. *Commun. Math. Phys.* **94**, 289–291 (1984).
- [31] Dembo, A.; Zeitouni, O.: *Large deviations and applications*. Handbook of Stochastic Analysis and Applications (edited by D. Kannan et al.). New York: Marcel Dekker. Stat., Textb. Monogr. 163, pp. 361–416 (2001).
- [32] Deuschel, J.-D.; Stroock, D.: *Large Deviations*. Rev. ed. Pure and Applied Mathematics, 137. Boston: Academic Press (1989).
- [33] Donsker, M.; Varadhan S.: *Asymptotic for the Wiener sausage*. *Commun. Pure Appl. Math.* **28**, 525–565 (1975).
- [34] Droese, J.; Kirsch, W.: *The effect of boundary conditions on the density of states for random Schrödinger operators*. *Stochastic Processes Appl.* **23**, 169–175 (1986).
- [35] Eastham, M.S.P.: *The Spectral Theory of Periodic Differential Equations*. Texts in Mathematics. Edinburgh–London: Scottish Academic Press (1973).
- [36] Englisch, H.; Kirsch, W.; Schröder, M.; Simon, B.: *The density of surface states*. *Phys. Rev. Letters* **61**, 1261–1262 (1988).
- [37] Englisch, H.; Kirsch, W.; Schröder, M.; Simon, B.: *Random Hamiltonians ergodic in all but one direction*. *Commun. Math. Phys.* **128**, 613–625 (1990).
- [38] Erdős, L.: *Lifschitz tail in a magnetic field: The nonclassical regime*. *Probab. Theory Relat. Fields* **112**, 321–371 (1998).

- [39] Erdős, L.: *Lifshitz tail in a magnetic field: Coexistence of classical and quantum behavior in the borderline case*. Probab. Theory Relat. Fields **121**, 219–236 (2001).
- [40] Feller, W.: *An Introduction to Probability Theory and Its Applications. Vol II*. 2nd ed. Wiley Series in Probability and Mathematical Statistics. New York: John Wiley and Sons (1971).
- [41] Fischer, W.; Hupfer, T.; Leschke, H.; Müller, P.: *Existence of the density of states for multi-dimensional continuum Schrödinger operators with Gaussian random potentials*. Commun. Math. Phys. **190**, 133–141 (1997).
- [42] Fröhlich, J.; Martinelli, F.; Scoppola, E.; Spencer, T.: *Constructive proof of localization in the Anderson tight binding model*. Commun. Math. Phys. **101**, 21–46 (1985).
- [43] Fröhlich, J.; Spencer, T.: *Absence of diffusion in the Anderson tight binding model for large disorder or low energy*. Commun. Math. Phys. **88**, 151–184 (1983).
- [44] Fukushima, M.: *On asymptotics of spectra of Schrödinger equations*. Aspects Statistiques et Aspects Physiques des Processus Gaussiens (Saint-Flour, 1980). Colloq. Int. C.N.R.S., 307, pp. 335–347 (1981).
- [45] Fukushima, M.; Nagai, H.; Nakao, S.: *On an asymptotic property of spectra of a random difference operator*. Proc. Japan Acad. **51**, 100–102 (1975).
- [46] Fürstenberg, H.; Kesten, H.: *Products of random matrices*, Ann. Math. Stat. **31**, 457–469 (1960).
- [47] Gärtner, J.; König, W.: *Moment asymptotics for the continuous parabolic Anderson model*. Ann. Appl. Probab. **10**, 192–217 (2000).
- [48] Gärtner, J.; Molchanov, S.: *Parabolic problems for the Anderson model. I: Intermittency and related topics*. Commun. Math. Phys. **132**, 613–655 (1990).
- [49] Gärtner, J.; Molchanov, S.: *Parabolic problems for the Anderson model. II: Second-order asymptotics and structure of high peaks*. Probab. Theory Relat. Fields **111**, 17–55 (1998).
- [50] Gärtner, J.; Molchanov, S.: *Moment asymptotics and Lifshitz tails for the parabolic Anderson model*. Stochastic Models (Ottawa, 1998). Edited by Luis G. Gorostiza et al. Providence, RI: American Mathematical Society for the Canadian Mathematical Society. CMS Conf. Proc. 26, pp. 141–157 (2000).
- [51] Goldstein, M.; Schlag, W.: *Hölder continuity of the integrated density of states for quasi-periodic Schrödinger equations and averages of shifts of subharmonic functions*. Ann. of Math. (2) **154**, 155–203 (2001).
- [52] Goldstein, M.; Schlag, W.: *On the formation of gaps in the spectrum of Schrödinger operators with quasi-periodic potentials*. In this Festschrift.
- [53] Goldstein, M.; Schlag, W.: *Fine properties of the integrated density of states and a quantitative separation property of the Dirichlet eigenvalues*. Preprint 2005.
- [54] Grimmett, G.: *Percolation*, 2nd ed. New York: Springer (1999).
- [55] Halperin, B.: *Properties of a particle in a one-dimensional random potential*. Adv. Chem. Phys. **31**, 123–177 (1967).
- [56] Herbert, D.; Jones, R.: *Localized states in disordered systems*. J. Phys. C **4**, 1145–1161 (1971).
- [57] Hislop, P.D.; Klopp, F.: *The integrated density of states for some random operators with nonsign definite potentials*. J. Funct. Anal. **195**, 12–47 (2002).
- [58] den Hollander, F.: *Large Deviations*. Fields Institute Monographs, 14. Providence, RI: American Mathematical Society (2000).
- [59] Hundertmark, D.; Killip, R.; Nakamura, S.; Stollmann, P.; Veselić, I.: *Bounds on the spectral shift function and the density of states*. Comm. Math. Phys. **262**, 489–503 (2005).
- [60] Hundertmark, D.; Kirsch, W.; Warzel, S.: *Classical magnetic Lifshitz tails in three space dimensions: Impurity potentials with slow anisotropic decay*. Markov Process. Relat. Fields **9**, 651–660 (2003).
- [61] Hundertmark, D.; Simon, B.: *An optimal L^p -bound on the Krein spectral shift function*. J. Anal. Math. **87**, 199–208 (2002).
- [62] Hupfer, T.; Leschke, H.; Müller, P.; Warzel, S.: *Existence and uniqueness of the integrated density of states for Schrödinger operators with magnetic fields and unbounded random potentials*. Rev. Math. Phys. **13**, 1547–1581 (2001).
- [63] Hupfer, T.; Leschke, H.; Müller, P.; Warzel, S.: *The absolute continuity of the integrated density of states for magnetic Schrödinger operators with certain unbounded random potentials*. Commun. Math. Phys. **221**, 229–254 (2001).

- [64] Hupfer, T.; Leschke, H.; Warzel, S.: *Poissonian obstacles with Gaussian walls discriminate between classical and quantum Lifshits tailing in magnetic fields*. J. Stat. Phys. **97**, 725–750 (1999).
- [65] Javrjan, V.A.: *A certain inverse problem for Sturm-Liouville operators*. (Russian) Izv. Akad. Nauk Ar-mjan. SSR Ser. Mat. **6**, 246–251 (1971).
- [66] Kingman, J.: *Subadditive ergodic theory*. Ann. Probab. **1**, 883–909 (1973).
- [67] Kirsch, W.: *Small perturbations and the eigenvalues of the Laplacian on large bounded domains*. Proc. Amer. Math. Soc. **101**, 509–512 (1987).
- [68] Kirsch, W.: *Random Schrödinger operators*. Schrödinger Operators (Sønderborg, Denmark, 1988). Lect. Notes in Phys., 345, pp. 264–370 (1989).
- [69] Kirsch, W.: *Wegner estimates and Anderson localization for alloy-type potentials*. Math. Z. **221**, 507–512 (1996).
- [70] Kirsch, W.: *An invitation to random Schrödinger operators*. Preprint (2006).
- [71] Kirsch, W.; Khorunzhiy, O.; Müller, P.: *Lifshits tails for spectra of Erdős–Rényi random graphs*. Ann. Appl. Probab. **16**, 295–309 (2006).
- [72] Kirsch, W.; Klopp, F.: *The band-edge behavior of the density of surfacic states*. Math. Phys. Anal. Geom. To appear.
- [73] Kirsch, W.; Leschke, H.; Warzel, S.: *On Lifshitz tails for homogeneous magnetic fields in three dimensions*. Unpublished.
- [74] Kirsch, W.; Martinelli, F.: *On the density of states of Schrödinger operators with a random potential*. J. Phys. A **15**, 2139–2156 (1982).
- [75] Kirsch, W.; Martinelli, F.: *On the spectrum of Schrödinger operators with a random potential*. Commun. Math. Phys. **85**, 329–350 (1982).
- [76] Kirsch, W.; Martinelli, F.: *On the ergodic properties of the spectrum of general random operators*. J. Reine Angew. Math. **334**, 141–156 (1982).
- [77] Kirsch, W.; Martinelli, F.: *Large deviations and Lifshitz singularity of the integrated density of states of random Hamiltonians*. Commun. Math. Phys. **89**, 27–40 (1983).
- [78] Kirsch, W.; Müller, P.: *Spectral properties of the Laplacian on bond-percolation graphs*. Math. Z. **252**, 899–916 (2006).
- [79] Kirsch, W.; Nitzschner, F.: *Lifshitz-tails and non-Lifshitz-tails for one-dimensional random point interactions*. Oper. Theory, Adv. Appl. **46**, 171–178 (1990).
- [80] Kirsch, W.; Simon, B.: *Lifshitz tails for periodic plus random potentials*. J. Stat. Phys. **42**, 799–808 (1986).
- [81] Kirsch, W.; Simon, B.: *Comparison theorems for the gap of Schrödinger operators*. J. Funct. Anal. **75**, 396–410 (1987).
- [82] Kirsch, W.; Stollmann, P.; Stolz, G.: *Anderson localization for random Schrödinger operators with long range interactions*. Commun. Math. Phys. **195**, 495–507 (1998).
- [83] Kirsch, W.; Stollmann, P.; Stolz, G.: *Localization for random perturbations of periodic Schrödinger operators*. Random Oper. Stoch. Equ. **6**, 241–268 (1998).
- [84] Kirsch, W.; Veselic, I.: *Existence of the density of states for one-dimensional alloy-type potentials with small support*. Mathematical Results in Quantum Mechanics (Taxco, Mexico, 2001). Rdated by Ricardo Weder et al. Contemp. Math. **307**, 171–176 (2002). Providence, RI: American Mathematical Society.
- [85] Kirsch, W.; Veselic, I.: *Wegner estimate for sparse and other generalized alloy type potentials*. Proc. Indian Acad. Sci. (Math. Sci.) **112**, 131–146 (2002).
- [86] Kirsch, W.; Warzel, S.: *Lifshits tails caused by anisotropic decay: the emergence of a quantum-classical regime*. Math. Phys. Anal. Geom. **8**, 257–285 (2005).
- [87] Kirsch, W.; Warzel, S.: *Anderson localization and Lifshits tails for random surface potentials*. J. Funct. Anal. **230**, 222–250 (2006).
- [88] Klassert, S.: *Unschärfe, der supersymmetrische Replica Trick und Lokalisierung*, Diploma Thesis, Johann Wolfgang Goethe-Universität, Frankfurt (2001).
- [89] Kleespies, F.; Stollmann, P.: *Lifshitz asymptotics and localization for random quantum waveguides*. Rev. Math. Phys. **12**, 1345–1365 (2000).
- [90] Klein, A.: *The supersymmetric replica trick and smoothness of the density of states for random Schrödinger operators*. Operator Theory, Operator Algebras and Applications, (Durham, NH, 1988). Proc. Symp. Pure Math., 51, Part 1, pp. 315–331 (1990).

- [91] Klein, A.; Lacroix, J.; Speis, A.: *Regularity of the density of states in the Anderson model on a strip for potentials with singular continuous distributions*. J. Stat. Phys. **57**, 65–88 (1989).
- [92] Klein, A.; Martinelli, F.; Perez, J.F.: *A rigorous replica trick approach to Anderson localization in one dimension*. Commun. Math. Phys. **106**, 623–633 (1986).
- [93] Klein, A.; Speis, A.: *Smoothness of the density of states in the Anderson model on a one-dimensional strip*. Ann. Phys. **183**, 352–398 (1988).
- [94] Klein, A.; Speis, A.: *Regularity of the invariant measure and of the density of states in the one-dimensional Anderson model*. J. Funct. Anal. **88**, 211–227 (1990).
- [95] Klopp, F.: *Localization for some continuous random Schrödinger operators*. Commun. Math. Phys. **167**, 553–569 (1995).
- [96] Klopp, F.: *Band edge behavior of the integrated density of states of random Jacobi matrices in dimension 1*. J. Stat. Phys. **90**, 927–947 (1998).
- [97] Klopp, F.: *Internal Lifshitz tails for random perturbations of periodic Schrödinger operators*. Duke Math. J. **98**, 335–396 (1999).
- [98] Klopp, F.: *Precise high energy asymptotics for the integrated density of states of an unbounded random Jacobi matrix*. Rev. Math. Phys. **12**, 575–620 (2000).
- [99] Klopp, F.: *Weak disorder localization and Lifshitz tails: Continuous Hamiltonians*. Ann. Henri Poincaré **3**, 711–737 (2002).
- [100] Klopp, F.: *Lifshitz tails for random perturbations of periodic Schrödinger operators*. Proc. Indian Acad. Sci. (Math. Sci.) **112**, 147–162 (2002).
- [101] Klopp, F.; Pastur, L.: *Lifshitz tails for random Schrödinger operators with negative singular Poisson potential*. Commun. Math. Phys. **206**, 57–103 (1999).
- [102] Klopp, F.; Raikov, G.: *Lifshitz tails in constant magnetic fields*. To appear in Commun. Math. Phys.
- [103] Klopp, F.; Ralston, J.: *Endpoints of the spectrum of periodic operators are generically simple*. Methods Appl. Anal. **7**, 459–463 (2000).
- [104] Klopp, F.; Wolff, T.: *Internal Lifshitz tails for random Schrödinger operators*. XIIIth International Congress on Mathematical Physics (Imperial College, London, 2000). pp. 425–433. Edited by A. Grigoryan et al. Boston: International Press (2001).
- [105] Klopp, F.; Wolff, T.: *Lifshitz tails for 2-dimensional random Schrödinger operators*. J. Anal. Math. **88**, 63–147 (2002).
- [106] Kostykin, V.; Schrader, R.: *The density of states and the spectral shift density of random Schrödinger operators*. Rev. Math. Phys. **12**, 807–847 (2000).
- [107] Kostykin, V.; Schrader, R.: *Regularity of the surface density of states*. J. Funct. Anal. **187**, 227–246 (2001).
- [108] Kostykin, V.; Veselic, I.: *On the Lipschitz continuity of the integrated density of states for sign-indefinite potentials*. Math. Z. **252**, 367–392 (2006).
- [109] Kotani, S.: *Lyapunov exponents and spectra for one-dimensional random Schrödinger operators*. Random Matrices and Their Applications (Brunswick, Maine, 1984). Contemp. Math. **50**, 277–286 (1986).
- [110] Kotani, S.; Simon, B.: *Localization in general one-dimensional random systems. II. Continuum Schrödinger operators*. Commun. Math. Phys. **112**, 103–119 (1987).
- [111] Krengel, U.: *Ergodic Theorems*. De Gruyter Studies in Mathematics, 6. Berlin–New York: Walter de Gruyter (1985).
- [112] Kuchment, P.: *Floquet Theory for Partial Differential Equations*. Operator Theory: Advances and Applications, 60. Basel: Birkhäuser Verlag (1993).
- [113] Lamperti, J.: *Stochastic Processes. A Survey of the Mathematical Theory*. Applied Mathematical Sciences, Vol. 23. New York–Heidelberg: Springer (1977).
- [114] LePage, E.: *Répartition d'état d'un opérateur de Schrödinger aléatoire. Distribution empirique des valeurs propres d'une matrice de Jacobi*. (State distribution of a random Schrödinger operator. Empirical distribution of the eigenvalues of a Jacobi matrix) Probability Measures on Groups, VII (Oberwolfach, 1983). Lecture Notes in Math., 1064, pp. 309–367. Berlin: Springer (1984).
- [115] Leschke, H.; Müller, P.; Warzel, S.: *A survey of rigorous results on random Schrödinger operators for amorphous solids*. Interacting Stochastic Systems. Edited by Jean-Dominique Deuschel et al. pp. 119–151. Berlin: Springer (2005).

- [116] Leschke, H.; Warzel, S.: *Quantum-classical transitions in Lifshitz tails with magnetic fields*. Phys. Rev. Lett. **92**, 086402 (2004).
- [117] Lifshitz, I.: *Structure of the energy spectrum of impurity bands in disordered solid solutions*. Soviet Physics JETP **17**, 1159–1170 (1963).
- [118] Lifshitz, I.: *Energy spectrum structure and quantum states of disordered condensed systems*. Soviet Physics Usp. **7**, 549–573 (1965).
- [119] Lifshitz, I.; Gredeskul, S.A.; Pastur, L.A.: *Introduction to the Theory of Disordered Systems*. New York: Wiley (1988).
- [120] March, P.; Sznitman, A.-S.: *Some connections between excursion theory and the discrete Schrödinger equation with random potentials*. Probab. Theory Relat. Fields **75**, 111–53 (1987).
- [121] Martinelli, F.; Holden, H.: *On absence of diffusion near the bottom of the spectrum for a random Schrödinger operator on $L^2(\mathbb{R}^d)$* . Commun. Math. Phys. **93**, 197–217 (1984).
- [122] Martinelli, F.; Micheli L.: *On the large coupling constant behavior of the Lyapunov exponent of a binary alloy*. J. Stat. Phys. **48**, 1–18 (1986).
- [123] McKean, H.P.; Trubowitz, E.: *Hill's surfaces and their theta functions*. Bull. Amer. Math. Soc. **84**, 1042–1085 (1978).
- [124] Mehta, M.L.: *Random Matrices*. Boston: Academic Press (1991).
- [125] Metzger, B.: *Asymptotische Eigenschaften im Wechselspiel von Diffusion und Wellenausbreitung in zufälligen Medien*, Ph.D. thesis, TU Chemnitz (2005).
- [126] Metzger, B.: *Lifshitz asymptotics for the Anderson model: On the transition from classical to quantum regime*. In preparation.
- [127] Metzger, B.: *Tauberian estimates from below and above*. In preparation.
- [128] Mezincescu, G.A.: *Internal Lifshitz singularities of disordered finite-difference Schrödinger operators*. Commun. Math. Phys. **103**, 167–176 (1986).
- [129] Mezincescu, G.A.: *Lifshitz singularities for periodic operators plus random potentials*. J. Stat. Phys. **49**, 1181–1190 (1987).
- [130] Mezincescu, G.A.: *Internal Lifshitz singularities for one-dimensional Schrödinger operators*. Commun. Math. Phys. **158**, 315–325 (1993).
- [131] Minami, N.: *Local fluctuation of the spectrum of a multidimensional Anderson tight binding model*. Commun. Math. Phys. **177**, 709–725 (1996).
- [132] Molchanov, S.A.: *The local structure of the spectrum of the one-dimensional Schrödinger operator*. Commun. Math. Phys. **78**, 429–446 (1981).
- [133] Müller, P.; Stollmann, P.: *Spectral asymptotics of the Laplacian on supercritical bond-percolation graphs*. Preprint (2005).
- [134] Najjar, H.: *Asymptotique de la densité d'états intégrée des modèles aléatoires continus*. Thèse Université Paris 13 (2000).
- [135] Najjar, H.: *Lifshitz tails for random acoustic operators*. J. Math. Phys. **44**, 1842–1867 (2003).
- [136] Najjar, H.: *Asymptotic behavior of the integrated density of states of acoustic operators with random long range perturbations*. J. Stat. Phys. **115**, 977–996 (2004).
- [137] Nakamura, S.: *Spectral shift function for trapping energies in the semiclassical limit*. Commun. Math. Phys. **208**, 173–193 (1999).
- [138] Nakamura, S.: *Lifshitz tail for 2D discrete Schrödinger operator with random magnetic field*. Ann. Henri Poincaré **1**, 823–835 (2000).
- [139] Nakamura, S.: *Lifshitz tail for Schrödinger operator with random magnetic field*. Commun. Math. Phys. **214**, 565–572 (2000).
- [140] Nakamura, S.: *A remark on the Dirichlet–Neumann decoupling and the integrated density of states*. J. Funct. Anal. **179**, 136–152 (2001).
- [141] Nakamura, S.: *A remark on the Lifshitz tail for Schrödinger operator with random magnetic field*. Proc. Indian Acad. Sci. (Math. Sci.) **112**, 183–187 (2002).
- [142] Nakao, S.: *On the spectral distribution of the Schrödinger operator with random potential*. Jap. J. Math. (N.S.) **3**, 111–139 (1977).
- [143] Pastur, L.A.: *Selfaveragability of the number of states of the Schrödinger equation with a random potential*. (Russian) Mat. Fiz. i Funkcional. Anal. Vyp. **2**, 111–116, 238 (1971).
- [144] Pastur, L.A.: *On the distribution of the eigenvalues of the Schrödinger equation with a random potential*. Funct. Anal. Appl. **6**, 163–165 (1972).
- [145] Pastur, L.A.: *Spectra of random selfadjoint operators*. Russian Math. Surv. **28**, 1–67 (1973).

- [146] Pastur, L.A.: *Behavior of some Wiener integrals as $t \rightarrow \infty$ and the density of states of Schrödinger equations with random potential*. Theor. Math. Phys. **32**, 615–620 (1978).
- [147] Pastur, L.A.: *Spectral properties of disordered systems in the one-body approximation*. Commun. Math. Phys. **75**, 179–196 (1980).
- [148] Pastur, L.; Figotin, A.: *Spectra of Random and Almost-Periodic Operators*. Grundlehren der Mathematischen Wissenschaften, 297. Berlin: Springer-Verlag (1992).
- [149] Reed, M.; Simon, B.: *Methods of Modern Mathematical Physics. II: Fourier Analysis, Self-Adjointness*. New York: Academic Press (1975).
- [150] Reed M.; Simon B.: *Methods of Modern Mathematical Physics. IV: Analysis of Operators*. New York: Academic Press (1978).
- [151] Reed, M.; Simon, B.: *Methods of Modern Mathematical Physics. III: Scattering Theory*. New York: Academic Press (1979).
- [152] Reed, M.; Simon, B.: *Methods of Modern Mathematical Physics. I: Functional Analysis*. Revised and enlarged ed. New York: Academic Press (1980).
- [153] Shubin, C.; Vakilian, R.; Wolff, T.: *Some harmonic analysis questions suggested by Anderson–Bernoulli models*. Geom. Funct. Anal. **8**, 932–964 (1998).
- [154] Simon, B.: *Functional Integration and Quantum Physics*. Pure and Applied Mathematics, 86. New York: Academic Press (1979); 2nd ed. Providence, RI: AMS Chelsea Publishing (2005).
- [155] Simon, B.: *Schrödinger semigroups*. Bull. Amer. Math. Soc. (N.S.) **7**, 447–526 (1982).
- [156] Simon, B.: *Lifshitz tails for the Anderson model*. J. Phys. **38**, 65–76 (1985).
- [157] Simon, B.: *Internal Lifshitz tails*. J. Stat. Phys. **46**, 911–918 (1987).
- [158] Simon, B.: *Spectral analysis of rank one perturbations and applications*. Mathematical Quantum Theory II: Schrödinger Operators (Vancouver, 1993). Edited by J. Feldman et al. Providence, RI: American Mathematical Society. CRM Proc. Lect. Notes. **8**, 109–149 (1995).
- [159] Simon, B.: *Spectral averaging and the Krein spectral shift*. Proc. Amer. Math. Soc. **126**, 1409–1413 (1998).
- [160] Simon, B.: *Schrödinger operators in the twenty-first century*, XIIIth International Congress on Mathematical Physics (Imperial College, London, 2000). Edited by A. Fokas et al. pp. 283–288. Boston: International Press (2001).
- [161] Simon, B.; Taylor, M.: *Harmonic analysis on $SL(2, \mathbb{R})$ and smoothness of the density of states in the one-dimensional Anderson model*, Commun. Math. Phys. **101**, 1–19 (1985).
- [162] Simon, B.; Taylor, M.; Wolff, T.: *Some rigorous results for the Anderson model*, Phys. Rev. Lett. **54**, 1589–1592 (1985).
- [163] Simon, B.; Wolff, T.: *Singular continuous spectrum under rank one perturbations and localization for random Hamiltonians*. Commun. Pure Appl. Math. **39**, 75–90 (1986).
- [164] Stollmann, P.: *Lifshitz asymptotics via linear coupling of disorder*. Math. Phys. Anal. Geom. **2**, 279–289 (1999).
- [165] Stollmann, P.: *Wegner estimates and localization for continuum Anderson models with some singular distributions*. Arch. Math. **75**, 307–311 (2000).
- [166] Stollmann, P.: *Caught by disorder*. Bound States in Random Media. Progress in Mathematical Physics, 20. Boston: Birkhäuser (2001).
- [167] Sznitman, A.-S.: *Lifshitz tail and Wiener sausage on hyperbolic space*. Commun. Pure Appl. Math. **42**, 1033–1065 (1989).
- [168] Sznitman, A.-S.: *Lifshitz tail and Wiener sausage*. J. Funct. Anal. **94**, 223–246, 247–272 (1990).
- [169] Sznitman, A.-S.: *Lifshitz tail on hyperbolic space: Neumann conditions*. Commun. Pure Appl. Math. **43**, 1–30 (1990).
- [170] Sznitman, A.-S.: *Brownian Motion, Obstacles and Random Media*. Springer Monographs in Mathematics. Berlin: Springer (1998).
- [171] Thouless, D.J.: *Electrons in disordered systems and the theory of localization*. Phys. Rep. **13**, 93–142 (1974).
- [172] Ueki, N.: *On spectra of random Schrödinger operators with magnetic fields*. Osaka J. Math. **31**, 177–187 (1994).
- [173] Ueki, N.: *Simple examples of Lifshitz tails in Gaussian random magnetic fields*. Ann. Henri Poincaré **1**, 473–498 (2000).

- [174] Ueki, N.: *Lifschitz tail of a random Pauli Hamiltonian with an anomalous moment*. Jap. J. Math. (N.S.) **28**, 261–286 (2002).
- [175] Veselic, I.: *Localization for random perturbations of periodic Schrödinger operators with regular Floquet eigenvalues*. Ann. Henri Poincaré **3**, 389–409 (2002).
- [176] Veselic, I.: *Spectral analysis of percolation Hamiltonians*. Math. Ann. **331**, 841–865 (2005).
- [177] Veselic, I.: *Integrated density and Wegner estimates for random Schrödinger operators*. Spectral Theory of Schrödinger Operators (Mexico, 2001). Edited by Rafael del Río et al. Providence, RI: American Mathematical Society (AMS). Contemp. Math. **340**, 97–183 (2004).
- [178] Warzel, S.: *On Lifshitz tails in magnetic fields*. Doctoral thesis, Universität Erlangen–Nürnberg, Logos, Berlin (2001).
- [179] Wegner, F.: *Bounds on the density of states in disordered systems*. Z. Phys. B **44**, 9–15 (1981).
- [180] Yafaev, D.: *Mathematical scattering theory. General theory*. Translations of Mathematical Monographs, 105. Providence, RI: American Mathematical Society (1992).

INSTITUT FÜR MATHEMATIK, RUHR-UNIVERSITÄT BOCHUM, D-44780 BOCHUM, GERMANY

E-mail address: `werner.kirsch@ruhr-uni-bochum.de`

E-mail address: `bernd.metzger@ruhr-uni-bochum.de`

Exotic Spectra: A Review of Barry Simon's Central Contributions

Yoram Last

Dedicated to Barry Simon on the occasion of his 60th birthday

ABSTRACT. We review some of Barry Simon's central contributions concerning what is often called exotic spectral properties. These include phenomena such as Cantor spectrum, thick point spectrum and singular continuous spectrum.

CONTENTS

1. Introduction
 2. The Almost Periodic Flu
 3. Thick Point Spectrum
 4. The Singular Continuous Spectrum Revolution
- References

1. Introduction

“General wisdom used to say that Schrödinger operators should have absolutely continuous spectrum plus some discrete point spectrum, while singular continuous spectrum is a pathology that should not occur in examples with V bounded.”

The above quote starts Section 10.4 of the 1987 book “*Schrödinger Operators*” by Cycon, Froese, Kirsch and Simon [15]. A few lines below, it further states: “Another correction to the ‘general picture’ is that point spectrum may be dense in some region of the spectrum rather than being a discrete set.” These statements

2000 *Mathematics Subject Classification.* 81Q10, 47B15, 82B44.

Key words and phrases. Barry Simon, Schrödinger operators, singular continuous spectrum, absolutely continuous spectrum, pure point spectrum, transient spectrum, recurrent spectrum, thick point spectrum, Cantor spectrum, almost periodic potentials.

Supported in part by The Israel Science Foundation (Grant No. 188/02) and by Grant No. 2002068 from the United States–Israel Binational Science Foundation (BSF), Jerusalem, Israel.

are made as background to the introduction of yet a third exotic spectral phenomenon, which is the occurrence of *Cantor spectrum*, and to the characterization of absolutely continuous spectrum associated with Cantor spectrum as *recurrent*.

The book [15] is mostly a summary of a summer school course that Simon gave in 1982. Section 10.4 mainly discusses ideas that were introduced in the seminal 1981 Avron–Simon paper “*Transient and recurrent spectrum*” [3] along with some relevant results of Avron–Simon [5], Bellissard–Simon [11], and Chulaevsky [13]. In particular, the above quotations echo similar statements previously made in [3].

For insight regarding the “general wisdom” of the time, it may be illuminating to look at the 1978 fourth (and last) volume of “*Methods of Modern Mathematical Physics*” by Reed and Simon [49]. This volume is called “*Analysis of Operators*” and its Section XIII.6 starts with the following statement: “Spectral analysis of an operator A concentrates on identifying the five sets $\sigma_{\text{ess}}(A)$, $\sigma_{\text{disc}}(A)$, $\sigma_{\text{ac}}(A)$, $\sigma_{\text{sing}}(A)$, $\sigma_{\text{pp}}(A)$.” This section is called “*The absence of singular continuous spectrum I*” and it is followed by similarly titled sections up to “*The absence of singular continuous spectrum IV*.” These “absence of singular continuous spectrum” sections occupy roughly 16% of the volume and point at the main role played by singular continuous spectrum in pre-1978 spectral theory: It was a non-occurring phenomenon which complicated life by requiring some effort to prove it did not occur.

The Avron–Simon paper [3] posed a significant challenge to the “general wisdom” of its time. First, it extended the above “five sets” by defining four new spectral types (*recurrent absolutely continuous*, *transient absolutely continuous*, *thick pure point* and *thin pure point*). Second, it made the prediction that what they called “extraordinary” spectra (and we call here *exotic spectra*) “will become more and more commonly encountered than one might have thought!” Indeed, Avron–Simon point out this prediction as a potential objection to their choice of the term “extraordinary.” The fact that their prediction materialized so fully is why we decided to reject their proposed terminology and adopt “exotic spectra” instead.

The historical background concerning exotic spectra which preceded [3] consisted of several fairly isolated results. In particular, one should mention the 1977 Goldsheid–Molchanov–Pastur [25] proof of Anderson localization [1] (namely, the occurrence of pure point spectrum with eigenvalues dense in an interval) in a random Schrödinger operator. They considered a continuous one-dimensional operator of the form $-\Delta + V$ on $L^2(\mathbb{R})$ with a certain type of random potential V .

Even earlier, the Ishii–Pastur theorem (see [15]) indicated that some random one-dimensional Schrödinger operators have no absolutely continuous spectrum in spite of their spectrum being an interval (which says they must have either thick point spectrum or singular continuous spectrum or both). Anderson localization for discrete Schrödinger operators with i.i.d.r.v. potentials was proven in 1980 by Kunz–Souillard [43]. Another notable result of the era is Pearson’s seminal 1978 paper [47], which gave an explicit construction of a one-dimensional Schrödinger operator having purely singular continuous spectrum.

[3] was essentially part of a series of papers concerned mostly with almost periodic Schrödinger operators that Simon wrote with several coauthors in the early 1980’s [4, 3, 5, 6, 7, 11, 14, 16, 55, 56]. These works were predated by the 1975 paper of Dinaburg–Sinai [20], which established absolutely continuous spectrum for some almost periodic Schrödinger operators, and by several papers of

Shubin (see [53]). They were paralleled by a number of works by other authors (e.g., [9, 10, 12, 13, 21, 22, 37, 38, 46, 52]).

The phenomenon of sudden broad interest in almost periodic Schrödinger operators has been named the “almost periodic flu” by Simon and is very nicely presented in his review paper [54], which is itself often called “the almost periodic flu paper.” [3] seems to have been primarily motivated by the Avron–Simon discovery [5] of Cantor absolutely continuous spectrum for certain almost periodic Schrödinger operators, but, while other papers in the series were focused purely on almost periodic, or very similar, potentials, [3] took a more general perspective: It looked at the potential implications of the new spectral phenomenon to quantum mechanics, identified natural mathematical structures which arise in this context and, most notably, identified the connection between the newly discovered spectral phenomenon for almost periodic operators and the prior findings of Goldsheid–Molchanov–Pastur [25] and Pearson [47]. We thus believe that it is [3], more than any other single work, that marks the beginning of a new era in spectral analysis: the era of exotic spectra.

We note that by classical inverse spectral theory, one should actually expect the full spectral richness allowed by measure theory to find its way into Schrödinger operators. In particular, we have

THEOREM 1.1 (Gel’fand–Levitan [24]). *Given any finite Borel measure ν on $[a, b] \subset \mathbb{R}$, there exists a continuous half-line Schrödinger operator for which the spectral measure coincides with ν on $[a, b]$.*

Thus, one could argue that the “general wisdom” expecting only “absolutely continuous spectrum plus some discrete point spectrum” was never really justified. Indeed, in looking back, it seems to be very much the consequence of mathematical physicists concentrating their attention on operators arising from atomic and molecular physics, mathematical problems associated with scattering experiments, periodic problems, etc., namely, on problems which happen to have such spectra.

The exotic spectra era can thus be at least partially attributed to mathematical physicists starting to look into some rich problems of modern condensed matter physics which yield rich spectral phenomena. Indeed, several developments in condensed matter physics in the early 1980’s contributed to the growing interest in almost periodic Schrödinger operators and exotic spectra. The 1980 discovery of the integer quantum Hall effect by von Klitzing [40] (for which he got the Nobel Prize in 1985), led to a beautiful theory by Thouless, Kohmoto, Nightingale and den Nijs [66], which explains the quantization of charge transport in this effect as connected with certain topological invariants. Central to their theory is the use of the almost Mathieu operator as a model for Bloch electrons in a magnetic field (in which case the frequency α is proportional to the magnetic flux; see below).

Another strong source of interest in almost periodic problems came from the 1984 discovery of quasicrystals by Shechtman et al. [48], as almost periodic Schrödinger operators provide elementary models for electronic properties in such media. Yet another motivating development occurred in the context of quantum chaos theory, notably in works of Fishman, Grempel, and Prange [23, 30, 31], as discrete one-dimensional Schrödinger operators with rich potentials (and, in particular, almost periodic ones) appeared in studies of dynamics of some elementary quantum models and, in particular, in studies aimed towards distinguishing quantum from classical dynamics in chaotic systems.

Interestingly, while the above developments in physics certainly contributed to broader interest in problems connected with exotic spectra, the timeline doesn't point to a simple scenario of mathematical physicists following the footsteps of physicists. Anderson localization was discovered in 1958 [1] and has been an active field of research by condensed matter physicists throughout the 1960's and 1970's, yet the Goldsheid–Molchanov–Pastur paper [25] came only in 1977. On the other hand, the “almost periodic flu” was in full motion *before* the above developments in physics that made it all the more interesting.

It is also interesting to note that Pearson's example [47] of a one-dimensional Schrödinger operator having purely singular continuous spectrum isn't directly connected with anything that has been of much interest to physics. The same is true for quite a few other models that have been considered in the context of exotic spectra. Thus, it appears that at least some of the exotic spectra era is associated with the willingness of spectral analysts, largely headed by Simon, to look at problems that are interesting from a mathematical perspective and to loosen some of the ties with physics. The fact that the “almost periodic flu” was very soon followed by discoveries that made it more interesting from a physics perspective than what one might have initially thought is an intriguing historical sidelight.

Another important phase in the field of exotic spectra came with Simon's “singular continuous spectrum revolution” which started around 1994. The core of this revolution has been the realization that singular continuous spectrum is a much more common phenomenon than what was previously thought and that it is, in fact, a generic phenomenon for broad classes of operators. In particular, the revolution supplied numerous new examples for operators with singular continuous spectrum. The revolution started with the three papers [19, 36, 61] by Simon and coauthors, along with Gordon's work [27, 28], which independently obtained roughly the same results as [19] a little earlier (also see the announcement paper [17], which summarizes the central findings of [19, 36, 61]). These papers soon inspired many more works by Simon and coauthors, as well as by others.

Alongside the central theme of identifying more and more operators with singular continuous spectrum, some of the focus (e.g., in [18, 44]) has also been on improving the understanding of singular continuous spectrum itself, namely, on issues such as the implications of such spectra to dynamical properties of quantum systems, natural ways of distinguishing different types of singular continuous spectrum, etc. While the main spike of the revolution can be timed, roughly, to the 1994–1996 period, related work continued throughout the 1990's and continues to this day.

Simon's work in the area of exotic spectra occupies roughly fifty-five research papers written in the 1981–1999 period. They make up a vast collection of results, and this article makes no attempt to achieve any sort of systematic coverage of all this work. Instead, we focus on a relatively small subset of these results, which tend to emphasize certain aspects of Simon's work and/or to fit in certain historical contexts. Many of Simon's important contributions to the subject have been left out.

The rest of this paper is organized as follows. In Section 2 we discuss the almost periodic flu, in Section 3 we review some of Simon's central contributions to thick point spectrum and in Section 4 we discuss the singular continuous spectrum revolution.

2. The Almost Periodic Flu

Simon’s 1982 review paper [54] starts with the following:

“In many years, flu sweeps the world. The actual strain varies from year to year; some years it has been Hong Kong flu, some years swine flu. In 1981, it was the almost periodic flu!”

The paper then moves to counting specific contributions by Avron and Simon [4, 3, 5, 6, 7], Bellissard and Testard [12], Bellissard et al. [9], Chulaevsky [13], Johnson [37], Moser [46], Johnson and Moser [38], and Sarnak [52].

This colorful opening led to the phenomenon of sudden broad interest in almost periodic Schrödinger operators becoming known as the “almost periodic flu” and to [54] becoming known as the “almost periodic flu paper.” The flu paper came in the midst of the actual flu season and was closely followed by additional important developments such as Kotani theory [41, 55] and the paper by Deift–Simon [16] that built upon it.

The first of Simon’s flu season papers was the Avron–Simon paper [3], which has already been mentioned above. The central theme of this paper has been to extend the scope of spectral analysis by making the following definitions of new spectral types. Given a self-adjoint operator A on a separable Hilbert space \mathcal{H} and $\psi \in \mathcal{H}$, we denote by μ_ψ the spectral measure for A and ψ . It is the unique Borel measure on \mathbb{R} obeying

$$\langle \psi, f(A)\psi \rangle = \int f(x) d\mu_\psi(x)$$

for any bounded Borel function f . We further denote by $\hat{\mu}_\psi$ the Fourier transform of μ_ψ , namely,

$$\hat{\mu}_\psi(t) \equiv \int e^{-itx} d\mu_\psi(x).$$

Avron–Simon [3] defined the *transient subspace*, \mathcal{H}_{tac} , by

$$\mathcal{H}_{\text{tac}} = \overline{\{\psi \mid \hat{\mu}_\psi \in L^1\}},$$

where $\overline{}$ denotes closure. This should be compared with the well-known fact that the absolutely continuous subspace, \mathcal{H}_{ac} , obeys $\mathcal{H}_{\text{ac}} = \overline{\{\psi \mid \hat{\mu}_\psi \in L^2\}}$. Since $|\hat{\mu}_\psi(t)|^2$ coincides with the quantum mechanical survival probability of ψ , \mathcal{H}_{tac} is a closed subspace of \mathcal{H}_{ac} which is made of vectors that have the fastest escape rate from their original position under the Schrödinger time evolution. The *recurrent subspace*, \mathcal{H}_{r} , is then defined by $\mathcal{H}_{\text{r}} = \mathcal{H}_{\text{tac}}^\perp$, and the *recurrent absolutely continuous subspace*, \mathcal{H}_{rac} , by $\mathcal{H}_{\text{rac}} = \mathcal{H}_{\text{tac}}^\perp \cap \mathcal{H}_{\text{ac}}$. Corresponding spectra are defined by $\sigma_{\text{tac}}(A) = \sigma(A \upharpoonright \mathcal{H}_{\text{tac}})$ and $\sigma_{\text{rac}}(A) = \sigma(A \upharpoonright \mathcal{H}_{\text{rac}})$. σ_{tac} is called *transient absolutely continuous spectrum* and σ_{rac} is called *recurrent absolutely continuous spectrum*.

Avron–Simon showed that $\mathcal{H}_{\text{tac}} = P_{\text{ei}}\mathcal{H}_{\text{ac}}$, where P_{ei} is the spectral projection on the essential interior of the essential support of the absolutely continuous part of the spectral measure class of A . This implies that the absolutely continuous spectrum of the free Laplacian (on $L^2(\mathbb{R}^d)$ or $\ell^2(\mathbb{Z}^d)$) and of similar problems such as periodic Schrödinger operators and atomic Hamiltonians is purely transient. If, however, A has nowhere dense spectrum, then $\sigma_{\text{tac}} = \emptyset$, and so any absolutely continuous spectrum of operators having Cantor spectrum is purely recurrent.

[3] also provided the following distinction between thin and thick point spectrum (σ_{pp} denotes the closure of the set of eigenvalues): $\lambda \in \sigma_{\text{pp}}$ is said to be in

the *thin point spectrum* if $(\lambda - \epsilon, \lambda + \epsilon) \cap \sigma_{pp}$ is countable for some $\epsilon > 0$. λ is said to be in the *thick point spectrum* if $(\lambda - \epsilon, \lambda + \epsilon) \cap \sigma_{pp}$ is uncountable for every $\epsilon > 0$. Avron–Simon showed that the thin point spectrum is countable and the thick point spectrum is a perfect set (namely, a closed set with no isolated points) which is empty if and only if σ_{pp} is countable.

Additional central flu season results by Simon and coauthors included the following:

The Avron–Simon paper [5] studied one-dimensional Schrödinger operators with limit periodic potentials, namely, potentials which are norm-limits of periodic potentials. A typical example (for the continuous Schrödinger operator case) is $V(x) = \sum_{j=1}^{\infty} a_j \cos(x/2^j)$ with $\sum_{j=1}^{\infty} |a_j| < \infty$. They considered the space of all such potentials with the ℓ^{∞} (namely, operator norm) topology and proved that for a dense G_{δ} set of such potentials the spectrum must be a Cantor set. Moreover, for a dense set, the spectrum must be both a Cantor set and purely absolutely continuous and it is thus purely recurrent absolutely continuous. Similar results also hold for some subsets of limit periodic potentials. We note that roughly the same results were independently obtained around the same time by Chulaevsky [13] and that a partial result (Cantor spectrum for a dense set) was also independently obtained by Moser [46].

The Avron–Simon paper [6] studied the almost Mathieu operator, which is the discrete Schrödinger operator $H_{\lambda, \alpha, \theta}$ on $\ell^2(\mathbb{Z})$ given by

$$(H_{\lambda, \alpha, \theta} \psi)(n) = \psi(n+1) + \psi(n-1) + \lambda \cos(2\pi \alpha n + \theta) \psi(n).$$

They showed, by utilizing a result of Gordon [26], that if $|\lambda| > 2$ and α is a Liouville number (namely, an irrational for which there is a sequence of rationals obeying $|\alpha - p_n/q_n| < n^{-q_n}$) then $H_{\lambda, \alpha, \theta}$ has purely singular continuous spectrum for Lebesgue a.e. θ . This provided a second concrete example (after Pearson's [47]) for a Schrödinger operator with purely singular continuous spectrum.

The Avron–Simon paper [7] and Craig–Simon paper [14] focused on the general theory of almost periodic Schrödinger operators and particularly on the spectrum, Lyapunov exponent and density of states. These papers prove many fundamental results (some of which were also independently obtained by others around the same time). Among the results are a rigorous proof of the Thouless formula $\gamma(E) = \int \ln |E - E'| dk(E')$, which connects the Lyapunov exponent $\gamma(E)$ with the density of states $dk(E)$, subharmonicity of the Lyapunov exponent and log-Hölder continuity of the density of states. [7] also had some results for the almost Mathieu operator, including a rigorous version of the Aubry duality [2] saying that for irrational α , $\sigma(\alpha, \lambda) = (\lambda/2)\sigma(\alpha, 4/\lambda)$, where $\sigma(\alpha, \lambda)$ is the spectrum of $H_{\lambda, \alpha, \theta}$ (it is independent of θ).

The Bellissard–Simon paper [11] provided the first rigorous result concerning Cantor spectrum for the almost Mathieu operator. They showed that $\sigma(\alpha, \lambda)$ is a Cantor set for a dense G_{δ} set of pairs (α, λ) . (It is known by now that the spectrum of $H_{\lambda, \alpha, \theta}$ is a Cantor set whenever α is irrational, as first conjectured in 1964 by Azbel [8].)

The Deift–Simon paper [16] built on Kotani theory [41], a version of which for discrete Schrödinger operators has been worked out by Simon in [55]. The Deift–Simon paper obtained several fundamental results for ergodic one-dimensional Schrödinger operators, most of which concern the absolutely continuous spectrum and its essential support, which, by Kotani theory, coincides with the set

$\{E \mid \gamma(E) = 0\}$. One of their results is the remarkable inequality (we give the versions of the results for discrete Schrödinger operators)

$$\frac{dk}{dE} \Big|_{\{E \mid \gamma(E)=0\}} \geq \frac{1}{2\pi \sin(\pi k)}$$

which says that the restriction of $\frac{dk}{dE}$ to the set $\{E \mid \gamma(E) = 0\}$ is bounded from below by the value of $\frac{dk}{dE}$ for the free Laplacian at corresponding values of the integrated density of states. It also implies

$$|\{E \mid \gamma(E) = 0\}| \leq 4,$$

where $|\cdot|$ denotes Lebesgue measure. Another result is the averaged boundedness of absolutely continuous spectrum eigenfunctions, more explicitly, the existence, for Lebesgue a.e. $E \in \{E \mid \gamma(E) = 0\}$, of two linearly independent solutions u_{\pm} of the corresponding Schrödinger equation, each of which is the complex conjugate of the other, normalized to have Wronskian $-2i$ and obeying

$$\lim_{L \rightarrow \infty} \frac{1}{L} \sum_{n=1}^L |u_{\pm}(n)|^2 \leq 2\pi \frac{dk}{dE}.$$

Yet another result of [16] is the mutual singularity of the singular parts of spectral measures for different realizations of an ergodic potential.

The paper which can be naturally considered as Simon’s last flu season paper is [56], which studies the Maryland model. This is the discrete Schrödinger operator $\tilde{H}_{\lambda,\alpha,\theta}$ on $\ell^2(\mathbb{Z})$ given by

$$(\tilde{H}_{\lambda,\alpha,\theta}\psi)(n) = \psi(n + 1) + \psi(n - 1) + \lambda \tan(\pi\alpha n + \theta)\psi(n).$$

The same name is also used for the multidimensional analog on $\ell^2(\mathbb{Z}^d)$ of this operator. $\tilde{H}_{\lambda,\alpha,\theta}$ was discovered by Fishman–Grepel–Prange [23, 30] (in the University of Maryland, from which it got its name) and has the remarkable property of being roughly precisely solvable. It exhibits singular continuous spectrum if α is a Liouville number and thick pure point spectrum (the spectrum as a set equals \mathbb{R} for any irrational α) with precisely computable eigenvalues if α is an irrational with typical Diophantine properties. The point spectrum result extends also to the multidimensional case. [56] extended and made rigorous some of the original results of the Maryland team. Similar work was done independently around the same time by Figotin–Pastur [21].

Aside from Simon being the leading worker on almost periodic spectral theory around the flu season (in terms of having the largest volume of results), he also had great impact on the field in terms of drawing the map for future progress. The flu paper [54] had a list of thirteen open problems and conjectures, of which five were devoted to the almost Mathieu operator and the rest were more general. Some of these were also repeated in Simon’s 1984 “*Fifteen problems in mathematical physics*” paper [57]. These problems helped to inspire a considerable amount of ongoing work in the 24 years that passed since [54]. The community of contributors to the field has been quite diverse, ranging from well-established world-class mathematicians to young entrants who were in various stages between starting kindergarten to finishing high school when [54] was written.

While probably no period since the spike of the flu season in 1981–1982 matched the level of almost periodic activity of that time, the overall progress made since

then is quite vast. In particular, most of the original open problems of [54] are solved by now (although certainly not all of them). Some of the conjectures turned out to be false! Considerable progress in the field continues as we write.

3. Thick Point Spectrum

Thick point spectrum tends to largely coincide with the phenomenon of Anderson localization [1]. The most central family of operators for which this phenomenon has been studied consists of various random multidimensional Schrödinger operators (with or without magnetic fields). Anderson's paper [1] gave some freedom regarding what may be rightfully called "the Anderson model." For simplicity of exposition, we use the name *Anderson model* for the (semi-concrete) random operator H_ω on $\ell^2(\mathbb{Z}^d)$ given by

$$H_\omega = \Delta + \sum_{n \in \mathbb{Z}^d} \lambda_n(\omega) \langle \delta_n, \cdot \rangle \delta_n,$$

where the $\lambda_n(\omega)$'s are independent, identically distributed random variables (i.i.d.r.v.) with uniform distribution in an interval and $\{\delta_n\}_{n \in \mathbb{Z}^d}$ is the natural basis of $\ell^2(\mathbb{Z}^d)$. Various variants and analogs have also been considered, including variants with different distributions, analogous continuous Schrödinger operators on $L^2(\mathbb{R}^d)$, discrete operators with more than just next-neighbor interactions, etc. While all of the results discussed below for "the Anderson model" are valid for more general families of random operators, the precise range of validity tends to be different for different results. We will thus keep things simple by discussing results almost exclusively for the above semi-concrete H_ω . We say that an Anderson model operator H_ω is "strongly coupled" if the interval over which the $\lambda_n(\omega)$'s are distributed is "large."

Anderson localization has been the first exotic spectral phenomenon to be discovered (both by physicists and in terms of proving rigorous mathematical results) and also the one which has drawn the most attention (the largest number of workers and papers) over the years. In a sense, one can say that the field of exotic spectra has been largely dominated by Anderson localization. Unlike almost periodic Schrödinger operators and the singular continuous spectrum revolution that we discuss in the next section, where the current state of those fields would be hard to imagine without Simon's crucial contributions, Simon's involvement in studies of Anderson localization was relatively minor and it is likely that the current state of the field would have been similar to what it is even without him. Nevertheless, Simon had quite a few papers and results involving point spectrum and Anderson localization and some of them are very important. The purpose of this section is to point out a small subset of these results.

Probably the most famous of Simon's contributions to Anderson localization is the 1986 Simon–Wolff criterion obtained by Simon–Wolff in [65]. It says that an Anderson model Hamiltonian H_ω has only point spectrum in an energy interval I , for a.e. ω , if and only if

$$\sum_{n \in \mathbb{Z}^d} |\langle \delta_n, (H_\omega - E - i0)^{-1} \delta_0 \rangle|^2 < \infty$$

for a.e. ω and Lebesgue a.e. $E \in I$. For this to hold, the distribution of the $\lambda_n(\omega)$'s need not be uniform in an interval, but it does need to be absolutely continuous with respect to the Lebesgue measure. The Simon–Wolff criterion thus reduces

the problem of establishing Anderson localization to proving appropriate square summability of resolvent matrix elements. This provides considerable simplification to many proofs of Anderson localization, and the Simon–Wolff criterion continues to play an important role in many settings of proving Anderson localization.

Another notable result is the absence of ballistic motion for point spectrum proven by Simon in [58]. It says that if H is a discrete Schrödinger operator on $\ell^2(\mathbb{Z}^d)$, X^2 is the squared position operator on this space, namely, $(X^2\psi)(n) = n^2\psi(n)$, and H has only point spectrum, then for any vector ψ in the domain of $|X|$,

$$\frac{1}{t^2} \langle \psi, e^{iHt} X^2 e^{-iHt} \psi \rangle \rightarrow 0$$

as $t \rightarrow \infty$. This shows that operators with purely point spectrum cannot induce ballistic propagation of initially localized wavepackets. (Note that ballistic motion is always an upper bound on the propagation rate of wavepackets for discrete Schrödinger operators, regardless of spectral properties.) [58] also obtained similar results for continuous Schrödinger operators on $L^2(\mathbb{R}^d)$, as long as the potential is sufficiently close to being bounded from below.

Yet another notable result was given in Simon’s paper [60] on cyclic vectors in the Anderson model. This paper established that for the Anderson model H_ω , with probability one, each of the δ_n vectors is a cyclic vector for the restriction of H_ω to its pure point subspace. In particular, this says that the δ_n vectors are cyclic vectors for H_ω whenever it has only point spectrum and that all of the point spectrum of H_ω must be simple (namely, the probability of H_ω having any degenerate eigenvalues is zero).

The final set of results we discuss in this section is from the work of del Rio–Jitomirskaya–Last–Simon [18]. They have shown, by constructing an explicit example, that from the fact that a Schrödinger operator has only point spectrum with exponentially localized eigenvectors, one cannot conclude anything for the growth rate of $\langle \psi, e^{iHt} X^2 e^{-iHt} \psi \rangle$ beyond the absence of strict ballistic motion as discussed above. That is, for such an operator one can still have

$$\limsup_{t \rightarrow \infty} \frac{1}{t^{2-\epsilon}} \langle \psi, e^{iHt} X^2 e^{-iHt} \psi \rangle = \infty$$

for any $\epsilon > 0$. Since the Anderson model is known to exhibit dynamical localization (namely, $\langle \psi, e^{iH_\omega t} X^2 e^{-iH_\omega t} \psi \rangle$ is bounded with probability one), this called for extending the understanding of its precise spectral characteristics. To achieve this goal, [18] introduced the following definitions: A self-adjoint operator A on $\ell^2(\mathbb{Z}^d)$ is said to have SULE (Semi-Uniformly Localized Eigenvectors) if it has only point spectrum and there exists a constant $\gamma > 0$ such that for any $b > 0$, there exists a constant $C(b) > 0$, such that for any eigenvector ψ_s of A one can find $n(s) \in \mathbb{Z}^d$, so that

$$|\langle \delta_k, \psi_s \rangle| \leq C(b) e^{b|n(s)| - \gamma|k - n(s)|}$$

for all $k \in \mathbb{Z}^d$. A is said to exhibit SUDL (Semi-Uniform Dynamical Localization) if there exists a constant $\tilde{\gamma} > 0$, such that for any $b > 0$, there exists a constant $\tilde{C}(b) > 0$, so that

$$\limsup_t |\langle \delta_n, e^{-itA} \delta_\ell \rangle| \leq \tilde{C}(b) e^{b|\ell| - \tilde{\gamma}|n - \ell|}.$$

[18] proved that SULE implies SUDL and that SUDL, along with simple spectrum, implies SULE. In particular, for operators on $\ell^2(\mathbb{Z}^d)$ with simple spectrum, SULE

\Leftrightarrow SUDL. It also proved that a sufficiently strongly coupled Anderson model H_ω exhibits SUDL and thus it also has SULE. Some additional results of [18] are discussed in Section 4 below.

4. The Singular Continuous Spectrum Revolution

Simon's singular continuous spectrum revolution started around 1994, when he discovered that singular continuous spectrum is a much more common phenomenon than what was previously thought, and proved results establishing it as a topologically generic phenomenon for many classes of operators. The initial results inspired much more work by Simon and coauthors, as well as by others, and within a few years the landscape concerning operators with singular continuous spectrum was drastically changed.

The most notable paper of the revolution is probably Simon's paper [61], which is sometimes called "the Wonderland paper," following how Simon named one of its central theorems. This paper focused on general self-adjoint operators and its central results provide conditions for topological families of operators to exhibit singular continuous spectrum for dense G_δ sets. The most central theorem of [61] is probably the following:

THEOREM 4.1 (Simon's Wonderland theorem). *Let X be a complete metric space of self-adjoint operators in which convergence implies strong resolvent convergence. Suppose*

- (a) $\{A \mid A \text{ has purely absolutely continuous spectrum}\}$ is dense in X ;
- (b) $\{A \mid A \text{ has purely point spectrum}\}$ is dense in X .

Then for a dense G_δ set of A 's, A has only singular continuous spectrum.

An illuminating corollary of Theorem 4.1 is that in a natural topological sense, a generic strongly coupled Anderson model operator has only singular continuous spectrum (as opposed to having only point spectrum with probability one). More precisely, if the exact same set of potentials of the Anderson model is considered with the natural product topology rather than as a probability space, then for a dense G_δ set of potentials in this topology, the corresponding operator has only singular continuous spectrum.

Another notable paper is the del Rio–Makarov–Simon paper [19], which studied the genericity of singular continuous spectrum in the context of rank one perturbations (see [59]). Its central result, which was also independently obtained a little earlier by Gordon [27, 28], is the following:

THEOREM 4.2. *Let A be a self-adjoint operator with a cyclic vector φ . Suppose $[a, b] \subset \sigma(A)$ and $\sigma_{ac}(A) \cap [a, b] = \emptyset$. Then for a dense G_δ set of λ 's, $A + \lambda\langle \varphi, \cdot \rangle \varphi$ has purely singular continuous spectrum on (a, b) .*

An important corollary of Theorem 4.2 (which also uses the cyclicity of δ_n vectors in the Anderson model discussed in Section 3) is that if one considers a typical realization of the strongly coupled Anderson model H_ω , which has only point spectrum, and continuously changes the value of the potential at a single point of \mathbb{Z}^d , then for a dense G_δ set of potential values at this point, the corresponding operator has only singular continuous spectrum. In particular, this says that the phenomenon of Anderson localization is extremely unstable, since the point spectrum can be converted into singular continuous spectrum by making arbitrarily small changes to the value of the potential at a single point.

The third paper of the trio which started the revolution is the Jitomirskaya–Simon paper [36], which studied genericity of singular continuous spectrum in the almost periodic context. They proved the following:

THEOREM 4.3. *For $|\lambda| > 2$ and any irrational α , there is a dense G_δ set of θ 's for which the almost Mathieu operator on $\ell^2(\mathbb{Z})$, given by*

$$(H_{\lambda,\alpha,\theta}\psi)(n) = \psi(n + 1) + \psi(n - 1) + \lambda \cos(2\pi\alpha n + \theta)\psi(n) ,$$

has only singular continuous spectrum.

Theorem 4.3 can be considered as an almost periodic analog of Theorem 4.2, since there are many other similarities between making rank one perturbations and changing realizations within the hull of one-dimensional almost periodic potentials. We note that the actual central result of [36] ensured the absence of eigenvalues, for an appropriate G_δ set in the hull, for any even almost periodic potential. The application to the almost Mathieu operator with $|\lambda| > 2$ is done in order to ensure the absence of absolutely continuous spectrum.

Another major paper of the revolution was the del Rio–Jitomirskaya–Last–Simon paper [18]. Some of its results were already discussed in Section 3. A central role in [18] was played by the fact that singular continuous spectra can be naturally decomposed into many spectral sub-types by using Hausdorff measures and dimensions. The measure-theoretic foundations for such decompositions go back at least to the works of Rogers–Taylor [50, 51] and they were introduced into spectral theory by Last [44], who was impacted by the singular continuous spectrum revolution along with Guarneri's seminal papers on quantum dynamics [32, 33] and Avron–Simon [3].

Recall that for any subset S of \mathbb{R} and $\alpha \in [0, 1]$, the α -dimensional Hausdorff measure, h^α , is given by

$$h^\alpha(S) \equiv \lim_{\delta \rightarrow 0} \inf_{\delta\text{-covers}} \sum_{\nu=1}^{\infty} |b_\nu|^\alpha ,$$

where a δ -cover is a cover of S by a countable collection of intervals, $S \subset \bigcup_{\nu=1}^{\infty} b_\nu$, such that for each ν the length of b_ν is at most δ . h^α is an outer measure on \mathbb{R} whose restriction to Borel sets is a Borel measure. h^1 coincides with the Lebesgue measure and h^0 is the counting measure, such that the family $\{h^\alpha \mid 0 \leq \alpha \leq 1\}$ can be viewed as a way of continuously interpolating between the counting measure and the Lebesgue measure. Given any nonempty set $S \subseteq \mathbb{R}$, there exists a unique $\alpha(S) \in [0, 1]$, called the Hausdorff dimension of S , such that $h^\alpha(S) = 0$ for any $\alpha > \alpha(S)$, and $h^\alpha(S) = \infty$ for any $\alpha < \alpha(S)$.

The following basic notions and facts stem from the Rogers–Taylor theory [50, 51]: Given α , a measure μ is called α -continuous (αc) if $\mu(S) = 0$ for every set S with $h^\alpha(S) = 0$. It is called α -singular (αs) if it is supported on some set S with $h^\alpha(S) = 0$. μ is said to be one-dimensional (αd) if it is α -continuous for every $\alpha < 1$. It is said to be zero-dimensional (αd) if it is α -singular for every $\alpha > 0$. A measure μ is said to have exact dimension α if, for every $\epsilon > 0$, it is both $(\alpha - \epsilon)$ -continuous and $(\alpha + \epsilon)$ -singular. Given a (positive, finite) measure μ and $\alpha \in [0, 1]$, one defines

$$D_\mu^\alpha(x) \equiv \limsup_{\epsilon \rightarrow 0} \frac{\mu((x - \epsilon, x + \epsilon))}{(2\epsilon)^\alpha}$$

and $T_\infty \equiv \{x \mid D_\mu^\alpha(x) = \infty\}$. The restriction $\mu(T_\infty \cap \cdot) \equiv \mu_{\alpha s}$ is α -singular, and $\mu((\mathbb{R} \setminus T_\infty) \cap \cdot) \equiv \mu_{\alpha c}$ is α -continuous. Thus, each measure decomposes uniquely into an α -continuous part and an α -singular part: $\mu = \mu_{\alpha c} + \mu_{\alpha s}$.

Consider now a separable Hilbert space \mathcal{H} and a self-adjoint operator H . By letting $\mathcal{H}_{\alpha c} \equiv \{\psi \mid \mu_\psi \text{ is } \alpha\text{-continuous}\}$ and $\mathcal{H}_{\alpha s} \equiv \{\psi \mid \mu_\psi \text{ is } \alpha\text{-singular}\}$, one obtains a decomposition $\mathcal{H} = \mathcal{H}_{\alpha c} \oplus \mathcal{H}_{\alpha s}$, of \mathcal{H} into mutually orthogonal closed subspaces which are invariant under H . The α -continuous spectrum ($\sigma_{\alpha c}$) and α -singular spectrum ($\sigma_{\alpha s}$) are then naturally defined as the spectra of the restrictions of H to the corresponding subspaces. Thus, the standard spectral theoretical scheme which uses the Lebesgue decomposition of a Borel measure into absolutely-continuous, singular-continuous, and pure-point parts can be extended to include further decompositions with respect to Hausdorff measures. As described in [44], the full picture is richer, and for every dimension $\alpha \in (0, 1)$, there is a natural unique decomposition (of finite Borel measures and thus also of \mathcal{H}) into five parts: one below the dimension α , one above it, and three within it—of which the middle one is absolutely-continuous with respect to h^α .

A major focus in [18] was rank one perturbations and attempting to understand the above discussed instability of Anderson localization under such small perturbations. While [18] showed that it is fully possible for a rank one perturbation to change the spectral type all the way from point spectrum to one-dimensional spectrum, it also established some spectral semi-stability for the Anderson model (in fact, for any self-adjoint operator on $\ell^2(\mathbb{Z}^d)$ that has SULE).

Explicitly, [18] showed that the singular continuous spectrum which is obtained in the Anderson model by changing the value of the potential at a single point must be purely zero-dimensional (namely, the spectral measures are supported on a set of zero Hausdorff dimension). This implies that if one focuses on the spectral dimension rather than on distinguishing point spectrum from continuous spectrum, the situation appears to be stable, since changing the value of the potential at a single point never changes the fact that the spectrum is zero-dimensional. This semi-stability result was given two proofs in [18]. One obtained it directly as a spectral result, while the other deduced it (using results of [44]) as a corollary of a dynamical result controlling the growth rate of $\langle \psi, e^{iHt} X^2 e^{-iHt} \psi \rangle$. That is, while $\langle \psi, e^{iHt} X^2 e^{-iHt} \psi \rangle$ cannot be bounded if H has continuous spectrum, [18] shows that if H is obtained from an operator with SULE by changing the value of the potential at a single point, then $\langle \psi, e^{iHt} X^2 e^{-iHt} \psi \rangle$ cannot grow with t faster than logarithmically.

Another interesting result in [18] is the fact that if an ergodic operator on $\ell^2(\mathbb{Z}^d)$ has what they called ULE (Uniformly Localized Eigenvectors), which means that all its eigenvectors can be fitted by shifting under a single exponential envelope, then this localization property is stable and there cannot be any potential in the support of the corresponding ergodic measure for which the spectrum is continuous. Thus, while the occurrence of singular continuous spectrum in the Anderson model is a zero-probability event, it is nevertheless saying something important about the probabilistic problem: With probability one, the Anderson model doesn't have ULE (and so its localization properties cannot be much stronger than SULE).

Another interesting paper of the revolution was Simon's paper [63], which studied, among other things, singular continuous spectrum of one-dimensional sparse barrier potentials. These include, in particular, the potential used by Pearson [47]

to construct the first explicit example with purely singular continuous spectrum. [63] proves that whenever such potentials are sufficiently sparse, the corresponding spectrum must be purely one-dimensional (namely, the spectral measure gives no weight to sets of Hausdorff dimension less than one). This roughly says that the singular continuous spectrum in Pearson's example [47] is one-dimensional. This should be contrasted with results of Last [44] and Jitomirskaya–Last [35] saying that all of the singular continuous spectrum which occurs for the almost Mathieu operator with coupling $|\lambda| > 2$ is purely zero-dimensional (this is roughly an almost periodic analog of the above discussed semi-stability result of [18] for the Anderson model). Thus, while the first concrete example of a Schrödinger operator with purely singular continuous spectrum had purely one-dimensional spectrum, the second example (of Avron–Simon [6]) had zero-dimensional spectrum.

To briefly mention a few more of Simon's revolution era results: Hof–Knill–Simon [34] established purely singular continuous spectrum for generic subsets in several families of discrete one-dimensional Schrödinger operators with potentials taking finitely many values. Simon–Stolz [64] established singular continuous spectrum for sufficiently sparse one-dimensional barrier potentials with growing barriers. Simon [62] proved the generic occurrence of purely singular continuous spectrum for certain topological families of graph Laplacians and Laplace–Beltrami operators. Gordon–Jitomirskaya–Last–Simon [29] established a new version of the Aubry duality for the almost Mathieu operator and proved that at the critical coupling $|\lambda| = 2$, it has purely singular continuous spectrum for Lebesgue a.e. pair α, θ .

Last–Simon [45] obtained some fundamental results for spectral analysis of one-dimensional Schrödinger operators, which yielded new proofs for many known results as well as important new results. Their results include characterizations of the absolutely continuous spectrum in terms of the behavior of eigenfunctions and transfer matrices, from which they also deduce that the absolutely continuous spectrum must be contained in the intersection of absolutely continuous spectra of a natural family of limiting operators. Among the applications of these results are a proof that the absolutely continuous spectrum of almost periodic Schrödinger operators is constant everywhere on the hull (which was also independently obtained around the same time by Kotani [42]), and the absence of absolutely continuous spectrum for strongly coupled (namely, any $|\lambda| > 2$) discrete potentials of the form $V(n) = \lambda \cos(n^\beta)$ with any $\beta > 1$.

Kiselev–Last–Simon [39] used modified Prüfer and EFGP transforms to study one-dimensional Schrödinger operators with decaying potentials. Among their results is a variant of Pearson's result [47] yielding purely singular continuous spectrum for suitable decaying sparse barrier potentials. They also obtain singular continuous spectrum with precisely computable fractional spectral Hausdorff dimensions for certain random decaying potentials.

References

- [1] P. W. Anderson, *Absence of diffusion in certain random lattices*, Phys. Rev. **109** (1958), 1492–1505
- [2] S. Aubry and G. André, *Analyticity breaking and Anderson localization in incommensurate lattices*, Ann. Israel Phys. Soc. **3** (1980), 133–164
- [3] J. Avron and B. Simon, *Transient and recurrent spectrum*, J. Funct. Anal. **43** (1981), 1–31

- [4] J. Avron and B. Simon, *Almost periodic Hill's equation and the rings of Saturn*, Phys. Rev. Lett. **46** (1981), 1166–1168
- [5] J. Avron and B. Simon, *Almost periodic Schrödinger operators, I. Limit periodic potentials*, Commun. Math. Phys. **82** (1982), 101–120
- [6] J. Avron and B. Simon, *Singular continuous spectrum for a class of almost periodic Jacobi matrices*, Bull. Amer. Math. Soc. **6** (1982), 81–85
- [7] J. Avron and B. Simon, *Almost periodic Schrödinger operators, II. The integrated density of states*, Duke Math. J. **50** (1983), 369–391
- [8] M. Ya. Azbel, *Energy spectrum of a conduction electron in a magnetic field*, Sov. Phys. JETP **19** (1964), 634–645
- [9] J. Bellissard, A. Formoso, R. Lima, and D. Testard, *Quasi-periodic interaction with a metal-insulator-transition*, Phys. Rev. **B 26** (1982), 3024–3030
- [10] J. Bellissard, R. Lima, and D. Testard, *A metal-insulator transition for the almost Mathieu model*, Commun. Math. Phys. **88** (1983), 207–234
- [11] J. Bellissard and B. Simon, *Cantor spectrum for the almost Mathieu equation*, J. Funct. Anal. **48** (1982), 408–419
- [12] J. Bellissard and D. Testard, preprint, 1980
- [13] V. Chulaevsky, *Perturbations of a Schrödinger operator with periodic potential*, (Russian) Uspekhi Mat. Nauk **36** (1981), 203–204
- [14] W. Craig and B. Simon, *Subharmonicity of the Lyapunov index*, Duke Math. J. **50** (1983), 551–560
- [15] H. L. Cycon, R. G. Froese, W. Kirsch, and B. Simon, *Schrödinger Operators with Application to Quantum Mechanics and Global Geometry*, Springer, Berlin, Heidelberg, 1987
- [16] P. Deift and B. Simon, *Almost periodic Schrödinger operators, III. The absolutely continuous spectrum in one dimension*, Commun. Math. Phys. **90** (1983), 389–411
- [17] R. del Rio, S. Jitomirskaya, N. Makarov, and B. Simon, *Singular continuous spectrum is generic*, Bull. Amer. Math. Soc. **31** (1994), 208–212
- [18] R. del Rio, S. Jitomirskaya, Y. Last, and B. Simon, *Operators with singular continuous spectrum, IV. Hausdorff dimensions, rank one perturbations, and localization*, J. d'Analyse Math. **69** (1996), 153–200
- [19] R. del Rio, N. Makarov, and B. Simon, *Operators with singular continuous spectrum: II. Rank one operators*, Commun. Math. Phys. **165** (1994), 59–67
- [20] E. Dinaburg and Ya. Sinai, *The one-dimensional Schrödinger equation with a quasi-periodic potential*, Funct. Anal. Appl. **9** (1975), 279–289
- [21] A. Figotin and L. Pastur, *An exactly solvable model of a multidimensional incommensurate structure*, Commun. Math. Phys. **95** (1984), 401–425
- [22] A. Figotin and L. Pastur, *The positivity of Lyapunov exponent and absence of the absolutely continuous spectrum for the almost-Mathieu equation*, J. Math. Phys. **25** (1984), 774–777
- [23] S. Fishman, D. R. Grempel, and R. E. Prange, *Chaos, quantum recurrences, and Anderson localization*, Phys. Rev. Lett. **49** (1982), 509–512
- [24] I. M. Gel'fand and B. M. Levitan, *On the determination of a differential equation from its spectral function*, (Russian) Izv. Akad. Nauk SSR. Ser. Mat. **15** (1951), 309–360; English transl. in Amer. Math. Soc. Transl. Ser. 2 **1** (1955), 253–304
- [25] I. Goldsheid, S. Molchanov, and L. Pastur, *A pure point spectrum of the stochastic one-dimensional Schrödinger equation*, Funct. Anal. Appl. **11** (1977), 1–10
- [26] A. Gordon, *On the point spectrum of one-dimensional Schrödinger operators*, Usp. Math. Nauk **31** (1976), 257–258
- [27] A. Ya. Gordon, *Exceptional values of the boundary phase for the Schrödinger equation on the semi-axis*, (Russian) Uspekhi Mat. Nauk **47** (1992), 211–212; English transl. in Russian Math. Surveys **47** (1992), 260–261
- [28] A. Ya. Gordon, *Pure point spectrum under 1-parameter perturbations and instability of Anderson localization*, Commun. Math. Phys. **164** (1994), 489–505
- [29] A. Gordon, S. Jitomirskaya, Y. Last, and B. Simon, *Duality and singular continuous spectrum in the almost Mathieu equation*, Acta Math. **178** (1997), 169–183
- [30] D. R. Grempel, S. Fishman, and R. E. Prange, *Localization in an incommensurate potential: An exactly solvable model*, Phys. Rev. Lett. **49** (1982), 833–836
- [31] D. R. Grempel and R. E. Prange, *Quantum dynamics of a nonintegrable system*, Phys. Rev. **A 29** (1984), 1639–1647

- [32] I. Guarneri, *Spectral properties of quantum diffusion on discrete lattices*, Europhys. Lett. **10** (1989), 95–100
- [33] I. Guarneri, *On an estimate concerning quantum diffusion in the presence of a fractal spectrum*, Europhys. Lett. **21** (1993), 729–733
- [34] A. Hof, O. Knill, and B. Simon, *Singular continuous spectrum for palindromic Schrödinger operators*, Commun. Math. Phys. **174** (1995), 149–159
- [35] S. Jitomirskaya and Y. Last, *Power law subordinacy and singular spectra, II. Line operators*, Commun. Math. Phys. **211** (2000), 643–658
- [36] S. Jitomirskaya and B. Simon, *Operators with singular continuous spectrum, III. Almost periodic Schrödinger operators*, Commun. Math. Phys. **165** (1994), 201–205
- [37] R. Johnson, preprints, 1980
- [38] R. Johnson and J. Moser, *The rotation number for almost periodic potentials*, Commun. Math. Phys. **84** (1982), 403–438
- [39] A. Kiselev, Y. Last, and B. Simon, *Modified Prüfer and EFGP transforms and the spectral analysis of one-dimensional Schrödinger operators*, Commun. Math. Phys. **194** (1998), 1–45
- [40] K. von Klitzing, G. Dorda, and M. Pepper, *New method for high-accuracy determination of the fine-structure constant based on quantized Hall resistance*, Phys. Rev. Lett. **45** (1980), 494–497
- [41] S. Kotani, *Ljaponov indices determine absolutely continuous spectra of stationary one-dimensional Schrödinger operators*, in Stochastic Analysis (edited by K. Ito), pp. 225–248, North-Holland, Amsterdam, 1984
- [42] S. Kotani, *Generalized Floquet theory for stationary Schrödinger operators in one dimension*, Chaos, Solitons & Fractals **8** (1997), 1817–1854
- [43] H. Kunz and B. Souillard, *Sur le spectre des operateurs aux différences finies aleatoires*, Commun. Math. Phys. **78** (1980), 201–246
- [44] Y. Last, *Quantum dynamics and decompositions of singular continuous spectra*, J. Funct. Anal. **142** (1996), 406–445
- [45] Y. Last and B. Simon, *Eigenfunctions, transfer matrices, and absolutely continuous spectrum of one-dimensional Schrödinger operators*, Invent. Math. **135** (1999), 329–367
- [46] J. Moser, *An example of a Schrödinger equation with almost periodic potential and nowhere dense spectrum*, Comment. Math. Helv. **56** (1981), 198–224
- [47] D. B. Pearson, *Singular continuous measures in scattering theory*, Commun. Math. Phys. **60** (1978), 13–36
- [48] D. Shechtman, I. Blech, D. Gratias, and J. W. Cahn, *Metallic phase with long-range orientational order and no translational symmetry*, Phys. Rev. Lett. **53** (1984), 1951–1953
- [49] M. Reed and B. Simon, *Methods of Modern Mathematical Physics, IV. Analysis of Operators*, Academic Press, New York, 1978
- [50] C. A. Rogers and S. J. Taylor, *The analysis of additive set functions in Euclidean space*, Acta Math. **101** (1959), 273–302
- [51] C. A. Rogers and S. J. Taylor, *Additive set functions in Euclidean space. II*, Acta Math. **109** (1963), 207–240
- [52] P. Sarnak, *Spectral behavior of quasiperiodic potentials*, Commun. Math. Phys. **84** (1982), 377–401
- [53] M. A. Shubin, *Spectral theory and the index of elliptic operators with almost-periodic coefficients*, (Russian) Uspekhi Mat. Nauk **34** (1979), 95–135
- [54] B. Simon, *Almost periodic Schrödinger operators: A review*, Adv. Appl. Math. **3** (1982), 463–490
- [55] B. Simon, *Kotani theory for one dimensional stochastic Jacobi matrices*, Commun. Math. Phys. **89** (1983), 227–234
- [56] B. Simon, *Almost periodic Schrödinger operators, IV. The Maryland model*, Ann. Phys. **159** (1985), 157–183
- [57] B. Simon, *Fifteen problems in mathematical physics*, in Perspectives in Mathematics. Anniversary of Oberwolfach 1984 (edited by W. Jäger, J. Moser, and R. Remmert), pp. 423–454, Birkhäuser, Basel, 1984
- [58] B. Simon, *Absence of ballistic motion*, Commun. Math. Phys. **134** (1990), 209–212
- [59] B. Simon, *Spectral analysis and rank one perturbations and applications*, CRM Lecture Notes Vol. 8 (J. Feldman, R. Froese, L. Rosen, eds.), pp. 109–149, Amer. Math. Soc., Providence, RI, 1995; reprinted in B. Simon, *Trace Ideals and Their Applications*, second

- edition, *Mathematical Surveys and Monographs*, Vol. 120, American Mathematical Society, Providence, RI, 2005
- [60] B. Simon, *Cyclic vectors in the Anderson model*, *Rev. Math. Phys.* **6** (1994), 1183–1185
 - [61] B. Simon, *Operators with singular continuous spectrum: I. General operators*, *Ann. of Math.* **141** (1995), 131–145
 - [62] B. Simon, *Operators with singular continuous spectrum, VI. Graph Laplacians and Laplace-Beltrami operators*, *Proc. Amer. Math. Soc.* **124** (1996), 1177–1182
 - [63] B. Simon, *Operators with singular continuous spectrum, VII. Examples with borderline time decay*, *Commun. Math. Phys.* **176** (1996), 713–722
 - [64] B. Simon and G. Stolz, *Operators with singular continuous spectrum, V. Sparse potentials*, *Proc. Amer. Math. Soc.* **124** (1996), 2073–2080
 - [65] B. Simon and T. Wolff, *Singular continuous spectrum under rank one perturbations and localization for random Hamiltonians*, *Commun. Pure Appl. Math.* **39** (1986), 75–90
 - [66] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, *Quantized Hall conductance in a two-dimensional periodic potential*, *Phys. Rev. Lett.* **49** (1982), 405–408

INSTITUTE OF MATHEMATICS, THE HEBREW UNIVERSITY, 91904 JERUSALEM, ISRAEL
E-mail address: ylast@math.huji.ac.il

Orthogonal Polynomials, Inverse Spectral Theory

Riemann–Hilbert Methods in the Theory of Orthogonal Polynomials

Percy Deift

*To Barry Simon, on his 60th birthday,
mathematician extraordinaire, teacher and friend*

ABSTRACT. In this paper we describe various applications of the Riemann–Hilbert method to the theory of orthogonal polynomials on the line and on the circle.

CONTENTS

1. Introduction
 2. Applications of (\mathbb{R}, ν) and (S^1, ν) : Identities, Equations and Formulae
 3. Applications of (\mathbb{R}, ν) and (S^1, ν) : Asymptotics
 4. Related Areas
- References

1. Introduction

In this paper $d\mu$ denotes either a Borel measure on \mathbb{R} with finite moments

$$\int_{\mathbb{R}} |x|^m d\mu(x) < \infty, \quad m = 0, 1, 2, \dots \quad (1)$$

or a finite Borel measure on the unit circle S^1

$$\int_{S^1} d\mu(\theta) < \infty. \quad (2)$$

In addition, unless stated explicitly otherwise, we will always assume that $d\mu$ is a *nontrivial probability measure*, i.e. $\text{supp}(d\mu)$ is infinite and the integral of $d\mu$ is 1.

Let

$$p_n(x) = k_n x^n + \dots, \quad k_n > 0, \quad n = 0, 1, 2, \dots \quad (3)$$

2000 *Mathematics Subject Classification.* 42C05, 30E05, 30E20, 35Q15.

Key words and phrases. Riemann–Hilbert problems, orthogonal polynomials, steepest-descent method.

$$\phi_n(z) = \kappa_n z^n + \dots, \quad \kappa_n > 0, \quad n = 0, 1, 2, \dots \tag{4}$$

denote the orthonormal polynomials (OP's) with respect to $d\mu$ on \mathbb{R} and S^1 respectively (see [60]),

$$\int_{\mathbb{R}} p_n(x)p_m(x)d\mu(x) = \int_{S^1} \overline{\phi_n(e^{i\theta})} \phi_m(e^{i\theta})d\mu(\theta) = \delta_{n,m}, \quad n, m \geq 0. \tag{5}$$

The fact that $d\mu$ is nontrivial implies, in particular, that the p_n 's and the ϕ_n 's, exist and are unique for all $n \geq 0$.

As is well known, the p_n 's satisfy a three-term recurrence relation

$$b_{n-1}p_{n-1}(x) + a_n p_n(x) + b_n p_{n+1}(x) = x p_n(x), \quad n \geq 0 \tag{6}$$

where

$$a_n \in \mathbb{R}, \quad b_n > 0, \quad n \geq 0 \tag{7}$$

and $b_{-1} \equiv 0$. Similarly the ϕ_n 's satisfy the Szegő recurrence relation

$$\sqrt{1 - |\alpha_n|^2} \phi_{n+1}(z) = z\phi_n(z) - \bar{\alpha}_n \phi_n^*(z), \quad n \geq 0 \tag{8}$$

where

$$\alpha_n \in \mathbb{C}, \quad |\alpha_n| < 1, \quad n \geq 0 \tag{9}$$

and for any polynomial $q(z)$ of degree n ,

$$q^*(z) \equiv z^n \overline{q(1/\bar{z})} \tag{10}$$

denotes the so-called *reverse polynomial*. Following [59], we call the α_n 's *Verblunsky coefficients*. A simple computation shows that

$$\alpha_n = -\frac{1}{\kappa_{n+1}} \overline{\phi_{n+1}(0)}, \quad n \geq 0. \tag{11}$$

On \mathbb{R} we define the $(n + 1) \times (n + 1)$ *Hankel determinant*

$$D_n = \det \left(\int_{\mathbb{R}} x^{j+k} d\mu(x) \right)_{0 \leq j, k \leq n}, \quad n \geq 0, \tag{12}$$

and on S^1 we similarly define the $(n + 1) \times (n + 1)$ *Toeplitz determinant*

$$\Delta_n = \det \left(\int_{S^1} e^{-i(j-k)\theta} d\mu(\theta) \right)_{0 \leq j, k \leq n}, \quad n \geq 0. \tag{13}$$

The determinants D_n and Δ_n are closely related to the OP's $\{p_n\}, \{\phi_n\}$ respectively. Indeed, one has (see, e.g., [60])

$$\frac{D_{n-1}}{D_n} = k_n^2, \quad \frac{\Delta_{n-1}}{\Delta_n} = \kappa_n^2, \quad n \geq 1. \tag{14}$$

Given $d\mu$, the study of the algebraic and asymptotic properties of the quantities

$$a_n, b_n, p_n(x), k_n, \alpha_n, \phi_n(z), \kappa_n,$$

and also

$$D_n \text{ and } \Delta_n,$$

constitutes the core of the classical theory of orthogonal polynomials.

The three-term relation (6) can be re-written in the form

$$Lp(z) = zp(z), \quad p(z) = (p_0(z), p_1(z), p_2(z), \dots)^T, \tag{15}$$

where L is an infinite Jacobi matrix, i.e. L is symmetric and tridiagonal

$$L = \begin{pmatrix} a_0 & b_0 & & & \\ b_0 & a_1 & b_1 & 0 & \\ & b_1 & a_2 & \ddots & \\ 0 & & \ddots & \ddots & \end{pmatrix} \tag{16}$$

with $b_i > 0, i \geq 0$. In the case that $d\mu$ has compact support on \mathbb{R} , the operator L is bounded on

$$\ell_2^+ = \left\{ u = (u_0, u_1, \dots)^T : \sum_{i=0}^{\infty} |u_i|^2 < \infty \right\}.$$

Let

$$F : \{d\mu \text{ on } \mathbb{R} : \text{supp}(d\mu) \text{ compact}\} \rightarrow \{\text{bounded Jacobi matrices on } \ell_2^+\}$$

denote the map taking $d\mu \mapsto L$. Conversely, if L is a bounded Jacobi matrix then, in particular, L is self-adjoint, and we let $d\mu$ denote the spectral measure associated with L in the cyclic subspace generated by L and e_0 , where $e_0 = (1, 0, 0, \dots)^T \in \ell_2^+$. Thus

$$\left(e_0, \frac{1}{L - \lambda} e_0 \right) = \int \frac{d\mu(x)}{x - \lambda}, \quad \lambda \in \mathbb{C} \setminus \mathbb{R} \tag{17}$$

and it follows further that $d\mu$ has compact support. Let

$$\hat{F} : \{\text{bounded Jacobi matrices on } \ell_2^+\} \rightarrow \{d\mu \text{ on } \mathbb{R} : \text{supp}(d\mu) \text{ compact}\}$$

denote the map taking L to $d\mu$. The basic fact of the matter (see, for example, [1], [58], and also [25]) is that F and \hat{F} are inverse to each other, $F \circ \hat{F} = id, \hat{F} \circ F = id$. From this point of view the (classical) orthogonal polynomial problem is the inverse spectral component of a spectral/inverse spectral problem. If the support of $d\mu$ is not compact, then the situation is similar, but the relation between $d\mu$ and L is more complicated because L is now an unbounded operator and we must distinguish between different self-adjoint extensions of L (see [1, 58] for more details).

In the case of measures $d\mu$ on the unit circle, the role of the Jacobi matrices is played by so-called CMV matrices C (see [59]). Such matrices C are unitary in ℓ_2^+ and pentadiagonal, and have the form

$$C = LM \tag{18}$$

where L and M are block diagonal

$$L = \text{diag}(\Theta_0, \Theta_2, \Theta_4, \dots), \quad M = \text{diag}(1, \Theta_1, \Theta_3, \dots) \tag{19}$$

with

$$\Theta_j = \begin{pmatrix} \bar{\alpha}_j & \rho_j \\ \rho_j & -\alpha_j \end{pmatrix}, \quad j \geq 0. \tag{20}$$

Here

$$|\alpha_j| < 1, \quad j \geq 0 \tag{21}$$

and

$$\rho_j = \sqrt{1 - |\alpha_j|^2}. \tag{22}$$

CMV matrices are named for Cantero, Moral and Velázquez [20], but in fact they appeared earlier in the literature (see, in particular, [67]). Let

$$\psi : \{d\mu \text{ on } S^1\} \rightarrow \{\text{CMV matrices}\}$$

denote the map taking $d\mu \rightarrow C$, the CMV matrix constructed from the Verblunsky coefficients $\alpha_j = \alpha_j(d\mu)$, $j \geq 0$, of $d\mu$, according to (18), (19) and (20). Conversely, given a CMV matrix C , let $d\mu$ be the spectral measure associated with C in the cyclic subspace generated by C , $C^* = C^{-1}$ and e_0 . Let

$$\hat{\psi} : \{\text{CMV matrices}\} \rightarrow \{d\mu \text{ on } S^1\}$$

denote the map taking C to $d\mu$. Then, as above (see [59]), ψ and $\hat{\psi}$ are inverse to each other, and we see again that the classical orthogonal polynomial problem on S^1 is the inverse spectral component of a spectral/inverse spectral problem.

The techniques used to analyze the direct spectral maps, \hat{F} and $\hat{\psi}$, are generally very different from the techniques used to analyze the inverse spectral maps, F or ψ , though sometimes there is some overlap (see, e.g., [31]). It is also interesting to note that in the solution of integrable systems, one needs knowledge of *both* \hat{F} and F (or $\hat{\psi}$ and ψ). For example, the Toda lattice induces a flow $L_0 \mapsto L = L(t)$ on Jacobi matrices ([43])

$$\begin{aligned} \frac{dL}{dt} &= B(L)L - LB(L) \\ L(t=0) &= L_0 \end{aligned} \tag{23}$$

where

$$L = \begin{pmatrix} a_0 & b_0 & & 0 \\ b_0 & a_1 & b_1 & \\ & b_1 & a_2 & \ddots \\ 0 & & \ddots & \ddots \end{pmatrix}, \quad B(L) = \begin{pmatrix} 0 & b_0 & & 0 \\ -b_0 & 0 & b_1 & \\ & -b_1 & 0 & \ddots \\ 0 & & \ddots & \ddots \end{pmatrix},$$

and the solution of (23) is given by the following well-known procedure ([56]):

$$L_0 \xrightarrow{\hat{F}} d\mu_0 = \hat{F}(L_0) \rightarrow d\mu_t(\lambda) = \frac{e^{2\lambda t} d\mu_0(\lambda)}{\int_{\mathbb{R}} e^{2xt} d\mu_0(x)} \xrightarrow{F} L(t) = F(d\mu_t).$$

The analysis of \hat{F} and $\hat{\psi}$ has benefited greatly from the powerful developments that have taken place over many years in the spectral theory of Schrödinger operators and their discrete analogs, reaching, over the last twenty years or so, and in the case of one dimension, a state of great precision. Here Barry Simon and his school have played a decisive role, and we refer the reader to [59], in particular, Part 2. The systematic analysis of F begins with the classic memoir of Stieltjes 1894–1895. Up till that point, a great deal of information had been obtained concerning particular polynomials, such as Legendre polynomials, Jacobi polynomials, Hermite polynomials, etc., but a unified point of view based on the orthogonality relation (5) had not yet emerged. The analysis of ψ began in 1920, when Szegö initiated the systematic study of polynomials orthogonal with respect to a measure on S^1 , as in (5). Szegö’s work in turn has led to many remarkable developments by researchers from all over the world, particularly the former USSR, Europe and the USA. We refer the reader to Simon’s book [59], where these developments are discussed in great detail together with many fascinating anecdotes concerning their discovery. Starting in the early 1950’s with the celebrated work of Gel’fand and Levitan, various techniques were developed to recover one-dimensional Schrödinger operators from their spectral measures. In the 1970’s, techniques based on the inverse-Schrödinger method (see [21] and [43]) started to play a role in the analysis of F and ψ .

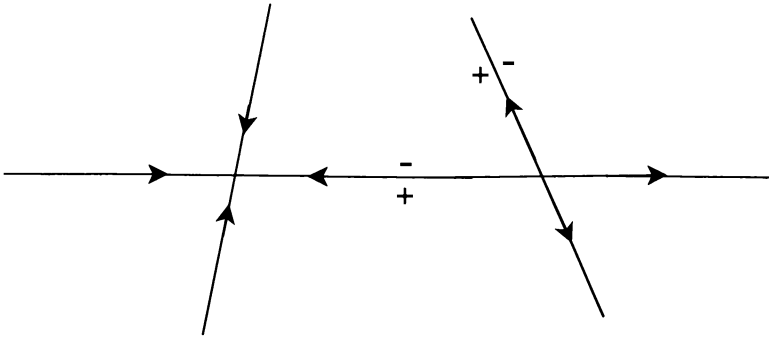


FIGURE 1. The contour Σ

The goal of this paper is to describe one of these techniques, which is different from the techniques in [21] or [43], and which has proved extremely fruitful, *viz.*, the Riemann–Hilbert (RH) method, also referred to as the Riemann–Hilbert Problem (RHP). The scope of the paper is limited to describing results for F and ψ obtained by RHP. Some of the results that we describe are quite standard and are included only for purposes of illustration. Other results, particularly asymptotic results, have been obtained, so far, only through RH methods. For a full up-to-date discussion of what is known about F and ψ , including the seminal contributions of Golinskii, Ismail, Khrushchev, Lubinsky, Nevai, Rakhmanov, Saff, Totik and many others, we again refer the reader to [58] and [59].

To begin, let Σ be an oriented contour in the complex plane \mathbb{C} (see Figure 1). By convention, if we move along the contour in the direction of the orientation, the (+)-side (resp. (–)-side) of the contour lies to the left (resp. right) (see again Figure 1). A $k \times k$ jump matrix v on Σ is a mapping from $\Sigma \rightarrow Gl(k, \mathbb{C})$ such that $v, v^{-1} \in L^\infty(\Sigma)$. We say that an $\ell \times k$ -valued matrix function $m(z)$ is a solution of the RHP (Σ, v) if

- (a) $m(z)$ is analytic in $\mathbb{C} \setminus \Sigma$
- (b) $m_+(z) = m_-(z)v(z), z \in \Sigma$, where $m_\pm(z) = \lim_{z' \rightarrow z, z \in (\pm)\text{-side}} m(z')$

If in addition $\ell = k$ and

- (c) $m(z) \rightarrow I$ as $z \rightarrow \infty$,

we say that m is a solution of the *normalized* RHP (Σ, v) .

Many technical issues arise. For example, in what sense do the limits m_\pm exist? In what sense does $m(z) \rightarrow I$ in (c)? How should one understand (b) at points of self-intersection in Σ ? Under what assumptions on Σ and v does a solution $m(z)$ exist? And if we normalize as in (c), is the solution unique? We will not consider such issues here and in the text that follows, and we simply refer the reader to [22] and the references therein for a general discussion of RHP’s (see also [41] for more recent information, and [10] for a discussion of points of self-intersections). In this paper we will consider almost exclusively problems with solutions $m(z)$ that are analytic in $\mathbb{C} \setminus \Sigma$ and continuous up to the boundary and at $z = \infty$. For such solutions, the limits in (b) and (c) are taken pointwise. Furthermore, for the problems we consider, the solution of the normalized RHP will always exist and be unique.

At the analytical level, a normalized RHP is equivalent to a problem for coupled singular integral equations on Σ . This is seen as follows.

Let C^Σ denote the Cauchy operator on Σ ,

$$C^\Sigma h(z) \equiv \int_\Sigma \frac{h(s)}{s-z} \frac{ds}{2\pi i}, \quad z \in \mathbb{C} \setminus \Sigma \tag{24}$$

with boundary values

$$(C^\Sigma_\pm h)(z) = \lim_{z' \in (\pm)\text{-side}} (C^\Sigma h)(z'), \quad z \in \Sigma. \tag{25}$$

Under reasonable conditions on Σ , $C^\Sigma_\pm \in \mathcal{L}(L^p(\Sigma))$, the bounded operators from $L^p(\Sigma) \rightarrow L^p(\Sigma)$, for any $1 < p < \infty$, and we have the relation

$$C^\Sigma_+ - C^\Sigma_- = 1. \tag{26}$$

Let

$$v(z) = (v_-(z))^{-1} v_+(z), \quad z \in \Sigma \tag{27}$$

be any pointwise factorization of v where

$$v_\pm, (v_\pm)^{-1} \in L^\infty(\Sigma). \tag{28}$$

Set

$$\begin{cases} w_+ = v_+ - I, & w_- = I - v_- \\ w = (w_+, w_-) \end{cases} \tag{29}$$

and define the singular integral operator on Σ

$$C^\Sigma_w h \equiv C^\Sigma_+(hw_-) + C^\Sigma_-(hw_+) \tag{30}$$

for row k -vectors h . As $w_\pm \in L^\infty(\Sigma)$, $C^\Sigma_w \in \mathcal{L}(L^p(\Sigma))$, $1 < p < \infty$. Suppose in addition that

$$w_\pm \in L^p(\Sigma) \text{ for some } 1 < p < \infty, \tag{31}$$

and consider the equation for a $k \times k$ -matrix function μ

$$(1 - C^\Sigma_w)\mu = I \tag{32}$$

in $I + L^p(\Sigma)$, or more precisely,

$$(1 - C^\Sigma_w)\nu = C^\Sigma_w I = C^\Sigma_+ w_- + C^\Sigma_- w_+ \in L^p(\Sigma) \tag{33}$$

where

$$\mu = I + \nu, \quad \nu \in L^p. \tag{34}$$

If a solution $\mu = I + \nu$ of (32)–(34) exists, set

$$m(z) = I + C^\Sigma(\mu(w_+ + w_-))(z). \tag{35}$$

Then a simple calculation shows that $m_\pm = \mu v_\pm$, and hence $m_+ = m_- v$, and as $m(z) \rightarrow I$ as $z \rightarrow \infty$, we see that (35) gives a solution of the normalized RHP (Σ, v) . Thus the normalized RHP (Σ, v) reduces to the analysis of the singular integral equations (32).

The connection between the OP problem and the RHP is due to Fokas, Its and Kitaev [44]. Let

$$P_n = \frac{1}{k_n} p_n = x^n + \dots, \quad n \geq 0 \tag{36}$$

denote the monic orthogonal polynomials associated with a measure

$$d\mu(x) = w(x)dx, \quad w(x) \geq 0 \tag{37}$$

absolutely continuous with respect to Lebesgue measure on \mathbb{R} , with $x^j w(x) \in H^1(\mathbb{R})$, the first Sobolev space, for all $j \geq 0$. Let $\Sigma = \mathbb{R}$, oriented from $-\infty$ to $+\infty$, and equipped with jump matrix

$$v = v(x) = \begin{pmatrix} 1 & w(x) \\ 0 & 1 \end{pmatrix}, \quad -\infty < x < \infty. \tag{38}$$

Finally, for any $n \geq 0$, let $X^{(n)} = (X_{ij}^{(n)})_{1 \leq i, j \leq 2}$ solve the RHP (\mathbb{R}, v)

$$\begin{aligned} X^{(n)}(z) & \text{ analytic in } \mathbb{C} \setminus \mathbb{R} \\ X_+^{(n)}(z) & = X_-^{(n)}(z)v(z), \quad z \in \mathbb{R} \end{aligned} \tag{39}$$

normalized so that

$$X^{(n)}(z) \begin{pmatrix} z^{-n} & 0 \\ 0 & z^n \end{pmatrix} \rightarrow I \text{ as } z \rightarrow \infty.$$

Then ([44], in addition see [25]) direct computation shows that

$$X^{(n)}(z) = \begin{pmatrix} P_n(z) & C(P_n w)(z) \\ -2\pi i k_{n-1}^2 P_{n-1}(z) & -2\pi i k_{n-1}^2 C(P_{n-1} w)(z) \end{pmatrix} \tag{40}$$

where $C = C^{\mathbb{R}}$ denotes the Cauchy operator on $\Sigma = \mathbb{R}$. In particular,

$$P_n(z) = X_{11}^{(n)}(z). \tag{41}$$

Furthermore, if $X_1^{(n)}$ denotes the residue of $X^{(n)} \begin{pmatrix} z^{-n} & 0 \\ 0 & z^n \end{pmatrix}$ at infinity,

$$X^{(n)}(z) \begin{pmatrix} z^{-n} & 0 \\ 0 & z^n \end{pmatrix} = I + \frac{X_1^{(n)}}{z} + O\left(\frac{1}{z^2}\right),$$

then

$$k_{n-1}^2 = -\frac{1}{2\pi i} (X_1^{(n)})_{21} \tag{42}$$

and in the notation of (6)

$$a_n = (X_1^{(n)})_{11} - (X_1^{(n+1)})_{11} \tag{43}$$

$$b_{n-1}^2 = (X_1^{(n)})_{12} (X_1^{(n+1)})_{21} \tag{44}$$

Also by (14) and (42),

$$\frac{D_{n-1}}{D_n} = -\frac{1}{2\pi i} (X_1^{(n+1)})_{21} \tag{45}$$

Thus all the basic quantities of interest in the OP problem can be read off from the solution $X^{(n)}$ of the RHP (\mathbb{R}, v) above.

On the unit circle, the situation is similar. Let

$$\Phi_n = \frac{1}{\kappa_n} \phi_n = z^n + \dots, \quad n \geq 0 \tag{46}$$

denote the monic orthogonal polynomials associated with a measure

$$d\mu(\theta) = \omega(\theta) \frac{d\theta}{2\pi} \tag{47}$$

absolutely continuous with respect to Lebesgue measure on S^1 with $\omega(\theta) \in H^1(S^1)$, $\omega(\theta) = \omega(\theta + 2\pi)$. Fix $n \geq 0$ and let $\Sigma = S^1$, oriented counterclockwise. Equip S^1 with the jump matrix

$$v = v(\theta) = \begin{pmatrix} 1 & \omega(\theta)z^{-n} \\ 0 & 1 \end{pmatrix}, \quad z = e^{i\theta} \tag{48}$$

and let $Y^{(n)} = (Y_{ij}^{(n)})_{1 \leq i, j \leq 2}$ solve the RHP (S^1, v)

- $Y^{(n)}(z)$ analytic in $\mathbb{C} \setminus S^1$ (49)

- $Y_+^{(n)}(z) = Y_-^{(n)}(z)v(\theta)$, $z = e^{i\theta} \in S^1$ (50)

normalized so that

- $Y^{(n)}(z) \begin{pmatrix} z^{-n} & 0 \\ 0 & z^n \end{pmatrix} \rightarrow I$ as $z \rightarrow \infty$. (51)

Then again (cf. [4]) direct computation shows that

$$Y^{(n)}(z) = \begin{pmatrix} \Phi_n(z) & C(\Phi_n \omega/S^n)(z) \\ -\kappa_{n-1}^2 \Phi_{n-1}^*(z) & -\kappa_{n-1}^2 C(\Phi_{n-1}^* \omega/S^n)(z) \end{pmatrix} \tag{52}$$

where $C = C^{S^1}$ denotes the Cauchy operator on S^1 and Φ_{n-1}^* is the reverse polynomial as in (10). In particular,

$$\Phi_n(z) = Y_{11}^{(n)}(z) \tag{53}$$

and hence by (11),

$$\alpha_{n-1} = -\overline{Y_{11}^{(n)}(z=0)}. \tag{54}$$

Also

$$\kappa_{n-1}^2 = -Y_{21}^{(n)}(z=0) \tag{55}$$

and hence

$$\frac{\Delta_{n-2}}{\Delta_{n-1}} = -Y_{21}^{(n)}(z=0). \tag{56}$$

Again we see that all basic quantities in the OP problem on the circle are expressed in terms of the solution $Y^{(n)}$ of the RHP (S^1, v) .

The outline of the paper is as follows. In Section 2 we show how to use the RHP's (\mathbb{R}, v) and (S^1, v) above to derive various identities, equations and formulae for the OP problem. In Section 3 we describe the application of the steepest descent method of Deift–Zhou for RHP's to asymptotic problems for OP's. Finally, in Section 4 we describe the application of RH ideas to areas related to the OP problem, such as random matrix theory, multi-orthogonal polynomials, orthogonal Laurent polynomials, and the rarefaction problem for the Toda lattice.

Technical Remark. In most of the paper we will be considering probability measures with some degree of smoothness as in (37) and (47) above. For such weights we then use the RHP's to derive, in particular, various identities such as (6), (8), (85), etc. If $d\mu(x)$ is an arbitrary probability measure on \mathbb{R} with finite moments, or $d\mu(\theta)$ is a probability measure on S^1 , we can approximate $d\mu(x)$ and $d\mu(\theta)$ appropriately with smooth measures $d\mu_\epsilon(x)$ and $d\mu_\epsilon(\theta)$ respectively: For such measures (6), (8), (85), etc., are true, and letting $\epsilon \downarrow 0$ we conclude that these identities are true, as they should be, for all measures $d\mu(x)$ and $d\mu(\theta)$ as above. Similar considerations apply at many points in the paper and we leave the details to the interested reader.

2. Applications of (\mathbb{R}, v) and (S^1, v) : Identities, Equations and Formulae

The applications of Riemann–Hilbert techniques to OP’s are principally of two types:

- (a) algebraic
- (b) asymptotic.

Under (a), the goal is to derive identities, equations and useful formulae for the OP problems. Under (b), the goal is to determine the asymptotic behavior of the OP’s p_n, P_n, ϕ_n, Φ_n as $n \rightarrow \infty$: Here one considers the case where the weight $\omega(x)$ is independent of n , as well as the case where $\omega(x)$ depends on n in a prescribed fashion (see (106) below). We consider (a) in this section, and (b) in the next.

Regarding (a), there is a general methodology, which may be traced all the way back to the original work of Gel’fand and Levitan, and which may be stated loosely as follows: If the jump matrix for a RHP is independent of a parameter, then differentiation with respect to that parameter (or taking differences in the discrete case) leads to an equation/identity.

We illustrate this methodology, which may be viewed as the analog for RHP’s of the celebrated theorem of Noether on conserved quantities for dynamical systems, first in the case of the defocusing Nonlinear Schrödinger Equation (NLS). In 1975 Shabat observed that the inverse scattering problem for the one-dimensional Schrödinger equation could be rephrased as a RHP. Because of the connection between Schrödinger operators and the Korteweg–de Vries (KdV) equation, this meant that KdV, and by extension all 1 + 1-dimensional integrable systems, could be solved by a RHP. In the case of defocusing NLS, Shabat’s observation amounts to the following (see, e.g., [41]). Let $q(x, t)$ be the solution of NLS on the line

$$\begin{cases} iq_t + q_{xx} - 2|q|^2q = 0 \\ q(x, t = 0) = q_0(x) \end{cases} \tag{57}$$

where $q_0(x) \rightarrow 0$ sufficiently rapidly as $|x| \rightarrow \infty$. Just as KdV is associated with the Schrödinger operator, NLS is associated with a first order, two-by-two scattering problem

$$\frac{d\psi}{dx} = i\frac{z}{2}\sigma_3\psi + \begin{pmatrix} 0 & iq \\ -i\bar{q} & 0 \end{pmatrix} \psi, \quad -\infty < x < \infty \tag{58}$$

where $\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ is the third Pauli matrix. Let $r = r(z)$ be the reflection coefficient for (58) with $q = q_0$. The map $\hat{R} : q \mapsto r$ is the analog for NLS of the OP maps \hat{F} and $\hat{\phi}$. Now, for fixed x and t , let $m = m(z; x, t)$ be the solution of the normalized RHP $(\mathbb{R}, v_{x,t})$ where \mathbb{R} is oriented from $-\infty$ to $+\infty$ and

$$\begin{cases} v_{x,t}(z) = \begin{pmatrix} 1 - |r(z)|^2 & re^{i\theta} \\ -\bar{r}e^{i\theta} & 1 \end{pmatrix}, & z \in \mathbb{R}. \\ \theta = xz - tz^2 \end{cases} \tag{59}$$

Let $m_1(x, t)$ be the residue of m at $z = \infty$,

$$m(z; x, t) = I + \frac{m_1(x, t)}{z} + O\left(\frac{1}{z^2}\right).$$

Then

$$q(x, t) = -i(m_1(x, t))_{12}. \tag{60}$$

How does one prove (60)? At the functorial level, \hat{R} is really a map from the category of differential operators to the category of RHP's,

$$L(q) \mapsto q \mapsto r \mapsto v_{x,t}$$

where $L(q) = i\sigma_3 \frac{d}{dx} + \begin{pmatrix} 0 & iq \\ -i\bar{q} & 0 \end{pmatrix}$, and so the key question becomes: "How is the differential operator encoded into the formalism of RHP's?"

To answer this question, observe that

$$\psi = \psi(z; x, t) \equiv m(z; x, t)e^{i\frac{g}{2}\sigma_3} \tag{61}$$

solves the RHP

- $\psi(z; x, t)$ analytic on $\mathbb{C} \setminus \mathbb{R}$
- $\psi_+ = \psi_- \begin{pmatrix} 1 - |r(z)|^2 & r(z) \\ -\bar{r}(z) & 1 \end{pmatrix}$, $z \in \mathbb{R}$

where the jump matrix is now independent of x and t . Differentiating with respect to x , we obtain

$$\psi_{x+} = \psi_{x-} \begin{pmatrix} 1 - |r(z)|^2 & r(z) \\ -\bar{r}(z) & 1 \end{pmatrix}$$

from which it follows that $T \equiv \psi_x \psi^{-1}$ has no jump across \mathbb{R} , and hence is entire. But as $z \rightarrow \infty$,

$$\begin{aligned} T &= m_x m^{-1} + m \frac{iz}{2} \sigma_3 m^{-1} \\ &= iz \frac{\sigma_3}{2} + A + O\left(\frac{1}{z}\right) \end{aligned}$$

for some constant matrix A . By Liouville, we must then have $T = iz \frac{\sigma_3}{2} + A$ or

$$\psi_x = iz \frac{\sigma_3}{2} \psi + A\psi. \tag{62}$$

Simple symmetry considerations imply that A is of the form $\begin{pmatrix} 0 & q \\ \bar{q} & 0 \end{pmatrix}$, and hence we recover the differential equation (58). Differentiating the ψ -RHP with respect to t yields similarly an equation of the form

$$\psi_t = B\psi \tag{63}$$

for some explicit matrix $B = B(z, q, q_x)$. Cross-differentiating (62) and (63), $(\psi_x)_t = (\psi_t)_x$, then yields the NLS equation (57). It is in this way in general that identities and differential relationships are encoded into the RHP.

To apply the above methodology to OP's, consider the solution $X^{(n)}$ of the RHP (\mathbb{R}, v) above. Observing that $X^{(n+1)}$ satisfies the same jump relation as $X^{(n)}$ across \mathbb{R} , we conclude as before that $T \equiv X^{(n+1)}(X^{(n)})^{-1}$ is entire. But

$$\begin{aligned} T &= X^{(n+1)}(z)(X^{(n)}(z))^{-1} \\ &= \left[\left(I + \frac{X_1^{(n+1)}}{z} + O\left(\frac{1}{z^2}\right) \right) z^{(n+1)\sigma_3} \right] \left[\left(I + \frac{X_1^{(n)}}{z} + O\left(\frac{1}{z^2}\right) \right) z^{n\sigma_3} \right]^{-1} \\ &= z \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + X_1^{(n+1)} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} X_1^{(n)} + O\left(\frac{1}{z}\right) \end{aligned}$$

and again by Liouville we conclude that

$$X^{(n+1)}(z) = \left(z \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + X_1^{(n+1)} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} X_1^{(n)} \right) X^{(n)}(z) \tag{64}$$

from which the three-term recurrence relation (6) now follows by a simple computation. Moreover, if we replace the weight $\omega(x)$ with $\omega_t(x) = \frac{e^{2xt}\omega(x)}{\int_{\mathbb{R}} e^{2st}\omega(s)ds}$, then

$$W^{(n)}(z; t) \equiv X^{(n)}(z; t)e^{(tz+g(t))\sigma_3}, \quad g(t) \equiv -\frac{1}{2} \ln \int_{\mathbb{R}} e^{2st}\omega(s)ds, \quad (65)$$

solves the RHP (\mathbb{R}, v) with jump matrix $v = \begin{pmatrix} 1 & \omega(x) \\ 0 & 1 \end{pmatrix}$ independent of t . Differentiating with respect to t , we obtain as above a differential equation for $W^{(n)}$

$$\frac{d}{dt}W^{(n)} = ((z + \dot{g})\sigma_3 + X_1^{(n)}\sigma_3 - \sigma_3X_1^{(n)})W^{(n)}.$$

Using Γ to denote the shift operator, $\Gamma W^{(n)} = W^{(n+1)}$, equation (64) takes the form

$$\Gamma W^{(n)} = \left(z \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + X_1^{(n+1)} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} X_1^{(n)} \right) W^{(n)}. \quad (66)$$

Cross-“differentiating” (65) and (66), $\frac{d}{dt}\Gamma W^{(n)} = \Gamma \frac{dW^{(n)}}{dt}$, one is led immediately to the Toda flow (23).

In another direction, if $\omega(x) = e^{-V(x)}$, $V(x) = \gamma_m x^{2m} + \dots$, $\gamma_m > 0$, then $U^{(n)} \equiv X^{(n)}e^{\frac{1}{2}V(x)\sigma_3}$ satisfies a jump relation across \mathbb{R} with jump matrix $v = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, which is independent of z , and by the above general methodology this leads to a differential equation for $U^{(n)}$ with respect to z , $\frac{dU^{(n)}}{dz} = DU^{(n)}$, for some explicit D . Cross-“differentiation”, $\frac{d}{dz}\Gamma U^{(n)} = \Gamma \frac{dU^{(n)}}{dz}$, then leads to so-called “string equations” for the recurrence coefficients a_n, b_n .

Applying the above methodology to the RHP (S^1, v) for OP’s on the unit circle, we obtain, in particular, simple and direct proofs of Szegő recurrence, Geronimus’ Theorem on the Schur iterates, and the Pinter–Nevai formula (see [59], and below). Indeed, let $Y^{(n)}$ solve the RHP (S^1, v) above. Then one observes that $V^{(n)} \equiv Y^{(n+1)} \begin{pmatrix} 1 & 0 \\ 0 & z \end{pmatrix}$ satisfies the same jump relation as $Y^{(n)}$ across S^1 ,

$$V_+^{(n)} = V_-^{(n)} \begin{pmatrix} 1 & \omega z^{-n} \\ 0 & 1 \end{pmatrix},$$

and hence $V^{(n)}(Y^{(n)})^{-1}$ is entire. As before, this leads to an equation for $V^{(n)}$ and $Y^{(n)}$, which takes the form

$$Y^{(n+1)} \begin{pmatrix} 1 & 0 \\ 0 & z \end{pmatrix} = V^{(n)} = \begin{pmatrix} z + \hat{a}_n & \hat{b}_n \\ \hat{c}_n & 1 \end{pmatrix} Y^{(n)} \quad (67)$$

for suitable constants $\hat{a}_n, \hat{b}_n, \hat{c}_n$. Furthermore $(\det Y^{(n)})_+ = (\det Y^{(n)})_- \det v = (\det Y^{(n)})_-$, and so $\det Y^{(n)}$ is entire. But $\det Y^{(n)} = \det \left(Y^{(n)} \begin{pmatrix} z^{-n} & 0 \\ 0 & z^n \end{pmatrix} \right) \rightarrow 1$ as $z \rightarrow \infty$, and hence $\det Y^{(n)} \equiv 1$. Taking determinants of both sides of (67), we find the relation

$$\hat{a}_n = \hat{b}_n \hat{c}_n. \quad (68)$$

From the first column of (67) we obtain the relations

$$\Phi_{n+1} = (z + \hat{a}_n)\Phi_n - \kappa_{n-1}^2 \hat{b}_n \Phi_{n-1}^* \quad (69)$$

$$-\kappa_n^2 \Phi_n^* = \hat{c}_n \Phi_n - \kappa_{n-1}^2 \Phi_{n-1}^*. \quad (70)$$

Eliminating Φ_{n-1}^* , we obtain the Szegő recurrence relation (8)

$$\Phi_{n+1} = z\Phi_n - \bar{\alpha}_n \Phi_n^* \quad (71)$$

with ϕ_n replaced by Φ_n , and with Verblunsky coefficient

$$\alpha_n = \kappa_n^2 \overline{\hat{b}_n}. \quad (72)$$

Letting $z \rightarrow \infty$ in (70), we find

$$\hat{c}_n = \kappa_n^2 \alpha_{n-1} \quad (73)$$

and hence by (68)

$$\hat{a}_n = \bar{\alpha}_n \alpha_{n-1}. \quad (74)$$

Now consider the second column in (67). Setting

$$r_n = C(\Phi_n \omega s^{-n}), \quad t_n = C(\Phi_n^* \omega s^{-n-1}) \quad (75)$$

and using (72), (73) and (74), we obtain as in (69) and (70)

$$zr_{n+1} = (z + \bar{\alpha}_n \alpha_{n-1})r_n - \bar{\alpha}_n \left(\frac{\kappa_{n-1}}{\kappa_n} \right)^2 t_{n-1} \quad (76)$$

$$-z\kappa_n^2 t_n = \kappa_n^2 \alpha_{n-1} r_n - \kappa_{n-1}^2 t_{n-1}. \quad (77)$$

Eliminating t_{n-1} as we eliminated Φ_{n-1}^* above, (76) and (77) reduce to

$$r_{n+1} = r_n - \bar{\alpha}_n t_n \quad (78)$$

$$zt_{n+1} = -\alpha_n r_n + t_n. \quad (79)$$

Defining

$$f_n \equiv t_n / r_n \quad (80)$$

and using (78) and (79), we obtain the recurrence relation

$$zf_{n+1} = \frac{f_n - \alpha_n}{1 - \bar{\alpha}_n f_n}, \quad n \geq 0. \quad (81)$$

In particular, for $z = 0$, we see that

$$\alpha_n = f_n(0) \quad (82)$$

and so (81) can be written in the form

$$zf_{n+1} = \frac{f_n - f_n(0)}{1 - f_n(0)f_n}, \quad n \geq 0. \quad (83)$$

Finally, observe that

$$f_0(z) = \frac{t_0}{r_0} = \frac{\int_{S^1} \frac{\pi_0^* \omega \frac{ds}{2\pi i s}}{s-z}}{\int_{S^1} \frac{\pi_0 \omega \frac{ds}{2\pi i}}{s-z}} = \frac{\int_{S^1} \frac{d\mu(\theta)}{s-z}}{\int_{S^1} s \frac{d\mu(\theta)}{s-z}}, \quad s = e^{i\theta} \quad (84)$$

where $d\mu(\theta) = \omega(\theta) \frac{d\theta}{2\pi}$.

Geronimus' Theorem (see [59]) states the following: Let

$$F(z) = \int_{S^1} \frac{s+z}{s-z} d\mu(\theta)$$

be the Carathéodory function for $d\mu$ and let $f_{\text{Schur}} \equiv \frac{1}{z} \frac{F(z)-1}{F(z)+1}$ be the associated Schur function. Let $(f_n)_{n \geq 0}$ solve the recurrence relation (83) with $f_n|_{n=0} = f_{\text{Schur}}$. Then

$$f_n(0) = \alpha_n, \quad n \geq 0$$

where $\{\alpha_n\}_{n \geq 0}$ are the Verblunsky coefficients for $d\mu$.

However, a simple computation shows that f_{Schur} is precisely f_0 in (84): Hence, using the general methodology for RHP’s as above, we have proved Geronimus’ Theorem. Moreover, we have the following formula for the Schur iterates:

$$f_n(z) = \frac{t_n}{r_n} = \frac{\int_{S^1} \frac{\Phi_n^* s^{-n}}{s-z} d\mu(\theta)}{\int_{S^1} \frac{\Phi_n s^{-n+1}}{s-z} d\mu(\theta)}, \quad n \geq 0 \tag{85}$$

which reduces simply, using (3.2.52) and (2.2.53) [59], to Golinskii’s formula ([59], Thm. 3.2.7).

Finally we note from [59], (1.3.79), together with the simple identity $\int \frac{s}{s-z} d\mu = 1 + z \int \frac{d\mu}{s-z}$, that

$$f_n = \frac{f_{\text{Schur}} B_{n-1} - A_{n-1}}{z B_{n-1}^* - z A_{n-1}^* f_{\text{Schur}}} = \frac{(B_{n-1} - z A_{n-1}) \int \frac{d\mu}{s-z} - A_n}{(z B_{n-1}^* - A_{n-1}^*) \int \frac{s d\mu}{s-z} + A_{n-1}^*} \tag{86}$$

where A_{n-1}, B_{n-1} are the Wall polynomials. But from (85), we obtain

$$f_n(z) = \frac{z^n \int_{S^1} \left(\frac{\Phi_n^*(s) s^{-n} - \Phi_n^*(z) z^{-n}}{s-z} \right) d\mu(\theta) + \Phi_n^*(z) \int \frac{d\mu(\theta)}{s-z}}{z^n \int_{S^1} \left(\frac{\Phi_n(s) s^{-n} - \Phi_n(z) z^{-n}}{s-z} \right) s d\mu(\theta) + \Phi_n(z) \int s \frac{d\mu(\theta)}{s-z}}.$$

Comparing with (86) we obtain

$$\Phi_n^*(z) = B_{n-1} - z A_{n-1} \tag{87}$$

or equivalently

$$\Phi_n(z) = z B_{n-1}^* - A_{n-1}^* \tag{88}$$

which is the Pinter–Nevai formula (see [59]) relating the OP’s to the Wall polynomials.

In addition to the formulae and identities obtained above for OP’s using the RHP’s (\mathbb{R}, ν) and (S^1, ν) , one can, using RHP’s closely related to (\mathbb{R}, ν) and (S^1, ν) , derive formulae for Toeplitz and Hankel determinants, or more precisely “relative” Toeplitz and Hankel determinants, that are particularly useful for asymptotic analysis. The asymptotic analysis of Toeplitz and Hankel determinants, dating back at least to the work of Szegö in 1915, is of considerable, and continuing, mathematical and physical interest, and we refer the reader to [9, 42] and the references therein for more information and recent results. The “relative” determinant formulae are as follows.

Let $\omega_1(x), \omega_2(x) \geq 0$ be two weights on \mathbb{R} and let $D_n(\omega_1 \omega_2), D_n(\omega_2)$ be the Hankel determinants associated with the measures $\omega_1(x) \omega_2(x) dx$ and $\omega_2(x) dx$ respectively. (Here we do not require $\omega_1 \omega_2 dx$ and $\omega_2 dx$ to be probability measures.) Then

$$\ln \frac{D_n(\omega_1 \omega_2)}{D_n(\omega_2)} = \int_0^1 dt \int_{\mathbb{R}} R_t(x) \left(\frac{d}{dt} \ln \omega_t(x) \right) dx \tag{89}$$

where $\omega_t = 1 - t + t \omega_1(x)$, $0 \leq t \leq 1$, and R_t is expressed in terms of the solution $X_t^{(n+1)} = ((X_t^{(n+1)})_{ij})_{1 \leq i, j \leq 2}$ of the RHP $(\mathbb{R}, \nu_t(x))$ in (39) with

$$\nu_t(x) = \begin{pmatrix} 1 & \omega_t(x) \omega_2(x) \\ 0 & 1 \end{pmatrix},$$

as follows:

$$R_t(x) = \frac{1}{2\pi i} \left((X_t^{(n+1)})_{11} (X_t^{(n+1)})'_{21} - (X_t^{(n+1)})'_{11} (X_t^{(n+1)})_{21} \right) \omega_t \omega_2. \tag{90}$$

Similarly, if $\omega_1(\theta), \omega_2(\theta) \geq 0$ are two weights on S^1 , with associated Toeplitz determinants $\Delta_n(\omega_1\omega_2), \Delta_n(\omega_2)$ respectively, then

$$\ln \frac{\Delta_n(\omega_1\omega_2)}{\Delta_n(\omega_2)} = \int_0^1 dt \int_{S^1} R_t(\theta) \frac{d}{dt} \ln \omega_t(\theta) \frac{d\theta}{2\pi} \tag{91}$$

where $\omega_t(\theta) = 1 - t + t\omega_1(\theta), 0 \leq t \leq 1$, and $R_t(\theta)$ is expressed in terms of the solution $Y_t^{(n+1)} = ((Y_t^{(n+1)})_{ij})_{1 \leq i, j \leq 2}$ of the RHP $(S^1, v_t(\theta))$ in (50) with

$$v_t(\theta) = \begin{pmatrix} 1 & \omega_t(\theta)\omega_2(\theta)z^{-(n+1)} \\ 0 & 1 \end{pmatrix}, \quad z = e^{i\theta},$$

as follows:

$$R_t(\theta) = \left((Y_t^{(n+1)})_{11} (Y_t^{(n+1)})'_{21} - (Y_t^{(n+1)})'_{11} (Y_t^{(n+1)})_{21} \right) \frac{\omega_t \omega_2}{z^n} \tag{92}$$

where $' \equiv \frac{d}{dz}$.

The functions $R_t(x), R_t(\theta)$ have the interpretation as 1-point functions

$$R_t(x) = (n+1) \int_{x_i \in \mathbb{R}, 1 \leq i \leq n} d\mu(x, x_1, x_2, \dots, x_n) \tag{93}$$

$$R_t(\theta) = (n+1) \int_{\theta_i \in S^1, 1 \leq i \leq n} d\mu(\theta, \theta_1, \dots, \theta_n) \tag{94}$$

for the random particle ensembles (see [55]) with distributions

$$d\mu(x_0, x_1, \dots, x_n) = (1/Z_{\mathbb{R}}) \prod_{0 \leq j < k \leq n} (x_i - x_j)^2 \prod_{j=0}^n (\omega_t \omega_2)(x_j) dx_0 dx_1 \dots dx_n \tag{95}$$

and

$$d\mu(\theta_0, \theta_1, \dots, \theta_n) = (1/Z_{S^1}) \prod_{0 \leq i < j \leq n} |e^{i\theta_j} - e^{i\theta_k}|^2 \prod_{j=0}^n (\omega_t \omega_2)(\theta_j) d\theta_0 d\theta_1 \dots d\theta_n \tag{96}$$

where $Z_{\mathbb{R}}, Z_{S^1}$ are normalization constants.

Note that on S^1 we can set $\omega_2 = 1$, so that $\Delta_n(\omega_2) = 1$ and (91) gives us a formula, first derived in [24], purely for $\Delta_n(\omega_1)$. In the non-compact situation on \mathbb{R} , this clearly cannot be done and we must always work with relative determinants as in (89).

Formulae (89) and (91) are due to Deift [26], and may be proved by generalizing the proof of (91) given in [24] for the case $\omega_2 = 1$. A key ingredient in the proof is the notion of an *integrable operator*: If Σ is an oriented contour in \mathbb{C} , we say that an operator K acting on $L^p(\Sigma), 1 < p < \infty$, is *integrable* if it has a kernel of the form

$$K(z, z') = \frac{\sum_{j=1}^{\ell} f_j(z)g_j(z')}{z - z'}, \quad z, z' \in \Sigma \tag{97}$$

for some functions $f_j, g_k \in L^\infty(\Sigma), 1 \leq j, k \leq \ell$. Special examples of integrable operators appeared in the 1960's in the work of McCoy, Tracy and others, and elements of the general theory were discovered by Sakhnovich in the late 60's, but the full general theory of such operators is due to Its, Izegin, Korepin and Slavnov [45] in 1990. Integrable operators have many useful properties (see, e.g., [24]). In particular, if K is integrable as in (97) above, then so is $(1 - K)^{-1} - 1$,

$$(1 - K)^{-1} = 1 + \frac{\sum_{j=1}^{\ell} F_j(z)G_j(z')}{z - z'}$$

for suitable $F_j, G_k, 1 \leq j, k \leq l$. Furthermore, quite remarkably, the functions $F = (F_1, \dots, F_l)^T, G = (G_1, \dots, G_l)^T$ can be computed in terms of the solution of a canonical, auxiliary RHP. Indeed, define the jump matrix $v = I - 2\pi i f g^T$ on Σ , where $f = (f_1, \dots, f_l)^T, g = (g_1, \dots, g_l)^T$, and assume for simplicity that $\sum_{j=1}^l f_j(z)g_j(z) = 0, z \in \Sigma$. Then, if m solves the normalized RHP (Σ, v) , we have

$$F = m_{\pm} f \quad \text{and} \quad G = (m_{\pm}^T)^{-1} g. \tag{98}$$

The proofs of (89) and (91) proceed by expressing the relative determinants $\frac{D_n(\omega_1\omega_2)}{D_n(\omega_2)}, \frac{\Delta_n(\omega_1\omega_2)}{\Delta_n(\omega_2)}$ in terms of Fredholm determinants of integrable operators K ,

$$\begin{aligned} \ln \det(1 - K) &= \int_0^1 \frac{d}{dt} \ln \det(1 - tK) \\ &= - \int_0^1 \text{tr} \left(\frac{1}{1 - tK} K \right) dt \end{aligned}$$

and then using (98) to express $(1 - tK)^{-1}K = ((1 - tK)^{-1} - 1)/t$ in terms of the solution of the auxiliary RHP associated to tK . We shall say more about (89) and (91) in what follows.

3. Applications of (\mathbb{R}, v) and (S^1, v) : Asymptotics

In this section we consider the asymptotics of OP's, denoted (b) in Section 2. In Section 2, the goal was to show how a variety of identities, equations and formulae, mostly classical and well-known, follow from a single, basic methodology in RHP's. Here the goal is to describe new results on the asymptotics of OP's that follow from the RH method, utilizing in particular the non-linear, non-commutative, steepest descent method introduced in [38] in 1993. Although much was known (see [60]) about the detailed asymptotic behavior of classical OP's, like Hermite, Laguerre, Jacobi polynomials, etc., both on and off the contour of orthogonality, little was known about the detailed asymptotics of OP's with respect to general weights. The main tool that makes possible the detailed analysis of the asymptotics of classical OP's is the existence of integral representations for these polynomials, to which the classical method of steepest descent can be applied (see, e.g., [60, Section 8.71]). For general weights, one may view the RHP's $(\mathbb{R}, v(x))$ and $(S^1, v(\theta))$ as non-commutative analogs of these integral representations, with the non-commutative steepest descent method now playing the role of the classical steepest descent method.

We now describe the steepest descent method for RHP's in broad outline. (Unfortunately, we do not have sufficient space in this article to describe the method in detail.) In the case of NLS (cf. (59) and (60)), we write the solution $q(x, t)$ of the Cauchy problem for NLS as a functional f , say, of the data $re^{i\theta}$,

$$q(x, t) = f(re^{i\theta}). \tag{99}$$

From (35) and (60) we see that

$$f(re^{i\theta}) = \left(\int_{\mathbb{R}} \mu(s; x, t) (\omega_+ + \omega_-) \frac{ds}{2\pi} \right)_{12}. \tag{100}$$

Using the factorization

$$v_{x,t} = \begin{pmatrix} 1 & -re^{i\theta} \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 \\ -\bar{r}e^{-\theta} & 1 \end{pmatrix} \tag{101}$$

(cf. (27)), so that

$$w_+ = \begin{pmatrix} 0 & 0 \\ -\bar{r}e^{-\theta} & 0 \end{pmatrix}, \quad w_- = \begin{pmatrix} 0 & -\bar{r}e^{-i\theta} \\ 0 & 0 \end{pmatrix} \tag{102}$$

we obtain

$$q(x,t) = \left(\int_{\mathbb{R}} ((I - C_{\omega}^{\mathbb{R}})^{-1} I)(\omega_+ + \omega_-) \frac{ds}{2\pi} \right)_{12}. \tag{103}$$

For r “small,” we have

$$\begin{aligned} q(x,t) &= \left(\int_{\mathbb{R}} ((I + \text{“small”})(\omega_+ + \omega_-) \frac{ds}{2\pi} \right)_{12} \\ &= \int_{\mathbb{R}} r(s) e^{i(xs - ts^2)} \frac{ds}{2\pi} + \text{“small”} \end{aligned}$$

indicating that the classical steepest descent method can be applied as $t \rightarrow \infty$. However, when r is no longer “small,” we see from the non-linear dependence of $q(x,t)$ on r in (103), and from the matrix nature of the problem, that a non-linear, non-commutative version of the steepest descent method is required, and this is the kind of method that was introduced in [38]. In the classical steepest descent method, the integral localizes as $t \rightarrow \infty$ to a small neighborhood of the stationary phase point(s), $\theta'(z_0) = 0$, $z_0 = x/2t$ in the case of NLS, and an explicit asymptotic formula for the solution is then obtained by evaluating a Gaussian integral: in the fully non-linear case (see [39] [41]) the RHP $(\mathbb{R}, v_{x,t})$ localizes to a RHP in the neighborhood of the stationary phase point $z_0 = x/2t$, and an asymptotic form for the solution

$$q(x,t) \sim \frac{1}{t^{1/2}} \alpha(z_0) e^{i(tz_0^2 - \beta(z_0) \ln t)} \tag{104}$$

is then obtained by solving this local RHP explicitly (in terms of parabolic cylinder functions, as it turns out). The asymptotic form (104) was first obtained by Zakharov and Manakov [68] by other means. In situations where there is more than one stationary phase point, for example, for MKdV, where $\theta = xz + 4tz^3$ with stationary phase points $\pm z_0 = \pm \sqrt{-x/12t}$, the long-time behavior of solutions of MKdV (see [38]) is a superposition of NLS-like contributions from $+z_0$ and $-z_0$, as long as these points remain separated, i.e., $\frac{-x}{t} > c > 0$. However, in the space-time region where $-x/12t \rightarrow 0$, and hence $+z_0 \rightarrow -z_0$, one is in a non-linear “caustic” region which is manifested by the solution taking the form of a self-similar oscillation, $q(x,t) \sim \frac{1}{(3t)^{1/3}} u(x/(3t)^{1/3})$, where u is a solution of the Painlevé II equation $u''(t) = tu + 2u^3$ (see [38]).

Up till this point, the RH asymptotic theory proceeded as a non-linear analog of the classical steepest descent method in which all the phenomena that arose could be viewed as non-linear counterparts of phenomena that had already arisen in the linear, scalar situation. However, with the analysis of the collisionless shock region for KdV (see [36, 40]), and the analysis of the asymptotic behavior of solutions of the Painlevé II equation, it began to be clear that there were phenomena

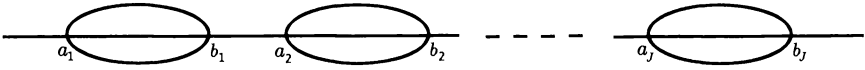


FIGURE 2. The contour $\hat{\Sigma}$

inherent in the non-linear steepest descent method that had no analog in the classical situation. Most importantly, it became clear that instead of stationary phase points, one could have “stationary phase lines” in which case all the points on some interval in \mathbb{C} contributed equally to the asymptotic behavior of the solution of the problem. Moreover, in place of modulated linear oscillations as in (104), one would now have genuinely non-linear oscillations described in terms of Jacobi’s sn and cn functions, etc. A systematic extension of the steepest descent method to allow for such “stationary phase lines” and genuinely non-linear oscillations was presented by Deift, Venakides and Zhou [37] in the context of their work on the zero dispersion problem. Soon thereafter, using the methods in [37] together with recent developments in the theory of logarithmic potentials with external fields (see [57], and also [32]), the authors in [34] derived so-called Plancherel–Rotach asymptotics for OP’s with measures of the form

$$e^{-V(x)} dx, \quad V(x) = \gamma x^{2m} + \delta x^{2m-1} + \dots \quad \gamma > 0, \tag{105}$$

and in [33], for measures of the form

$$e^{-nQ(x)} dx, \quad Q(x)/\ln|x| \rightarrow +\infty \quad \text{as } |x| \rightarrow \infty, \tag{106}$$

where $Q(x)$ is real analytic on \mathbb{R} . As described in [34] one obtains as $n \rightarrow \infty$ precise pointwise asymptotics for the OP’s $P_n(z)$ for all $z \in \mathbb{C}$, as well as detailed asymptotics for a_n, b_n, γ_n and the zeros of $p_n(z)$. In the special case $e^{-n(x^4 - tx^2)} dx$, Bleher and Its [12] obtained asymptotics for the associated OP’s using RH techniques and a mixture of steepest descent/isomonodromy ideas.

In broad outline the method proceeds as follows. For weights $e^{-V(x)}$ as above one first scales $x \rightarrow xn^{1/2m}$ so that $e^{-V(x)} \rightarrow e^{-nV_n(x)}$, where $V_n(x) = \gamma x^{2m} + \frac{\delta}{n^{1/2m}} x^{2m-1} + \dots$. Next, one considers the so-called equilibrium measure $d\mu_{\text{eq}}$ for the logarithmic potential problem associated with OP’s (see [57]). By [32], for weights $e^{-nV_n(x)}$ or $e^{-nQ(x)}$ as above, $d\mu_{\text{eq}}$ is supported on a finite union of disjoint intervals $\cup_{i=1}^J (a_i, b_i)$, $J < \infty$ (in the case $e^{-nV_n(x)}$, $J = 1$). Next one introduces the so-called “ g ” function, $g(z) \equiv \int_{\mathbb{R}} \ln(z-s) d\mu_{\text{eq}}(s) \sim \ln z$ as $z \rightarrow \infty$. Along with $d\mu_{\text{eq}}$, the logarithmic potential problem also produces a Lagrange multiplier ℓ , and we set $\tilde{X}^{(n)} \equiv e^{\frac{n\ell}{2}\sigma_3} X^{(n)}(z) e^{-ng(z)\sigma_3} e^{-\frac{n\ell}{2}\sigma_3}$. One observes that $\tilde{X}^{(n)}$ now solves a *normalized* RHP (\mathbb{R}, \tilde{v}) for some explicit jump matrix \tilde{v} . In the key step, the RHP for $\tilde{X}^{(n)}$ is now deformed to a RHP on a contour $\hat{\Sigma}$ of the form shown in Figure 2. By the properties of $g(z)$, or more properly, the properties of $d\mu_{\text{eq}}$, it turns out that as $n \rightarrow \infty$, \hat{v} , the jump matrix for the deformed RHP on $\hat{\Sigma}$, converges

$$\hat{v}(z) \rightarrow I \tag{107}$$

exponentially for all $z \in \hat{\Sigma} \setminus \cup_{i=1}^J [a_i, b_i]$. Thus as $n \rightarrow \infty$, the RHP reduces to a limiting RHP on the union of intervals $\cup_{i=1}^J [a_i, b_i]$. On each of the intervals (a_i, b_i) , $\hat{v}(z)$ has the simple form $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and this limiting RHP can be solved explicitly in terms of the function theory on the hyper-elliptic Riemann surface obtained

by gluing together two copies of $\mathbb{C} \setminus \cup_{i=1}^J (a_i, b_i)$ in the standard way. However, the convergence rate in (107) is not uniform, becoming slower and slower as z approaches the end points a_i, b_i . The natural topology for RHP's is convergence for the coefficients of \hat{v} in $L^p \cap L^\infty(\hat{\Sigma})$ (cf. (30)/(33)), and the lack of uniform convergence in (107) constitutes the major technical difficulty in implementing the steepest descent method as described above. We refer the reader to [33, 34] for more details.

We now consider the relative determinant formulae (89) and (91) and their associated RHP's $(\mathbb{R}, v_t(x))$ and $(S^1, v_t(\theta))$ respectively. The celebrated strong Szegő limit theorem, in the definitive form due to Ibragimov (see [59] for many proofs and much historical discussion) states that if $d\mu(\theta) = e^{-V(\theta)} \frac{d\theta}{2\pi}$, and $V(\theta)$ has Fourier coefficients $\{\hat{V}_k\}$ satisfying $\sum_{k=1}^\infty k|\hat{V}_k|^2 < \infty$, then as $n \rightarrow \infty$

$$\ln \Delta_n(e^{-V}) = (n + 1)\hat{V}_0 + \sum_{k=1}^\infty k|\hat{V}_k|^2 + o(1). \tag{108}$$

In addition to the many proofs in [59], (108) can also be proved, under certain additional smoothness assumptions on $V(\theta)$, by applying the steepest descent method to the RHP's $(S^1, v_t(\theta))$, $0 < t < 1$. The situation is simpler than in [33, 34], but the argument in this situation is particularly illustrative of the emergence of a “stationary phase line”; details are given in [24]. There is also a version of the strong Szegő limit theorem for block Toeplitz determinants (see [64, 65], and also [18] for more recent results). In the block Toeplitz case, the analog of (108) contains a certain Fredholm determinant which is difficult to evaluate in elementary terms. In certain cases the method in [24] extends to the block Toeplitz case and, quite surprisingly, the term corresponding to this Fredholm determinant is evaluated automatically (see [46]).

In [7] (see also [8]) the authors state the following analog of the Szegő strong limit theorem for the case of Hankel matrices. Let $\omega_2 = e^{-x^2}$ and let $\omega_1(x) > 0$ have the property that $\omega_1(x) \rightarrow 1$ sufficiently rapidly as $|x| \rightarrow \infty$. Then as $n \rightarrow \infty$,

$$\ln \frac{D_n(\omega_1\omega_2)}{D_n(\omega_2)} = \frac{\sqrt{2(n+1)}}{\pi} \int_{\mathbb{R}} \ln \omega_1(x) dx + \frac{1}{4\pi} \int_{\mathbb{R}} |k| |\hat{f}(k)|^2 dk + o(1) \tag{109}$$

where $\hat{f}(k) = \frac{1}{\sqrt{2\pi}} \int (\ln \omega_1(x)) e^{-ikx} dx$. Using (89), this result can also be proved ([27]) using the steepest descent method, not only for $\omega_2 = e^{-x^2}$, but also for more general weights, $\omega_2 = e^{-V(x)}$, $V(x) = \gamma x^{2m} + \dots$, $\gamma > 0$, as above.

Riemann–Hilbert techniques and the RH method are useful not only for asymptotic evaluation, but also for estimation. For example, let $\omega(\theta) \in L^\infty(S^1)$ be a bounded weight on S^1 with Fourier coefficients $\omega_k = \int_{-\pi}^\pi e^{-ik\theta} \omega(\theta) \frac{d\theta}{2\pi}$, $k \in \mathbb{Z}$. Let $((T(\omega))_{jk})_{j,k=0}^\infty = (\omega_{j-k})_{j,k=0}^\infty$ denote the Toeplitz matrix associated with ω acting on $\ell_2^+ = \{u = (u_0, u_1, \dots) : \sum_{k=0}^\infty |u_k|^2 < \infty\}$, and let $((T_n(\omega))_{jk})_{j,k=0}^n = (\omega_{j-k})_{j,k=0}^n$ denote the leading $(n + 1) \times (n + 1)$ section of $T(\omega)$. If ω is in the Wiener space $W^0 = \{\omega : \sum_{j=-\infty}^\infty |\omega_j| < \infty\}$ with $\omega(\theta) > 0$, then, by a well-known theorem of Krein, $(T(\omega))^{-1}$ exists as a bounded operator in ℓ_2^+ . The question is the following: How closely does $(T_n(\omega))^{-1}$ approximate $(T(\omega))^{-1}$ for n large? Let $\nu = (\nu_k)_{k \in \mathbb{Z}}$ be a Beurling weight (see, e.g., [59]): Thus $\nu_k \geq 1$, $\nu_k = \nu_{-k}$ and

$\nu_{j+k} \leq \nu_j \nu_k$ for all $j, k \in \mathbb{Z}$. In particular, $((1 + |k|)^\ell)_{k \in \mathbb{Z}}$, $\ell > 0$, and $(e^{\alpha|k|})_{k \in \mathbb{Z}}$, $\alpha > 0$, are Beurling weights. Define the Beurling class

$$W_\nu = \left\{ \omega \in L^1(S^1) : \sum_{j \in \mathbb{Z}} \nu_j |\omega_j| < \infty \right\}.$$

Clearly $W_\nu \subset W^0$ for any Beurling weight ν . Let $\omega \in W_\nu$ for some ν and assume in addition, for simplicity, that the weights increase on \mathbb{Z}_+ , i.e., $\nu_j \leq \nu_k$ for $0 \leq j < k$. Then the following is true [35]: for n sufficiently large and $0 \leq j, k \leq n$,

$$\left| (T_n(\omega))_{jk}^{-1} - (T(\omega))_{jk}^{-1} \right| \leq c_\nu(\omega) \min(\nu_{n+1-k}^{-1}, \nu_{n+1-j}^{-1}) \tag{110}$$

for some constant $c_\nu(\omega)$. Thus for $0 \leq j, k \leq n$, $(T_n(\omega))_{jk}^{-1}$ is a good approximation to $(T(\omega))_{jk}^{-1}$, apart from the lower right corner $j \sim k \sim n$. This estimate is a generalization of an earlier estimate due essentially to Widom (see [19] for references and further discussion). The proof of (110) in [35] uses RH techniques in an essential way closely related to the proof of (86). The paper also contains other results for orthogonal polynomials on the unit circle, including a new RH proof of the reverse statement in Baxter’s theorem (cf. [59]). Interestingly, the Borodin–Okounkov operator [15], or more properly, the Borodin–Okounkov–Case–Geronimo operator, which has emerged recently as a powerful tool in the analysis of Toeplitz determinants, arises naturally in the analysis in [35].

The steepest descent method for varying weights $\omega(x) = e^{-nQ(x)}$ in [33] can also be applied to orthogonal polynomials on the unit circle with varying weights $\omega(\theta) = e^{-nQ(\theta)}$. For example, in their analysis of the length $l_n = l_n(\pi)$ of the longest increasing subsequence of a random permutation π on n letters, the authors in [4] prove that

$$\lim_{n \rightarrow \infty} \text{Prob} \left(\frac{l_n - 2\sqrt{n}}{n^{1/6}} \leq t \right) = F_2(t) \tag{111}$$

where $F_2(t)$ is the Tracy–Widom distribution function for the largest eigenvalue of a random matrix from the Gaussian Unitary Ensemble. The proof of (111) in [4] reduces, by a formula of Gessel, to the analysis of the Toeplitz determinant $\Delta_{n-1}(e^{s \cos \theta})$ where $s = (n + 1)(1 - \frac{t}{2^{1/3}(n+1)^{2/3}})$ as $n \rightarrow \infty$, and where t is the same as in (111). As indicated above, the method of [4] is modeled on the RH steepest descent method in [33]. The same RH problem with weight $e^{s \cos \theta}$ on S^1 also appears in the work of Baik and Rains [6] in their analysis of monotone subsequences of involutions.

The steepest descent method for OP’s $\{\phi_n\}$ on the unit circle can also be used to obtain detailed information on the zeros of the ϕ_n ’s as $n \rightarrow \infty$ (see [52]). In a further development, the authors of [53] have introduced an extension of the steepest descent method to non-analytic weights, obtaining in particular new results for the zeros of OP’s on the unit circle for such weights.

Throughout this paper we have restricted our attention to measures that are smooth as in (37) and (47). The OP problem for general measures $d\mu$ is then analyzed (cf. Technical Remark above) by approximating the measure appropriately by smooth measures $d\mu_\epsilon$, and then taking the limit as $\epsilon \rightarrow 0$. This approach works well for the derivation of equations, formulae, etc., but for asymptotic questions one clearly needs a different approach. Recently, remarkable connections have been

discovered ([47]) between various combinatorial problems—random growth models, random word problems, tiling problems—and certain polynomials orthogonal with respect to discrete measures. The polynomials that arise include the classical Meixner, Charlier, Krawtchouk and Hahn polynomials (see [60]). Related discoveries have also been made in the representation of the infinite dimensional symmetric and unitary groups [16, 17]. The Meixner, Charlier and Krawtchouk polynomials all have convenient integral representations (see [60]) and their asymptotic behavior can be read off using the classical method of steepest descent. This is unfortunately not the case for the Hahn polynomials (such polynomials are needed in particular to describe the tiling of hexagons by rhombi). It turns out, however, that discrete OP problems can be rephrased in terms of a discrete RHP, which is an analogue of the continuous case, and which was introduced by Borodin, along with a theory of discrete integrable operators, in [14].

In a significant further development of the non-linear steepest descent method, the authors in [5] extended the method to a wide class of discrete RHP's which includes the discrete RHP for the Hahn polynomials (as well as the other three discrete OP systems mentioned above). The relevant limit here is when the order of the OP's p_n becomes large and simultaneously the spacing between the points in the measures goes to zero at a prescribed rate (see [5]). In this way the authors are able to analyze the Hahn polynomials asymptotically, proving *en route* a conjecture of Johansson in [47] that for hexagonal tiling, the so-called “arctic circle” of [23] exhibits Tracy–Widom fluctuations as in (111) above. In [17] the authors also consider an asymptotic problem for Hahn polynomials using a discrete RHP, but the relevant limit is different from that in [5].

Many researchers are currently involved in the application of RH techniques to the theory of OP's. In addition to those mentioned above, the list includes Chen, Claeys, Kapaev, Kitaev, Kuijlaars, van Assche and Vanlessen, amongst many others. Because of space limitations, however, we unfortunately cannot describe their work in any detail, and we must refer the reader to the literature.

4. Related Areas

In this final section we will describe, very briefly, various areas related to OP's in which the RH method plays a role.

We first consider random matrix theory (RMT), which has been a major source of questions and challenges to OP theorists for over forty years (see, e.g., [55] and [25]). The situation is as follows. A Unitary Ensemble (UE) is an ensemble of $N \times N$ Hermitian matrices $\{M = M^*\}$ with probability distribution

$$P_N(M) dM = \frac{1}{Z_N} e^{-\text{tr} W(M)} dM \quad (112)$$

where

- dM denotes Lebesgue measure on the algebraically independent elements of M .
- $W(x)$ is a real-valued function that goes to $+\infty$ as $|x| \rightarrow \infty$. The case $W(x) = x^2$ gives rise to the Gaussian Unitary Ensemble (GUE).
- Z_N is a normalization coefficient.

“Unitary” refers to the fact that the distribution (112) is invariant under unitary conjugation, $M \rightarrow UMU^*$, U unitary. The Universality Conjecture for

UE's (see [55] and [25]) states, in particular, the following: Given W , if $J_N = c_N + s_N(-t, t)$ is a suitably centered and scaled interval in \mathbb{R} , then as $N \rightarrow \infty$, $P(J_N) = \text{Prob}(M : M \text{ has no eigenvalues in } J_N)$ converges to a *universal limit* independent of W ,

$$\lim_{N \rightarrow \infty} P(J_N) = \det(1 - S_t) \tag{113}$$

where S_t is the trace class operator with kernel $S_t(x, y) = \frac{\sin \pi(x-y)}{\pi(x-y)}$ acting in $L^2(-t, t)$. The specific form of the weight $e^{-W(x)} dx$ is reflected only in the precise values of c_N and s_N . OP's enter the picture because of the celebrated result of Gaudin and Mehta (see [55]) that if $B \subset \mathbb{R}$ is a Borel set, then

$$\text{Prob}(M : M \text{ has no eigenvalues in } B) = \det(1 - K_{N,B}) \tag{114}$$

where $K_{N,B}$ is the finite rank operator with kernel

$$K_N(x, y) = \sum_{j=0}^{N-1} p_j(x)p_j(y)e^{-\frac{1}{2}W(x)}e^{-\frac{1}{2}W(y)} \tag{115}$$

acting on $L^2(B)$, and $\{p_j\}_{j \geq 0}$ are the orthonormal polynomials (3) with respect to the weight $e^{-W(x)} dx$. Hence the question of proving universality as in (113) becomes a question of deriving the appropriate asymptotics for OP's, and this is the main scientific content of [34, 33, 25] and [12]. Of course, if the weight $e^{-W(x)} dx$ is classical, e.g., $W(x) = x^2$, and the asymptotics of the associated polynomials $\{p_j\}_{j \geq 0}$ can be derived from an integral representation, then universality for these ensembles can be proved without recourse to the RH steepest descent method, and this has been done by various authors (see [34, 33] for references to the literature).

Orthogonal ensembles (OE's) of $N \times N$ real symmetric matrices $\{M = \bar{M} = M^T\}$ and Symplectic Ensembles (SE's) of $2N \times 2N$ Hermitian self-dual matrices $\{M = M^*, JM J^T = M^T\}$, where $J = \text{diag}(\tau, \dots, \tau)$, $\tau = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, equipped with invariant weights analogous to (112), are more difficult to analyze. Firstly, in the place of determinantal expressions as in (114), one obtains Pfaffians (see [55] for classical ensembles, [61] for the general case)

$$\text{Prob}(M : M \text{ has no eigenvalues in } B) = (\det(1 - \hat{K}_{N,B}))^{1/2}, \tag{116}$$

and, moreover, the operators $\hat{K}_{N,B}$ are now 2×2 matrix operators with kernels $(\hat{K}_{N,ij}(x, y))_{1 \leq i, j \leq 2}$, $x, y \in B$. In contrast to (115), these kernels are most naturally expressed in terms of certain skew-orthogonal polynomials (see [55]), but for general weights $e^{-W(x)} dx$ the asymptotic behavior of such polynomials is not known. However, Widom [66] has shown that if W'/W is rational, then $(\hat{K}_{N,ij}(x, y))$ can be expressed conveniently in terms of the orthonormal polynomials $\{p_j\}_{j \geq 0}$ with respect to the weight $e^{-W(x)} dx$, so again, as in the unitary case, the question of universality of OE's and SE's becomes a question of analyzing the asymptotic behavior of OP's. The expressions for $(\hat{K}_{N,ij}(x, y))_{1 \leq i, j \leq 2}$ are now more cumbersome than (115) and significant new technical issues arise, but nevertheless, using the asymptotic analysis in [34] as a basic ingredient, it is indeed possible to use Widom's formulae in [66] to prove universality for OE's and SE's with weights of the form $e^{-V(x)} dx$, $V(x) = \gamma x^{2m} + \dots$, $\gamma > 0$. This is the content of [28] and [29].

Biorthogonal polynomials $\pi_k(x) = x^k + \dots, \sigma_j(y) = y^j + \dots, k, j \geq 0,$

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \pi_k(x) \sigma_j(y) e^{-V(x)-W(y)+2\tau xy} dx dy = 0 \quad \text{if } j \neq k, \tag{117}$$

arise in the analysis of the theory of coupled random matrices. Here $V(x)$ and $W(y)$ grow sufficiently rapidly as $|x|, |y| \rightarrow \infty,$ and $\tau \neq 0.$ Various RH problems have been proposed to analyze these polynomials (see, in particular, [11, 48, 49] and the references therein), but the analysis of the asymptotic behavior of these RHP's is still at a preliminary stage.

For $m \geq 2,$ let $n = (n_1, n_2, \dots, n_m)$ be a vector of non-negative integers, and let $\omega_1(x) \geq 0, \dots, \omega_m(x) \geq 0$ be weights on \mathbb{R} with finite moments. Let $|n| = n_1 + \dots + n_m.$ Multiple orthogonal polynomials (see [2]) of type I are polynomials $A_n^{(k)}$ for $k = 1, 2, \dots, m,$ $\deg A_n^{(k)} \leq n_k - 1$ such that the function

$$H_n(x) = \sum_{k=1}^m A_n^{(k)}(x) \omega_k(x)$$

satisfies

$$\int_{\mathbb{R}} x^j H_n(x) dx = \begin{cases} 0, & \text{for } j = 0, \dots, |n| - 2; \\ 1, & \text{for } j = |n| - 1. \end{cases} \tag{118}$$

Multiple orthogonal polynomials $L_n(x)$ of type II are monic polynomials of degree $|n|$ satisfying

$$\int_{\mathbb{R}} L_n(x) x^k \omega_j(x) dx = 0 \quad \text{for } k = 0, \dots, n_j - 1, j = 1, \dots, m. \tag{119}$$

Multiple orthogonal polynomials were first introduced by Hermite in his proof of the transcendence of $e.$ In 2000, van Assche, Geronimo and Kuijlaars [62] showed that multiple orthogonal polynomial problems of types I and II could be rephrased as RHP's analogous to the RHP of Fokas, Its and Kitaev for ordinary OP's, and they used these RHP's to derive various properties and relations for the multiple OP's. In the last year or two significant progress has been made in extending and applying the steepest descent method to RHP's which arise from multiple OP's in special cases. We mention, in particular, [13, 3] and [51, 50] and the references therein: In the first two papers the authors consider a random matrix ensemble $P_N(M) dM = \frac{1}{Z_N} e^{-N \operatorname{tr}(\frac{1}{2} M^2 - AM)} dM,$ with external source $A,$ first analyzed by Pastur, Brézin–Hikami, and later by Zinn–Justin. Under certain conditions on $A,$ they show that the ensemble can be analyzed as $N \rightarrow \infty$ in terms of a 3×3 RHP to which an extension of the non-linear steepest descent method can be applied: A new phenomenon now occurs in the analysis, which the authors term a “global opening of lenses” (see [3]). In the second two papers the authors analyze type I and type II Hermite–Padé approximations to the exponential function, which they are again able to control by applying an extension of the steepest descent to a 3×3 RHP.

Riemann–Hilbert techniques can also be used to analyze the asymptotics of so-called *orthogonal Laurent polynomials.* Such polynomials arise in the following way. Let $V(x)$ be a real-analytic function on $\mathbb{R} \setminus \{0\}$ with the property

$$\lim_{|x| \rightarrow \infty} \frac{V(x)}{\ln|x|} = \lim_{|x| \rightarrow 0} \frac{V(x)}{\ln(|x|^{-1})} = +\infty.$$

Orthogonalization of the ordered basis $\{1, z^{-1}, z, z^{-2}, z^2, \dots\}$ with respect to the pairing $(f, g) \mapsto \int_{\mathbb{R}} f(s)g(s)e^{-NV(s)} ds$ leads to the even degree and odd degree orthonormal Laurent polynomials $\{\phi_m\}_{m \geq 0}$: $\phi_{2n}(z) = \xi_{-n}^{(2n)}z^{-n} + \dots + \xi_n^{(2n)}z^n$, $\xi_n^{(2n)} > 0$, $\phi_{2n+1}(z) = \xi_{-n-1}^{(2n+1)}z^{-n-1} + \dots + \xi_n^{(2n+1)}z^n$, $\xi_{-n-1}^{(2n+1)} > 0$. Recently, McLaughlin, Vartanian and Zhou (see [54] and the references therein) have used RHP-steepest descent methods to analyze the asymptotic behavior of the Laurent polynomials $\phi_{2n}(z), \phi_{2n+1}(z)$ and their associated norming constants $\xi_n^{(2n)}, \xi_{-n-1}^{(2n+1)}$ in the limit as $N \rightarrow \infty, N/n \rightarrow 1$. The work of McLaughlin et al. involves significant extensions of the steepest descent method: Such extensions are needed in order to overcome the new difficulties introduced into the problem by the singularity of the potential $V(x)$ at $x = 0$.

Finally, there are problems in which the asymptotic behavior of the system at hand is described by OP's. This happens, in particular, in the case of the so-called Toda rarefaction problem (see [30]). Here one considers the initial-boundary value problem for the Toda lattice

$$\ddot{x}_n = e^{x_{n-1}-x_n} - e^{x_n-x_{n+1}}, \quad n \geq 1 \tag{120}$$

where for some $\alpha > 0$

$$\begin{cases} x_n(0) = \alpha n, & n \geq 1; \\ \dot{x}_n(0) = 0, & n \geq 1, \end{cases} \tag{121}$$

and the driving particle moves with a fixed velocity $2a$

$$x_0(t) = 2at, \quad t \geq 0. \tag{122}$$

Making the change of variables $x_n \rightarrow \alpha n + y_n$ one sees that, apart from rescaling time, one can always assume without loss of generality that $\alpha = 0$ in (121). One may think of (120)–(122) as a cylinder of particles $\{x_n\}_{n \geq 1}$ driven by a piston x_0 . If $a > 0$, one has the (Toda) shock problem ([63]) and if $a < 0$, one has the (Toda) rarefaction problem. In the rarefaction problem, if $|a|$ is sufficiently large ($|a| > 1$ turns out to be the critical region), one expects that the piston will separate from the “gas” $\{x_n\}_{n \geq 1}$ and cavitation will occur. This is indeed what happens: if $a < -1$, the authors in [30] show, using the RH steepest descent method, that as $t \rightarrow \infty$, the solution of the Toda lattice splits into two parts, I+II. Part I models the cavitation and Part II is an exponentially decreasing error term. Quite remarkably, Part I is constructed from the solution of an associated OP problem, which turns out to be the Fokas, Its, Kitaev RHP in disguise. We refer the reader to [30] for details.

Acknowledgements. The work of the author was supported in part by DMS Grants No. 0296084 and No. 0500923 and also by a Friends of the Institute Visiting Membership at the Institute for Advanced Study at Princeton, Spring 2006. The author would also like to thank Irina Nenciu for many comments and much help.

References

- [1] Akhiezer, N. I.: *The Classical Moment Problem and Some Related Questions on Analysis*, Translated by N. Kemmer, Hafner, New York, 1965.
- [2] Aptekarev, A. I.: Multiple-orthogonal polynomials, *J. Comput. Appl. Math.* **99** (1998), 423–447.
- [3] Aptekarev, A. I., Bleher, P. M. and Kuijlaars, A. B. J.: Large n limit source. Part II, *Comm. Math. Phys.* **259** (2005), 367–389.

- [4] Baik, J., Deift, P. and Johansson, K.: On the distribution of the length of the longest increasing subsequence of random permutations, *J. Amer. Math. Soc.* **12** (1999), 1119–1178.
- [5] Baik, J., Kriecherbauer, T., McLaughlin, K. T-R. and Miller, P. D.: Uniform asymptotics for polynomials orthogonal with respect to a general class of discrete weights and universality results for associated ensembles, to appear in *Annals of Math. Studies*, math.CA/0310278.
- [6] Baik, J. and Rains, E.: The asymptotics of monotone subsequences of involutions, *Duke Math. J.* **109** (2001), 205–281.
- [7] Basor, E. L., Chen, Y. and Widom, H.: Hankel determinants as Fredholm determinants, *Random Matrix Models and Their Applications*, Math. Sci. Res. Inst. Publ. **40**. Cambridge University Press, Cambridge, 2001, 21–29.
- [8] Basor, E. L., Chen, Y. and Widom, H.: Determinants for Hankel matrices, *J. Func. Anal.* **179** (2001), 214–234.
- [9] Basor, E. L. and Widom, H.: Wiener–Hopf determinants and Fisher–Hartwig symbols, *Operator Theoretical Methods and Applications to Mathematical Physics*, Oper. Theory Adv. Appl. **147**. Birkhäuser, Basel, 2004, 131–149.
- [10] Beals, R., Deift, P. and Tomei, C.: Direct and inverse scattering on the line, *Math. Surveys and Monographs* **28**. American Mathematical Society, Providence, Rhode Island, 1988.
- [11] Bertola, M., Eynard, B. and Harnad, J.: Differential systems for bi-orthogonal polynomials appearing in 2-matrix models and the associated Riemann–Hilbert problem, *Comm. Math. Phys.* **243** (2003), 193–240.
- [12] Bleher, P. M. and Its, A. R.: Semiclassical asymptotics of orthogonal polynomials, Riemann–Hilbert problems, and universality in the matrix model, *Ann. of Math. (2)* **150** (1999), 185–266.
- [13] Bleher, P. M. and Kuijlaars, A. B. J.: Large n limit of Gaussian random matrices with external source. Part I, *Comm. Math. Phys.* **252** (2004), 43–46.
- [14] Borodin, A.: Riemann–Hilbert problem and the discrete Bessel kernel, *Int. Math. Res. Not.* **9** (2000), 467–494.
- [15] Borodin, A. and Okounkov, A. A.: Fredholm determinant formula for Toeplitz determinants, *Integral Equations Operator Theory* **37** (2000), 386–396.
- [16] Borodin, A. and Olshanski, G.: Distributions on partitions, point processes, and the hypergeometric kernel, *Comm. Math. Phys.* **211** (2000), 335–358.
- [17] Borodin, A. and Olshanski, G.: Harmonic analysis on the infinite-dimensional unitary group and determinantal point processes, *Ann. of Math.* **161** (2005), 1319–1422.
- [18] Böttcher, A.: On the determinant formulas by Borodin, Okounkov, Baik, Deift and Rains, *Toeplitz Matrices and Singular Integral Equations* (Pobershaw, 2001), Oper. Theory Adv. Appl., **135**. Birkhäuser, Basel, 2002, 91–99.
- [19] Böttcher, A. and Silbermann, B.: *Introduction to Large Truncated Toeplitz Matrices*, Springer-Verlag, New York, 1999.
- [20] Cantero, M. J., Moral, L. and Velázquez, L.: Five-diagonal matrices and zeros of orthogonal polynomials on the unit circle, *Linear Algebra Appl.* **362** (2003), 29–56.
- [21] Case, K. M. and Geronimo, J. S.: Scattering theory and polynomials orthogonal on the unit circle, *J. Math. Phys.* **20** (1979), 299–310.
- [22] Clancey, K. and Gohberg, I.: *Factorization of Matrix Functions and Singular Integral Operators*, Operator Theory: Advances and Applications, **3**. Birkhäuser, Basel, Boston, 1981.
- [23] Cohn, H., Larsen, M. and Propp, J.: The shape of a typical boxed plane partition, *New York J. Math.* **4** (1998), 137–165.
- [24] Deift, P.: Integrable operators, *Amer. Math. Soc. Transl. (2)* **189** (1999), 69–84.
- [25] Deift, P.: *Orthogonal Polynomials and Random Matrices: A Riemann–Hilbert approach*, Courant Lecture Notes **3**. American Mathematical Society, Providence, Rhode Island, 2000.
- [26] Deift, P.: Unpublished.
- [27] Deift, P.: Unpublished.
- [28] Deift, P. and Gioev, D.: Universality in random matrix theory for orthogonal and symplectic ensembles, preprint, math-ph/0411075.

- [29] Deift, P. and Gioev, D.: Universality at the edge of the spectrum for unitary, orthogonal and symplectic ensembles of random matrices, to appear in *Comm. Pure Appl. Math.*, math-ph/0507023.
- [30] Deift, P., Kamvissis, S., Kriecherbauer, T. and Zhou, X.: The Toda rarefaction problem, *Comm. Pure Appl. Math.* **49** (1996), 35–83.
- [31] Deift, P. and Killip, R.: On the absolutely continuous spectrum of one-dimensional Schrödinger operators with square summable potentials, *Comm. Math. Phys.* **10** (2000), 702–731.
- [32] Deift, P. A., Kriecherbauer, T. and McLaughlin, K. T-R.: New results for the asymptotics of orthogonal polynomials and related problems via the Lax–Levermore method. Announcement in *Proceedings of Symposia in Applied Mathematics* **54** (1998) 87–104. Full text in *J. Approx. Theory* **95** (1998), 388–475, under the title, New results on the equilibrium measure for logarithmic potentials in the presence of an external field.
- [33] Deift, P. A., Kriecherbauer, T., McLaughlin, K. T-R., Venakides, S. and Zhou, X.: Uniform asymptotics for polynomials orthogonal with respect to varying exponential weights and applications to universality questions in random matrix theory, *Comm. Pure Appl. Math.* **52** (1999), 1335–1425.
- [34] Deift, P. A., Kriecherbauer, T., McLaughlin, K. T-R., Venakides, S. and Zhou, X.: Strong asymptotics of orthogonal polynomials with respect to exponential weights, *Comm. Pure Appl. Math.* **52** (1999), 1491–1552.
- [35] Deift, P. and Ostensson, J.: A Riemann–Hilbert approach to some theorems on Toeplitz operators and orthogonal polynomials, *J. Approx. Theory* **139** (2006), 144–171.
- [36] Deift, P., Venakides, S. and Zhou, X.: The collisionless shock region for the long-time behavior of solutions of the KdV equation, *Comm. Pure Appl. Math.* **47** (1994) 199–206.
- [37] Deift, P., Venakides, S. and Zhou, X.: New results in small dispersion KdV by an extension of the steepest descent method for Riemann–Hilbert problems, *Int. Math. Res. Not.* **6** (1997), 285–299. *Summary of Proc. Natl. Acad. Sci.*, **95** (1998), 450–454, under the title, An extension of the steepest descent method for Riemann–Hilbert problems: The small dispersion limit of the Korteweg de Vries (KdV) equation.
- [38] Deift, P. and Zhou, X.: A steepest descent method for oscillatory Riemann–Hilbert problems. Asymptotics for the MKdV equation, *Ann. of Math.* **137** (1993), 295–368.
- [39] Deift, P. and Zhou, X.: Long-time behavior of the non-focusing nonlinear Schrödinger equation: A case study, *New Series: Lectures in Mathematical Sciences*, **5**. University of Tokyo, 1994.
- [40] Deift, P. and Zhou, X.: Asymptotics for the Painlevé II equation, *Comm. Pure Appl. Math.* **48** (1995), 277–337.
- [41] Deift, P. and Zhou, X.: Long-time asymptotics for solutions of the nonlinear Schrödinger equation with initial data in a weighted Sobolev space, *Comm. Pure Appl. Math.* **56** (2003), 1029–1077.
- [42] Ehrhardt, T.: A status report on the asymptotic behavior of Toeplitz determinants with Fisher–Hartwig singularities, *Oper. Theory Adv. Appl.* **124** (2001), 217–241.
- [43] Flaschka, H.: The Toda lattice, I. Existence of integrals, *Phys. Rev. B (3)* **9** (1974), 1924–1925.
- [44] Fokas, A. S., Its, A. R. and Kitaev, A. V.: The isomonodromy approach to matrix models in 2D quantum gravity, *Comm. Math. Phys.* **147** (1992), 395–430. Announcement in *Uspekhi Mat. Nauk* **45** (6) (276) (1990), 135–136; translation in *Russian Math. Surveys* **45** (6) (1990), 155–157.
- [45] Its, A. R., Izergin, A. G., Korepin, V. E. and Slavnov, N. A.: Differential equations for quantum correlation functions, *Int. J. Mod. Phys. B* **4** (1990), 1003–1037; The quantum correlation function as the τ function of classical differential equations, *Important Developments in Soliton Theory*, A. S. Fokas and V. E. Zakharov (eds.), Springer-Verlag, Berlin, 1993, 407–417.
- [46] Its, A. R., Jin, B.-Q. and Korepin, V. E.: Entanglement in XY spin chain, preprint, quant-ph/0409027.
- [47] Johansson, K.: Discrete orthogonal polynomial ensembles and Plancherel measure, *Ann. of Math. (2)* **153** (2001), 259–296.
- [48] Kapaev, A. A.: A Riemann–Hilbert problem for bi-orthogonal polynomials, *J. Phys. A* **36** (2003) 4629–4640.

- [49] Kuijlaars, A. B. J. and McLaughlin, K. T-R.: A Riemann–Hilbert problem for biorthogonal polynomials, *J. Comput. Appl. Math.* **178** (2005), 313–320.
- [50] Kuijlaars, A. B. J., Stahl, H., van Assche, W. and Wielonsky, F.: Type II Hermite–Padé approximation to the exponential function, preprint, math.CA/0510278.
- [51] Kuijlaars, A. B. J., van Assche, W. and Wielonsky, F.: Quadratic Hermite–Padé approximation to the exponential function: A Riemann–Hilbert approach, *Constr. Approx.* **21** (2005), 351–412.
- [52] Martinez-Finkelshtein, A., McLaughlin, K. T-R. and Saff, E. B.: Szegő orthogonal polynomials with respect to an analytic weight: Canonical representation and strong asymptotics, to appear in *Constr. Approx.*, math.CA/0502300.
- [53] McLaughlin, K. T-R. and Miller, P. D.: The $\bar{\partial}$ steepest descent method and the asymptotic behavior of polynomials orthogonal on the unit circle with fixed and exponentially varying nonanalytic weights, *Int. Math. Res. Papers* **2006**, Article ID 48673, 78 pages, 2006.
- [54] McLaughlin, K. T.-R., Vartanian, A. H. and Zhou, X.: Asymptotics of Laurent polynomials of even degree orthogonal with respect to varying exponential weights, *Int. Math. Res. Papers* **2006**, Article ID 62815, 216 pages, 2006.
- [55] Mehta, M.: *Random Matrices*, third edition, Pure and Appl. Math. (Amsterdam), **142**. Elsevier/Academic Press, Amsterdam, 2004.
- [56] Moser, J.: Finitely many mass points on the line under the influence of an exponential potential: An integrable system, *Dynamical Systems, Theory and Applications* (Recontres, Battelle Res. Inst., Seattle, 1974), Lect. Notes in Phys. **38**. Springer, Berlin, 1975, 467–497.
- [57] Saff, E. B. and Totik, V.: *Logarithmic Potentials and External Fields*, Springer, New York-Berlin, 1997.
- [58] Simon, B.: The classical moment problem as a self-adjoint finite difference operator, *Advances in Math.* **137** (1998), 82–203.
- [59] Simon, B.: *Orthogonal polynomials on the Unit Circle. Part 1: Classical Theory; Part 2: Spectral Theory*, Amer. Math. Soc. Colloquium Publications **54**. American Mathematical Society, Providence, Rhode Island, 2005.
- [60] Szegő, G.: *Orthogonal Polynomials*, fourth edition, American Mathematical Society, Providence, Rhode Island, 1975.
- [61] Tracy, C. A. and Widom, H.: Correlation functions, cluster functions, and spacing distributions for random matrices, *J. Statist. Phys.* **92** (1998), 809–835.
- [62] van Assche, W., Geronimo, J. S. and Kuijlaars, A. B. J.: Riemann–Hilbert problems for multiple-orthogonal polynomials, *Special Function 2000: Current Perspective and Future Directions* (J. Bustoz et al., eds.), NATO Science Series II Mathematics, Physics and Chemistry **30**. Kluwer, Dordrecht, 2001, 23–59.
- [63] Venakides, S., Deift, P. and Oba, R.: The Toda shock problem, *Comm. Pure Appl. Math.* **44** (1991), 1171–1242.
- [64] Widom, H.: Asymptotic behavior of block Toeplitz matrices and determinants, *Advances in Math.* **13** (1974), 284–322.
- [65] Widom, H.: Asymptotic behavior of block Toeplitz matrices and determinants. II, *Advances in Math.* **21** (1976), 1–29.
- [66] Widom, H.: On the relation between orthogonal, symplectic and unitary matrix ensembles, *J. Statist. Phys.* **94** (1999), 347–363.
- [67] Watkins, D. S.: Some perspectives on the eigenvalue problem, *SIAM Review* **35** (1993), 430–471.
- [68] Zakharov, V. E. and Manakov, S. V.: Asymptotic behavior of nonlinear wave systems by the inverse method, *Sov. Phys. JETP* **44** (1976), 106–112.

COURANT INSTITUTE OF MATHEMATICAL SCIENCES, NEW YORK UNIVERSITY, 251 MERCER STREET, NEW YORK, NY 10012, U.S.A.

E-mail address: deift@cims.nyu.edu

Inverse Spectral Theory as Influenced by Barry Simon

Fritz Gesztesy

*Dedicated with great pleasure to Barry Simon, mentor and friend,
on the occasion of his 60th birthday*

ABSTRACT. We survey Barry Simon's principal contributions to the field of inverse spectral theory in connection with one-dimensional Schrödinger and Jacobi operators.

CONTENTS

1. Introduction
 2. The Dirichlet Spectral Deformation Method
 3. Renormalized Oscillation Theory
 4. Trace Formulas for Schrödinger and Jacobi Operators:
The ξ Function
 5. Various Uniqueness Theorems in Inverse Spectral Theory
 6. The Crown Jewel: Simon's New Approach
to Inverse Spectral Theory
- References

1. Introduction

This Festschrift contribution is devoted to a survey of Barry Simon's principal contributions to the area of inverse spectral theory for one-dimensional Schrödinger and Jacobi operators. We decided to put the emphasis on the following five groups of topics:

.

2000 *Mathematics Subject Classification*. Primary 34A55, 34B20, 34B24; Secondary 34C10, 34L05.

Key words and phrases. Spectral deformations, oscillation theory, trace formulas, Weyl–Titchmarsh theory, uniqueness results, inverse spectral theory.

- The Dirichlet spectral deformation method

A general spectral deformation method applicable to Schrödinger, Jacobi, and Dirac-type operators in one dimension, which can be used to insert eigenvalues into spectral gaps of arbitrary background operators but is also an ideal technique to construct isospectral (in fact, unitarily equivalent) sets of operators starting from a given base operator.

- Renormalized oscillation theory

Renormalized oscillation theory formulated in terms of Wronskians of appropriate solutions, rather than solutions themselves, applies, in particular, to energies above the essential spectrum where real-valued solutions exhibit infinitely many zeros and traditional eigenvalue counting methods would naively lead to $\infty - \infty$. While not directly related to inverse spectral methods, we chose to include this topic because of its fundamental importance to the Dirichlet spectral deformation method.

- The xi function and trace formulas for Schrödinger and Jacobi operators

The xi function, that is, essentially, the argument of the diagonal Green's function, which also takes on the role of a particular spectral shift function, is an ideal tool to derive a hierarchy of (higher-order) trace formulas for one-dimensional Schrödinger and Jacobi operators. The latter are the natural extensions of the well-known trace formulas for periodic and algebro-geometric finite-band potential coefficients to arbitrary coefficients. The xi function provides a tool for direct and inverse spectral theory.

- Uniqueness theorems in inverse spectral theorem

Starting from the Borg–Marchenko uniqueness theorem, the basic uniqueness result for Schrödinger and Jacobi operators in terms of the Weyl–Titchmarsh m -coefficient, a number of uniqueness results are discussed. The latter include the Borg-type two-spectra results as well as Hochstadt–Lieberman-type results with mixed prescribed data. In addition to these traditional inverse spectral problems, several new types of inverse spectral problems are addressed.

- Simon's new approach to inverse spectral theory

In some sense, Simon's new approach to inverse spectral theory for half-line problems, based on a particular representation of the Weyl–Titchmarsh m -function as a finite Laplace transform with control about the error term, can be viewed as a continuum analog of the continued fraction approach (based on the Riccati equation) to the inverse spectral problem for semi-infinite Jacobi matrices (the actual details, however, differ considerably). Among a variety of spectral-theoretic results, this leads to a formulation of the half-line inverse spectral problem alternative to that of Gel'fand and Levitan. In addition, it leads to a fundamental new result, the local Borg–Marchenko uniqueness theorem.

Each individual section focuses on a particular paper or a group of papers to be surveyed, representing the five items just discussed. Since this survey is fairly long, it was our intention to write each section in such a manner that it can be read independently.

Only self-adjoint Schrödinger and Jacobi operators are considered in this survey. In particular, all potential coefficients V and Jacobi matrix coefficients a and b are

assumed to be real-valued throughout this survey (although we occasionally remind the reader of this assumption).

Certainly, this is not a survey of the state of the art of inverse spectral theory for one-dimensional Schrödinger and Jacobi operators. Rather, it focuses exclusively on Barry Simon’s contributions to and influence exerted on the field. The bibliography, although quite long, is far from complete and only reflects the particular purpose of this survey. The references included are typically of two kinds: First, background references that were used by Barry Simon and his coworkers in writing a particular paper. Such references are distributed throughout the particular survey of one of his papers. Second, at the end of each such survey we refer to more recent references which complement the results of the particular paper in question.

It was in April of 1983 that Barry and I first met in person at Caltech and started our collaboration. Barry became a mentor and then a friend, and it is fair to say he has had a profound influence on my career since that time. Working with Barry has been exciting and most rewarding for me. Happy Birthday, Barry, and many more such anniversaries!

2. The Dirichlet Spectral Deformation Method

In this section we describe some of the principal results of the paper:

[100] F. Gesztesy, B. Simon, and G. Teschl, *Spectral deformations of one-dimensional Schrödinger operators*, *J. Analyse Math.* **70**, 267–324 (1996).

Spectral deformations of Schrödinger operators in $L^2(\mathbb{R})$, isospectral and certain classes of non-isospectral ones, have attracted a lot of interest over the past three decades due to their prominent role in connection with a variety of topics, including the Korteweg–de Vries (KdV) hierarchy, inverse spectral problems, supersymmetric quantum mechanical models, level comparison theorems, etc. In fact, the construction of N -soliton solutions of the KdV hierarchy (and more generally, the construction of solitons relative to reflectionless backgrounds) is a typical example of a non-isospectral deformation of $H = -\frac{d^2}{dx^2}$ in $L^2(\mathbb{R})$ since the resulting deformation $\tilde{H} = -\frac{d^2}{dx^2} + \tilde{V}$ acquires an additional point spectrum $\{\lambda_1, \dots, \lambda_N\} \subset (-\infty, 0)$, $N \in \mathbb{N}$, such that

$$\sigma(\tilde{H}) = \sigma(H) \cup \{\lambda_1, \dots, \lambda_N\}$$

($\sigma(\cdot)$ abbreviating the spectrum). In the N -soliton context (ignoring the KdV time parameter for simplicity), \tilde{V} is of the explicit form

$$\tilde{V}(x) = -2 \frac{d^2}{dx^2} \ln[W(\Psi_1(x), \dots, \Psi_N(x))], \quad x \in \mathbb{R}, \tag{2.1}$$

where $W(f_1, \dots, f_N)$ denotes the Wronskian of f_1, \dots, f_N and the functions Ψ_j , $j = 1, \dots, N$, are given by

$$\begin{aligned} \Psi_j(x) &= (-1)^{j+1} e^{-\kappa_j x} + \alpha_j e^{\kappa_j x}, \quad x \in \mathbb{R}, \\ 0 &< \kappa_1 < \dots < \kappa_N, \quad \alpha_j > 0, \quad j = 1, \dots, N. \end{aligned}$$

The Wronski-type formula in (2.1) is typical also for general background potentials and typical for the Crum–Darboux-type commutation approach [44], [48] (cf. [90] and the references therein for general backgrounds) which underlies all standard

spectral deformation methods for one-dimensional Schrödinger operators such as single and double commutation, and the Dirichlet deformation method presented in this section.

On the other hand, the solution of the inverse periodic problem and the corresponding solution of the algebro-geometric quasi-periodic finite-band inverse problem for the KdV hierarchy (and certain almost-periodic limiting situations thereof) are intimately connected with isospectral (in fact, unitary) deformations of a given base (background) operator $H = -\frac{d^2}{dx^2} + V$. Although not a complete bibliography on applications of spectral deformations in mathematical physics, the interested reader may consult [14], [20], [22], [28], [44], [46], [48], [59, Sect. 4.3], [60], [61], [63], [64], [66], [68], [72], [73], [102], [74], [78], [79, App. G], [90], [100], [103], [105], [112, Ch. 2, App. A], [132], [135], [159, Sect. 6.6], [174]–[178], [192, Chs. 3, 4], [194], [207], [216], [217], [218], [219], and the numerous references cited therein.

The main motivation for writing [100] originated in our interest in inverse spectral problems. As pointed out later (see Remarks 2.5–2.7), spectral deformation methods can provide crucial insights into the isospectral class of a given base potential V , and in some cases can even determine the whole isospectral class of such potentials. A particularly interesting open problem in inverse spectral theory concerns the characterization of the isospectral class of potentials V with purely discrete spectra (e.g., the harmonic oscillator $V(x) = x^2$, cf. [32]–[36], [99], [160], [180], [193]).

The principal result in [100], reviewed in this section (cf. Theorem 2.4 (i)), provides a complete spectral characterization of a new method of constructing isospectral (in fact, unitary) deformations of general Schrödinger operators $H = -\frac{d^2}{dx^2} + V$ in $L^2(\mathbb{R})$. The technique is connected to Dirichlet data, that is, to the spectrum of the operator H^D on $L^2((-\infty, x_0]) \oplus L^2([x_0, \infty))$ with a Dirichlet boundary condition at x_0 . The transformation moves a single eigenvalue of H^D and perhaps flips the half-line (i.e., $(-\infty, x_0)$ to (x_0, ∞) , or vice versa) to which the Dirichlet eigenvalue belongs. On the remainder of the spectrum, the transformation is realized by a unitary operator.

To describe the Dirichlet deformation method (DDM) as developed in [100] in some detail, we suppose that $V \in L^1_{loc}(\mathbb{R})$ is real-valued and introduce the differential expression $\tau = -\frac{d^2}{dx^2} + V(x)$, $x \in \mathbb{R}$. Assuming τ to be in the limit point case at $\pm\infty$ (for the general case we refer to [100]) one defines the self-adjoint base (i.e., background) operator H in $L^2(\mathbb{R})$ by

$$Hf = \tau f, \quad f \in \text{dom}(H) = \{g \in L^2(\mathbb{R}) \mid g, g' \in AC_{loc}(\mathbb{R}); \tau g \in L^2(\mathbb{R})\}. \quad (2.2)$$

Here $W(f, g)(x) = f(x)g'(x) - f'(x)g(x)$ denotes the Wronskian of $f, g \in AC_{loc}(\mathbb{R})$ (the set of locally absolutely continuous functions on \mathbb{R}). Given H and a fixed reference point $x_0 \in \mathbb{R}$, we introduce the associated Dirichlet operator $H^D_{x_0}$ in $L^2(\mathbb{R})$ by

$$H^D_{x_0} f = \tau f, \quad f \in \text{dom}(H^D_{x_0}) = \{g \in L^2(\mathbb{R}) \mid g \in AC_{loc}(\mathbb{R}), g' \in AC_{loc}(\mathbb{R} \setminus \{x_0\}); \lim_{\epsilon \downarrow 0} g(x_0 \pm \epsilon) = 0; \tau g \in L^2(\mathbb{R})\}.$$

Clearly, $H^D_{x_0}$ decomposes into $H^D_{x_0} = H^D_{-,x_0} \oplus H^D_{+,x_0}$ with respect to the orthogonal decomposition $L^2(\mathbb{R}) = L^2((-\infty, x_0]) \oplus L^2([x_0, \infty))$. Moreover, for any $\mu \in$

$\sigma_d(H_{x_0}^D) \setminus \sigma(H)$ ($\sigma_d(\cdot)$, the discrete spectrum, $\sigma(\cdot)$ and $\sigma_{\text{ess}}(\cdot)$, the spectrum and essential spectrum, respectively), we introduce the Dirichlet datum

$$(\mu, \sigma) \in \{\sigma_d(H_{x_0}^D) \setminus \sigma_d(H)\} \times \{-, +\}, \tag{2.3}$$

which identifies μ as a discrete Dirichlet eigenvalue on the interval $(x_0, \sigma\infty)$, that is, $\mu \in \sigma_d(H_{\sigma, x_0}^D)$, $\sigma \in \{-, +\}$ (but excludes it from being simultaneously a Dirichlet eigenvalue on $(x_0, -\sigma\infty)$).

Next, we pick a fixed spectral gap (E_0, E_1) of H , the endpoints of which (without loss of generality) belong to the spectrum of H ,

$$(E_0, E_1) \subset \mathbb{R} \setminus \sigma(H), \quad E_0, E_1 \in \sigma(H)$$

and choose a discrete eigenvalue μ of $H_{x_0}^D$ in the closure of that spectral gap,

$$\mu \in \sigma_d(H_{x_0}^D) \cap [E_0, E_1] \tag{2.4}$$

(we note there is at most one such μ since $(H_{x_0}^D - z)^{-1}$ is a rank-one perturbation of $(H - z)^{-1}$). According to (2.3), this either gives rise to a Dirichlet datum

$$(\mu, \sigma) \in (E_0, E_1) \times \{-, +\}, \tag{2.5}$$

or else to a discrete eigenvalue of H_{-, x_0}^D and H_{+, x_0}^D , that is,

$$\mu \in \{E_0, E_1\} \cap \sigma_d(H) \cap \sigma_d(H_{-, x_0}^D) \cap \sigma_d(H_{+, x_0}^D) \tag{2.6}$$

since the eigenfunction of H associated with μ has a zero at x_0 . In particular, since $(H_{x_0}^D - z)^{-1}$ is a rank-one perturbation of $(H - z)^{-1}$, one infers

$$\sigma_{\text{ess}}(H_{x_0}^D) = \sigma_{\text{ess}}(H),$$

and thus, $\mu \in \{E_0, E_1\} \cap \sigma_{\text{ess}}(H)$ is excluded by assumption (2.4). Hence, the case distinctions (2.5) and (2.6) are exhaustive.

In addition to μ as in (2.4)–(2.6), we also need to introduce $\tilde{\mu} \in [E_0, E_1]$ and $\tilde{\sigma} \in \{-, +\}$ as follows: Either

$$(\tilde{\mu}, \tilde{\sigma}) \in (E_0, E_1) \times \{-, +\},$$

or else

$$\tilde{\mu} \in \{E_0, E_1\} \cap \sigma_d(H).$$

Given H , one introduces Weyl–Titchmarsh-type solutions $\psi_{\pm}(z, x)$ of $(\tau - z)\psi(z) = 0$ by

$$\begin{aligned} \psi_{\pm}(z, \cdot) &\in L^2((R, \pm\infty)), \quad R \in \mathbb{R}, \\ \lim_{x \rightarrow \pm\infty} W(\psi_{\pm}(z), g)(x) &= 0 \quad \text{for all } g \in \text{dom}(H). \end{aligned}$$

If $\psi_{\pm}(z, x)$ exist, they are unique up to constant multiples. In particular, $\psi_{\pm}(z, x)$ exist for $z \in \mathbb{C} \setminus \sigma_{\text{ess}}(H)$ and we can (and will) assume them to be holomorphic with respect to $z \in \mathbb{C} \setminus \sigma(H)$ and real-valued for $z \in \mathbb{R}$ (cf. the discussion in connection with (3.1)).

Given $\psi_{\sigma}(\mu, x)$ and $\psi_{-\tilde{\sigma}}(\tilde{\mu}, x)$, one defines

$$W_{(\tilde{\mu}, \tilde{\sigma})}(x) = \begin{cases} (\tilde{\mu} - \mu)^{-1} W(\psi_{\sigma}(\mu), \psi_{-\tilde{\sigma}}(\tilde{\mu}))(x), & \mu, \tilde{\mu} \in [E_0, E_1], \quad \tilde{\mu} \neq \mu, \\ -\sigma \int_{\sigma\infty}^x dx' \psi_{\sigma}(\mu, x')^2, & (\tilde{\mu}, \tilde{\sigma}) = (\mu, -\sigma), \quad \mu \in (E_0, E_1), \end{cases}$$

and the associated Dirichlet deformation

$$\tilde{\tau}_{(\tilde{\mu}, \tilde{\sigma})} = -\frac{d^2}{dx^2} + \tilde{V}_{(\tilde{\mu}, \tilde{\sigma})}(x),$$

$$\begin{aligned} \tilde{V}_{(\tilde{\mu}, \tilde{\sigma})}(x) &= V(x) - 2\{\ln[W_{(\tilde{\mu}, \tilde{\sigma})}(x)]\}'', \quad x \in \mathbb{R}, \\ \mu, \tilde{\mu} &\in [E_0, E_1], \quad \mu \neq \tilde{\mu} \text{ or } (\tilde{\mu}, \tilde{\sigma}) = (\mu, -\sigma), \quad \mu \in (E_0, E_1). \end{aligned} \tag{2.7}$$

As discussed in Section 3, $W_{(\tilde{\mu}, \tilde{\sigma})}(x) \neq 0$, $x \in \mathbb{R}$, and hence (2.7) yields a well-defined potential $\tilde{V}_{(\tilde{\mu}, \tilde{\sigma})} \in L^1_{\text{loc}}(\mathbb{R})$.

In the remaining cases $(\tilde{\mu}, \tilde{\sigma}) = (\mu, \sigma)$, $\mu \in [E_0, E_1]$, and $\mu = \tilde{\mu} \in \{E_0, E_1\} \cap \sigma_d(H)$, we define $\tilde{V}_{(\tilde{\mu}, \tilde{\sigma})} = V$ which represents the trivial deformation of V (i.e., none at all), and for notational simplicity these trivial cases are excluded in the remainder of this section. For obvious reasons we will allude to (2.7) as the Dirichlet deformation method in the following.

If $\tilde{\mu} \in \sigma_d(H)$, then $\psi_-(\tilde{\mu}) = c\psi_+(\tilde{\mu})$ for some $c \in \mathbb{R} \setminus \{0\}$, showing that $W_{(\tilde{\mu}, \tilde{\sigma})}(x)$ and hence, $\tilde{V}_{(\tilde{\mu}, \tilde{\sigma})}(x)$ in (2.7) becomes independent of σ or $\tilde{\sigma}$. In this case we shall occasionally use a more appropriate notation and write $\tilde{V}_{\tilde{\mu}}$ and $\tilde{\tau}_{\tilde{\mu}}$ (instead of $\tilde{V}_{(\tilde{\mu}, \tilde{\sigma})}$ and $\tilde{\tau}_{(\tilde{\mu}, \tilde{\sigma})}$).

For later reference, we now summarize our basic assumptions on V , μ , and $\tilde{\mu}$ in the following hypothesis.

HYPOTHESIS 2.1. *Suppose $V \in L^1_{\text{loc}}(\mathbb{R})$ to be real-valued. In addition, we assume*

$$\begin{aligned} (E_0, E_1) &\subset \mathbb{R} \setminus \sigma(H), \quad E_0, E_1 \in \sigma(H), \\ \mu &\in \sigma_d(H_{x_0}^D), \quad (\mu, \sigma) \in (E_0, E_1) \times \{-, +\} \text{ or } \mu \in \{E_0, E_1\} \cap \sigma_d(H), \\ (\tilde{\mu}, \tilde{\sigma}) &\in (E_0, E_1) \times \{-, +\} \text{ or } \tilde{\mu} \in \{E_0, E_1\} \cap \sigma_d(H), \\ \mu, \tilde{\mu} &\in [E_0, E_1], \quad \mu \neq \tilde{\mu} \text{ or } (\tilde{\mu}, \tilde{\sigma}) = (\mu, -\sigma), \quad \mu \in (E_0, E_1). \end{aligned}$$

Next, introducing the following solutions of $(\tilde{\tau}_{(\tilde{\mu}, \tilde{\sigma})} - z)\tilde{\psi}(z) = 0$,

$$\begin{aligned} \tilde{\psi}_{-\sigma}(\mu, x) &= \psi_{-\tilde{\sigma}}(\tilde{\mu}, x) / W_{(\tilde{\mu}, \tilde{\sigma})}(x), \\ \tilde{\psi}_{\tilde{\sigma}}(\tilde{\mu}, x) &= \psi_{\sigma}(\mu, x) / W_{(\tilde{\mu}, \tilde{\sigma})}(x), \quad \tilde{\psi}_{\tilde{\sigma}}(\tilde{\mu}, x_0) = 0, \end{aligned}$$

one infers

$$(\tilde{\tau}_{(\tilde{\mu}, \tilde{\sigma})}\tilde{\psi}_{-\sigma}(\mu))(x) = \mu\tilde{\psi}_{-\sigma}(\mu, x), \quad (\tilde{\tau}_{(\tilde{\mu}, \tilde{\sigma})}\tilde{\psi}_{\tilde{\sigma}}(\tilde{\mu}))(x) = \tilde{\mu}\tilde{\psi}_{\tilde{\sigma}}(\tilde{\mu}, x).$$

The Dirichlet deformation operator $\tilde{H}_{(\tilde{\mu}, \tilde{\sigma})}$ associated with $\tilde{\tau}_{(\tilde{\mu}, \tilde{\sigma})}$ in (2.7) is then defined as follows:

$$\begin{aligned} \tilde{H}_{(\tilde{\mu}, \tilde{\sigma})}f &= \tilde{\tau}_{(\tilde{\mu}, \tilde{\sigma})}f, \quad f \in \text{dom}(\tilde{H}_{(\tilde{\mu}, \tilde{\sigma})}) = \{g \in L^2(\mathbb{R}) \mid g, g' \in AC_{\text{loc}}(\mathbb{R}); \\ &g \text{ satisfies the b.c. in (2.9); } \tilde{\tau}_{(\tilde{\mu}, \tilde{\sigma})}g \in L^2(\mathbb{R})\}. \end{aligned} \tag{2.8}$$

The boundary conditions (b.c.) alluded to in (2.8) are chosen as follows:

$$\begin{aligned} \lim_{x \rightarrow \tilde{\sigma}\infty} W(\tilde{\psi}_{\tilde{\sigma}}(\tilde{\mu}), g)(x) &= 0 \text{ if } \tilde{\tau}_{(\tilde{\mu}, \tilde{\sigma})} \text{ is l.c. at } \tilde{\sigma}\infty, \\ \lim_{x \rightarrow -\tilde{\sigma}\infty} W(\tilde{\psi}_{-\sigma}(\mu), g)(x) &= 0 \text{ if } \tilde{\tau}_{(\tilde{\mu}, \tilde{\sigma})} \text{ is l.c. at } -\tilde{\sigma}\infty. \end{aligned} \tag{2.9}$$

Here we abbreviate the limit point and limit circle cases by l.p. and l.c., respectively. As usual, the boundary condition at $\omega\infty$ in (2.8) is omitted if $\tilde{\tau}_{(\tilde{\mu}, \tilde{\sigma})}$ is l.p. at $\omega\infty$, $\omega \in \{-, +\}$.

For future reference we note that in analogy to the Dirichlet operators $H_{x_0}^D$, H_{\pm, x_0}^D introduced in connection with the operator H , one can also introduce the corresponding Dirichlet operators $\tilde{H}_{(\tilde{\mu}, \tilde{\sigma}), x_0}^D$, $\tilde{H}_{(\tilde{\mu}, \tilde{\sigma}), \pm, x_0}^D$ associated with $\tilde{H}_{(\tilde{\mu}, \tilde{\sigma})}$.

Next, we turn to the Weyl–Titchmarsh m -functions for the Dirichlet deformation operator $\tilde{H}_{(\tilde{\mu}, \tilde{\sigma})}$ and relate them to those of H .

Let $\phi(z, x), \theta(z, x)$ be the standard fundamental system of solutions of $(\tau - z)\psi(z) = 0, z \in \mathbb{C}$ defined by $\phi(z, x_0) = \theta'(z, x_0) = 0, \phi'(z, x_0) = \theta(z, x_0) = 1, z \in \mathbb{C}$, and denote by $\tilde{\theta}_{(\tilde{\mu}, \tilde{\sigma})}(z, x), \tilde{\phi}_{(\tilde{\mu}, \tilde{\sigma})}(z, x)$ the analogously normalized fundamental system of solutions of $(\tilde{\tau}_{(\tilde{\mu}, \tilde{\sigma})} - z)\tilde{\psi}(z) = 0, z \in \mathbb{C}$, at x_0 . One then has

$$m_\sigma(z, x_0) = \psi'_\sigma(z, x_0)/\psi_\sigma(z, x_0), \quad \sigma \in \{-, +\}, z \in \mathbb{C} \setminus \mathbb{R},$$

where $m_\sigma(\cdot, x_0)$ denotes the Weyl–Titchmarsh m -function of H with respect to the half-line $(x_0, \sigma\infty), \sigma \in \{-, +\}$. For the corresponding half-line Weyl–Titchmarsh m -functions of $\tilde{H}_{(\tilde{\mu}, \tilde{\sigma})}$ in terms of those of H , one then obtains the following result.

THEOREM 2.2. *Assume Hypothesis 2.1 and $z \in \mathbb{C} \setminus \mathbb{R}$. Let H and $\tilde{H}_{(\tilde{\mu}, \tilde{\sigma})}$ be given by (2.2) and (2.8), respectively, and denote by m_\pm and $\tilde{m}_{(\tilde{\mu}, \tilde{\sigma}), \pm}$ the corresponding m -functions associated with the half-lines $(x_0, \pm\infty)$. Then,*

$$\tilde{m}_{(\tilde{\mu}, \tilde{\sigma}), \pm}(z, x_0) = \frac{z - \mu}{z - \tilde{\mu}} m_\pm(z, x_0) - \frac{\tilde{\mu} - \mu}{z - \tilde{\mu}} m_{-\tilde{\sigma}}(\tilde{\mu}, x_0), \quad \tilde{\mu} \neq \mu,$$

$$\tilde{m}_{(\tilde{\mu}, \tilde{\sigma}), \pm}(z, x_0) = m_\pm(z, x_0) - \left(\int_{\sigma\infty}^{x_0} dx \phi(\mu, x)^2 \right)^{-1} \frac{1}{z - \mu}, \quad (\tilde{\mu}, \tilde{\sigma}) = (\mu, -\sigma).$$

Given the fundamental relation between $\tilde{m}_{(\tilde{\mu}, \tilde{\sigma}), \pm}$ and m_\pm in Theorem 2.2, one can now readily derive the ensuing relation between the corresponding spectral functions $\tilde{\rho}_{(\tilde{\mu}, \tilde{\sigma}), \pm}$ and ρ_\pm associated with the half-line Dirichlet operators $\tilde{H}_{(\tilde{\mu}, \tilde{\sigma}), \pm, x_0}^D$ and H_{\pm, x_0}^D . For this and a complete spectral characterization of $\tilde{H}_{(\tilde{\mu}, \tilde{\sigma}), \pm, x_0}^D$ in terms of H_{\pm, x_0}^D we refer to [100].

Next we turn to the principal results of [100] including explicit computations of the Weyl–Titchmarsh and spectral matrices of $\tilde{H}_{(\tilde{\mu}, \tilde{\sigma})}$ in terms of those of H and a complete spectral characterization of $\tilde{H}_{(\tilde{\mu}, \tilde{\sigma})}$ and $\tilde{H}_{(\tilde{\mu}, \tilde{\sigma}), x_0}^D$ in terms of H and $H_{x_0}^D$.

We start with the Weyl–Titchmarsh matrices for H and $\tilde{H}_{(\tilde{\mu}, \tilde{\sigma})}$. To fix notation, we introduce the Weyl–Titchmarsh M -matrix in \mathbb{C}^2 associated with H by

$$M(z, x_0) = (M_{p,q}(z, x_0))_{1 \leq p, q \leq 2} = [m_-(z, x_0) - m_+(z, x_0)]^{-1} \times \begin{pmatrix} m_-(z, x_0)m_+(z, x_0) & [m_-(z, x_0) + m_+(z, x_0)]/2 \\ [m_-(z, x_0) + m_+(z, x_0)]/2 & 1 \end{pmatrix}, \quad z \in \mathbb{C} \setminus \mathbb{R},$$

and similarly $\tilde{M}_{(\tilde{\mu}, \tilde{\sigma})}$ in connection with $\tilde{H}_{(\tilde{\mu}, \tilde{\sigma})}$. An application of Theorem 2.2 then yields

THEOREM 2.3. *Assume Hypothesis 2.1 and $z \in \mathbb{C} \setminus \mathbb{R}$. Let H and $\tilde{H}_{(\tilde{\mu}, \tilde{\sigma})}$ be given by (2.3) and (2.32), respectively. Then the corresponding Weyl–Titchmarsh matrices M and $\tilde{M}_{(\tilde{\mu}, \tilde{\sigma})}$ are related by*

$$\begin{aligned} \tilde{M}_{(\tilde{\mu}, \tilde{\sigma}), 1,1}(z, x_0) &= \frac{z - \mu}{z - \tilde{\mu}} M_{1,1}(z, x_0) - 2 \frac{\tilde{\mu} - \mu}{z - \tilde{\mu}} m_{-\tilde{\sigma}}(\tilde{\mu}, x_0) M_{1,2}(z, x_0) \\ &\quad + \frac{(\tilde{\mu} - \mu)^2}{(z - \mu)(z - \tilde{\mu})} m_{-\tilde{\sigma}}(\tilde{\mu}, x_0)^2 M_{2,2}(z, x_0), \\ \tilde{M}_{(\tilde{\mu}, \tilde{\sigma}), 1,2}(z, x_0) &= M_{1,2}(z, x_0) - \frac{\tilde{\mu} - \mu}{z - \mu} m_{-\tilde{\sigma}}(\tilde{\mu}, x_0) M_{2,2}(z, x_0), \end{aligned}$$

$$\tilde{M}_{(\tilde{\mu}, \tilde{\sigma}), 2, 2}(z, x_0) = \frac{z - \tilde{\mu}}{z - \mu} M_{2, 2}(z, x_0), \quad \tilde{\mu} \neq \mu.$$

Given the basic connection between $\tilde{M}_{(\tilde{\mu}, \tilde{\sigma})}$ and M in Theorem 2.3, one can now proceed to derive the analogous relations between the spectral matrices $\tilde{\rho}_{(\tilde{\mu}, \tilde{\sigma})}$ and ρ associated with $\tilde{H}_{(\tilde{\mu}, \tilde{\sigma})}$ and H , respectively (cf. [100] for details).

The principal spectral deformation result of [100] then reads as follows.

THEOREM 2.4. *Assume Hypothesis 2.1. Then,*

(i) *Suppose $\mu, \tilde{\mu} \in (E_0, E_1)$. Then $\tilde{H}_{(\tilde{\mu}, \tilde{\sigma})}$ and H are unitarily equivalent. Moreover, $\tilde{H}_{(\tilde{\mu}, \tilde{\sigma}), x_0}^D$ and $H_{x_0}^D$, restricted to the orthogonal complements of the one-dimensional eigenspaces corresponding to $\tilde{\mu}$ and μ , are unitarily equivalent.*

(ii) *Assume $\mu \in \{E_0, E_1\} \cap \sigma_d(H)$, $\tilde{\mu} \in (E_0, E_1)$. Then,*

$$\sigma_{(p)}(\tilde{H}_{(\tilde{\mu}, \tilde{\sigma})}) = \sigma_{(p)}(H) \setminus \{\mu\}, \quad \sigma_{(p)}(\tilde{H}_{(\tilde{\mu}, \tilde{\sigma}), x_0}^D) = \{\sigma_{(p)}(H_{x_0}^D) \setminus \{\mu\}\} \cup \{\tilde{\mu}\}.$$

(iii) *Suppose $\mu \in (E_0, E_1)$, $\tilde{\mu} \in \{E_0, E_1\} \cap \sigma_d(H)$. Then,*

$$\sigma_{(p)}(\tilde{H}_{\tilde{\mu}}) = \sigma_{(p)}(H) \setminus \{\tilde{\mu}\}, \quad \sigma_{(p)}(\tilde{H}_{\tilde{\mu}, x_0}^D) = \sigma_{(p)}(H_{x_0}^D) \setminus \{\mu\}, \quad \tilde{\mu} \notin \sigma(\tilde{H}_{\tilde{\mu}, x_0}^D).$$

(iv) *Assume $\mu, \tilde{\mu} \in \{E_0, E_1\} \cap \sigma_d(H)$, $\mu \neq \tilde{\mu}$. Then,*

$$\sigma_{(p)}(\tilde{H}_{\tilde{\mu}}) = \sigma_{(p)}(H) \setminus \{E_0, E_1\}, \quad \sigma_{(p)}(\tilde{H}_{\tilde{\mu}, x_0}^D) = \sigma_{(p)}(H_{x_0}^D) \setminus \{\mu\}, \quad \tilde{\mu} \notin \sigma(\tilde{H}_{\tilde{\mu}, x_0}^D).$$

In cases (ii)–(iv), the corresponding pairs of operators, restricted to the obvious orthogonal complements of the eigenspaces corresponding to μ and/or $\tilde{\mu}$, are unitarily equivalent. In particular,

$$\sigma_{\text{ess, ac}, \mu, \text{sc}}(\tilde{H}_{(\tilde{\mu}, \tilde{\sigma})}) = \sigma_{\text{ess, ac}, \text{sc}}(\tilde{H}_{(\tilde{\mu}, \tilde{\sigma}), x_0}^D) = \sigma_{\text{ess, ac}, \text{sc}}(H_{x_0}^D) = \sigma_{\text{ess, ac}, \text{sc}}(H).$$

REMARK 2.5. (i) Perhaps the most important consequence of Theorem 2.4 (i), from an inverse spectral point of view, is the fact that *any* finite number of deformations of Dirichlet data within spectral gaps of V yields a potential \tilde{V} in the isospectral class of V . No further constraints on $(\mu_j, \sigma_j), (\tilde{\mu}_j, \tilde{\sigma}_j)$, other than $(\mu_j, \sigma_j), (\tilde{\mu}_j, \tilde{\sigma}_j) \in (E_{j-1}, E_j) \times \{-, +\}, (E_{j-1}, E_j) \in \mathbb{R} \setminus \sigma(H), j = 1, \dots, N, N \in \mathbb{N}$, are involved.

(ii) The isospectral property (i) in Theorem 2.4, in the special case of periodic potentials, was first proved by Finkel, Isaacson, and Trubowitz [63]. Further results can be found in Buys and Finkel [28] and Iwasaki [132] (see also [46], [48], [174], [175], [176], [177]). Similar constructions for Schrödinger operators on a compact interval can be found in Pöschel and Trubowitz [192, Chs. 3, 4] and Ralston and Trubowitz [194].

(iii) Let $\mu \in (E_0, E_1)$. Then the (isospectral) Dirichlet deformation $(\mu, \sigma) \rightarrow (\mu, -\sigma)$ is precisely the isospectral case of the double commutation method (cf. [74], [78], [100, App. B], [103]). It simply flips the Dirichlet eigenvalue μ on the half-line $(x_0, \sigma\infty)$ to the other half-line $(x_0, -\sigma\infty)$. In the special case where $V(x)$ is periodic, this procedure was first used by McKean and van Moerbeke [178].

(iv) The topology of these Dirichlet data strongly depends on the nature of the endpoints E_0, E_1 of a particular spectral gap. For instance, in cases like the periodic one, different spectral gaps are separated by intervals of absolutely continuous spectrum and the two intervals $[E_0, E_1]$ together with $\sigma \in \{-, +\}$ can be identified with a circle (upon identifying the two intervals as two rims of a cut). Globally this then leads to a product of circles, that is, a torus. In particular, the

Dirichlet eigenvalues in different spectral gaps can be prescribed independently of each other. The situation is entirely different if an endpoint, say E_1 , belongs to the discrete spectrum of H . In this case there are two neighboring spectral gaps (E_0, E_1) and (E_1, E_2) and the two Dirichlet eigenvalues $\mu_j \in (E_{j-1}, E_j)$, $j = 1, 2$, are not independent of each other. In fact, if one of μ_1 or μ_2 approaches E_1 , then necessarily so does the other. The topology is then not a product of circles. For instance, a closer analysis of the case of N -soliton potentials in (2.1) then illustrates that the appropriate coordinates parametrizing the N -soliton isospectral class are $\alpha_j \in (0, \infty)$ (compare also with positive norming constants), which results globally in a product of open half-lines.

REMARK 2.6. In certain cases where the base (background) potential V is reflectionless (see, e.g., [106]) and H is bounded from below and has no singularly continuous spectrum, the isospectral class $\text{Iso}(V)$ of V (the set of all reflectionless \tilde{V} 's such that $\sigma(\tilde{H}) = \sigma(H)$) is completely characterized by the distribution of Dirichlet (initial) data $(\mu_{j+1}(x_0), \sigma_{j+1}(x_0)) \in [E_j, E_{j+1}] \times \{-, +\}$, $j \in J$, in nontrivial spectral gaps of H . Here $x_0 \in \mathbb{R}$ is a fixed reference point and $J = \{0, 1, \dots, N-1\}$, $N \in \mathbb{N}$, or $j \in J = \mathbb{N}_0 (= \mathbb{N} \cup \{0\})$ parametrizes these nontrivial spectral gaps (E_j, E_{j+1}) of H (the trivial one being $(-\infty, \inf \sigma(H))$). Prime examples of this type are periodic potentials, algebro-geometric quasi-periodic finite-band potentials (cf. [15, Ch. 3], [79, Ch. 1], [159, Chs. 8–12], [172, Ch. 4], [183, Ch. II]), and certain limiting cases thereof (e.g., soliton potentials). In these cases, an iteration of the Dirichlet deformation method, in the sense that $(\mu_{j+1}(x_0), \sigma_{j+1}(x_0)) \rightarrow (\tilde{\mu}_{j+1}(x_0), \tilde{\sigma}_{j+1}(x_0))$ within $[E_j, E_{j+1}] \times \{-, +\}$ for each $j \in J$, independently of each other, yields an explicit realization of the underlying isospectral class $\text{Iso}(V)$ of reflectionless potentials with base V . In the periodic case, this was first proved by Finkel, Isaacson, and Trubowitz [63] (see also [28], [132]). More precisely, the inclusion of limiting cases $\mu_{j+1}(x_0) \in \{E_j, E_{j+1}\} \cap \sigma_{\text{ess}}(H)$ requires a special argument (since it is excluded by Hypothesis 2.1) but this can be provided in the special cases at hand.

REMARK 2.7. Another case of primary interest concerns potentials V with purely discrete spectra bounded from below, that is,

$$\sigma(H) = \sigma_d(H) = \{E_j\}_{j \in \mathbb{N}_0}, \quad -\infty < E_0, E_j < E_{j+1}, j \in \mathbb{N}_0, \quad \sigma_{\text{ess}}(H) = \emptyset.$$

(For simplicity, one may think in terms of the harmonic oscillator $V(x) = x^2$, [32]–[36], [99], [160], [180], [193].) In this case, either

$$(\mu_{j+1}(x_0), \sigma_{j+1}(x_0)) \in (E_j, E_{j+1}) \times \{-, +\} \text{ or } \mu_{j+1}(x_0) = E_{j+1} = \mu_{j+2}(x_0),$$

that is, Dirichlet eigenvalues necessarily meet in pairs whenever they hit an eigenvalue of H . The following trace formula for V in terms of $\sigma(H) = \{E_j\}_{j \in \mathbb{N}_0}$ and $\sigma(H_x^D) = \{\mu_j(x)\}_{j \in \mathbb{N}}$ (with H_y^D the Dirichlet operator associated with $\tau = -\frac{d^2}{dx^2} + V(x)$ and a Dirichlet boundary condition at $x = y$), proved in [91] (cf. Section 4),

$$V(x) = E_0 + \lim_{\alpha \downarrow 0} \alpha^{-1} \sum_{j=1}^{\infty} \left(2e^{-\alpha \mu_j(x)} - e^{-\alpha E_j} - e^{-\alpha E_{j+1}} \right), \tag{2.10}$$

then shows one crucial difference to the periodic-type cases mentioned previously. Unlike in the periodic case, though, the initial Dirichlet eigenvalues $\mu_{j+1}(x_0)$ cannot be prescribed arbitrarily in the spectral gaps (E_j, E_{j+1}) of H . Indeed, the fact

that the Abelian regularization in the trace formula (2.10) for $V(x)$ converges to a limit restricts the asymptotic distribution of $\mu_{j+1}(x) \in [E_j, E_{j+1}]$ as $j \rightarrow \infty$. For instance, consider $V(x) = x^2 - 1$, then $E_j = 2j$, $j \in \mathbb{N}_0$. The choice $\mu_j(x_0) = 2j - \gamma$, $\gamma \in (0, 1)$, then yields for the Abelian regularization on the right-hand side of (2.10),

$$\lim_{\alpha \downarrow 0} \alpha^{-1} \sum_{j=1}^{\infty} \left(2e^{-\alpha(2j-\gamma)} - e^{-\alpha 2j} - e^{-\alpha(2j-2)} \right) = \lim_{\alpha \downarrow 0} 2[(\gamma - 1) + O(\alpha)] \frac{1}{1 - e^{-2\alpha}} = \infty.$$

Put differently, our choice of $\mu_j(x_0) = 2j - \gamma$, $\gamma \in (0, 1)$, was not an admissible choice of Dirichlet eigenvalues for the (shifted) harmonic oscillator potential $V(x) = x^2 - 1$. However, as stressed in Remark 2.5 (i), one of the fundamental consequences of [100] concerns the fact that there is no such restriction for any finite number of spectral gaps of H . In other words, only the tail end of the Dirichlet eigenvalues $\mu_{j+1}(x_0)$ as $j \rightarrow \infty$ is restricted (the precise nature of this restriction being unknown at this point), any finite number of them can be placed arbitrarily in the spectral gaps (E_j, E_{j+1}) (with the obvious “crossing” restrictions at the common boundary E_{j+1} of (E_j, E_{j+1}) and (E_{j+1}, E_{j+2})). The only other known restriction to date on Dirichlet initial data $(\mu_j(x_0), \sigma_j(x_0))$ is that $\sigma_j(x_0) = -$ and $\sigma_j(x_0) = +$ infinitely often, that is, both half-lines $(-\infty, x_0)$ and (x_0, ∞) support (naturally) infinitely many Dirichlet eigenvalues.

For various extensions of the results presented, including a careful discussion of limit point/limit circle properties of the Dirichlet deformation operators, iterations of DDM to insert finitely many eigenvalues in spectral gaps, applications to reflectionless Schrödinger operators, general Sturm–Liouville operators in a weighted L^2 -space, applications to short-range scattering theory, and a concise summary of single and double commutation methods, we refer to [100].

More recent references: An interesting refinement of Theorem 2.4 (i), in which a unitary operator relating $\tilde{H}_{(\tilde{\mu}, \tilde{\sigma})}$ and H is explicitly characterized, is due to Schmincke [217]. DDM for one-dimensional Jacobi and Dirac-type operators has been worked out by Teschl [238], [243, Ch. 11] (see also [104], [242]), [240].

3. Renormalized Oscillation Theory

In this section we summarize some of the principal results of the paper:

[101] F. Gesztesy, B. Simon, and G. Teschl, *Zeros of the Wronskian and renormalized oscillation theory*, Amer. J. Math. **118**, 571–594 (1996).

For over a hundred and fifty years, oscillation theorems for second-order differential equations have fascinated mathematicians. Originating with Sturm’s celebrated memoir [232], extended in a variety of ways by Bôcher [21] and others, a large body of material has been accumulated since then (thorough treatments can be found, e.g., in [42], [147], [203], [235], and the references therein). In [101] a new wrinkle to oscillation theory was added by showing that zeros of Wronskians can be used to count eigenvalues in situations where a naive use of oscillation theory would give $\infty - \infty$ (i.e., Wronskians lead to renormalized oscillation theory). In a nutshell, we will show in this section the following result for general Sturm–Liouville operators H in $L^2((a, b); r dx)$ with separated boundary conditions at a and b : If $E_{1,2} \in \mathbb{R}$ and if $u_{1,2}$ solve the differential equation $Hu_j = E_j u_j$, $j = 1, 2$

and respectively satisfy the boundary condition on the left/right, then the dimension of the spectral projection $P_{(E_1, E_2)}(H)$ of H equals the number of zeros of the Wronskian of u_1 and u_2 .

The main motivation in writing [101] had its origins in attempts to provide a general construction of isospectral potentials for one-dimensional Schrödinger operators following previous work by Finkel, Isaacson, and Trubowitz [63] (see also [28]) in the special case of periodic potentials. In fact, in the case of periodic Schrödinger operators H , the nonvanishing of $W(u_1, u_2)(x)$ for Floquet solutions $u_1 = \psi_{\varepsilon_1}(E_1)$, $u_2 = \psi_{\varepsilon_2}(E_2)$, $\varepsilon_{1,2} \in \{+, -\}$ of H_p , for E_1 and E_2 in the same spectral gap of H , is proved in [63]. The extension of these ideas to general one-dimensional Schrödinger operators was done in [100] and is reviewed in Section 2 of this survey. So while [101] is not directly related to the overarching inverse spectral theory topic of this survey, we decided to include it because of its relevance in connection with Section 2.

To set the stage, we consider Sturm–Liouville differential expressions of the form

$$(\tau u)(x) = r(x)^{-1}[-(p(x)u'(x))' + q(x)u(x)], \quad x \in (a, b), \quad -\infty \leq a < b \leq \infty$$

where

$$r, p^{-1}, q \in L^1_{\text{loc}}((a, b)) \text{ are real-valued and } r, p > 0 \text{ a.e. on } (a, b).$$

We shall use τ to describe the formal differentiation expression and H for the operator in $L^2((a, b); r dx)$ given by τ with separated boundary conditions at a and/or b .

If a (resp., b) is finite and q, p^{-1}, r are in addition integrable near a (resp., b), a (resp., b) is called a *regular end point*. τ (resp., H) is called *regular* if both a and b are regular. As is usual ([57, Sect. XIII.2], [182, Sect. 17], [249, Ch. 3]), we consider the local domain

$$D_{\text{loc}} = \{u \in AC_{\text{loc}}((a, b)) \mid pu' \in AC_{\text{loc}}((a, b)), \tau u \in L^2_{\text{loc}}((a, b); r dx)\},$$

where $AC_{\text{loc}}((a, b))$ is the set of locally absolutely continuous functions on (a, b) . General ODE theory shows that for any $E \in \mathbb{C}$, $x_0 \in (a, b)$, and $(\alpha, \beta) \in \mathbb{C}^2$, there is a unique $u \in D_{\text{loc}}$ such that $-(pu')' + qu - Eru = 0$ for a.e. $x \in (a, b)$ and $(u(x_0), (pu')(x_0)) = (\alpha, \beta)$.

The maximal and minimal operators are defined by taking

$$\text{dom}(T_{\text{max}}) = \{u \in L^2((a, b); r dx) \cap D_{\text{loc}} \mid \tau u \in L^2((a, b); r dx)\},$$

with

$$T_{\text{max}}u = \tau u.$$

T_{min} is the operator closure of $T_{\text{max}} \upharpoonright D_{\text{loc}} \cap \{u \text{ has compact support in } (a, b)\}$. Then T_{min} is symmetric and $T_{\text{min}}^* = T_{\text{max}}$.

According to Weyl’s theory of self-adjoint extensions ([57, Sect. XIII.6], [182, Sect. 18], [201, App. to X.1], [248, Section 8.4], [249, Chs. 4, 5]), the deficiency indices of T_{min} are $(0, 0)$ or $(1, 1)$ or $(2, 2)$ depending on whether it is limit point at both, one, or neither endpoint. Moreover, the self-adjoint extensions can be described in terms of Wronskians ([57, Sect. XIII.2], [182, Sects. 17, 18], [248, Sect. 8.4], [249, Ch. 3]). Define

$$W(u_1, u_2)(x) = u_1(x)(pu'_2)(x) - (pu'_1)(x)u_2(x).$$

Then if T_{\min} is limit point at both ends, $T_{\min} = T_{\max} = H$. If T_{\min} is limit point at b but not at a , for H any self-adjoint extension of T_{\min} , if φ_- is any function in $\text{dom}(H) \setminus \text{dom}(T_{\min})$, then

$$\text{dom}(H) = \{u \in \text{dom}(T_{\max}) \mid W(u, \varphi_-)(x) \rightarrow 0 \text{ as } x \downarrow a\}.$$

Finally, if u_1 is limit circle at both ends, the operators H with separated boundary conditions are those for which we can find $\varphi_{\pm} \in \text{dom}(H)$, $\varphi_+ = 0$ near a , $\varphi_- = 0$ near b , and $\varphi_{\pm} \in \text{dom}(H) \setminus \text{dom}(T_{\min})$. In that case,

$$\text{dom}(H) = \{u \in D(T_{\max}) \mid W(u, \varphi_-)(x) \rightarrow 0 \text{ as } x \downarrow a, W(u, \varphi_+)(x) \rightarrow 0 \text{ as } x \uparrow b\}.$$

Of course, if H is regular, we can just specify the boundary conditions by taking values at a, b since by regularity any $u \in \text{dom}(T_{\max})$ has u, pu' continuous on $[a, b]$. It follows from this analysis that

$$\text{if } u_{1,2} \in \text{dom}(H), \text{ then } W(u_1, u_2)(x) \rightarrow 0 \text{ as } x \rightarrow a \text{ or } b.$$

Such operators will be called SL operators (for Sturm–Liouville, but SL includes separated boundary conditions (if necessary)) and denoted by H .

It will be convenient to write $\ell_- = a, \ell_+ = b$.

Throughout this section we will denote by $\psi_{\pm}(z, x) \in D_{\text{loc}}$ solutions of $\tau\psi = z\psi$ so that $\psi_{\pm}(z, \cdot)$ is L^2 at ℓ_{\pm} and $\psi_{\pm}(z, \cdot)$ satisfies the appropriate boundary condition at ℓ_{\pm} in the sense that for any $u \in \text{dom}(H)$, $\lim_{x \rightarrow \ell_{\pm}} W(\psi_{\pm}(z), u)(x) = 0$. If $\psi_{\pm}(z, \cdot)$ exist, they are unique up to constant multiples. In particular, $\psi_{\pm}(z, \cdot)$ exist for z not in the essential spectrum of H and we can assume them to be holomorphic with respect to z in $\mathbb{C} \setminus \sigma(H)$ and real for $z \in \mathbb{R}$. One can choose

$$\psi_{\pm}(z, x) = ((H - z)^{-1} \chi_{(c,d)})(x) \text{ for } x < c \text{ and } x > d, \quad a < c < d < b \quad (3.1)$$

and uniquely continue $\psi_{\pm}(z, x)$ for $x > c$ and $x < d$. Here $(H - z)^{-1}$ denotes the resolvent of H and χ_{Ω} the characteristic function of the set $\Omega \subseteq \mathbb{R}$. Clearly we can include a finite number of isolated eigenvalues in the domain of holomorphy of ψ_{\pm} by removing the corresponding poles. Moreover, to simplify notations, all solutions u of $\tau u = Eu$ are understood to be not identically vanishing and solutions associated with real values of the spectral parameter E are assumed to be real-valued in this paper. Thus if E is real and in the resolvent set for H or an isolated eigenvalue, we are guaranteed there are solutions that satisfy the boundary conditions at a or b . If E is in the essential spectrum, it can happen that such solutions do not exist or it may happen that they do. In Theorems 3.15 and 3.16 below, we shall explicitly assume such solutions exist for the energies of interest. If these energies are not in the essential spectrum, that is automatically fulfilled.

The key idea in [101] is to look at zeros of the Wronskian. That zeros of the Wronskian are related to oscillation theory is indicated by an old paper of Leighton [151], who noted that if $u_j, pu'_j \in AC_{\text{loc}}((a, b))$, $j = 1, 2$ and u_1 and u_2 have a nonvanishing Wronskian $W(u_1, u_2)$ in (a, b) , then their zeros must intertwine each other. (In fact, pu'_1 must have opposite signs at consecutive zeros of u_1 , so by nonvanishing of W , u_2 must have opposite signs at consecutive zeros of u_1 as well. Interchanging the role of u_1 and u_2 yields strict interlacing of their zeros.) Moreover, let $E_1 < E_2$ and $\tau u_j = E_j u_j$, $j = 1, 2$. If x_0, x_1 are two consecutive zeros of u_1 , then the number of zeros of u_2 inside (x_0, x_1) is equal to the number of zeros of the Wronskian $W(u_1, u_2)$ plus one (cf. Theorem 3.20). Hence the Wronskian comes

with a built-in renormalization counting the additional zeros of u_2 in comparison to u_1 .

We let $W_0(u_1, u_2)$ be the number of zeros of the Wronskian in the open interval (a, b) not counting multiplicities of zeros. Given $E_1 < E_2$, we let $N_0(E_1, E_2) = \dim(\text{ran}(P_{(E_1, E_2)}(H)))$ be the dimension of the spectral projection $P_{(E_1, E_2)}(H)$ of H .

We begin by presenting two aspects of zeros of the Wronskian which are critical for the two halves of our proofs (i.e., for showing $N_0 \geq W_0$ and that $N_0 \leq W_0$). First, the vanishing of the Wronskian lets us patch solutions together:

LEMMA 3.1. *Suppose that $\psi_{+,j}, \psi_- \in D_{\text{loc}}$ and that $\psi_{+,j}$ and $\tau\psi_{+,j}$, $j = 1, 2$ are in $L^2((c, b))$ and that ψ_- and $\tau\psi_-$ are in $L^2((a, c))$ for all $c \in (a, b)$. Suppose, in addition, that $\psi_{+,j}$, $j = 1, 2$ satisfy the boundary condition defining H at b (i.e., $W(u, \psi_{+,j})(c) \rightarrow 0$ as $c \uparrow b$ for all $u \in \text{dom}(H)$) and similarly, that ψ_- satisfies the boundary condition at a . Then,*

(i) *If $W(\psi_{+,1}, \psi_{+,2})(c) = 0$ and $(\psi_{+,2}(c), (p\psi'_{+,2})(c)) \neq (0, 0)$, then there exists a γ such that*

$$\eta = \chi_{[c,b]}(\psi_{+,1} - \gamma\psi_{+,2}) \in \text{dom}(H)$$

and

$$H\eta = \chi_{[c,b]}(\tau\psi_{+,1} - \gamma\tau\psi_{+,2}).$$

(ii) *If $W(\psi_{+,1}, \psi_-)(c) = 0$ and $(\psi_-(c), (p\psi'_-)(c)) \neq (0, 0)$, then there is a γ such that*

$$\eta = \gamma\chi_{(a,c]}\psi_- + \chi_{(c,b)}\psi_{+,1} \in \text{dom}(H)$$

and

$$H\eta = \gamma\chi_{(a,c]}\tau\psi_- + \chi_{(c,b)}\tau\psi_{+,1}.$$

The second aspect connects zeros of the Wronskian to Prüfer variables ρ_u, θ_u (for u, pu' continuous) defined by

$$u(x) = \rho_u(x) \sin(\theta_u(x)), \quad (pu')(x) = \rho_u(x) \cos(\theta_u(x)).$$

If $(u(x), (pu')(x))$ is never $(0, 0)$, then ρ_u can be chosen positive and θ_u is uniquely determined once a value of $\theta_u(x_0)$ is chosen subject to the requirement that θ_u be continuous in x .

Notice that

$$W(u_1, u_2)(x) = \rho_{u_1}(x)\rho_{u_2}(x) \sin(\theta_{u_1}(x) - \theta_{u_2}(x)).$$

Thus, one obtains the following results.

LEMMA 3.2. *Suppose (u_j, pu'_j) , $j = 1, 2$ are never $(0, 0)$. Then $W(u_1, u_2)(x_0)$ is zero if and only if $\theta_{u_1}(x_0) = \theta_{u_2}(x_0) \pmod{\pi}$.*

In linking Prüfer variables to rotation numbers, an important role is played by the observation that because of

$$u(x) = \int_{x_0}^x dt \frac{\rho_u(t) \cos(\theta_u(t))}{p(t)},$$

$\theta_u(x_0) = 0 \pmod{\pi}$ implies $[\theta_u(x) - \theta_u(x_0)]/(x - x_0) > 0$ for $0 < |x - x_0|$ sufficiently small and hence for all $0 < |x - x_0|$ if $(u, pu') \neq (0, 0)$. (In fact, suppose $x_1 \neq x_0$ is the closest x such that $\theta_u(x_1) = \theta_u(x_0)$ then apply the local result at x_1 to obtain a contradiction.) We summarize:

LEMMA 3.3. (i) If $(u, pu') \neq (0, 0)$ then $\theta_u(x_0) = 0 \pmod{\pi}$ implies

$$[\theta_u(x) - \theta_u(x_0)] / (x - x_0) > 0$$

for $x \neq x_0$. In particular, if $\theta_u(c) \in [0, \pi)$ and u has n zeros in (c, d) , then $\theta_u(d - \epsilon) \in (n\pi, (n + 1)\pi)$ for sufficiently small $\epsilon > 0$.

(ii) Let $E_1 < E_2$ and assume that $u_{1,2}$ solve $\tau u_j = E_j u_j$, $j = 1, 2$. Let $\Delta(x) = \theta_{u_2}(x) - \theta_{u_1}(x)$. Then $\Delta(x_0) = 0 \pmod{\pi}$ implies $(\Delta(x) - \Delta(x_0)) / (x - x_0) > 0$ for $0 < |x - x_0|$.

REMARK 3.4. (i) Suppose r, p are continuous on (a, b) . If $\theta_{u_1}(x_0) = 0 \pmod{\pi}$ then $\theta_{u_1}(x) - \theta_{u_1}(x_0) = c_0(x - x_0) + o(x - x_0)$ with $c_0 > 0$. If $\Delta(x_0) = 0 \pmod{\pi}$ and $\theta_{u_1}(x_0) \neq 0 \pmod{\pi}$, then $\Delta(x) - \Delta(x_0) = c_1(x - x_0) + o(x - x_0)$ with $c_1 > 0$. If $\theta_{u_1}(x_0) = 0 = \Delta(x_0) \pmod{\pi}$, then $\Delta(x) - \Delta(x_0) = c_2(x - x_0)^3 + o(x - x_0)^3$ with $c_2 > 0$. Either way, Δ increases through x_0 . (In fact, $c_0 = p(x_0)^{-1}$, $c_1 = (E_2 - E_1)r(x_0) \sin^2(\theta_{u_1}(x_0))$ and $c_2 = \frac{1}{3}r(x_0)p(x_0)^{-2}(E_2 - E_1)$).

(ii) In other words, Lemma 3.3 implies that the integer parts of θ_u/π and $\Delta_{u,v}/\pi$ are increasing with respect to $x \in (a, b)$ (even though θ_u and $\Delta_{u,v}$ themselves are not necessarily monotone in x).

(iii) Let $E \in [E_1, E_2]$ and assume $[E_1, E_2]$ to be outside the essential spectrum of H . Then, for $x \in (a, b)$ fixed,

$$\frac{d\theta_{\psi_{\pm}}}{dE}(E, x) = -\frac{\int_x^{\ell_{\pm}} dt \psi_{\pm}(E, t)^2}{\rho_{\psi_{\pm}}(E, x)}$$

proves that $\mp\theta_{\psi_{\pm}}(E, x)$ is strictly increasing with respect to E .

We continue with some preparatory results in the regular case.

LEMMA 3.5. Assume H to be a regular SL operator.

(i) Let $u_{1,2}$ be eigenfunctions of H with eigenvalues $E_1 < E_2$ and let ℓ be the number of eigenvalues of H in (E_1, E_2) . Then $W(u_1, u_2)(x)$ has exactly ℓ zeros in (a, b) .

(ii) Let $E_1 \leq E_2$ be eigenvalues of H and suppose $[E_1, E_2]$ has ℓ eigenvalues. Then for $\epsilon \geq 0$ sufficiently small, $W_0(\psi_-(E_1 - \epsilon), \psi_+(E_2 + \epsilon)) = \ell$.

(iii) Let $E_3 < E_4 < E$ and u be any solution of $\tau u = Eu$. Then,

$$W_0(\psi_-(E_3), u) \geq W_0(\psi_-(E_4), u). \tag{3.2}$$

Similarly, if $E_3 > E_4 > E$ and u is any solution of $\tau u = Eu$, then (3.2) holds.

(iv) Item (iii) remains true if every ψ_- is replaced by a ψ_+ .

REMARK 3.6. (i) Since (E_1, E_2) has $\ell - 2$ eigenvalues, Lemma 3.5 (i) implies that the Wronskian $W(\psi_-(E_1), \psi_+(E_2))(x)$ has $\ell - 2$ zeros in (a, b) and clearly it has zeros at a and b . Essentially, Lemma 3.5 (ii) implies that replacing E_1 by $E_1 - \epsilon$ and E_2 by $E_2 + \epsilon$ moves the zeros at a, b inside (a, b) to give $\ell - 2 + 2 = \ell$ zeros.

(ii) Lemma 3.5 (iv) follows from Lemma 3.5 (iii) upon reflecting at some point $c \in (a, b)$, implying an interchange of ψ_+ and ψ_- .

Lemma 3.5 then implies the following result.

LEMMA 3.7. Let H be a regular SL operator and suppose $E_1 < E_2$. Then,

$$W_0(\psi_-(E_1), \psi_+(E_2)) \geq N_0(E_1, E_2).$$

Using the approach of Weidmann ([249, Ch. 14]) to control some limits, one can remove the assumption that H is regular in Lemma 3.7 as follows.

Fix functions $u_1, u_2 \in D_{\text{loc}}$. Pick $c_n \downarrow a, d_n \uparrow b$. Define \tilde{H}_n on $L^2((c_n, d_n); r dx)$ by imposing the following boundary conditions on $\eta \in \text{dom}(\tilde{H}_n)$

$$W(u_1, \eta)(c_n) = 0 = W(u_2, \eta)(d_n).$$

On $L^2((a, b); r dx) = L^2((a, c_n); r dx) \oplus L^2((c_n, d_n); r dx) \oplus L^2((d_n, b); r dx)$ take $H_n = \alpha I \oplus \tilde{H}_n \oplus \alpha I$ with α a fixed real constant. Then Weidmann proves:

LEMMA 3.8. *Suppose that either H is limit point at a or that u_1 is a $\psi_-(E, x)$ for some E and similarly, that either H is limit point at b or u_2 is a $\psi_+(E', x)$ for some E' . Then H_n converges to H in strong resolvent sense as $n \rightarrow \infty$.*

The idea of Weidmann’s proof is that it suffices to find a core D_0 of H such that for every $\eta \in D_0$ there exists an $n_0 \in \mathbb{N}$ with $\eta \in D_0$ for $n \geq n_0$ and $H_n \eta \rightarrow H \eta$ as n tends to infinity (see [248, Theorem 9.16 (i)]). If H is limit point at both ends, take $\eta \in D_0 = \{u \in D_{\text{loc}} \mid \text{supp}(u) \text{ compact in } (a, b)\}$. Otherwise, pick $\tilde{u}_1, \tilde{u}_2 \in \text{dom}(H)$ with $\tilde{u}_2 = u_2$ near b and $\tilde{u}_2 = 0$ near a and with $\tilde{u}_1 = u_1$ near a and $\tilde{u}_1 = 0$ near b . Then pick $\eta \in D_0 + \text{span}[\tilde{u}_1, \tilde{u}_2]$ which one can show is a core for H ([249, Ch. 14]).

Secondly, one uses the following fact:

LEMMA 3.9. *Let $A_n \rightarrow A$ in strong resolvent sense as $n \rightarrow \infty$. Then,*

$$\dim(\text{ran}(P_{(E_1, E_2)}(A))) \leq \varliminf_{n \rightarrow \infty} \dim(\text{ran}(P_{(E_1, E_2)}(A_n))).$$

Combining Lemmas 3.7–3.9 then yields the following result.

THEOREM 3.10. *Let $E_1 < E_2$. If $u_1 = \psi_-(E_1)$ and either $u_2 = \psi_+(E_2)$ or $\tau u_2 = E_2 u_2$ and H is limit point at b . Then,*

$$W_0(u_1, u_2) \geq \dim(\text{ran}((P_{(E_1, E_2)}(H)))).$$

Next, we indicate how the following result can be proved:

THEOREM 3.11. *Let $E_1 < E_2$. Let either $u_1 = \psi_+(E_1)$ or $u_1 = \psi_-(E_1)$ and either $u_2 = \psi_+(E_2)$ or $u_2 = \psi_-(E_2)$. Then,*

$$W_0(u_1, u_2) \leq \dim(\text{ran}(P_{(E_1, E_2)}(H))). \tag{3.3}$$

Let $E_1 < E_2$. Suppose first that $u_1 = \psi_-(E_1)$ and $u_2 = \psi_+(E_2)$. Let x_1, \dots, x_m be zeros of $W(u_1, u_2)(x)$. Suppose we can prove that $\dim P_{(E_1, E_2)}(H) \geq m$. If $W_0(u_1, u_2) = m$, this proves (3.3). If $W_0 = \infty$, we can take m arbitrarily large, and again (3.3) holds. Define

$$\eta_j(x) = \begin{cases} u_1(x), & x \leq x_j, \\ \gamma_j u_2(x), & x \geq x_j, \end{cases} \quad 1 \leq j \leq m,$$

where γ_j is defined such that $\eta_j \in \text{dom}(H)$ by Lemma 3.1. Let

$$\tilde{\eta}_j(x) = \begin{cases} u_1(x), & x \leq x_j, \\ -\gamma_j u_2(x), & x > x_j, \end{cases} \quad 1 \leq j \leq m.$$

If E_2 is an eigenvalue of H , we define in addition $\eta_0 = u_2 = -\tilde{\eta}_0, x_0 = a$ and if E_1 is an eigenvalue of H , $\eta_{m+1} = u_1 = \tilde{\eta}_{m+1}, x_{m+1} = b$.

LEMMA 3.12. $\langle \eta_j, \eta_k \rangle = \langle \tilde{\eta}_j, \tilde{\eta}_k \rangle$ for all j, k , where $\langle \cdot, \cdot \rangle$ is the $L^2((a, b); r dx)$ inner product.

Notice that by (3.2),

$$\left(H - \frac{E_2 + E_1}{2} \right) \eta_j = \left(\frac{E_1 - E_2}{2} \right) \tilde{\eta}_j.$$

This result and Lemma 3.12 imply the following lemma.

LEMMA 3.13. If η is in the span of the η_j , then

$$\left\| \left(H - \frac{E_2 + E_1}{2} \right) \eta \right\| = \frac{|E_2 - E_1|}{2} \|\eta\|.$$

Thus, $\dim(\text{ran}(P_{[E_1, E_2]}(H))) \geq \dim(\text{span}(\{\eta_j\}))$. But u_1 and u_2 are independent on each interval (since their Wronskian is nonconstant) and so the η_j are linearly independent. This proves Theorem 3.11 in the $\psi_-(E_1), \psi_+(E_2)$ case. The case $u_1 = \psi_-(E_1)$, $u_2 = \psi_-(E_2)$ is proved similarly. The cases $u_1 = \psi_+(E_1)$, $u_2 = \psi_\pm(E_2)$ can be obtained by reflection.

Next one proves the following result.

THEOREM 3.14. Let $E_1 \neq E_2$. Let $\tau u_j = E_j u_j$, $j = 1, 2$, $\tau v_2 = E_2 v_2$ with u_2 linearly independent of v_2 . Then the zeros of $W(u_1, u_2)$ interlace the zeros of $W(u_1, v_2)$ and vice versa (in the sense that there is exactly one zero of one function in between two zeros of the other). In particular, $|W_0(u_1, u_2) - W_0(u_1, v_2)| \leq 1$.

Theorems 3.10, 3.11, and 3.14 then yield the following two theorems, the principal results of [101]:

THEOREM 3.15. Suppose $E_1 < E_2$. Let $u_1 = \psi_-(E_1)$ and $u_2 = \psi_+(E_2)$. Then,

$$W_0(u_1, u_2) = N_0(E_1, E_2).$$

THEOREM 3.16. Suppose $E_1 < E_2$. Let $u_1 = \psi_-(E_1)$ and $u_2 = \psi_-(E_2)$. Then either

$$W_0(u_1, u_2) = N_0(E_1, E_2), \tag{3.4}$$

or

$$W_0(u_1, u_2) = N_0(E_1, E_2) - 1. \tag{3.5}$$

If either $N_0 = 0$ or H is limit point at b , then (3.4) holds.

One infers that if b is a regular point and $E_2 > e > E_1$ with e an eigenvalue and $|E_2 - E_1|$ is small, then (3.5) holds rather than (3.4). One also sees that if $u_{1,2}$ are arbitrary solutions of $\tau u_j = E_j u_j$, $j = 1, 2$, then, in general, $|W_0 - N_0| \leq 2$ (this means that if one of the quantities is infinite, the other is as well) and we note that any of $0, \pm 1, \pm 2$ can occur for $W_0 - N_0$. Especially, if either E_1 or E_2 is in the interior of the essential spectrum of H (or $\dim(\text{ran}(P_{(E_1, E_2)}(H))) = \infty$), then $W_0(u_1, u_2) = \infty$ for any u_1 and u_2 satisfying $\tau u_j = E_j u_j$, $j = 1, 2$ (cf. Theorem 3.19).

REMARK 3.17. Of course, by reflecting about a point $c \in (a, b)$, Theorems 3.10, 3.15, and 3.16 hold for $u_1 = \psi_+(E_1)$ and $u_2 = \psi_-(E_2)$ (and either $N_0 = 0$ or H is limit point at a in the corresponding analog of Theorem 3.16 yields (3.4)) and similarly, $\tau u_2 = E_2 u_2$ and H is limit point at a yields the conclusion in the corresponding analog of Theorem 3.10.

We add a few more results proved in [101].

By applying Theorem 3.14 twice, one concludes

THEOREM 3.18. *Let $E_1 \neq E_2$. Let u_1, u_2, v_1, v_2 be the linearly independent functions with $\tau u_j = E_j u_j$ and $\tau v_j = E_j v_j$. Then,*

$$|W_0(u_1, u_2) - W_0(v_1, v_2)| \leq 2.$$

THEOREM 3.19. *If $\dim(\text{ran}(P_{(E_1, E_2)}(H))) = \infty$, then $W_0(u_1, u_2) = \infty$ for any u_1 and u_2 satisfying $\tau u_j = E_j u_j$, $j = 1, 2$.*

THEOREM 3.20. *Let $E_1 < E_2$. Let $\tau u_j = E_j u_j$, $j = 1, 2$. If $a < x_0 < x_1 < b$ are zeros of u_1 or of $W(u_1, u_2)(\cdot)$, then the number of zeros of u_2 inside (x_0, x_1) equals the number of zeros of $W(u_1, u_2)(\cdot)$ inside (x_0, x_1) plus the number of zeros of u_1 inside (x_0, x_1) plus one.*

The following result is of special interest in connection with the problem of whether the total number of eigenvalues of H in one of its essential spectral gaps is finite or infinite. In particular, the energies E_1, E_2 in Theorem 7.5 below may lie in the essential spectrum of H . For this purpose we consider an auxiliary Dirichlet operator $H_{x_0}^D$, $x_0 \in (a, b)$ associated with H . $H_{x_0}^D$ is obtained by taking the direct sum of the restrictions $H_{x_0, \pm}^D$ of H to (a, x_0) (resp., (x_0, b)) with a Dirichlet boundary condition at x_0 . We emphasize that the Dirichlet boundary conditions can be replaced by boundary conditions of the type $\lim_{\epsilon \downarrow 0} [u'(x_0 \pm \epsilon) + \beta u(x_0 \pm \epsilon)] = 0$, $\beta \in \mathbb{R}$.

THEOREM 3.21. *Let $E_1 < E_2$. Let $\tau u_j = E_j u_j$, $\tau s_j = E_j s_j$, and $s_j(E_j, x_0) = 0$, $j = 1, 2$. Then,*

$$\dim(\text{ran}(P_{(E_1, E_2)}(H))) < \infty \text{ if and only if } W_0(u_1, u_2) < \infty,$$

$$\begin{aligned} \dim(\text{ran}(P_{(E_1, E_2)}(H))) - 1 &\leq \dim(\text{ran}(P_{(E_1, E_2)}(H_{x_0}^D))) \\ &\leq \dim(\text{ran}(P_{(E_1, E_2)}(H))) + 2, \end{aligned}$$

$$W_0(s_1, s_2) - 1 \leq \dim(\text{ran}(P_{(E_1, E_2)}(H_{x_0}^D))) \leq W_0(s_1, s_2) + 1.$$

For an application of this circle of ideas to the notion of the density of states, we refer to [101].

More recent references: Oscillation and renormalized oscillation theory was also put in perspective by Simon's contribution [229] to the the Festschrift [8] in honor of Sturm and Liouville. Renormalized oscillation theory for one-dimensional Jacobi and Dirac-type operators was developed by Teschl [237] (see also [243, Sect. 4.3]) and [241]. For an interesting application of some of the results in [101] to the stability theory of complete minimal surfaces, we refer to a paper by Schmidt [215]. For additional results on oscillation theory, critical coupling constants, and eigenvalue asymptotics, we refer to Schmidt [214].

4. Trace Formulas for Schrödinger and Jacobi Operators: The ξ Function

In this section we summarize some of the principal results of the following papers:

[80] F. Gesztesy, H. Holden, and B. Simon, *Absolute summability of the trace relation for certain Schrödinger operators*, Commun. Math. Phys. **168**, 137–161 (1995).

[81] F. Gesztesy, H. Holden, B. Simon, and Z. Zhao, *Trace formulae and inverse spectral theory for Schrödinger operators*, Bull. Amer. Math. Soc. **29**, 250–255 (1993).

[82] F. Gesztesy, H. Holden, B. Simon, and Z. Zhao, *Higher order trace relations for Schrödinger operators*, Rev. Math. Phys. **7**, 893–922 (1995).

[83] F. Gesztesy, H. Holden, B. Simon, and Z. Zhao, *A trace formula for multidimensional Schrödinger operators*, J. Funct. Anal. **141**, 449–465 (1996).

[91] F. Gesztesy and B. Simon, *The ξ function*, Acta Math. **176**, 49–71 (1996).

We start with [91]. One of the principal goals in [91] was to introduce a special function $\xi(\cdot, \cdot)$ on $\mathbb{R} \times \mathbb{R}$ associated with one-dimensional Schrödinger operators H (and Jacobi operators h) which led to a generalization of the known trace formula for periodic Schrödinger operators for general potentials V and established ξ as a new tool in the spectral theory of one-dimensional Schrödinger operators and (multi-dimensional) Jacobi operators.

To illustrate this point we recall the well-known trace formula for periodic potentials V of period $a > 0$. Then, by Floquet theory (see, e.g., [58], [163], [202]),

$$\sigma(H) = [E_0, E_1] \cup [E_2, E_3] \cup \dots$$

a set of bands. If V is $C^1(\mathbb{R})$, one can show that the sum of the gap sizes is finite, that is,

$$\sum_{n=1}^{\infty} |E_{2n} - E_{2n-1}| < \infty. \quad (4.1)$$

For fixed y , let H_y be the operator $-\frac{d^2}{dx^2} + V$ in $L^2([y, y+a])$ with Dirichlet boundary conditions $u(y) = u(y+a) = 0$. Its spectrum is discrete, that is, there are eigenvalues $\{\mu_n(y)\}_{n=1}^{\infty}$ with

$$E_{2n-1} \leq \mu_n(y) \leq E_{2n}. \quad (4.2)$$

The trace formula for V then reads

$$V(y) = E_0 + \sum_{n=1}^{\infty} [E_{2n} + E_{2n-1} - 2\mu_n(y)]. \quad (4.3)$$

By (4.2), $|E_{2n} + E_{2n-1} - 2\mu_n(y)| \leq |E_{2n} - E_{2n-1}|$ so (4.1) implies the convergence of the sum in (4.3). An elegant direct proof of (4.3) can be found, for instance, in [225, Sect. 26].

The earliest trace formula for Schrödinger operators was found on a finite interval in 1953 by Gel'fand and Levitan [71] with later contributions by Dikii [55], Gel'fand [69], Halberg and Kramer [113], and Gilbert and Kramer [109]. The first

trace formula for periodic V was obtained in 1965 by Hochstadt [118], who showed that for finite-band potentials

$$V(x) - V(0) = 2 \sum_{n=1}^g [\mu_n(0) - \mu_n(x)].$$

Dubrovin [9] then proved (4.3) for finite-band potentials. The general formula (4.3) under the hypothesis that V is periodic and in $C^\infty(\mathbb{R})$ was proved in 1975 by Flaschka [65] and McKean and van Moerbeke [178], and later for general $C^3(\mathbb{R})$ potentials by Trubowitz [246]. Formula (4.3) is a key element of the solution of inverse spectral problems for periodic potentials [56], [65], [118], [159, Ch. 11], [172, Sect. 4.3], [178], [179], [246].

There have been two classes of potentials for which (4.3) has been extended. Certain almost periodic potentials are studied in Craig [43], Levitan [158], [159, Ch. 11], and Kotani–Krishna [141].

In 1979, Deift and Trubowitz [48] proved that if $V(x)$ decays sufficiently rapidly at infinity and $-\frac{d^2}{dx^2} + V$ has no negative eigenvalues, then

$$V(x) = \frac{2i}{\pi} \int_{-\infty}^{\infty} dk k \ln \left[1 + R(k) \frac{f_+(x, k)}{f_-(x, k)} \right] \tag{4.4}$$

(where $f_{\pm}(x, k)$ are the Jost functions at energy $E = k^2$ and $R(k)$ is a reflection coefficient) which can be shown to be an analog of (4.3). Previously, Venakides [247] studied a trace formula for V , a positive smooth potential of compact support, by writing (4.3) for the periodic potential $V_L(x) = \sum_{n=-\infty}^{\infty} V(x + nL)$ and then taking L to ∞ . He found an integral formula which is precisely (4.4) (although, this was not identified as such in [247]).

The basic definition of ξ depends on the theory of the Lifshits–Krein spectral shift function [146]. If A and B are self-adjoint operators bounded from below, that is, $A \geq \eta$, $B \geq \eta$ for some real η , and so that $[(A + i)^{-1} - (B + i)^{-1}]$ is trace class, then there exists a measurable function $\xi(\lambda)$ associated with the pair (B, A) so that

$$\text{Tr}[f(A) - f(B)] = - \int_{\mathbb{R}} d\lambda f'(\lambda) \xi(\lambda) \tag{4.5}$$

for a class of functions f which are sufficiently smooth and which decay sufficiently rapidly at infinity, and, in particular, for $f(\lambda) = e^{-t\lambda}$ for any $t > 0$; and so that

$$\xi(\lambda) = 0 \text{ for } \lambda < \eta. \tag{4.6}$$

Moreover, (4.5) and (4.6) uniquely determine $\xi(\lambda)$ for a.e. λ . Moreover, if $[(A + i)^{-1} - (B + i)^{-1}]$ is rank n , then $|\xi(\lambda)| \leq n$ and if $B \geq A$, then $\xi(\lambda) \geq 0$.

For the rank-one case of importance in this paper, an extensive study of ξ can be found in [227].

Let V be a continuous function on \mathbb{R} which is bounded from below. Let $H = -\frac{d^2}{dx^2} + V$ in $L^2(\mathbb{R})$ which is essentially self-adjoint on $C_0^\infty(\mathbb{R})$ and let H_x^D be the operator on $L^2((-\infty, x)) \oplus L^2((x, \infty))$ with $u(x) = 0$ Dirichlet boundary conditions. Then $[(H_x^D + i)^{-1} - (H + i)^{-1}]$ is rank one, so there results a spectral shift function $\xi(\lambda, x)$ for the pair (H_x^D, H) which, in particular, satisfies,

$$\text{Tr}[e^{-tH} - e^{-tH_x^D}] = t \int_0^\infty d\lambda e^{-t\lambda} \xi(\lambda, x). \tag{4.7}$$

While ξ is defined in terms of H and H_x^D , there is a formula that only involves H , or more precisely, the Green's function $G(z, x, y)$ of H defined by

$$((H - z)^{-1}f)(x) = \int_{\mathbb{R}} dy G(z, x, y)f(y), \quad \text{Im}(z) \neq 0.$$

Then by general principles, $\lim_{\epsilon \downarrow 0} G(\lambda + i\epsilon, x, y)$ exists for a.e. $\lambda \in \mathbb{R}$, and

$$\xi(\lambda, x) = \frac{1}{\pi} \text{Arg}(\lim_{\epsilon \downarrow 0} G(\lambda + i\epsilon, x, x)). \tag{4.8}$$

This is *formally* equivalent to formulas that Krein [146] has for ξ but in a singular setting (i.e., corresponding to an infinite coupling constant). With this definition out of the way, we can state the general trace formula derived in [91]:

THEOREM 4.1. *Suppose V is a continuous function bounded from below on \mathbb{R} . Let $\xi(\lambda, x)$ be the spectral shift function for the pair (H_x^D, H) with H_x^D the operator on $L^2((-\infty, x)) \oplus L^2((x, \infty))$ obtained from $H = -\frac{d^2}{dx^2} + V$ in $L^2(\mathbb{R})$ with a Dirichlet boundary condition at x . Let $E_0 \leq \inf \sigma(H)$. Then*

$$V(x) = \lim_{\alpha \downarrow 0} \left[E_0 + \int_{E_0}^{\infty} d\lambda e^{-\alpha\lambda} [1 - 2\xi(\lambda, x)] \right]. \tag{4.9}$$

In particular, if $\int_{E_0}^{\infty} d\lambda |1 - 2\xi(\lambda, x)| < \infty$, then

$$V(x) = E_0 + \int_{E_0}^{\infty} d\lambda [1 - 2\xi(\lambda, x)].$$

We note that the trace formula extends to real-valued potentials $V \in L^1_{\text{loc}}(\mathbb{R})$ as long as H stays bounded from below (it then is in the limit point case at $\pm\infty$). Equation (4.9) then holds at all Lebesgue points of V and hence for a.e. $x \in \mathbb{R}$.

For certain almost periodic potentials, Craig [43] used a regularization similar to the α -regularization in (4.9).

Basically, (4.9) follows from (4.7) and an asymptotic formula,

$$\text{Tr}[e^{-tH} - e^{-tH_x^D}] = \frac{1}{2}[1 - tV(x) + o(t)]. \tag{4.10}$$

We present a few examples next:

EXAMPLE 4.2. Pick a constant $C \in \mathbb{R}$ such that $V(x) = C$ for all $x \in \mathbb{R}$. Then $G(\lambda, x, x) = (C - \lambda)^{-1/2}/2$ and hence one infers that $\text{Arg}(G(\lambda, x, x)) = 0$ (resp., $\pi/2$) if $\lambda < C$ (resp., $\lambda > C$). Thus, by (4.8), $\xi(\lambda, x) = \frac{1}{2}$ on (C, ∞) and $\xi(\lambda, x) = 0$ on $(-\infty, C)$. When $\xi = \frac{1}{2}$ on a subset of $\sigma(H)$, that set does not contribute to the integral in (4.9) and one verifies for $E_0 \leq C$,

$$V(x) = E_0 + \int_{E_0}^C d\lambda = E_0 + (C - E_0) = C, \quad x \in \mathbb{R}.$$

EXAMPLE 4.3. Suppose that $V(x) \rightarrow \infty$ as $|x| \rightarrow \infty$. Then H has eigenvalues $E_0 < E_1 < E_2 < \dots$ and H_x^D has eigenvalues $\{\mu_j(x)\}_{j=1}^{\infty}$ with $E_{j-1} \leq \mu_j(x) \leq E_j$. We have

$$\xi(\lambda, x) = \begin{cases} 1, & E_{j-1} < \lambda < \mu_j(x), \\ 0, & \lambda < E_0 \text{ or } \mu_j(x) < \lambda < E_j. \end{cases}$$

Thus (4.9) becomes:

$$V(x) = E_0 + \lim_{\alpha \downarrow 0} \left[\sum_{j=1}^{\infty} (2e^{-\alpha\mu_j(x)} - e^{-\alpha E_j} - e^{-\alpha E_{j-1}}) / \alpha \right]. \tag{4.11}$$

If we could take α to zero inside the sum, we would get

$$V(x) = E_0 + \sum_{j=1}^{\infty} [E_j + E_{j-1} - 2\mu_j(x)] \quad (\text{formal!}) \tag{4.12}$$

which is just a limit of the periodic formula (4.3) in the limit of vanishing band widths. (4.11) is just a kind of Abelianized summation procedure applied to (4.12).

As a special case of this example, consider $V(x) = x^2 - 1$. Then $E_j = 2j$ and $\{\mu_j(0)\}$ is the set $\{2, 2, 6, 6, 10, 10, 14, 14, \dots\}$ of j odd eigenvalues, each doubled. Thus (4.12) is the formal sum

$$-1 = -2 + 2 - 2 + 2 \dots \quad (\text{formal!})$$

with (4.11)

$$\begin{aligned} -1 &= \lim_{\alpha \downarrow 0} [(e^{-2\alpha} - 1) / \alpha] [1 - e^{-2\alpha} + e^{-4\alpha} \dots] \\ &= \lim_{\alpha \downarrow 0} [(e^{-2\alpha} - 1) / \alpha] [1 / (1 + e^{-2\alpha})] \end{aligned}$$

with a true Abelian summation.

EXAMPLE 4.4. Suppose V is periodic. Let $E_j, \mu_j(x)$ be the band edges and Dirichlet eigenvalues as in (4.2) and (4.3). Then it follows from the fact that the two Floquet solutions are complex conjugates of each other on the spectrum of H , and the Wronskian is antisymmetric in its argument ($W(f, g) = -W(g, f)$), that $g(\lambda, x)$ is purely imaginary on $\sigma(H)$; that is, $\xi(\lambda, x) = \frac{1}{2}$ there, so

$$\xi(\lambda, x) = \begin{cases} \frac{1}{2}, & E_{2n} < \lambda < E_{2n+1}, \\ 1, & E_{2n-1} < \lambda < \mu_n(x), \\ 0, & \mu_n(x) < \lambda < E_{2n}. \end{cases}$$

(This fact has also been used by Deift and Simon [47] and Kotani [140] and also follows from the fact that $g(\lambda, x) = G(\lambda + i0, x, x) = -[m_+(\lambda, x) + m_-(\lambda, x)]^{-1}$ in terms of the Weyl–Titchmarsh m -functions.) Thus,

$$\int_{E_0}^{\infty} d\lambda |1 - 2\xi(\lambda, x)| = \sum_{n=1}^{\infty} |E_{2n} - E_{2n-1}|$$

is finite if (4.1) holds. In that case one can take the limit inside the integral in (4.9) and so recover (4.3).

EXAMPLE 4.5. In [80] it is proved that if V is short-range, that is, $V \in H^{2,1}(\mathbb{R})$, then, $\int_{E_0}^{\infty} d\lambda |1 - 2\xi(\lambda, x)| < \infty$ and one can take the limit in (4.9) inside the integral. This recovers Venakides’ result [247] with an explicit form for ξ in terms of the Green’s function (see Theorem 4.1). Similarly, one can treat short-range perturbations W of periodic background potentials V (modeling scattering off defects or impurities, described by W , in one-dimensional solids) and “cascading” potentials, that is, potentials approaching different spatial asymptotes sufficiently fast (cf. [80] for details).

Next we mention a striking inverse spectral application of the trace formula (4.9) to a celebrated two-spectra inverse spectral theorem due to Borg [24]:

THEOREM 4.6. *Let $V \in L^1_{\text{loc}}(\mathbb{R})$ be real-valued and periodic. Let $H = -\frac{d^2}{dx^2} + V$ be the associated self-adjoint Schrödinger operator in $L^2(\mathbb{R})$ and suppose that $\sigma(H) = [E_0, \infty)$ for some $E_0 \in \mathbb{R}$. Then*

$$V(x) = E_0 \text{ for a.e. } x \in \mathbb{R}.$$

Given the trace formula (4.9) (observing the a.e. extension noted after Theorem 4.1) and using the fact that for all $x \in \mathbb{R}$ and a.e. $\lambda > E_0$, $\xi(\lambda, x) = \frac{1}{2}$ (cf. Example 4.12), the proof of Borg’s Theorem 4.6 is effectively reduced to just a one-line argument (as was observed in [39]). In addition, the new proof permits one to replace periodic by reflectionless potentials and hence applies to algebro-geometric quasi-periodic (KdV) potentials and certain classes of almost periodic potentials.

Now we turn to an analog for Theorem 4.1 for Jacobi operators. This turns out to be a special case of the following result.

THEOREM 4.7. *Let A be a bounded self-adjoint operator in some complex separable Hilbert space \mathcal{H} with $\alpha = \inf \sigma(A)$, $\beta = \sup \sigma(A)$. Let $\varphi \in \mathcal{H}$, $\|\varphi\|_{\mathcal{H}} = 1$ be an arbitrary unit vector in \mathcal{H} and let $\xi(\lambda)$ be the spectral shift function for the pair (A_∞, A) , where A_∞ is defined by*

$$(A_\infty - z)^{-1} = (A - z)^{-1} - (\varphi, (A - z)^{-1}\varphi)^{-1}((A - \bar{z})^{-1}\varphi, \cdot)(A - z)^{-1}\varphi.$$

Then for any $E_- \leq \alpha$ and $E_+ \geq \beta$:

$$\begin{aligned} (\varphi, A\varphi) &= E_- + \int_{E_-}^{E_+} d\lambda [1 - \xi(\lambda)] = E_+ - \int_{E_-}^{E_+} d\lambda \xi(\lambda) \\ &= \frac{1}{2}(E_+ + E_-) + \frac{1}{2} \int_{E_-}^{E_+} d\lambda [1 - 2\xi(\lambda)]. \end{aligned}$$

COROLLARY 4.8. *Let H be a Jacobi matrix on $\ell^2(\mathbb{Z}^\nu)$, that is, for a bounded function V on \mathbb{Z}^ν ,*

$$(Hu)(n) = \sum_{|n-m|=1} u(m) + V(n)u(n), \quad n \in \mathbb{Z}^\nu. \tag{4.13}$$

For $r \in \mathbb{Z}^\nu$, let H_r^D be the operator on $L^2(\mathbb{Z}^\nu \setminus \{r\})$ given by (4.13) with $u(r) = 0$ boundary conditions. Let $\xi(\lambda, r)$ be the spectral shift function for the pair (H_r^D, H) . Then

$$\begin{aligned} V(r) &= E_- + \int_{E_-}^{E_+} d\lambda [1 - \xi(\lambda, r)] = E_+ - \int_{E_-}^{E_+} d\lambda \xi(\lambda, r) \\ &= \frac{1}{2}(E_+ + E_-) + \frac{1}{2} \int_{E_-}^{E_+} d\lambda [1 - 2\xi(\lambda, r)] \end{aligned} \tag{4.14}$$

for any $E_- \leq \inf \sigma(H)$, $E_+ \geq \sup \sigma(H)$.

Only when $\nu = 1$ does this have an interpretation as a formula using Dirichlet problems on the half-line.

Next, we look at some applications to absolutely continuous spectra. In particular, we will point out that $\xi(\lambda, x)$ for a single fixed $x \in \mathbb{R}$ determines the absolutely continuous spectrum of a one-dimensional Schrödinger or Jacobi operator. We begin with a result that holds for higher-dimensional Jacobi operators as well:

LEMMA 4.9. (i) For an arbitrary Jacobi matrix, H , on \mathbb{Z}^ν , $\cup_{j \in \mathbb{Z}^\nu} \{\lambda \in \mathbb{R} \mid 0 < \xi(\lambda, j) < 1\}$ is an essential support for the absolutely continuous spectrum of H .

(ii) For a one-dimensional Schrödinger operator, $H = -\frac{d^2}{dx^2} + V$ bounded from below, $\cup_{x \in \mathbb{Q}} \{\lambda \in \mathbb{R} \mid 0 < \xi(\lambda, x) < 1\}$ is an essential support for the absolutely continuous spectrum of H .

REMARK 4.10. We recall that every absolutely continuous measure, $d\mu$, has the form $\int f(E) dE$. $S = \{E \in \mathbb{R} \mid f(E) \neq 0\}$ is called an essential support. Any Borel set which differs from S by sets of zero Lebesgue measure is also called an essential support. If A is a self-adjoint operator on \mathcal{H} and φ_n , an orthonormal basis for \mathcal{H} , and $d\mu_n$, the spectral measure for the pair, A, φ_n (i.e., $(\varphi_n, e^{isA} \varphi_n) = \int_{\mathbb{R}} e^{isE} d\mu_n(E)$) and if $d\mu_n^{ac}$ is the absolutely continuous component of $d\mu_n$ with S_n its essential support, then $\cup_n S_n$ is the essential support of the absolutely continuous spectrum for A .

In one dimension though, a single x suffices:

THEOREM 4.11. For one-dimensional Schrödinger (resp., Jacobi) operators, $\{\lambda \in \mathbb{R} \mid 0 < \xi(\lambda, x) < 1\}$ is an essential support for the absolutely continuous measure for any fixed $x \in \mathbb{R}$ (resp., \mathbb{Z}).

These results are of particular interest because of their implications for a special kind of semi-continuity of the spectrum.

DEFINITION 4.12. Let $\{V_n\}, V$ be continuous potentials on \mathbb{R} (resp., on \mathbb{Z}). We say that V_n converges to V locally as $n \rightarrow \infty$ if and only if

- (i) $\inf_{(n,x) \in \mathbb{N} \times \mathbb{R}} V_n(x) > -\infty$ (\mathbb{R} case) or $\sup_{(n,j) \in \mathbb{N} \times \mathbb{Z}} |V_n(j)| < \infty$ (\mathbb{Z} case).
- (ii) For each $R < \infty$, $\sup_{|x| \leq R} |V_n(x) - V(x)| \rightarrow 0$ as $n \rightarrow \infty$.

LEMMA 4.13. If $V_n \rightarrow V$ locally as $n \rightarrow \infty$ and H_n, H are the corresponding Schrödinger operators (resp., Jacobi matrices), then $(H_n - z)^{-1} \rightarrow (H - z)^{-1}$ strongly for $\text{Im } z \neq 0$ as $n \rightarrow \infty$.

THEOREM 4.14. If $V_n \rightarrow V$ locally as $n \rightarrow \infty$ and $\xi_n(\lambda, x), \xi(\lambda, x)$ are the corresponding xi functions for fixed x , then $\xi_n(\lambda, x) d\lambda$ converges to $\xi(\lambda, x) d\lambda$ weakly in the sense that for any $f \in L^1(\mathbb{R}; d\lambda)$,

$$\int_{\mathbb{R}} d\lambda f(\lambda) \xi_n(\lambda, x) \rightarrow \int_{\mathbb{R}} d\lambda f(\lambda) \xi(\lambda, x) \text{ as } n \rightarrow \infty.$$

DEFINITION 4.15. For any H , let $|S_{ac}(H)|$ denote the Lebesgue measure of the essential support of the absolutely continuous spectrum of H .

THEOREM 4.16 (For one-dimensional Schrödinger or Jacobi operators). Suppose $V_n \rightarrow V$ locally as $n \rightarrow \infty$ and each V_n is periodic. Then for any interval $(a, b) \subset \mathbb{R}$,

$$|(a, b) \cap S_{ac}| \geq \overline{\lim}_{n \rightarrow \infty} |(a, b) \cap S_{ac}(H_n)|.$$

We note that the periods of V_n need not be fixed; indeed, almost periodic potentials are allowed.

EXAMPLE 4.17. Let α_n be a sequence of rationals and $\alpha = \lim_{n \rightarrow \infty} \alpha_n$. Let H_n be the Jacobi matrix with potential $\lambda \cos(2\pi\alpha_n + \theta)$ for λ, θ fixed. Then [2] have shown for $|\lambda| \leq 2, |S_n| \geq 4 - 2|\lambda|$. It follows from Theorem 4.16 that $|S| \geq 4 - 2|\lambda|$,

providing a new proof (and a strengthening) of a result of Last [148]. At present much more is known about this example and the interested reader may want to consult the survey by Last [149] for additional results.

EXAMPLE 4.18. Let $\{a_m\}_{m \in \mathbb{N}}$ be a sequence with $s = \sum_{m=1}^{\infty} 2^m |a_m| < 2$. Let $V(n) = \sum_{m=1}^{\infty} a_m \cos(2\pi n/2^m)$, a limit periodic potential on \mathbb{Z} . Let h be the corresponding Jacobi matrix, then one can show that $|\sigma_{ac}(h)| \geq 2(2 - s)$.

* * *

Next we very briefly turn to higher-order trace formulas derived in [82] obtained by higher-order expansions in (4.10) as $t \downarrow 0$. For simplicity we now assume that $V \in C^\infty(\mathbb{R})$ is bounded from below. Then (4.10) can be extended to

$$\text{Tr}[e^{tH_x^D} - e^{-tH}] \underset{t \downarrow 0}{\sim} \sum_{j=0}^{\infty} s_j(x) t^j, \quad x \in \mathbb{R}.$$

Similarly, one has,

$$\begin{aligned} \text{Tr}[(H_x^D - z)^{-1} - (H - z)^{-1}] \underset{z \downarrow -\infty}{\sim} \sum_{j=0}^{\infty} r_j(x) z^{-j-1}, \\ r_0(x) = 1/2, \quad r_1(x) = V(x)/2, \quad x \in \mathbb{R} \end{aligned}$$

and one can show that

$$s_j(x) = (-1)^{j+1} (j!)^{-1} r_j(x), \quad j \in \mathbb{N} \cup \{0\}.$$

In particular, $r_j(x)$ and $s_j(x)$ are the celebrated KdV invariants (up to inessential numerical factors). They can be computed recursively (see, e.g., [82]). The higher-order analogs of (4.9) then read

$$\begin{aligned} s_0(x) &= -\frac{1}{2}, \\ s_j(x) &= \frac{(-1)^{j+1}}{j!} \left\{ \frac{E_0^j}{2} + j \lim_{t \downarrow 0} \int_{E_0}^{\infty} d\lambda e^{-t\lambda} \lambda^{j-1} \left[\frac{1}{2} - \xi(\lambda, x) \right] \right\} \quad j \in \mathbb{N}, x \in \mathbb{R}. \end{aligned}$$

and similarly using a resolvent rather than a heat kernel regularization,

$$\begin{aligned} r_1(x) &= \frac{1}{2} V(x) = \frac{E_0}{2} + \lim_{z \rightarrow i\infty} \int_{E_0}^{\infty} d\lambda \frac{z^2}{(\lambda - z)^2} \left[\frac{1}{2} - \xi(\lambda, x) \right], \\ r_j(x) &= \frac{E_0^j}{2} + \lim_{z \rightarrow i\infty} \int_{E_0}^{\infty} d\lambda \frac{z^{j+1}}{(\lambda - z)^{j+1}} j(-\lambda)^{j-1} \left[\frac{1}{2} - \xi(\lambda, x) \right], \quad j \in \mathbb{N}, x \in \mathbb{R}. \end{aligned}$$

In the special periodic case, the corresponding extension of (4.3) then reads

$$2(-1)^{j+1} j! s_j(x) = 2r_j(x) = E_0^j + \sum_{n=1}^{\infty} [E_{2n-1}^j + E_{2n}^j - 2\mu_n(x)^j], \quad j \in \mathbb{N}, x \in \mathbb{R}.$$

The latter formulas were originally found in [65] and [178].

We also note that the use of the Dirichlet boundary boundary condition $u(x) = 0$ and hence the choice of the Dirichlet operator H_x^D in connection with (4.7) can be replaced by any self-adjoint boundary condition of the type $u'(x_{\pm}) + \beta u(x_{\pm}) = 0$, $\beta \in \mathbb{R}$, and the corresponding Schrödinger operator H_x^β in $L^2((-\infty, x)) \oplus L^2((x, \infty))$. This is worked out in detail in [82].

Additional results on trace formulas for Schrödinger operators were presented in [75], [76], [77], [86], [88].

More recent references: Important extensions of the trace formula (4.9), including the case of Schrödinger operators unbounded from below, were discussed by Rybkin [208], [209]. Further discussions of the trace formula (4.14) for Jacobi operators can be found in [94] and Teschl [239], [243, Ch. 6]. An extension of Corollary 4.8 to Schrödinger operators on a countable set was discussed by Shirai [223].

Removal of the resolvent regularization limit in the above trace formula for r_1 (resp., V) under optimal conditions on V has been studied by Rybkin [209], [211] (the latter reference offers necessary and sufficient conditions for absolute summability of the trace formula).

A certain multi-dimensional variant of these trace formulas, inspired by work of Lax [150], was established in [83] (see also [76]).

Matrix-valued extensions of the trace formula for Schrödinger, Dirac-type, and Jacobi matrices, as well as Borg and Hochstadt-type theorems were studied in [16], [37], [38], [39], [40], [77], and [89].

Trace formulas and an ensuing general Borg-type theorem for CMV operators (i.e., in connection with orthogonal polynomials on the unit circle, cf. [229]) appeared in [107].

An application of ξ -function ideas to obtain Weyl-type asymptotics using ζ -function regularizations of determinants of certain operators on complete Riemannian manifolds can be found in Carron [29].

Theorem 4.11 was used in [85] to solve an inverse spectral problem for Jacobi matrices and most recently in [106] in connection with proving purely absolutely continuous spectrum of a class of reflectionless Schrödinger operators with homogeneous spectrum. It has also recently been discussed in [53, Sect. 1.5].

5. Various Uniqueness Theorems in Inverse Spectral Theory

In this section we summarize some of the principal results of the following papers:

[50] R. del Rio, F. Gesztesy, and B. Simon, *Inverse spectral analysis with partial information on the potential, III. Updating boundary conditions*, Intl. Math. Research Notices **1997**, No. 15, 751–758.

[51] R. del Rio, F. Gesztesy, and B. Simon, *Corrections and Addendum to “Inverse spectral analysis with partial information on the potential, III. Updating boundary conditions”*, Intl. Math. Research Notices **1999**, No. 11, 623–625.

[92] F. Gesztesy and B. Simon, *Uniqueness theorems in inverse spectral theory for one-dimensional Schrödinger operators*, Trans. Amer. Math. Soc. **348**, 349–373 (1996).

[93] F. Gesztesy and B. Simon, *Inverse spectral analysis with partial information on the potential, I. The case of an a.c. component in the spectrum*, Helv. Phys. Acta **70**, 66–71, 1997.

[94] F. Gesztesy and B. Simon, *m -functions and inverse spectral analysis for finite and semi-infinite Jacobi matrices*, J. Analyse Math. **73**, 267–297 (1997).

[95] F. Gesztesy and B. Simon, *On the determination of a potential from three spectra*, in *Differential Operators and Spectral Theory*, V. Buslaev, M. Solomyak, and D. Yafaev (eds.), Amer. Math. Soc. Transl. Ser. 2, **189**, 85–92 (1999).

[96] F. Gesztesy and B. Simon, *Inverse spectral analysis with partial information on the potential, II. The case of discrete spectrum*, Trans. Amer. Math. Soc. **352**, 2765–2787 (2000).

One can argue that inverse spectral theory, especially, the case of uniqueness theorems in inverse spectral theory, started with the paper by Ambarzumian [7] in 1929 and was turned into a full-fledged discipline by the seminal 1946 paper by Borg [24]. Ambarzumian proved the special uniqueness theorem that if the eigenvalues of a Schrödinger operator $-\frac{d^2}{dx^2} + V$ in $L^2([0, \pi])$ with Neumann boundary conditions at the endpoints $x = 0$ and $x = \pi$ coincide with the sequence of numbers n^2 , $n = 0, 1, 2, \dots$, then $V = 0$ a.e. on $[0, \pi]$. This result is very special. Indeed, Borg showed that for more general boundary conditions, one set of eigenvalues, in general (i.e., in the absence of symmetries of V), is insufficient to determine V uniquely. Moreover, he described in great detail when two spectra guarantee unique determination of the potential V . In this section we will discuss Borg's celebrated two-spectra uniqueness result and many of its extension due to Gasymov, Hald, Hochstadt, Levitan, Lieberman, Marchenko, and Simon and collaborators.

We start with paper [92]. It contains a variety of new uniqueness theorems for potentials V in one-dimensional Schrödinger operators $-\frac{d^2}{dx^2} + V$ on \mathbb{R} and on the half-line $\mathbb{R}_+ = [0, \infty)$ in terms of appropriate spectral shift functions introduced in a series of papers describing new trace formulas for V on \mathbb{R} [80], [81], [82], [91] and on \mathbb{R}_+ [76]. In particular, it contains a generalization of a well-known uniqueness theorem of Borg and Marchenko for Schrödinger operators on the half-line with purely discrete spectra to arbitrary spectral types and a new uniqueness result for Schrödinger operators with confining potentials on the entire real line.

Turning to the half-line case first, we recall one of the principal uniqueness results proved in [92], which extends a well-known theorem of Borg [25] and Marchenko [171] in the special case of purely discrete spectra to arbitrary spectral types. We suppose

$$V \in L^1([0, R]) \text{ for all } R > 0, \quad V \text{ real-valued,} \quad (5.1)$$

and introduce the differential expression

$$\tau_+ = -\frac{d^2}{dx^2} + V(x), \quad x \geq 0,$$

for simplicity assuming that τ is in the limit point case at ∞ . (We refer to [92] for a general treatment that includes the limit circle case.) Associated with τ_+ one introduces the following self-adjoint operator $H_{+, \alpha}$ in $L^2(\mathbb{R}_+)$.

$$\begin{aligned} H_{+, \alpha} f &= \tau_+ f, \quad \alpha \in [0, \pi), \\ f &\in \text{dom}(H_{+, \alpha}) = \{g \in L^2(\mathbb{R}_+) \mid g, g' \in AC([0, R]) \text{ for all } R > 0; \\ &\quad \sin(\alpha)g'(0_+) + \cos(\alpha)g(0_+) = 0; \tau_+ g \in L^2(\mathbb{R}_+)\}. \end{aligned} \quad (5.2)$$

Then $H_{+, \alpha}$ has uniform spectral multiplicity one.

Next we introduce the fundamental system $\phi_\alpha(z, x), \theta_\alpha(z, x), z \in \mathbb{C}$, of solutions of $\tau_+\psi(z, x) = z\psi(z, x), x \geq 0$, satisfying

$$\phi_\alpha(z, 0) = -\theta'_\alpha(z, 0) = -\sin(\alpha), \quad \phi'_\alpha(z, 0) = \theta_\alpha(z, 0) = \cos(\alpha)$$

such that $W(\theta_\alpha(z), \phi_\alpha(z)) = 1$. Furthermore, let $\psi_{+,\alpha}(z, x), z \in \mathbb{C} \setminus \mathbb{R}$ be the unique solution of $\tau\psi(z) = \psi(z)$ which satisfies

$$\psi_{+,\alpha}(z, \cdot) \in L^2(\mathbb{R}_+), \quad \sin(\alpha)\psi'_{+,\alpha}(z, 0_+) + \cos(\alpha)\psi_{+,\alpha}(z, 0_+) = 1.$$

$\psi_{+,\alpha}$ is of the form

$$\psi_{+,\alpha}(z, x) = \theta_\alpha(z, x) + m_{+,\alpha}(z)\phi_\alpha(z, x)$$

with $m_{+,\alpha}(z)$ the half-line Weyl–Titchmarsh m -function. Being a Herglotz function (i.e., an analytic function in the open upper-half plane that maps the latter to itself), $m_{+,\alpha}(z)$ has the following representation in terms of a positive measure $d\rho_{+,\alpha}$,

$$m_{+,\alpha} = \begin{cases} a_{+,\alpha} + \int_{\mathbb{R}} \left[\frac{1}{\lambda-z} - \frac{\lambda}{1+\lambda^2} \right] d\rho_{+,\alpha}(\lambda), & \alpha \in [0, \pi), \\ \cot(\alpha) + \int_{\mathbb{R}} (\lambda-z)^{-1} d\rho_{+,\alpha}(\lambda), & \alpha \in (0, \pi). \end{cases}$$

The basic uniqueness criterion for Schrödinger operators on the half-line $[0, \infty)$, due to Marchenko [171], that we shall rely on repeatedly in the following, can be stated as follows.

THEOREM 5.1. *Suppose $\alpha_1, \alpha_2 \in [0, \pi), \alpha_1 \neq \alpha_2$ and define $H_{+,j,\alpha_j}, m_{+,j,\alpha_j}, \rho_{+,j,\alpha_j}$ associated with the differential expressions $\tau_j = -\frac{d^2}{dx^2} + V_j(x), x \geq 0$, where $V_j, j = 1, 2$ satisfy assumption (5.1). Then the following three assertions are equivalent:*

- (i) $m_{+,1,\alpha_1}(z) = m_{+,2,\alpha_2}(z), z \in \mathbb{C}_+$.
- (ii) $\rho_{+,1,\alpha_1}((-\infty, \lambda]) = \rho_{+,2,\alpha_2}((-\infty, \lambda]), \lambda \in \mathbb{R}$.
- (iii) $\alpha_1 = \alpha_2$ and $V_1(x) = V_2(x)$ for a.e. $x \geq 0$.

Next we relate Green’s functions for different boundary conditions at $x = 0$.

LEMMA 5.2. *Let $\alpha_j \in [0, \pi), j = 1, 2, x, x' \in \mathbb{R}_+$, and $z \in \mathbb{C} \setminus \{\sigma(H_{+,\alpha_1}) \cup \sigma(H_{+,\alpha_2})\}$. Then,*

$$\begin{aligned} G_{+,\alpha_2}(z, x, x') - G_{+,\alpha_1}(z, x, x') &= -\frac{\psi_{+,\alpha_1}(z, x)\psi_{+,\alpha_1}(z, x')}{\cot(\alpha_2 - \alpha_1) + m_{+,\alpha_1}(z)}, \\ \frac{G_{+,\alpha_2}(z, 0, 0)}{G_{+,\alpha_1}(z, 0, 0)} &= \frac{1}{(\beta_1 - \beta_2) \sin^2(\alpha_1)[\cot(\alpha_2 - \alpha_1) + m_{+,\alpha_1}(z)]} \\ &= (\beta_1 - \beta_2) \sin^2(\alpha_2)[\cot(\alpha_2 - \alpha_1) - m_{+,\alpha_2}(z)], \quad \beta_j = \cot(\alpha_j), j = 1, 2 \\ \text{Tr}[(H_{+,\alpha_2} - z)^{-1} - (H_{+,\alpha_1} - z)^{-1}] &= -\frac{d}{dz} \ln[\cot(\alpha_2 - \alpha_1) + m_{+,\alpha_1}(z)] \\ &= \frac{d}{dz} \ln[\cot(\alpha_2 - \alpha_1) - m_{+,\alpha_2}(z)]. \end{aligned}$$

Since $m_{+,\alpha}(z)$ is a Herglotz function, we may now introduce spectral shift function [27] $\xi_{\alpha_1, \alpha_2}(\lambda)$ for the pair $(H_{+,\alpha_2}, H_{+,\alpha_1})$ via the exponential Herglotz representation of $m_{+,\alpha}(z)$ (cf. [12])

$$\cot(\alpha_2 - \alpha_1) + m_{+,\alpha_1}(z) = \exp \left\{ \text{Re}[\ln(\cot(\alpha_2 - \alpha_1) + m_{+,\alpha_1}(i))] \right\}$$

$$+ \int_{\mathbb{R}} \left[\frac{1}{\lambda - z} - \frac{\lambda}{1 + \lambda^2} \right] \xi_{\alpha_1, \alpha_2}(\lambda) d\lambda \Bigg\}, \quad 0 \leq \alpha_1 < \alpha_2 < \pi, z \in \mathbb{C} \setminus \mathbb{R}.$$

This is extended to all $\alpha_1, \alpha_2 \in [0, \pi)$ by

$$\xi_{\alpha, \alpha}(\lambda) = 0, \quad \xi_{\alpha_2, \alpha_1}(\lambda) = -\xi_{\alpha_1, \alpha_2}(\lambda) \text{ for a.e. } \lambda \in \mathbb{R}.$$

Next we summarize a few properties of $\xi_{\alpha_1, \alpha_2}(\lambda)$.

LEMMA 5.3. (i) Suppose $0 \leq \alpha_1 < \alpha_2 < \pi$. Then for a.e. $\lambda \in \mathbb{R}$,

$$\xi_{\alpha_1, \alpha_2}(\lambda) = \begin{cases} \lim_{\epsilon \downarrow 0} \pi^{-1} \text{Im} \{ \ln [\cot(\alpha_2 - \alpha_1) + m_{+, \alpha_1}(\lambda + i\epsilon)] \}, \\ - \lim_{\epsilon \downarrow 0} \pi^{-1} \text{Im} \{ \ln [\cot(\alpha_2 - \alpha_1) - m_{+, \alpha_2}(\lambda + i\epsilon)] \}, \\ \lim_{\epsilon \downarrow 0} \pi^{-1} \text{Im} \left\{ \ln \left[\frac{1}{\sin(\alpha_1)} \frac{G_{+, \alpha_1}(\lambda + i\epsilon, 0, 0)}{G_{+, \alpha_2}(\lambda + i\epsilon, 0, 0)} \right] \right\}. \end{cases}$$

(For $\alpha_1 = 0$, $G_{+, \alpha_1}(\lambda + i\epsilon, 0, 0)/\sin(\alpha_1)$ has to be replaced by -1 in the last expression.) Moreover,

$$0 \leq \xi_{\alpha_1, \alpha_2}(\lambda) \leq 1 \text{ a.e.}$$

(ii) Let $\alpha_j \in [0, \pi)$, $1 \leq j \leq 3$. Then the ‘‘chain rule’’

$$\xi_{\alpha_1, \alpha_3}(\lambda) = \xi_{\alpha_1, \alpha_2}(\lambda) + \xi_{\alpha_2, \alpha_3}(\lambda)$$

holds for a.e. $\lambda \in \mathbb{R}$.

(iii) For all $\alpha_1, \alpha_2 \in [0, \pi)$,

$$\xi_{\alpha_1, \alpha_2} \in L^1(\mathbb{R}; (1 + \lambda^2)^{-1} d\lambda).$$

(iv) Assume $\alpha_1, \alpha_2 \in [0, \pi)$, $\alpha_1 \neq \alpha_2$. Then,

$$\xi_{\alpha_1, \alpha_2} \in L^1(\mathbb{R}; (1 + |\lambda|)^{-1} d\lambda) \text{ if and only if } \alpha_1, \alpha_2 \in (0, \pi).$$

(v) For all $\alpha_1, \alpha_2 \in [0, \pi)$,

$$\text{Tr}[(H_{+, \alpha_2} - z)^{-1} - (H_{+, \alpha_1} - z)^{-1}] = - \int_{\mathbb{R}} \frac{d\lambda \xi_{\alpha_1, \alpha_2}(\lambda)}{(\lambda - z)^{-2}}.$$

We note that $\xi_{\alpha_1, \alpha_2}(\lambda)$ (for $\alpha_1, \alpha_2 \in (0, \pi)$) has been introduced by Javrjan [133], [134]. In particular, he proved Lemma 5.2 (iii) and Lemma 5.3 (v) in the non-Dirichlet cases where $0 < \alpha_1, \alpha_2 < \pi$.

Given these preliminaries, we are now able to state the main uniqueness result for half-line Schrödinger operators of [92].

THEOREM 5.4. Suppose V_j satisfy assumption (5.1) and define $H_{+, j, \alpha_j, \ell}$, $j, \ell = 1, 2$, associated with the differential expressions $\tau_j = -\frac{d^2}{dx^2} + V_j(x)$, $x \geq 0$, $j = 1, 2$, where $\alpha_{j, \ell} \in [0, \pi)$, $\ell = 1, 2$, and we suppose $0 \leq \alpha_{1,1} < \alpha_{1,2} < \pi$, $0 \leq \alpha_{2,1} < \alpha_{2,2} < \pi$. In addition, let $\xi_{j, \alpha_{j,1}, \alpha_{j,2}}$, $j = 1, 2$ be the spectral shift function for the pair $(H_{+, j, \alpha_{j,1}}, H_{+, j, \alpha_{j,2}})$. Then the following two assertions are equivalent:

(i) $\xi_{1, \alpha_{1,1}, \alpha_{1,2}}(\lambda) = \xi_{2, \alpha_{2,1}, \alpha_{2,2}}(\lambda)$ for a.e. $\lambda \in \mathbb{R}$.

(ii) $\alpha_{1,1} = \alpha_{2,1}$, $\alpha_{1,2} = \alpha_{2,2}$, and $V_1(x) = V_2(x)$ for a.e. $x \geq 0$.

As a corollary, one obtains a well-known uniqueness result originally due to Borg [25, Theorem 1] and Marchenko [171, Theorem 2.3.2] (see also [161]).

COROLLARY 5.5. Define $\tau_j, H_{+,j,\alpha}, \alpha \in [0, \pi)$ as in Theorem 5.4. Assume in addition that $H_{+,1,\alpha_1}$ and $H_{+,2,\alpha_2}$ have purely discrete spectra for some (and hence for all) $\alpha_j \in [0, \pi)$, that is,

$$\sigma_{\text{ess}}(H_{+,j,\alpha_j}) = \emptyset \text{ for some } \alpha_j \in [0, \pi), j = 1, 2.$$

Then the following two assertions are equivalent:

- (i) $\sigma(H_{+,1,\alpha_{1,1}}) = \sigma(H_{+,2,\alpha_{2,1}}), \sigma(H_{+,1,\alpha_{1,2}}) = \sigma(H_{+,2,\alpha_{2,2}}), \alpha_{j,\ell} \in [0, \pi), j, \ell = 1, 2, \sin(\alpha_{1,1} - \alpha_{1,2}) \neq 0.$
- (ii) $\alpha_{1,1} = \alpha_{2,1}, \alpha_{1,2} = \alpha_{2,2},$ and $V_1(x) = V_2(x)$ for a.e. $x \geq 0.$

Roughly speaking, Corollary 5.5 implies that two sets of purely discrete spectra $\sigma(H_{+,\alpha_1}), \sigma(H_{+,\alpha_2})$ associated with distinct boundary conditions at $x = 0$ (but a fixed boundary condition (if any) at $+\infty$), that is, $\sin(\alpha_2 - \alpha_1) \neq 0$, uniquely determine V a.e. The first main result in [92], Theorem 5.4, removes all a priori spectral hypotheses and shows that the spectral shift function $\xi_{\alpha_1,\alpha_2}(\lambda)$ for the pair $(H_{+,\alpha_2}, H_{+,\alpha_1})$ with distinct boundary conditions at $x = 0, \sin(\alpha_2 - \alpha_1) \neq 0$, uniquely determines V a.e. This illustrates that Theorem 5.4 is the natural generalization of Borg’s and Marchenko’s theorems from the discrete spectrum case to arbitrary spectral types.

Now we turn to uniqueness results for Schrödinger operators on the whole real line. We shall rely on the notation $\tau, \phi_\alpha, \theta_\alpha, \psi_{\pm,\alpha}, m_{\pm,\alpha}, d\rho_{\pm,\alpha}$, which are defined in complete analogy to the half-line case (with $x \in \mathbb{R}$), and we shall assume

$$V \in L^1_{\text{loc}}(\mathbb{R}), \quad V \text{ real-valued.} \tag{5.3}$$

Following [82], we introduce, in addition, the following family of self-adjoint operators H_y^β in $L^2(\mathbb{R})$,

$$\begin{aligned} H_y^\beta f &= \tau f, \quad \beta \in \mathbb{R} \cup \{\infty\}, \quad y \in \mathbb{R}, \\ \text{dom}(H_y^\beta) &= \{g \in L^2(\mathbb{R} \mid g, g' \in AC([y, \pm R]) \text{ for all } R > 0; g'(y_\pm) + \beta g(y_\pm) = 0; \\ &\quad \lim_{R \rightarrow \pm\infty} W(f_\pm(z_\pm), g)(R) = 0; \tau g \in L^2(\mathbb{R})\}. \end{aligned}$$

Thus $H_y^D := H_y^\infty$ (resp., $H_y^N := H_y^0$) corresponds to the Schrödinger operator with an additional Dirichlet (resp., Neumann) boundary condition at y . In obvious notation, H_y^β decomposes into the direct sum of half-line operators

$$H_y^\beta = H_{-,y}^\beta \oplus H_{+,y}^\beta$$

with respect to $L^2(\mathbb{R}) = L^2((-\infty, y]) \oplus L^2([y, \infty))$. In particular, $H_{+,y}^\beta$ equals $H_{+,\alpha}$ for $\beta = \cot(\alpha)$ (and $y = 0$) in our notation (5.2) and, as done in Sections 2 and 4, the reference point y will be added as a subscript to obtain $\theta_{\alpha,y}(z, x), \phi_{\alpha,y}(z, x), \psi_{\pm,\alpha,y}(z, x), m_{\pm,\alpha,y}(z), M_{\alpha,y}(z)$, etc.

Next, we recall a few results from [82]. With $G(z, x, x')$ and $G_y^\beta(z, x, x')$ the Green’s functions of H and H_y^β , one obtains (for $z \in \mathbb{C} \setminus \{\sigma(H_y^\infty) \cup \sigma(H)\}$)

$$\begin{aligned} G_y^\beta(z, x, x') &= G(z, x, x') - \frac{(\beta + \partial_2)G(z, x, y)(\beta + \partial_1)G(z, y, x')}{(\beta + \partial_1)(\beta + \partial_2)G(z, y, y)}, \\ &\quad \beta \in \mathbb{R}, z \in \mathbb{C} \setminus \{\sigma(H_y^\beta) \cup \sigma(H)\}, \\ G_y^\infty(z, x, x') &= G(z, x, x') - G(z, y, y)^{-1}G(z, x, y)G(z, y, x'). \end{aligned}$$

Here we abbreviated

$$\begin{aligned} \partial_1 G(z, y, x') &= \partial_x G(z, x, x')|_{x=y}, & \partial_2(G, z, x, y) &= \partial_{x'} G(z, x, x')|_{x'=y}, \\ \partial_1 \partial_2 G(z, y, y) &= \partial_x \partial_{x'} G(z, x, x')|_{x=y=x'}, & \text{etc.} \\ \partial_1 G(z, y, x) &= \partial_2 G(z, x, y), & x &\neq y. \end{aligned}$$

As a consequence,

$$\begin{aligned} \text{Tr}[(H_y^\beta - z)^{-1} - (H - z)^{-1}] &= -\frac{d}{dz} \ln[(\beta + \partial_1)(\beta + \partial_2)G(z, y, y)], & \beta \in \mathbb{R}, \\ \text{Tr}[(H_y^\infty - z)^{-1} - (H - z)^{-1}] &= -\frac{d}{dz} \ln[G(z, y, y)]. \end{aligned}$$

In analogy to the Herglotz property of $G(z, y, y)$, $(\beta + \partial_1)(\beta + \partial_2)G(z, y, y)$ is also Herglotz for each $y \in \mathbb{R}$. Hence, both admit exponential representations of the form

$$\begin{aligned} G(z, y, y) &= \exp\left\{c_\infty + \int_{\mathbb{R}} d\lambda \left[\frac{1}{\lambda - z} - \frac{\lambda}{1 + \lambda^2}\right] \xi^\infty(\lambda, y)\right\}, \\ c_\infty \in \mathbb{R}, & \quad 0 \leq \xi^\infty(\lambda, y) \leq 1 \text{ a.e.}, \\ \xi^\infty(\lambda, y) &= \lim_{\epsilon \downarrow 0} \pi^{-1} \text{Im}\{\ln[G(\lambda + i\epsilon, y, y)]\} \text{ for a.e. } \lambda \in \mathbb{R}, \\ (\beta + \partial_1)(\beta + \partial_2)G(z, y, y) &= \exp\left\{c_\beta + \int_{\mathbb{R}} d\lambda \left[\frac{1}{\lambda - z} - \frac{\lambda}{1 + \lambda^2}\right] [\xi^\beta(\lambda, y) + 1]\right\}, \\ c_\beta \in \mathbb{R}, & \quad -1 \leq \xi^\beta(\lambda, y) \leq 0 \text{ a.e.}, \quad \beta \in \mathbb{R}, \\ \xi^\beta(\lambda, y) &= \lim_{\epsilon \downarrow 0} \pi^{-1} \text{Im}\{\ln[(\beta + \partial_1)(\beta + \partial_2)G(\lambda + i\epsilon, y, y)]\} - 1, \quad \beta \in \mathbb{R} \end{aligned}$$

for each $y \in \mathbb{R}$. Moreover,

$$\text{Tr}[(H_y^\beta - z)^{-1} - (H - z)^{-1}] = - \int_{\mathbb{R}} d\lambda (\lambda - z)^{-2} \xi^\beta(\lambda, y), \quad \beta \in \mathbb{R} \cup \{\infty\}.$$

Applying the basic uniqueness criterion for Schrödinger operators to both half-lines $(-\infty, y]$ and $[y, \infty)$ then yields the following principal characterization result for Schrödinger operators on \mathbb{R} first proved in [92].

THEOREM 5.6. *Let $\beta_1, \beta_2 \in \mathbb{R} \cup \{\infty\}$, $\beta_1 \neq \beta_2$, and $x_0 \in \mathbb{R}$. Then the following assertions hold:*

- (i) $\xi^{\beta_1}(\lambda, x_0)$ and $\xi^{\beta_2}(\lambda, x_0)$ for a.e. $\lambda \in \mathbb{R}$ uniquely determine $V(x)$ for a.e. $x \in \mathbb{R}$ if the pair (β_1, β_2) differs from $(0, \infty)$, $(\infty, 0)$.
- (ii) If $(\beta_1, \beta_2) = (0, \infty)$ or $(\infty, 0)$, assume in addition that τ is in the limit point case at $+\infty$ and $-\infty$. Then $\xi^\infty(\lambda, x_0)$ and $\xi^0(\lambda, x_0)$ for a.e. $\lambda \in \mathbb{R}$ uniquely determine V a.e. up to reflection symmetry with respect to x_0 ; that is, both $V(x)$, $V(2x_0 - x)$ for a.e. $x \in \mathbb{R}$ correspond to $\xi^\infty(\lambda, x_0)$ and $\xi^0(\lambda, x_0)$ for a.e. $\lambda \in \mathbb{R}$.

COROLLARY 5.7. *Suppose τ is in the limit point case at $+\infty$ and $-\infty$ and let $\beta \in \mathbb{R} \cup \{\infty\}$ and $x_0 \in \mathbb{R}$. Then $\xi^\beta(\lambda, x_0)$ for a.e. $\lambda \in \mathbb{R}$ uniquely determines $V(x)$ for a.e. $x \in \mathbb{R}$ if and only if V is reflection symmetric with respect to x_0 , that is, $V(2x_0 - x) = V(x)$ a.e.*

In view of Corollary 5.5, it seems appropriate to formulate Theorem 5.6 in the special case of purely discrete spectra.

COROLLARY 5.8. *Suppose H (and hence H_y^β for all $y \in \mathbb{R}$, $\beta \in \mathbb{R} \cup \{\infty\}$) has purely discrete spectrum, that is, $\sigma_{\text{ess}}(H) = \emptyset$ and let $\beta_1, \beta_2 \in \mathbb{R} \cup \{\infty\}$, $\beta_1 \neq \beta_2$, and $x_0 \in \mathbb{R}$.*

(i) *$\sigma(H)$, $\sigma(H_{x_0}^{\beta_j})$, $j = 1, 2$ uniquely determine V a.e. if the pair (β_1, β_2) differs from $(0, \infty)$ and $(\infty, 0)$.*

(ii) *If $(\beta_1, \beta_2) = (0, \infty)$ or $(\infty, 0)$, assume in addition that τ is in the limit point case at $+\infty$ and $-\infty$. Then $\sigma(H)$, $\sigma(H_{x_0}^\infty)$, and $\sigma(H_{x_0}^0)$ uniquely determine V a.e. up to reflection symmetry with respect to x_0 , that is, both $V(x)$ and $\widehat{V}(x) = V(2x_0 - x)$ for a.e. $x \in \mathbb{R}$ correspond to $\sigma(H) = \sigma(\widehat{H})$, $\sigma(H_{x_0}^\infty) = \sigma(\widehat{H}_{x_0}^\infty)$, and $\sigma(H_{x_0}^0) = \sigma(\widehat{H}_{x_0}^0)$. Here, in obvious notation, \widehat{H} , $\widehat{H}_{x_0}^\infty$, $\widehat{H}_{x_0}^0$ correspond to $\widehat{\tau} = -\frac{d^2}{dx^2} + \widehat{V}(x)$, $x \in \mathbb{R}$.*

(iii) *Suppose τ is in the limit point case at $+\infty$ and $-\infty$ and let $\beta \in \mathbb{R} \cup \{\infty\}$. Then $\sigma(H)$ and $\sigma(H_{x_0}^\beta)$ uniquely determine V a.e. if and only if V is reflection symmetric with respect to x_0 .*

(iv) *Suppose that V is reflection symmetric with respect to x_0 and that τ is nonoscillatory at $+\infty$ and $-\infty$. Then V is uniquely determined a.e. by $\sigma(H)$ in the sense that V is the only potential symmetric with respect to x_0 with spectrum $\sigma(H)$.*

Of course, Corollary 5.8 (iii) is implied by the result of Borg [5] and Marchenko [32] (see Corollary 5.5 with $\alpha_1 = 0$, $\alpha_2 = \pi/2$).

Thus far, we dealt exclusively with ξ -functions and spectra in connection with uniqueness theorems. A variety of further uniqueness results can be obtained by invoking alternative information such as the left/right distribution of $\lambda_n^\beta(x_0)$ (i.e., whether $\lambda_n^\beta(x_0)$ is an eigenvalue of H_{-,x_0}^β in $L^2((-\infty, x_0])$ or of H_{+,x_0}^β in $L^2([x_0, \infty))$) and/or associated norming constants. For details we refer to the discussion in [92].

More recent references: Uniqueness theorems related to Theorem 5.4 in the short-range case with spectral shift data replaced by scattering data were studied by Aktosun and Weder [5], [6]. Analogs of Corollaries 5.5 and 5.8 (i) for Jacobi operators were derived by Teschl [239].

* * *

Next we focus on [96], which discussed results where the discrete spectrum (or partial information on the discrete spectrum) and partial information on the potential V of a one-dimensional Schrödinger operator $H = -\frac{d^2}{dx^2} + q$ determines the potential completely. Included are theorems for finite intervals and for the whole line. In particular, a new type of inverse spectral problem involving fractions of the eigenvalues of H on a finite interval and knowledge of V over a corresponding fraction of the interval was posed and solved in [96]. The methods employed in [96] rest on Weyl–Titchmarsh m -function techniques starting with the basic Borg–Marchenko uniqueness result (cf. Theorems 5.15 and 5.16) and densities of zeros of a class of entire functions since the m -functions are meromorphic functions in this context.

In 1978, Hochstadt and Lieberman [123] proved the following remarkable theorem:

THEOREM 5.9. *Let $h_0 \in \mathbb{R}$, $h_1 \in \mathbb{R} \cup \{\infty\}$ and assume $V_1, V_2 \in L^1((0, 1))$ to be real-valued. Consider the Schrödinger operators H_1, H_2 in $L^2([0, 1])$ given by*

$$H_j = -\frac{d^2}{dx^2} + V_j, \quad j = 1, 2,$$

with the boundary conditions

$$\begin{aligned} u'(0) + h_0 u(0) &= 0, \\ u'(1) + h_1 u(1) &= 0. \end{aligned} \tag{5.4}$$

Let $\sigma(H_j) = \{\lambda_{j,n}\}$ be the (necessarily simple) spectra of $H_j, j = 1, 2$. Suppose that $V_1 = V_2$ a.e. on $[0, \frac{1}{2}]$ and that $\lambda_{1,n} = \lambda_{2,n}$ for all n . Then $V_1 = V_2$ a.e. on $[0, 1]$.

Here, in obvious notation, $h_1 = \infty$ in (5.4) singles out the Dirichlet boundary condition $u(1) = 0$.

For each $\varepsilon > 0$, there are simple examples where $V_1 = V_2$ on $[0, (\frac{1}{2}) - \varepsilon]$ and $\sigma(H_1) = \sigma(H_2)$ but $V_1 \neq V_2$. (Choose $h_0 = -h_1, V_1(x) = 0$ for $x \in (0, (\frac{1}{2}) - \varepsilon] \cup [\frac{1}{2}, 1]$ and nonzero on $((\frac{1}{2}) - \varepsilon, \frac{1}{2})$, and $V_2(x) = V_1(1 - x)$. See also Theorem I' in the appendix of [234].)

Later refinements of Theorem 5.9 in [115], [234] (see also the summary in [233]) showed that the boundary condition for H_1 and H_2 at $x = 1$ need not be assumed a priori to be the same, and that if V is continuous, then one only needs $\lambda_{1,n} = \lambda_{2,m(n)}$ for all values of n but one. The same boundary condition for H_1 and H_2 at $x = 0$, however, is crucial for Theorem 1.1 to hold (see [115], [49]).

Moreover, analogs of Theorem 5.9 for certain Schrödinger operators are considered in [136] (see also [192, Ch. 4]). Reconstruction techniques for V in this context are discussed in [206].

Our purpose in [96] was to provide a new approach to Theorem 5.9 that we felt was more transparent and, moreover, capable of vast generalizations. To state our generalizations, we will introduce a shorthand notation to paraphrase Theorem 5.9 by saying “ V on $[0, \frac{1}{2}]$ and the eigenvalues of H uniquely determine V .” This is just a shorthand notation for saying $V_1 = V_2$ a.e. if the obvious conditions hold.

Unless explicitly stated otherwise, all potentials V, V_1 , and V_2 will be real-valued and in $L^1([0, 1])$ for the remainder of this paper. Moreover, to avoid too many case distinctions we shall assume $h_0, h_1 \in \mathbb{R}$ in (5.4). In particular, for $h_0, h_1 \in \mathbb{R}$ we index the corresponding eigenvalues λ_n of H by $n \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$. The case of Dirichlet boundary conditions, where $h_0 = \infty$ and/or $h_1 = \infty$, has been dealt with in detail in [96, Appendix A].

Here is a summary of the generalizations proved for Schrödinger operators on $[0, 1]$ in [96]:

THEOREM 5.10. *Let $H = -\frac{d^2}{dx^2} + V$ in $L^2([0, 1])$ with boundary conditions (5.4) and $h_0, h_1 \in \mathbb{R}$. Suppose V is $C^{2k}(((\frac{1}{2}) - \varepsilon, (\frac{1}{2}) + \varepsilon))$ for some $k = 0, 1, \dots$ and for some $\varepsilon > 0$. Then V on $[0, \frac{1}{2}]$, h_0 , and all the eigenvalues of H except for $(k + 1)$ uniquely determine h_1 and V on all of $[0, 1]$.*

REMARK 5.11. (i). The case $k = 0$ in Theorem 5.10 is due to Hald [115].

(ii) In the non-shorthand form of this theorem (cf. the paragraphs preceding Theorem 5.10), we mean that both V_1 and V_2 are C^{2k} near $x = \frac{1}{2}$.

(iii) One need not know which eigenvalues are missing. Since the eigenvalues asymptotically satisfy

$$\lambda_n = (\pi n)^2 + 2(h_1 - h_0) + \int_0^1 dx V(x) + o(1) \text{ as } n \rightarrow \infty,$$

given a set of candidates for the spectrum, one can tell how many are missing.

(iv) For the sake of completeness we mention the precise definition of H in $L^2([0, 1])$ for real-valued $V \in L^1([0, 1])$ and boundary condition parameters $h_0, h_1 \in \mathbb{R} \cup \{\infty\}$ in (1.1):

$$H = -\frac{d^2}{dx^2} + V, \quad \text{dom}(H) = \{g \in L^2([0, 1]) \mid g, g' \in AC([0, 1]); (-g'' + Vg) \in L^2([0, 1]); \quad (5.5)$$

$$g'(0) + h_0g(0) = 0, g'(1) + h_1g(1) = 0\},$$

where $AC([0, 1])$ denotes the set of absolutely continuous functions on $[0, 1]$ and $h_{x_0} = \infty$ represents the Dirichlet boundary condition $g(x_0) = 0$ for $x_0 \in \{0, 1\}$ in (5.5).

By means of explicit examples, it has been shown in Section 3 of [96], that Theorem 5.10 is optimal in the sense that if V is only assumed to be C^{2k-1} near $x = \frac{1}{2}$ for some $k \geq 1$, then it is not uniquely determined by $V \upharpoonright [0, \frac{1}{2}]$ and all the eigenvalues but $(k + 1)$.

Theorem 5.10 works because the condition that V is C^{2k} near $x = \frac{1}{2}$ gives us partial information about V on $[\frac{1}{2}, 1]$; indeed, we know the values of

$$V(\frac{1}{2}), V'(\frac{1}{2}), \dots, V^{(2k)}(\frac{1}{2})$$

computed on $[\frac{1}{2}, 1]$ since one can compute them on $[0, \frac{1}{2}]$. This suggests that knowing V on more than $[0, \frac{1}{2}]$ should let one dispense with a finite density of eigenvalues. That this is indeed the case is the content of the following theorem. (We denote by $\#\{\dots\}$ the cardinality of the set $\{\dots\}$.)

THEOREM 5.12. *Let $H = -\frac{d^2}{dx^2} + V$ in $L^2([0, 1])$ with boundary conditions (5.4) and $h_0, h_1 \in \mathbb{R}$. Then V on $[0, (1 + \alpha)/2]$ for some $\alpha \in (0, 1)$, h_0 , and a subset $S \subseteq \sigma(H)$ of all the eigenvalues $\sigma(H)$ of H satisfying*

$$\#\{\lambda \in S \mid \lambda \leq \lambda_0\} \geq (1 - \alpha)\#\{\lambda \in \sigma(H) \mid \lambda \leq \lambda_0\} + (\alpha/2) \quad (5.6)$$

for all sufficiently large $\lambda_0 \in \mathbb{R}$, uniquely determine h_1 and V on all of $[0, 1]$.

REMARK 5.13. (i) As a typical example, knowing slightly more than half the eigenvalues and knowing V on $[0, \frac{3}{4}]$ determines V uniquely on all of $[0, 1]$. To the best of our knowledge, Theorem 5.12 introduced and solved a new type of inverse spectral problem.

(ii) As in the case $\alpha = 0$, one has an extension of the same type as Theorem 5.10. Explicitly, if V is assumed to be C^{2k} near $x = (1 + \alpha)/2$, we only need

$$\#\{\lambda \in S \mid \lambda \leq \lambda_0\} \geq (1 - \alpha)\#\{\lambda \in \sigma(H) \mid \lambda \leq \lambda_0\} + (\alpha/2) - (k + 1)$$

instead of (5.6).

One can also derive results about problems on all of \mathbb{R} .

THEOREM 5.14. *Suppose that $V \in L^1_{loc}(\mathbb{R})$ satisfies the following two conditions:*

- (i) $V(x) \geq C|x|^{2+\varepsilon} - D$ for some $C, \varepsilon, D > 0$.
- (ii) $V(-x) \geq V(x) \quad x \geq 0$.

Then V on $[0, \infty)$ and the spectrum of $H = -\frac{d^2}{dx^2} + V$ in $L^2(\mathbb{R})$ uniquely determine V on all of \mathbb{R} .

Hochstadt–Lieberman [123] used the details of the inverse spectral theory in their proof. In a sense, we only used in [96] the main uniqueness theorem of that theory due to Marchenko [171], which we now describe. For $V \in L^1([a, b])$ real-valued, $-\infty < a < b < \infty$, consider $-u'' + Vu = zu$ with the boundary condition

$$u'(b) + h_b u(b) = 0 \tag{5.7}$$

at $x = b$. Let $u_+(z, x)$ denote the solution of this equation, normalized, say, by $u_+(z, b) = 1$. The m_+ -function is then defined by

$$m_+(z, a) = \frac{u'_+(z, a)}{u_+(z, a)}. \tag{5.8}$$

Similarly, given a boundary condition at $x = a$,

$$u'(a) + h_a u(a) = 0, \tag{5.9}$$

we define the solution $u_-(z, x)$ of $-u'' + Vu = zu$ normalized by $u_-(z, a) = 1$ and then define

$$m_-(z, b) = \frac{u'_-(z, b)}{u_-(z, b)}. \tag{5.10}$$

In our present context where $-\infty < a < b < \infty$, m_{\pm} are even meromorphic on \mathbb{C} . Moreover,

$$\text{Im}(z) > 0 \text{ implies } \text{Im}(m_-(z, b)) < 0, \text{Im}(m_+(z, a)) > 0.$$

Marchenko’s [171] fundamental uniqueness theorem of inverse spectral theory then reads as follows:

THEOREM 5.15. *$m_+(z, a)$ uniquely determines h_b as well as V a.e. on $[a, b]$.*

If $V \in L^1_{loc}([a, \infty))$ is real-valued (with $|a| < \infty$) and $-\frac{d^2}{dx^2} + V$ is in the limit point case at infinity, one can still define a unique $m_+(z, a)$ function but now for $\text{Im}(z) \neq 0$ rather than all $z \in \mathbb{C}$. For such z , there is a unique function $u_+(z, \cdot)$ which is L^2 at infinity (unique up to an overall scale factor which drops out of $m_+(z, a)$ defined by (5.8)). Again, one has the following uniqueness result independently proved by Borg [25] and Marchenko [171].

THEOREM 5.16. *$m_+(z, a)$ uniquely determines V a.e. on $[a, \infty)$.*

It is useful to have $m_-(z, b)$ because of the following basic fact:

THEOREM 5.17. *Let $H = -\frac{d^2}{dx^2} + V$ be a Schrödinger operator in $L^2([a, b])$ with boundary conditions (5.7) and (5.9) and let $G(z, x, y)$ be the integral kernel of $(H - z)^{-1}$. Suppose $c \in (a, b)$ and let $m_+(z, c)$ be the corresponding m_+ -function for $[c, b]$ and $m_-(z, c)$ the m_- -function for $[a, c]$. Then*

$$G(z, c, c) = \frac{1}{m_-(z, c) - m_+(z, c)}. \tag{5.11}$$

Theorems 5.15 and 5.16 are deep facts; Theorem 5.17 is an elementary calculation following from the explicit formula for the integral kernel of $(H - z)^{-1}$,

$$G(z, x, y) = \frac{u_-(z, \min(x, y))u_+(z, \max(x, y))}{W(u_-(z), u_+(z))(x)},$$

where as usual $W(f, g)(x) = f'(x)g(x) - f(x)g'(x)$ denotes the Wronskian of f and g . An analog of Theorem 5.17 holds in case $[a, b]$ is replaced by $(-\infty, \infty)$.

We can now describe the strategy of our proofs of Theorems 5.9–5.14. $G(z, c, c)$ has poles at the eigenvalues of H (this is not quite true; see below), so by (5.11), at eigenvalues λ_n of H :

$$m_+(\lambda_n, c) = m_-(\lambda_n, c). \tag{5.12}$$

If we know V on a left partial interval $[a, c]$ and we know some eigenvalue λ_n , then we know $m_-(z, c)$ exactly; so by (5.12), we know the value of $m_+(\lambda_n, c)$ at the point λ_n . Below we indicate when knowing the values of $f(\lambda_n)$ of an analytic function of the type of the m -functions uniquely determines $f(z)$. If $m_+(z, c)$ is determined, then by Theorem 5.15, V is determined on $[a, b]$ and so is h_b .

So the logic of the argument for a theorem like Theorem 5.9 is the following:

- (i) V on $[0, \frac{1}{2}]$ and h_0 determine $m_-(z, \frac{1}{2})$ by direct spectral theory.
- (ii) The λ_n and (5.12) determine $m_+(\lambda_n, \frac{1}{2})$, and then by suitable theorems in complex analysis, $m_+(z, \frac{1}{2})$ is uniquely determined for all z .
- (iii) $m_+(z, \frac{1}{2})$ uniquely determines V (a.e.) on $[\frac{1}{2}, 1]$ and h_1 by inverse spectral theory.

It is clear from this approach why h_0 is required and h_1 is free in the context of Theorem 5.9 (see [49] for examples where h_1 and $V \upharpoonright [0, \frac{1}{2}]$ do not determine V); without h_0 we cannot compute $m_-(z, \frac{1}{2})$ and so start the process.

As indicated before (5.12), $G(z, c, c)$ may not have a pole at an eigenvalue λ_n of H . It will if $u_n(c) \neq 0$, but if $u_n(c) = 0$, then $G(z, c, c) = 0$ rather than ∞ . Here u_n denotes the eigenfunction of H associated with the (necessarily simple) eigenvalue λ_n . Nevertheless, (5.12) holds at points where $u_n(c) = 0$ since then $u_-(c) = u_+(c) = 0$, and so both sides of (5.12) are infinite. (In spite of (5.12), $m_- - m_+$ is also infinite at $z = \lambda_n$ and so $G(\lambda_n, c, c) = 0$.) We summarize this discussion next:

THEOREM 5.18. *For any $c \in (a, b)$, (5.12) holds at any eigenvalue λ_n of $H_{[a,b]}$ (with the possibility of both sides of (5.12) being infinite).*

More recent references: A new inverse nodal problem was reduced to Theorem 5.12 by Yang [251]. A substantial generalization of Theorem 5.14, replacing condition (i) by H being bounded from below with purely discrete spectrum, was proved by Khodakovsky [137], [138]. He also found other variants of Theorem 5.14.

* * *

We end our survey of [96] by briefly indicating the uniqueness theorems for entire functions needed in the proofs of Theorems 5.9–5.14. In discussing extensions of Hochstadt’s discrete (finite matrix) version [122] of the Hochstadt–Lieberman theorem in [94], we made use of the following simple lemma which is an elementary consequence of the fact that any polynomial of degree d with $d + 1$ zeros must be the zero polynomial:

LEMMA 5.19. Suppose $f_1 = P_1/Q_1$ and $f_2 = P_2/Q_2$ are two rational fractions where the polynomials satisfy $\deg(P_1) = \deg(P_2)$ and $\deg(Q_1) = \deg(Q_2)$. Suppose that $d = \deg(P_1) + \deg(Q_1)$ and that $f_1(z_n) = f_2(z_n)$ for $d + 1$ distinct points $\{z_n\}_{n=1}^{d+1} \in \mathbb{C}$. Then $f_1 = f_2$.

In the context of [96], one is interested in entire functions of the form

$$f(z) = C \prod_{n=0}^{\infty} \left(1 - \frac{z}{x_n}\right), \tag{5.13}$$

where $0 < x_0 < x_1 < \dots$ is a suitable sequence of positive numbers which are the zeros of f and C is some complex constant.

Given a sequence $\{x_n\}_{n=0}^{\infty}$ of positive reals, we define

$$N(t) = \#\{n \in \mathbb{N} \cup \{0\} \mid x_n < t\}.$$

We recall the following basic theorem (see, e.g., [152, Ch. I], [173, Sects. II.48, II.49]):

THEOREM 5.20. Fix $0 < \rho_0 < 1$. Then:

(i) If $\{x_n\}_{n=0}^{\infty}$ is a sequence of positive reals with

$$\sum_{n=0}^{\infty} x_n^{-\rho} < \infty \text{ for all } \rho > \rho_0, \tag{5.14}$$

then the product in (5.13) defines an entire function f with

$$|f(z)| \leq C_1 \exp(C_2 |z|^\rho) \text{ for all } \rho > \rho_0. \tag{5.15}$$

(ii) Conversely, if f is an entire function satisfying (5.15) with all its (complex) zeros on $(0, \infty)$, then its zeros $\{x_n\}_{n=0}^{\infty}$ satisfy (5.14), and f has the canonical product expansion (5.13).

Moreover, (5.14) holds if and only if

$$N(t) \leq C |t|^\rho \text{ for all } \rho > \rho_0. \tag{5.16}$$

Given this theorem, we single out the following definition.

DEFINITION 5.21. A function f is called of m -type if and only if f is an entire function satisfying (5.15) (of order $0 < \rho < 1$ in the usual definition) with all the zeros of f on $(0, \infty)$.

Our choice of “ m -type” in Definition 5.21 comes from the fact that in many cases we discuss in this paper, the m -function is a ratio of functions of m -type. By Theorem 5.20, f in Definition 5.21 has the form (5.13) and $N(t)$, which we will denote as $N_f(t)$, satisfies (5.16).

LEMMA 5.22. Let f be a function of m -type. Then there exists a $0 < \rho < 1$ and a sequence $\{R_k\}_{k=1}^{\infty}$, $R_k \rightarrow \infty$ as $k \rightarrow \infty$, so that

$$\inf\{|f(z)| \mid |z| = R_k\} \geq C_1 \exp(-C_2 R_k^\rho).$$

LEMMA 5.23. Let F be an entire function that satisfies the following two conditions:

(i) $\sup_{|z|=R_k} |F(z)| \leq C_1 \exp(C_2 R_k^\rho)$ for some $0 \leq \rho < 1$, $C_1, C_2 > 0$, and some sequence $R_k \rightarrow \infty$ as $k \rightarrow \infty$.

(ii) $\lim_{|x| \rightarrow \infty; x \in \mathbb{R}} |F(ix)| = 0$.
 Then $F \equiv 0$.

Lemmas 5.22 and 5.23 finally yield the following result.

THEOREM 5.24. *Let f_1, f_2, g be three functions of m -type so that the following two conditions hold:*

- (i) $f_1(z) = f_2(z)$ at any point z with $g(z) = 0$.
- (ii) For all sufficiently large t ,

$$\max(N_{f_1}(t), N_{f_2}(t)) \leq N_g(t) - 1.$$

Then, $f_1 = f_2$.

* * *

Refinements of the results of [96] can be found in [50], [51]. Here we just mention the following facts.

THEOREM 5.25. *Let $H_1(h_0), H_2(h_0)$ be associated with two potentials V_1, V_2 on $[0, 1]$ and two potentially distinct boundary conditions $h_1^{(1)}, h_1^{(2)} \in \mathbb{R}$ at $x = 1$. Suppose that $\{(\lambda_n, h_0^{(n)})\}_{n \in \mathbb{N}_0}$ is a sequence of pairs with $\lambda_0 < \lambda_1 < \dots \rightarrow \infty$ and $h_0^{(n)} \in \mathbb{R} \cup \{\infty\}$ so that both $H_1(h_0^{(n)})$ and $H_2(h_0^{(n)})$ have eigenvalues at λ_n . Suppose that*

$$\sum_{n=0}^{\infty} \frac{(\lambda_n - \frac{1}{4}\pi^2 n^2)_+}{n^2} < \infty$$

holds. Then $V_1 = V_2$ a.e. on $[0, 1]$ and $h_1^{(1)} = h_1^{(2)}$.

This implies Borg's celebrated two-spectra uniqueness result [24] (see also, [154], [161], [159, Ch. 3], [171]):

COROLLARY 5.26. *Fix $h_0^{(1)}, h_0^{(2)} \in \mathbb{R}$. Then all the eigenvalues of $H(h_0^{(1)})$ and all the eigenvalues of $H(h_0^{(2)})$, save one, uniquely determine V a.e. on $[0, 1]$.*

It also implies the following amusing result:

COROLLARY 5.27. *Let $h_0^{(1)}, h_0^{(2)}, h_0^{(3)} \in \mathbb{R}$ and denote by $\sigma_j = \sigma(H(h_0^{(j)}))$ the spectra of $H(h_0^{(j)})$, $j = 1, 2, 3$. Assume $S_j \subseteq \sigma_j$, $j = 1, 2, 3$ and suppose that for all sufficiently large $\lambda_0 > 0$ one has*

$$\#\{\lambda \in \{S_1 \cup S_2 \cup S_3\} \text{ with } \lambda \leq \lambda_0\} \geq \frac{2}{3} \#\{\lambda \in \{\sigma_1 \cup \sigma_2 \cup \sigma_3\} \text{ with } \lambda \leq \lambda_0\} - 1.$$

Then V is uniquely determined a.e. on $[0, 1]$.

In particular, two-thirds of three spectra determine V .

More recent references: Further refinements of Corollary 5.27, involving N spectra, were proved by Horváth [124] (he also studies the corresponding analog for a Dirac-type operator). Optimal and nearly optimal conditions for a set of eigenvalues to determine the potential in terms of closedness properties of the exponential system corresponding to the known eigenvalues (implying Theorem 5.25 and a generalization thereof) were also derived by Horváth [125]. For an interesting half-line problem related to this circle of ideas we also refer to Horváth [126]. A variant of Theorem 5.25 was discussed by Ramm [197], [198]. Hochstadt–Lieberman-type problems for Schrödinger operators including a reconstruction algorithm have been

presented by L. Sakhnovich. The analog of the two-spectra result, Corollary 5.26, including a reconstruction algorithm, for a class of singular potentials has been discussed by Hryniv and Mykytyuk [127], [129] (see also [131]). They also studied Hochstadt–Lieberman-type results for such a class of singular potentials in [130]. Hochstadt–Lieberman-type results for a class of Dirac-type operators relevant to the AKNS system were published by del Rio and Grébert [52]. Borg- and Hochstadt–Lieberman-type inverse problems for systems including matrix-valued Schrödinger and Dirac-type equations, were studied in depth by M. Malamud [164]–[167]. He also studied Borg-type theorems for n th-order scalar equations [168]. Borg- and Hochstadt–Lieberman-type inverse problems for matrix-valued Schrödinger operators were also studied by Shen [220]. He also considered Borg-type inverse problems for Schrödinger operators with weights [221].

Additional results on determining the potential uniquely from spectra associated to three intervals of the type $[0, 1]$, $[0, a]$, and $[a, 1]$ for some $a \in (0, 1)$ (and similarly for whole-line problems with purely discrete spectra) can be found in [95]. This has been inspired by work of Pivovrachik [184], who also addressed the reconstruction algorithm from three spectra in the symmetric case $a = 1/2$ (see also [185], [188], [189]). He also considered the analogous Sturm–Liouville problem applicable to a smooth inhomogeneous partially damped string in [186] and extended some of these results to Sturm–Liouville equations on graphs in [187], [190]. Uniqueness and characterization problems for a class of singular Sturm–Liouville problems associated with three spectra were studied by Hryniv and Mykytyuk [128]. The reconstruction of a finite Jacobi matrix from three of its spectra was presented by Michor and Teschl [181].

* * *

These results are related to two other papers: In [94], we considered, among other topics, analogs of Theorems 5.9 and 5.12 for finite tri-diagonal (Jacobi) matrices, extending a result in [122]. The approach there is very similar to the current one except that the somewhat subtle theorems on zeros of entire functions in this paper are replaced by the elementary fact that a polynomial of degree at most N with $N + 1$ zeros must be identically zero. In [93], we considered results related to Theorem 5.14 in that for Schrödinger operators on $(-\infty, \infty)$, “spectral” information plus the potential on one of the half-lines determine the potential on all of $(-\infty, \infty)$. In that paper, we considered situations where there are scattering states for some set of energies and the “spectral” data are given by a reflection coefficient on a set of positive Lebesgue measure in the a.c. spectrum of H . The approach is not as close to this paper as is [94], but m -function techniques (see also [92]) are critical in all three papers.

More recent references: For additional results on inverse scattering with partial information on the potential we refer to Aktosun and Papanicolaou [2], Aktosun and Sacks [3], Aktosun and Weder [4], and the references therein.

* * *

We conclude this section by briefly describing some of the results in [94], where inverse spectral analysis for finite and semi-infinite Jacobi operators H was studied. While discussing a variety of topics (including trace formulas), we also provided a new proof of a result of Hochstadt [122] and its extension, which can be viewed as the discrete analog of the Hochstadt and Lieberman result in [123]. Moreover, we

solved the inverse spectral problem for $(\delta_n, (H - z)^{-1}\delta_n)$ in the case of finite Jacobi matrices. As mentioned earlier, the tools we apply are grounded in m -function techniques.

Explicitly, [94] studied finite $N \times N$ matrices of the form:

$$H = \begin{pmatrix} b_1 & a_1 & 0 & 0 & \cdot & \cdot & \cdot \\ a_1 & b_2 & a_2 & 0 & \cdot & \cdot & \cdot \\ 0 & a_2 & b_3 & a_3 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 0 & a_{N-1} & b_N \end{pmatrix}$$

and the semi-infinite analog H defined on

$$\ell^2(\mathbb{N}) = \left\{ u = (u(1), u(2), \dots) \mid \sum_{n=1}^{\infty} |u(n)|^2 < \infty \right\}$$

given by

$$(Hu)(n) = \begin{cases} a_n u(n+1) + b_n u(n) + a_{n-1} u(n-1), & n \geq 2, \\ a_1 u(2) + b_1 u(1), & n = 1. \end{cases}$$

In both cases, we assume $a_n, b_n \in \mathbb{R}$ with $a_n > 0$. To avoid inessential technical complications, we will only consider the case where $\sup_n [|a_n| + |b_n|] < \infty$ in which case H is a map from ℓ^2 to ℓ^2 and defines a bounded self-adjoint operator. In the semi-infinite case, we will set $N = \infty$. It will also be useful to consider the b 's and a 's as a single sequence $b_1, a_1, b_2, a_2, \dots = c_1, c_2, \dots$, that is,

$$c_{2n-1} = b_n, \quad c_{2n} = a_n, \quad n \in \mathbb{N}.$$

Concerning the recovery of a finite Jacobi matrix from parts of the matrix and additional spectral information (i.e., mixed data), Hochstadt [122] proved the following remarkable theorem.

THEOREM 5.28. *Let $N \in \mathbb{N}$. Suppose that c_{N+1}, \dots, c_{2N-1} are known, as well as the eigenvalues $\lambda_1, \dots, \lambda_N$ of H . Then c_1, \dots, c_N are uniquely determined.*

The discrete Hochstadt–Lieberman-type theorem proved in [94] reads as follows.

THEOREM 5.29. *Suppose that $1 \leq j \leq N$ and c_{j+1}, \dots, c_{2N-1} are known, as well as j of the eigenvalues. Then c_1, \dots, c_j are uniquely determined.*

We emphasize that one need *not* know which of the j eigenvalues one has.

Borg [24] proved the celebrated theorem that the spectra for two boundary conditions of a bounded interval regular Schrödinger operator uniquely determine the potential. Later refinements (see, e.g., [25], [120], [153], [154], [161], [171]) imply that they even determine the two boundary conditions.

Next, we consider analogs of this result for a finite Jacobi matrix. Such analogs were first considered by Hochstadt [119], [121] (see also [23], [67], [110], [111], [114], [122]). The results below are adaptations of known results for the continuum or the semi-infinite case, but the ability to determine parameters by counting sheds light on facts like the one that the lowest eigenvalue in the Borg result is not needed under certain circumstances.

Given H , an $N \times N$ Jacobi matrix, one defines $H(b)$ to be the Jacobi matrix where all a 's and b 's are the same as H , except b_1 is replaced by $b_1 + b$, that is,

$$H(b) = H + b(\delta_1, \cdot)\delta_1.$$

THEOREM 5.30. *The eigenvalues $\lambda_1, \dots, \lambda_N$ of H , together with b and $N - 1$ eigenvalues $\lambda(b)_1, \dots, \lambda(b)_{N-1}$ of $H(b)$, determine H uniquely.*

Again it is irrelevant which $N - 1$ eigenvalues of the N eigenvalues of $H(b)$ are known.

THEOREM 5.31. *The eigenvalues $\lambda_1, \dots, \lambda_N$ of H , together with the N eigenvalues $\lambda(b)_1, \dots, \lambda(b)_N$ of some $H(b)$ (with b unknown), determine H and b .*

REMARK 5.32. Since

$$b = \text{Tr}(H(b) - H) = \sum_{j=1}^N (\lambda(b)_j - \lambda_j),$$

we can a priori deduce b from the $\lambda(b)$'s and λ 's and so deduce Theorem 5.31 from Theorem 5.30. We note that the parameter counting works out. In Theorem 5.30, $2n - 1$ eigenvalues determine $2n - 1$ parameters; and in Theorem 5.31, $2n$ eigenvalues determine $2n$ parameters.

The basic inverse spectral theorem for finite Jacobi matrices shows that $(\delta_1, (H - z)^{-1}\delta_1)$ determines H uniquely. In [94] we considered $N \in \mathbb{N}$, $1 \leq n \leq N$, and asked whether $(\delta_n, (H - z)^{-1}\delta_n)$ determines H uniquely. For notational convenience, we occasionally allude to $G(z, n, n)$ as the n, n Green's function in the remainder of this section. The $n = 1$ result can be summarized via:

THEOREM 5.33. *$(\delta_1, (H - z)^{-1}\delta_1)$ has the form $\sum_{j=1}^N \alpha_j (\lambda_j - z)^{-1}$ with $\lambda_1 < \dots < \lambda_N$, $\sum_{j=1}^N \alpha_j = 1$ and each $\alpha_j > 0$. Every such sum arises as the $1, 1$ Green's function of an H and of exactly one such H .*

For general n , define $\tilde{n} = \min(n, N + 1 - n)$. Then the following theorems were proved in [94]:

THEOREM 5.34. *$(\delta_n, (H - z)^{-1}\delta_n)$ has the form $\sum_{j=1}^k \alpha_j (\lambda_j - z)^{-1}$ with k one of $N, N - 1, \dots, N - \tilde{n} + 1$ and $\lambda_1 < \dots < \lambda_k$, $\sum_{j=1}^k \alpha_j = 1$ and each $\alpha_j > 0$. Every such sum arises as the n, n Green's function of at least one H .*

THEOREM 5.35. *If $k = N$, then precisely $\binom{N-1}{n-1}$ operators H yield the given n, n Green's function.*

THEOREM 5.36. *If $k < N$, then infinitely many Jacobi matrices H yield the given n, n Green's function. Indeed, the inverse spectral family is then a collection of $\binom{k-1}{N-k} \binom{k-1-N+k}{n-1-N+k}$ disjoint manifolds, each of dimension $N - k$ and diffeomorphic to an $(N - k)$ -dimensional open ball.*

More recent references: Additional geometric information in connection with Theorem 5.36 and a version for off-diagonal Green's functions were studied by Gibson [108]. Borg- and discrete Hochstadt-Lieberman-type results for generalized (i.e., certain tri-diagonal block) Jacobi matrices were studied by Derevyagin [54] (see also Shieh [222]). The case of non-self-adjoint Jacobi matrices with a rank-one imaginary part, and an extension of Hochstadt-Lieberman-type results to this

situation was recently discussed by Arlinskiĭ and Tsekanovskiĭ [11]. An extension of results of Hochstadt [121] to the case of normal matrices was found by S. Malamud [169]. A detailed treatment of two-spectra inverse problems of semi-infinite Jacobi operators, including reconstruction, has recently been presented by Silva and Weder [224].

6. The Crown Jewel: Simon's New Approach to Inverse Spectral Theory

In this section we summarize some of the principal results of the following papers:

- [228] B. Simon, *A new approach to inverse spectral theory, I. Fundamental formalism*, Ann. Math. **150**, 1029–1057 (1999).
- [97] F. Gesztesy and B. Simon, *A new approach to inverse spectral theory, II. General real potentials and the connection to the spectral measure*, Ann. Math. **152**, 593–643 (2000).
- [200] A. Ramm and B. Simon, *A new approach to inverse spectral theory, III. Short range potentials*, J. Analyse Math. **80**, 319–334 (2000).
- [98] F. Gesztesy and B. Simon, *On local Borg–Marchenko uniqueness results*, Commun. Math. Phys. **211**, 273–287 (2000).

As the heading of this section suggests, we are approaching the pinnacle of Barry Simon's contributions to inverse scattering theory thus far: In his spectacular paper [228], he single-handedly developed a new approach to inverse spectral theory for Schrödinger operators on a half-line, by starting from a particular representation of the Weyl–Titchmarsh m -function as a finite Laplace-type transform with control over the error term. In addition to establishing this feat, it also led to a completely unexpected uniqueness result for Weyl–Titchmarsh functions, what is now called the local Borg–Marchenko uniqueness theorem, but which really should have been named Simon's local uniqueness theorem. The inverse spectral approach for Schrödinger operators on a half-line (including a reconstruction algorithm for the potential) originated with the celebrated paper [70] by Gel'fand and Levitan in 1951 and an independent approach by Krein [142] in the same year, followed by a seminal contribution [171] by Marchenko in 1952. The Borg–Marchenko uniqueness result was first published by Marchenko [170] in 1950 but Borg apparently had it in 1949 and it was independently published by Borg [25] and again by Marchenko [171] in 1952. Both results, the uniqueness theorem and the Gel'fand–Levitan (reconstruction) formalism, remained pillars of the inverse spectral theory that withstood any reformulation or improvement for nearly fifty years. Hence it was an incredible achievement by Barry Simon to have changed the inverse spectral landscape by offering such a reformulation of inverse spectral theory and in the very same paper [228] to have been able to substantially improve the Borg–Marchenko uniqueness theorem from a global to a local version.

We start by highlighting the approach in Simon's paper [228] and then switch to a more detailed treatment of some aspects of the theory by borrowing from [97].

Inverse spectral methods have been actively studied in the past years both via their relevance in a variety of applications and due to their connection with integrable evolution equations such as the KdV equation. In this section, however,

we will not deal with the full-line inverse spectral approach relevant to integrable equations but exclusively focus on inverse spectral theory for half-line Schrödinger operators. In this particular context, a major role is played by the Gel'fand–Levitan equations [70] (see also, [30, Chs. 3, 4], [31], [143], [144], [145], [159, Ch. 2], [171], [172, Ch. 2], [182, Ch. VIII], [204], [236], [244]). The goal in Simon's paper [228] was to present a new approach to their basic results. In particular, he introduced a new basic object, the A -function (see (6.20) below), the remarkable equation (6.23) it satisfies, and illustrated its fundamental importance with several new results including improved asymptotic expansions of the Weyl–Titchmarsh m -function in the high-energy regime and the local uniqueness result.

To present some of these new results, we will first describe the major players in this game. One is concerned with self-adjoint differential operators on either $L^2([0, b])$ with $b < \infty$, or $L^2([0, \infty))$ associated with differential expressions of the form

$$-\frac{d^2}{dx^2} + V(x), \quad x \in (0, b). \tag{6.1}$$

If b is finite, we suppose

$$\int_0^b dx |V(x)| < \infty$$

and place a boundary condition

$$u'(b) + hu(b) = 0 \tag{6.2}$$

at b , where $h \in \mathbb{R} \cup \{\infty\}$ with $h = \infty$ shorthand for the Dirichlet boundary condition $u(b) = 0$. If $b = \infty$, we suppose

$$\int_y^{y+1} dx |V(x)| < \infty \text{ for all } y \geq 0$$

and

$$\sup_{y>0} \int_y^{y+1} dx \max(V(x), 0) < \infty. \tag{6.3}$$

Under condition (6.3), it is known that (6.1) is limit point at infinity [201, App. to Sect. X.1]. In addition, a fixed self-adjoint boundary condition at $x = 0$ is assumed when talking about the self-adjoint operator associated with (6.1).

In either case, for each $z \in \mathbb{C} \setminus [\beta, \infty)$ with $-\beta$ sufficiently large, there is a unique solution (up to an overall constant), $u(z, x)$, of $-u'' + Vu = zu$ which satisfies (6.2) at b if $b < \infty$ or which is L^2 at ∞ if $b = \infty$. The principal m -function $m(z)$ is defined by

$$m(z) = \frac{u'(z, 0)}{u(z, 0)}.$$

If we replace b by $b_1 = b - x_0$ with $x_0 \in (0, b)$ and let $V(s) = V(x_0 + s)$ for $s \in (0, b_1)$, we get a new m -function we will denote by $m(z, x_0)$. It is given by

$$m(z, x) = \frac{u'(z, x)}{u(z, x)}.$$

$m(z, x)$ satisfies the Riccati-type equation

$$\frac{d}{dx} m(z, x) = V(x) - z - m^2(z, x). \tag{6.4}$$

Obviously, $m(z, x)$ depends only on V on (x, b) (and on h if $b < \infty$). A basic result of the inverse spectral theory says that the converse is true as was shown independently by Borg [25] and Marchenko [171] in 1952:

THEOREM 6.1. *m uniquely determines V . Explicitly, if V_j are potentials with corresponding m -functions m_j , $j = 1, 2$, and $m_1 = m_2$, then $V_1 = V_2$ a.e. (including $h_1 = h_2$).*

In 1999, Simon [228] spectacularly improved this to obtain a local version of the Borg–Marchenko uniqueness result as follows:

THEOREM 6.2. *If (V_1, b_1, h_1) , (V_2, b_2, h_2) are two potentials and $a < \min(b_1, b_2)$ and if*

$$V_1(x) = V_2(x) \text{ on } (0, a), \tag{6.5}$$

then as $\kappa \rightarrow \infty$,

$$m_1(-\kappa^2) - m_2(-\kappa^2) = \tilde{O}(e^{-2\kappa a}). \tag{6.6}$$

Conversely, if (6.6) holds, then (6.5) holds.

In (6.6), we use the symbol \tilde{O} defined by

$$f = \tilde{O}(g) \text{ as } x \rightarrow x_0 \text{ (where } \lim_{x \rightarrow x_0} g(x) = 0)$$

$$\text{if and only if } \lim_{x \rightarrow x_0} \frac{|f(x)|}{|g(x)|^{1-\varepsilon}} = 0 \text{ for all } \varepsilon > 0.$$

From a results point of view, this local version of the Borg–Marchenko uniqueness theorem was the most significant new result in Simon’s paper [228], but a major thrust of this paper was the new set of methods introduced which led to a new approach of the inverse spectral problem. Theorem 6.2 implies that V is determined by the asymptotics of $m(-\kappa^2)$ as $\kappa \rightarrow \infty$. One can also read off differences of the boundary condition from these asymptotics. Moreover, the following result is proved in [228]:

THEOREM 6.3. *Let (V_1, b_1, h_1) , (V_2, b_2, h_2) be two potentials and suppose that*

$$b_1 = b_2 \equiv b < \infty, \quad |h_1| + |h_2| < \infty, \quad V_1(x) = V_2(x) \text{ on } (0, b). \tag{6.7}$$

Then

$$\lim_{\kappa \rightarrow \infty} e^{2b\kappa} |m_1(-\kappa^2) - m_2(-\kappa^2)| = 4(h_1 - h_2). \tag{6.8}$$

Conversely, if (6.8) holds for some $b < \infty$ with a limit in $(0, \infty)$, then (6.7) holds.

To understand Simon’s new approach, it is useful to recall briefly the two approaches to the inverse problem for Jacobi matrices on $\ell^2(\mathbb{N}_0)$ [19, Ch. VII], [94], [231]:

$$A = \begin{pmatrix} b_0 & a_0 & 0 & 0 & \dots \\ a_0 & b_1 & a_1 & 0 & \dots \\ 0 & a_1 & b_2 & a_2 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

with $a_j > 0, b_j \in \mathbb{R}$. Here the m -function is just $(\delta_0, (A - z)^{-1} \delta_0) = m(z)$ and, more generally, $m_n(z) = (\delta_n, (A^{(n)} - z)^{-1} \delta_n)$ with $A^{(n)}$ on $\ell^2(\{n, n + 1, \dots\})$ obtained by truncating the first n rows and n columns of A . Here δ_n is the Kronecker vector, that is, the vector with 1 in slot n and 0 in other slots. The fundamental theorem in this case is that $m(z) \equiv m_0(z)$ determines the b_n ’s and a_n ’s.

$m_n(z)$ satisfies an analog of the Riccati equation (6.4):

$$a_n^2 m_{n+1}(z) = b_n - z - \frac{1}{m_n(z)}. \tag{6.9}$$

One solution of the inverse problem is to turn (6.9) around to see that

$$m_n(z)^{-1} = -z + b_n - a_n^2 m_{n+1}(z) \tag{6.10}$$

which, first of all, implies that as $z \rightarrow \infty$, $m_n(z) = -z^{-1} + O(z^{-2})$, so (6.10) implies

$$m_n(z)^{-1} = -z + b_n + a_n^2 z^{-1} + O(z^{-2}). \tag{6.11}$$

Thus, (6.11) for $n = 0$ yields b_0 and a_0^2 and so $m_1(z)$ by (6.9), and then an obvious induction yields successive b_k , a_k^2 , and $m_{k+1}(z)$.

A second solution involves orthogonal polynomials. Let $P_n(z)$ be the eigensolutions of the formal $(A - z)P_n = 0$ with boundary conditions $P_{-1}(z) = 0$, $P_0(z) = 1$. Explicitly,

$$P_{n+1}(z) = a_n^{-1}[(z - b_n)P_n(z)] - a_{n-1}P_{n-1}. \tag{6.12}$$

Let $d\rho$ be the spectral measure for A and vector δ_0 so that

$$m(z) = \int_{\mathbb{R}} \frac{d\rho(\lambda)}{\lambda - z}. \tag{6.13}$$

Then one can show that

$$\int_{\mathbb{R}} d\mu(\lambda) P_n(\lambda)P_m(\lambda) = \delta_{n,m}, \quad n, m \in \mathbb{N}_0. \tag{6.14}$$

Thus, $P_n(z)$ is a polynomial of degree n with positive leading coefficients determined by (6.14). These orthonormal polynomials are determined via Gram-Schmidt from ρ and by (6.13) from m . Once one has the polynomials P_n , one can determine the a 's and b 's from equation (6.12).

Of course, these approaches via Riccati equation and orthogonal polynomials are not completely disjoint. The Riccati solution gives the a_n 's and b_n 's as continued fractions and the connection between continued fractions and orthogonal polynomials played a fundamental role in Stieltjes' work [231] on the moment problem in 1895.

The Gel'fand-Levitan approach to the continuum case (cf. [70], [159, Ch. 2], [171], [172, Ch. 2]) is a direct analog of this orthogonal polynomial case. One looks at solutions $U(k, x)$ of

$$-U''(k, x) + V(x)U(k, x) = k^2U(k, x) \tag{6.15}$$

satisfying $U(k, 0) = 1$, $U'(k, 0) = ik$, and proves that they satisfy a representation

$$U(k, x) = e^{ikx} + \int_{-x}^x dy K(x, y)e^{iky}, \tag{6.16}$$

the analog of $P_n(z) = cz^n + \text{lower order}$. One defines $s(k, x) = (2ik)^{-1}[U(k, x) - U(-k, x)]$ which satisfies (6.15) with $s(k, 0_+) = 0$, $s'(k, 0_+) = 1$.

The spectral measure $d\rho$ associated to $m(z)$ by

$$d\rho(\lambda) = (2\pi)^{-1} \lim_{\varepsilon \downarrow 0} [\text{Im}(m(\lambda + i\varepsilon))] d\lambda$$

satisfies

$$\int_{\mathbb{R}} d\rho(k^2) s(k, x)s(k, y) = \delta(x - y), \tag{6.17}$$

at least formally. (6.16) and (6.17) yield an integral equation for K depending only on $d\rho$ and then once one has K , one can find U and hence V via (6.15) (or via another relation between K and V).

The principal goal in [228] was to present a new approach to the continuum case, that is, an analog of the Riccati equation approach to the discrete inverse problem. The simple idea for this is attractive but has a difficulty to overcome. $m(z, x)$ determines $V(x)$, at least if V is continuous by the known asymptotics ([45], [210]):

$$m(-\kappa^2, x) = -\kappa - \frac{V(x)}{2\kappa} + o(\kappa^{-1}). \tag{6.18}$$

We can therefore think of (6.4) with V defined by (6.18) as an evolution equation for m . The idea is that using a suitable underlying space and uniqueness theorem for solutions of differential equations, (6.4) should uniquely determine m for all positive x , and hence $V(x)$ by (6.18).

To understand the difficulty, consider a potential $V(x)$ on the whole real line. There are then functions $u_{\pm}(z, x)$ defined for $z \in \mathbb{C} \setminus [\beta, \infty)$ which are L^2 at $\pm\infty$ and two m -functions $m_{\pm}(z, x) = u'_{\pm}(z, x)/u_{\pm}(z, x)$. Both satisfy (6.4), yet $m_+(z, 0)$ determines and is determined by V on $(0, \infty)$ while $m_-(z, 0)$ has the same relation to V on $(-\infty, 0)$. Put differently, $m_+(z, 0)$ determines $m_+(z, x)$ for $x > 0$ but not at all for $x < 0$. m_- is the reverse. So uniqueness for (6.4) is one-sided and either side is possible! That this does not make the scheme hopeless is connected with the fact that m_- does not satisfy (6.18), but rather

$$m_-(-\kappa^2, x) = \kappa + \frac{V(x)}{2\kappa} + o(\kappa^{-1}). \tag{6.19}$$

We will see the one-sidedness of the solvability is intimately connected with the sign of the leading $\pm\kappa$ term in (6.18) and (6.19).

The key object in this new approach is a function $A(\alpha)$ defined for $\alpha \in (0, b)$ related to m by

$$m(-\kappa^2) = -\kappa - \int_0^a d\alpha A(\alpha)e^{-2\alpha\kappa} + \tilde{O}(e^{-2a\kappa}) \tag{6.20}$$

as $\kappa \rightarrow \infty$. We have written $A(\alpha)$ as a function of a single variable but we will allow similar dependence on other variables. Since $m(-\kappa^2, x)$ is also an m -function, (6.20) has an analog with a function $A(\alpha, x)$.

By uniqueness of inverse Laplace transforms (see [228, Appendix 2, Theorem A.2.2]), (6.20) and m near $-\infty$ uniquely determine $A(\alpha)$.

Not only will (6.20) hold but, in a sense, $A(\alpha)$ is close to $V(\alpha)$. Explicitly, one can prove the following result:

THEOREM 6.4. *Let m be the m -function of the potential V . Then there is a function $A \in L^1([0, b])$ if $b < \infty$ and $A \in L^1([0, a])$ for all $a < \infty$ if $b = \infty$ so that (6.20) holds for any $a \leq b$ with $a < \infty$. $A(\alpha)$ only depends on $V(y)$ for $y \in [0, \alpha]$. Moreover, $A(\alpha) = V(\alpha) + E(\alpha)$ where $E(\alpha)$ is continuous and satisfies*

$$|E(\alpha)| \leq \left(\int_0^\alpha dy |V(y)| \right)^2 \exp\left(\alpha \int_0^\alpha dy |V(y)| \right).$$

Restoring the x -dependence, we see that $A(\alpha, x) = V(\alpha + x) + E(\alpha, x)$ where

$$\lim_{\alpha \downarrow 0} \sup_{0 \leq x \leq a} |E(\alpha, x)| = 0$$

for any $a > 0$, so

$$\lim_{\alpha \downarrow 0} A(\alpha, x) = V(x), \tag{6.21}$$

where this holds in general in the L^1 -sense. If V is continuous, (6.21) holds pointwise. In general, (6.21) will hold at any point of right Lebesgue continuity of V .

Because E is continuous, A determines any discontinuities or singularities of V . More is true. If V is C^k , then E is C^{k+2} in α , and so A determines k th-order kinks in V . Much more is true and one can also prove the following result:

THEOREM 6.5. *V on $[0, a]$ is only a function of A on $[0, a]$. Explicitly, if V_1, V_2 are two potentials, let A_1, A_2 be their A -functions. If $a < b_1, a < b_2$, and $A_1(\alpha) = A_2(\alpha)$ for $\alpha \in [0, a]$, then $V_1(x) = V_2(x)$ for $x \in [0, a]$.*

Theorems 6.4 and 6.5 imply Theorem 6.2.

As noted, the singularities of V come from singularities of A . A boundary condition is a kind of singularity, so one might hope that boundary conditions correspond to very singular A . In essence, we will see that this is the case—there are delta-function and delta-prime singularities at $\alpha = b$. Explicitly, one can prove the following result:

THEOREM 6.6. *Let m be the m -function for a potential V with $b < \infty$. Then for $a < 2b$,*

$$m(-\kappa^2) = -\kappa - \int_0^a d\alpha A(\alpha) e^{-2\alpha\kappa} - A_1 \kappa e^{-2\kappa b} - B_1 e^{-2\kappa b} + \tilde{O}(e^{-2a\kappa}), \tag{6.22}$$

where the following facts hold:

- (a) If $h = \infty$, then $A_1 = 2, B_1 = -2 \int_0^b V(y) dy$.
- (b) If $|h| < \infty$, then $A_1 = -2, B_1 = 2[2h + \int_0^b V(y) dy]$.

This implies Theorem 6.3.

The reconstruction theorem, Theorem 6.5, depends on the differential equation that $A(\alpha, x)$ satisfies. Remarkably, V drops out of the translation of (6.4) to the equation for A :

$$\frac{\partial A(\alpha, x)}{\partial x} = \frac{\partial A(\alpha, x)}{\partial \alpha} + \int_0^\alpha d\beta A(\beta, x) A(\alpha - \beta, x). \tag{6.23}$$

If V is C^1 , the equation holds in classical sense. For general V , it holds in a variety of weaker senses. Either way, $A(\alpha, 0)$ for $\alpha \in [0, a]$ determines $A(\alpha, x)$ for all x, α with $\alpha > 0$ and $0 < x + \alpha < a$. (6.21) then determines $V(x)$ for $x \in [0, a]$. That is the essence from which uniqueness comes. We will return to this circle of ideas later on when discussing Simon’s approach to the inverse spectral problem in detail.

* * *

Now we switch to [97] and take a closer look at some of the concepts introduced in [228]. In particular, we continue the study of the A -amplitude associated to half-line Schrödinger operators, $-\frac{d^2}{dx^2} + V$ in $L^2([0, b))$, $b \leq \infty$. A is related to the Weyl–Titchmarsh m -function via $m(-\kappa^2) = -\kappa - \int_0^a d\alpha A(\alpha) e^{-2\alpha\kappa} + O(e^{-(2a-\varepsilon)\kappa})$ for all $\varepsilon > 0$. Three main issues will be discussed:

- First, we describe how to extend the theory to general V in $L^1([0, a])$ for all $a > 0$, including V ’s which are limit circle at infinity.

- Second, the following relation between the A -amplitude and the spectral measure ρ :

$$A(\alpha) = -2 \int_{-\infty}^{\infty} d\rho(\lambda) \lambda^{-\frac{1}{2}} \sin(2\alpha\sqrt{\lambda}),$$

will be discussed. (Since the integral is divergent, this formula has to be properly interpreted.)

- Third, a Laplace transform representation for m without error term in the case $b < \infty$ will be presented.

We consider Schrödinger operators

$$-\frac{d^2}{dx^2} + V \tag{6.24}$$

in $L^2([0, b])$ for $0 < b < \infty$ or $b = \infty$ and real-valued locally integrable V . There are essentially four distinct cases.

Case 1. $b < \infty$. We suppose $V \in L^1([0, b])$. We then pick $h \in \mathbb{R} \cup \{\infty\}$ and add the boundary condition at b

$$u'(b_-) + hu(b_-) = 0, \tag{6.25}$$

where $h = \infty$ is shorthand for the Dirichlet boundary condition $u(b_-) = 0$.

For Cases 2–4, $b = \infty$ and

$$\int_0^a dx |V(x)| < \infty \quad \text{for all } a < \infty. \tag{6.26}$$

Case 2. V is “essentially” bounded from below in the sense that

$$\sup_{a>0} \left(\int_a^{a+1} dx \max(-V(x), 0) \right) < \infty. \tag{6.27}$$

Examples include $V(x) = c(x + 1)^\beta$ for $c > 0$ and all $\beta \in \mathbb{R}$ or $V(x) = -c(x + 1)^\beta$ for all $c > 0$ and $\beta \leq 0$.

Case 3. (6.27) fails but (6.24) is limit point at ∞ (see, e.g., [41, Ch. 9], [201, Sect. X.1] for a discussion of limit point/limit circle), that is, for each $z \in \mathbb{C}_+ = \{z \in \mathbb{C} \mid \text{Im}(z) > 0\}$,

$$-u'' + Vu = zu \tag{6.28}$$

has a unique solution, up to a multiplicative constant, which is L^2 at ∞ . An example is $V(x) = -c(x + 1)^\beta$ for $c > 0$ and $0 < \beta \leq 2$.

Case 4. (6.24) is limit circle at infinity, that is, every solution of (6.28) is $L^2([0, \infty))$ at infinity if $z \in \mathbb{C}_+$. We then pick a boundary condition by choosing a nonzero solution u_0 of (6.28) for $z = i$. Other functions u satisfying the associated boundary condition at infinity then are supposed to satisfy

$$\lim_{x \rightarrow \infty} W(u_0, u)(x) = \lim_{x \rightarrow \infty} [u_0(x)u'(x) - u'_0(x)u(x)] = 0. \tag{6.29}$$

Examples include $V(x) = -c(x + 1)^\beta$ for $c > 0$ and $\beta > 2$.

The Weyl–Titchmarsh m -function, $m(z)$, is defined for $z \in \mathbb{C}_+$ as follows. Fix $z \in \mathbb{C}_+$. Let $u(x, z)$ be a nonzero solution of (6.28) which satisfies the boundary condition at b . In Case 1, that means u satisfies (6.25); in Case 4, it satisfies (6.29);

and in Cases 2–3, it satisfies $\int_R^\infty |u(z, x)|^2 dx < \infty$ for some (and hence for all) $R \geq 0$. Then,

$$m(z) = \frac{u'(z, 0_+)}{u(z, 0_+)} \tag{6.30}$$

and, more generally,

$$m(z, x) = \frac{u'(z, x)}{u(z, x)}. \tag{6.31}$$

$m(z, x)$ satisfies the Riccati equation (with $m' = \partial m / \partial x$),

$$m'(z, x) = V(x) - z - m(z, x)^2. \tag{6.32}$$

m is an analytic function of z for $z \in \mathbb{C}_+$, and the following properties hold:

Case 1. m is meromorphic in \mathbb{C} with a discrete set $\lambda_1 < \lambda_2 < \dots$ of poles on \mathbb{R} (and none on $(-\infty, \lambda_1)$).

Case 2. For some $\beta \in \mathbb{R}$, m has an analytic continuation to $\mathbb{C} \setminus [\beta, \infty)$ with m real on $(-\infty, \beta)$.

Case 3. In general, m cannot be continued beyond \mathbb{C}_+ (there exist V 's where m has a dense set of polar singularities on \mathbb{R}).

Case 4. m is meromorphic in \mathbb{C} with a discrete set of poles (and zeros) on \mathbb{R} with limit points at both $+\infty$ and $-\infty$.

Moreover,

$$\text{if } z \in \mathbb{C}_+ \text{ then } m(z, x) \in \mathbb{C}_+,$$

so m admits the Herglotz representation,

$$m(z) = \text{Re}(m(i)) + \int_{\mathbb{R}} d\rho(\lambda) \left[\frac{1}{\lambda - z} - \frac{\lambda}{1 + \lambda^2} \right], \quad z \in \mathbb{C} \setminus \mathbb{R}, \tag{6.33}$$

where ρ is a positive measure called the spectral measure, which satisfies

$$\int_{\mathbb{R}} \frac{d\rho(\lambda)}{1 + |\lambda|^2} < \infty, \tag{6.34}$$

$$d\rho(\lambda) = \text{w-lim}_{\varepsilon \downarrow 0} \frac{1}{\pi} \text{Im}(m(\lambda + i\varepsilon)) d\lambda, \tag{6.35}$$

where w-lim is meant in distributional sense.

All these properties of m are well known (see, e.g. [162, Ch. 2]).

In (6.33), the constant $\text{Re}(m(i))$ is determined by the result of Everitt [62] that for each $\varepsilon > 0$,

$$m(-\kappa^2) = -\kappa + o(1) \quad \text{as } |\kappa| \rightarrow \infty \text{ with } -\frac{\pi}{2} + \varepsilon < \arg(\kappa) < -\varepsilon < 0. \tag{6.36}$$

Atkinson [13] improved (6.36) to read,

$$m(-\kappa^2) = -\kappa + \int_0^{a_0} d\alpha V(\alpha) e^{-2\alpha\kappa} + o(\kappa^{-1}) \tag{6.37}$$

again as $|\kappa| \rightarrow \infty$ with $-\frac{\pi}{2} + \varepsilon < \arg(\kappa) < -\varepsilon < 0$ (actually, he allows $\arg(\kappa) \rightarrow 0$ as $|\kappa| \rightarrow \infty$ as long as $\text{Re}(\kappa) > 0$ and $\text{Im}(\kappa) > -\exp(-D|\kappa|)$ for suitable D). In (6.37), a_0 is any fixed $a_0 > 0$.

One of the main results in [97] was to go way beyond the two leading orders in (6.37).

THEOREM 6.7. *There exists a function $A(\alpha)$ for $\alpha \in [0, b)$ so that $A \in L^1([0, a])$ for all $a < b$ and*

$$m(-\kappa^2) = -\kappa - \int_0^a d\alpha A(\alpha)e^{-2\alpha\kappa} + \tilde{O}(e^{-2\alpha\kappa}) \tag{6.38}$$

as $|\kappa| \rightarrow \infty$ with $-\frac{\pi}{2} + \varepsilon < \arg(\kappa) < -\varepsilon < 0$. Here we say $f = \tilde{O}(g)$ if $g \rightarrow 0$ and for all $\varepsilon > 0$, $(\frac{f}{g})|g|^\varepsilon \rightarrow 0$ as $|\kappa| \rightarrow \infty$. Moreover, $A - q$ is continuous and

$$|(A - q)(\alpha)| \leq \left[\int_0^\alpha dx |V(x)| \right]^2 \exp \left(\alpha \int_0^\alpha dx |V(x)| \right). \tag{6.39}$$

This result was proved in Cases 1 and 2 in [228]. The proof of this result if one only assumes (6.26) (i.e., in Cases 3 and 4) has been provided in [97].

Actually, in [228], (6.38) was proved in Cases 1 and 2 for κ real with $|\kappa| \rightarrow \infty$. The proof in [97], assuming only condition (6.26), includes Case 2 in the general κ -region $\arg(\kappa) \in (-\frac{\pi}{2} + \varepsilon, -\varepsilon)$ and, as can be shown, the proof also holds in this region for Case 1.

REMARK 6.8. At first sight, it may appear that Theorem 6.7, as stated, does not imply the κ real result of [228], but if the spectral measure ρ of (6.33) has $\text{supp}(\rho) \subseteq [a, \infty)$ for some $a \in \mathbb{R}$, (6.38) extends to all κ in $|\arg(\kappa)| < \frac{\pi}{2} - \varepsilon$, $|\kappa| \geq a + 1$. To see this, one notes by (6.33) that $m'(z)$ is bounded away from $[a, \infty)$ so one has the a priori bound $|m(z)| \leq C|z|$ in the region $\text{Re}(z) < a - 1$. This bound and a Phragmén–Lindelöf argument let one extend (6.38) to the real κ axis.

The following is a basic result from [228]:

THEOREM 6.9 (Theorem 2.1 of [228]). *Let $V \in L^1([0, \infty))$. Then there exists a function A on $(0, \infty)$ so that $A - V$ is continuous and satisfies (6.39) such that for $\text{Re}(\kappa) > \|V\|_1/2$,*

$$m(-\kappa^2) = -\kappa - \int_0^\infty d\alpha A(\alpha)e^{-2\alpha\kappa}. \tag{6.40}$$

REMARK 6.10. In [228], this is only stated for κ real with $\kappa > \|V\|_1/2$, but (6.39) implies that $|A(\alpha) - V(\alpha)| \leq \|V\|_1^2 \exp(\alpha\|V\|_1)$ so the right-hand side of (6.40) converges to an analytic function in $\text{Re}(\kappa) > \|V\|_1/2$. Since $m(z)$ is analytic in $\mathbb{C} \setminus [\alpha, \infty)$ for suitable α , we have equality in $\{\kappa \in \mathbb{C} \mid \text{Re}(\kappa) > \|V\|_1/2\}$ by analyticity.

Theorem 6.7 in all cases follows from Theorem 6.9 and the following result which was proved in [97]:

THEOREM 6.11. *Let V_1, V_2 be potentials defined on $(0, b_j)$ with $b_j > a$ for $j = 1, 2$. Suppose that $V_1 = V_2$ on $[0, a]$. Then in the region $\arg(\kappa) \in (-\frac{\pi}{2} + \varepsilon, -\varepsilon)$, $|\kappa| \geq K_0$, we have that*

$$|m_1(-\kappa^2) - m_2(-\kappa^2)| \leq C_{\varepsilon, \delta} \exp(-2a\text{Re}(\kappa)), \tag{6.41}$$

where $C_{\varepsilon, \delta}$ depends only on ε, δ , and $\sup_{0 \leq x \leq a} (\int_x^{x+\delta} dy |V_j(y)|)$, where $\delta > 0$ is any number so that $a + \delta \leq b_j, j = 1, 2$.

REMARK 6.12. (i) An important consequence of Theorem 6.11 is that if $V_1(x) = V_2(x)$ for $x \in [0, a]$, then $A_1(\alpha) = A_2(\alpha)$ for $\alpha \in [0, a]$. Thus, $A(\alpha)$ is only a function of V on $[0, \alpha]$. (We emphasize that, conversely, one can show that also $V(x)$ is only a function of A on $[0, x]$.)

(ii) This implies Theorem 6.7 by taking $V_1 = V$ and $V_2 = V\chi_{[0,a]}$ and using Theorem 6.9 on V_2 .

(iii) The actual proof implies (6.41) on a larger region than $\arg(\kappa) \in (-\frac{\pi}{2} + \epsilon, -\epsilon)$. Basically, one needs $\text{Im}(\kappa) \geq -C_1 \exp(-C_2|\kappa|)$ as $\text{Re}(\kappa) \rightarrow \infty$.

The basic connection between the spectral measure $d\rho$ and the A -amplitude established in [97] says

$$A(\alpha) = -2 \int_{-\infty}^{\infty} d\rho(\lambda) \lambda^{-\frac{1}{2}} \sin(2\alpha\sqrt{\lambda}). \tag{6.42}$$

However, the integral in (6.42) is not convergent. Indeed, the asymptotics (6.36) imply that $\int_0^R d\rho(\lambda) \sim \frac{2}{3\pi} R^{\frac{3}{2}}$ so (6.42) is never absolutely convergent. Thus, (6.42) has to be suitably interpreted.

We will indicate how to demonstrate (6.42) as a distributional relation, smeared in α on both sides by a function $f \in C_0^\infty((0, \infty))$. This holds for all V 's in Cases 1–4. Finally, we will discuss an Abelianized version of (6.42), namely,

$$A(\alpha) = -2 \lim_{\epsilon \downarrow 0} \int_{-\infty}^{\infty} d\rho(\lambda) e^{-\epsilon\lambda} \lambda^{-\frac{1}{2}} \sin(2\alpha\sqrt{\lambda}) \tag{6.43}$$

at any point, α , of Lebesgue continuity for V . (6.43) is proved only for a restricted class of V 's including Cases 1 and 2 and those V 's satisfying

$$V(x) \geq -Cx^2, \quad x \geq R$$

for some $R > 0, C > 0$, which are always in the limit point case at infinity. Subsequently, we will use (6.43) as the point of departure for relating $A(\alpha)$ to scattering data.

In order to prove (6.42) for finite b , one needs to analyze the finite b case extending (6.38) to all a , including $a = \infty$ (by allowing A to have δ and δ' singularities at multiples of b). This was originally done in [228] for κ real and positive and $a < \infty$. We now need results in the entire region $\text{Re}(\kappa) \geq K_0$. Explicitly, the following was proved in [97]:

THEOREM 6.13. *In Case 1, there are A_n, B_n for $n = 1, 2, \dots$, and a function $A(\alpha)$ on $(0, \infty)$ with*

$$|A_n| \leq C, \quad |B_n| \leq Cn,$$

$$\int_0^a d\alpha |A(\alpha)| \leq C \exp(K_0|a|) \text{ so that for } \text{Re}(\kappa) > \frac{1}{2}K_0:$$

$$m(-\kappa^2) = -\kappa - \sum_{n=1}^{\infty} A_n \kappa e^{-2\kappa bn} - \sum_{n=1}^{\infty} B_n e^{-2\kappa bn} - \int_0^{\infty} d\alpha A(\alpha) e^{-2\alpha\kappa}.$$

(6.42) can be used to obtain a priori bounds on $\int_{-R}^0 d\rho(\lambda)$ as $R \rightarrow \infty$.

Now we turn to more details, and start by illustrating how to use the Riccati equation and a priori control on m_j to obtain exponentially small estimates on $m_1 - m_2$.

LEMMA 6.14. Let m_1, m_2 be two absolutely continuous functions on $[a, b]$ so that for some $Q \in L^1([a, b])$,

$$m'_j(x) = Q(x) - m_j(x)^2, \quad j = 1, 2, \quad x \in (a, b). \tag{6.44}$$

Then

$$[m_1(a) - m_2(a)] = [m_1(b) - m_2(b)] \exp \left(\int_a^b dy [m_1(y) + m_2(y)] \right).$$

As an immediate corollary, one obtains the following result (which implies Theorem 6.11):

THEOREM 6.15. Let $m_j(-\kappa^2, x)$ be functions defined for $x \in [a, b]$ and $\kappa \in K$ some region of \mathbb{C} . Suppose that for each κ in K , m_j is absolutely continuous in x and satisfies (note that V is the same for m_1 and m_2),

$$m'_j(-\kappa^2, x) = V(x) + \kappa^2 - m_j(-\kappa^2, x)^2, \quad j = 1, 2.$$

Suppose C is such that for each $x \in [a, b]$ and $\kappa \in K$,

$$|m_j(-\kappa^2, x) + \kappa| \leq C, \quad j = 1, 2, \tag{6.45}$$

then

$$|m_1(-\kappa^2, a) - m_2(-\kappa^2, a)| \leq 2C \exp[-2(b - a)[\text{Re}(\kappa) - C]]. \tag{6.46}$$

Theorem 6.15 places importance on a priori bounds of the form (6.45). Fortunately, by modifying ideas of Atkinson [13], we can obtain estimates of this form as long as $\text{Im}(\kappa)$ is bounded away from zero.

Atkinson's method allows one to estimate $|m(-\kappa^2) + \kappa|$ in two steps. We will fix some $a < b$ finite and define $m_0(-\kappa^2)$ by solving

$$m'_0(-\kappa^2, x) = V(x) + \kappa^2 - m_0(-\kappa^2, x)^2, \tag{6.47a}$$

$$m_0(-\kappa^2, a) = -\kappa \tag{6.47b}$$

and then setting

$$m_0(-\kappa^2) := m_0(-\kappa^2, 0_+). \tag{6.47c}$$

One then proves the following result.

LEMMA 6.16. There is a $C > 0$ depending only on V and a universal constant $E > 0$ so that if $\text{Re}(\kappa) \geq C$ and $\text{Im}(\kappa) \neq 0$, then

$$|m(-\kappa^2) - m_0(-\kappa^2)| \leq E \frac{|\kappa|^2}{|\text{Im}(\kappa)|} e^{-2a\text{Re}(\kappa)}. \tag{6.48}$$

In fact, one can take

$$C = \max \left(a^{-1} \ln(6), 4 \int_0^a dx |V(x)| \right), \quad E = \frac{3 \cdot 2 \cdot 12^2}{5}.$$

LEMMA 6.17. There exist constants D_1 and D_2 (depending only on a and V), so that for $\text{Re}(\kappa) > D_1$,

$$|m_0(-\kappa^2) + \kappa| \leq D_2.$$

Indeed, one can take

$$D_1 = D_2 = 2 \int_0^a dx |V(x)|.$$

These lemmas together with Theorem 6.9 yield the following explicit form of Theorem 6.11.

THEOREM 6.18. *Let V_1, V_2 be defined on $(0, b_j)$ with $b_j > a$ for $j = 1, 2$. Suppose that $V_1 = V_2$ on $[0, a]$. Pick δ so that $a + \delta \leq \min(b_1, b_2)$ and let $\eta = \sup_{0 \leq x \leq a; j=1,2} (\int_x^{x+\delta} dy |V_j(y)|)$. Then if $\text{Re}(\kappa) \geq \max(4\eta, \delta^{-1} \ln(6))$ and $\text{Im}(\kappa) \neq 0$, one obtains*

$$|m_1(-\kappa^2) - m_2(-\kappa^2)| \leq 2g(\kappa) \exp(-2a[\text{Re}(\kappa) - g(\kappa)]),$$

where

$$g(\kappa) = 2\eta + \frac{864}{5} \frac{|\kappa|^2}{|\text{Im}(\kappa)|} e^{-2\delta \text{Re}(\kappa)}.$$

REMARK 6.19. (i) To obtain Theorem 6.11, we need only note that in the region $\arg(\kappa) \in (-\frac{\pi}{2} + \varepsilon, -\varepsilon)$, $|\kappa| \geq K_0$, g is bounded.

(ii) We need not require that $\arg(\kappa) < -\varepsilon$ to obtain g bounded. It suffices, for example, that $\text{Re}(\kappa) \geq |\text{Im}(\kappa)| \geq e^{-\alpha \text{Re}(\kappa)}$ for some $\alpha < 2\delta$.

(iii) For g to be bounded, we need not require that $\arg(\kappa) > -\frac{\pi}{2} + \varepsilon$. It suffices that $|\text{Im}(\kappa)| \geq \text{Re}(\kappa) \geq \alpha \ln[|\text{Im}(\kappa)|]$ for some $\alpha > (2\delta)^{-1}$. Unfortunately, this does not include the region $\text{Im}(-\kappa^2) = c$, $\text{Re}(-\kappa^2) \rightarrow \infty$, where $\text{Re}(\kappa)$ goes to zero as $|\kappa|^{-1}$. However, as $\text{Re}(-\kappa^2) \rightarrow \infty$, we only need that $|\text{Im}(-\kappa^2)| \geq 2\alpha|\kappa|\ln(|\kappa|)$.

Next, we turn to finite b representations with no errors: Theorem 6.9 implies that if $b = \infty$ and $V \in L^1([0, \infty))$, then (6.40) holds, a Laplace transform representation for m without errors. It is, of course, of direct interest that such a formula holds, but we are especially interested in a particular consequence of it—namely, that it implies that the formula (6.38) with error holds in the region $\text{Re}(\kappa) > K_0$ with error uniformly bounded in $\text{Im}(\kappa)$; that is, one proves the following result:

THEOREM 6.20. *If $V \in L^1([0, \infty))$ and $\text{Re}(\kappa) > \|V\|_1/2$, then for all $a > 0$:*

$$\left| m(-\kappa^2) + \kappa + \int_0^a d\alpha A(\alpha) e^{-2\alpha\kappa} \right| \leq \left[\|V\|_1 + \frac{\|V\|_1^2 e^{a\|V\|_1}}{2\text{Re}(\kappa) - \|V\|_1} \right] e^{-2a\text{Re}(\kappa)}. \tag{6.49}$$

The principal goal is to prove an analog of this result in the case $b < \infty$. To do so, we will need to first prove an analog of (6.40) in case $b < \infty$ —something of interest in its own right. The idea will be to mimic the proof of Theorem 2 from [228] but use the finite b , $V^{(0)}(x) = 0, x \geq 0$ Green’s function where [228] used the infinite b Green’s function. The basic idea is simple, but the arithmetic is a bit involved.

We will start with the $h = \infty$ case. Three functions for $V^{(0)}(x) = 0, x \geq 0$ are significant. First, the kernel of the resolvent $(-\frac{d^2}{dx^2} + \kappa^2)^{-1}$ with $u(0_+) = u(b_-) = 0$ boundary conditions. By an elementary calculation (see, e.g., [228, Sect. 5]), it has the form

$$G_{h=\infty}^{(0)}(-\kappa^2, x, y) = \frac{\sinh(\kappa x_<)}{\kappa} \left[\frac{e^{-\kappa x_>} - e^{-\kappa(2b-x_>)}}{1 - e^{-2\kappa b}} \right], \tag{6.50}$$

with $x_< = \min(x, y)$, $x_> = \max(x, y)$.

The second function is

$$\psi_{h=\infty}^{(0)}(-\kappa^2, x) = \lim_{y \downarrow 0} \frac{\partial G_{h=\infty}^{(0)}}{\partial y}(-\kappa^2, x, y) = \frac{e^{-\kappa x} - e^{-\kappa(2b-x)}}{1 - e^{-2\kappa b}} \tag{6.51}$$

and finally (notice that $\psi_{h=\infty}^{(0)}(-\kappa^2, 0_+) = 1$ and $\psi_{h=\infty}^{(0)}$ satisfies the equations $-\psi'' = -\kappa^2\psi$ and $\psi(-\kappa^2, b_-) = 0$):

$$m_{h=\infty}^{(0)}(-\kappa^2) = \psi_{h=\infty}^{(0)'}(-\kappa^2, 0_+) = -\frac{\kappa + \kappa e^{-2\kappa b}}{1 - e^{-2\kappa b}}. \tag{6.52}$$

In (6.52), prime means d/dx .

Now fix $V \in C_0^\infty((0, b))$. The pair of formulas

$$\left(-\frac{d^2}{dx^2} + V + \kappa^2\right)^{-1} = \sum_{n=0}^\infty (-1)^n \left(-\frac{d^2}{dx^2} + \kappa^2\right)^{-1} \left[V \left(-\frac{d^2}{dx^2} + \kappa^2\right)^{-1}\right]^n$$

and

$$m(-\kappa^2) = \lim_{x < y; y \downarrow 0} \frac{\partial^2 G(-\kappa^2, x, y)}{\partial x \partial y}$$

yields the following expansion for the m -function of $-\frac{d^2}{dx^2} + V$ with $u(b_-) = 0$ boundary conditions.

LEMMA 6.21. *Let $V \in C_0^\infty((0, b))$, $b < \infty$. Then*

$$m(-\kappa^2) = \sum_{n=0}^\infty M_n(-\kappa^2; V), \tag{6.53}$$

where

$$M_0(-\kappa^2; V) = m_{h=\infty}^{(0)}(-\kappa^2), \tag{6.54}$$

$$M_1(-\kappa^2; V) = -\int_0^b V(x) \psi_{h=\infty}^{(0)}(-\kappa^2, x)^2 dx, \tag{6.55}$$

and for $n \geq 2$,

$$\begin{aligned} M_n(-\kappa^2; V) &= (-1)^n \int_0^b dx_1 \dots \int_0^b dx_n V(x_1) \dots V(x_n) \\ &\times \psi_{h=\infty}^{(0)}(-\kappa^2, x_1) \psi_{h=\infty}^{(0)}(-\kappa^2, x_n) \prod_{j=1}^{n-1} G_{h=\infty}^{(0)}(-\kappa^2, x_j, x_{j+1}). \end{aligned} \tag{6.56}$$

The precise region of convergence is unimportant since one can expand regions by analytic continuation. For now, we note it certainly converges in the region κ real with $\kappa^2 > \|V\|_\infty$.

Writing each term in (6.53) as a Laplace transform then yields the following result:

THEOREM 6.22 (Theorem 6.13 for $h = \infty$). *Let $b < \infty$, $h = \infty$, and $V \in L^1([0, b])$. Then for $\text{Re}(\kappa) > \|V\|_1/2$, we have that*

$$m(-\kappa^2) = -\kappa - \sum_{j=1}^\infty A_j \kappa e^{-2\kappa b_j} - \sum_{j=1}^\infty B_j e^{-2\kappa b_j} - \int_0^\infty d\alpha A(\alpha) e^{-2\alpha \kappa}, \tag{6.57}$$

where

$$A_j = 2, \quad B_j = -2j \int_0^b dx V(x), \quad j \in \mathbb{N},$$

$$|A(\alpha) - A_1(\alpha)| \leq \frac{(2\alpha + b)(2\alpha + 2b)}{2b^2} \|V\|_1^2 \exp(\alpha \|V\|_1)$$

with A_1 given by

$$A_1(\alpha) = \begin{cases} V(\alpha), & 0 \leq \alpha < b, \\ (n + 1)V(\alpha - nb) + nq((n + 1)b - \alpha), & nb \leq \alpha < (n + 1)b, \ n \in \mathbb{N}. \end{cases}$$

In particular, for all $a \in (0, b)$,

$$\int_0^a d\alpha |A(\alpha)| \leq C(b, \|V\|_1)(1 + a^2) \exp(a\|V\|_1).$$

This implies the following estimate:

COROLLARY 6.23. *If $V \in L^1([0, \infty))$ and $\operatorname{Re}(\kappa) \geq \frac{1}{2}\|V\|_1 + \varepsilon$, then for all $a \in (0, b)$, $b < \infty$, we have that*

$$\left| m(-\kappa^2) + \kappa + \int_0^a d\alpha A(\alpha)e^{-2\alpha\kappa} \right| \leq C(a, \varepsilon)e^{-2a\operatorname{Re}(\kappa)},$$

where $C(a, \varepsilon)$ depends only on a and ε (and $\|V\|_1$) but not on $\operatorname{Im}(\kappa)$.

Next, we turn to the case $h \in \mathbb{R}$. Then (6.50)–(6.52) become

$$G_h^{(0)}(-\kappa^2, x, y) = \frac{\sinh(\kappa x)}{\kappa} \psi_h^{(0)}(-\kappa^2, x), \tag{6.58}$$

$$\psi_h^{(0)}(-\kappa^2, x) = \left[\frac{e^{-\kappa x} + \zeta(h, \kappa)e^{-\kappa(2b-x)}}{1 + \zeta(h, \kappa)e^{-2b\kappa}} \right], \tag{6.59}$$

$$m_h^{(0)}(-\kappa^2) = -\kappa + 2\kappa \frac{\zeta(h, \kappa)e^{-2\kappa b}}{1 + \zeta(h, \kappa)e^{-2\kappa b}}, \tag{6.60}$$

where

$$\zeta(h, \kappa) = \frac{\kappa - h}{\kappa + h}. \tag{6.61}$$

This then leads to the following result:

THEOREM 6.24 (Theorem 6.13 for general $h \in \mathbb{R}$). *Let $b < \infty$, $|h| < \infty$, and $V \in L^1([0, b])$. Then for $\operatorname{Re}(\kappa) > \frac{1}{2}D_1[\|V\|_1 + |h| + b^{-1} + 1]$ for a suitable universal constant D_1 , (6.57) holds, where*

$$A_j = 2(-1)^j, \quad B_j = 2(-1)^{j+1}j \left[2h + \int_0^b dx V(x) \right], \tag{6.62}$$

$$|A(\alpha) - V(\alpha)| \leq \|V\|_1^2 \exp(\alpha\|V\|_1) \tag{6.63}$$

if $|\alpha| < b$, and for any $a > 0$,

$$\int_0^a d\alpha |A(\alpha)| \leq D_2(b, \|V\|_1, h) \exp(D_1 a(\|V\|_1 + |h| + b^{-1} + 1)). \tag{6.64}$$

Hence, one obtains the following estimate:

COROLLARY 6.25. *Fix $b < \infty$, $V \in L^1([0, b])$, and $|h| < \infty$. Fix $a < b$. Then there exist positive constants C and K_0 so that for all complex κ with $\operatorname{Re}(\kappa) > K_0$,*

$$\left| m(-\kappa^2) + \kappa + \int_0^a d\alpha A(\alpha)e^{-2\alpha\kappa} \right| \leq Ce^{-2a\kappa}.$$

Next we return to the relation between A and ρ and discuss a first distributional form of this relation: Our primary goal in the following is to discuss a formula which formally says that

$$A(\alpha) = -2 \int_{-\infty}^{\infty} d\rho(\lambda) \lambda^{-\frac{1}{2}} \sin(2\alpha\sqrt{\lambda}), \tag{6.65}$$

where for $\lambda \leq 0$, we define

$$\lambda^{-\frac{1}{2}} \sin(2\alpha\sqrt{\lambda}) = \begin{cases} 2\alpha, & \text{if } \lambda = 0, \\ (-\lambda)^{-\frac{1}{2}} \sinh(2\alpha\sqrt{-\lambda}), & \text{if } \lambda < 0. \end{cases}$$

In a certain sense which will become clear, the left-hand side of (6.65) should be $A(\alpha) - A(-\alpha) + \delta'(\alpha)$.

To understand (6.65) at a formal level, note the basic formulas,

$$m(-\kappa^2) = -\kappa - \int_0^{\infty} d\alpha A(\alpha) e^{-2\alpha\kappa}, \tag{6.66}$$

$$m(-\kappa^2) = \text{Re}(m(i)) + \int_{-\infty}^{\infty} d\rho(\lambda) \left[\frac{1}{\lambda + \kappa^2} - \frac{\lambda}{1 + \lambda^2} \right], \tag{6.67}$$

and

$$(\lambda + \kappa^2)^{-1} = 2 \int_0^{\infty} d\alpha \lambda^{-\frac{1}{2}} \sin(2\alpha\sqrt{\lambda}) e^{-2\alpha\kappa}, \tag{6.68}$$

which is an elementary integral if $\kappa > 0$ and $\lambda > 0$. Plug (6.68) into (6.67), formally interchange the order of integrations, and (6.66) should only hold if (6.65) does. However, a closer examination of this procedure reveals that the interchange of the order of integrations is not justified and indeed (6.65) is not true as a simple integral since, $\int_0^R d\rho(\lambda) \underset{R \rightarrow \infty}{\sim} \frac{2}{3\pi} R^{\frac{3}{2}}$, which implies that (6.65) is not absolutely convergent. We will even see later that the integral sometimes fails to be conditionally convergent.

Our primary method for understanding (6.65) is as a distributional statement, that is, it will hold when smeared in α for α in $(0, b)$. We discuss this next if $V \in L^1([0, \infty))$ or if $b < \infty$. Later it will be extended to all V (i.e., all Cases 1–4) by a limiting argument. Subsequently, we will study (6.65) as a pointwise statement, where the integral is defined as an Abelian limit.

Suppose $b < \infty$ or $b = \infty$ and $V \in L^1([0, b))$. Fix $a < b$ and $f \in C_0^\infty((0, a))$. Define

$$m_a(-\kappa^2) := -\kappa - \int_0^a d\alpha A(\alpha) e^{-2\alpha\kappa} \tag{6.69}$$

for $\text{Re}(\kappa) \geq 0$. Fix κ_0 real and let

$$g(y, \kappa_0, a) := m_a(-(\kappa_0 + iy)^2),$$

with κ_0, a as real parameters and $y \in \mathbb{R}$ a variable. As usual, define the Fourier transform by (initially for smooth functions and then by duality for tempered distributions [201, Ch. IX])

$$\hat{F}(k) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} dy e^{-iky} F(y), \quad \check{F}(k) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} dy e^{iky} F(y). \tag{6.70}$$

Then by (6.69),

$$\hat{g}(k, \kappa_0, a) = -\sqrt{2\pi} \kappa_0 \delta(k) - \sqrt{2\pi} \delta'(k) - \frac{\sqrt{2\pi}}{2} e^{-k\kappa_0} A\left(\frac{k}{2}\right) \chi_{(0, 2a)}(k). \tag{6.71}$$

Thus, since $f(0_+) = f'(0_+) = 0$, in fact, f has support away from 0 and a ,

$$\begin{aligned} \int_0^a d\alpha A(\alpha)f(\alpha) &= -\frac{2}{\sqrt{2\pi}} \int_0^a d\alpha \widehat{g}(2\alpha, \kappa_0, a)e^{2\alpha\kappa_0} f(\alpha) \\ &= -\frac{1}{\sqrt{2\pi}} \int_0^{2a} d\alpha \widehat{g}(\alpha, \kappa_0, a)e^{\alpha\kappa_0} f\left(\frac{\alpha}{2}\right) \\ &= -\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} dy g(y, \kappa_0, a)\check{F}(y, \kappa_0), \end{aligned} \tag{6.72}$$

where we have used the unitarity of $\widehat{\cdot}$ and

$$\check{F}(y, \kappa_0) = \frac{1}{\sqrt{2\pi}} \int_0^{2a} d\alpha e^{\alpha(\kappa_0+iy)} f\left(\frac{\alpha}{2}\right) = \frac{2}{\sqrt{2\pi}} \int_0^a d\alpha e^{2\alpha(\kappa_0+iy)} f(\alpha). \tag{6.73}$$

Notice that

$$|\check{F}(y, \kappa_0)| \leq C e^{2(a-\varepsilon)\kappa_0} (1 + |y|^2)^{-1} \tag{6.74}$$

since f is smooth and supported in $(0, a - \varepsilon)$ for some $\varepsilon > 0$.

By Theorem 6.20 and Corollary 6.25,

$$|m_a(-(\kappa_0 + iy)^2) - m(-(\kappa_0 + iy)^2)| \leq C e^{-2a\kappa_0} \tag{6.75}$$

for large κ_0 , uniformly in y . From (6.72), (6.74), and (6.75), one concludes the following fact:

LEMMA 6.26. *Let $f \in C_0^\infty((0, a))$ with $0 < a < b$ and $V \in L^1([0, b])$. Then*

$$\int_0^a d\alpha A(\alpha)f(\alpha) = \lim_{\kappa_0 \uparrow \infty} \left[-\frac{1}{\pi} \int_{\mathbb{R}} dy m(-(\kappa_0 + iy)^2) \int_0^a d\alpha e^{2\alpha(\kappa_0+iy)} f(\alpha) \right]. \tag{6.76}$$

As a function of y , for κ_0 fixed, the alpha integral is $O((1 + y^2)^{-N})$ for all N because f is C^∞ . Now define

$$\tilde{m}_R(-\kappa^2) = \left[c_R + \int_{\lambda \leq R} \frac{d\rho(\lambda)}{\lambda + \kappa^2} \right], \tag{6.77}$$

where c_R is chosen so that $\tilde{m}_R \xrightarrow{R \rightarrow \infty} m$. Because $\int_{\mathbb{R}} \frac{d\rho(\lambda)}{1+\lambda^2} < \infty$, the convergence is uniform in y for κ_0 fixed and sufficiently large. Thus, in (6.76) we can replace m by m_R and take a limit (first $R \rightarrow \infty$ and then $\kappa_0 \uparrow \infty$). Since $f(0_+) = 0$, the $\int dy c_R d\alpha$ -integrand is zero. Moreover, we can now interchange the $dy d\alpha$ and $d\rho(\lambda)$ integrals. The result is that

$$\begin{aligned} \int_0^a d\alpha A(\alpha)f(\alpha) &= \lim_{\kappa_0 \uparrow \infty} \lim_{R \rightarrow \infty} \int_{\lambda \leq R} d\rho(\lambda) \\ &\quad \times \left[\int_0^a d\alpha e^{2\alpha\kappa_0} f(\alpha) \left[-\frac{1}{\pi} \int_{\mathbb{R}} \frac{dy e^{2\alpha iy}}{(\kappa_0 + iy)^2 + \lambda} \right] \right]. \end{aligned} \tag{6.78}$$

In the case at hand, $d\rho$ is bounded below, say $\lambda \geq -K_0$. As long as we take $\kappa_0 > K_0$, the poles of $(\kappa_0 + iy)^2 + \lambda$ occur in the upper half-plane

$$y_{\pm} = i\kappa_0 \pm \sqrt{\lambda}.$$

Closing the contour in the upper plane, we find that if $\lambda \geq -K_0$,

$$-\frac{1}{\pi} \int_{\mathbb{R}} \frac{dy e^{2\alpha iy}}{(\kappa_0 + iy)^2 + \lambda} = -2e^{-2\alpha\kappa_0} \frac{\sin(2\alpha\sqrt{\lambda})}{\sqrt{\lambda}}.$$

Thus (6.78) becomes

$$\int_0^a d\alpha A(\alpha)f(\alpha) = -2 \lim_{\kappa_0 \uparrow \infty} \lim_{R \rightarrow \infty} \int_{\lambda \leq R} d\rho(\lambda) \left[\int_0^a d\alpha f(\alpha) \frac{\sin(2\alpha\sqrt{\lambda})}{\sqrt{\lambda}} \right].$$

κ_0 has dropped out and the α integral is bounded by $C(1 + \lambda^2)^{-1}$, so one can take the limit as $R \rightarrow \infty$ since $\int_{\mathbb{R}} \frac{d\rho(\lambda)}{1 + \lambda^2} < \infty$. One is therefore led to the following result.

THEOREM 6.27. *Let $f \in C_0^\infty((0, a))$ with $a < b$ and either $b < \infty$ or $V \in L^1([0, \infty))$ with $b = \infty$. Then*

$$\int_0^a d\alpha A(\alpha)f(\alpha) = -2 \int_{\mathbb{R}} d\rho(\lambda) \left[\int_0^a d\alpha f(\alpha) \frac{\sin(2\alpha\sqrt{\lambda})}{\sqrt{\lambda}} \right]. \tag{6.79}$$

One can strengthen this in two ways. First, one wants to allow $a > b$ if $b < \infty$. As long as A is interpreted as a distribution with δ and δ' functions at $\alpha = nb$, this is easy. One also wants to allow f to have a nonzero derivative at $\alpha = 0$. The net result is described in the next theorem:

THEOREM 6.28. *Let $f \in C_0^\infty(\mathbb{R})$ with $f(-\alpha) = -f(\alpha)$, $\alpha \in \mathbb{R}$ and either $b < \infty$ or $V \in L^1([0, \infty))$ with $b = \infty$. Then*

$$-2 \int_{\mathbb{R}} d\rho(\lambda) \left[\int_{-\infty}^{\infty} d\alpha f(\alpha) \frac{\sin(2\alpha\sqrt{\lambda})}{\sqrt{\lambda}} \right] = \int_{-\infty}^{\infty} d\alpha \tilde{A}(\alpha)f(\alpha), \tag{6.80}$$

where \tilde{A} is the distribution

$$\tilde{A}(\alpha) = \chi_{(0, \infty)}(\alpha)A(\alpha) - \chi_{(-\infty, 0)}(\alpha)A(-\alpha) + \delta'(\alpha) \tag{6.81a}$$

if $b = \infty$ and

$$\begin{aligned} \tilde{A}(\alpha) &= \chi_{(0, \infty)}(\alpha)A(\alpha) - \chi_{(-\infty, 0)}(\alpha)A(-\alpha) + \delta'(\alpha) \\ &\quad + \sum_{j=1}^{\infty} B_j [\delta(\alpha - 2bj) - \delta(\alpha + 2bj)] \\ &\quad + \sum_{j=1}^{\infty} \frac{1}{2} A_j [\delta'(\alpha - 2bj) + \delta'(\alpha + 2bj)] \end{aligned} \tag{6.81b}$$

if $b < \infty$, where A_j, B_j are h dependent and given in Theorems 6.22 and 6.24.

Next we change the subject temporarily and turn to bounds on $\int_0^{\pm R} d\rho(\lambda)$ which are of independent interest: As we will see, (6.36) implies asymptotic results on $\int_{-R}^R d\rho(\lambda)$, and (6.65) will show that $\int_{-\infty}^0 e^{b\sqrt{-\lambda}} d\rho(\lambda) < \infty$ for all $b > 0$ and more. It follows from (6.67) that

$$\text{Im}(m(ia)) = a \int_{\mathbb{R}} \frac{d\rho(\lambda)}{\lambda^2 + a^2}, \quad a > 0.$$

Thus, Everitt's result (6.36) implies that

$$\lim_{a \rightarrow \infty} a^{\frac{1}{2}} \int_{\mathbb{R}} \frac{d\rho(\lambda)}{\lambda^2 + a^2} = 2^{-\frac{1}{2}}.$$

Standard Tauberian arguments (see, e.g., in [225, Sect. III.10], which in this case shows that on even functions $R^{\frac{3}{2}} d\rho(\lambda/R) \xrightarrow{R \rightarrow \infty} (2\pi)^{-1} |\lambda|^{\frac{1}{2}} d\lambda$) then imply the following result:

THEOREM 6.29.

$$\lim_{R \rightarrow \infty} R^{-\frac{3}{2}} \int_{-R}^R d\rho(\lambda) = \frac{2}{3\pi}. \tag{6.82}$$

REMARK 6.30. (i) This holds in all cases (1–4) we consider here, including some with $\text{supp}(d\rho)$ unbounded below.

(ii) Since one can show that $\int_{-\infty}^0 d\rho$ is bounded, one can replace \int_{-R}^R by \int_0^R in (6.82).

Next, we recall the following a priori bound that follows from Lemmas 6.16 and 6.17:

LEMMA 6.31. *Let $d\rho$ be the spectral measure for a Schrödinger operator in Cases 1–4. Fix $a < b$. Then there is a constant C_a depending only on a and $\int_0^a dy |V(y)|$ so that*

$$\int_{\mathbb{R}} \frac{d\rho(\lambda)}{1 + \lambda^2} \leq C_a. \tag{6.83}$$

The goal is to bound $\int_{-\infty}^0 e^{2\alpha\sqrt{-\lambda}} d\rho(\lambda)$ for any $\alpha < b$ and to find an explicit bound in terms of $\sup_{0 \leq x \leq \alpha+1} [-V(y)]$ when that sup is finite. As a preliminary, we need the following result from the standard limit circle theory [41, Sect. 9.4].

LEMMA 6.32. *Let $b = \infty$ and let $d\rho$ be the spectral measure for some Schrödinger operator in Cases 2–4. Let $d\rho_{R,h}$ be the spectral measure for the problem with $b = R < \infty, h$ and potential equal to $V(x)$ for $x \leq R$. Then there exists $h(R)$ so that*

$$d\rho_{R,h(R)} \xrightarrow{R \rightarrow \infty} d\rho,$$

when smeared with any function f of compact support. In particular, if $f \geq 0$, then

$$\int_{\mathbb{R}} d\rho(\lambda) f(\lambda) \leq \overline{\lim}_{R \rightarrow \infty} \int_{\mathbb{R}} d\rho_{R,h(R)}(\lambda) f(\lambda).$$

This result implies that we need only obtain bounds for $b < \infty$ (where we have already proved (6.79)).

LEMMA 6.33. *If ρ_1 has support in $[-E_0, \infty)$, $E_0 > 0$, then*

$$\int_{-\infty}^0 e^{\gamma\sqrt{-\lambda}} d\rho_1(\lambda) \leq e^{\gamma\sqrt{E_0}} (1 + E_0^2) \int_{-\infty}^0 \frac{d\rho_1(\lambda)}{1 + \lambda^2}. \tag{6.84}$$

Lemmas 6.31, 6.32, and 6.33 imply the following result.

THEOREM 6.34. *Let ρ be the spectral measure for some Schrödinger operator in Cases 2–4. Let*

$$E(\alpha_0) := - \inf \left\{ \int_0^{\alpha_0+1} dx (|\varphi'_n(x)|^2 + V(x)|\varphi(x)|^2) \mid \varphi \in C_0^\infty((0, \alpha_0 + 1)), \int_0^{\alpha_0+1} dx |\varphi(x)|^2 \leq 1 \right\}.$$

Then for all $\delta > 0$ and $\alpha_0 > 0$,

$$\alpha_0 \delta \int_{-\infty}^0 e^{2(1-\delta)\alpha_0\sqrt{-\lambda}} d\rho(\lambda) \leq \left[C_1(1 + \alpha_0) + C_2(1 + E(\alpha_0)^2)e^{2(\alpha_0+1)\sqrt{E(\alpha_0)}} \right], \tag{6.85}$$

where C_1, C_2 only depend on $\int_0^1 dx |V(x)|$. In particular,

$$\int_{-\infty}^0 e^{B\sqrt{-\lambda}} d\rho(\lambda) < \infty \tag{6.86}$$

for all $B < \infty$.

As a special case, suppose $V(x) \geq -C(x + 1)^2$. Then $E(\alpha_0) \geq -C(\alpha_0 + 2)^2$ and we see that

$$\int_{-\infty}^0 e^{B\sqrt{-\lambda}} d\rho(\lambda) \leq D_1 e^{D_2 B^2}. \tag{6.87}$$

This implies the next result.

THEOREM 6.35. *If $d\rho$ is the spectral measure for a potential which satisfies*

$$V(x) \geq -Cx^2, \quad x \geq R \tag{6.88}$$

for some $R > 0, C > 0$, then for $\varepsilon > 0$ sufficiently small,

$$\int_{-\infty}^0 e^{-\varepsilon\lambda} d\rho(\lambda) < \infty. \tag{6.89}$$

If in addition $V \in L^1([0, \infty))$, then the corresponding Schrödinger operator is bounded from below and hence $d\rho$ has compact support on $(-\infty, 0]$. This fact will be useful later in the scattering-theoretic context.

The estimate (6.86), in the case of non-Dirichlet boundary conditions at $x = 0_+$, appears to be due to Marchenko [171]. Since it is a fundamental ingredient in the inverse spectral problem, it generated considerable attention; see, for instance, [70], [155], [156], [157], [161], [171], [172, Sect. 2.4]. The case of a Dirichlet boundary at $x = 0_+$ was studied in detail by Levitan [157]. These authors, in addition to studying the spectral asymptotics of $\rho(\lambda)$ as $\lambda \downarrow -\infty$, were also particularly interested in the asymptotics of $\rho(\lambda)$ as $\lambda \uparrow \infty$ and established Theorem 6.29. In the latter context, we also refer to Bennewitz [17], Harris [116], and the literature cited therein. In contrast to these activities, we were not able to find estimates of the type (6.85) (which implies (6.86)) and (6.89) in the literature.

At this point one can return to the relation between A and ρ and discuss a second distributional form of this relation which extends Theorem 6.27 to all four cases.

THEOREM 6.36. *Let $f \in C_0^\infty((0, \infty))$ and suppose $b = \infty$. Assume V satisfies (6.26) and let $d\rho$ be the associated spectral measure and A the associated A -function. Then (6.80) and (6.81) hold.*

Next we establish a third relation between A and ρ and turn to Abelian limits: For $f \in C_0^\infty(\mathbb{R})$, define for $\lambda \in \mathbb{R}$,

$$Q(f)(\lambda) = \int_{-\infty}^{\infty} d\alpha f(\alpha) \frac{\sin(2\alpha\sqrt{\lambda})}{\sqrt{\lambda}} \tag{6.90}$$

and then

$$T(f) = -2 \int_{\mathbb{R}} d\rho(\lambda) Q(f)(\lambda) \tag{6.91}$$

$$= \int_{-\infty}^{\infty} d\alpha \tilde{A}(\alpha) f(\alpha). \tag{6.92}$$

Relations (6.80) and (6.81) show that for $f \in C_0^\infty(\mathbb{R})$, the two expressions (6.91) and (6.92) define the same $T(f)$. This was proved for odd f 's but both integrals vanish for even f 's. Now one wants to use (6.91) to extend to a large class of f , but needs to exercise some care not to use (6.92), except for $f \in C_0^\infty(\mathbb{R})$.

$Q(f)$ can be defined as long as f satisfies

$$|f(\alpha)| \leq C_k e^{-k|\alpha|}, \quad \alpha \in \mathbb{R} \tag{6.93}$$

for all $k > 0$. In particular, a simple calculation shows that

$$f(\alpha) = (\pi\varepsilon)^{-\frac{1}{2}} [e^{-(\alpha-\alpha_0)^2/\varepsilon}] \text{ implies } Q(f)(\lambda) = \frac{\sin(2\alpha_0\sqrt{\lambda})}{\sqrt{\lambda}} e^{-\varepsilon\lambda}. \tag{6.94}$$

We use $f(\alpha, \alpha_0, \varepsilon)$ for the function f in (6.94).

For $\lambda \geq 0$, repeated integrations by parts show that

$$|Q(f)(\lambda)| \leq C(1 + \lambda^2)^{-1} \left[\|f\|_1 + \left\| \frac{d^3 f}{d\alpha^3} \right\|_1 \right], \tag{6.95}$$

where $\|\cdot\|_1$ represents the $L^1(\mathbb{R})$ -norm. Moreover, essentially by repeating the calculation that led to (6.94), one sees that for $\lambda \leq 0$,

$$|Q(f)(\lambda)| \leq C e^{\varepsilon|\lambda|} \|e^{+\alpha^2/\varepsilon} f\|_\infty. \tag{6.96}$$

One then concludes the following result.

LEMMA 6.37. *If $\int_{\mathbb{R}} (1 + \lambda^2)^{-1} d\rho(\lambda) < \infty$ (always true!) and $\int_{-\infty}^0 e^{-\varepsilon_0\lambda} d\rho(\lambda) < \infty$ (see Theorem 6.35 and the remark following its proof), then using (6.91), $T(\cdot)$ can be extended to functions $f \in C^3(\mathbb{R})$ that satisfy $e^{\alpha^2/\varepsilon_0} f \in L^\infty(\mathbb{R})$ for some $\varepsilon_0 > 0$ and $\frac{d^3 f}{d\alpha^3} \in L^1(\mathbb{R})$, and moreover,*

$$|T(f)| \leq C \left[\left\| \frac{d^3 f}{d\alpha^3} \right\|_1 + \|e^{\alpha^2/\varepsilon_0} f\|_\infty \right] := C \|f\|_{\varepsilon_0}. \tag{6.97}$$

Next, fix α_0 and $\varepsilon_0 > 0$ so that $\int_{-\infty}^0 e^{-\varepsilon_0\lambda} d\rho(\lambda) < \infty$. If $0 < \varepsilon < \varepsilon_0$, $f(\alpha, \alpha_0, \varepsilon)$ satisfies $\|f\|_{\varepsilon_0} < \infty$ so we can define $T(f)$. Fix $g \in C_0^\infty(\mathbb{R})$ with $g := 1$ on $(-2\alpha_0, 2\alpha_0)$. Then $\|f(\cdot, \alpha_0, \varepsilon)(1 - g)\|_{\varepsilon_0} \rightarrow 0$ as $\varepsilon \downarrow 0$. So

$$\lim_{\varepsilon \downarrow 0} T(f(\cdot, \alpha_0, \varepsilon)) = \lim_{\varepsilon \downarrow 0} T(gf(\cdot, \alpha_0, \varepsilon)).$$

For gf , we can use the expression (6.92). f is approximately $\delta(\alpha - \alpha_0)$ so standard estimates show that if α_0 is a point of Lebesgue continuity of $\tilde{A}(\alpha)$, then

$$\int_{-\infty}^{\infty} d\alpha f(\alpha, \alpha_0, \varepsilon) g(\alpha) \tilde{A}(\alpha) \xrightarrow{\varepsilon \downarrow 0} \tilde{A}(\alpha_0).$$

Since $A - g$ is continuous, points of Lebesgue continuity of A exactly are points of Lebesgue continuity of V . Thus, one obtains the following theorem.

THEOREM 6.38. *Suppose either $b < \infty$ and $V \in L^1([0, b])$ or $b = \infty$, and then either $V \in L^1([0, \infty))$ or $V \in L^1([0, a])$ for all $a > 0$ and*

$$V(x) \geq -Cx^2, \quad x \geq R$$

for some $R > 0, C > 0$. Let $\alpha_0 \in (0, b)$ and be a point of Lebesgue continuity of V . Then

$$A(\alpha_0) = -2 \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}} d\rho(\lambda) e^{-\varepsilon\lambda} \frac{\sin(2\alpha_0\sqrt{\lambda})}{\sqrt{\lambda}}. \tag{6.98}$$

Finally, we specialize (6.98) to the scattering-theoretic setting. Assuming $V \in L^1([0, \infty); (1+x) dx)$, the corresponding Jost solution $f(z, x)$ is defined by

$$f(z, x) = e^{i\sqrt{z}x} - \int_x^\infty dx' \frac{\sin(\sqrt{z}(x-x'))}{\sqrt{z}} V(x')f(z, x'), \quad \text{Im}(\sqrt{z}) \geq 0, \tag{6.99}$$

and the corresponding Jost function, $F(\sqrt{z})$, and scattering matrix, $S(\lambda), \lambda \geq 0$, then read

$$F(\sqrt{z}) = f(z, 0_+), \tag{6.100}$$

$$S(\lambda) = \overline{F(\sqrt{\lambda})} / F(\sqrt{\lambda}), \quad \lambda \geq 0. \tag{6.101}$$

The spectrum of the Schrödinger operator H in $L^2([0, \infty))$ associated with the differential expression $-\frac{d^2}{dx^2} + V(x)$ and a Dirichlet boundary condition at $x = 0_+$ (cf. (6.128) for precise details) is simple and of the type

$$\sigma(H) = \{-\kappa_j^2 < 0\}_{j \in J} \cup [0, \infty).$$

Here J is a finite (possibly empty) index set, $\kappa_j > 0, j \in J$, and the essential spectrum is purely absolutely continuous. The corresponding spectral measure explicitly reads

$$d\rho(\lambda) = \begin{cases} \pi^{-1} |F(\sqrt{\lambda})|^{-2} \sqrt{\lambda} d\lambda, & \lambda \geq 0, \\ \sum_{j \in J} c_j \delta(\lambda + \kappa_j^2) d\lambda, & \lambda < 0, \end{cases} \tag{6.102}$$

where

$$c_j = \|\varphi(-\kappa_j^2, \cdot)\|_2^{-2}, \quad j \in J \tag{6.103}$$

are the norming constants associated with the eigenvalues $\lambda_j = -\kappa_j^2 < 0$. Here the regular solution $\varphi(z, x)$ of $-\psi''(z, x) + [V(x) - z]\psi(z, x) = 0$ (defined by $\varphi(z, 0_+) = 0, \varphi'(z, 0_+) = 1$) and $f(z, x)$ in (6.99) are linearly dependent precisely for $z = -\kappa_j^2, j \in J$.

Since

$$|F(\sqrt{\lambda})| = \prod_{j \in J} \left(1 + \frac{\kappa_j^2}{\lambda}\right) \exp\left(\frac{1}{\pi} P \int_0^\infty \frac{d\lambda' \delta(\lambda')}{\lambda - \lambda'}\right), \quad \lambda \geq 0,$$

where $P \int_0^\infty$ denotes the principal value symbol and $\delta(\lambda)$ the corresponding scattering phase shift, that is, $S(\lambda) = \exp(2i\delta(\lambda)), \delta(\lambda) \xrightarrow{\lambda \uparrow \infty} 0$, the scattering data

$$\{-\kappa_j^2, c_j\}_{j \in J} \cup \{S(\lambda)\}_{\lambda \geq 0}$$

uniquely determine the spectral measure (6.102) and hence $A(\alpha)$. Inserting (6.102) into (6.98) then yields the following expression for $A(\alpha)$ in terms of scattering data.

THEOREM 6.39. *Suppose that $V \in L^1([0, \infty); (1 + x) dx)$. Then*

$$\begin{aligned}
 A(\alpha) &= -2 \sum_{j \in J} c_j \kappa_j^{-1} \sinh(2\alpha \kappa_j) \\
 &\quad - 2\pi^{-1} \lim_{\varepsilon \downarrow 0} \int_0^\infty d\lambda e^{-\varepsilon \lambda} |F(\sqrt{\lambda})|^{-2} \sin(2\alpha \sqrt{\lambda})
 \end{aligned}
 \tag{6.104}$$

at points $\alpha \geq 0$ of Lebesgue continuity of V .

REMARK 6.40. In great generality $|F(k)| \rightarrow 1$ as $k \rightarrow \infty$, so one cannot take the limit in ε inside the integral in (6.104). In general, though, one can replace $|F(\sqrt{\lambda})|^{-2}$ by $(|F(\sqrt{\lambda})|^{-2} - 1) \equiv X(\lambda)$ and ask if one can take a limit there. As long as V is $C^2((0, \infty))$ with $V'' \in L^1([0, \infty))$, it is not hard to see that as $\lambda \rightarrow \infty$

$$X(\lambda) = -\frac{V(0)}{2\lambda} + O(\lambda^{-2}).$$

Thus, if $V(0) = 0$, then

$$\begin{aligned}
 A(\alpha) &= -2 \sum_{j \in J} c_j \kappa_j^{-1} \sinh(2\alpha \kappa_j) \\
 &\quad - 2\pi^{-1} \int_0^\infty d\lambda (|F(\sqrt{\lambda})|^{-2} - 1) \sin(2\alpha \sqrt{\lambda}).
 \end{aligned}
 \tag{6.105}$$

The integral in (6.105) is only conditionally convergent if $V(0) \neq 0$.

We note that in the present case, where $V \in L^1([0, \infty); (1 + x) dx)$, the representation (6.40) of the m -function in terms of the $A(\alpha)$ -amplitude was considered in a paper by Ramm [195] (see also [196, pp. 288–291]).

We add a few more remarks in the scattering-theoretic setting. Assuming $V \in L^1([0, \infty); (1 + x) dx)$, one sees that

$$|F(k)| \underset{k \uparrow \infty}{=} 1 + o(k^{-1})
 \tag{6.106}$$

(cf. [31, eq. II.4.13] and apply the Riemann–Lebesgue lemma; actually, one only needs $V \in L^1([0, \infty))$ for the asymptotic results on $F(k)$ as $k \uparrow \infty$ but we will ignore this refinement in the following). A comparison of (6.106) and (6.104) then clearly demonstrates the necessity of an Abelian limit in (6.104). Even replacing $d\rho$ in (6.98) by $d\sigma = d\rho - d\rho^{(0)}$, that is, effectively replacing $|F(\sqrt{\lambda})|^{-2}$ by $[|F(\sqrt{\lambda})|^{-2} - 1]$ in (6.104), still does not necessarily produce an absolutely convergent integral in (6.104).

The latter situation changes upon increasing the smoothness properties of V since, for example, assuming $V \in L^1([0, \infty); (1 + x) dx)$, $V' \in L^1([0, \infty))$, yields

$$|F(k)|^{-2} - 1 \underset{k \uparrow \infty}{=} O(k^{-2})$$

as detailed high-energy considerations (cf. [87]) reveal. Indeed, if $V'' \in L^1([0, \infty))$, then the integral one gets is absolutely convergent if and only if $V(0) = 0$.

As a final issue related to the representation (6.65), we discuss the issue of bounds on A when $|V(x)| \leq Cx^2$. One has two general bounds on A : the estimate of [228] (see (6.39)),

$$|A(\alpha) - V(\alpha)| \leq \left[\int_0^\alpha dy |V(y)| \right]^2 \exp \left[\alpha \int_0^\alpha dy |V(y)| \right],
 \tag{6.107}$$

and the estimate in Theorem 6.42,

$$|A(\alpha)| \leq \frac{\gamma(\alpha)}{\alpha} I_1(2\alpha\gamma(\alpha)), \tag{6.108}$$

where $|\gamma(\alpha)| = \sup_{0 \leq x \leq \alpha} |V(x)|^{1/2}$ and $I_1(\cdot)$ is the modified Bessel function of order one (cf., e.g., [1], Ch. 9). Since ([1], p. 375)

$$0 \leq I_1(x) \leq e^x, \quad x \geq 0, \tag{6.109}$$

one concludes that

$$|A(\alpha)| \leq \sqrt{C} e^{2\sqrt{C}} \alpha^2 \tag{6.110}$$

if $|V(x)| \leq Cx^2$.

We continue with a discussion of the case of constant V :

EXAMPLE 6.41. If $b = \infty$ and $V(x) = V_0, x \geq 0$, then if $V_0 > 0$,

$$A(\alpha) = \frac{V_0^{1/2}}{\alpha} J_1(2\alpha V_0^{1/2}), \tag{6.111}$$

where $J_1(\cdot)$ is the Bessel function of order one (cf., e.g., [1], Ch. 9); and if $V_0 < 0$,

$$A(\alpha) = \frac{(-V_0)^{1/2}}{\alpha} I_1(2\alpha(-V_0)^{1/2}), \tag{6.112}$$

with $I_1(\cdot)$ the corresponding modified Bessel function.

This example is important because of the following monotonicity property:

THEOREM 6.42. Let $|V_1(x)| \leq -V_2(x)$ on $[0, a]$ with $a \leq \min(b_1, b_2)$. Then,

$$|A_1(\alpha)| \leq -A_2(\alpha) \text{ on } [0, a].$$

In particular, for any V satisfying $\sup_{0 \leq x \leq \alpha} |V(x)| < \infty$, one obtains

$$|A(\alpha)| \leq \frac{\gamma(\alpha)}{\alpha} I_1(2\alpha\gamma(\alpha)), \tag{6.113}$$

where

$$\gamma(\alpha) = \sup_{0 \leq x \leq \alpha} (|V(x)|^{1/2}). \tag{6.114}$$

In particular, (6.109) implies

$$|A(\alpha)| \leq \alpha^{-1} \gamma(\alpha) e^{2\alpha\gamma(\alpha)}, \tag{6.115}$$

and if V is bounded,

$$|A(\alpha)| \leq \alpha^{-1} \|V\|_\infty^{1/2} \exp(2\alpha\|V\|_\infty^{1/2}). \tag{6.116}$$

For α small, (6.115) is a poor estimate and one should use (6.107) which implies that $|A(\alpha)| \leq \|V\|_\infty + \alpha^2 \|V\|_\infty^2 e^{\alpha^2 \|V\|_\infty}$.

This lets one prove the following result:

THEOREM 6.43. Let $h = \infty$ and $V \in L^\infty([0, \infty))$. Suppose $\kappa^2 > \|V\|_\infty$. Then

$$m(-\kappa^2) = -\kappa - \int_0^\infty d\alpha A(\alpha) e^{-2\alpha\kappa} \tag{6.117}$$

(with an absolutely convergent integral and no error term).

REMARK 6.44. We recall (cf. (6.40)) that the representation (6.117) also holds with $A \in L^1([0, a])$ for all $a > 0$ and as an absolutely convergent integral for $\text{Re}(\kappa) > \|V\|_1/2$ if $V \in L^1([0, \infty))$. This fact will be used below.

The case of Bargmann potentials has been worked out in [97, Sect. 11] and explicit formulas for the A -function have been obtained.

We end this survey of [97] and [228] by recalling the major thrust of [228]—the connection between A and the inverse spectral theory. Namely, there is an $A(\alpha, x)$ function associated to $m(z, x)$ by

$$m(-\kappa^2, x) = -\kappa - \int_0^a d\alpha A(\alpha, x)e^{-2\alpha\kappa} + \tilde{O}(e^{-2\alpha\kappa}) \tag{6.118}$$

for $a < b - x$. This, of course, follows from Theorem 6.7 by translating the origin. The point is that A satisfies the simple differential equation in distributional sense

$$\frac{\partial A}{\partial x}(\alpha, x) = \frac{\partial A}{\partial \alpha}(\alpha, x) + \int_0^a d\beta A(\alpha - \beta, x)A(\beta, x). \tag{6.119}$$

This is proved in [228] for $V \in L^1([0, a])$ (and some other V 's) and so holds in the generality of [97] since Theorem 6.11 implies $A(\alpha, x)$ for $\alpha + x \leq a$ is only a function of $V(y)$ for $y \in [0, a]$.

Moreover, by (6.39), one has

$$\lim_{\alpha \downarrow 0} |A(\alpha, x) - V(\alpha + x)| = 0 \tag{6.120}$$

uniformly in x on compact subsets of the real line, so by the uniqueness theorem for solutions of (6.119) (proved in [228]), A on $[0, a]$ determines V on $[0, a]$.

In the limit circle case, there is an additional issue to discuss. Namely, that $m(z, x = 0)$ determines the boundary condition at ∞ . This is because, as we just discussed, m determines A which determines V on $[0, \infty)$. $m(z, 0_+)$ and V determine $m(z, x)$ by the Riccati equation. Once we know m , we can recover $u(i, x) = \exp(\int_0^x m(i, y) dy)$, and so the particular solution that defined the boundary condition at ∞ .

Thus, the inverse spectral theory aspects of the framework easily extend to the general case of potentials considered in [97].

To turn this into an inverse spectral approach alternative to and fully equivalent to that of Gel'fand and Levitan, one needs to settle necessary and sufficient conditions for solvability of the differential equation (6.119) in terms of an initial condition $A(\alpha, 0_+) = A_0(\alpha)$, that is, in terms of properties of A_0 . This final step was accomplished by Remling [205] and we briefly describe its major elements next.

Remling's first result is of local nature and determines a necessary and sufficient condition on A to be the A -function of a potential V . Assuming $V \in L^1([0, b])$ for all $b > 0$, he introduces the set

$$\mathcal{A}_b = \{A \in L^1([0, b]) \mid A \text{ real-valued, } I + \mathcal{K}_A > 0\}, \tag{6.121}$$

where

$$(\mathcal{K}_A f)(\alpha) = \int_0^b d\beta K(\alpha, \beta)f(\beta), \quad \alpha \in [0, b], f \in L^2([0, b]),$$

$$K(\alpha, \beta) = [\phi(\alpha - \beta) - \phi(\alpha + \beta)]/2, \quad \phi(\alpha) = \int_0^{|\alpha|/2} d\gamma A(\gamma), \quad \alpha, \beta \in [0, b].$$

Based on his reformulation of the Gel'fand–Levitan approach in terms of de Branges spaces in [204], Remling obtained the following characterization of A -functions:

THEOREM 6.45. \mathcal{A}_b is precisely the set of A -functions in

$$m(-\kappa^2) = -\kappa - \int_0^a d\alpha A(\alpha)e^{-2\alpha\kappa} + \tilde{O}(e^{-2\alpha\kappa}) \text{ for all } a < b.$$

Equivalently, given $A_0 \in L^1([0, b])$, there exists a potential $V \in L^1([0, b])$ such that A_0 is the A -function of V if and only if $A_0 \in \mathcal{A}_b$.

(We recall that all potentials V in this survey are assumed to be real-valued.)

As a second result, Remling also proved in [205] that the positivity condition in (6.121) is necessary and sufficient to solve (6.119) on $\Delta_b = \{(\alpha, x) \in \mathbb{R}^2 \mid \alpha \in [0, b - x], x \in [0, b]\}$ given an initial condition $A(\cdot, 0_+) = A_0 \in L^1([0, b])$. The potential V can then be read off from

$$V(x) = A(0_+, x) \text{ for } x \in [0, b]. \tag{6.122}$$

Necessity of this positivity condition had been established independently by Keel and Simon (unpublished). To make this precise, it pays to slightly rewrite (6.119) as follows: Let

$$B(\alpha, x) = A(\alpha - x, x) - A_0(\alpha), \quad (\alpha, x) \in \tilde{\Delta}_b, \tag{6.123}$$

where

$$\tilde{\Delta}_b = \{(\alpha, x) \in \mathbb{R}^2 \mid 0 \leq x \leq \alpha \leq b\}.$$

Then (6.119) together with the initial condition $A(\cdot, 0_+) = A_0 \in L^1([0, b])$ becomes

$$\begin{aligned} B(\alpha, x) &= \int_0^x dy \int_0^{\alpha-y} d\beta [B(y + \beta, y) + A_0(y + \beta)][B(\alpha - \beta, y) + A_0(\alpha - \beta)] \\ B(\alpha, 0_+) &= 0, \quad (\alpha, x) \in \tilde{\Delta}_b. \end{aligned} \tag{6.124}$$

If A is actually the A -function of a potential, then $B \in C(\tilde{\Delta}_b)$ by [228, Theorem 2.1]. Remling [205] then proves the following result:

THEOREM 6.46. Suppose $A_0 \in L^1([0, b])$. Then (6.124) has a solution $B \in C(\tilde{\Delta}_b)$ if and only if $A_0 \in \mathcal{A}_b$.

This brings Simon’s inverse approach to full circle and one can envision the following two scenarios. First, Simon’s inverse A -function approach, as complemented by Remling [205]:

$$\begin{aligned} A_0 \in \mathcal{A}_b &\xrightarrow{\text{by (6.124)}} B(\alpha, x), \quad (\alpha, x) \in \tilde{\Delta}_b \xrightarrow{\text{by (6.123)}} A(\alpha, x), \quad (\alpha, x) \in \Delta_b \\ &\xrightarrow{\text{by (6.122)}} V = A(0_+, \cdot) \in L^1([0, b]). \end{aligned} \tag{6.125}$$

Second, denote by \mathcal{R} the set of spectral functions ρ associated with self-adjoint half-line Schrödinger operators with a Dirichlet boundary condition at $x = 0$ and a self-adjoint boundary condition (6.29) at infinity (if any, i.e., if (6.24) is in the limit circle case at ∞). For characterizations of \mathcal{R} we refer, for instance, to [161], [159, Ch. 2], [172, Ch. 2]. Then combining (6.125) with (6.80) yields Simon’s inverse spectral approach as an alternative to that by Gel’fand and Levitan:

$$\begin{aligned}
 \rho \in \mathcal{R} &\xrightarrow{\text{by (6.80)}} A_0 \in \mathcal{A}_b \text{ for all } b > 0 \\
 &\xrightarrow{\text{by (6.124)}} B(\alpha, x), (\alpha, x) \in \tilde{\Delta}_b \text{ for all } b > 0 \\
 &\xrightarrow{\text{by (6.123)}} A(\alpha, x), (\alpha, x) \in \Delta_b \text{ for all } b > 0 \\
 &\xrightarrow{\text{by (6.122)}} V = A(0_+, \cdot) \in L^1([0, b]) \text{ for all } b > 0.
 \end{aligned}
 \tag{6.126}$$

More recent references: Local solvability and a necessary condition for global solvability of the A -equation (6.119) were recently discussed by Zhang [252], [253]. Connections between the A -amplitude and the scattering transform for Schrödinger operators on the real line have been discussed by Hitrik [117].

★ ★ ★

Next we briefly quote the main results by Ramm and Simon [200]. The primary goal in this paper was to study A as an interesting object in its own right and, in particular, using ideas implicit in Ramm [195] to obtain detailed information on the behavior of $A(\alpha)$ as $\alpha \rightarrow \infty$ when V decays sufficiently fast as $x \rightarrow \infty$. Indeed, for potentials decaying rapidly enough, Ramm [195] stated the representation (6.117) (actually, (6.40)), but no proof was given (nor was there any connection of the function A to the inverse problem for V). In [195] the inverse problem of finding the potential from the knowledge of the m -function has been solved for short-range potentials. A more detailed discussion of the result in [195] can be found in [198], [199].

Throughout [200] it is assumed that

$$\int_0^\infty (1+x) dx |V(x)| < \infty
 \tag{6.127}$$

and the Dirichlet-type Schrödinger operator H in $L^2([0, \infty))$ defined by

$$\begin{aligned}
 Hf = -f'' + Vf, \quad f \in \text{dom}(H) = \{u \in L^2([0, \infty)) \mid u, u' \in AC_{\text{loc}}([0, b]) \\
 \text{for all } b > 0; u(0_+) = 0; (-g'' + Vg) \in L^2([0, \infty))\}
 \end{aligned}
 \tag{6.128}$$

is considered.

More generally, for $n \in \mathbb{N}_0$, $B \leq 0$ and $\ell \geq 0$, the space $C_n^{B, \ell}$ of all functions q with $n - 1$ classical derivatives and $q^{(n)} \in L^1([0, \infty))$ so that

$$\int_0^\infty (1+x)^\ell e^{-Bx} dx |q^{(j)}(x)| < \infty$$

for $j = 0, 1, \dots, n$. Thus, (6.127) says $V \in C_{n=0}^{B=0, \ell=1}$.

Under condition (6.127), general principles (see, e.g., [172, Ch. 3]) imply that for all $\kappa \in \mathbb{C}$ with $\text{Re}(\kappa) \geq 0$, there is a unique solution $F(\kappa, x)$ of $-f'' + Vf = -\kappa^2 f$ normalized so that $F(\kappa, x) = e^{-\kappa x}(1 + o(1))$ as $x \rightarrow \infty$. We set $F(\kappa) := F(\kappa, 0_+)$. Except for the change of variables $\kappa = -ik$, $F(\kappa, x)$ and $F(\kappa)$ are the standard Jost solution and Jost function. Both $F(\kappa, x)$ and $F(\kappa)$ are analytic with respect to κ in $\{\kappa \in \mathbb{C} \mid \text{Re}(\kappa) > 0\}$. If $V \in C_n^{B, \ell}$ for any n, ℓ and $B < 0$, then $F(\kappa, x)$ and $F(\kappa)$ have analytic continuations into the region $\text{Re}(\kappa) > B/2$.

The following is easy to see and well known (cf. [172, Ch. 3]):

- (1) The zeros of F in $\{\kappa \in \mathbb{C} \mid \operatorname{Re}(\kappa) > 0\}$ occur precisely at those points κ_j with $-\kappa_j^2$ an eigenvalue of the operator H and each such zero is simple.
- (2) F has no zeros in $\{\kappa \in \mathbb{C} \mid \operatorname{Re}(\kappa) = 0, \kappa \neq 0\}$.
- (3) If $F(0) = 0$ and $V \in C_{n=0}^{B=0, \ell=2}$, then F is C^1 and $F'(0) \neq 0$. If $F(0) = 0$, we say that H has a *zero energy resonance*.

If F can be analytically continued to $\{\kappa \in \mathbb{C} \mid \operatorname{Re}(\kappa) > B/2\}$ for $B < 0$, then zeros of F in $\{\kappa \in \mathbb{C} \mid \operatorname{Re}(\kappa) < 0\}$ are called *resonances* of H . They occur in complex conjugate pairs (since F is real on the real axis). If $F'(\kappa_0) \neq 0$ at a zero κ_0 , we say that κ_0 is a simple resonance. Resonances need not be simple if $\operatorname{Re}(\kappa_0) < 0$ although they are generically simple.

The result stated in [195] can be phrased as follows:

THEOREM 6.47. *Suppose that V satisfies (6.127) (i.e., it lies in $C_{n=0}^{B=0, \ell=1}$) and that H does not have a zero energy resonance. Let $\{-\kappa_j^2\}_{j=1}^J$ be the negative eigenvalues of H with $\kappa_j > 0$. Then*

$$A(\alpha) = \sum_{j=1}^J B_j e^{2\alpha\kappa_j} + g(\alpha), \tag{6.129}$$

where $g \in L^1([0, \infty))$. In particular, if H has no eigenvalues and no zero energy resonance (e.g., if $V \geq 0$), then $A \in L^1([0, \infty))$.

REMARK 6.48. (i) The result stated in [195] assumes implicitly that there is no zero energy resonance. Details can be found in [198].

(ii) If $A \in L^1([0, \infty))$, then the representation (6.117) (resp., (6.40)) can be analytically continued to the entire region $\operatorname{Re}(\kappa) \geq 0$.

(iii) If u_j is the eigenfunction of H corresponding to the eigenvalue $-\kappa_j^2$, normalized by $\|u_j\|_2 = 1$, then

$$B_j = -\frac{|u'_j(0_+)|^2}{\kappa_j}.$$

This follows from (6.104) and the fact that

$$d\rho(\lambda) \upharpoonright (-\infty, 0) = \sum_{j=1}^J |u'_j(0_+)|^2 \delta(\lambda + \kappa_j^2) d\lambda.$$

To handle zero energy resonances of H , one needs an extra two powers of decay (just as (6.28) says more or less that $|V(x)|$ is bounded by $O(x^{-2-\epsilon})$, the condition in the next theorem says that $|V(x)|$ is more or less $O(x^{-4-\epsilon})$):

THEOREM 6.49. *Let $V \in C_{n=0}^{B=0, \ell=3}$. Suppose that H has a zero energy resonance and negative eigenvalues at $\{-\kappa_j^2\}_{j=1}^J$ with $\kappa_j > 0$. Then*

$$A(\alpha) = B_0 + \sum_{j=1}^J B_j e^{2\alpha\kappa_j} + g(\alpha), \tag{6.130}$$

where $g \in L^1([0, \infty))$.

These results are special cases of the following theorem:

THEOREM 6.50. *Let $V \in C_n^{B=0, \ell}$ where $\ell \geq 1$ and, if H has a zero energy resonance, then $\ell \geq 3$. Then (6.129) (resp., (6.130)) if there is a zero energy resonance holds, where $g \in C_n^{B=0, \ell-1}$ (resp., $C_n^{B=0, \ell-3}$).*

Finally, for $B < 0$, the following result was proved in [200]:

THEOREM 6.51. *Let $V \in C_n^{B, \ell=0}$ with $B < 0$. Let $\tilde{B} \in (B, 0)$ and let $\{-\kappa_j^2\}_{j=1}^J$ with $\kappa_j > 0$ be the negative eigenvalues of H , $\{\lambda_j\}_{j=1}^M$ with $\lambda_j \leq 0$ the real resonances (a.k.a. anti-bound states) of H , and $\{\mu_j \pm i\nu_j\}_{j=1}^N$ the complex resonances of H with $\tilde{B} \leq \mu_j < 0$ and $\nu_j > 0$. Suppose each resonance is simple. Then for suitable $\{B_j\}_{j=1}^J, \{C_j\}_{j=1}^M, \{D_j\}_{j=1}^N, \{\theta_j\}_{j=1}^N$, one obtains*

$$A(\alpha) = \sum_{j=1}^J B_j e^{2\alpha\kappa_j} + \sum_{j=1}^M C_j e^{2\alpha\lambda_j} + \sum_{j=1}^N D_j e^{2\mu_j\alpha} \cos(2\nu_j\alpha + \theta_j) + \tilde{g}(\alpha),$$

where $\tilde{g} \in C_n^{\tilde{B}, \ell=0}$. In particular, if H has no negative eigenvalues, the rate of decay of A is determined by the resonance with the least negative value of λ or μ .

★ ★ ★

We conclude this section with a brief look at the principal results in [98].

Let $H_j = -\frac{d^2}{dx^2} + V_j, V_j \in L^1([0, b])$ for all $b > 0, V_j$ real-valued, $j = 1, 2$, be two self-adjoint operators in $L^2([0, \infty))$ with a Dirichlet boundary condition at $x = 0_+$. Let $m_j(z), z \in \mathbb{C} \setminus \mathbb{R}$ be the Weyl–Titchmarsh m -functions associated with $H_j, j = 1, 2$. The main purpose of [98] was to provide a short proof of the following local uniqueness theorem in the spectral theory of one-dimensional Schrödinger operators, originally obtained by Simon [228], but under slightly more general assumptions than in [228].

We summarize the principal results of [98] as follows:

THEOREM 6.52. (i) *Let $a > 0, 0 < \varepsilon < \pi/2$ and suppose that*

$$|m_1(z) - m_2(z)|_{|z| \rightarrow \infty} = O(e^{-2\text{Im}(z^{1/2})a})$$

along the ray $\arg(z) = \pi - \varepsilon$. Then

$$V_1(x) = V_2(x) \text{ for a.e. } x \in [0, a].$$

(ii) *Conversely, let $\arg(z) \in (\varepsilon, \pi - \varepsilon)$ for some $0 < \varepsilon < \pi$ and suppose $a > 0$. If*

$$V_1(x) = V_2(x) \text{ for a.e. } x \in [0, a],$$

then

$$|m_1(z) - m_2(z)|_{|z| \rightarrow \infty} = O(e^{-2\text{Im}(z^{1/2})a}). \tag{6.131}$$

(iii) *In addition, suppose that $H_j, j = 1, 2$, are bounded from below. Then (6.131) extends to all $\arg(z) \in (\varepsilon, \pi]$.*

COROLLARY 6.53. *Let $0 < \varepsilon < \pi/2$ and suppose that for all $a > 0$,*

$$|m_1(z) - m_2(z)|_{|z| \rightarrow \infty} = O(e^{-2\text{Im}(z^{1/2})a})$$

along the ray $\arg(z) = \pi - \varepsilon$. Then

$$V_1(x) = V_2(x) \text{ for a.e. } x \in [0, \infty).$$

Theorem 6.52 and Corollary 6.53 follow by combining some of the Riccati equation methods in [97] with properties of transformation operators (cf. [172, Sect. 3.1]) and a uniqueness theorem for finite Laplace transforms [228, Lemma A.2.1].

In particular, Corollary 6.53 represents a considerable strengthening of the original Borg–Marchenko uniqueness result [25], [170], [171]:

THEOREM 6.54. *Suppose*

$$m_1(z) = m_2(z), \quad z \in \mathbb{C} \setminus \mathbb{R},$$

then

$$V_1(x) = V_2(x) \text{ for a.e. } x \in [0, \infty).$$

REMARK 6.55. (i) Marchenko [170] first published Theorem 6.54 in 1950. His extensive treatise on spectral theory of one-dimensional Schrödinger operators [171], repeating the proof of his uniqueness theorem, then appeared in 1952, which also marked the appearance of Borg’s proof of the uniqueness theorem [25] (apparently, based on his lecture at the 11th Scandinavian Congress of Mathematicians held at Trondheim, Norway, in 1949).

We emphasize that Borg and Marchenko also treat the general case of non-Dirichlet boundary conditions at $x = 0_+$ (see also item (vi) below). Moreover, Marchenko simultaneously discussed the half-line and finite interval case (cf. item (vii) below).

(ii) As pointed out by Levitan [159] in his Notes to Chapter 2, Borg and Marchenko were actually preceded by Tikhonov [245] in 1949, who proved a special case of Theorem 6.54 in connection with the string equation (and hence under certain additional hypotheses on V_j).

(iii) Since Weyl–Titchmarsh functions m are uniquely related to the spectral measure $d\rho$ of H by the standard Herglotz representation theorem, (6.33), Theorem 6.54 is equivalent to the following statement: Denote by $d\rho_j$ the spectral measures of H_j , $j = 1, 2$. Then

$$d\rho_1 = d\rho_2 \text{ implies } V_1 = V_2 \text{ a.e. on } [0, \infty).$$

In fact, Marchenko took the spectral measures $d\rho_j$ as his point of departure while Borg focused on the Weyl–Titchmarsh functions m_j .

(iv) The Borg–Marchenko uniqueness result, Theorem 6.54 (but not the strengthened version, Corollary 6.53), under the additional condition of short-range potentials V_j satisfying $V_j \in L^1([0, \infty); (1+x) dx)$, $j = 1, 2$, can also be proved using Property C, a device used by Ramm [197], [198] in a variety of uniqueness results.

(v) The ray $\arg(z) = \pi - \varepsilon$, $0 < \varepsilon < \pi/2$ chosen in Theorem 6.52(i) and Corollary 6.53 is of no particular importance. A limit taken along any non-self-intersecting curve \mathcal{C} going to infinity in the sector $\arg(z) \in ((\pi/2) + \varepsilon, \pi - \varepsilon)$ will do since we can apply the Phragmén–Lindelöf principle ([191, Part III, Sect. 6.5]) to the region enclosed by \mathcal{C} and its complex conjugate $\overline{\mathcal{C}}$.

(vi) For simplicity of exposition, we only discussed the Dirichlet boundary condition

$$u(0_+) = 0$$

in the Schrödinger operator H . Everything extends to the the general boundary condition

$$u'(0_+) + hu(0_+) = 0, \quad h \in \mathbb{R},$$

and we refer to [98, Remark 2.9] for details.

(*vii*) Similarly, the case of a finite interval problem on $[0, b]$, $b \in (0, \infty)$, instead of the half-line $[0, \infty)$ in Theorem 6.52 (*i*), with $0 < a < b$, and a self-adjoint boundary condition at $x = b_-$ of the type

$$u'(b_-) + h_b u(b_-) = 0, \quad h_b \in \mathbb{R},$$

can be discussed (cf. [98, Remark 2.10]).

While we have separately described a few extensions in Remarks 6.55 (*v*)–(*vii*), it is clear that they can all be combined at once.

Without going into further details, we also mention that [98] contains the analog of the local Borg–Marchenko uniqueness result, Theorem 6.52 (*i*) for Schrödinger operators on the real line. In addition, the case of half-line Jacobi operators and half-line matrix-valued Schrödinger operators was dealt with in [98].

More recent references: An even shorter proof of Theorem 6.52 (*i*), close in spirit to Borg’s original paper [25], was found by Bennewitz [18]. Still other proofs were presented by Horváth [124] and Knudsen [139]. Various local and global uniqueness results for matrix-valued Schrödinger, Dirac-type, and Jacobi operators were considered in [84]. The analog of the local Borg–Marchenko theorem for certain Dirac-type systems was also studied by A. Sakhnovich [212]. The matrix-valued weighted Sturm–Liouville case has further been studied by Andersson [9]. He also studied uniqueness questions for certain scalar higher-order differential operators in [10]. A local Borg–Marchenko theorem for complex-valued potentials has been proved by Brown, Peacock, and Weikard [26]. The case of semi-infinite Jacobi operators with complex-valued coefficients was studied by Weikard [250]. A (global) uniqueness result for trees in terms of the (generalized) Dirichlet-to-Neumann map was found by Brown and Weikard [27].

Acknowledgments. I’m grateful to Mark Ashbaugh and the anonymous referee for a critical reading of this manuscript. I am also indebted to T. Tombrello for the kind invitation to visit the California Institute of Technology, Pasadena, CA, during the fall semester of 2005, where parts of this paper were written. The great hospitality of the Department of Mathematics at Caltech is gratefully acknowledged. Moreover, I thank the Research Council and the Office of Research of the University of Missouri-Columbia for granting me a research leave for the academic year 2005–2006.

References

- [1] M. Abramowitz and I. A. Stegun (eds.), *Handbook of Mathematical Functions*, 9th printing, Dover, New York, 1972.
- [2] T. Aktosun and V. G. Papanicolaou, *Recovery of a potential from the ratio of reflection and transmission coefficients*, J. Math. Phys. **44**, 4875–4883 (2003).
- [3] T. Aktosun and P. E. Sacks, *Inverse problem on the line without phase information*, Inverse Probl. **14**, 211–224 (1998).
- [4] T. Aktosun and R. Weder, *Inverse scattering with partial information on the potential*, J. Math. Anal. Appl. **270**, 247–266 (2002).
- [5] T. Aktosun and R. Weder, *Inverse spectral-scattering problem with two sets of discrete spectra for the radial Schrödinger equation*, Inverse Probl. **22**, 89–114 (2006).

- [6] T. Aktosun and R. Weder, *The Borg–Marchenko theorem with a continuous spectrum, in Recent Advances in Differential Equations and Mathematical Physics*, N. Chernov, Yu. Karpeshina, I. W. Knowles, R. T. Lewis, and R. Weikard (eds.), *Contemp. Math.* **412**, 15–30 (2006).
- [7] V. Ambarzumian, *Über eine Frage der Eigenwerttheorie*, *Z. Phys.* **53**, 690–695 (1929).
- [8] W. O. Amrein, A. M. Hinz, and D. B. Pearson (eds.), *Sturm–Liouville Theory: Past and Present*, Birkhäuser, Basel, 2005.
- [9] E. Andersson, *On the M -function and Borg–Marchenko theorems for vector-valued Sturm–Liouville equations*, *J. Math. Phys.* **44**, 6077–6100 (2003).
- [10] E. Andersson, *A uniqueness theorem in the inverse spectral theory of a certain higher-order ordinary differential equation*, *J. London Math. Soc.* (2) **72**, 169–184 (2005).
- [11] Yu. Arlinskii and E. Tsekanovskii, *Non-self-adjoint Jacobi matrices with a rank-one imaginary part*, *J. Funct. Anal.*, to appear.
- [12] N. Aronszajn and W. F. Donoghue, *On exponential representations of analytic functions in the upper half-plane with positive imaginary part*, *J. Analyse Math.* **5**, 321–388 (1956–57).
- [13] F. V. Atkinson, *On the location of the Weyl circles*, *Proc. Roy. Soc. Edinburgh* **88A**, 345–356 (1981).
- [14] B. Baumgartner, *Level comparison theorems*, *Ann. Phys.* **168**, 484–526 (1986).
- [15] E. D. Belokolos, A. I. Bobenko, V. Z. Enol'skii, A. R. Its, and V. B. Matveev, *Algebro-Geometric Approach to Nonlinear Integrable Equations*, Springer, Berlin, 1994.
- [16] E. D. Belokolos, F. Gesztesy, K. A. Makarov, and L. A. Sakhnovich, *Matrix-valued generalizations of the theorems of Borg and Hochstadt*, in *Evolution Equations*, G. Ruiz Goldstein, R. Nagel, and S. Romanelli (eds.), *Lecture Notes in Pure and Applied Mathematics*, Vol. 234, Marcel Dekker, New York, 2003, pp. 1–34.
- [17] C. Bennewitz, *Spectral asymptotics for Sturm–Liouville equations*, *Proc. London Math. Soc.* (3) **59**, 294–338 (1989).
- [18] C. Bennewitz, *A proof of the local Borg–Marchenko theorem*, *Commun. Math. Phys.* **218**, 131–132 (2001).
- [19] Ju. M. Berezanskii, *Expansions in Eigenfunctions of Self-Adjoint Operators*, *Transl. Math. Monographs* **17**, Amer. Math. Soc., Providence, RI, 1968.
- [20] R. F. Bikbaev and R. A. Sharipov, *Asymptotics at $t \rightarrow \infty$ of the solution to the Cauchy problem for the Korteweg–de Vries equation in the class of potentials with finite-gap behavior as $x \rightarrow \pm\infty$* , *Theoret. Math. Phys.* **78**, 244–252 (1989).
- [21] M. Bôcher *Leçons de Méthodes de Sturm*, Gauthier–Villars, Paris, 1917.
- [22] D. Bollé, F. Gesztesy, H. Grosse, W. Schweiger, and B. Simon, *Witten index, axial anomaly, and Krein's spectral shift function in supersymmetric quantum mechanics*, *J. Math. Phys.* **28**, 1512–1525 (1987).
- [23] C. de Boor and G. H. Golub, *The numerically stable reconstruction of a Jacobi matrix from spectral data*, *Lin. Algebra Appl.* **21**, 245–260 (1978).
- [24] G. Borg, *Eine Umkehrung der Sturm–Liouilleschen Eigenwertaufgabe*, *Acta Math.* **78**, 1–96 (1946).
- [25] G. Borg, *Uniqueness theorems in the spectral theory of $y'' + (\lambda - q(x))y = 0$* , *Proc. 11th Scandinavian Congress of Mathematicians*, Johan Grundt Tanums Forlag, Oslo, 1952, pp. 276–287.
- [26] B. M. Brown, R. A. Peacock, and R. Weikard, *A local Borg–Marchenko theorem for complex potentials*, *J. Comp. Appl. Math.* **148**, 115–131 (2002).
- [27] B. M. Brown and R. Weikard, *A Borg–Levinson theorem for trees*, *Proc. Roy. Soc. London Math. A* **461**, 3231–3243 (2005).
- [28] M. Buys and A. Finkel, *The inverse periodic problem for Hill's equation with a finite-gap potential*, *J. Diff. Eq.* **55**, 257–275 (1984).
- [29] G. Carron, *Déterminant relatif et la fonction ξ* , *Amer. J. Math.* **124**, 307–352 (2002).
- [30] K. Chadan, D. Colton, L. Päivärinta, and W. Rundell, *An Introduction to Inverse Scattering and Inverse Spectral Problems*, SIAM, Philadelphia, 1997.
- [31] K. Chadan and P. C. Sabatier, *Inverse Problems in Quantum Scattering Theory*, 2nd ed., Springer, New York, 1989.
- [32] D. S. Chelkak, *Approximation in the space of spectral data of a perturbed harmonic oscillator*, *J. Math. Sci.* **117**, 4260–4269 (2003).

- [33] D. S. Chelkak, *Asymptotics of spectral data of a harmonic oscillator perturbed by a potential*, J. Math. Sci. **129**, 4053–4082 (2005).
- [34] D. Chelkak, P. Kargaev, and E. Korotyaev, *An inverse problem for an harmonic oscillator perturbed by potential: Uniqueness*, Lett. Math. Phys. **64**, 7–21 (2003).
- [35] D. Chelkak, P. Kargaev, and E. Korotyaev, *Inverse problem for harmonic oscillator perturbed by potential, characterization*, Commun. Math. Phys. **249**, 133–196 (2004).
- [36] D. Chelkak and E. Korotyaev, *The inverse problem for perturbed harmonic oscillator on the half-line*, preprint, 2005.
- [37] S. Clark and F. Gesztesy, *Weyl–Titchmarsh M -function asymptotics for matrix-valued Schrödinger operators*, Proc. London Math. Soc. **82**, 701–724 (2001).
- [38] S. Clark and F. Gesztesy, *Weyl–Titchmarsh M -function asymptotics and Borg-type theorems for Dirac operators*, Trans. Amer. Math. Soc. **354**, 3475–3534 (2002).
- [39] S. Clark, F. Gesztesy, H. Holden, and B. M. Levitan, *Borg-type theorems for matrix-valued Schrödinger operators*, J. Diff. Eq. **167**, 181–210 (2000).
- [40] S. Clark, F. Gesztesy, and W. Renger, *Trace formulas and Borg-type theorems for matrix-valued Jacobi and Dirac finite difference operators*, J. Diff. Eq. **219**, 144–182 (2005).
- [41] E. A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw–Hill, New York, 1955.
- [42] W. A. Coppel *Disconjugacy*, Lecture Notes in Math. **220**, Springer, Berlin, 1971.
- [43] W. Craig, *The trace formula for Schrödinger operators on the line*, Commun. Math. Phys. **126**, 379–407 (1989).
- [44] M. M. Crum, *Associated Sturm–Liouville systems*, Quart. J. Math. Oxford (2) **6**, 121–127 (1955).
- [45] A. A. Danielyan and B. M. Levitan, *On the asymptotic behavior of the Weyl–Titchmarsh m -function*, Math. USSR Izv. **36**, 487–496 (1991).
- [46] P. A. Deift, *Applications of a commutation formula*, Duke Math. J. **45**, 267–310 (1978).
- [47] P. Deift and B. Simon, *Almost periodic Schrödinger operators III. The absolutely continuous spectrum in one dimension*, Commun. Math. Phys. **90**, 389–411 (1983).
- [48] P. Deift and E. Trubowitz, *Inverse scattering on the line*, Commun. Pure Appl. Math. **32**, 121–251 (1979).
- [49] R. del Rio Castillo, *On boundary conditions of an inverse Sturm–Liouville problem*, SIAM J. Appl. Math. **50**, 1745–1751 (1990).
- [50] R. del Rio, F. Gesztesy, and B. Simon, *Inverse spectral analysis with partial information on the potential, III. Updating boundary conditions*, Intl. Math. Research Notices **1997**, No. 15, 751–758.
- [51] R. del Rio, F. Gesztesy, and B. Simon, *Corrections and Addendum to “Inverse spectral analysis with partial information on the potential, III. Updating boundary conditions”*, Intl. Math. Research Notices **1999**, No. 11, 623–625.
- [52] R. del Rio and B. Grébert, *Inverse spectral results for AKNS systems with partial information on the potentials*, Math. Phys. Anal. Geom. **4**, 229–244 (2001).
- [53] M. Demuth and M. Krishna, *Determining Spectra in Quantum Theory*, Birkhäuser, Boston, 2005.
- [54] M. S. Derevyagin, *Borg-type theorems for generalized Jacobi matrices and trace formulas*, Meth. Funct. Anal. Topology, to appear.
- [55] L.A. Dikii, *Trace formulas for Sturm–Liouville differential operators*, Amer. Math. Soc. Trans. Ser. (2) **18**, 81–115 (1961).
- [56] B. A. Dubrovin, *Periodic problems for the Korteweg–de Vries equation in the class of finite band potentials*, Funct. Anal. Appl. **9**, 215–223 (1975).
- [57] N. Dunford and J. T. Schwartz, *Linear Operators, Part II: Spectral Theory*, Wiley–Interscience, New York, 1988.
- [58] M. S. P. Eastham, *The Spectral Theory of Periodic Differential Equations*, Scottish Academic Press, Edinburgh, 1973.
- [59] M. S. P. Eastham and H. Kalf, *Schrödinger-Type Operators with Continuous Spectra*, Pitman, Boston, 1982.
- [60] F. Ehlers and H. Knörrer, *An algebro-geometric interpretation of the Bäcklund transformation of the Korteweg–de Vries equation*, Comment. Math. Helv. **57**, 1–10 (1982).
- [61] N. M. Ercolani and H. Flaschka, *The geometry of the Hill equation and of the Neumann system*, Phil. Trans. Roy. Soc. London **A315**, 405–422 (1985).

- [62] W. N. Everitt, *On a property of the m -coefficient of a second-order linear differential equation*, J. London Math. Soc. **4**, 443–457 (1972).
- [63] A. Finkel, E. Isaacson, and E. Trubowitz, *An explicit solution of the inverse periodic problem for Hill's equation*, SIAM J. Math. Anal. **18**, 46–53 (1987).
- [64] N. E. Firsova, *On solution of the Cauchy problem for the Korteweg–de Vries equation with initial data the sum of a periodic and a rapidly decreasing function*, Math. USSR Sbornik **63**, 257–265 (1989).
- [65] H. Flaschka, *On the inverse problem for Hill's operator*, Arch. Rat. Mech. Anal. **59**, 293–309 (1975).
- [66] H. Flaschka and D. W. McLaughlin, *Some comments on Bäcklund transformations, canonical transformations, and the inverse scattering method*, in “Bäcklund Transformations, the Inverse Scattering Method, Solitons, and their Applications”, R. M. Miura (ed.), Lecture Notes in Math. **515**, Springer, Berlin, 1976, pp. 252–295.
- [67] L. Fu and H. Hochstadt, *Inverse theorems for Jacobi matrices*, J. Math. Anal. Appl. **47**, 162–168 (1974).
- [68] C. S. Gardner, J. M. Greene, M. D. Kruskal, and R. M. Miura, *Korteweg–de Vries equation and generalizations. VI. Methods for exact solution*, Commun. Pure Appl. Math. **27**, 97–133 (1974).
- [69] I. M. Gel'fand, *On identities for eigenvalues of a second order differential operator*, Uspek. Mat. Nauk **11:1**, 191–198 (1956) (Russian); Engl. transl. in Izrail M. Gelfand, *Collected Papers Vol. I*, S. G. Gindikin, V. W. Guillemin, A. A. Kirillov, B. Kostant, and S. Sternberg (eds.), Springer, Berlin, 1987, pp. 510–517.
- [70] I. M. Gel'fand and B. M. Levitan, *On the determination of a differential equation from its special function*, Izv. Akad. Nauk SSR. Ser. Mat. **15**, 309–360 (1951) (Russian); Engl. transl. in Amer. Math. Soc. Transl. Ser. 2, **1**, 253–304 (1955).
- [71] I. M. Gel'fand and B. M. Levitan, *On a simple identity for eigenvalues of a second order differential operator*, Dokl. Akad. Nauk SSSR **88**, 593–596 (1953) (Russian); Engl. transl. in Izrail M. Gelfand, *Collected Papers Vol. I*, S. G. Gindikin, V. W. Guillemin, A. A. Kirillov, B. Kostant, and S. Sternberg (eds.), Springer, Berlin, 1987, pp. 457–461.
- [72] F. Gesztesy, *On the modified Korteweg–de Vries equation, in Differential Equations with Applications in Biology, Physics, and Engineering*, J. A. Goldstein, F. Kappel, and W. Schappacher (eds.), Marcel Dekker, New York, 1991, pp. 139–183.
- [73] F. Gesztesy, *Quasi-periodic, finite-gap solutions of the modified Korteweg–de Vries equation*, in *Ideas and Methods in Mathematical Analysis, Stochastics, and Applications*, S. Albeverio, J. E. Fenstad, H. Holden, and T. Lindstrøm (eds.), Vol. 1, Cambridge Univ. Press, Cambridge, 1992, pp. 428–471.
- [74] F. Gesztesy, *A complete spectral characterization of the double commutation method*, J. Funct. Anal. **117**, 401–446 (1993).
- [75] F. Gesztesy, *New trace formulas for Schrödinger operators*, in *Evolution Equations*, G. Ferreira, G. Ruiz Goldstein, and F. Neubrander (eds.), Marcel Dekker, New York, 1995, pp. 201–221.
- [76] F. Gesztesy and H. Holden, *On new trace formulae for Schrödinger operators*, Acta Appl. Math. **39**, 315–333 (1995).
- [77] F. Gesztesy and H. Holden, *On trace formulas for Schrödinger-type operators*, in *Multi-particle Quantum Scattering with Applications to Nuclear, Atomic, and Molecular Physics*, D. G. Truhlar and B. Simon (eds.), Springer, New York, 1997, pp. 121–145.
- [78] F. Gesztesy and H. Holden, *Darboux-type transformations and hyperelliptic curves*, J. reine angew. Math. **527**, 151–183 (2000).
- [79] F. Gesztesy and H. Holden, *Soliton Equations and Their Algebro-Geometric Solutions. Vol. I: (1 + 1)-Dimensional Continuous Models*, Cambridge Studies in Advanced Mathematics, Vol. 79, Cambridge Univ. Press, Cambridge, 2003.
- [80] F. Gesztesy, H. Holden, and B. Simon, *Absolute summability of the trace relation for certain Schrödinger operators*, Commun. Math. Phys. **168**, 137–161 (1995).
- [81] F. Gesztesy, H. Holden, B. Simon, and Z. Zhao, *Trace formulae and inverse spectral theory for Schrödinger operators*, Bull. Amer. Math. Soc. **29**, 250–255 (1993).
- [82] F. Gesztesy, H. Holden, B. Simon, and Z. Zhao, *Higher order trace relations for Schrödinger operators*, Rev. Math. Phys. **7**, 893–922 (1995).

- [83] F. Gesztesy, H. Holden, B. Simon, and Z. Zhao, *A trace formula for multidimensional Schrödinger operators*, J. Funct. Anal. **141**, 449–465 (1996).
- [84] F. Gesztesy, A. Kiselev, and K. A. Makarov, *Uniqueness results for matrix-valued Schrödinger, Jacobi, and Dirac-type operators*, Math. Nachr. **239–240**, 103–145 (2002).
- [85] F. Gesztesy, M. Krishna, and G. Teschl, *On isospectral sets of Jacobi operators*, Commun. Math. Phys. **181**, 631–645 (1996).
- [86] F. Gesztesy and K. A. Makarov, *Some applications of the spectral shift operator*, in *Operator Theory and its Applications*, A. G. Ramm, P. N. Shivakumar, and A. V. Strauss (eds.), Fields Institute Communications, Vol. 25, Amer. Math. Society, Providence, RI, 2000, pp. 267–292.
- [87] F. Gesztesy, W. Plessas, and B. Thaller, *On the high-energy behavior of scattering phase shifts for Coulomb-like potentials*, J. Phys. A **13**, 2659–2671 (1980).
- [88] F. Gesztesy, R. Ratnaseelan, and G. Teschl, *The KdV hierarchy and associated trace formulas*, in *Recent Developments in Operator Theory and its Applications*, I. Gohberg, P. Lancaster, and P. N. Shivakumar (eds.), Operator Theory: Advances and Applications, Vol. 87, Birkhäuser, Basel, 1996, pp. 125–163.
- [89] F. Gesztesy and L. A. Sakhnovich, *A class of matrix-valued Schrödinger operators with prescribed finite-band spectra*, in *Reproducing Kernel Hilbert Spaces, Positivity, System Theory and Related Topics*, D. Alpay (ed.), Operator Theory: Advances and Applications, Vol. 143, Birkhäuser, Basel, 2003, pp. 213–253.
- [90] F. Gesztesy, W. Schweiger, and B. Simon, *Commutation methods applied to the mKdV-equation*, Trans. Amer. Math. Soc. **324**, 465–525 (1991).
- [91] F. Gesztesy and B. Simon, *The xi function*, Acta Math. **176**, 49–71 (1996).
- [92] F. Gesztesy and B. Simon, *Uniqueness theorems in inverse spectral theory for one-dimensional Schrödinger operators*, Trans. Amer. Math. Soc. **348**, 349–373 (1996).
- [93] F. Gesztesy and B. Simon, *Inverse spectral analysis with partial information on the potential, I. The case of an a.c. component in the spectrum*, Helv. Phys. Acta **70**, 66–71 (1997).
- [94] F. Gesztesy and B. Simon, *m-functions and inverse spectral analysis for finite and semi-infinite Jacobi matrices*, J. Analyse Math. **73**, 267–297 (1997).
- [95] F. Gesztesy and B. Simon, *On the determination of a potential from three spectra*, in *Differential Operators and Spectral Theory*, V. Buslaev, M. Solomyak, and D. Yafaev (eds.), Amer. Math. Soc. Transl. Ser. 2, **189**, 85–92 (1999).
- [96] F. Gesztesy and B. Simon, *Inverse spectral analysis with partial information on the potential, II. The case of discrete spectrum*, Trans. Amer. Math. Soc. **352**, 2765–2787 (2000).
- [97] F. Gesztesy and B. Simon, *A new approach to inverse spectral theory, II. General real potentials and the connection to the spectral measure*, Ann. Math. **152**, 593–643 (2000).
- [98] F. Gesztesy and B. Simon, *On local Borg–Marchenko uniqueness results*, Commun. Math. Phys. **211**, 273–287 (2000).
- [99] F. Gesztesy and B. Simon, *Connectedness of the isospectral manifold for one-dimensional half-line Schrödinger operators*, J. Stat. Phys. **116**, 361–365 (2004).
- [100] F. Gesztesy, B. Simon, and G. Teschl, *Spectral deformations of one-dimensional Schrödinger operators*, J. Analyse Math. **70**, 267–324 (1996).
- [101] F. Gesztesy, B. Simon, and G. Teschl, *Zeros of the Wronskian and renormalized oscillation theory*, Amer. J. Math. **118**, 571–594 (1996).
- [102] F. Gesztesy and R. Svirsky, *(m)KdV-Solitons on the background of quasi-periodic finite-gap solutions*, Memoirs Amer. Math. Soc. **118** (563), 1–88 (1995).
- [103] F. Gesztesy and G. Teschl, *On the double commutation method*, Proc. Amer. Math. Soc. **124**, 1831–1840 (1996).
- [104] F. Gesztesy and G. Teschl, *Commutation methods for Jacobi operators*, J. Diff. Eq. **128**, 252–299 (1996).
- [105] F. Gesztesy and R. Weikard, *Spectral deformations and soliton equations*, in *Differential Equations with Applications in Mathematical Physics*, W. F. Ames, E. M. Harrell, and J. V. Herod (eds.), Academic Press, Boston, 1993, pp. 101–139.
- [106] F. Gesztesy and P. Yuditskii, *Spectral properties of a class of reflectionless Schrödinger operators*, J. Funct. Anal., to appear.
- [107] F. Gesztesy and M. Zinchenko, *A Borg-type theorem associated with orthogonal polynomials on the unit circle*, J. London Math. Soc., to appear.

- [108] P. C. Gibson, *Inverse spectral theory of finite Jacobi matrices*, Trans. Amer. Math. Soc. **354**, 4703–4749 (2002).
- [109] R. C. Gilbert and V. A. Kramer, *Trace formulas for powers of a Sturm–Liouville operator*, Canad. J. Math. **16**, 412–422 (1964).
- [110] W. B. Gragg and W. J. Harrod, *The numerically stable reconstruction of Jacobi matrices from spectral data*, Num. Math., **44**, 317–335 (1984).
- [111] L. J. Gray and D. G. Wilson, *Construction of a Jacobi matrix from spectral data*, Lin. Algebra Appl. **14**, 131–134 (1976).
- [112] H. Grosse and A. Martin, *Particle Physics and the Schrödinger Equation*, Cambridge Univ. Press, Cambridge, 1997.
- [113] C. J. A. Halberg and V. A. Kramer, *A generalization of the trace concept*, Duke Math. J. **27**, 607–617 (1960).
- [114] O. Hald, *Inverse eigenvalue problems for Jacobi matrices*, Lin. Algebra Appl. **14**, 63–85 (1976).
- [115] O. H. Hald, *Inverse eigenvalue problem for the mantle*, Geophys. J. R. Astr. Soc. **62**, 41–48 (1980).
- [116] B. J. Harris, *The form of the spectral functions associated with Sturm–Liouville problems with continuous spectrum*, Mathematika **44**, 149–161 (1997).
- [117] M. Hitrik, *Properties of the scattering transform on the real line*, J. Math. Anal. Appl. **258**, 223–243 (2001).
- [118] H. Hochstadt, *On the determination of a Hill's equation from its spectrum*, Arch. Rat. Mech. Anal. **19**, 353–362 (1965).
- [119] H. Hochstadt, *On some inverse problems in matrix theory*, Arch. Math. **18**, 201–207 (1967).
- [120] H. Hochstadt, *On inverse problems associated with second-order differential operators*, Acta Math. **119**, 173–192 (1967).
- [121] H. Hochstadt, *On the construction of a Jacobi matrix from spectral data*, Lin. Algebra Appl. **8**, 435–446 (1974).
- [122] H. Hochstadt, *On the construction of a Jacobi matrix from mixed given data*, Lin. Algebra Appl. **28**, 113–115 (1979).
- [123] H. Hochstadt and B. Lieberman, *An inverse Sturm–Liouville problem with mixed given data*, SIAM J. Appl. Math. **34**, 676–680 (1978).
- [124] M. Horváth, *On the inverse spectral theory of Schrödinger and Dirac operators*, Trans. Amer. Math. Soc. **353**, 4155–4171 (2001).
- [125] M. Horváth, *Inverse spectral problems and closed exponential systems*, Ann. Math. **162**, 885–918 (2005).
- [126] M. Horváth, *Inverse scattering with fixed energy and an inverse eigenvalue problem on the half-line*, Trans. Amer. Math. Soc. **358**, 5161–5177 (2006).
- [127] R. O. Hryniv and Ya. V. Mykytyuk, *Inverse spectral problems for Sturm–Liouville operators with singular potentials*, Inverse Probl. **19**, 665–684 (2003).
- [128] R. O. Hryniv and Ya. V. Mykytyuk, *Inverse spectral problems for Sturm–Liouville operators with singular potentials, III. Reconstruction by three spectra*, J. Math. Anal. Appl. **284**, 626–646 (2003).
- [129] R. O. Hryniv and Ya. V. Mykytyuk, *Inverse spectral problems for Sturm–Liouville operators with singular potentials, II. Reconstruction by two spectra*, in *Functional Analysis and its Applications*, V. Kadets and W. Zelazko (eds.), North–Holland Math. Stud. Vol. 197, Elsevier, Amsterdam, 2004, pp. 97–114.
- [130] R. O. Hryniv and Ya. V. Mykytyuk, *Half-inverse spectral problems for Sturm–Liouville operators with singular potentials*, Inverse Probl. **20**, 1423–1444 (2004).
- [131] R. O. Hryniv and Ya. V. Mykytyuk, *Inverse spectral problems for Sturm–Liouville operators with singular potentials, IV. Potentials in the Sobolev scale.*, preprint, 2004.
- [132] K. Iwasaki, *Inverse problem for Sturm–Liouville and Hill's equations*, Ann. Mat. Pura Appl. (4) **149**, 185–206 (1987).
- [133] V. A. Javřjan, *On the regularized trace of the difference between two singular Sturm–Liouville operators*, Sov. Math. Dokl. **7**, 888–891 (1966).
- [134] V. A. Javřjan, *A certain inverse problem for Sturm–Liouville operators*, Izv. Akad. Nauk Armjan. SSR Ser. Mat. **6**, 246–251 (1971) (Russian).
- [135] I. Kay and H. E. Moses, *Reflectionless transmission through dielectrics and scattering potentials*, J. Appl. Phys. **27**, 1503–1508 (1956).

- [136] A. B. Khaled, *Problème inverse de Sturm–Liouville associé à un opérateur différentiel singulier*, C. R. Acad. Sc. Paris **299**, 221–224 (1984).
- [137] A. M. Khodakovsky, *Inverse Spectral Problem with Partial Information on the Potential*, Ph.D. Thesis, California Institute of Technology, Pasadena, CA, 1999.
- [138] A. Khodakovsky, *Inverse spectral problem with partial information on the potential: The case of the whole real line*, Commun. Math. Phys. **210**, 399–411 (2000).
- [139] K. Knudsen, *On a local uniqueness result for the inverse Sturm–Liouville problem*, Ark. Mat. **39**, 361–373 (2001).
- [140] S. Kotani, *Ljapunov indices determine absolutely continuous spectra of stationary random one-dimensional Schrödinger operators*, in *Stochastic Analysis*, K. Ito (ed.), North–Holland, Amsterdam, 1984, pp. 225–247.
- [141] S. Kotani and M. Krishna, *Almost periodicity of some random potentials*, J. Funct. Anal. **78**, 390–405 (1988).
- [142] M. G. Krein, *Solution of the inverse Sturm–Liouville problem*, Dokl. Akad. Nauk SSSR **76**, 21–24 (1951) (Russian).
- [143] M. G. Krein, *On the transfer function of a one-dimensional boundary value problem of the second-order*, Dokl. Akad. Nauk SSSR **88**, 405–408 (1953) (Russian).
- [144] M. G. Krein, *On some cases of effective determination of the density of an inhomogeneous cord from its spectral function*, Dokl. Akad. Nauk SSSR **93**, 617–620 (1953) (Russian).
- [145] M. G. Krein, *On a method of effective solution of an inverse boundary problem*, Dokl. Akad. Nauk SSSR **94**, 987–990 (1954) (Russian).
- [146] M. G. Krein, *Perturbation determinants and a formula for the traces of unitary and self-adjoint operators*, Sov. Math. Dokl. **3**, 707–710 (1962).
- [147] K. Kreith, *Oscillation Theory*, Lecture Notes in Math. 324, Springer, Berlin, 1973.
- [148] Y. Last, *A relation between a.c. spectrum of ergodic Jacobi matrices and the spectra of periodic approximants*, Commun. Math. Phys. **151**, 183–192 (1993).
- [149] Y. Last, *Spectral theory of Sturm–Liouville operators on infinite intervals: A review of recent developments*, in *Sturm–Liouville Theory: Past and Present*, W. O. Amrein, A. M. Hinz, and D. B. Pearson (eds.), Birkhäuser, Basel, 2005, pp. 99–120.
- [150] P. Lax, *Trace formulas for the Schrödinger operator*, Commun. Pure Appl. Math. **47**, 503–512 (1994).
- [151] W. Leighton, *On self-adjoint differential equations of second order*, J. London Math. Soc. **27**, 37–47 (1952).
- [152] B. Ja. Levin, *Distribution of Zeros of Entire Functions*, rev. ed., Amer. Math. Soc., Providence, RI, 1980.
- [153] N. Levinson, *The inverse Sturm–Liouville problem*, Mat. Tidskr. B, 25–30 (1949).
- [154] B. M. Levitan, *On the determination of a Sturm–Liouville equation by two spectra*, Amer. Math. Soc. Transl. Ser. 2, **68**, 1–20 (1968).
- [155] B. M. Levitan, *On the asymptotic behavior of the spectral function of a self-adjoint second order differential equation*, Amer. Math. Soc. Transl. Ser. 2, **101**, 192–221 (1973).
- [156] B. M. Levitan, *On the asymptotic behavior of the spectral function of a self-adjoint differential equation of the second order and on expansion in eigenfunctions*, Amer. Math. Soc. Transl. Ser. 2, **102**, 191–229 (1973).
- [157] B. M. Levitan, *On the asymptotic behavior of the spectral function of a self-adjoint differential equation of the second order and on expansion in eigenfunctions. II*, Amer. Math. Soc. Transl. Ser. 2, **110**, 165–188 (1977).
- [158] B. M. Levitan, *On the closure of the set of finite-zone potentials*, Math. USSR Sbornik **51**, 67–89 (1985).
- [159] B. M. Levitan, *Inverse Sturm–Liouville Problems*, VNU Science Press, Utrecht, 1987.
- [160] B. M. Levitan, *Sturm–Liouville operators on the whole line, with the same discrete spectrum*, Math. USSR Sbornik **60**, 77–106 (1988).
- [161] B. M. Levitan and M. G. Gasymov, *Determination of a differential equation by two of its spectra*, Russian Math. Surv. **19:2**, 1–63 (1964).
- [162] B. M. Levitan and I. S. Sargsjan, *Introduction to Spectral Theory*, Transl. Math. Monographs **39**, Amer. Math. Soc., Providence, RI, 1975.
- [163] W. Magnus and S. Winkler, *Hill's Equation*, Dover, New York, 1979.
- [164] M. M. Malamud, *Uniqueness questions in inverse problems for systems of ordinary differential equations on a finite interval*, Trans. Moscow Math. Soc. **60**, 173–224 (1999).

- [165] M. M. Malamud, *Borg type theorems for first-order systems on a finite interval*, *Funct. Anal. Appl.* **33**, 64–68 (1999).
- [166] M. M. Malamud, *Uniqueness of the matrix Sturm–Liouville equation given a part of the monodromy matrix, and Borg type results*, in *Sturm–Liouville Theory: Past and Present*, W. O. Amrein, A. M. Hinz, and D. B. Pearson (eds.), Birkhäuser, Basel, 2005, pp. 237–270.
- [167] M. M. Malamud, *Borg type results for Dirac systems on a finite interval*, *Russ. J. Math. Phys.*, to appear.
- [168] M. M. Malamud, *On the unique recovery of the n th order ordinary differential equation by a part of the monodromy matrix*, preprint, 2005.
- [169] S. M. Malamud, *Inverse spectral problem for normal matrices and the Gauss–Lucas theorem*, *Trans. Amer. Math. Soc.* **357**, 4043–4064 (2004).
- [170] V. A. Marchenko, *Certain problems in the theory of second-order differential operators*, *Doklady Akad. Nauk SSSR* **72**, 457–460 (1950) (Russian).
- [171] V. A. Marchenko, *Some questions in the theory of one-dimensional linear differential operators of the second order, I*, *Trudy Moskov. Mat. Obšč.* **1**, 327–420 (1952) (Russian); Engl. transl. in *Amer. Math. Soc. Transl. Ser. 2*, **101**, 1–104 (1973).
- [172] V. A. Marchenko, *Sturm–Liouville Operators and Applications*, Birkhäuser, Basel, 1986.
- [173] A. I. Markushevich, *Theory of Functions of a Complex Variable*, 2nd ed., Chelsea, New York, 1985.
- [174] H. P. McKean, *Variation on a theme of Jacobi*, *Commun. Pure Appl. Math.* **38**, 669–678 (1985).
- [175] H. P. McKean, *Geometry of KdV (1): Addition and the unimodular spectral classes*, *Rev. Mat. Iberoamericana* **2**, 235–261 (1986).
- [176] H. P. McKean, *Geometry of KdV (2): Three examples*, *J. Stat. Phys.* **46**, 1115–1143 (1987).
- [177] H. P. McKean, *Geometry of KdV (3): Determinants and unimodular isospectral flows*, *Commun. Pure Appl. Math.* **45**, 389–415 (1992).
- [178] H. P. McKean and P. van Moerbeke, *The spectrum of Hill’s equation*, *Invent. Math.* **30**, 217–274 (1975).
- [179] H. P. McKean and E. Trubowitz, *Hill’s operator and hyperelliptic function theory in the presence of infinitely many branch points*, *Commun. Pure Appl. Math.* **29**, 143–226 (1976).
- [180] H. P. McKean and E. Trubowitz, *The spectral class of the quantum-mechanical harmonic oscillator*, *Commun. Math. Phys.* **82**, 471–495 (1982).
- [181] J. Michor and G. Teschl, *Reconstructing Jacobi matrices from three spectra*, in *Spectral Methods for Operators of Mathematical Physics*, J. Janas, P. Kurasov, and S. Naboko (eds.), *Operator Theory: Advances and Applications*, Vol. 154, Birkhäuser, Basel, 2004, pp. 151–154.
- [182] M. A. Naimark, *Linear Differential Operators, Part II*, F. Ungar, New York, 1968.
- [183] S. Novikov, S. V. Manakov, L. P. Pitaevskii, and V. E. Zakharov, *Theory of Solitons*, Consultants Bureau, New York, 1984.
- [184] V. Pivovarchik, *An inverse Sturm–Liouville problem by three spectra*, *Integr. equ. oper. theory* **34**, 234–243 (1999).
- [185] V. N. Pivovarchik, *Reconstruction of the potential of the Sturm–Liouville equation from three spectra of boundary value problems*, *Funct. Anal. Appl.* **33**, 233–235 (1999).
- [186] V. Pivovarchik, *A particular case of the inverse problem for the Sturm–Liouville equation with parameter-dependent potential*, in *Mathematical Results in Quantum Mechanics*, J. Dittrich, P. Exner, and M. Tater (eds.), *Operator Theory: Advances and Applications*, Vol. 108, 1999, pp. 335–341.
- [187] V. Pivovarchik, *Inverse problem for the Sturm–Liouville equation on a simple graph*, *SIAM J. Math. Anal.* **32**, 801–819 (2000).
- [188] V. Pivovarchik, *Direct and inverse three-point Sturm–Liouville problem with parameter-dependent boundary conditions*, *Asymptot. Anal.* **26**, 219–238 (2001).
- [189] V. Pivovarchik, *Recovering a part of potential by partial information on spectra of boundary problems*, *Opuscula Math.* **25**, 131–137 (2005).
- [190] V. Pivovarchik, *Inverse problem for the Sturm–Liouville equation on a star-shaped graph*, *Math. Nachrichten*, to appear.
- [191] G. Pólya and G. Szegő, *Problems and Theorems in Analysis I*, Springer, Berlin, 1972.
- [192] J. Pöschel and E. Trubowitz, *Inverse Spectral Theory*, Academic Press, Boston, 1987.

- [193] A. Pushnitski and I. Sorrell, *High energy asymptotics and trace formulas for the perturbed harmonic oscillator*, Ann. H. Poincaré **7**, 1–17 (2005).
- [194] J. Ralston and E. Trubowitz, *Isospectral sets for boundary value problems on the unit interval*, Ergodic Th. Dynam. Syst. **8**, 301–358 (1988).
- [195] A. G. Ramm, *Recovery of the potential from the I-function*, C. R. Math. Rep. Acad. Sci. Canada **IX**, 177–182 (1987).
- [196] A. G. Ramm, *Multidimensional Inverse Scattering Problems*, Longman, New York, 1992.
- [197] A. G. Ramm, *Property C for ordinary differential equations and applications to inverse scattering*, Z. Analysis Anwendungen **18**, 331–348 (1999).
- [198] A. G. Ramm, *Property C for ODE and applications to inverse problems*, in *Operator Theory and its Applications*, A. G. Ramm, P. N. Shivakumar, and A. V. Strauss (eds.), Fields Inst. Commun. Ser., Vol. 25, Amer. Math. Soc., Providence, RI, 2000, pp. 15–75.
- [199] A. G. Ramm, *Reconstruction of the potential from I-function*, J. Inverse Ill-Posed Probl. **10**, 385–394 (2002).
- [200] A. Ramm and B. Simon, *A new approach to inverse spectral theory, III. Short range potentials*, J. Analyse Math. **80**, 319–334 (2000).
- [201] M. Reed and B. Simon *Methods of Mathematical Physics. II. Fourier Analysis, Self-Adjointness*, Academic Press, New York, 1975.
- [202] M. Reed and B. Simon, *Methods of Modern Mathematical Physics. IV: Analysis of Operators*, Academic Press, New York, 1978.
- [203] W. T. Reid, *Sturmian Theory for Ordinary Differential Equations*, Springer, New York, 1980.
- [204] C. Remling, *Schrödinger operators and de Branges spaces*, J. Funct. Anal. **196**, 323–394 (2002).
- [205] C. Remling, *Inverse spectral theory for one-dimensional Schrödinger operators: The A function*, Math. Z. **245**, 597–617 (2003).
- [206] W. Rundell and P. E. Sacks, *Reconstruction techniques for classical inverse Sturm–Liouville problems*, Math. Comp. **58**, 161–183 (1992).
- [207] A. V. Rybin and M. A. Sall', *Solitons of the Korteweg–de Vries equation on the background of a known solution*, Theoret. Math. Phys. **63**, 545–550 (1985).
- [208] A. Rybkin, *The KdV invariants and Herglotz functions*, Diff. Integral. Eq. **14**, 493–512 (2001).
- [209] A. Rybkin, *On the trace approach to the inverse scattering problem in dimension one*, SIAM J. Math. Anal. **32**, 1248–1264 (2001).
- [210] A. Rybkin, *Some new and old asymptotic representations of the Jost solution and the Weyl m-function for Schrödinger operators on the line*, Bull. London Math. Soc. **34**, 61–72 (2002).
- [211] A. Rybkin, *Necessary and sufficient conditions for absolute summability of the trace formulas for certain one dimensional Schrödinger operators*, Proc. Amer. Math. Soc. **131**, 219–229 (2002).
- [212] A. Sakhnovich, *Dirac type and canonical systems: Spectral and Weyl–Titchmarsh matrix functions, direct and inverse problems*, Inverse Probl. **18**, 331–348 (2002).
- [213] L. Sakhnovich, *Half-inverse problems on the finite interval*, Inverse Probl. **17**, 527–532 (2001).
- [214] K. M. Schmidt, *Critical coupling constants and eigenvalue asymptotics of perturbed periodic Sturm–Liouville operators*, Commun. Math. Phys. **211**, 465–485 (2000).
- [215] K. M. Schmidt, *An application of Gesztesy–Simon–Teschl oscillation theory to a problem in differential geometry*, J. Math. Anal. Appl. **261**, 61–71 (2001).
- [216] U.-W. Schmincke, *On Schrödinger's factorization method for Sturm–Liouville operators*, Proc. Roy. Soc. Edinburgh **80A**, 67–84 (1978).
- [217] U.-W. Schmincke, *On a paper by Gesztesy, Simon, and Teschl concerning isospectral deformations of ordinary Schrödinger operators*, J. Math. Anal. Appl. **277**, 51–78 (2003).
- [218] R. A. Sharipov, *Finite-zone analogues of N-multiplet solutions of the Korteweg–de Vries equation*, Russ. Math. Surv. **41:5**, 165–166 (1986).
- [219] R. A. Sharipov, *Soliton multiplets of the Korteweg–de Vries equation*, Sov. Phys. Dokl. **32**, 121–123 (1987).
- [220] C.-L. Shen, *Some inverse spectral problems for vectorial Sturm–Liouville equations*, Inverse Probl. **17**, 1253–1294 (2001).

- [221] C.-L. Shen, *On the Liouville transformation and some inverse spectral problems*, Inverse Probl. **21**, 591–614 (2005).
- [222] C.-T. Shieh, *Some inverse problems on Jacobi matrices*, Inverse Probl. **20**, 589–600 (2004).
- [223] T. Shirai, *A trace formula for discrete Schrödinger operators*, Publ. Res. Inst. Math. Sci. **34**, 27–41 (1998).
- [224] L. O. Silva and R. Weder, *On the two spectra inverse problem for semi-infinite Jacobi matrices*, preprint, 2005.
- [225] B. Simon, *Functional Integration and Quantum Physics*, Academic Press, New York, 1979; and AMS Chelsea Publishing, 2nd ed., Amer. Math. Soc., Providence, RI, 2005.
- [226] B. Simon, *Trace Ideals and Their Applications*, London Math. Soc. Lecture Note Series, Vol. 35, Cambridge University Press, Cambridge, 1979; and Mathematical Surveys and Monographs, Vol. 120, 2nd ed., Amer. Math. Soc., Providence, RI, 2005.
- [227] B. Simon, Spectral analysis of rank one perturbations and applications, in *Proc. Mathematical Quantum Theory II: Schrödinger Operators*, J. Feldman, R. Froese, and L. M. Rosen (eds.), CRM Proceedings and Lecture Notes Vol. 8, Amer. Math. Soc., Providence, RI, 1995, pp. 109–149; reprinted in the 2nd ed. of [226].
- [228] B. Simon, *A new approach to inverse spectral theory, I. Fundamental formalism*, Ann. Math. **150**, 1029–1057 (1999).
- [229] B. Simon, *Orthogonal Polynomials on the Unit Circle, Part 1: Classical Theory, Part 2: Spectral Theory*, AMS Colloquium Publication Series, Vol. 54, Providence, RI, 2005.
- [230] B. Simon, *Sturm oscillation and comparison theorems*, in *Sturm–Liouville Theory: Past and Present*, W. O. Amrein, A. M. Hinz, and D. B. Pearson (eds.), Birkhäuser, Basel, 2005, pp. 29–43.
- [231] T. Stieltjes, *Recherches sur les fractions continues*, Ann. Fac. Sci. Toulouse, **8**, 1–122 (1894); **9**, A1–47 (1895).
- [232] J. C. F. Sturm, *Mémoire sur les équations différentielles linéaires du second ordre*, J. Math. Pures Appl. **1**, 106–186 (1836).
- [233] T. Suzuki, *Deformation formulas and their applications to spectral and evolutionary inverse problems*, Lecture Notes in Num. Appl. Anal. **5**, 289–311 (1982).
- [234] T. Suzuki, *Inverse problems for heat equations on compact intervals and on circles, I*, J. Math. Soc. Japan **38**, 39–65 (1986).
- [235] C. A. Swanson, *Comparison and Oscillation Theory of Linear Differential Equations*, Academic Press, New York, 1968.
- [236] W. Symes, *Inverse boundary value problems and a theorem of Gel'fand and Levitan*, J. Math. Anal. Appl. **71**, 379–402 (1979).
- [237] G. Teschl, *Oscillation theory and renormalized oscillation theory for Jacobi operators*, J. Diff. Eq. **129**, 532–558 (1996).
- [238] G. Teschl, *Spectral deformations of Jacobi operators*, J. reine angew. Math. **491**, 1–15 (1997).
- [239] G. Teschl, *Trace formulas and inverse spectral theory for Jacobi operators*, Commun. Math. Phys. **196**, 175–202 (1998).
- [240] G. Teschl, *Deforming the point spectra of one-dimensional Dirac operators*, Proc. Amer. Math. Soc. **126**, 2873–2881 (1998).
- [241] G. Teschl, *Renormalized oscillation theory for Dirac operators*, Proc. Amer. Math. Soc. **126**, 1685–1695 (1998).
- [242] G. Teschl, *On the Toda and Kac–van Moerbeke hierarchies*, Math. Z. **231**, 325–344 (1999).
- [243] G. Teschl, *Jacobi Operators and Completely Integrable Nonlinear Lattices*, Math. Surv. Monographs, Vol. 72, Amer. Math. Soc., Providence, RI, 2000.
- [244] C. Thurlow, *A generalisation of the inverse spectral theorem of Levitan and Gasymov*, Proc. Roy. Soc. Edinburgh **84A**, 185–196 (1979).
- [245] A. N. Tikhonov, *On the uniqueness of the problem of electric prospecting*, Doklady Akad. Nauk SSSR **69**, 797–800 (1949) (Russian).
- [246] E. Trubowitz, *The inverse problem for periodic potentials*, Commun. Pure Appl. Math. **30**, 321–337 (1977).
- [247] S. Venakides, *The infinite period limit of the inverse formalism for periodic potentials*, Commun. Pure Appl. Math. **41**, 3–17 (1988).
- [248] J. Weidmann, *Linear Operators in Hilbert Spaces*, Springer, New York, 1980.
- [249] J. Weidmann, *Spectral Theory of Ordinary Differential Operators*, Lecture Notes in Math. 1258, Springer, Berlin, 1987.

- [250] R. Weikard, *A local Borg–Marchenko theorem for difference equations with complex coefficients*, in *Partial Differential Equations and Inverse Problems*, C. Conca, R. Manásevich, G. Uhlmann, and M. S. Vogelius (eds.), *Contemp. Math.* **362**, 403–410 (2004).
- [251] X.-F. Yang, *A new inverse nodal problem*, *J. Diff. Eq.* **169**, 633–653 (2001).
- [252] Y. Zhang, *The One Dimensional Inverse Problem and New Integrable Dynamical Systems*, Ph.D. Thesis, Univ. of Rochester, NY, 2003.
- [253] Y. Zhang, *Solvability of a class of integro-differential equations and connections to one dimensional inverse problems*, *J. Math. Anal. Appl.* **321**, 286–298 (2006).

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MISSOURI, COLUMBIA, MO 65211, U.S.A.

E-mail address: fritz@math.missouri.edu

URL: <http://www.math.missouri.edu/personnel/faculty/gesztesyf.html>

Orthogonal Polynomials: From Jacobi to Simon

Leonid Golinskii and Vilmos Totik

Dedicated to Barry Simon on his 60th birthday

ABSTRACT. This is a survey of orthogonal polynomials of a single (real or complex) variable. It consists of two parts: the general theory and orthogonal polynomials on the unit circle. Some introductory material, some historical facts, some motivations and some heuristics are given, but the main emphasis is on the different aspects of the theory of orthogonal polynomials, on the major trends in it and on Barry Simon's contribution to the field.

CONTENTS

1. Introduction

Part 1. GENERAL THEORY

2. Orthogonal Polynomials
3. Classical Orthogonal Polynomials
4. Where Do Orthogonal Polynomials Come From?
5. Some Questions Leading to Classical Orthogonal Polynomials
6. Heuristics
7. General Orthogonal Polynomials
8. Strong and Ratio Asymptotics
9. Exponential and Freud Weights
10. Sobolev Orthogonality
11. Non-Hermitian Orthogonality
12. Multiple Orthogonality
13. Matrix Orthogonal Polynomials

Part 2. ORTHOGONAL POLYNOMIALS ON THE UNIT CIRCLE

14. Definitions and Basic Properties
15. Schur, Geronimus, Khrushchev
16. Szegő's Theory and Extensions

2000 *Mathematics Subject Classification.* 42C05, 33C47.

Key words and phrases. orthogonal polynomials, orthogonal polynomials on the unit circle.

17. CMV Matrices

Part 3. BARRY SIMON'S CONTRIBUTIONS

- 18. Analysis of CMV Matrices
- 19. Zeros
- 20. Spectral Theory in Special Classes
- 21. Periodic Verblunsky Coefficients
- 22. Random Coefficients
- 23. Miscellanea
- References

1. Introduction

Originally we were asked to write two separate papers: one on Barry Simon's work and one on the state of the art in the theory of orthogonal polynomials. However, Simon's work on orthogonal polynomials is so fresh and fundamental that it will be, for quite some time, the state of the art of the theory on the unit circle. This conflict could have only been resolved in a joint article.

This work is meant for non-experts, and therefore it contains some introductory material. We tried to list most of the actively researched fields, but because of space limitation, we have one or two pages for areas where dozens of papers and several books had been published. As a result our account is necessarily incomplete.

The connection of orthogonal polynomials with other branches of mathematics is truly impressive. Without even trying to be complete, we mention continued fractions, operator theory (Jacobi operators), moment problems, analytic functions (Bieberbach's conjecture), interpolation, Padé approximation, quadrature, approximation theory, numerical analysis, electrostatics, statistical quantum mechanics, special functions, number theory (irrationality and transcendence), graph theory (matching numbers), combinatorics, random matrices, stochastic processes (birth and death processes, prediction theory), data sorting and compression, Radon transform and computer tomography.

The theory of orthogonal polynomials can be divided into two main but only loosely related parts. The two parts have many things in common, and the dividing line is quite blurred—it is more or less along algebra vs. analysis. One of the parts is the algebraic aspect of the theory, which has close connections with special functions, combinatorics and algebra, and it is mainly devoted to concrete orthogonal systems or hierarchies of systems such as the Jacobi, Hahn, Askey–Wilson, etc. polynomials. All the discrete polynomials and the q -analogues of classical ones belong to this theory. We will not treat this part; the interested reader can consult three recent excellent monographs [51, 32, 4]. Much of the present state of the theory of orthogonal polynomials of several variables also lies close to this algebraic part of the theory. To discuss them would take us too far from our main direction, rather we refer the reader to the recent book [28].

The other part is the analytical aspect of the theory. Its methods are analytical, and it deals with questions that are typical in analysis, or questions that have emerged in and are related to other parts of mathematical analysis. General properties fill a smaller part of the analytic theory, and the greater part falls into

two main and extremely rich branches: orthogonal polynomials on the real line and on the circle. The richness is due to some special features of the real line and the circle. Classical real orthogonal polynomials, sometimes in other forms like continued fractions, can be traced back to the 18th century, but their rapid development occurred in the 19th and early 20th century. Orthogonal polynomials on the unit circle is much younger, and their existence is largely due to Szegő and Geronimus in the first half of the 20th century. Simon’s recent treatise [106, 107] (see also [100]) summarizes and greatly extends what has happened since then.

The organization of the present article is as follows. First, in Part 1, we give a brief outline of general and real orthogonal polynomials. Then we elaborate on some recent trends and the state of the art of this branch of the analytic theory. Simon’s contributions to real orthogonal polynomials will be mentioned in this part. After that, in Part 2, we move to orthogonal polynomials on the circle, and, finally, Part 3 lists many of Simon’s contributions.

Each of us has opted for his own style of exposition. Part 1, prepared by the second author, deals mostly with the state of the art in orthogonal polynomials—except orthogonal polynomials on the unit circle—and covers areas/results that are from a period of over a hundred years and from a large number of people; therefore the style there is somewhat informal. In contrast, Part 3 discusses mostly achievements of Barry Simon, and there, more formal statements are given.

Acknowledgements. Both authors participate in the project INTAS 03-51-6637 that supported a visit of the first author to Szeged, during which the outline of this paper was laid out. Research of the second author was supported by NSF grant DMS-040650 and OTKA T049448, TS44782.

Part 1. GENERAL THEORY

2. Orthogonal Polynomials

2.1. Orthogonal Polynomials With Respect to Measures. Let μ be a positive Borel measure on the complex plane with infinite support for which

$$\int |z|^m d\mu(z) < \infty$$

for all $m > 0$. There are unique polynomials

$$p_n(z) = p_n(\mu, z) = \kappa_n z^n + \dots, \quad \kappa_n > 0, \quad n = 0, 1, \dots$$

which form an orthonormal system in $L^2(\mu)$, i.e.,

$$\int p_m \overline{p_n} d\mu = \begin{cases} 0 & \text{if } m \neq n \\ 1 & \text{if } m = n. \end{cases}$$

These p_n ’s are called the orthonormal polynomials corresponding to μ . κ_n is the leading coefficient, and $p_n(z)/\kappa_n = z^n + \dots$ is called the monic orthogonal polynomial. The leading coefficients play a special and important role in the theory; many properties depend on their behavior. When $d\mu(x) = w(x)dx$ on some interval, say, then we talk about orthogonal polynomials with respect to the weight function w .

The p_n ’s can be easily generated: all we have to do is to make sure that

$$\int \frac{p_n(z)}{\kappa_n} \overline{z^k} d\mu(z) = 0, \quad k = 0, 1, \dots, n - 1,$$

which is an $n \times n$ system of equations for the coefficients of $p_n(z)/\kappa_n$ with matrix $(\sigma_{i,j})_{i,j=0}^{n-1}$, where

$$\sigma_{i,j} = \int z^i \overline{z^j} d\mu(z)$$

are the complex moments of μ . This matrix is non-singular, so the system has a unique solution, and finally κ_n comes from normalization.

In particular, the complex moments already determine the polynomials. In terms of them one can write up explicit determinantal formulae:

$$p_n(z) = \frac{1}{\sqrt{D_{n-1}D_n}} \begin{vmatrix} \sigma_{0,0} & \sigma_{0,1} & \cdots & \sigma_{0,n-1} & 1 \\ \sigma_{1,0} & \sigma_{1,1} & \cdots & \sigma_{1,n-1} & z \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \sigma_{n-1,0} & \sigma_{n-1,1} & \cdots & \sigma_{n-1,n-1} & z^{n-1} \\ \sigma_{n,0} & \sigma_{n,1} & \cdots & \sigma_{n,n-1} & z^n \end{vmatrix} \quad (2.1)$$

where

$$D_n = |\sigma_{i,j}|_{i,j=0}^n \quad (2.2)$$

are the so-called Gram determinants.

Note that if μ is supported on the real line then

$$\sigma_{i,j} = \int x^{i+j} d\mu(x) =: \alpha_{i+j},$$

so $D_n = |\alpha_{i+j}|_{i,j=0}^n$ is a Hankel determinant, while if μ is supported on the unit circle then

$$\sigma_{i,j} = \int z^{i-j} d\mu(z) =: \beta_{i-j},$$

so $D_n = |\beta_{i-j}|_{i,j=0}^n$ is a Toeplitz determinant. In these two important cases the orthogonal polynomials have many special properties that are missing in the general theory. For example, in the real case, i.e., if μ is supported on the real line, the p_n 's obey a three-term recurrence formula

$$xp_n(x) = a_n p_{n+1}(x) + b_n p_n(x) + a_{n-1} p_{n-1}(x), \quad (2.3)$$

where

$$a_n = \frac{\kappa_n}{\kappa_{n+1}} > 0, \quad b_n = \int xp_n^2(x) d\mu(x),$$

and conversely, any system of polynomials satisfying (2.3) with real $a_n > 0$, b_n is an orthonormal system with respect to a (not necessarily unique) measure on the real line (Favard's theorem). In the real case the zeros of p_n are real and simple and the zeros of p_n and p_{n+1} interlace, i.e., in between any two zeros of p_{n+1} , there is a zero of p_n . We emphasize that the three-term recurrence is a very special property of real orthogonal polynomials, and it is due to the fact that in this case the polynomials are real, hence

$$\int xp_n(x) \overline{p_m(x)} d\mu(x) = \int p_n(x) (xp_m(x)) d\mu(x) = 0$$

for $m < n - 1$. This three-term recurrence is missing in the general case, and it is replaced by a different recurrence for polynomials on the circle (see Part 2).

For example, in the real case the three-term recurrence implies for the reproducing kernel the Christoffel–Darboux formula

$$\sum_{k=0}^n p_k(x)p_k(t) = \frac{\kappa_n}{\kappa_{n+1}} \frac{p_{n+1}(x)p_n(t) - p_n(x)p_{n+1}(t)}{x - t}.$$

The starting values of the recurrence (2.3) are $p_{-1} \equiv 0$, $p_0 = (\mu(\mathbf{C}))^{-1/2}$. If one starts from $q_{-1} = -1$, $q_0 \equiv 0$ and use the same recurrence (with $a_{-1} = 1$)

$$xq_n(x) = a_n q_{n+1}(x) + b_n q_n(x) + a_{n-1} q_{n-1}(x), \tag{2.4}$$

then q_n is of degree $n - 1$, and by Favard’s theorem the different q_n ’s are orthogonal with respect to some measure. The q_n ’s are called orthogonal polynomials of the second kind (sometimes for p_n we say that they are of the first kind). They can also be written in the form

$$q_n(z) = (\mu(\mathbf{C}))^{-1/2} \int \frac{p_n(z) - p_n(x)}{z - x} d\mu(x).$$

2.2. The Riemann–Hilbert Approach. Let μ still be supported on the real line, and suppose that it is of the form $d\mu(t) = w(t)dt$ with some smooth function w . A new approach to generating orthogonal polynomials that has turned out to be of great importance was given in the early 1990’s by Fokas, Its and Kitaev [31]. Consider 2×2 matrices

$$Y(z) = \begin{pmatrix} Y_{11}(z) & Y_{12}(z) \\ Y_{21}(z) & Y_{22}(z) \end{pmatrix}$$

where the Y_{ij} are analytic functions on $\mathbf{C} \setminus \mathbf{R}$, and solve for such matrices the following matrix-valued Riemann–Hilbert problem:

1. For all $x \in \mathbf{R}$

$$Y_+(x) = Y_-(x) \begin{pmatrix} 1 & w(x) \\ 0 & 1 \end{pmatrix}$$

where Y_+ (resp. Y_-) is the limit of $Y(z)$ as z tends to x from the upper (resp. lower) half plane, and

- 2.

$$Y(z) = \left(I + O\left(\frac{1}{z}\right) \right) \begin{pmatrix} z^n & 0 \\ 0 & z^{-n} \end{pmatrix}$$

at infinity, where I denotes the identity matrix.

There is a unique solution $Y(z)$, and its entry $Y_{11}(z)$ is precisely the monic polynomial $p_n(\mu, z)/\kappa_n$. The other entries can also be explicitly written in terms of p_n and p_{n-1} . Furthermore, κ_n and the recurrence coefficients a_n, b_n can be expressed from the entries of Y_1 , where Y_1 is the matrix in

$$Y(z) \begin{pmatrix} z^{-n} & 0 \\ 0 & z^n \end{pmatrix} = I + Y_1 \frac{1}{z} + O\left(\frac{1}{z^2}\right).$$

For details on this Riemann–Hilbert approach, see [20] or [21] in this Festschrift.

2.3. Orthogonal Polynomials With Respect to Inner Products. Sometimes one talks about orthogonal polynomials with respect to an inner product $\langle \cdot, \cdot \rangle$ which is defined on some linear space containing all polynomials. In this case orthogonality means $\langle p_n, p_m \rangle = 0$ for $m \neq n$. When the inner product is positive definite in the sense that $\langle p, p \rangle = 0$ only for the zero polynomial p , then the aforementioned orthogonalization process can be used, and with $\sigma_{i,j} = \langle x^i, x^j \rangle$, the determinantal formula (2.1) is still valid. The same is true if the Gram determinants (2.2) are different from zero. However, if this is not so, e.g., in the so-called non-Hermitian orthogonality (see Section 11), then these cannot be used. In this case we write

$$p_n(z) = \gamma_n z^n + \gamma_{n-1} z^{n-1} + \dots,$$

and make sure that p_n is orthogonal to all powers z^k , $0 \leq k < n$, i.e., solve the homogeneous system of equations

$$\sum_{j=0}^n \gamma_j \sigma_{j,k} = 0, \quad k = 0, \dots, n - 1$$

for $\gamma_0, \gamma_1, \dots, \gamma_n$. Since the number of unknowns is bigger than the number of equations, there is always a non-trivial solution, which gives rise to non-trivial orthogonal polynomials. However, we cannot assert any more $\gamma_n \neq 0$, so the degree of p_n may be smaller than n , and there may be several choices for p_n . Still, in applications where non-Hermitian orthogonality is used, these p_n play the role of orthogonal polynomials.

2.4. Varying Weights. In the last twenty-five years orthogonal polynomials with respect to varying measures have played a significant role in several problems; see, e.g., the sections on exponential and Freud weights or on random matrices in Section 4. In forming them one has a sequence of measures μ_n (generally with some particular behavior), and for each n one forms the orthogonal system $\{p_k(\mu_n, z)\}_{k=0}^\infty$. In most cases one needs the behavior of $p_n(\mu_n, z)$ or that of $p_{n \pm k}(\mu_n, z)$ with some fixed k .

2.5. Matrix Orthogonal Polynomials. Orthogonality of matrix polynomials (i.e., when the entries of the fixed size matrix are polynomials of degree $n = 0, 1, \dots$ and orthogonality is with respect to a matrix measure) is a very active area which shows extreme richness compared to the scalar case. See Section 13 for a short discussion.

3. Classical Orthogonal Polynomials

These are

- Jacobi polynomials $P_n^{(\alpha, \beta)}$, $\alpha, \beta > -1$, orthogonal with respect to the weight $(1 - x)^\alpha(1 + x)^\beta$ on $[-1, 1]$
- Laguerre polynomials $L_n^{(\alpha)}$, $\alpha > -1$ with orthogonality weight $x^\alpha e^{-x}$ on $[0, \infty)$,
- Hermite polynomials H_n orthogonal with respect to e^{-x^2} on $(-\infty, \infty)$.

In the literature various normalizations are used for them.

They are very special, for they possess many properties that no other orthogonal polynomial system does. In particular,

- they have derivatives which form again an orthogonal polynomial system, e.g., the derivative of $P_n^{(\alpha,\beta)}$ is a constant multiple of $P_{n-1}^{(\alpha+1,\beta+1)}$:

$$(P_n^{(\alpha,\beta)})'(x) = \frac{1}{2} (n + \alpha + \beta + 1) P_{n-1}^{(\alpha+1,\beta+1)}(x),$$

- they all possess a Rodrigues-type formula

$$P_n(x) = \frac{1}{d_n w(x)} \frac{d^n}{dx^n} \{w(x)\sigma(x)^n\},$$

where w is the weight function and σ is a polynomial that is independent of n , e.g.,

$$L_n^{(\alpha)}(x) = e^x x^{-\alpha} \frac{1}{n!} \frac{d^n}{dx^n} (e^{-x} x^{n+\alpha}),$$

- they satisfy a differential-difference relation of the form

$$\pi(x)P_n'(x) = (\alpha_n x + \beta_n)P_n(x) + \gamma_n P_{n-1}(x),$$

e.g.,

$$x(L_n^{(\alpha)})'(x) = nL_n^{(\alpha)}(x) - (n + \alpha)L_{n-1}^{(\alpha)}(x),$$

- they satisfy a non-linear equation of the form

$$\sigma(x) (P_n(x)P_{n-1}(x))' = (\alpha_n x + \beta_n)P_n(x)P_{n-1}(x) + \gamma_n P_n^2(x) + \delta_n P_{n-1}^2(x),$$

with some constants $\alpha_n, \beta_n, \gamma_n, \delta_n$, and σ a polynomial of degree at most 2, e.g.,

$$(H_n(x)H_{n-1}(x))' = 2xH_n(x)H_{n-1}(x) - H_n^2(x) + 2nH_{n-1}^2(x).$$

Now every one of these has a converse, namely, if a system of orthogonal polynomials possesses any of these properties, then it is (up to a change of variables) one of the classical systems [3]. See also Bochner’s result in the next section claiming that the classical orthogonal polynomials are essentially the only polynomial (not just orthogonal polynomial) systems that satisfy a certain second-order differential equation.

Classical orthogonal polynomials are also special in the sense that they possess a relatively simple

- second-order differential equation, e.g.,

$$xy'' + (\alpha + 1 - x)y' + ny = 0$$

for $L_n^{(\alpha)}$

- generating function, e.g.,

$$\sum_n \frac{H_n(x)}{n!} w^n = \exp(2xw - w^2)$$

- integral representation, e.g.,

$$(1 - x)^\alpha (1 + x)^\beta P_n^{(\alpha,\beta)}(x) = \frac{(-1)^n}{2^{n+1}\pi i} \int (1 - t)^{n+\alpha} (1 + t)^{n+\beta} (t - x)^{-n-1} dt$$

over an appropriate contour,

and these are powerful tools to study their behavior. For all these results, see [122].

4. Where Do Orthogonal Polynomials Come From?

In this section we mention a few selected areas where orthogonal polynomials naturally arise.

4.1. Continued Fractions. Continued fractions played an extremely important role in the development of several branches of mathematics, but their significance has been unjustly diminished in modern mathematics. A continued fraction is of the form

$$\frac{B_1}{A_1 + \frac{B_2}{A_2 + \dots}},$$

and its n -th convergent is

$$\frac{S_n}{R_n} = \frac{B_1}{A_1 + \frac{B_2}{A_2 + \dots + \frac{B_n}{A_n}}}, \quad n = 1, 2, \dots$$

The value of the continued fraction is the limit of its convergents. The denominators and numerators of the convergents satisfy the three-term recurrence relations

$$\begin{aligned} R_n &= A_n R_{n-1} + B_n R_{n-2}, & R_0 &\equiv 1, R_{-1} \equiv 0 \\ S_n &= A_n S_{n-1} + B_n S_{n-2}, & S_0 &\equiv 0, S_{-1} \equiv 1, \end{aligned}$$

which immediately connects continued fractions with three-term recurrences and hence with orthogonal polynomials.

But the connection is deeper than just this formal observation. Many elementary functions (like $z - \sqrt{z^2 - 1}$) have a continued fraction development where the B_n 's are constants while the A_n 's are linear functions, in which case the convergents are ratios of some orthogonal polynomials of the second and first kind. An important example is that of Cauchy transforms of measures μ with compact support on the real line (the so-called Markov functions):

$$f(z) = \int \frac{d\mu(x)}{x - z} = -\frac{\alpha_0}{z} - \frac{\alpha_1}{z^2} - \dots \tag{4.1}$$

The coefficients α_j in the development of (4.1) are the moments

$$\alpha_j = \int x^j d\mu(x), \quad j = 0, 1, \dots$$

of the measure μ . The continued fraction development

$$f(z) \sim \frac{B_1}{z - A_1 + \frac{B_2}{z - A_2 + \dots}}$$

of f at infinity converges locally uniformly outside the smallest interval that contains the support of μ (Markov's theorem).

As has been mentioned, the numerators $S_n(z)$ and the denominators $R_n(z)$ of the n -th convergents

$$\frac{S_n(z)}{R_n(z)} = \frac{B_1}{z - A_1 + \frac{B_2}{z - A_2 + \dots + \frac{B_n}{z - A_n}}}, \quad n = 1, 2, \dots$$

satisfy the recurrence relations

$$\begin{aligned} R_n(z) &= (z - A_n)R_{n-1}(z) + B_n R_{n-2}(z), & R_0 &\equiv 1, R_{-1} \equiv 0 \\ S_n(z) &= (z - A_n)S_{n-1}(z) + B_n S_{n-2}(z), & S_0 &\equiv 0, S_{-1} \equiv 1. \end{aligned} \tag{4.2}$$

These are precisely the recurrence formulae for the monic orthogonal polynomials of the first and second kind with respect to μ . Hence the n -th convergent is $cq_n(z)/p_n(z)$ with $c = \mu(\mathbf{C})^{1/2}$. See [122, pp. 54–57] as well as [56] and the numerous references there.

4.2. Padé Approximation and Rational Interpolation. With the preceding notation, the rational function $S_n(z)/R_n(z) = cq_n(z)/p_n(z)$ with $c = \mu(\mathbf{C})^{1/2}$ of numerator degree $n - 1$ and of denominator degree n interpolates $f(z)$ at infinity in the order $2n$. This is the analogue (called $[n - 1/n]$ Padé approximation) of the n -th Taylor polynomial (which interpolates the function in the order n) for rational functions. The advantage of Padé approximation over Taylor polynomials lies in the fact that the poles of Padé approximants may imitate the singularities of the function in question, while Taylor polynomials are good only up to the first singularity. The error in $[n - 1/n]$ Padé approximation has the form

$$f(z) - c \frac{q_n(z)}{p_n(z)} = \frac{1}{p_n^2(z)} \int \frac{p_n^2(x)}{x - z} d\mu(x).$$

Orthogonal polynomials appear in more general rational interpolation (called multipoint Padé approximation) to Markov functions; see, e.g., [117, Sect. 6.1].

4.3. Moment Problem. The moments of a measure μ , $\mu(\mathbf{C}) = 1$, supported on the real line, are

$$\alpha_n = \int x^n d\mu(x), \quad n = 0, 1, \dots$$

The Hamburger moment problem is to determine if a sequence $\{\alpha_n\}$ (with normalization $\alpha_0 = 1$) of real numbers is the moment sequence of a measure with infinite support, and if this measure is unique (the Stieltjes moment problem asks the same, but for measures on $[0, \infty)$). The existence is easy: $\{\alpha_n\}$ are the moments of some measure supported on \mathbf{R} if and only if all the Hankel determinants $|\alpha_{i+j}|_{i,j=0}^m$, $m = 0, 1, \dots$ are positive. The unicity (usually called determinacy) depends on the behavior of the orthogonal polynomials (2.1) defined from the moments $\sigma_{i,j} = \alpha_{i+j}$. In fact, there are different measures with the same moments α_j if and only if there is a non-real z_0 with $\sum_n |p_n(z_0)|^2 < \infty$, which in turn is equivalent to $\sum_n |p_n(z)|^2 < \infty$ for all $z \in \mathbf{C}$. Furthermore, the Cauchy transforms of all solutions ν of the moment problem have the parametric form

$$\int \frac{d\nu(x)}{z - x} = \frac{C(z)F(z) + A(z)}{D(z)F(z) + B(z)},$$

where F is an arbitrary analytic function (the parameter) mapping the upper half plane \mathbf{C}_+ into $\overline{\mathbf{C}}_+ \cup \{\infty\}$, and A, B, C and D have explicit representation in terms of the first and second kind orthogonal polynomials p_n and q_n :

$$\begin{aligned} A(z) &= z \sum_n q_n(0)q_n(z); & B(z) &= -1 + z \sum_n q_n(0)p_n(z); \\ C(z) &= 1 + z \sum_n p_n(0)q_n(z); & D(z) &= z \sum_n p_n(0)p_n(z). \end{aligned}$$

For all these results and for an operator theoretic approach to the moment problem, see Simon’s survey [96] (in particular, Theorems 3 and 4.14).

4.4. Jacobi Matrices and Spectral Theory of Self-Adjoint Operators.

Tridiagonal, so-called Jacobi matrices

$$J = \begin{pmatrix} b_0 & a_0 & 0 & 0 & \cdots \\ a_0 & b_1 & a_1 & 0 & \cdots \\ 0 & a_1 & b_2 & a_2 & \cdots \\ 0 & 0 & a_2 & b_2 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

with bounded $a_n > 0$ and bounded real b_n define a self-adjoint bounded operator J in l_2 , a so-called Jacobi operator. These are the discrete analogue of second-order linear differential operators of Schrödinger type on the half line. Every bounded self-adjoint operator with a cyclic vector is a Jacobi operator in an appropriate base.

The formal eigenequation $J\pi = \lambda\pi$ is equivalent to the three-term recurrence

$$\begin{aligned} a_{n-1}\pi_{n-1} + b_n\pi_n + a_n\pi_{n+1} &= \lambda\pi_n, & n = 1, 2, \dots \\ b_0\pi_0 + a_0\pi_1 &= \lambda\pi_0, & \pi_0 = 1. \end{aligned}$$

Thus, $\pi_n(\lambda)$ is of degree n in λ .

By the spectral theorem, J , as a self-adjoint operator having a cyclic vector $((1, 0, 0, \dots))$, is unitarily equivalent to multiplication by x in some $L^2(\mu)$ with μ having compact support on the real line. More precisely, if $p_n(x) = p_n(\mu, x)$ are the orthonormal polynomials with respect to μ , and U maps the unit vector $e_n = (0, \dots, 0, 1, 0, \dots)$ into p_n , then U can be extended into a unitary operator from l_2 onto $L^2(\mu)$, and if $Sf(x) = xf(x)$ is the multiplication operator by x in $L^2(\mu)$, then $J = U^{-1}SU$. The recurrence coefficients for $p_n(\mu, x)$ are precisely the a_n 's and b_n 's from the Jacobi matrix, i.e., $p_n(x) = cp_n(x)$ with some fixed constant c . These show that Jacobi operators are equivalent to multiplication by x in $L^2(\mu)$ spaces if the particular basis $\{p_n(\mu)\}$ is used (see, e.g., [21, Ch. 2]).

The truncated $n \times n$ matrix

$$J_n = \begin{pmatrix} b_0 & a_0 & 0 & 0 & \cdots \\ a_0 & b_1 & a_1 & 0 & \cdots \\ 0 & a_1 & b_2 & a_2 & \cdots \\ \vdots & \vdots & \vdots & \ddots & \cdots \\ 0 & 0 & 0 & a_{n-2} & b_{n-1} \end{pmatrix}$$

has n real and distinct eigenvalues, which turn out to be the zeros of p_n , i.e., the monic polynomial $p_n(z)/\kappa_n$ is the characteristic polynomial of J_n .

4.5. Quadrature. For a measure μ , an n -point quadrature is a set of points

x_1, \dots, x_n and a set of associated numbers $\lambda_1, \dots, \lambda_n$. It is expected that

$$\int f d\mu \sim \sum_{k=1}^n \lambda_k f(x_k)$$

in some sense for as large class of functions as possible. Often the accuracy of the quadrature is measured by its exactness, which is defined as the largest m such that the quadrature is exact for all polynomials of degree at most m , i.e., m is the

largest number with the property that

$$\int x^j d\mu(x) = \sum_{k=1}^n \lambda_k x_k^j \quad \text{for all } 0 \leq j \leq m.$$

For μ with support on the real line and for quadrature based on n points, this exactness m cannot be larger than $2n - 1$, and this optimal value $2n - 1$ is attained if and only if x_1, \dots, x_n are the zeros of the orthonormal polynomial $p_n(\mu, x)$ corresponding to μ , and the so-called Cotes numbers λ_k are chosen to be

$$\lambda_k = \left(\sum_{j=0}^n p_j(\mu, x_k)^2 \right)^{-1}.$$

See [122, Ch. XV].

4.6. Random Matrices. Some statistical-mechanical models in quantum systems use random matrices. Let \mathcal{H}_n be the set of all $n \times n$ Hermitian matrices $M = (m_{i,j})_{i,j=1}^n$, and let there be given a probability distribution on \mathcal{H}_n of the form

$$P_n(M)dM = D_n^{-1} \exp(-n\text{Tr}\{V(M)\})dM,$$

where $V(\lambda)$, $\lambda \in \mathbf{R}$, is a real-valued function that increases sufficiently fast at infinity (typically an even polynomial in quantum field theory applications), $\text{Tr}\{H\}$ denotes the trace of the matrix H ,

$$dM = \prod_{k=1}^n dm_{k,k} \prod_{k < j} d\Re m_{k,j} d\Im m_{k,j}$$

is the ‘‘Lebesgue’’ measure for Hermitian matrices, and D_n is a normalizing constant so that the total integral of $P_n(M)dM$ is one.

Every matrix $M \in \mathcal{H}_n$ has n real eigenvalues which carry physical information on the system when it is in the state described by M . The quantity

$$N_n(\mathbf{D}) = \frac{\#\{\text{eigenvalues in } \mathbf{D}\}}{n}$$

is the random variable that equals the normalized number of eigenvalues in the interval \mathbf{D} . This model is known as the unitary ensemble associated with V .

Let $p_j(w^n, x)$ be the orthonormal polynomials with respect to the varying weight $w^n(x)$, $w(x) = \exp(-V(x))$. Then the joint probability density of the eigenvalues can be written in the form

$$d_n \left| p_{i-1}(w^n, \lambda_j) w^{n/2}(\lambda_j) \right|_{1 \leq i, j \leq n}^2$$

where d_n is a normalizing constant built up from the leading coefficients of the $p_j(w^n, \cdot)$. With the so-called weighted reproducing kernel

$$K_n(t, s) = \sum_{j=0}^{n-1} p_j(w^n, t) w^{n/2}(t) p_j(w^n, s) w^{n/2}(s),$$

it can also be written in the form

$$\frac{1}{n!} |K_n(\lambda_i, \lambda_j)|_{1 \leq i, j \leq n}.$$

In particular, for the expected number of eigenvalues in an interval \mathbf{D} , we have

$$EN_n(\mathbf{D}) = \int_{\mathbf{D}} \frac{K_n(\lambda, \lambda)}{n} d\lambda,$$

where $1/K_n(\lambda, \lambda)$ is known in the theory of orthogonal polynomials as the n -th (weighted) Christoffel function associated with the weight w^n , while the limit of the left-hand side (as $n \rightarrow \infty$) is known as the density of states. See, e.g., [77] and [86].

5. Some Questions Leading to Classical Orthogonal Polynomials

There are almost an infinite number of problems where classical orthogonal polynomials emerge. Let us just mention a few.

5.1. Electrostatics. Put to 1 and -1 two positive charges p and q , and with these fixed charges put n positive unit charge on $[-1, 1]$ to the points x_1, \dots, x_n . On the plane the Coulomb force is proportional with the reciprocal of the distance, and so a charge generates a logarithmic potential field. Therefore, the mutual energy of all these charges is

$$I(x_1, \dots, x_n) = p \sum_{j=1}^n \log \frac{1}{|1 - x_j|} + q \sum_{j=1}^n \log \frac{1}{|1 + x_j|} + \sum_{i < j} \log \frac{1}{|x_i - x_j|},$$

and the equilibrium problem asks for finding x_1, \dots, x_n for which this energy is minimal. The unique minimum occurs (see [122, Sect. 6.7]) for the zeros of the Jacobi polynomial $P_n^{(2p-1, 2q-1)}$ orthogonal with respect to the weight $(1-x)^{2p-1}(1+x)^{2q-1}$.

There is a similar characterization of the zeros of Laguerre and Hermite polynomials, and even of more general orthogonal polynomials (for the latter, see [51, Sect. 3.5]).

5.2. Polynomial Solutions of Eigenvalue Problems. Consider the eigenvalue problem,

$$f(x) \frac{d^2}{dx^2} y(x) + g(x) \frac{d}{dx} y(x) + h(x)y(x) = \lambda y(x),$$

where f, g, h are fixed polynomials and λ is a free constant, and it is required that this has a polynomial solution of exact degree n for all $n = 0, 1, \dots$, for which the corresponding λ and $y(x)$ will be denoted by λ_n and $y_n(x)$, respectively. Bochner's theorem [14] states that except for some trivial solutions of the form $y(x) = ax^n + bx^m$ and for some polynomials related to Bessel functions, the only solutions are (in all of them we can take $h(x) = 0$):

- Jacobi polynomials $P_n^{(\alpha, \beta)}$ ($f(x) = 1 - x^2$, $g(x) = \beta - \alpha - x(\alpha + \beta + 2)$, $\lambda_n = -n(n + \alpha + \beta + 1)$)
- Laguerre polynomials $L_n^{(\alpha)}$ ($f(x) = x$, $g(x) = 1 + \alpha - x$, $\lambda_n = -n$) and
- Hermite polynomials $H_n(x)$ ($f(x) = 1$, $g(x) = -2x$, $\lambda_n = -2n$).

5.3. Harmonic Analysis on Spheres and Balls. Harmonic analysis on spheres and balls in \mathbf{R}^d is based on spherical harmonics, i.e., harmonic homogeneous polynomials. In this theory special Jacobi polynomials, so-called ultraspherical or Gegenbauer polynomials $P_n^{(\alpha)}$, play a fundamental role—they are orthogonal with respect to the weight $(1 - x^2)^{\alpha-1/2}$.

Let S^{d-1} be the unit sphere in \mathbf{R}^d and let \mathcal{H}_n^d be the restriction to S^{d-1} of all harmonic polynomials $Q(x_1, \dots, x_n)$ of d variables that are homogeneous of degree n , i.e.,

$$\sum_{k=1}^n \frac{\partial^2}{\partial x_k^2} Q = 0, \quad Q(\lambda x_1, \dots, \lambda x_n) = \lambda^n Q(x_1, \dots, x_n), \quad \lambda > 0.$$

The dimension of \mathcal{H}_n^d is

$$\binom{n+d-1}{d-1} - \binom{n+d-3}{d-1},$$

and an orthogonal basis in it can be produced as follows. With $\rho = x_{d-1}^2 + x_d^2$, let $g_{s,0} = \rho^s P_s^{(0)}(x_{d-1}/\rho)$ and $g_{s,1} = x_d \rho^s P_s^{(1)}(x_{d-1}/\rho)$. With $n_d = 0$ or $n_d = 1$, consider all multi-indices $\mathbf{n} = (n_1, n_2, \dots, n_d)$ such that $n_1 + \dots + n_d = n$, and if for such a multi-index we define

$$Y_{\mathbf{n}}(x_1, \dots, x_d) = g_{n_{d-1}, n_d} \prod_{j=1}^{d-2} \left((x_j^2 + \dots + x_d^2)^{n_j} P_{n_j}^{(\lambda_j)}(x_j(x_j^2 + \dots + x_d^2)^{-1/2}) \right),$$

then these $Y_{\mathbf{n}}$ constitute an orthogonal basis in \mathcal{H}_n^d (see, e.g., [28, p. 35]).

5.4. Approximation Theory. In the literature, expansions of functions into classical orthogonal polynomial series are second only to trigonometric expansions, and numerous works have been devoted to their convergence and approximation properties; see, e.g., [122, Ch. XIII].

The Chebyshev polynomials given by $\cos(n \arccos x)$ are orthogonal on $[-1, 1]$ with respect to the weight $w(x) = (1 - x^2)^{-1/2}$. These directly correspond to trigonometric functions, and expansions into them have virtually the same properties as trigonometric Fourier expansions. But there are many other aspects of approximation where Chebyshev polynomials appear. If one considers, e.g., the best approximation on $[-1, 1]$ of x^n by polynomials $P_{n-1}(x)$ of smaller degree, then the smallest error appears when $x^n - P_{n-1}(x) = 2^{1-n} \cos(n \arccos x)$ is the monic n -th Chebyshev polynomial.

The monic orthogonal polynomials $p_n(\mu)/\kappa_n$ are the solutions to the extremal problem

$$\int |P_n|^2 d\mu \rightarrow \min, \tag{5.1}$$

where the minimum is taken for all monic polynomials of degree n . This extremal property makes orthogonal polynomials, particularly Chebyshev polynomials, indispensable tools in approximation theory.

Lagrange interpolation and its various generalizations like Hermite, Hermite-Fejér interpolation, etc. is mostly done on the zeros of some orthogonal polynomials. In fact, these nodes are often close to optimal in the sense that the Lebesgue constant increases in the optimal rate. In many cases interpolation on zeros of

orthogonal polynomials have special properties due to explicitly calculable expressions. Recall, e.g., Fejér’s result that if P_{2n-1} is the unique polynomial of degree at most $2n - 1$ that interpolates a continuous function f at the nodes of the n -th Chebyshev polynomial and that has zero derivative at each of these nodes, then P_{2n-1} uniformly converges to f on $[-1, 1]$ as $n \rightarrow \infty$. For the role of orthogonal polynomials in interpolation, see [121] and [72].

6. Heuristics

In this section we do not state precise results; we just want to indicate some heuristics on the behavior of orthogonal polynomials. For the concepts below, as well as for a more precise form of some of the heuristics, see the following sections, in particular, Section 7.

As we have just seen, the monic orthogonal polynomials $p_n(\mu)/\kappa_n$ minimize the $L^2(\mu)$ norm in (5.1). Therefore, the polynomials try to be small where the measure is large, e.g., one expects the zeros to cluster at the support $S(\mu)$ of μ . The example of arc measure on the unit circle, for which the orthogonal polynomials are z^n , shows, however, that this is not true (due to the fact that the complement of the support is not connected). The statement is true when the support lies on \mathbf{R} or on some systems of arcs, and also in the general case when, instead of the support, one considers the polynomial convex hull of the support of μ : on any compact set outside the polynomial convex hull, there can only be a fixed number of zeros of $p_n(\mu)$ for every n . When the complement of $S(\mu)$ is connected and $S(\mu)$ has no interior, then the distribution of the zeros shows a remarkable universality and indifference with respect to the size of μ . In many situations the distribution of the zeros is the equilibrium distribution of the support $S(\mu)$. When $S(\mu) = [-1, 1]$, this means that under very weak assumptions, the zero distribution is always the arcsine distribution $dx/\pi\sqrt{1 - x^2}$.

The $L^2(\mu)$ minimality of $p_n(\mu)/\kappa_n$ in the sense of (5.1) is something like minimality in L^∞ norm on $S(\mu)$. Therefore, $p_n(\mu)/\kappa_n$ should behave like the monic polynomial T_n minimizing the L^∞ norm on $S(\mu)$ (the so-called Chebyshev polynomials for $S(\mu)$). Since

$$\frac{1}{n} \log |T_n(z)| = \int \log |z - t| d\nu_n(t)$$

where ν_n has mass $1/n$ at each zero of T_n , in the limit the behavior should be like

$$U^\nu(z) = \int \log |z - t| d\nu(t), \tag{6.1}$$

where ν is the probability measure on $S(\mu)$ for which the maximum of U^ν on $S(\mu)$ is as small as possible (this is the so-called equilibrium measure of $S(\mu)$). More generally, if $d\nu = d\nu_n = w^n(x)dx$ is a varying weight in the specified way, then the same reasoning leads to a behavior like (6.1), but now ν is a measure for which the supremum of $U^\nu(z) + \log w(z)$ is as small as possible (weighted equilibrium measure).

Universal behavior can also be seen for the polynomials themselves. Usually they obey

$$\frac{1}{n} \log |p_n(\mu, z)| \rightarrow g_{\mathbf{C} \setminus S(\mu)}(z, \infty), \quad z \notin S(\mu) \tag{6.2}$$

where $g_{\mathbf{C} \setminus S(\mu)}(z, \infty)$ is the Green’s function with pole at infinity associated with the complement of the support. When the unbounded component of the complement of $S(\mu)$ is simply connected, then in that component, often there is a finer asymptotic behavior of $p_n(\mu)$ of the form

$$p_n(z) \sim d_n g_\mu(z) \Phi(z)^n, \quad z \notin S(\mu) \tag{6.3}$$

where Φ is the mapping function that maps $\mathbf{C} \setminus S(\mu)$ conformally onto the outside of the unit disk, and g_μ is a function (might be called the generalized Szegő function) that depends on μ . Such a fine asymptotic is restricted to the simply connected case; see, e.g., Section 8.

Asymptotics on orthogonal polynomials have a hierarchy, and the different types of asymptotics usually require the measure to be sufficiently strong with different degree on its support. Consider first the case of compact support $S(\mu)$. The weakest is n -th root asymptotics stating the behavior (6.2) for $|p_n(\mu, z)|^{1/n}$ outside the support of the measure. It is mostly equivalent with a corresponding distribution of the zeros, as well as asymptotical minimal behavior of $\kappa_n^{1/n}$. It holds under very weak assumption on the measure, roughly stating that the logarithmic capacity where $\mu' > 0$ (derivative with respect to equilibrium measure) be the same as the capacity of $S(\mu)$. Ratio asymptotics, i.e., asymptotic behavior of $p_{n+1}(\mu, z)/p_n(\mu, z)$, is stronger, and is equivalent with asymptotics for the ratio κ_{n+1}/κ_n of consecutive leading coefficients. It can only hold when $\mathbf{C} \setminus S(\mu)$ (more precisely, its unbounded component) is simply connected, and in this case it is enough that $\mu' > 0$ almost everywhere with respect to the equilibrium measure of the support of μ (see Section 8). Finally, strong asymptotics of the form (6.3) needs roughly that $\log \mu'$ be integrable (Szegő condition, see Section 8).

All these are outside the support. On the support the orthogonal polynomials are of oscillatory behavior, and in the real case under smoothness assumptions on the measure, often a so-called Plancherel–Rotach-type asymptotic formula

$$p_n(\mu, x) \sim d_n g(x) \sin(nh(x) + H(x))$$

holds, where g, h, H are fixed functions. Here $h(x)$ is directly linked with the zeros, h'/π is precisely the distribution of the zeros. When $S(\mu) = [-1, 1]$ and the measure is smooth, then $h(x) = \arccos x$.

When $S(\mu)$ is not of compact support (like Laguerre, Hermite or Freud weights), then usually the zeros are spreading out, and one has to rescale them to $[-1, 1]$ (or to $[0, 1]$) to get a distribution, which is mostly *not* the arcsine distribution. In similar fashion, various asymptotics hold for the polynomials only after the corresponding rescaling.

7. General Orthogonal Polynomials

In this section μ is always of compact support $S(\mu)$. For all the results below, see [117] and the references there.

The energy $V(K)$ of a compact set K is defined as the infimum of

$$I(\nu) = \iint \log \frac{1}{|x - t|} d\nu(x) d\nu(t) \tag{7.1}$$

where the infimum is taken for all positive Borel measures on K with total mass 1. The logarithmic capacity is then $\text{cap}(K) = e^{-V(K)}$. For the leading coefficients κ_n

of the orthonormal polynomials $p_n(\mu)$, we have

$$\frac{1}{\text{cap}(S(\mu))} \leq \liminf_{n \rightarrow \infty} \kappa_n^{1/n}. \tag{7.2}$$

When $\text{cap}(K)$ is positive, then there is a unique measure $\nu = \omega_K$ minimizing the energy in (7.1), and this measure is called the equilibrium measure of K . The Green’s function $g_{\mathbf{C} \setminus K}(z, \infty)$ with pole at infinity of $\mathbf{C} \setminus K$ can then be defined as

$$g_{\mathbf{C} \setminus K}(z, \infty) = \log \frac{1}{\text{cap}(K)} - \int \log \frac{1}{|z - t|} d\omega_K(t). \tag{7.3}$$

We have for all μ (with $\text{cap}(S(\mu)) > 0$) the estimate

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log |p_n(\mu, z)|^{1/n} \geq g_{\mathbf{C} \setminus S(\mu)}(z, \infty) \tag{7.4}$$

locally uniformly outside the convex hull of $S(\mu)$, while in the convex hull but outside the polynomial convex hull $\text{Pc}(S(\mu))$ (see below), (7.4) is true quasi-everywhere (i.e., with the exception of a set of zero capacity). The same is true on the outer boundary of $S(\mu)$, which is defined as the boundary $\partial\Omega$ of the unbounded component Ω of the complement $\mathbf{C} \setminus S(\mu)$, namely, for quasi-every $z \in \partial\Omega$,

$$\liminf_{n \rightarrow \infty} |p_n(\mu, z)|^{1/n} \geq 1.$$

All these estimates are sharp.

The zeros of $p_n(\mu)$ lie in the convex hull of $S(\mu)$. When Ω is the unbounded component of the complement $\mathbf{C} \setminus S(\mu)$, then $\text{Pc}(S(\mu)) = \mathbf{C} \setminus \Omega$ is called the polynomial convex hull of $S(\mu)$ (it is the union of $S(\mu)$ with all the “holes” in it, i.e., with the bounded components of $\mathbf{C} \setminus S(\mu)$). Now the zeros cluster on $\text{Pc}(S(\mu))$ in the sense that for any compact subset K of Ω , there is a number N_K independent of n , such that $p_n(\mu)$ can have at most N_K zeros in K . For example, if μ is supported on the real line, then $\text{Pc}(S(\mu)) = S(\mu)$, and if K is a closed interval disjoint from the support, then there is at most one zero in K . Denisov and Simon [26] showed that if $x_0 \in \mathbf{R}$ is not in the support, then for some $\delta > 0$ and all n , either p_n or p_{n+1} has no zero in $(x_0 - \delta, x_0 + \delta)$. Note that if μ is a symmetric measure on $[-1, -\frac{1}{2}] \cup [\frac{1}{2}, 1]$, then $p_{2n+1}(0) = 0$ for all n , so the result is sharp.

In [26] Denisov and Simon focused on attracting properties of isolated points of the supp μ . Let z_0 be an isolated point of $S(\mu)$, such that its distance from the convex hull of $S(\mu) \setminus \{z_0\}$ is $\delta > 0$. Then p_n has at most one zero in the disk $\{|z - z_0| < \delta/3\}$. It is also clear that for any symmetric measure μ with $S(\mu) = [-1, -\frac{1}{2}] \cup \{0\} \cup [\frac{1}{2}, 1]$, the polynomials $p_{2n}(\mu)$ have two zeros near 0 (in this case $\delta = 0$). Moreover, if μ lies on the unit circle, then there exist two positive constants C and a and a zero z_n of p_n such that $|z_n - z_0| \leq Ce^{-an}$.

Next put a unit mass to every zero of $p_n(\mu)$ (counting multiplicity), this gives the so-called counting measure $\nu_{p_n(\mu)}$ on the zero set. Zero distribution amounts to finding the limit behavior of $\frac{1}{n} \nu_{p_n(\mu)}$. The normalized arc measure on the unit circle (for which $p_n(\mu, z) = z^n$) shows that if the interior of the polynomial convex hull $\text{Pc}(S(\mu))$ is not empty, then the zeros may be far away from the outer boundary $\partial\Omega$, where the equilibrium measure $\omega_{S(\mu)}$ is supported. Thus, assume that $\text{Pc}(S(\mu))$ has empty interior and also that there is no Borel set of capacity zero and full μ -measure (the case when this is not true is rather pathological; almost anything can happen

with the zeros then). In this case,

$$\lim_{n \rightarrow \infty} \kappa_n^{1/n} = \log \frac{1}{\text{cap}(S(\mu))} \tag{7.5}$$

if and only if

$$\lim \frac{1}{n} \nu_{p_n(\mu)} = \omega_{S(\mu)}$$

in weak * sense, i.e., asymptotically minimal behavior of $\kappa_n^{1/n}$ (see (7.2)) is equivalent to the fact that the zero distribution is the equilibrium distribution.

(7.5) is called regular limit behavior, and in this case we write $\mu \in \mathbf{Reg}$. Thus, the important class **Reg** is defined by the property (7.5). If Ω is a regular set with respect to the Dirichlet problem, then $\mu \in \mathbf{Reg}$ is equivalent to either of

- $\lim_{n \rightarrow \infty} \|p_n(\mu)\|_{\text{sup}, S(\mu)}^{1/n} = 1$
- For any sequence $\{P_n\}$ of polynomials of degree $n = 1, 2, \dots$

$$\lim_{n \rightarrow \infty} \left(\frac{\|P_n\|_{\text{sup}, S(\mu)}}{\|P_n\|_{L^2(\mu)}} \right)^{1/n} = 1.$$

The last statement expresses the fact that in n -th root sense the $L^2(\mu)$ and L^∞ norms (on $S(\mu)$) are asymptotically the same.

All equivalent formulations of $\mu \in \mathbf{Reg}$ point to a certain “thickness” of μ on its support. Regularity is an important property, and it is desirable to know “thickness” conditions under which it is true. Several regularity criteria are known, e.g., either of the conditions

- all Borel sets $B \subseteq S(\mu)$ with full measure (i.e with $\mu(B) = \mu(S(\mu))$) have capacity $\text{cap}(B) = \text{cap}(S(\mu))$
- $d\mu/d\omega_{S(\mu)} > 0$ (Radon–Nikodym derivative) $\omega_{S(\mu)}$ -almost everywhere

is sufficient for $\mu \in \mathbf{Reg}$. Regularity holds under fairly weak assumptions on the measure, e.g., if $S(\mu) = [0, 1]$, and

$$\liminf_{r \rightarrow 0} r \log \mu([x - r, x + r]) \geq 0$$

for almost every $x \in [0, 1]$ (i.e., if μ is not exponentially small around almost every point), then $\mu \in \mathbf{Reg}$.

No necessary and sufficient condition for regularity in terms of the size of the measure μ is known. The only existing necessary condition is for the case $S(\mu) = [0, 1]$, and it reads that for every $\eta > 0$,

$$\lim_{n \rightarrow \infty} \text{cap}(\{x \mid \mu([x - 1/n, x + 1/n]) \geq e^{-\eta n}\}) = \frac{1}{4}$$

(here $\frac{1}{4}$ is the capacity of $[0, 1]$).

8. Strong and Ratio Asymptotics

Let μ be supported on $[-1, 1]$ and suppose that the so-called Szegő condition

$$\int_{-1}^1 \frac{\log \mu'(t)}{\sqrt{1-t^2}} dt > -\infty \tag{8.1}$$

holds, where μ' is the Radon–Nikodym derivative of μ with respect to linear Lebesgue measure. Note that this condition means that the integral is finite, for it cannot be ∞ . It expresses a certain denseness of μ , and under this condition, Szegő proved several asymptotics for the corresponding orthonormal polynomials $p_n(\mu)$. This

theory was developed on the unit circle and then was translated into the real line. The Szegő function associated with μ is

$$D_\mu(z) = \exp \left(\sqrt{z^2 - 1} \frac{1}{2\pi} \int_{-1}^1 \frac{\log \mu'(t)}{z - t} \frac{dt}{\sqrt{1 - t^2}} \right) \tag{8.2}$$

and it is the outer function in the Hardy space on $\mathbf{C} \setminus [-1, 1]$ with boundary values $|D_\mu(x)|^2 = \mu'(x)$. Outside $[-1, 1]$ the asymptotic formula

$$p_n(\mu, z) = (1 + o(1)) \frac{1}{\sqrt{2\pi}} (z + \sqrt{z^2 - 1})^n D_\mu(z)^{-1} \tag{8.3}$$

holds locally uniformly; in particular, the leading coefficient κ_n of $p_n(\mu)$ is of the form

$$\kappa_n = (1 + o(1)) \frac{2^n}{\sqrt{2\pi}} \exp \left(\frac{-1}{2\pi} \int_{-1}^1 \frac{\log \mu'(t)}{\sqrt{1 - t^2}} dt \right). \tag{8.4}$$

For all these results, see [122, Ch. 6]. The Szegő condition is also necessary for these results, e.g., an asymptotic formula like (8.3) or (8.4) is equivalent to (8.1).

If one assumes weaker conditions, then necessarily weaker results will follow. A large and important class of measures is the Nevai class $M(b, a)$ (see [80]), for which the coefficients in the three-term recurrence

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

satisfy $a_n \rightarrow a, b_n \rightarrow b$. This is equivalent to ratio asymptotics

$$\lim_{n \rightarrow \infty} \frac{p_{n+1}(z)}{p_n(z)} = \frac{z - b + \sqrt{(z - b)^2 - 4a^2}}{2}$$

for large z (actually, away from the support of μ), and the monograph [80] contains a very detailed treatment of orthogonal polynomials in this class. Simon [101] showed that if the limit of $p_{n+1}(z)/p_n(z)$ exists at a single non-real z , then $\mu \in M(b, a)$ for some a, b .

The classes $M(b, a)$ are scaled versions of each other, and the most important condition ensuring $M(0, \frac{1}{2})$ is given in Rakhmanov’s theorem [92]: if μ is supported in $[-1, 1]$ and $\mu' > 0$ almost everywhere on $[-1, 1]$, then $\mu \in M(0, \frac{1}{2})$. Conversely, Blumenthal’s theorem [13] states that $\mu \in M(0, \frac{1}{2})$ implies that the support of μ is $[-1, 1]$ plus at most countably many points that converge to ± 1 . Thus, in this respect, the extension of Rakhmanov’s theorem given in [24] by Denisov is of importance: if $\mu' > 0$ almost everywhere on $[-1, 1]$ and outside $[-1, 1]$ the measure μ has at most countably many mass points converging to ± 1 , then $\mu \in M(0, \frac{1}{2})$. However, $M(0, \frac{1}{2})$ contains many other measures not just those that are in these theorems; e.g., in [23] a continuous singular measure in the Nevai class was exhibited, and the result in [126] shows that the Nevai class contains practically all kinds of measures allowed by Blumenthal’s theorem.

There are a number of papers where the condition $b_n \rightarrow 0, a_n \rightarrow \frac{1}{2}$ is strengthened. See, e.g., the paper [57] of Killip and Simon, where a complete characterization of

$$\sum_n (a_n - \frac{1}{2})^2 + b_n^2 < \infty$$

is given in terms of the generating measure.

Under Rakhmanov’s condition, $\text{supp}(\mu) = [-1, 1]$, $\mu' > 0$ a.e., some parts of Szegő’s theory can be proven in a weaker form (see, e.g., [74, 75]). In these the Turán determinants

$$T_n(x) = p_n^2(x) - p_{n-1}(x)p_{n+1}(x)$$

play a significant role. In fact, given any interval $\mathbf{D} \subset (-1, 1)$, the Turán determinant T_n is positive on \mathbf{D} for all large n , and $T_n^{-1}(x)dx$ converges in weak- $*$ sense to $d\mu$ on \mathbf{D} . Furthermore, the absolutely continuous part μ' can also be separately recovered from T_n :

$$\lim_{n \rightarrow \infty} \int \left| T_n(x)\mu'(x) - \frac{2}{\pi}(1-x^2)^{1/2} \right| dx = 0.$$

Under Rakhmanov’s condition, we also have weak convergence, e.g.,

$$\lim_{n \rightarrow \infty} \int f(x)p_n^2(x)\mu'(x)dx = \frac{1}{\pi} \int_{-1}^1 \frac{f(x)}{\sqrt{1-x^2}} dx \tag{8.5}$$

for any continuous function f . Pointwise we only know a highly oscillatory behavior: for almost all $x \in [-1, 1]$,

$$\begin{aligned} \limsup_{n \rightarrow \infty} p_n(x) &\geq \frac{2}{\pi} (\mu'(x))^{-1/2} (1-x^2)^{-1/4}, \\ \liminf_{n \rightarrow \infty} p_n(x) &\leq -\frac{2}{\pi} (\mu'(x))^{-1/2} (1-x^2)^{-1/4}, \end{aligned}$$

and if $E_n(\varepsilon)$ is the set of points $x \in [-1, 1]$ where

$$|p_n(x)| \geq (1 + \varepsilon) \frac{2}{\pi} (\mu'(x))^{-1/2} (1-x^2)^{-1/4},$$

then $|E_n(\varepsilon)| \rightarrow 0$ for all $\varepsilon > 0$. However, it is not true that the sequence $\{p_n(\mu, x)\}$ is pointwise bounded, since for every $\varepsilon > 0$ there is a weight function $w > 1$ on $[-1, 1]$ such that $p_n(0)/n^{1/2-\varepsilon}$ is unbounded (see [93]).

Simon [101] extended (8.5) by showing that if the recurrence coefficients satisfy $b_n \rightarrow b$, $a_{2n+1} \rightarrow a'$ and $a_{2n} \rightarrow a''$, then there is an explicitly calculated measure ρ depending only on b, a', a'' such that

$$\lim_{n \rightarrow \infty} \int f(x)p_n^2(x)\mu'(x)dx = \int_{-1}^1 f(x)d\rho(x) \tag{8.6}$$

for any continuous function f , and conversely, if (8.6) exists for $f(x) = x, x^2, x^4$, then $b_n \rightarrow b$, $a_{2n+1} \rightarrow a'$ and $a_{2n} \rightarrow a''$ with some b, a', a'' .

Szegő’s theory can be extended to measures lying on a single Jordan curve or arc J (see [52] where also additional outside lying mass points are allowed), in which case the role of $z + \sqrt{z^2 - 1}$ in (8.3) is played by the conformal map Φ of $\mathbf{C} \setminus J$ onto the exterior of the unit disk, and the role of 2^n in (8.4) is played by the reciprocal of the logarithmic capacity of J (see Section 7). Things change considerably if the measure is supported on a set J consisting of two or more smooth curve or arc components J_1, \dots, J_m . A general feature of this case is that $\kappa_n \text{cap}(J)^n$ does not have a limit—its limit points fill a whole interval (though if some associated harmonic measures are all rational then the limit points may form a finite set). The polynomials themselves have asymptotic form

$$\frac{p_n(z)}{\kappa_n} = \text{cap}(J)^n \Phi(z)^n (F_n(z) + o(1))$$

uniformly away from J , where Φ is the (multi-valued) complex Green’s function of the complement $\mathbf{C} \setminus J$, and where F_n is the solution of an L^2 -extremal problem involving analytic functions belonging to some class Γ_n . The functions F in Γ_n are determined by an H^2 condition plus an argument condition: if the change of the argument of Φ as we go around J_k is $\gamma_k 2\pi$ modulo 2π , then for the functions in Γ_n the change of the argument around J_k is $-n\gamma_k 2\pi$ modulo 2π . Now the point is that these function classes Γ_n change with n , and hence so does F_n , and that is the reason that a single asymptotic formula like (8.4) or (8.3) does not hold. The fundamentals of the theory were laid in Widom’s paper [133]; and since then many results have been obtained by Peherstorfer and his collaborators, as well as Aptekarev, Geronimo, Suetin and Van Assche. The theory has deep connections with function theory, the theory of Abelian integrals and the theory of elliptic functions. We refer the reader to the papers [6, 38, 87, 88, 89, 90, 91, 119, 120].

The Christoffel functions

$$\lambda_n^{-1}(\mu, x) = \sum_{k=0}^n p_k(\mu, x)^2$$

behave somewhat more regularly than the orthogonal polynomials. In [76] it was shown that if μ is supported on $[-1, 1]$, it belongs to the **Reg** class there (see Section 7) and $\log \mu'$ is integrable over an interval $I \subset [-1, 1]$, then for almost all $x \in I$,

$$\lim_{n \rightarrow \infty} n\lambda_n(x) = \pi \sqrt{1 - x^2} \mu'(x).$$

This result is true [127] in the form

$$\lim_{n \rightarrow \infty} n\lambda_n(x) = \frac{d\mu(x)}{d\omega_{\text{supp}(\mu)}(x)}, \quad \text{a.e. } x \in I$$

when the support is a general compact subset of \mathbf{R} , $\mu \in \mathbf{Reg}$ and $\log \mu' \in L^1(I)$.

Often only a rough estimate is needed for Christoffel functions, and such an estimate is provided in [73]: if w is supported on $[-1, 1]$ and it is a doubling weight, i.e.,

$$\int_{2I} w \leq L \int_I w$$

for all $I \subset [-1, 1]$, where $2I$ is the twice enlarged I , then uniformly on $[-1, 1]$,

$$\lambda_n(x) \sim \int_{\Delta_n(x)} w; \quad \Delta_n(x) = \frac{\sqrt{1 - x^2}}{n} + \frac{1}{n^2}.$$

9. Exponential and Freud Weights

These are weight functions of the form $e^{-2Q(x)}$, where x is on the real line or on some subinterval of it. For simplicity we shall first assume that Q is even. We get Freud weights when $Q(x) = |x|^\alpha$, $\alpha > 0$, $x \in \mathbf{R}$ and Erdős weights if Q tends to infinity faster than any polynomial as $|x| \rightarrow \infty$. Freud started to investigate these weights in the sixties and seventies, but they independently appeared also in the Russian literature and in statistical physics. One can safely say that some of Freud’s problems and the work of Nevai and Rakhmanov were the primary cause of the sudden revitalization of the theory of orthogonal polynomials since the early 1980’s. In the last twenty years, Lubinsky with coauthors have conducted systematic studies on exponential weights, see, e.g., [61, 62, 65, 66, 128]; we should mention the

names Levin, Saff, Van Assche, Rakhmanov and Mhaskar. In the mid-1990's a new stimulus came from the Riemann–Hilbert approach that was used together with the steepest descent method by Deift and his collaborators ([20]) to give complete asymptotics when Q is analytic.

One can roughly say that because of the fast vanishing of the weight around infinity, things happen on a finite subinterval $[-a_n, a_n]$ (depending on the degree of the polynomials), and on $[-a_n, a_n]$ techniques developed for $[-1, 1]$ are applied. For Freud weights one can also make the substitution $x \rightarrow n^{1/\lambda}x$ and go to orthogonality with respect to the varying weight $e^{-n|x|^\lambda}$, in which case things are automatically reduced to a finite interval which is the support of a weighted energy problem.

a_n are the so-called Mhaskar–Rakhmanov–Saff numbers with definition

$$n = \frac{2}{\pi} \int_0^1 \frac{a_n t Q'(a_n t)}{\sqrt{1-t^2}} dt. \tag{9.1}$$

The zeros of $p_n(w^2)$, $w(x) = \exp(-Q(x))$ are spreading out and the largest zero is very close to a_n , which tends to ∞ .

To describe the distribution of the zeros and the behavior of the polynomials, one has to make appropriate contractions. Let us consider first the case of Freud weight $w(x) = \exp(-|x|^\alpha)$, and let p_n be the n -th orthogonal polynomial with respect to w^2 (on $(-\infty, \infty)$). In this case

$$a_n = n^{1/\alpha} \gamma_\alpha, \quad \gamma_\alpha^\alpha := \Gamma\left(\frac{\alpha}{2}\right) \Gamma\left(\frac{1}{2}\right) / 2\Gamma\left(\frac{\alpha}{2} + \frac{1}{2}\right).$$

Thus, for the largest zero $x_{n,n}$ we have $x_{n,n}/n^{1/\alpha} \rightarrow \gamma_\alpha$ as $n \rightarrow \infty$, and to describe zero distribution we divide (contract) all zeros $x_{n,i}$ by $n^{1/\alpha} \gamma_\alpha$. These contracted zeros asymptotically have the distribution

$$\frac{d\mu_w(t)}{dt} := \frac{\alpha}{\pi} \int_{|t|}^1 \frac{u^{\alpha-1}}{\sqrt{u^2-t^2}} du, \quad t \in [-1, 1]. \tag{9.2}$$

This measure μ_w minimizes the weighted energy

$$\iint \log \frac{1}{|x-t|} d\mu(x) d\mu(t) + 2 \int Q d\mu \tag{9.3}$$

among all probability measures compactly supported on \mathbf{R} . It is a general feature of exponential weights that the behavior of zeros or the polynomials is governed by the solution of a weighted energy problem (weighted equilibrium measures). If κ_n is the leading coefficient of p_n , i.e., $p_n(z) = \kappa_n z^n + \dots$, then

$$\lim_{n \rightarrow \infty} \kappa_n \pi^{1/2} 2^{-n} e^{-n/\alpha} n^{(n+1/2)/\alpha} = 1,$$

and we have

$$\lim_{n \rightarrow \infty} |p_n(n^{1/\alpha} \gamma_\alpha z)|^{1/n} = \exp \left(\log |z + \sqrt{z^2 - 1}| + \operatorname{Re} \int_0^1 \frac{z u^{\alpha-1}}{\sqrt{z^2 - u^2}} du \right)$$

locally uniformly outside $[-1, 1]$. This latter one is so-called n -th root asymptotics, while the former one is strong asymptotics. Strong asymptotics for $p_n(z)$ on different parts of the complex plane was given in [59] using the Hilbert–Riemann approach (see also [21] in this Festschrift).

Things become more complicated for non-Freud weights, but the corresponding results are of the same flavor. In this case the weight is not necessarily symmetric, but under some conditions (like Q being convex or $xQ'(x) \nearrow$ for $x > 0$ and an

analogous condition for $x < 0$), the support of the relevant weighted equilibrium measure is an interval, and the definition of the Mhaskar–Rakhmanov–Saff numbers $a_{\pm n}$ is

$$n = \frac{1}{\pi} \int_{a_{-n}}^{a_n} \frac{xQ'(x)}{\sqrt{(x - a_{-n})(a_n - x)}} dx,$$

$$0 = \frac{1}{\pi} \int_{a_{-n}}^{a_n} \frac{Q'(x)}{\sqrt{(x - a_{-n})(a_n - x)}} dx.$$

Now one solves the weighted equilibrium problem (9.3) for all measures μ with total mass n , and if μ_n is the solution, then $[a_{-n}, a_n]$ is the support of μ_n and μ_n/n will play the role of the measure μ_w from (9.2) above.

The weight does not even have to be defined on all \mathbf{R} , e.g., in [62] a theory was developed that simultaneously includes far-reaching generalizations of non-symmetric Freud, Erdős and Pollaczek weights (the latter are defined on $[-1, 1]$ and vanish at high order at ± 1).

10. Sobolev Orthogonality

In Sobolev orthogonality we consider orthogonality with respect to an inner product

$$(f, g) = \sum_{k=0}^r \int f^{(k)} \overline{g^{(k)}} d\mu_k, \tag{10.1}$$

where μ_k are given positive measures. There are several motivations for this kind of orthogonality, perhaps the most natural one is smooth data fitting. The Spanish school around Marcellán, Lopez and Martinez-Finkelshtein has been particularly active in developing this area (see the surveys [68, 69, 70] and the references there).

In this section let $Q_n(z) = z^n + \dots$ denote the monic orthogonal polynomial with respect to the Sobolev inner product (10.1), and $q_n(\mu_k)$ the monic orthogonal polynomials with respect to the measure μ_k .

Most arguments for the standard theory fail in this case, e.g., it is no longer true that the zeros lie in the convex hull of the support of the measures $\mu_k, k = 0, 1, \dots, r$. It is not even known if the zeros are bounded if all the measures μ_k have compact support. Nonetheless, for the case $r = 1$, and $\mu_0, \mu_1 \in \mathbf{Reg}$ (see Section 7), it was shown in [33] that the asymptotic distribution of the zeros of the derivative Q'_n is the equilibrium measure $\omega_{E_0 \cup E_1}$, where E_i is the support of $\mu_i, i = 0, 1$ (which also have to be assumed to be regular). If, in addition, $E_0 \subseteq E_1$, then the asymptotic zero distribution of Q_n is ω_{E_0} .

In general, both the algebraic and the asymptotic/analytic situation is quite complicated, and there are essentially two important cases which have been understood to a satisfactory degree.

Case I: The discrete case. In this case μ_0 is some “strong” measure, e.g., from the Nevai class $M(b, a)$ (see Section 8), and μ_1, \dots, μ_k are finite discrete measures. It turns out then that the situation is similar to adding these discrete measures to μ_0 (the new measure will also be in the same Nevai class), and considering standard orthogonality with respect to this new measure. For example, if $r = 1$, then

$$\lim_{n \rightarrow \infty} \frac{Q_n(z)}{q_n(\mu_0 + \mu_1, z)} = 1$$

holds uniformly on compact subsets of $\mathbf{C} \setminus \text{supp}(\mu_0 + \mu_1)$. Thus, the Sobolev orthogonal polynomials differ from those of the measure μ_0 , but not more than what happens when adding mass points to μ_0 .

In this discrete case the Q_n 's satisfy a higher-order recurrence relation, hence this case is also related to matrix orthogonality (see the end of Section 13).

Case II: The Szegő case. Suppose now that μ_0, \dots, μ_k are all supported on the same smooth curve or arc J , and they satisfy Szegő's condition there (see Section 8). In this case the k -th derivative of Q_n satisfies locally uniformly in the complement of J the asymptotic formula

$$\lim_{n \rightarrow \infty} \frac{Q_n^{(k)}(z)}{n^k q_{n-k}(\mu_r, z)} = \frac{1}{[\Phi'(z)]^{m-k}},$$

where Φ is the conformal map that maps $\mathbf{C} \setminus J$ onto the complement of the unit disk. That is, in this case the measures μ_0, \dots, μ_{r-1} do not appear in the asymptotic formula, only μ_r matters. The reason for this is the following: $Q = Q_n$ minimizes

$$(Q, Q) = \sum_{k=0}^r \int |Q^{(k)}|^2 d\mu_k \tag{10.2}$$

among all monic polynomials of degree n , while $q = q_{n-k}(\mu_k)$ minimizes

$$\int |q|^2 d\mu_k$$

among all monic polynomials of degree $n - k$. But the polynomial $Q_n^{(k)}(t) = n(n - 1) \cdots (n - k + 1)t^{n-k} + \dots$ is a monic polynomial times the factor $n(n - 1) \cdots (n - k + 1) \sim n^k$, and this factor is dominant for $k = r$, so everything else will be negligible. There are results for compensation of this n^k factor which lead to Sobolev orthogonality with respect to varying measures.

Under the much less restrictive assumption that $\mu_0 \in \mathbf{Reg}$ and the other measures μ_k are supported in the support E of μ_0 , it is true ([63]) that the asymptotic zero distribution of $Q_n^{(k)}$ is the equilibrium measure ω_E for all k ,

$$\lim_{n \rightarrow \infty} \|Q_n^{(k)}\|_{\text{sup}, E}^{1/n} = \text{cap}(E),$$

and hence, away from the zeros in the unbounded component of the complement of E , we have

$$\lim_{n \rightarrow \infty} |Q_n^k(z)|^{1/n} = e^{g_{\mathbf{C} \setminus E}(z)}$$

where $g_{\mathbf{C} \setminus E}$ is the Green's function for this unbounded component.

The techniques developed for exponential weights and for Sobolev orthogonality were combined in [36] to prove strong asymptotics for Sobolev orthogonal polynomials when $r = 1$ and $\mu_0 = \mu_1$ are exponential weights.

11. Non-Hermitian Orthogonality

We refer to non-Hermitian orthogonality in either of these cases:

- the measure μ is non-positive or even complex-valued and we consider p_n with

$$\int p_n(z) \overline{z^k} d\mu = 0, \quad k = 0, 1, \dots, n - 1 \tag{11.1}$$

- μ is again non-positive or complex-valued, or positive but lies on a complex curve or arc and orthogonality is considered without complex conjugation, i.e., if

$$\int p_n(z)z^k d\mu = 0, \quad k = 0, 1, \dots, n - 1. \tag{11.2}$$

More generally, one could consider non-positive inner products, but we shall restrict our attention to complex measures and orthogonality (11.2).

As an example, consider the diagonal Padé approximant to the Cauchy transform

$$f(z) = \int \frac{d\mu(t)}{z - t}$$

of a signed or complex-valued measure, i.e., consider polynomials p_n and q_n of degree at most n such that

$$f(z)p_n(z) - q_n(z) = O(z^{-n-1})$$

at infinity. Then p_n satisfies the non-Hermitian orthogonality relation

$$\int p_n(x)x^j d\mu(x) = 0, \quad j = 0, 1, \dots, n - 1. \tag{11.3}$$

In this non-Hermitian case even the Gram–Schmidt process may fail, and then p_n is defined as the solution of the orthogonality condition (11.1) (resp. (11.2)), which gives a system of homogeneous equations for the coefficients of p_n . Thus, p_n may be of smaller than n degree, and things can go pretty wild with this kind of orthogonality, e.g., in the simple case,

$$d\mu(x) = (x - \cos \pi\alpha_1)(x - \cos \pi\alpha_2)(1 - x^2)^{-1/2}dx, \quad x \in [-1, 1]$$

with $0 < \alpha_1 < \alpha_2 < 1$ rationally independent algebraic numbers, the zeros of p_n from (11.3) are dense on the whole complex plane (compare this with the fact that for positive μ all zeros lie in $[-1, 1]$). In [116] it was shown that it is possible to construct a complex measure μ on $[-1, 1]$, such that for an arbitrarily prescribed asymptotic behavior some subsequence $\{p_{n_k}\}$ will have this zero behavior. Nonetheless, the asymptotic distribution of the zeros is again the equilibrium distribution of the support of μ under regularity conditions on μ . For example, this is the case if $|\mu|$ belongs to the **Reg** class (see Section 7), and the argument of μ , i.e., $d\mu(t)/d|\mu|(t)$ is of bounded variation [9]. In [114, 115, 116] Stahl obtained asymptotics for non-Hermitian orthogonal polynomials even for varying measures and gave several applications of them to Padé approximation. When the measure μ is of the form $d\mu(x) = g(x)(1 - x^2)^{-1/2}dx$, $x \in [-1, 1]$ with an analytic g , for $z \in \mathbf{C} \setminus [-1, 1]$, the strong asymptotic formula of the form

$$\frac{p_n(z)}{\kappa_n} = (1 + o(1)) \frac{(z + \sqrt{z^2 - 1})^n}{2^n} D_\mu(z)^{-1} \exp\left(\frac{1}{2\pi} \int_{-1}^1 \frac{\log \mu'(t)}{\sqrt{1 - t^2}} dt\right)$$

(with D_μ the Szegő function (8.2)) was proved by Nuttall [84, 85], Gonchar and Suetin [49]. For a recent Riemann–Hilbert approach, see the paper [8] by Aptekarev and Van Assche. A similar result holds on the support of the measure, as well as for the case of varying weights; see [8].

12. Multiple Orthogonality

Multiple orthogonality comes from simultaneous Padé approximation. It is a relatively new area where we have to mention the names of Nikishin, Sorokin, Gonchar and Rakhmanov, Aptekarev, Kuijlaars and Van Assche (see the survey [129] by Van Assche and the references there and the paper [48]).

On \mathbf{R} let there be given r measures μ_1, \dots, μ_r with finite moments and infinite support, and consider multi-indices $\underline{n} = (n_1, \dots, n_r)$ of non-negative integers with norm $|\underline{n}| = n_1 + \dots + n_r$. There are two types of multiple orthogonality corresponding to the appropriate Hermite–Padé approximation.

In type I we are looking for polynomials $Q_{\underline{n},j}$ of degree $n_j - 1$ for each $j = 1, \dots, r$ such that

$$\sum_{j=1}^r \int x^k Q_{\underline{n},j}(x) d\mu_j(x) = 0, \quad k = 0, 1, \dots, |\underline{n}| - 2.$$

These orthogonality relations give $|\underline{n}| - 1$ homogeneous linear equations for the $|\underline{n}|$ coefficients of the r polynomials $Q_{\underline{n},j}$, so there is a non-trivial solution. If the rank of the system is $|\underline{n}| - 1$, then the solution is unique up to a multiplicative factor, in which case the index \underline{n} is called normal. This happens precisely if each $Q_{\underline{n},j}$ is of exact degree $n_j - 1$.

In type II we are looking for a single polynomial $P_{\underline{n}}$ of degree $|\underline{n}|$ such that

$$\begin{aligned} \int x^k P_{\underline{n}}(x) d\mu_1(x) &= 0, & k = 1, \dots, n_1 - 1 \\ &\vdots \\ \int x^k P_{\underline{n}}(x) d\mu_r(x) &= 0, & k = 1, \dots, n_r - 1. \end{aligned}$$

These are $|\underline{n}|$ homogeneous linear equations for the $|\underline{n}| + 1$ coefficients of $P_{\underline{n}}$, and again if the solution is unique up to a multiplicative constant, then \underline{n} is called normal. This is again equivalent to $P_{\underline{n}}$ being of exact degree $|\underline{n}|$.

\underline{n} is normal for type I orthogonality precisely when it is normal for type II, so we just speak of normality. This is the case, for example, if the μ_j 's are supported on intervals $[a_j, b_j]$ that are disjoint except perhaps for their endpoints; in fact, in this case $P_{\underline{n}}$ has n_j simple zeros on (a_j, b_j) .

To describe recurrence formulae, let $\underline{e}_j = (0, \dots, 1, \dots, 0)$ where the single 1 entry is at position j . Under the normality assumption, if $P_{\underline{n}}$ is the monic orthogonal polynomial, then for any k ,

$$xP_{\underline{n}}(x) = P_{\underline{n}+\underline{e}_k}(x) + a_{\underline{n},0}P_{\underline{n}}(x) + \sum_{j=1}^r a_{\underline{n},j}P_{\underline{n}-\underline{e}_j}(x).$$

Another recurrence formula is

$$xP_{\underline{n}}(x) = P_{\underline{n}+\underline{e}_k}(x) + b_{\underline{n},0}P_{\underline{n}}(x) + \sum_{j=1}^r b_{\underline{n},j}P_{\underline{n}-\underline{e}_{\pi(1)}-\dots-\underline{e}_{\pi(j)}}(x),$$

where $\pi(1), \dots, \pi(r)$ is an arbitrary but fixed permutation of $1, 2, \dots, r$. The orthogonal polynomials with different indices are strongly related to one another, e.g., $P_{\underline{n}+\underline{e}_k}(x) - P_{\underline{n}+\underline{e}_l}(x)$ is a constant multiple of $P_{\underline{n}}(x)$.

If $d\mu_j = w_j d\mu$, then similar recurrence relations hold in case of type I orthogonality for

$$Q_{\underline{n}}(x) = \sum_{j=1}^r Q_{\underline{n},j}(x)w_j(x).$$

Also, type I and type II are related by a biorthogonality property:

$$\int P_{\underline{n}}Q_{\underline{m}}d\mu = 0$$

except for the case when $\underline{m} = \underline{n} + \underline{e}_k$ for some k , and then the previous integral is not zero (under normality condition).

To describe an analogue of the Christoffel–Darboux formula, let $\{\underline{n}_j\}$ be a sequence of multi-indices such that \underline{n}_0 is the identically 0 multi-index, and \underline{n}_{j+1} coincides with \underline{n}_j except for one component which is 1 larger than the corresponding component of \underline{n}_j . Set $P_j = P_{\underline{n}_j}$, $Q_j = Q_{\underline{n}_{j+1}}$, $\underline{n} = \underline{n}_n$, and $(\underline{n})_j$ denoting the j -th component of the multi-index \underline{n}

$$h_{\underline{n}}^{(j)} = \int P_{\underline{n}}(x)x^{(\underline{n})_j}d\mu_j(x).$$

Then [18] with $\underline{n} = \underline{n}_n$,

$$(x - y) \sum_{k=0}^{n-1} P_k(x)Q_k(y) = P_{\underline{n}}(x)Q_{\underline{n}}(y) - \sum_{j=1}^r \frac{h_{\underline{n}}^{(j)}}{h_{\underline{n}-\underline{e}_j}^{(j)}} P_{\underline{n}-\underline{e}_j}(x)Q_{\underline{n}+\underline{e}_j}(y).$$

Thus, the left-hand side depends only on $\underline{n} = \underline{n}_n$ and not on the particular choice of the sequence \underline{n}_j leading to it.

There is an approach [130] to both types of multiple orthogonality in terms of the matrix-valued Riemann–Hilbert problem for $(r + 1) \times (r + 1)$ matrices $Y = (Y_{ij}(z))_{i,j=0}^r$.

Asymptotic behavior of multiple orthogonal polynomials is not fully understood yet, due to the interaction of the different measures. For all the existing results, see [5, 129] and Van Assche’s Chapter 23 in [51] and the references there.

13. Matrix Orthogonal Polynomials

In the last twenty years the fundamentals of matrix orthogonal polynomials have been developed mainly by Duran and his coauthors (see also the work [7] by Aptekarev and Nikishin). The theory shows many similarities with the scalar case, but there is an unexpected richness which is still to be explored.

For all the results in this section, see [64] and [29] and the numerous references there.

An $N \times N$ matrix

$$P(t) = \begin{pmatrix} p_{11}(t) & \cdots & p_{1N}(t) \\ \vdots & \ddots & \vdots \\ p_{N1}(t) & \cdots & p_{NN}(t) \end{pmatrix}$$

with polynomial entries $p_{ij}(t)$ of degree at most n is called a matrix polynomial of degree at most n . Alternatively, one can write

$$P(t) = C_n t^n + \cdots + C_0$$

with numerical matrices C_n, \dots, C_0 of size $N \times N$.

$t = a$ is called a zero of P if $P(a)$ is singular, and the multiplicity of a is the multiplicity of a as a zero of $\det P(a)$. When the leading coefficient matrix C_n is non-singular, then P has nN zeros counting multiplicity.

From now on we fix the dimension to be N , but the degree n can be any natural number. I will denote the $N \times N$ unit matrix and 0 stands for all kinds of zeros (numerical or matrix).

A matrix

$$W(t) = \begin{pmatrix} \mu_{11}(t) & \cdots & \mu_{1N}(t) \\ \vdots & \ddots & \vdots \\ \mu_{N1}(t) & \cdots & \mu_{NN}(t) \end{pmatrix}$$

of measures defined on (or part of) the real line is positive definite if for any Borel set E the numerical matrix $W(E)$ is positive semidefinite. We assume that all moments of W are finite. With such a matrix we can define a matrix inner product on the space of $N \times N$ matrix polynomials via

$$(P, Q) = \int P(t)dW(t)Q^*(t),$$

and if (P, P) is non-singular for any P with non-singular leading coefficient, then just as in the scalar case one can generate a sequence $\{P_n\}_{n=0}^\infty$ of matrix polynomials of degree $n = 0, 1, \dots$ which are orthonormal with respect to W :

$$\int P_n(t)dW(t)P_m^*(t) = \begin{cases} 0 & \text{if } n \neq m \\ I & \text{if } n = m, \end{cases}$$

and here P_n has non-singular leading coefficient matrix. The sequence $\{P_n\}$ is determined only up to multiplication on the left by unitary matrices, i.e., if U_n are unitary matrices, then the polynomials $U_n P_n$ also form an orthonormal system with respect to W .

Just as in the scalar case, these orthogonal polynomials satisfy a three-term recurrence relation

$$tP_n(t) = A_{n+1}P_{n+1}(t) + B_nP_n(t) + A_n^*P_{n-1}(t), \quad n \geq 0, \tag{13.1}$$

where A_n are non-singular matrices, and B_n are Hermitian. Conversely, the analogue of Favard’s theorem is also true: if a sequence of matrix polynomials $\{P_n\}$ of corresponding degree $n = 0, 1, 2, \dots$ satisfy (13.1) with non-singular A_n and Hermitian B_n , then there is a positive definite measure matrix W such that P_n are orthonormal with respect to W .

The three-term recurrence formula easily yields the Christoffel–Darboux formula:

$$(w - z) \sum_{k=0}^{n-1} P_k^*(z)P_k(w) = P_{n-1}^*(z)A_nP_n(w) - P_n^*(z)A_n^*P_{n-1}(w),$$

from which, e.g., it follows that

$$P_{n-1}^*(z)A_nP_n(z) - P_n^*(z)A_n^*P_{n-1}(z) = 0$$

$$\sum_{k=0}^{n-1} P_k^*(z)P_k(z) = P_{n-1}^*(z)A_nP_n'(z) - P_n^*(z)A_n^*P_{n-1}'(z).$$

The orthogonal polynomials Q_n of the second kind

$$Q_n(t) = \int \frac{P_n(t) - P_n(x)}{t - x} dW(x), \quad n = 1, 2, \dots$$

also satisfy the same recurrence and are orthogonal with respect to some other matrix measure. For them we have

$$P_{n-1}^*(t)A_nQ_n(t) - P_n^*(z)A_n^*Q_{n-1}(t) \equiv I$$

and

$$Q_n(t)P_{n-1}^*(t) - P_n(t)Q_{n-1}^*(t) \equiv A_n^{-1}.$$

With the recurrence coefficient matrices A_n, B_n , one can form the block Jacobi matrix

$$J = \begin{pmatrix} B_0 & A_0 & 0 & 0 & \dots \\ A_0^* & B_1 & A_1 & 0 & \dots \\ 0 & A_1^* & B_2 & A_2 & \dots \\ 0 & 0 & A_2^* & B_2 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

The zeros of P_n are real and they are the eigenvalues (with the same multiplicity) of the N -truncated block Jacobi matrix (which is of size nN). If a is a zero then its multiplicity p is at most N , the rank of $P_n(a)$ is $N - p$, and the space of those vectors v for which $P_n(a)v = 0$ is of dimension p . If we write $x_{n,k}, 1 \leq k \leq m$ for the different zeros of P_n , and l_k is the multiplicity of $x_{n,k}$, then the matrices

$$\Gamma_k = \frac{1}{(\det(P_n(t))^{(l_k)}(x_{n,k}))} (\text{Adj}(P_n(t)))^{(l_k-1)}(x_{n,k})Q_n(x_{n,k}), \quad 1 \leq k \leq m$$

are positive semidefinite of rank l_k , and with them the matrix quadrature formula

$$\int P(t)dW(t) = \sum_{k=1}^m P(x_{n,k})\Gamma_{n,k}$$

holds for all matrix polynomials P of degree at most $2n - 1$.

If the matrix of orthogonality is diagonal (or similar to a diagonal matrix) with diagonal entries μ_i , then the orthogonal matrix polynomials are also diagonal with i -th entry equal to $p_n(\mu_i)$, the n -th orthogonal polynomial with respect to μ_i . Many matrix orthogonal polynomials in the literature can be reduced to this scalar case. However, recently some remarkably rich non-reducible families have been obtained by Duran and Grünbaum (see [29] and the references there), which may play the role of the classical orthogonal polynomials in higher dimension. They found families of matrix orthogonal polynomials that satisfy second-order (matrix) differential equations just like the classical orthogonal polynomials. Their starting point was a symmetry property between the orthogonality measure matrix and a second-order differential operator. They worked out several explicit examples. Here is one of them: $N = 2$, the measure matrix is

$$H(t) = e^{-t^2} \begin{pmatrix} 1 + |a|^2t^4 & at^2 \\ \bar{a}t^2 & 1 \end{pmatrix}, \quad t \in \mathbf{R},$$

where $a \in \mathbf{C} \setminus \{0\}$ is a free parameter. The corresponding $P_n(t)$ satisfies

$$P_N''(t) + P_n'(t) \begin{pmatrix} -2t & 4at \\ 0 & -2t \end{pmatrix} + P_n(t) \begin{pmatrix} -4 & 2a \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} -2n - 4 & 2a(2n + 1) \\ 0 & -2n \end{pmatrix} P_n(t)$$

There is an explicit Rodrigues-type representation for the polynomials themselves, and the three-term recurrence (13.1) holds with $B_n = 0$,

$$A_{n+1} = \sqrt{\frac{n+1}{2}} \begin{pmatrix} \gamma_{n+3}/\gamma_{n+2} & a\gamma_{n+2}\gamma_{n+1} \\ 0 & \gamma_n/\gamma_{n+1} \end{pmatrix},$$

where

$$\gamma_n^2 = 1 + \frac{|a|^2}{2} \binom{n}{2}.$$

Matrix orthogonality is closely connected to $(2N+1)$ -term recurrences for scalar polynomials. To describe this we need the following operators on polynomials p : if $p(t) = \sum_k a_k t^k$, then

$$R_{N,m}(p) = \sum_s a_{sN+m} t^s,$$

i.e., from a polynomial the operator $R_{N,m}$ takes those powers where the exponent is congruent to m modulo N , removes the common factor t^m and changes t^N to t .

Now suppose that $\{p_n\}_{n=0}^\infty$ is a sequence of scalar polynomials of corresponding degree $n = 0, 1, \dots$ and suppose that this sequence satisfies a $(2N + 1)$ -term recurrence relation

$$t^N p_n(t) = c_{n,0} p_n(t) + \sum_{k=1}^N (\overline{c_{n,k}} p_{n-k}(t) + c_{n+k,k} p_{n+k}(t)),$$

where $c_{n,0}$ is real, $c_{n,N} \neq 0$ (and $p_k(t) \equiv 0$ for $k < 0$). Then

$$P_n(t) = \begin{pmatrix} R_{N,0}(p_{nN}) & \cdots & R_{N,N-1}(p_{nN}) \\ R_{N,0}(p_{nN+1}) & \cdots & R_{N,N-1}(p_{nN+1}) \\ \vdots & \ddots & \vdots \\ R_{N,0}(p_{nN+N-1}) & \cdots & R_{N,N-1}(p_{nN+N-1}) \end{pmatrix}$$

is a sequence of matrix orthogonal polynomials with respect to a positive definite measure matrix. Conversely, if $P_n = (P_{n,m,j})_{m,j=0}^{N-1}$ is a sequence of orthonormal matrix polynomials, then the scalar polynomials

$$p_{nN+m}(t) = \sum_{j=0}^{N-1} t^j P_{n,m,j}(t^N), \quad 0 \leq m < N, \quad n = 0, 1, 2, \dots$$

satisfy a $(2N + 1)$ -recurrence relation of the above form.

Part 2. ORTHOGONAL POLYNOMIALS ON THE UNIT CIRCLE

In what follows we shall use Simon’s abbreviation OPUC for orthogonal polynomials on the unit circle.

14. Definitions and Basic Properties

14.1. Orthogonality. The unit circle \mathbf{T} is by far the simplest closed curve on the complex plane with a number of additional properties, so polynomials orthogonal with respect to measures on \mathbf{T} are of specific interest.

If μ is a non-trivial probability measure on \mathbf{T} (i.e., not supported on a finite set), the monic orthogonal polynomials $\Phi_n(z, \mu)$ (or Φ_n if μ is understood) are uniquely determined by

$$\Phi_n(z) = \prod_{j=1}^n (z - z_{n,j}), \quad \int_{\mathbf{T}} \zeta^{-j} \Phi_n(\zeta) d\mu = 0, \quad j = 0, 1, \dots, n - 1 \quad (14.1)$$

so in the Hilbert space $L^2_\mu(\mathbf{T})$, $\langle \Phi_n, \Phi_m \rangle = 0, n \neq m$. The orthonormal polynomials φ_n are $\varphi_n = \Phi_n / \|\Phi_n\|$. The orthonormal set $\{\varphi_n\}_{n \geq 0}$ may not be a basis in $L^2_\mu(\mathbf{T})$ (e.g., if $\mu = dm$ is the normalized Lebesgue measure, then $\varphi_n = \zeta^n$ and ζ^{-1} is orthogonal to all φ_n). A celebrated result of Szegő–Kolmogorov–Krein (see Theorem 14.2 below) states that $\{\varphi_n\}$ is a basis in $L^2_\mu(\mathbf{T})$ if and only if $\log \mu' \notin L^1(\mathbf{T})$, where μ' is the Radon–Nikodym derivative of μ with respect to dm .

Clearly, (14.1) and the fact that the polynomials of degree n have dimension $n + 1$ implies

$$\deg(P) = n, \quad P \perp \zeta^j, \quad j = 0, 1, \dots, n - 1 \Rightarrow P = c\Phi_n. \quad (14.2)$$

On $L^2_\mu(\mathbf{T})$ the anti-unitary map $f^*(\zeta) := \zeta^n \overline{f(\zeta)}$ (which depends on n) is naturally defined. The set of polynomials of degree at most n is left invariant:

$$P(z) = \sum_{j=0}^n p_j z^j \Rightarrow P^*(z) = \sum_{j=0}^n \bar{p}_{n-j} z^j. \quad (14.3)$$

(14.2) now implies

$$\deg(P) = n, \quad P \perp \zeta^j, \quad j = 1, \dots, n \Rightarrow P = c\Phi_n^*. \quad (14.4)$$

14.2. Szegő Recurrences and Verblunsky Coefficients. A key feature of the unit circle is that the multiplication operator $Uf = zf$ in $L^2_\mu(\mathbf{T})$ is unitary. So the difference $\Phi_{n+1}(z) - z\Phi_n(z)$ is of degree n and orthogonal to ζ^j for $j = 1, 2, \dots, n$, and by (14.4),

$$\Phi_{n+1}(z) = z\Phi_n(z) - \bar{\alpha}_n \Phi_n^*(z) \quad (14.5)$$

with some complex numbers α_n , called the *Verblunsky coefficients*. (14.5) is known as the *Szegő recurrences* after its first occurrence in the Szegő book [122]. (14.5) at $z = 0$ implies

$$\alpha_n = -\overline{\Phi_{n+1}(0)}. \quad (14.6)$$

Applying (14.3) to (14.5) yields

$$\Phi_{n+1}^*(z) = \Phi_n^*(z) - \alpha_n z \Phi_n(z). \quad (14.7)$$

It follows from the unitarity of U and $\Phi_n^* \perp \Phi_{n+1}$ that

$$\|\Phi_{n+1}\|^2 = (1 - |\alpha_n|^2) \|\Phi_n\|^2, \quad \|\Phi_n\|^2 = \prod_{j=0}^{n-1} (1 - |\alpha_j|^2), \quad (14.8)$$

and so $|\alpha_n| < 1$. Since it arises often, define

$$\rho_j := \sqrt{1 - |\alpha_j|^2}, \quad 0 < \rho \leq 1, \quad |\alpha_j|^2 + \rho_j^2 = 1. \quad (14.9)$$

Using (14.8) one can get the recursion relations for φ_n written in matrix form

$$\begin{pmatrix} \varphi_{n+1}(z) \\ \varphi_{n+1}^*(z) \end{pmatrix} = A(z, \alpha_n) \begin{pmatrix} \varphi_n(z) \\ \varphi_n^*(z) \end{pmatrix}, \quad A(z, \alpha) = \frac{1}{\rho} \begin{pmatrix} z & -\bar{\alpha} \\ -z\alpha & 1 \end{pmatrix} \quad (14.10)$$

or

$$\begin{pmatrix} \varphi_{n+1}(z) \\ \varphi_{n+1}^*(z) \end{pmatrix} = T_{n+1}(z) \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad T_p(z) = A(z, \alpha_{p-1}) \dots A(z, \alpha_0) \quad (14.11)$$

known as the *transfer matrix*.

Let \mathbf{D}^∞ be the set of complex sequences $\{\alpha_j\}_{j=0}^\infty$ with $|\alpha_j| < 1$. The map \mathcal{S} from $\mu \rightarrow \{\alpha_j(\mu)\}_{j=0}^\infty$ is a well-defined map from the set \mathcal{P} of non-trivial probability measures on \mathbf{T} to \mathbf{D}^∞ . The following fundamental result is proved in [131].

THEOREM 14.1 (Verblunsky’s Theorem). \mathcal{S} is a bijection.

As a matter of fact, \mathcal{S} is a homeomorphism if \mathcal{P} is given the weak- $*$ topology and in \mathbf{D}^∞ the topology of component convergence is considered.

The following result is usually attributed to Szegő, Kolmogorov and Krein.

THEOREM 14.2. For any non-trivial measure μ , the following are equivalent:

- (i) $\lim_{n \rightarrow \infty} \|\Phi_n\| = 0$
- (ii) $\sum_{n=0}^\infty |\alpha_n|^2 = \infty$
- (iii) $\{\varphi_n\}_{n=0}^\infty$ is a basis for $L^2_\mu(\mathbf{T})$
- (iv) $\int_{\mathbf{T}} \log \mu' \, d\mu = -\infty$, i.e., $\log \mu' \notin L^1(\mathbf{T})$.

14.3. Bernstein–Szegő Approximation. An interesting problem is to identify measures μ with finite sequences of Verblunsky coefficients: $\alpha_j(\mu) = 0$ for all large enough j .

THEOREM 14.3. Let μ be a non-trivial probability measure on \mathbf{T} with orthonormal polynomials φ_n . Let

$$\mu_n := \frac{d\mu}{|\varphi_n(\zeta)|^2}. \quad (14.12)$$

Then μ_n are probability measures with

$$\alpha_j(\mu_n) = \alpha_j(\mu), \quad j = 0, 1, \dots, n - 1; \quad \alpha_j(\mu_n) = 0, \quad j \geq n. \quad (14.13)$$

This result is often credited to Geronimus [41] even though it was proven (in different terms) by Verblunsky [132] ten years earlier. Since, for each fixed j , $\alpha_j(\mu_n) \rightarrow \alpha_j(\mu)$ (indeed, they are equal for $n > j$), $\mu_n \rightarrow \mu$ weakly since \mathcal{S} is a homeomorphism. It was Verblunsky who also found the Carathéodory function for measures (14.12):

$$F(z, \mu_n) = \int_{\mathbf{T}} \frac{\zeta + z}{\zeta - z} d\mu_n(\zeta) = \frac{\psi_n^*(z)}{\varphi_n^*(z)},$$

where the second kind polynomials ψ_n are the orthonormal polynomials with respect to the measure μ_{-1} with $\alpha_j(\mu_{-1}) = -\alpha_j(\mu)$.

In fact, the measures with finite sequences of Verblunsky coefficients are exactly those of the form $\mu = c|P(\zeta)|^{-2} d\mu$, where c is picked to make μ a probability measure, and P is a polynomial of degree n with all zeros in \mathbf{D} .

15. Schur, Geronimus, Khrushchev

15.1. Schur Functions and Algorithm. Given a probability measure μ on \mathbf{T} , define the *Carathéodory function* by

$$F(z, \mu) := \int_{\mathbf{T}} \frac{\zeta + z}{\zeta - z} d\mu(\zeta) = 1 + 2 \sum_{n=1}^\infty \beta_n z^n, \quad \beta_n = \int_{\mathbf{T}} \zeta^{-n} d\mu \quad (15.1)$$

the moments of μ . F is an analytic function in \mathbf{D} which obeys $\Re F > 0$, $F(0) = 1$. The Schur function is then defined by

$$f(z, \mu) = \frac{F(z) - 1}{z(F(z) + 1)}, \quad F(z) = \frac{1 + zf(z)}{1 - zf(z)}, \quad (15.2)$$

which is an analytic function in \mathbf{D} with $\sup_{\mathbf{D}} |f(z)| \leq 1$. So a one-one correspondence can be easily set up between the three classes (probability measures, Carathéodory and Schur functions). Under this correspondence μ is trivial, that is, supported on a finite set, if and only if the associated Schur function is a finite Blaschke product.

Let us proceed with the Schur algorithm. Given a Schur function $f = f_0$ which is not a finite Blaschke product, define inductively

$$f_{n+1}(z) = \frac{f_n(z) - \gamma_n}{z(1 - \bar{\gamma}_n z f_n(z))}, \quad \gamma_n = f_n(0). \quad (15.3)$$

It is clear that the sequence $\{f_n\}$ is an infinite sequence of Schur functions (called the n -th Schur iterates) and neither of its terms is a finite Blaschke product. The numbers $\{\gamma_n\}$ are called the Schur parameters.

The fundamental paper of Schur [94] had appeared a few years before Szegő introduced the notion of orthogonal polynomials on the unit circle (OPUC). Amazingly, neither of them benefited from the results of the other. And only twenty years later did Geronimus [40] put these ideas together and come up with the following:

THEOREM 15.1 (Geronimus' Theorem). *Let μ be a non-trivial probability measure on \mathbf{T} , f its Schur function and $\gamma_n(f)$ the Schur parameters of f . Then $\gamma_n(f) = \alpha_n(\mu)$.*

The latter formula explains why a minus and conjugate is taken in (14.5). As a straightforward consequence of this result, we see that $\gamma_j(f) = \gamma_j(g)$ for all j implies $f = g$. Furthermore, a nice relation between the moments β_n from (15.1) of the measure (Taylor coefficients of the Carathéodory function) and the Schur parameters (Verblunsky coefficients) is given by

$$\beta_n = \alpha_{n-1} \prod_{j=0}^{n-2} (1 - |\alpha_j|^2) + \text{polynomial in } (\alpha_0, \bar{\alpha}_0, \dots, \alpha_{n-2}, \bar{\alpha}_{n-2}).$$

15.2. Khrushchev's Theory. In two remarkable papers [53, 55], Khrushchev found deep connections between Schur iterates and the structure of OPUC. A key input for the theory is

THEOREM 15.2 (Khrushchev's Formula). *The Schur function for the measure $|\varphi_n|^2 d\mu$ is given by the product $b_n(z)f_n(z)$, where f_n is the n -th Schur iterate, and b_n is the finite Blaschke product*

$$b_n(z) = \frac{\varphi_n(z)}{\varphi_n^*(z)}.$$

The most important consequence of Khrushchev's formula is

THEOREM 15.3. *The essential support of the a.c. part of μ is all of \mathbf{T} if and only if*

$$\lim_{n \rightarrow \infty} \int_{\mathbf{T}} |f_n(\zeta)|^2 dm = 0,$$

dm is the normalized Lebesgue measure on \mathbf{T} .

Here are some other important results of Khrushchev’s theory.

THEOREM 15.4.

$$* - \lim_{n \rightarrow \infty} |\varphi_n|^2 d\mu = dm \Leftrightarrow \lim_{n \rightarrow \infty} \alpha_{n+j} \bar{\alpha}_n = 0, \quad j = 1, 2, \dots$$

Define the n -th Schur approximate $f^{[n]}$ by

$$\gamma_j(f^{[n]}) = \gamma_j(f), \quad j = 1, 2, \dots, n \quad \gamma_j(f^{[n]}) = 0, \quad j \geq n.$$

THEOREM 15.5. *Let $f^{[n]}$ be the n -th Schur approximate. Then*

$$\lim_{n \rightarrow \infty} \int_{\mathbf{T}} |f^{[n]}(\zeta) - f(\zeta)|^2 dm = 0$$

if and only if either μ is purely singular or $\alpha_n(\mu) \rightarrow 0$.

16. Szegő’s Theory and Extensions

Szegő’s theorems may well be the most celebrated in OPUC. They have repeatedly served as a source for further development. For historical reasons one should state them in terms of Toeplitz determinants, $D_n(\mu)$. This is defined as the determinant of the $(n + 1) \times (n + 1)$ matrix $\{\beta_{k-j}\}_{0 \leq k, j \leq n}$ with moments β ’s given in (15.1). The invariance of such determinants under triangular change of basis implies (using (14.8))

$$D_n(\mu) = \prod_{j=0}^n \|\Phi_j\|^2 = \prod_{j=0}^{n-1} (1 - |\alpha_j|^2)^{n-j},$$

and so

$$S(\mu) = \lim_{n \rightarrow \infty} (D_n(\mu))^{1/n} = \prod_{j=0}^{\infty} (1 - |\alpha_j|^2),$$

$$G(\mu) = \lim_{n \rightarrow \infty} \frac{D_n(\mu)}{S^{n+1}(\mu)} = \prod_{j=0}^{\infty} (1 - |\alpha_j|^2)^{-j-1}.$$

S is always defined and is a non-negative number. G is defined as long as $S > 0$ and is finite if and only if $\sum_{j=0}^{\infty} j|\alpha_j|^2 < \infty$.

Szegő’s theorems express S and G in terms of the a.c. and singular components of the Lebesgue decomposition of μ : $\mu = w dm + \mu_s$, $w \in L^1(\mathbf{T})$.

THEOREM 16.1 (Szegő’s Theorem).

$$S(\mu) = \prod_{j=0}^{\infty} (1 - |\alpha_j|^2) = \exp \left(\frac{1}{2\pi} \int_{\mathbf{T}} \log w(\zeta) dm \right). \tag{16.1}$$

Szegő proved this when $\mu_s = 0$ in 1915 (in his very first paper!). The result does not depend on μ_s —this was shown by Verblunsky [132].

It is immediate from Szegő’s theorem that

$$\sum_{j=0}^{\infty} |\alpha_j|^2 < \infty \Leftrightarrow \log w \in L^1(\mathbf{T}). \tag{16.2}$$

The equivalent conditions (16.2) are called the *Szegő condition*, and the corresponding class of measures is known as the *Szegő class*. Within this class the *Szegő function*

$$D(z, w) = \exp \left(\frac{1}{4\pi} \int_{\mathbf{T}} \frac{\zeta + z}{\zeta - z} \log w(\zeta) \, dm \right), \quad |z| < 1 \tag{16.3}$$

is well defined. Standard boundary value theory implies $D(\zeta) = \lim_{r \uparrow 1} D(rz)$ exists almost everywhere and $|D(\zeta)|^2 = w(\zeta)$ a.e. The main asymptotic result (due to Szegő) claims that

$$\lim_{n \rightarrow \infty} \varphi_n^*(z) = D^{-1}(z)$$

uniformly on compact subsets of \mathbf{D} .

THEOREM 16.2 (Strong Szegő Theorem). *If $\mu_s = 0$ and the Szegő condition holds, then*

$$G(\mu) = \prod_{j=0}^{\infty} (1 - |\alpha_j|^2)^{-j-1} = \exp \left(\sum_{n=0}^{\infty} n |w_n|^2 \right),$$

where w_n are the Fourier coefficients of $\log w$.

For an up-to-date approach to this result, see [108].

In a series of papers [74, 75, 76], Máté, Nevai, and Totik extended some parts of Szegő’s theory to the cases where the Szegő condition fails. Their main result can be viewed as a comparative asymptotics.

THEOREM 16.3. *Let $\mu = wdm + \mu_s$ be a non-trivial probability measure on \mathbf{T} obeying $w > 0$ a.e. Consider another probability measure $\nu = g d\mu$ for a non-negative function $g \in L^1(\mu)$. Suppose next that there is a polynomial Q so that $g^{\pm} Q \in L^{\infty}(\mu)$. Then*

$$\lim_{n \rightarrow \infty} \frac{\varphi_n^*(z, \nu)}{\varphi_n^*(z, \mu)} = D(z, g^{-1}),$$

(D is defined in (16.3)), uniformly on compact subsets of the unit disk \mathbf{D} .

Another natural extension of Szegő’s theory deals with ratio asymptotics.

THEOREM 16.4. *Let μ be a non-trivial probability measure on \mathbf{T} . Suppose*

$$\lim_{n \rightarrow \infty} \frac{\Phi_{n+1}^*(z)}{\Phi_n^*(z)} = G(z) \tag{16.4}$$

exists uniformly on compact subsets of \mathbf{D} . Then either $G \equiv 1$ or

$$G(z) = G_{a,\lambda}(z) = \frac{1 + \lambda z + \sqrt{(1 - \lambda z)^2 + 4a^2 \lambda z}}{2}$$

for some $\lambda \in \mathbf{T}$ and $a \in (0, 1]$.

(16.4) holds with $G = G_{a,\lambda}$ if and only if $\alpha_n(\mu)$ obeys the López condition

$$\lim_{n \rightarrow \infty} |\alpha_n| = a, \quad \lim_{n \rightarrow \infty} \frac{\alpha_{n+1}}{\alpha_n} = \lambda.$$

In this case the essential support of μ is an arc, and (16.4) holds uniformly on compact subsets of $\mathbf{C} \setminus \text{supp } \mu$.

The first statement is due to Khrushchev [55] and the second one to Barrios and López [10].

The following extension of the above result, which can be viewed as *relative ratio asymptotics*, was proved in [47].

THEOREM 16.5. *Let μ and ν be two non-trivial probability measures on \mathbf{T} . Let $\{\alpha_n(\mu)\}$ and $\{\alpha_n(\nu)\}$, respectively, be their Verblunsky coefficients, and let $\Phi_n^*(\mu)$ and $\Phi_n^*(\nu)$, respectively, be their reversed monic orthogonal polynomials. Then*

$$\frac{\Phi_{n+1}^*(\mu, z)}{\Phi_n^*(\mu, z)} - \frac{\Phi_{n+1}^*(\nu, z)}{\Phi_n^*(\nu, z)} \rightarrow 0 \tag{16.5}$$

uniformly on compact subsets of \mathbf{D} as $n \rightarrow \infty$ if and only if for any $\ell \geq 1$,

$$\lim_{n \rightarrow \infty} [\alpha_n(\mu)\bar{\alpha}_{n-\ell}(\mu) - \alpha_n(\nu)\bar{\alpha}_{n-\ell}(\nu)] = 0. \tag{16.6}$$

17. CMV Matrices

One of the most interesting developments in the theory of OPUC in recent years is the discovery by Cantero, Moral, and Velázquez [16, 17] of a matrix realization for multiplication by ζ on $L^2_\mu(\mathbf{T})$ which is of finite band size (i.e., $|\langle \zeta\chi_m, \chi_n \rangle| = 0$ if $|m - n| > k$ for some k); in this case, $k = 2$ to be compared with $k = 1$ for the Jacobi matrices which correspond to the real line case. The CMV basis (complete, orthonormal system) $\{\chi_n\}$ is obtained by orthonormalizing the sequence $1, \zeta, \zeta^{-1}, \zeta^2, \zeta^{-2}, \dots$, and the matrix, called the *CMV matrix*,

$$\mathcal{C}(\mu) = \|\mathcal{C}_{n,m}\|_{m,n=0}^\infty = \langle \zeta\chi_m, \chi_n \rangle, \quad m, n \in \mathbf{Z}_+$$

is five-diagonal. Remarkably, the χ 's can be expressed in terms of φ 's and φ^* 's

$$\chi_{2n} = z^{-n}\varphi_{2n}^*(z), \quad \chi_{2n+1} = z^{-n}\varphi_{2n+1}(z), \quad n \in \mathbf{Z}_+,$$

and the matrix elements in terms of α 's and ρ 's: $\mathcal{C} = LM$ where L, M are block diagonal matrices

$$L = \text{Diag}(\Theta_0, \Theta_2, \Theta_4, \dots), \quad M = \text{Diag}(1, \Theta_1, \Theta_3, \dots) \tag{17.1}$$

with

$$\Theta_j = \begin{pmatrix} \bar{\alpha}_j & \rho_j \\ \rho_j & -\alpha_j \end{pmatrix}, \quad j = 0, 1, \dots \tag{17.2}$$

(the first block of M is 1×1). By \mathcal{C}_0 we denote the CMV matrix for the Lebesgue measure dm .

There is an important relation between CMV matrices and monic orthogonal polynomials akin to the well-known property of orthogonal polynomials on the real line:

$$\Phi_n(z) = \det(zI_n - \mathcal{C}^{(n)}), \tag{17.3}$$

where $\mathcal{C}^{(n)}$ is the principal $n \times n$ block of \mathcal{C} . The CMV representation provides one of the proofs of Verblunsky's theorem.

The natural extension of L, M, \mathcal{C} to doubly infinite matrices proves helpful in some problems related to periodic and stochastic Verblunsky coefficients. For the standard basis $\{e_j\}_{j \in \mathbf{Z}}$, denote by E_k the two-dimensional subspace spanned by $\{e_k, e_{k+1}\}$, so

$$\ell^2(\mathbf{Z}) = \bigoplus_{j \in \mathbf{Z}} E_{2j} = \bigoplus_{j \in \mathbf{Z}} E_{2j+1}.$$

Let the operator Θ_k act in E_k by (17.2). We come to

$$\hat{L} = \bigoplus_{j \in \mathbf{Z}} \Theta_{2j}, \quad \hat{M} = \bigoplus_{j \in \mathbf{Z}} \Theta_{2j+1}, \quad \hat{\mathcal{C}} = \hat{L}\hat{M}. \tag{17.4}$$

The Weyl–Titchmarsh theory for such CMV operators is developed in [42].

The CMV matrices play much the same role in the study of OPUC that Jacobi matrices do in orthogonal polynomials on the real line.

Part 3. BARRY SIMON’S CONTRIBUTIONS

18. Analysis of CMV Matrices

18.1. CMV Matrices and Spectral Analysis. Perturbation theory involves looking at similarities of measures when their Verblunsky coefficients are close in some suitable sense. The CMV matrices provide a powerful tool for the comparison of properties of two measures μ_1 and μ_2 if we have some information about $\alpha_n(\mu_2)$ as a perturbation of $\alpha_n(\mu_1)$. Of course, this idea is standard in the real line setting where Jacobi matrices do the job.

Put

$$\alpha_n(\mu_1, \mu_2) := |\alpha_n(\mu_1) - \alpha_n(\mu_2)|, \quad \rho_n(\mu_1, \mu_2) := |\rho_n(\mu_1) - \rho_n(\mu_2)|.$$

An easy estimate using the LM factorization shows with $\|\cdot\|_p$ the S_p trace ideal norm

LEMMA 18.1. *For all $1 \leq p \leq \infty$,*

$$\|\mathcal{C}(\mu_1) - \mathcal{C}(\mu_2)\|_p \leq 6 \left(\sum_{n=0}^{\infty} \alpha_n^p(\mu_1, \mu_2) + \rho_n^p(\mu_1, \mu_2) \right)^{1/p}. \tag{18.1}$$

For $p = \infty$, the right-hand side of (18.1) is interpreted as

$$\sup_n \max(\alpha_n(\mu_1, \mu_2), \rho_n(\mu_1, \mu_2)).$$

This result (joint with L. Golinskii) allows one to translate the ideas of Simon–Spencer to a new operator theoretic proof of the following result called Rakhmanov’s lemma: μ is purely singular whenever $\limsup_n |\alpha_n| = 1$.¹ It is also a key ingredient in proving the following (see [106, Sect. 4.3]).

THEOREM 18.2. *Let $\mu_j = w_j dm + \mu_{j,s}$, $j = 1, 2$.*

1. *If $\lim_{n \rightarrow \infty} \alpha_n(\mu_1, \mu_2) = 0$, then the supports of μ_1 and μ_2 have identical sets of limit points: $(\text{supp} \mu_1)' = (\text{supp} \mu_2)'$.*
2. *If $\sum_n \alpha_n(\mu_1, \mu_2) < \infty$, then up to sets of Lebesgue measure zero,*

$$\{\zeta : w_1(\zeta) \neq 0\} = \{\zeta : w_2(\zeta) \neq 0\}.$$

Concerning the essential spectrum of CMV operators, see [60].

18.2. CMV Matrices and the Szegő Function. An intimate relation between the CMV matrices and the Szegő functions (16.3) is presented in [106, Sect. 4.2]. Let

$$\log D(z) = \frac{1}{2} w_0 + \sum_{n=1}^{\infty} w_n z^n, \quad w_k = \int_{\mathbf{T}} \log w(\zeta) \zeta^{-k} dm \tag{18.2}$$

(w_k are the Fourier coefficients of $\log w$).

THEOREM 18.3. (i). *Assume that the Verblunsky coefficients $\{\alpha_n\} \in \ell^2$. Then*

1. *$\mathcal{C} - \mathcal{C}_0$ is Hilbert–Schmidt;*

¹A similar result for the Jacobi matrices appeared prior to Simon–Spencer in [27].

2.

$$\frac{D(0)}{D(z)} = \det_2 \left(\frac{I - z\bar{C}}{I - z\bar{C}_0} \right) e^{zw_1}, \quad w_1 = \alpha_0 - \sum_{n=1}^{\infty} \alpha_n \bar{\alpha}_{n-1}$$

with \det_2 being the renormalized determinant for the Hilbert–Schmidt class;

3.

$$\bar{w}_n = \frac{\text{Tr}(C^n - C_0^n)}{n}, \quad n \geq 2.$$

(ii). Assume that $\{\alpha_n\} \in \ell^1$. Then

1. $C - C_0$ is trace class;
- 2.

$$\frac{D(0)}{D(z)} = \det \left(\frac{I - z\bar{C}}{I - z\bar{C}_0} \right);$$

3.

$$\bar{w}_n = \frac{\text{Tr}(C^n - C_0^n)}{n}, \quad n \geq 1.$$

In either case,

$$\bar{w}_n = \sum_{j=0}^{\infty} \frac{(C^n)_{jj}}{n}.$$

19. Zeros

19.1. Limit Sets of Zeros. The structure of zero sets for OPUC is another fascinating topic of the theory. Given a non-trivial probability measure $\mu \in \mathcal{P}$, denote by $Z_n(\mu)$ the zero set for Φ_n :

$$Z_n(\mu) = \{z_{jn}\}_{j=1}^n, \quad |z_{nn}| \leq |z_{n-1,n}| \leq \dots \leq |z_{1,n}|, \quad \Phi_n(z_{jn}, \mu) = 0.$$

The basic property of zeros reads that $|z_{1,n}| < 1$, i.e., $Z_n(\mu) \subset \mathbf{D}$. Indeed, let $z_0 \in Z_n$ and define $P = \Phi_n/(z - z_0)$. Since $\deg P = n - 1$, $P \perp \Phi_n$ and so

$$\|P\|^2 = \|zP\|^2 = \|z_0P + \Phi_n\|^2 = |z_0|^2 \|P\|^2 + \|\Phi_n\|^2.$$

Hence $\|\Phi_n\|^2 = (1 - |z_0|^2)\|P\|^2$, as needed.

In 1988 Alfaro and Vigil [2] answering a question of Turán showed that for an arbitrary sequence of points $\{z_k\}$ in \mathbf{D} , there exists a unique measure $\mu \in \mathcal{P}$ with $\Phi_n(z_n, \mu) = 0$. Hence the total set of zeros of Φ_n 's $Z_\infty(\mu) = \cup_n Z_n(\mu)$ can be dense in \mathbf{D} . A vast generalization of the Alfaro–Vigil theorem is due to Simon–Totik [111].

THEOREM 19.1. *For an arbitrary sequence of points $\{z_k\}$ in \mathbf{D} and arbitrary sequence of positive integers $0 < m_1 < m_2 < \dots$, there exists $\mu \in \mathcal{P}$ such that $\Phi_{m_k}(z_j, \mu) = 0$ for $j = m_{k-1} + 1, \dots, m_k$.*

The following consequence of this result may seem kind of amazing. Let a measure μ belong to the class of non-trivial probability measures \mathcal{P} . Consider the sequence $\{\nu_n(\mu)\}_{n \geq 1}$ of normalized counting measures for zeros of Φ_n , i.e.,

$$\text{supp } \nu_n = Z_n, \quad \nu_n\{z_{jn}\} = \frac{l(z_{jn})}{n} \tag{19.1}$$

with $l(z_{jn})$ equal to the multiplicity of the zero z_{jn} . Let $\mathcal{M}_+(\bar{\mathbf{D}})$ be a space of probability measures on $\bar{\mathbf{D}}$ endowed with the weak-* topology. A measure $\mu \in \mathcal{P}$ is said to be *universal* if for each $\sigma \in \mathcal{M}_+(\bar{\mathbf{D}})$ there is a sequence of indices n_j such

that $\nu_n(\mu)$ converges to σ as $j \rightarrow \infty$ in weak-* topology. The existence of universal measures is proved in [111, Cor. 3].

It is known that zeros of $\Phi_n(\mu)$ cluster to the support of the orthogonality measure μ as long as the support is not the whole circle. The situation changes dramatically if $\text{supp } \mu = \mathbf{T}$ (see, e.g., the Lebesgue measure). By the Alfaro–Vigil, theorem zeros of Φ_n can cluster to all points of $\bar{\mathbf{D}}$. Denote by

$$Z_w(\mu) := \{z \in \bar{\mathbf{D}} : \liminf_{n \rightarrow \infty} \text{dist}(z, Z_n) = 0\}$$

the point set of limit points for the zeros of all Φ_n (weakly attracting points). Let $Z_w = \{Z_w(\mu)\}_{\mu \in \mathcal{P}}$ be the class of all such point sets. So, $\bar{\mathbf{D}} \in Z_w$. It turns out that Z_w is rich enough. More precisely, each compact subset K of \mathbf{D} belongs to Z_w , and the same is true if $K \supset \mathbf{T}$ ([111, Thm. 4]). On the other hand, $K = [\frac{1}{2}, 1]$ is not in Z_w .

Similarly, denote by

$$Z_s(\mu) := \{z \in \bar{\mathbf{D}} : \lim_{n \rightarrow \infty} \text{dist}(z, Z_n) = 0\}$$

the point set of strongly attracting points, and Z_s the class of all such point sets. The structure of the latter is quite different from that of Z_w . For instance, it is proved in [1] that if $0 \in Z_s(\mu)$ for some $\mu \in \mathcal{P}$, then $Z_s(\mu)$ is at most a countable set which converges to the origin. So, e.g., the disk $\{|z| \leq \frac{1}{2}\}$ is not in Z_s .

19.2. Mhaskar–Saff and Clock Theorems for Zeros. A remarkable theorem of Mhaskar and Saff [78] provides some information about the limit points (in the space $\mathcal{M}_+(\bar{\mathbf{D}})$) of the sequence of counting measures of zeros associated with a non-trivial probability measure $\mu \in \mathcal{P}$ in the case when Verblunsky coefficients tend to zero fast enough.

THEOREM 19.2 (Mhaskar–Saff). *Let*

$$A := \limsup_{n \rightarrow \infty} |\alpha_n(\mu)|^{1/n} = \lim_{j \rightarrow \infty} |\alpha_{n_j}(\mu)|^{1/n_j}.$$

Suppose that either $A < 1$ or $A = 1$ and $\sum_{j=0}^{n-1} |\alpha_j(\mu)| = o(n)$ as $n \rightarrow \infty$. Then $\{\nu_n(\mu)\}$ converges to the uniform measure on the circle of radius A .

Simon suggested a new approach to this result based on CMV matrices instead of potential theory. A key relation which links the two subjects is obvious from (17.3):

$$\frac{1}{n} \text{Tr}(\mathcal{C}^{(n)})^k = \int_{\mathbf{D}} z^k d\nu_n(\mu).$$

THEOREM 19.3. *Let $\sum_{j=0}^{n-1} |\alpha_n(\mu_1) - \alpha_n(\mu_2)| = o(n)$, $n \rightarrow \infty$. Then for each $k \in \mathbf{Z}_+$,*

$$\lim_{n \rightarrow \infty} \int_{\mathbf{D}} z^k (d\nu_n(\mu_1) - d\nu_n(\mu_2)) = 0.$$

In particular, if $\nu_n(\mu_1)$ tends to ν and γ is any limit point of $\nu_n(\mu_2)$, then

$$\int_{\mathbf{D}} z^k d\gamma = \int_{\mathbf{D}} z^k d\nu.$$

A crucial feature of the Mhaskar–Saff theorem is its universality. Under its assumption the angular distribution is the same. To get certain quantitative bounds on the distance between zeros, Simon studies various more stringent conditions, and among them the so-called *Barrios–López–Saff condition*

$$\alpha_n(\mu) = Cb^n + O((b\delta)^n); \quad C \in \mathbf{C} \setminus 0, \quad 0 < b, \delta, < 1. \tag{19.2}$$

The following result is proved in [110].

THEOREM 19.4. *Under the assumption (19.2), $\Phi_n(\mu)$ has a finite number J of “spurious” zeros off the circle $|z| = b$ for all large n . Furthermore, let for $j = 1, 2, \dots, n - J$,*

$$z_{jn} = |z_{jn}|e^{i\Theta_{jn}}; \quad 0 = \Theta_{0n} < \Theta_{1n} < \dots < \Theta_{n-J,n} < 2\pi = \Theta_{n-J+1,n}$$

be the other zeros. Then the limit relations hold

$$\sup_{1 \leq j \leq n-J} ||z_{jn}| - b| = O\left(\frac{\log n}{n}\right), \quad n \rightarrow \infty; \tag{19.3}$$

$$\sup_{1 \leq j \leq n-J} n \left| \Theta_{j+1,n} - \Theta_{jn} - \frac{2\pi}{n} \right| = o(1), \quad n \rightarrow \infty; \tag{19.4}$$

$$\frac{|z_{j+1,n}|}{|z_{jn}|} = 1 + O\left(\frac{1}{n \log n}\right), \quad n \rightarrow \infty. \tag{19.5}$$

Note that (19.4)–(19.5) imply $\lim_n n|z_{j+1,n} - z_{jn}| = 2\pi b$. Amazingly, the spurious zeros also follow the clock pattern!

In [102] Simon treats the more general case

$$\alpha_n(\mu) = \sum_{l=1}^m C_l e^{in\Theta_l} b^n + O((b\delta)^n).$$

20. Spectral Theory in Special Classes

20.1. High-Order Szegő Theorem. Simon came up with the idea of extending Szegő’s theorem by allowing “Pollaczek singularities” (so all quantities in (16.1) may be infinite). His result can be viewed as the first-order Szegő’s theorem: for any $\zeta_0 \in \mathbf{T}$,

$$|\zeta - \zeta_0|^2 \log w \in L^1(\mathbf{T}) \Leftrightarrow \sum_{j=0}^{\infty} |\alpha_{j+1} - \bar{\zeta}_0 \alpha_j|^2 + |\alpha_j|^4 < \infty.$$

Moreover, there is a precise formula for this case similar to the second equality in (16.1) [106, Sect. 2.8]. The second-order Szegő’s theorem appeared in [113]. Let $\zeta_k \in \mathbf{T}$, $k = 1, 2$. Then for $\zeta_1 \neq \zeta_2$

$$|\zeta - \zeta_1|^2 |\zeta - \zeta_2|^2 \log w \in L^1(\mathbf{T}) \Leftrightarrow \sum_{j=0}^{\infty} |\alpha_{j+2} - (\bar{\zeta}_1 + \bar{\zeta}_2)\alpha_{j+1} + \bar{\zeta}_1 \bar{\zeta}_2 \alpha_j|^2 + |\alpha_j|^4 < \infty,$$

and for $\zeta_1 = \zeta_2$,

$$|\zeta - \zeta_1|^4 \log w \in L^1(\mathbf{T}) \Leftrightarrow \sum_{j=0}^{\infty} |\alpha_{j+2} - 2\bar{\zeta}_1 \alpha_{j+1} + \bar{\zeta}_1^2 \alpha_j|^2 + |\alpha_j|^6 < \infty.$$

The general nice-looking conjecture called the *higher-order Szegő's theorem* is put forward in [106, Sect. 2.8]. Given $\zeta_k \in \mathbf{T}$, $k = 1, \dots, n$ and $\zeta_p \neq \zeta_q$, $p \neq q$, define a polynomial

$$P(\zeta) := \prod_{k=1}^n (\zeta - \zeta_k)^{m_k}, \quad m_k \in \mathbf{N}, \quad \bar{P}(\zeta) := \prod_{k=1}^n (\zeta - \bar{\zeta}_k)^{m_k},$$

and put $m = 1 + \max_k m_k$. Then

$$|P(\zeta)|^2 \log w \in L^1(\mathbf{T}) \Leftrightarrow (\bar{P}(S)) \{\alpha_j\} \in \ell^2 \text{ and } \{\alpha_j\} \in \ell^{2m},$$

where S is the shift operator: $S(\alpha_0, \alpha_1, \dots) = (\alpha_1, \alpha_2, \dots)$. A measure $\mu = wdm + \mu_s \in \mathcal{P}$ belongs to the polynomial Szegő class if $|P(\zeta)|^2 \log w \in L^1(\mathbf{T})$. For further advances concerning the polynomial Szegő class, in particular, the asymptotics inside the disk and L^2 on the circle, see [25].

The following particular case of Simon's conjecture which can be called the *higher-order Szegő's theorem in ℓ^4* has been proved recently in [47].

THEOREM 20.1. *Assume that $\{\alpha_j\} \in \ell^4$. Then*

$$|P(\zeta)|^2 \log w \in L^1(\mathbf{T}) \Leftrightarrow (\bar{P}(S)) \{\alpha_j\} \in \ell^2.$$

20.2. Baxter's Theorem. The celebrated paper [11] by Baxter appears to be one of the cornerstones of OPUC theory. He was interested in general complex-valued weight functions and corresponding non-Hermitian Toeplitz matrices of moments in connection with two sets of Verblunsky coefficients. Applied to the case of OPUC, his basic result looks as follows.

THEOREM 20.2 (Baxter). *Let $\mu = wdm + \mu_s$ be a non-trivial probability measure on \mathbf{T} with Verblunsky coefficients $\alpha_n(\mu)$ and moments $\beta_n(\mu)$. Then the following are equivalent:*

1. $\sum_{j=0}^\infty |\alpha_j(\mu)| < \infty$
2. $\sum_{j=0}^\infty |\beta_j(\mu)| < \infty$, $\mu_s = 0$, w is continuous, and $\min w(\zeta) > 0$.

A new approach to Baxter's theorem with some extensions is suggested in [106, Ch. 5].

A function $\nu(n)$, $n \in \mathbf{Z}$, is called a *Beurling weight* if

$$\nu(-n) = \nu(n), \quad \nu(n) \geq 1, \quad \nu(m+n) \leq \nu(m)\nu(n),$$

and a *strong Beurling weight* if in addition,

$$A(\nu) := \lim_{n \rightarrow \infty} \frac{\log \nu(n)}{n} = 0.$$

Examples include $\nu(n) = (1 + |n|)^\alpha$ (strong) and $\nu(n) = e^{\alpha|n|}$, $\alpha \geq 0$.

Given a Beurling weight ν , the *Beurling algebra* ℓ_ν is the Banach algebra of two-sided sequences $\{a(n)\}_{n \in \mathbf{Z}}$ with standard operations (addition and convolution) and

$$\|a\|_\nu := \sum_{n \in \mathbf{Z}} |a(n)|\nu(n) < \infty.$$

THEOREM 20.3. *Let ν be a strong Beurling weight. Then the following are equivalent:*

1. $\{\alpha_n(\mu)\} \in \ell_\nu$
2. $\{\beta_n(\mu)\} \in \ell_\nu$, $\mu_s = 0$, w is continuous, and $\min w(\zeta) > 0$.

A number of equivalent conditions are displayed in [106, Thm. 5.2.2]. As an immediate consequence of this result, Simon obtains

COROLLARY 20.4. *Let $k \in \mathbb{N}$. Then*

1. *If $\sum_n (n + 1)^k |\alpha_n(\mu)| < \infty$, then $\mu = wdm$ with $\inf w(\zeta) > 0$ and w a C^k function.*
2. *If $k \geq 2$ and $\mu = wdm$ with $\inf w(\zeta) > 0$, and w is a C^k function, then $\sum_n (n + 1)^\beta |\alpha_n(\mu)| < \infty$ for any $0 \leq \beta < k - 1$.*

In particular, $\mu = wdm$ with $\inf w(\zeta) > 0$, and w a C^∞ function if and only if for all k ,

$$|\alpha_n(\mu)| \leq C_k (n + 1)^{-k}.$$

The last statement is due to B. Golinskii [44]. For some further extensions of Baxter’s theorem, see [37].

20.3. B. Golinskii–Ibragimov Condition. In 1971 B. Golinskii and Ibragimov proved that if the Verblunsky coefficients $\{\alpha_n(\mu)\}$ of a measure μ obey $\sum_n n |\alpha_n(\mu)|^2 < \infty$ then μ is absolutely continuous. The extension of their ideas led Simon [97] to the class of measures for which $\sum_{n=0}^N n |\alpha_n(\mu)|^2$ diverges logarithmically. More precisely, Simon’s condition is

$$\sum_{n=0}^N n |\alpha_n(\mu)|^2 \leq A \log N + C \tag{20.1}$$

with A, C constants. He proved in [106, Thm. 6.1.7], that μ is absolutely continuous as long as (20.1) holds with some $A < \frac{1}{4}$. On the other hand, for any $A > \frac{1}{4}$, there are examples with $\mu_\sigma \neq 0$, that is, $\frac{1}{4}$ is optimal. As a matter of fact, for $A \geq \frac{1}{4}$, one should distinguish the pure point component μ_{pp} of $\mu_s = \mu_{pp} + \mu$ from the singular continuous one, μ . For instance, it is shown in [106, Cor. 2.7.6] that $\mu_{pp} = 0$ as long as (20.1) holds with $A = \frac{1}{4}$. It turns out that for measures (20.1) with some $A < \infty$, the support of μ_{pp} consists of at most $[4A]$ points, $[x]$ being an integer part of x (see [107, Theorem 10.12.7]). Recently, Damanik [19], answering a question of Simon, proved that, indeed, $\mu = 0$ whenever (20.1) holds for some $A < \infty$.

20.4. Sparse Verblunsky Coefficients. In [58] Kiselev, Last and Simon presented a thorough analysis of continuum and discrete Schrödinger operators with sparse potentials. In [107, Sect. 12.5] Simon found the analogues of these results for OPUC. A set of Verblunsky coefficients is said to be *sparse* if $\alpha_j \neq 0$ only for $j \in \{n_k\}_{k \geq 1}$ where $n_{k+1} - n_k \rightarrow \infty$ perhaps at some specific rate. The first result of the “sparse flavor” for OPUC is due to Khrushchev [53], who proved that under the so-called Máté–Nevai condition,

$$\lim_{n \rightarrow \infty} \alpha_n(\mu) \bar{\alpha}_{n-l}(\mu) = 0$$

for each fixed $l = 1, 2, \dots$ (which certainly holds for sparse sets and characterizes Rakhmanov’s class of measures) μ is singular as long as $\limsup_{n \rightarrow \infty} |\alpha_n(\mu)| > 0$.

For the proof of the next two results, see [107, Thms. 12.5.1, 12.5.2].

THEOREM 20.5. *Let $\{n_k\}$ be a monotone sequence of positive integers with $\liminf_{k \rightarrow \infty} n_{k+1}/n_k > 1$. Suppose*

$$\alpha_j(\mu) = 0, \quad j \notin \{n_k\}, \quad \sum_{j=0}^{\infty} |\alpha_k(\mu)|^2 < \infty.$$

Then $\text{supp } \mu = \mathbf{T}$ and $\mu_s = 0$, i.e., μ is absolutely continuous. Moreover, $\mu = wdm$ with $w^{\pm 1} \in L^p(\mathbf{T})$ for all $1 \leq p < \infty$.

THEOREM 20.6. *Let $\{n_k\}$ be a monotone sequence of positive integers with $\liminf_{k \rightarrow \infty} n_{k+1}/n_k = \infty$. Suppose*

$$\alpha_j(\mu) = 0, \quad j \notin \{n_k\}, \quad \lim_{j \rightarrow \infty} \alpha_j = 0, \quad \sum_{j=0}^{\infty} |\alpha_k(\mu)|^2 = \infty.$$

Then $\text{supp } \mu = \mathbf{T}$ and $\mu_{pp} = \mu_{ac} = 0$, i.e., μ is purely singular continuous.

20.5. Dense Embedded Point Spectrum. The story for Schrödinger operators goes back to Naboko [79], who constructed the operator $-\frac{d^2}{dx^2} + V(x)$ with V decaying only slightly slower than $|x|^{-1}$ but there is dense embedded point spectrum. Naboko’s method extends to OPUC; see [107, Sect. 12.3].

THEOREM 20.7. *Let $g(n)$ be an arbitrary function with $0 < g(n) \leq g(n + 1)$ and $g(n) \rightarrow \infty$ as $n \rightarrow \infty$. Let $\{\omega_j\}_{j \geq 0}$ be an arbitrary subset of \mathbf{T} which are multiplicatively rationally independent, that is, for no integers n_1, \dots, n_k other than zeros is it true that $\prod_{j=1}^k \omega_j^{n_j} = 1$. Then there exists a sequence $\{\alpha_j(\mu)\}$ of Verblunsky coefficients with*

$$|\alpha_n(\mu)| \leq \frac{g(n)}{n}$$

for all n so that μ has pure points at each ω_j .

Note that if $g(n) \leq n^{1/2-\epsilon}$, then $\alpha_n \in \ell^2$ so by Szegő’s theorem, $w > 0$ a.e., that is, the point spectrum is embedded in a.c. spectrum.

20.6. Fibonacci Subshifts. There is an extensive literature on subshifts for discrete Schrödinger operators. Simon [107, Sect. 12.8] has analyzed the OPUC analogues of the most heavily studied of these subshifts, defined as follows. Pick $\alpha \neq \beta$ in the open unit disk and let $F_1 = \alpha$, $F_2 = \alpha\beta$, and $F_{n+1} = F_n F_{n-1}$ for $n = 2, 3, \dots$. These are finite strings built up from α and β . As F_{n+1} starts with F_n , there is a limit $F(\alpha, \beta)$ which is a one-way infinite string $\alpha\alpha\beta\alpha\beta\alpha \dots$. Let $\mu(\alpha, \beta)$ be the measure that has Verblunsky coefficients (in an infinite string form) $\alpha_0\alpha_1 \dots = F(\alpha, \beta)$.

THEOREM 20.8. *The essential support of the measure $\mu(\alpha, \beta)$ is a closed perfect set of Lebesgue measure zero. For fixed α_0, β_0 and a.e. $\lambda \in \mathbf{T}$, the measure $\mu(\lambda\alpha_0, \lambda\beta_0)$ is a pure point measure, with each pure point isolated and the limit points of the pure points form a perfect set of Lebesgue measure zero.*

21. Periodic Verblunsky Coefficients

The theory of one-dimensional periodic Schrödinger operators (also known as Hill’s equations) and of periodic Jacobi matrices has been extensively developed. The theory up to the 1950’s is summarized in the Magnus–Winkler book [67]. There was an explosion of ideas following the KdV revolution, including spectrally invariant flows and Abelian functions on hyperelliptic Riemann surfaces. The ideas have been carried over to the discrete setting of orthogonal polynomials on the real line; see, e.g., Toda [125] and Flaschka–McLaughlin [30].

In the 1940's Geronimus [39] found the earliest results on OPUC with periodic Verblunsky coefficients, i.e., for some $p \geq 1$,

$$\alpha_{j+p} = \alpha_j, \quad j \in \mathbf{Z}_+. \tag{21.1}$$

In particular, the case $\alpha_j \equiv \alpha \in \mathbf{D} \setminus \{0\}$ yields OPUC called Geronimus polynomials (see [106, Ex. 1.6.12]). Many of the general features for OPUC obeying (21.1) appeared in a fundamental series of papers by Peherstorfer and collaborators.

The aforementioned literature on OPUC used little from the work on Hill's equation. A partial link is Geronimo–Johnson [35], which discussed almost periodic Verblunsky coefficients using Abelian functions. See also Geronimo–Gesztesy–Holden [34], which includes work on isospectral flows.

Some new results and approaches for periodic Verblunsky coefficients are presented in [107, Ch. 11].

21.1. Discriminant. For Schrödinger operators and Jacobi matrices, the discriminant is known to be just the trace of the transfer matrix, which has determinant one in this case. For OPUC, the transfer matrix T_p (14.11) has $\det(T_p(z)) = z^p$, so it is natural to define the *discriminant* by

$$\Delta(z) := \text{Tr}(z^{-p/2} T_p(z)) \tag{21.2}$$

which explains why p is normally assumed to be an even number. Δ is known to be real on \mathbf{T} , so

$$\Delta(1/\bar{z}) = \overline{\Delta(z)},$$

and $z^{-p/2} T_p(z)$ has eigenvalues $\frac{\Delta}{2} \pm i\sqrt{1 - (\frac{\Delta}{2})^2}$. In particular, these eigenvalues have magnitude 1, that is, $\sup_m \|T_{mp}(z)\| < \infty$ exactly when $\Delta \in [-2, 2]$.

THEOREM 21.1. *There exist $\{x_j\}_{j=1}^{2p}, \{y_j\}_{j=1}^{2p}$ with*

$$x_1 < y_1 \leq x_2 < y_2 \leq \dots \leq x_p < y_p \leq x_1 + 2\pi \equiv x_{p+1}$$

so that the solutions of $\Delta(z) = 2$ (resp. -2) are exactly $e^{ix_1}, e^{iy_2}, e^{ix_3}, \dots, e^{ix_p}$ (resp. $e^{iy_1}, e^{ix_2}, e^{iy_3}, \dots, e^{iy_p}$) and $\Delta(z) \in [-2, 2]$ exactly on the bands

$$B = \bigcup_{j=1}^p B_j, \quad B_j = \{e^{i\theta} : x_j \leq \theta \leq y_j\}. \tag{21.3}$$

B is the essential support of μ_{ac} and the only possible singular spectrum are mass points which can occur in open gaps (i.e., non-empty sets of the form $\{e^{i\theta} : y_j \leq \theta \leq x_{j+1}\}$) with at most one mass point in each gap.

Further properties of Δ are given in the following result.

THEOREM 21.2. (i) *For all non-zero $z \in \mathbf{C}$, the Lyapunov exponent is*

$$\gamma(z) = \lim_{n \rightarrow \infty} \|T_n(z)\|^{1/n} = \frac{1}{2} \log(z) + \frac{1}{p} \left| \frac{\Delta(z)}{2} + \sqrt{\frac{\Delta^2(z)}{4} - 1} \right|, \tag{21.4}$$

where the branch of square root is taken that maximizes the logarithm.

(ii) *The logarithmic capacity of B (21.3) is given by*

$$C_B = \prod_{j=0}^{p-1} (1 - |\alpha_j|^2)^{1/p}$$

and $-\lceil \gamma + \log C_B \rceil$ is the equilibrium potential for B .

(iii) *The density of zeros is the equilibrium measure for B and is given in terms of Δ by $d\nu = Vd\theta/2\pi$, where*

$$V(\theta) = \frac{1}{p} \frac{|\Delta'(e^{i\theta})|}{\sqrt{4 - \Delta^2(e^{i\theta})}}$$

on B_j , and $V = 0$ on the complement of B .

(iv) $\nu(B_j) = 1/p$.

For any $\{\alpha_j\}_{j=0}^{p-1} \in \mathbf{D}^p$, one can define a discriminant $\Delta(z, \{\alpha_j\}_{j=0}^{p-1})$ for the p -periodic Verblunsky coefficients that agree with $\{\alpha_j\}_{j=0}^{p-1}$ for $j = 0, \dots, p - 1$. It turns out (see [107, Thm. 11.10.2]) that the set $\{\alpha_j\}_{j=0}^{p-1} \in \mathbf{D}^p$ for which Δ has all gaps open, i.e., all B_j in (21.3) are disjoint, is a dense open set in \mathbf{D}^p (the case of generic potentials). [107] has two proofs of this result: one is based on Sard’s theorem and one is a perturbation theoretic calculation.

In 1946 Borg [15] proved two celebrated theorems for Schrödinger operators about the implication of closed gaps. In [107, Sect. 11.14] Simon gives the analogues of these results for OPUC.

THEOREM 21.3. (i) *Let α_j be a periodic sequence of Verblunsky coefficients. Suppose all gaps are closed, i.e., the support of the orthogonality measure is the whole unit circle. Then $\alpha_j \equiv 0$.*

(ii) *Let p be even and $\{\alpha_j\}$ has period $2p$. Then all gaps with $\Delta(z) = -2$ are closed if and only if $\alpha_{j+p} = \alpha_j$. All the gaps with $\Delta(z) = 2$ are closed if and only if $\alpha_{j+p} = -\alpha_j$.*

As shown in [43], Theorem 21.3(i) permits a non-trivial generalization. First, one can admit reflectionless Verblunsky coefficients from the outset and does not have to assume periodicity. Second, one can solve the inverse spectral problem (still in the more general reflectionless case) if the spectrum is a connected arc on the unit circle rather than the whole arc.

The isospectral results constitute one of the highlights of the OPUC theory.

THEOREM 21.4. *Let $\{\alpha_j\}_{j=0}^{p-1} \in \mathbf{D}^p$ such that $\Delta(z, \{\alpha_j\}_{j=0}^{p-1})$ has k open gaps. Then the set of all $\{\beta_j\}_{j=0}^{p-1} \in \mathbf{D}^p$ with $\Delta(z, \{\beta_j\}_{j=0}^{p-1}) = \Delta(z, \{\alpha_j\}_{j=0}^{p-1})$ is a k -dimensional torus.*

There is one important difference between OPUC and the Jacobi case. In the latter, the infinite gap does not count in the calculation of the dimension of the torus, so it has a dimension equal to the genus of the Riemann surface for the m -function. In the OPUC case, all gaps count and the torus has dimension one more than the genus.

The torus can be defined explicitly in terms of natural additional data associated to $\{\alpha_j\}_{j=0}^{p-1}$. One way to define the data is to analytically continue the Carathéodory function F for the periodic sequence. One cuts \mathbf{C} on the connected components of $\{|\Delta| \leq 2\}$ and forms the two-sheeted Riemann surface associated to $\sqrt{\Delta^2 - 4}$. On this surface F is meromorphic with exactly one pole on each “extended gap.” The ends of the gap are branch points and join the two images of the gap into a circle. The p points, one on each gap, are thus p -dimensional torus, and the refined version of the above result is that there is exactly one Carathéodory function associated to a period p set of Verblunsky coefficients with specified poles.

Alternately, the points in the gap are solutions of $\Phi_p = \Phi_p^*$ with sheets determined by whether the points are pure points of the associated measure or not.

22. Random Coefficients

Let Ω be the space $\prod_{n=0}^\infty \mathbf{D}$, σ_n measures on \mathbf{D} and $\sigma = \prod_{j=0}^\infty \sigma_j$ the product measure on Ω . In other words, sequences $\{\alpha_n\}_{n=0}^\infty$ of Verblunsky coefficients considered in this section are independent random variables. If in addition $\sigma_j = \sigma_0$ for all j , then they are identically distributed. One is interested in statements that hold when $\alpha_j = \omega_j$ for a.e. $\omega \in \Omega$ with respect to σ . The main result is that, typically, the associated measure μ_ω with $\alpha_j(\mu_\omega) = \omega_j$ is a pure point measure with pure points often dense in \mathbf{T} .

For $\lambda \in \mathbf{T}$, let $\mu_{\lambda,\omega}$ be the measure with $\alpha_j(\mu_{\lambda,\omega}) = \lambda\omega_j$. The main result for the i.i.d. case is [107, Thm. 12.6.3].

THEOREM 22.1. *Suppose that σ_0 is not supported at a single point. Then for a.e. $(\lambda, \omega) \in \mathbf{T} \times \Omega$ the measure $\mu_{\lambda,\omega}$ has pure point spectrum. If σ_0 is absolutely continuous with respect to the area Lebesgue measure on \mathbf{D} , then μ_ω is pure point for a.e. $\omega \in \Omega$.*

22.1. Decaying Random Verblunsky Coefficients. Decaying random potentials were studied starting with Simon [95] who found the first example of Jacobi matrices with $|a_n - 1| + |b_n| \rightarrow 0$ and μ purely singular. The pioneering results on decaying Verblunsky coefficients are due to Nikishin [83] and Teplyaev [123]. The detailed account of the subject is presented in [107, Sects. 12.6–7].

Now Verblunsky coefficients $\{\alpha_j(\omega)\}$ are assumed to be independent random but not necessarily identically distributed variables, which decay to zero in some sense; at a minimum, the mean value $\mathbf{E}(|\alpha_j(\omega)|^2) \rightarrow 0$. The main result states that there is no singular spectrum.

THEOREM 22.2. *Let the Verblunsky coefficients $\{\alpha_j(\omega)\}$ be independent random variables with $\mathbf{E}(\alpha_j(\omega)) = 0$ and*

$$\sum_{j=0}^\infty \mathbf{E}(|\alpha_j(\omega)|^2) < \infty,$$

i.e., the Szegő condition holds a.e. Then the corresponding orthogonality measure μ_ω is absolutely continuous for a.e. ω .

The situation changes dramatically for the slowly decaying Verblunsky coefficients. Assume that

$$\sup_{\omega,j} |\alpha_j(\omega)| < 1, \quad \lim_{j \rightarrow \infty} \sup_{\omega} |\alpha_j(\omega)| = 0, \tag{22.1}$$

$$\mathbf{E}(\alpha_j(\omega)) = \mathbf{E}(\alpha_j^2(\omega)) = 0, \tag{22.2}$$

and

$$\mathbf{E}(|\alpha_j(\omega)|^2)^{1/2} = \Gamma j^{-\gamma}, \quad j \geq j_0, \quad 0 < \gamma \leq \frac{1}{2}. \tag{22.3}$$

THEOREM 22.3. (i) *Suppose that independent Verblunsky coefficients $\{\alpha_j(\omega)\}$ satisfy (22.1)–(22.3) with $\gamma < \frac{1}{2}$. Then for a.e. pairs $(\lambda, \omega) \in \mathbf{T} \times \Omega$, the measure $\mu_{\lambda,\omega}$ is pure point and dense in \mathbf{T} .*

(ii) Suppose that (22.1)–(22.3) hold with $\gamma = \frac{1}{2}$ and in addition,

$$\sup_{\omega, j} n^{1/2} |\alpha_j(\omega)| < \infty.$$

Then for a.e. pairs $(\lambda, \omega) \in \mathbf{T} \times \Omega$, the measure $\mu_{\lambda, \omega}$ has dense pure point spectrum as long as $\Gamma > 1$. If $\Gamma \leq 1$, then for a.e. pairs $(\lambda, \omega) \in \mathbf{T} \times \Omega$, the measure $\mu_{\lambda, \omega}$ has purely singular continuous spectrum of exact Hausdorff dimension $1 - \Gamma^2$.

The distribution of the zeros of random paraorthogonal polynomials on the unit circle is studied in [118].

23. Miscellanea

23.1. Exponential Decay Estimates. In [81] Nevai and Totik proved that $\limsup_{n \rightarrow \infty} |\alpha_n(\mu)|^{1/n} = R^{-1} < 1$ if and only if μ is absolutely continuous and the reciprocal of the Szegő function (16.3) D^{-1} is analytic in the disk $\{|z| < R\}$, providing a formula for the exact rate of exponential decay in terms of properties of D^{-1} . Moreover, the following result is true (see [104, Thm. 2.1]).²

THEOREM 23.1. *Let $\limsup_{n \rightarrow \infty} |\alpha_n(\mu)|^{1/n} = R^{-1} < 1$ and define*

$$S(z) := - \sum_{j=0}^{\infty} \alpha_{j-1}(\mu) z^j, \quad r(z) := \overline{D(1/\bar{z})} D^{-1}(z), \tag{23.1}$$

so both S and D^{-1} are analytic in $\{|z| < R\}$. Then the difference $r - S$ is analytic in $\{z : 1 - \delta < |z| < R^3\}$ for some $\delta > 0$.

The point of this theorem is that both S and D^{-1} have singularities on the circle of radius R so the fact that the combination has analytic continuation is subtle.

To study the more specific exponential decay of Verblunsky coefficients, Simon came up with the following:

Definition. A sequence $\{A_n\}_{n \geq -1}$ of complex numbers is said to have an *asymptotic series with error R^{-n}* for some $R > 1$ if there exists a finite number of points $\{\mu_j\}_{j=1}^J$, $1 < |\mu_j| < R$, and polynomials $\{P_j\}_{j=1}^J$ so that

$$\limsup_{n \rightarrow \infty} \left| A_n - \sum_{j=1}^J \frac{P_j(n)}{\mu_j^{n+1}} \right|^{1/n} \leq R^{-1}. \tag{23.2}$$

A_n has a *complete asymptotic series* if it has an asymptotic series with error R^{-n} for all $R > 1$.

It is not hard to see that $\{A_n\}_{n \geq -1}$ has an asymptotic series with error R^{-n} if and only if

$$L(z) := \sum_{n=0}^{\infty} A_{n-1} z^n$$

is meromorphic in $\{|z| < R\}$ with a finite number of poles, all in $\{1 < |z| < R\}$. In particular, $\{A_n\}_{n \geq -1}$ has a complete asymptotic series if and only if L is an entire meromorphic function, i.e., meromorphic on the whole \mathbf{C} .

Here is the main result of [104].

²This result was proved independently in [22].

THEOREM 23.2. *Let μ be a non-trivial probability measure in \mathcal{P} with Verblunsky coefficients $\alpha_n(\mu)$. Then α_n has a complete asymptotic series if and only if μ is an absolutely continuous measure from the Szegő class and D^{-1} is an entire meromorphic function. The latter is equivalent to the function S (23.1) being an entire meromorphic function.*

Simon also provides the relation between μ_j 's that enter in the asymptotic series and the poles of D^{-1} . For some further extensions, see [71].

23.2. Rakhmanov's Theorem on an Arc. Let $\mu = wdm + \mu_s$ be a measure from \mathcal{P} . It was Rakhmanov [92] who proved that $w > 0$ a.e. on \mathbf{T} implies $\alpha_n(\mu) \rightarrow 0$ as $n \rightarrow \infty$. Subsequently, Bello and López [12] took this result over to an arc of the unit circle.

THEOREM 23.3 (Bello–López). *Let Δ_α be an arc on \mathbf{T}*

$$\Delta_\alpha := \{\zeta = e^{it} : \alpha \leq t \leq 2\pi - \alpha\}, \quad 0 < \alpha < \pi.$$

Suppose that $\text{supp } \mu = \Delta_\alpha$ and $w > 0$ a.e. on this arc. Then

$$\lim_{n \rightarrow \infty} |\alpha_n(\mu)| = \sin \frac{\alpha}{2}, \quad \lim_{n \rightarrow \infty} \alpha_n(\mu) \bar{\alpha}_{n-1}(\mu) = \sin^2 \frac{\alpha}{2}.$$

On the other hand, the recent result of Denisov [24] states

THEOREM 23.4 (Denisov). *Let $J = J(\{a_n\}, \{b_n\})$ be a Jacobi matrix with the spectral measure $\sigma = vdm + \sigma_s$. Suppose that $\text{ess supp } \sigma = [-2, 2]$ and $v > 0$ a.e. on $[-2, 2]$. Then $a_n \rightarrow 1$, $b_n \rightarrow 0$ as $n \rightarrow \infty$.*

The original Rakhmanov theorem (more precisely, its real line analogue) corresponds to the case $\text{supp } \sigma = [-2, 2]$. There is also an alternate proof of Denisov's theorem due to Nevai–Totik [82], which appeared to be a starting point for Simon to prove

THEOREM 23.5. *The result of Bello–López holds under the relaxed assumption $\text{ess supp } \mu = \Delta_\alpha$.*

23.3. Measures with VC's from ℓ^p , $p > 2$. For the class \mathcal{P} of non-trivial probability measures on \mathbf{T} , Totik [126] established the following

THEOREM 23.6. *Let $\mu \in \mathcal{P}$ with $\text{supp } \mu = \mathbf{T}$. Then there exists a measure ν mutually absolutely continuous with μ and such that $\alpha_n(\nu) \rightarrow 0$ as $n \rightarrow \infty$.*

Simon proved the stronger (see [106, Thm. 2.10.1])

THEOREM 23.7. *Let $\mu \in \mathcal{P}$ with $\text{supp } \mu = \mathbf{T}$. Then there exists a measure ν mutually absolutely continuous with μ and such that for all $p > 2$,*

$$\sum_{n=0}^{\infty} |\alpha_n(\nu)|^p < \infty.$$

Khrushchev [54] gave examples of singular continuous measure μ on \mathbf{T} with $\{\alpha_n(\mu)\} \in \ell^p$ for all $p > 2$.

23.4. Counting Eigenvalues in Gaps. Suppose that $\mu^{(0)}$ and μ are measures from \mathcal{P} with Verblunsky coefficients $\{\alpha_n^{(0)}\}$ and $\{\alpha_n\}$, respectively, and let $|\alpha_n - \alpha_n^{(0)}| \rightarrow 0$ with some information on the rate. If an open circular arc Γ is disjoint from $\sigma_{\text{ess}}(\mu^{(0)})$ then the same happens for $\sigma_{\text{ess}}(\mu)$ by Theorem 20.2. If the number of pure points of μ in Γ is infinite, they can only have the endpoints of Γ as limit points, and one can ask about the growth of the number of pure points near the endpoints of Γ .

The analogous problem for Schrödinger operators and Jacobi matrices is heavily studied. Given a Jacobi matrix $J(\{a_n\}_{n \geq 1}, \{b_n\}_{n \geq 1})$ with $a_n \rightarrow 1$ and $b_n \rightarrow 0$ as $n \rightarrow \infty$, denote by $N(J)$ a number of eigenvalues of J off $[-2, 2]$.

THEOREM 23.8 (Geronimo–Case, Chihara–Nevai). . *Suppose that*

$$\sum_{n=1}^{\infty} n(|b_n| + |a_n - 1|) < \infty.$$

Then $N(J) < \infty$.

Geronimo and later Hundertmark and Simon found a quantitative bound (see [50, Thm. A1])

$$N(J) \leq \sum_{n=1}^{\infty} (n|b_n| + (4n + 2)(a_n - 1)_+), \quad (x)_+ = \max(x, 0), \quad (23.3)$$

the result known as the *Bargman-type bound*. Next, denote by $\{E_n^\pm\}$ the eigenvalues of J with

$$E_1^+ > E_2^+ > \dots > 2 > -2 > \dots > E_2^- > E_1^-.$$

Then (see [50, Thm. 2])

$$\sum_{n=1}^{\infty} (|E_n^+ - 2|^{p-1/2}| + |E_n^- + 2|^{p-1/2}|) \leq C_p \left(\sum_{n=1}^{\infty} |b_n|^p + 4 \sum_{n=1}^{\infty} |a_n - 1|^p \right) \quad (23.4)$$

for all $p \geq 1$, where C_p is an explicitly given constant depending only on p . The estimate is usually called a *Lieb–Thirring-type bound*.

In [107, Sect. 12.2] Simon found similar bounds in the OPUC setting.

THEOREM 23.9. *Let $\eta_{j+N} = \eta_j$, $j \in \mathbf{Z}_+$, be a periodic sequence of complex numbers, and suppose the Verblunsky coefficients $\alpha_n(\mu)$ of a measure $\mu \in \mathcal{P}$ satisfy*

$$\sum_{j=0}^{\infty} j|\alpha_j(\mu) - \eta_j| < \infty.$$

Then μ has an essential support whose complement has at most N gaps, and each gap has only finitely many mass points.

THEOREM 23.10. *Suppose α 's and η 's are as in the above theorem, and*

$$\sum_{j=0}^{\infty} |\alpha_j(\mu) - \eta_j|^p < \infty$$

holds for some $p \geq 1$. Then for the mass points z_j in the gaps, we have

$$\sum_{z_j} \text{dist}(z_j, \text{ess supp } \mu)^q < \infty,$$

where $q > \frac{1}{2}$ if $p = 1$ and $q \geq p - \frac{1}{2}$ if $p > 1$.

Remark. In a recent paper [45] L. Golinskii found precise quantitative analogues of (23.3) and (23.4) in the OPUC setting for the case of a constant background ($N = 1$).

23.5. Jitomirskaya–Last Inequalities. In a fundamental paper intended to understand and extend the subordinacy theory of Gilbert–Pearson, Jitomirskaya and Last proved some basic inequalities about singularities of the m -function as the energy approaches the spectrum. In [107, Sect. 10.8] an analogue of their results for OPUC is established.

For a sequence of complex numbers $\{a_j\}_{j \geq 0}$ and a positive $0 < x < \infty$, define

$$\|a\|_x^2 := \sum_{j=0}^{[x]} |a_j|^2 + (x - [x])|a_{[x]+1}|^2,$$

so $\|a\|_n^2 = \sum_{j=0}^n |a_j|^2$ and $\|a\|_x^2$ is linearly interpolated in between. Let $\varphi_n(\zeta, \mu)$ and $\psi_n(\zeta, \mu)$ be the orthonormal polynomials of the first and second kind, respectively, with respect to μ . Since either $\{\varphi_n\}$ or $\{\psi_n\}$ is not in ℓ^2 for $\zeta \in \mathbf{T}$, the product $\|\varphi\|_x \|\psi\|_x$ increases monotonically from 1 to $+\infty$, $0 < x < \infty$. Hence for $\zeta \in \mathbf{T}$ and $0 \leq r < 1$, there is a unique solution $x(r)$ of

$$\|\varphi(\zeta)\|_{x(r)} \|\psi(\zeta)\|_{x(r)} = \frac{\sqrt{2}}{1 - r}.$$

The unit circle analogue of Jitomirskaya–Last inequalities takes the following form.

THEOREM 23.11. *Let $F = F(z, \mu)$ be the Carathéodory function of μ . Then*

$$A^{-1} \left[\frac{\|\psi(\zeta)\|_{x(r)}}{\|\varphi(\zeta)\|_{x(r)}} \right] \leq |F(r\zeta, \mu)| \leq A \left[\frac{\|\psi(\zeta)\|_{x(r)}}{\|\varphi(\zeta)\|_{x(r)}} \right],$$

where $1 < A < \infty$ is a universal constant.

COROLLARY 23.12. *Let $G \subset \mathbf{T}$ be a Borel set on the unit circle so that for all $\zeta \in G$, the transfer matrix T_n defined in (14.11) satisfies*

$$\sup_n \|T_n(\zeta)\| < \infty.$$

Then μ is absolutely continuous on G .

The OPUC analogues of the Gilbert–Pearson subordinacy theory appeared first in [46] and the Simon–Wolff theory in [124] (see also [46]) and were elaborated in [107, Sect. 10].

References

- [1] M. P. Alfaro, M. Bello, J. M. Montaner and J. L. Varona, Some asymptotic properties for orthogonal polynomials with respect to varying measures, *J. Approx. Theory* **135** (2005), 22–34.
- [2] M. P. Alfaro and L. Vigil, Solution of a problem of P. Turán, *J. Approx. Theory* **53** (1988), 195–197.
- [3] W. A. Al-Salam, Characterization theorems for orthogonal polynomials, *Orthogonal Polynomials: Theory and Practice*, (M. Ismail and P. Nevai, eds.), NATO ASI Series C, **294**, Kluwer, Dordrecht, 1990.
- [4] G. E. Andrews, R. Askey and R. Roy, *Special Functions*, Encyclopedia of Mathematics and Its Applications, Cambridge University Press, New York, 1999.
- [5] A. I. Aptekarev, Multiple orthogonal polynomials, *J. Comput. Appl. Math.* **99** (1988), 423–447.

- [6] A. I. Aptekarev, Asymptotic properties of polynomials orthogonal on a system of contours, and periodic motions of Toda chains, (Russian) *Mat. Sb. (N.S.)* **125(167)** (1984), 231–258.
- [7] A. I. Aptekarev and E. M. Nikishin, The scattering problem for a discrete Sturm–Liouville operator, (Russian) *Mat. Sb. (N.S.)* **121** (1983), 327–358.
- [8] A. I. Aptekarev and W. Van Assche, Scalar and matrix Riemann–Hilbert approach to the strong asymptotics of Padé approximants and complex orthogonal polynomials with varying weights, *J. Approx. Theory* **129** (2004), 129–166.
- [9] L. Baratchart, R. Küstner and V. Totik, Zero distributions via orthogonality, *Ann. Inst. Fourier (Grenoble)* **55** (2005), 1455–1499.
- [10] D. Barrios and G. López, Ratio asymptotics for polynomials orthogonal on arcs of the unit circle, *Constr. Approx.* **15** (1999), 1–31.
- [11] G. Baxter, A convergence equivalence related to polynomials orthogonal on the unit circle, *Trans. Amer. Math. Soc.* **99** (1961), 471–487.
- [12] M. Bello and G. López, Ratio and relative asymptotics of polynomials orthogonal on an arc of the unit circle, *J. Approx. Theory* **92** (1998), 216–244.
- [13] O. Blumenthal, *Über die Entwicklung einer willkürlichen Funktion nach den Nennern des Kettenbruches für $\int_{-\infty}^0 [\phi(\xi)/(z - \xi)] d\xi$* , Inaugural Dissertation, Göttingen, 1898.
- [14] S. Bochner, Über Sturm–Liouvillesche Polynomsysteme, *Math. Z.* **29** (1929), 730–736.
- [15] G. Borg, Eine Umkehrung der Sturm–Liouvilleschen Eigenwertaufgabe. Bestimmung der Differentialgleichung durch die Eigenwerte, *Acta Math.* **78** (1946), 1–96.
- [16] M. J. Cantero, L. Moral and L. Velázquez, Five-diagonal matrices and zeros of orthogonal polynomials on the unit circle, *Linear Algebra Appl.* **362** (2003), 29–56.
- [17] M. J. Cantero, L. Moral and L. Velázquez, Measures on the unit circle and unitary truncations of unitary operators, *J. Approx. Theory* **139** (2006), 430–468.
- [18] E. Daems and A. B. J. Kuijlaars, A Christoffel–Darboux formula for multiple orthogonal polynomials, *J. Approx. Theory* **130** (2004), 190–202.
- [19] D. Damanik, Verblunsky coefficients with Coulomb-type decay, *J. Approx. Theory* **139** (2006), 257–268.
- [20] P. Deift, *Orthogonal Polynomials and Random Matrices: A Riemann–Hilbert Approach*, Courant Lecture Notes in Mathematics, **3**, New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 1999.
- [21] P. Deift, Riemann–Hilbert methods in the theory of orthogonal polynomials, in this Festschrift.
- [22] P. Deift and J. Östensson, A Riemann–Hilbert approach to some theorems on Toeplitz operators and orthogonal polynomials, *J. Approx. Theory* **139** (2006), 144–171.
- [23] F. Delyon, B. Simon and B. Souillard, From power pure point to continuous spectrum in disordered systems, *Ann. Inst. H. Poincaré* **42** (1985), 283–309.
- [24] S. Denisov, On Rakhmanov’s theorem for Jacobi matrices, *Proc. Amer. Math. Soc.* **132** (2004), 847–852.
- [25] S. Denisov and S. Kupin, Asymptotics of the orthogonal polynomials for the Szegő class with a polynomial weight, *J. Approx. Theory* **139** (2006), 8–28.
- [26] S. Denisov and B. Simon, Zeros of orthogonal polynomials on the real line, *J. Approx. Theory* **121** (2003), 357–364.
- [27] J. Dombrowski, Quasitriangular matrices, *Proc. Amer. Math. Soc.* **69** (1978), 95–96.
- [28] C. F. Dunkl and Y. Xu, *Orthogonal Polynomials of Several Variables*, Encyclopedia of Mathematics and Its Applications, Cambridge University Press, New York, 2001.
- [29] A. J. Durán and F. A. Grünbaum, A survey on orthogonal matrix polynomials satisfying second order differential equations, *J. Comput. Appl. Math.* **178** (2005), 169–190.
- [30] H. Flaschka and D. W. McLaughlin, Canonically conjugate variables for the Korteweg–de Vries equation and the Toda lattice with periodic boundary conditions, *Progr. Theoret. Phys.* **55** (1976), 438–456.
- [31] A. S. Fokas, A. R. Its and A. V. Kitaev, The isomonodromy approach to matrix models in 2D quantum gravity, *Comm. Math. Phys.* **147** (1992), 395–430.
- [32] W. Gautschi, *Orthogonal Polynomials: Computation and Approximation*, Numerical Mathematics and Scientific Computation, Oxford University Press, New York, 2004.
- [33] W. Gautschi and A. B. J. Kuijlaars, Zeros and critical points of Sobolev orthogonal polynomials, *J. Approx. Theory* **91** (1997), 117–137.

- [34] J. S. Geronimo, F. Gesztesy and H. Holden, Algebro-geometric solutions of the Baxter–Szegő difference equation, *Comm. Math. Phys.* **258** (2005), 149–177.
- [35] J. S. Geronimo and R. Johnson, An inverse problem associated with polynomials orthogonal on the unit circle, *Comm. Math. Phys.* **193** (1998), 125–150.
- [36] J. S. Geronimo, D. S. Lubinsky and F. Marcellán, Asymptotics for Sobolev orthogonal polynomials for exponential weights, *Constr. Approx.* **22** (2005), 309–346.
- [37] J. S. Geronimo and A. Martínez-Finkelshtein, On extensions of a theorem of Baxter, *J. Approx. Theory* **139** (2006), 214–222.
- [38] J. S. Geronimo and W. Van Assche, Orthogonal polynomials on several intervals via a polynomial mapping, *Trans. Amer. Math. Soc.* **308** (1988), 559–581.
- [39] Ya. L. Geronimus, On the character of the solution of the moment problem in the case of the periodic in the limit associated fraction, (Russian) *Bull. Acad. Sci. USSR Math.* **5** (1941), 203–210.
- [40] Ya. L. Geronimus, On polynomials orthogonal on the circle, on trigonometric moment problem, and on allied Carathéodory and Schur functions, (Russian) *Mat. Sb.* **15** (1944), 99–130.
- [41] Ya. L. Geronimus, On the trigonometric moment problem, *Annals of Math.* **47** (1946), 742–761.
- [42] F. Gesztesy and M. Zinchenko, Weyl–Titchmarsh theory for CMV operators associated with orthogonal polynomials on the unit circle, *J. Approx. Theory* **139** (2006), 172–213.
- [43] F. Gesztesy and M. Zinchenko, A Borg-type theorem associated with orthogonal polynomials on the unit circle, to appear in *J. London Math. Soc.*
- [44] B. Golinskii, On a connection of the rate of decay of parameters of orthogonal polynomials with properties of the corresponding function of distribution, (Russian) *Izv. Akad. Nauk Armyan. SSR* **15** (1980), 127–144.
- [45] L. Golinskii, Lieb–Thirring and Bargmann-type inequalities for circular arc, *J. Approx. Theory* **139** (2006), 65–74.
- [46] L. Golinskii and P. Nevai, Szegő difference equations, transfer matrices and orthogonal polynomials on the unit circle, *Comm. Math. Phys.* **223** (2001), 223–259.
- [47] L. Golinskii and A. Zlatoš, Coefficients of orthogonal polynomials on the unit circle and higher-order Szegő theorems, preprint arXiv:math.CA/0509192, 2005.
- [48] A. A. Gonchar and E. A. Rakhmanov, On the convergence of simultaneous Padé approximants for systems of functions of Markov type, Number Theory, Mathematical Analysis and Their Applications. (Russian) *Trudy Mat. Inst. Steklov.* **157** (1981), 31–48.
- [49] A. A. Gonchar and S. P. Suetin, On Padé approximation of Markov type meromorphic functions, Current Problems in Mathematics, **5**, (Russian) Rossiiskaya Akademiya Nauk, Matematicheskii Institut im. V. A. Steklova, Moscow, 2004. 68 pp. (electronic).
- [50] D. Hundertmark and B. Simon, Lieb–Thirring inequalities for Jacobi matrices, *J. Approx. Theory* **118** (2002), 106–130.
- [51] M. E. H. Ismail, *Classical and Quantum Orthogonal Polynomials in One Variable*, Encyclopedia of Mathematics and Its Applications, **98**, Cambridge University Press, Cambridge, 2005.
- [52] V. Kaliaguine, A note on the asymptotics of orthogonal polynomials on a complex arc: the case of a measure with a discrete part, *J. Approx. Theory* **80** (1995), 138–145.
- [53] S. Khrushchev, Schur’s algorithm, orthogonal polynomials, and convergence of Wall’s fractions in $L^2(\mathbf{T})$, *J. Approx. Theory* **108** (2001), 161–248.
- [54] S. Khrushchev, A singular Riesz product in the Nevai class and inner functions with the Schur parameters in $\cap_{p>2} \ell^p$, *J. Approx. Theory* **108** (2001), 249–255.
- [55] S. Khrushchev, Classification theorems for general orthogonal polynomials on the unit circle, *J. Approx. Theory* **116** (2002), 268–342.
- [56] S. Khrushchev, Continued fractions and orthogonal polynomials on the unit circle, *J. Comput. Appl. Math.* **178** (2005), 267–303.
- [57] R. Killip and B. Simon, Sum rules for Jacobi matrices and their applications to spectral theory, *Annals of Math.* **158** (2003), 253–321.
- [58] A. Kiselev, Y. Last and B. Simon, Modified Prüfer and EFGP transforms and the spectral analysis of one-dimensional Schrödinger operators, *Comm. Math. Phys.* **194** (1998), 1–45.
- [59] T. Kriecherbauer and K. T.-R. McLaughlin, Strong asymptotics of polynomials orthogonal with respect to a Freud weight, *Int. Math. Res. Not.* **6** (1999), 299–333.

- [60] Y. Last and B. Simon, The essential spectrum of Schrödinger, Jacobi, and CMV operators, to appear in *J. Anal. Math.*
- [61] A. Levin and D. S. Lubinsky, Christoffel functions and orthogonal polynomials for exponential weights on $[-1, 1]$, *Mem. Amer. Math. Soc.* **111** (1994), no. 535, 146 pp.
- [62] A. Levin and D. S. Lubinsky, *Orthogonal Polynomials for Exponential Weights*, CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, 4, Springer-Verlag, New York, 2001.
- [63] G. López, H. Pijera-Cabrera and I. P. Izquierdo, Sobolev orthogonal polynomials in the complex plane, *J. Comput. Appl. Math.* **127** (2001), 219–230.
- [64] P. López-Rodríguez and A. J. Durán, Orthogonal matrix polynomials, Laredo Lectures on Orthogonal Polynomials and Special Functions, pp. 13–44, Adv. Theory Spec. Funct. Orthogonal Polynomials, Nova Sci. Publ., Hauppauge, NY, 2004.
- [65] D. S. Lubinsky, *Strong Asymptotics for Extremal Errors Associated with Erdős-type Weights*, Pitman Res. Notes, **202**, Longmans, Essex, 1989.
- [66] D. S. Lubinsky and E. B. Saff, *Strong Asymptotics for Extremal Polynomials Associated with Weights on $(-\infty, \infty)$* , Springer Lecture Notes, **1305**, Springer-Verlag, Berlin, 1988.
- [67] W. Magnus and S. Winkler, *Hill's Equation*, Interscience Tracts in Pure and Applied Mathematics, **20**, Interscience Publishers, New York, 1966.
- [68] F. Marcellán, M. Alfaro and M. L. Rezola, Orthogonal polynomials on Sobolev spaces: old and new directions, *J. Comput. Appl. Math.* **48** (1993), 113–131.
- [69] A. Martínez-Finkelshtein, Analytic aspects of Sobolev orthogonal polynomials, *J. Comput. Appl. Math.* **99** (1998), 491–510.
- [70] A. Martínez-Finkelshtein, Analytic properties of Sobolev orthogonal polynomials revisited, *J. Comput. Appl. Math.* **127** (2001), 255–266.
- [71] A. Martínez-Finkelshtein, Szegő polynomials: a view from the Riemann–Hilbert window, to appear in *Electronic Trans. Numer. Anal.*
- [72] G. Mastroianni and G. Milovanovic, *Approximation of Functions and Functionals with Application*, World Scientific, to appear.
- [73] G. Mastroianni and V. Totik, Polynomial inequalities with doubling and A_∞ weights, *Constr. Approx.* **16** (2000), 37–71.
- [74] A. Máté, P. Nevai and V. Totik, Extensions of Szegő's theory of orthogonal polynomials, *Constr. Approx.* **3** (1987), 51–72.
- [75] A. Máté, P. Nevai and V. Totik, Strong and weak convergence of orthogonal polynomials, *Amer. J. Math.* **109** (1987), 239–282.
- [76] A. Máté, P. Nevai and V. Totik, Szegő's extremum problem on the unit circle, *Annals of Math.* **134** (1991), 433–453.
- [77] M. Mehta, *Random Matrices*, second ed., Academic Press, Boston, 1991.
- [78] H. N. Mhaskar and E. B. Saff, On the distribution of zeros of polynomials orthogonal on the unit circle, *J. Approx. Theory* **63** (1990), 30–38.
- [79] S. N. Naboko, On the dense point spectrum of Schrödinger and Dirac operators, *Theoret. Math. Phys.* **68** (1986), 646–653.
- [80] P. Nevai, Orthogonal polynomials, *Mem. Amer. Math. Soc.* **18** (1979), no. 213, 185 pp.
- [81] P. Nevai and V. Totik, Orthogonal polynomials and their zeros, *Acta Sci. Math. (Szeged)* **53** (1989), 99–104.
- [82] P. Nevai and V. Totik, Denisov's theorem on recurrent coefficients, *J. Approx. Theory* **127** (2004), 240–245.
- [83] E. M. Nikishin, An estimate for orthogonal polynomials, (Russian) *Acta Sci. Math. (Szeged)* **48** (1985), 395–399.
- [84] J. Nuttall, Asymptotics for diagonal Hermite–Padé approximants, *J. Approx. Theory* **42** (1984), 299–386.
- [85] J. Nuttall, Padé polynomial asymptotics from a singular integral equation, *Constr. Approx.* **6** (1990), 157–166.
- [86] L. Pastur and A. Figotin, *Spectra of Random and Almost-Periodic Operators*, Grundlehren de mathematischen Wissenschaften, **297**, Springer-Verlag, Berlin, 1992.
- [87] F. Peherstorfer, On Bernstein–Szegő orthogonal polynomials on several intervals, *SIAM J. Math. Anal.* **21** (1990), 461–482.

- [88] F. Peherstorfer, On Bernstein–Szegő orthogonal polynomials on several intervals. II. Orthogonal polynomials with periodic recurrence coefficients, *J. Approx. Theory* **64** (1991), 123–161.
- [89] F. Peherstorfer, Orthogonal and extremal polynomials on several intervals, *J. Comput. Appl. Math.* **48** (1993), 187–205.
- [90] F. Peherstorfer, Zeros of polynomials orthogonal on several intervals, *Int. Math. Res. Not.* **7** (2003), 361–385.
- [91] F. Peherstorfer and R. Steinbauer, On polynomials orthogonal on several intervals, Special Functions (Torino, 1993), *Ann. Numer. Math.* **2** (1995), 353–370.
- [92] E. A. Rakhmanov, On the asymptotics of the ratio of orthogonal polynomials, *Math. USSR Sbornik* **46** (1983), 105–117; Russian original: *Mat. Sb.* **118** (1982), 104–117.
- [93] E. A. Rakhmanov, On V. A. Steklov’s problem in the theory of orthogonal polynomials, (Russian) *Dokl. Akad. Nauk SSSR* **254** (1980), 802–806.
- [94] I. Schur, Über Potenzreihen, die im Innern des Einheitskreises beschränkt sind, I, II, *J. Reine Angew. Math.* **147** (1917), 205–232; **148** (1918), 122–145.
- [95] B. Simon, Some Jacobi matrices with decaying potential and dense point spectrum, *Comm. Math. Phys.* **87** (1982), 253–258.
- [96] B. Simon, The classical moment problem as a self-adjoint finite difference operator, *Advances in Math.* **137** (1998), 82–203.
- [97] B. Simon, The Golinskii–Ibragimov method and a theorem of Damanik–Killip, *Int. Math. Res. Not.* **36** (2003), 1973–1986.
- [98] B. Simon, Analogs of the m -functions in the theory of orthogonal polynomials on the unit circle, *J. Comp. Appl. Math.* **171** (2004), 411–424.
- [99] B. Simon, A canonical factorization for meromorphic Herglotz functions on the disk and sum rules for Jacobi matrices, *J. Funct. Anal.* **214** (2004), 396–409.
- [100] B. Simon, Orthogonal polynomials on the unit circle: New results, *Int. Math. Res. Not.* **53** (2004), 2837–2880.
- [101] B. Simon, Ratio asymptotics and weak asymptotic measures for orthogonal polynomials on the real line, *J. Approx. Theory* **126** (2004), 198–217.
- [102] B. Simon, Fine structure of the zeros of orthogonal polynomials, II. OPUC with competing exponential decay, *J. Approx. Theory* **135** (2005), 125–139.
- [103] B. Simon, Fine structure of the zeros of orthogonal polynomials, III. Periodic recursion coefficients, *Comm. Pure Appl. Math.* **59** (2005) 1042–1062.
- [104] B. Simon, Meromorphic Szegő functions and asymptotic series for Verblunsky coefficients, *Acta Math.* **195** (2005), 267–285.
- [105] B. Simon, OPUC on one foot, *Bull. Amer. Math. Soc.* **42** (2005), 431–460.
- [106] B. Simon, *Orthogonal Polynomials on the Unit Circle, Part 1: Classical Theory*, AMS Colloquium Series, American Mathematical Society, Providence, RI, 2005.
- [107] B. Simon, *Orthogonal Polynomials on the Unit Circle, Part 2: Spectral Theory*, AMS Colloquium Series, American Mathematical Society, Providence, RI, 2005.
- [108] B. Simon, The sharp form of the strong Szegő theorem, *Contemp. Math.* **387** pp. 253–275, American Mathematical Society, Providence, RI, 2005.
- [109] B. Simon, Aizenman’s theorem for orthogonal polynomials on the unit circle, *Constr. Approx.* **23** (2006), 229–240.
- [110] B. Simon, Fine structure of the zeros of orthogonal polynomials, I. A tale of two pictures, Proc. Constructive Functions Tech-04, to appear in *Electronic Trans. Numer. Anal.*
- [111] B. Simon and V. Totik, Limits of zeros of orthogonal polynomials on the circle, *Math. Nachr.* **278** (2005), 1615–1620.
- [112] B. Simon and A. Zlatoš, Sum rules and the Szegő condition for orthogonal polynomials on the real line, *Comm. Math. Phys.* **242** (2003), 393–423.
- [113] B. Simon and A. Zlatoš, Higher-order Szegő theorems with two singular points, *J. Approx. Theory* **134** (2005), 114–129.
- [114] H. Stahl, Divergence of diagonal Padé approximants and the asymptotic behavior of orthogonal polynomials associated with nonpositive measures, *Constr. Approx.* **1** (1985), 249–270.
- [115] H. Stahl, Orthogonal polynomials with complex-valued weight function. I, II, *Constr. Approx.* **2** (1986), 225–240, 241–251.

- [116] H. Stahl, Orthogonal polynomials with respect to complex-valued measures, *Orthogonal Polynomials and Their Applications* (Erice, 1990), *IMACS Ann. Comput. Appl. Math.* **9** (1991), 139–154.
- [117] H. Stahl and V. Totik, *General Orthogonal Polynomials*, *Encyclopedia of Mathematics and Its Applications*, **43**, Cambridge University Press, New York, 1992.
- [118] M. Stoiciu, The statistical distribution of the zeros of random paraorthogonal polynomials on the unit circle, *J. Approx. Theory* **139** (2006), 29–64.
- [119] S. P. Suetin, On the uniform convergence of diagonal Padé approximants for hyperelliptic functions, *Sb. Math.* **191** (2000), 1339–1373; Russian original: *Mat. Sb.* **191** (2000), 81–114.
- [120] S. P. Suetin, Padé approximants and the effective analytic continuation of a power series, *Russian Math. Surveys* **57** (2002), 43–141; Russian original: *Uspekhi Mat. Nauk* **57** (2002), 45–142.
- [121] J. Szabados and P. Vértesi, *Interpolation of Functions*, World Scientific, Singapore, 1990.
- [122] G. Szegő, *Orthogonal Polynomials*, Coll. Publ., **XXIII**, American Mathematical Society, Providence, RI, 1975.
- [123] A. V. Teplyaev, Properties of polynomials that are orthogonal on the circle with random parameters, *J. Soviet Math.* **61** (1992), 1931–1935.
- [124] A. V. Teplyaev, The pure point spectrum of random orthogonal polynomials on the circle, *Soviet Math. Dokl.*, **44** (1992), 407–411.
- [125] M. Toda, *Theory of Nonlinear Lattices*, second edition, Springer Series in Solid-State Sciences, **20**, Springer-Verlag, Berlin, 1989.
- [126] V. Totik, Orthogonal polynomials with ratio asymptotics, *Proc. Amer. Math. Soc.* **114** (1992), 491–495.
- [127] V. Totik, Asymptotics for Christoffel functions for general measures on the real line, *J. d'Analyse Math.* **81** (2000), 283–303.
- [128] W. Van Assche, *Asymptotics for Orthogonal Polynomials*, Springer Lecture Notes, **1265**, Springer-Verlag, Berlin, 1987.
- [129] W. Van Assche, Padé and Hermite–Padé approximation and orthogonality, *Surveys in Approx. Theory* **2** (2006), 61–91.
- [130] W. Van Assche, J. S. Geronimo and A. B. J. Kuijlaars, Riemann–Hilbert problems for multiple orthogonal polynomials, *Special Functions 2000: Current Perspective and Future Directions* (Tempe, AZ), pp. 23–59, *NATO Sci. Ser. II Math. Phys. Chem.*, **30**, Kluwer, Dordrecht, 2001.
- [131] S. Verblunsky, On positive harmonic functions: A contribution to the algebra of Fourier series, *Proc. London Math. Soc.* **38** (1935), 125–157.
- [132] S. Verblunsky, On positive harmonic functions (second part), *Proc. London Math. Soc.*, **40** (1936), 290–320.
- [133] H. Widom, Extremal polynomials associated with a system of curves in the complex plane, *Adv. Math.* **3** (1969), 127–232.

(Leonid Golinskii) VERKIN INSTITUTE FOR LOW TEMPERATURE PHYSICS, 47 LENIN AVE., 61103 KHARKOV, UKRAINE

E-mail address: golinskii@ilt.kharkov.ua

(Vilmos Totik) BOLYAI INSTITUTE, UNIVERSITY OF SZEGED, SZEGED, ARADI V. TERE 1, 6720, HUNGARY; AND DEPARTMENT OF MATHEMATICS, UNIVERSITY OF SOUTH FLORIDA, TAMPA, FL, 33620, U.S.A.

E-mail address: totik@shell.cas.usf.edu

Orthogonal Polynomials: The First Minutes

Sergey Khrushchev

Dedicated to Barry Simon on his 60th birthday

ABSTRACT. This survey clarifies the role played by Wallis' infinite product and Brouncker's continued fraction in the discovery of general orthogonal polynomials by Chebyshev. Brouncker's continued fraction attracted the attention of Euler and seemingly was a motive for the discoveries of the Gamma and Beta functions. We give evidence indicating how close Euler was to the discovery of orthogonal polynomials and show the relationship of works by Chebyshev, Gauss, and Jacobi to these papers of Euler. It is interesting that Brouncker's polynomials were explicitly written in Wallis' *Arithmetica Infinitorum* in 1655. But it was only in 1977 that J. Wilson introduced a family of polynomials, now bearing his name, of which Brouncker's polynomials are a partial case.

Using Euler's methods we also give an analogue of Stirling's formula for Brouncker's continued fraction (where E_{2k} are Euler's numbers):

$$b(s) = s + \prod_{n=1}^{\infty} \left(\frac{(2n-1)^2}{2s} \right) \sim s \cdot \exp \left\{ - \sum_{k=1}^{\infty} \frac{E_{2k}}{2ks^{2k}} \right\}$$

CONTENTS

1. Introduction
 2. Arithmetica Infinitorum
 3. Recovery of Brouncker's Proof
 4. Euler
 5. The Asymptotic Expansion for $b(s)$
 6. Euler's Problem
 7. Gauss' Continued Fractions for ${}_2F_1$
 8. Gauss' Quadrature Formula
 9. Jacobi's Contribution
 10. Sturm's Theorem
 11. Chebyshev's Solution to Euler's Problem
 12. Back to Brouncker's Polynomials
- References

2000 *Mathematics Subject Classification.* 30B70 (primary), 01A45, 33B15, 33C47, 33D45.
Key words and phrases. continued fraction algorithm, orthogonal polynomials.

1. Introduction

It is customary to consider Chebyshev, Gauss, Jacobi, and Legendre as the main creators of the theory of orthogonal polynomials. However, their contributions were directly influenced by Brouncker and Wallis who, in March of 1655, made discoveries which influenced the development of analysis for the next hundred years. Namely, Wallis found an infinite product of rational numbers converging to $4/\pi$ and Brouncker gave a remarkable continued fraction for this quantity. Brouncker didn't leave much on how he proved his formula. However, some vague remarks remained in Section 191 of Wallis' *Arithmetica Infinitorum*. The only mathematician who understood the importance of these discoveries was Euler. He undertook serious efforts to recover Brouncker's proof and found several proofs, but they all used differential and integral calculus, tools which were not available to Brouncker. It is clear from Euler's remarks in [3] that he felt that the recovery of the original Brouncker's proof could open new perspectives for analysis. As usual, Euler was right. The purpose of this paper is to show that indeed Brouncker's proof leads to the discovery of orthogonal polynomials and, in fact, the polynomials listed in Wallis' *Arithmetica Infinitorum* in 1655 are the first orthogonal polynomials which were written explicitly.

The plan of the paper is the following. In Sections 2–3, we discuss Brouncker's formula and recover Brouncker's proof from Wallis' remarks. In Section 4 we present an application of Euler's ideology of Wallis' interpolation to the theory of the Gamma function and show how this topic is related to Brouncker's continued fraction. In Section 5 using Euler's differential method for obtaining continued fractions, we derive the asymptotic expansion for Brouncker's continued fraction in terms of Euler's numbers. In Section 6 we state Euler's problem on rational parameterization of orthogonal matrices which finally led Chebyshev to the discovery of general orthogonal polynomials and their most important properties. In Sections 7–8 we show how Euler's ideas in the field of continued fractions influenced the discoveries of Gauss. In Sections 9–10 we present Jacobi's and Sturm's contributions. In Section 11 we give an account of Chebyshev's method, and finally in Section 12 we summarize Chebyshev's contribution to the example of Brouncker's polynomials.

2. Arithmetica Infinitorum

By March of 1655 John Wallis had almost completed his *Arithmetica Infinitorum* in which he promoted an important method of interpolation. This was a great work. Very briefly, Wallis discovered that analytic formulas can be interpolated by their values at integer numbers. Nowadays these facts are established in complex analysis with uniqueness theorems. But in 1655, even differential calculus was not available. Nevertheless, Wallis successfully applied his interpolation to find formulas for the areas under many curves. Only one curve remained uncovered. It was the unit circle. In 1593 Viète found the formula

$$\frac{2}{\pi} = \frac{\sqrt{2}}{2} \cdot \frac{\sqrt{2+\sqrt{2}}}{2} \cdot \frac{\sqrt{2+\sqrt{2+\sqrt{2}}}}{2} \cdots \quad (1)$$

Since the multipliers in Viète's formula are algebraic irrationalities of increasing order, it was not the formula which could meet Wallis' requirements. Finally, in

March of 1655, Wallis obtained his now well-known formula

$$\frac{2}{\pi} = \frac{1 \cdot 3}{2 \cdot 2} \cdot \frac{3 \cdot 5}{4 \cdot 4} \cdot \frac{5 \cdot 7}{6 \cdot 6} \cdot \dots \cdot \frac{(2n-1) \cdot (2n+1)}{2n \cdot 2n} \cdot \dots \tag{2}$$

in which all multipliers are rational. Wallis shared this important discovery with William Brouncker [15], who gave the quadrature continued fraction for π :

$$\frac{4}{\pi} = 1 + \frac{1^2}{2} + \frac{3^2}{2} + \frac{5^2}{2} + \dots = 1 + \mathbf{K}_{n=1}^{\infty} \left(\frac{(2n-1)^2}{2} \right) . \tag{3}$$

Brouncker’s proof has never been published. However, Wallis included his comments on Brouncker’s proof at the end of *Arithmetica Infinitorum*. From these comments one can conclude that Brouncker considered the function

$$b(s) = s + \mathbf{K}_{n=1}^{\infty} \left(\frac{(2n-1)^2}{2s} \right) \tag{4}$$

and proved that it satisfies the following functional equation

$$b(s)b(s+2) = (s+1)^2 . \tag{5}$$

This functional equation looks especially attractive written in the form

$$\left(s - 1 + \frac{1^2}{2(s-1)} + \frac{3^2}{2(s-1)} + \dots \right) \times \left(s + 1 + \frac{1^2}{2(s+1)} + \frac{3^2}{2(s+1)} + \dots \right) = s^2 .$$

We give here the most important extract from the recent translation by Stedall of *Arithmetica Infinitorum* [22]:

The Noble Gentleman noticed that two consecutive odd numbers, if multiplied together, form a product which is the square of the intermediate even number minus one... He asked, therefore, by what ratio the factors must be increased to form a product, not those squares minus one, but equal to the squares themselves.

Algebraically these can be written as a transformation of

$$s(s+2) = (s+1)^2 - 1$$

into (5).

The important discovery of Brouncker remained untouched until May 1727, when Euler arrived in St. Petersburg with *Arithmetica Infinitorum*. Brouncker’s formula mentioned in it didn’t escape Euler’s attention. In fact, Euler undertook tremendous efforts to recover Brouncker’s proof. For instance in his main paper on continued fraction [3], Euler wrote:

Therefore for quite a long time I undertook great efforts to prove this Theorem so that its proof a priori can be related with this function; this research, in my opinion, is more difficult, but I believe it may result in great benefits. While any such research was condemned to failure, I regret most of all the fact that Brouncker’s method was nowhere present and most likely sank into oblivion.

3. Recovery of Brouncker’s Proof

Let us notice first [15] that Wallis showed Brouncker his infinite product. As it is clear from the following formula

$$\frac{1 \cdot 3}{2 \cdot 2} \cdot \frac{3 \cdot 5}{4 \cdot 4} \cdot \frac{5 \cdot 7}{6 \cdot 6} \cdot \dots \cdot \frac{(2n - 1) \cdot (2n + 1)}{2n \cdot 2n} = \frac{1 \cdot 3}{0} + \frac{2 \cdot 2}{0} + \frac{3 \cdot 5}{0} + \dots + \frac{(2n - 1) \cdot (2n + 1)}{0} + \frac{(2n) \cdot (2n)}{1}, \quad (6)$$

products and continued fractions are related. Since any formal infinite continued fraction with identically zero partial denominators diverges, something should be done to make them positive. This was known to Brouncker; see Wallis [22, p. 169, footnote 79]: ‘The first fraction, beginning with zero, oscillates between zero and infinity, but multiplied by the next fraction, beginning with 2, it is supposed to make 1.’

This suggests the idea of increasing s to $b(s)$ and $s + 2$ to $b(s + 2)$ in odd partial numerators of continued fraction (6), where b satisfies (5). Then to keep (6) valid, odd zero partial denominators in the right-hand side of (6) must become positive automatically. That is exactly what is needed to complete the proof. The fact that $s + 1$ is even is also helpful since it may provide necessary cancellations. Using (5) repeatedly, we may write

$$\begin{aligned} b(1) &= \frac{2^2}{b(3)} = \frac{2^2}{4^2} b(5) = \frac{2^2 \cdot 6^2}{4^2 b(7)} = \frac{2^2 \cdot 6^2}{4^2 \cdot 8^2} b(9) = \dots \\ &= \frac{2^2 \cdot 6^2 \cdot 10^2}{4^2 \cdot 8^2 \cdot 12^2} \dots \frac{(4n - 2)^2}{(4n)^2} b(4n + 1) = \frac{1^2 \cdot 3^2}{2^2 \cdot 4^2} \dots \frac{(2n - 1)^2}{(2n)^2} b(4n + 1) = \\ &= \frac{1 \cdot 3}{2^2} \cdot \frac{3 \cdot 5}{4^2} \cdot \frac{5 \cdot 7}{6^2} \cdot \dots \cdot \frac{(2n - 1)(2n + 1)}{(2n)^2} \cdot \frac{b(4n + 1)}{(2n + 1)}. \end{aligned} \quad (7)$$

Combined with Wallis’ formula, this implies

$$b(1) = \left(\frac{2}{\pi} + o(1) \right) \cdot \frac{b(4n + 1)}{(2n + 1)}. \quad (8)$$

Since $s + 2 < b(s + 2)$ and $b(s)b(s + 2) = (s + 1)^2$, we have

$$s < b(s) < \frac{s^2 + 2s + 1}{s + 2} = s + \frac{1}{2 + s}, \quad (9)$$

which together with (8) imply

$$b(1) = \lim_n \left(\frac{2}{\pi} + o(1) \right) \cdot \frac{b(4n + 1)}{(2n + 1)} = \frac{4}{\pi}. \quad (10)$$

It remains only to find a formula for $b(s)$. Inequality (9) suggests that $b(s)$ may be represented as a sum of an infinite power series in $1/s$. The analogy with the conversion of decimal numbers into regular continued fractions indicates that this method may be at least tried. It is not a big deal to obtain by iterations of (5) and (9), using the method of uncertain coefficients, that

$$b(s) = s + \frac{1}{2s} - \frac{9}{8s^3} + \frac{153}{16s^5} + O\left(\frac{1}{s^7}\right), \quad (11)$$

or equivalently

$$b(s) = s + \frac{8s^4 - 18s^2 + 153}{16s^5} + O\left(\frac{1}{s^7}\right). \tag{12}$$

Applying the Euclidean algorithm to the quotient of polynomials, we have

$$\begin{aligned} \frac{8s^4 - 18s^2 + 153}{16s^5} &= \frac{1}{2s + \frac{9(4s^3 - 34s)}{8s^4 - 18s^2 + 153}} = \frac{1}{2s + \frac{9}{\frac{8s^4 - 18s^2 + 153}{4s^3 - 34s}}} \\ &= \frac{1}{2s + \frac{9}{2s + \frac{25(2s^2 + 153/25)}{4s^3 - 34s}}} = \frac{1}{2s + \frac{9}{2s + \frac{25}{2s + \dots}}}. \end{aligned}$$

A remarkable property of the above calculations is that $1^2 = 1$, $3^2 = 9$, $5^2 = 25$, etc. appear automatically as common divisors of the coefficients of the polynomials in Euclid’s algorithm. The fraction $153/25$ occurs only because c_7 was not found. Increasing the number of terms in (11), we naturally arrive at the conclusion that (4) is a correct formula for b .

Having obtained (4), we may reverse the order of arguments and compute the following differences for first convergents P_n/Q_n of continued fraction (4)

$$\frac{P_0(s)}{Q_0(s)} \frac{P_0(s+2)}{Q_0(s+2)} - (s+1)^2 = s(s+2) - (s+1)^2 = \boxed{(-1) = O(1)};$$

$$\begin{aligned} \frac{P_1(s)}{Q_1(s)} \frac{P_1(s+2)}{Q_1(s+2)} - (s+1)^2 &= \frac{4s^4 + 16s^3 + 20s^2 + 8s + 9}{4s^2 + 8s} - \\ &- \frac{4s^4 + 16s^3 + 20s^2 + 8s}{4s^2 + 8s} = \boxed{\frac{9}{4s^2 + 8s} = O\left(\frac{1}{s^2}\right)}; \end{aligned}$$

$$\begin{aligned} \frac{P_2(s)}{Q_2(s)} \frac{P_2(s+2)}{Q_2(s+2)} - (s+1)^2 &= \\ \frac{16s^6 + 96s^5 + 280s^4 + 480s^3 + 649s^2 + 594s}{16s^4 + 64s^3 + 136s^2 + 144s + 225} - (s+1)^2 &= \\ = \boxed{\frac{-225}{16s^4 + 64s^3 + 136s^2 + 144s + 225} = O\left(\frac{1}{s^4}\right)}. \end{aligned}$$

One can find these very formulas in [22, pp. 169–170], where Wallis writes: ‘... which is less than the square $F^2 + 2F + 1$.¹ And thus it may be done as far as one likes; it will form a product which will be (in turn) now greater than, now less than, the given square (the difference, however, continually decreasing, as is clear), which was to be proved.’

To make Wallis’ comments more clear, we first show that the convergents $P_n(s)/Q_n(s)$ tend to some limit if $s > 0$ and $n \rightarrow +\infty$. Polynomials P_n and Q_n satisfy the predecessor of Szegö’s recursion:

$$\begin{aligned} P_n &= 2sP_{n-1} + (2n - 1)^2P_{n-2}, \\ Q_n &= 2sQ_{n-1} + (2n - 1)^2Q_{n-2}. \end{aligned} \tag{13}$$

¹In Wallis’ notation, $s = F$.

Formulas of type (13) valid for arbitrary continued fractions are now known as Euler–Wallis formulas. However, as it is clear from Wallis’ notes, it was Brouncker who showed these formulas to Wallis.

LEMMA 1. *Continued fraction (4) converges for $s \geq 1$:*

$$\lim_n \frac{P_n(s)}{Q_n(s)} = b(s) .$$

PROOF. An iterative application of (13) shows that

$$P_n Q_{n-1} - P_{n-1} Q_n = (-1)^{n-1} (2n - 1)!!^2 ,$$

which implies

$$\frac{Q_{n-1}}{P_{n-1}} - \frac{Q_n}{P_n} = \frac{(-1)^{n-1} (2n - 1)!!^2}{P_n P_{n-1}} .$$

Next

$$\begin{aligned} \frac{Q_{n-2}}{P_{n-2}} - \frac{Q_n}{P_n} &= \frac{Q_{n-2}}{P_{n-2}} - \frac{Q_{n-1}}{P_{n-1}} + \frac{Q_{n-1}}{P_{n-1}} - \frac{Q_n}{P_n} = \\ &= \frac{(-1)^{n-1} (2n - 3)!!^2}{P_{n-1}} \left\{ \frac{(2n - 1)^2}{P_n} - \frac{1}{P_{n-2}} \right\} = \frac{(-1)^n (2n - 3)!!^2 2s}{P_n P_{n-2}} . \end{aligned}$$

Hence, even convergents P_n/Q_n increase and odd convergents decrease. The first formula in (13) shows by induction that $P_n(1) = (2n + 1)!!$. Indeed, $P_0(1) = 1$, $P_1(1) = 2 \cdot 1^2 + 1 = 3$ and finally

$$P_n(1) = (2n - 1)!! (2 + (2n - 1)) = (2n + 1)!! . \tag{14}$$

Since $P_n(s) \geq P_n(1)$ for $s \geq 1$, we obtain for such s that

$$\left| \frac{Q_{n-1}}{P_{n-1}} - \frac{Q_n}{P_n} \right| \leq \frac{(2n - 1)!!^2}{(2n + 1)!! (2n - 1)!!} = \frac{1}{2n + 1} ,$$

implying the existence of $\lim_n Q_n/P_n$ (and hence of $\lim_n P_n/Q_n$) for $s \geq 1$. □

Since most likely Brouncker considered integer s , he probably didn’t prove the convergence for $1 > s > 0$. The easiest way to establish the convergence for all $s > 0$ is to apply a simple old theorem by Pringsheim [17], see [12, Ch. I, §5].

THEOREM 2. *Let $p_n > 0$, $q_n > 0$, and*

$$\sum_{n=1}^{\infty} \left(\frac{q_{n-1} q_n}{p_n} \right)^{1/2} = +\infty .$$

Then $q_0 + \mathbf{K}_{n=1}^{\infty} \left(\frac{p_n}{q_n} \right)$ converges.

In our case $q_n = 2s$ and $p_n = (2n - 1)^2$. Therefore the convergence of continued fraction (4) follows from the divergence of the harmonic series.

Formulas (13) also imply that the denominators in the boxed Brouncker’s formulas are products $Q_n(s)Q_n(s + 2)$ of polynomials with positive coefficients and therefore have positive coefficients as well.

For $n = 0, 1, 2$, Brouncker’s formulas show that

$$P_n(s)P_n(s + 2) - (s + 1)^2 Q_n(s)Q_n(s + 2) = c_n , \tag{15}$$

where c_n is a constant. If it is true for every n , then c_n can be easily found. Indeed, putting $s = -1$ in (15), we obtain $c_n = P_n(-1)P_n(1)$. By (14) we have

$P_n(1) = (2n + 1)!!$. Since $P_n(s)$ is odd for even n and is even for odd n , $c_n = -(-1)^n[(2n + 1)!!]^2$. It follows that for $s > 0$,

$$\frac{P_{2k}(s)}{Q_{2k}(s)} \cdot \frac{P_{2k}(s + 2)}{Q_{2k}(s + 2)} < (s + 1)^2 < \frac{P_{2k+1}(s)}{Q_{2k+1}(s)} \cdot \frac{P_{2k+1}(s + 2)}{Q_{2k+1}(s + 2)}.$$

Since continued fraction (4) converges and all its terms are positive, even convergents increase to b whereas odd convergents decrease to b . Passing to the limit in the above inequalities, we obtain that continued fraction $b(s)$ satisfies functional equation (5). The inequality $s < b(s)$ is clear from (4). Thus the proof is completed by the following technical lemma, which can be proved by induction (see its proof in [13]).

LEMMA 3. *Let $P_n(s)/Q_n(s)$ be the n -th convergent to Brouncker's continued fraction (4). Then*

$$P_n(s)P_n(s + 2) - (s + 1)^2Q_n(s)Q_n(s + 2) = (-1)^{n+1}[(2n + 1)!!]^2. \tag{16}$$

The following theorem describes all solutions to functional equation (5) exceeding s for $s > 0$. It also interpolates the relationship between Wallis' product and Brouncker's continued fraction to all $s \geq 0$.

THEOREM 4. *Let $b(s)$ be a function on $(0, +\infty)$ satisfying (5) and the inequality $s < b(s)$ for $s > C$, where C is some constant. Then*

$$b(s) = (s + 1) \prod_{n=1}^{\infty} \frac{(s + 4n - 3)(s + 4n + 1)}{(s + 4n - 1)^2} = s + \prod_{n=1}^{\infty} \left(\frac{(2n - 1)^2}{2s} \right) \tag{17}$$

for every positive s .

PROOF. Repeating Brouncker's arguments, we obtain

$$\begin{aligned} b(s) &= \frac{(s + 1)^2}{(s + 3)^2} b(s + 4) = \frac{(s + 1)^2}{(s + 3)^2} \cdot \frac{(s + 5)^2}{(s + 7)^2} b(s + 8) = \\ &= \frac{(s + 1)^2}{(s + 3)^2} \cdot \frac{(s + 5)^2}{(s + 7)^2} \cdot \dots \cdot \frac{(s + 4n - 3)^2}{(s + 4n - 1)^2} b(s + 4n) = \\ &= (s + 1) \frac{(s + 1)(s + 5)}{(s + 3)^2} \cdot \dots \cdot \frac{(s + 4n - 3)(s + 4n + 1)}{(s + 4n - 1)^2} \frac{b(s + 4n)}{(s + 4n + 1)}. \end{aligned}$$

Multipliers are grouped in accordance with the rule of Wallis' formula:

$$\frac{(s + 4n - 3)(s + 4n + 1)}{(s + 4n - 1)^2} = 1 - \frac{4}{(s + 4n - 1)^2},$$

which provides the convergence of the product for all $s \neq -(4n - 1)$. For positive s it converges to $b(s)$ since $s + 4n < b(s + 4n) < s + 4n + 1$ by the assumption that $s < b(s)$ for $s > C$. □

A simple analysis of the proof presented shows that by March 1665, Brouncker had at his disposal elements of the theory of positive continued fractions including recursion formulas (13). It looks therefore that Wallis' question fell into fertile soil. Brouncker proposed several new ideas but a very new and important idea here was that of considering a continuum of *asymptotic series*.

Definition. A formal power series

$$c_0 + \frac{c_1}{s} + \frac{c_2}{s^2} + \frac{c_3}{s^3} + \dots$$

is called an *asymptotic expansion* of a function $y(s)$ at ∞ if for every n

$$y(s) = \sum_{k=0}^n \frac{c_k}{s^k} + O\left(\frac{1}{s^{n+1}}\right), \quad s \rightarrow \infty.$$

Summing up Brouncker’s method (*Brouncker’s program*) puts together four different objects:

- 1) functional equation $y(s)y(s + 2) = (s + 1)^2$;
- 2) asymptotic series (11);
- 3) continued fraction (4) corresponding to this series;
- 4) infinite product (17) locating singular points of $y(s)$.

4. Euler

In his letter to Goldbach of October 13, 1729, Euler applied Brouncker’s program to answer a question posed by Daniel Bernoulli and Goldbach. The problem was to find a formula extending the factorial $n \rightarrow n! = 1 \cdot 2 \cdot \dots \cdot n$ to real values of n .

Arguing by analogy, one can search an extension $\Gamma(x)$ for $\Gamma(n + 1) = n!$ as a solution to

$$\Gamma(x + 1) = x\Gamma(x). \tag{18}$$

Let $0 \leq x < 1$. Iterating (18), we obtain

$$\Gamma(x) = \frac{\Gamma(x + n + 1)}{x(x + 1) \dots (x + n)}, \quad n \in \mathbb{Z}_+. \tag{19}$$

Now if one can find an asymptotic formula for $\Gamma(t + 1)$ as $t \rightarrow +\infty$, then it can be used to define $\Gamma(x)$ as a limit of elementary functions.

Definition. A positive sequence $c_0 = 1, c_1, \dots, c_n, \dots$ is called *convex* if $2c_k \leq c_{k+1} + c_{k-1}$ for $k = 1, 2, \dots$

The function $\log \Gamma(t + 1)$ interpolates the sequence $z_k = \log k!$ at $t = k$. The sequence $\{z_k\}_{k \geq 0}$ is convex. Since $\log \Gamma(t + 1)$ interpolates it, one may assume that the graph of $\log \Gamma(t + 1)$ is convex too. Comparing on the coordinate plane the slopes of the three chords determined by the points $(n - 1, z_{n-1}), (n, z_n)$, and $(n, z_n), (x + n, \log \Gamma(x + n + 1))$, and finally by $(n, z_n), (n + 1, z_{n+1})$, we obtain that

$$\log n \leq \frac{\log \Gamma(x + n + 1) - \log n!}{x} \leq \log(n + 1),$$

or equivalently,

$$n^x n! \leq \Gamma(x + n + 1) \leq (n + 1)^x n!. \tag{20}$$

The substitution of these inequalities in (19) shows that

$$\frac{n^x n!}{x(x + 1) \dots (x + n)} \leq \Gamma(x) \leq \frac{(n + 1)^x n!}{x(x + 1) \dots (x + n)},$$

which leads to Euler’s definition of the Gamma function:

$$\Gamma(x) = \lim_{n \rightarrow \infty} \frac{n^x n!}{x(x + 1) \dots (x + n)}. \tag{21}$$

The above arguments show that if (18) has a logarithmic convex solution, then it is defined by (21). Next,

$$\frac{n^x n!}{x(x+1)\dots(x+n)} = \frac{1}{x} \left(\frac{n}{n+1}\right)^x \prod_{j=1}^n \left(1 + \frac{x}{j}\right)^{-1} \left(1 + \frac{1}{j}\right)^x \tag{22}$$

implies the existence of the limit in (21), since

$$\left(1 + \frac{x}{j}\right)^{-1} \left(1 + \frac{1}{j}\right)^x = 1 + \frac{x(x-1)}{2j^2} + O\left(\frac{1}{j^3}\right).$$

Finally,

$$\frac{d^2}{dx^2} \log \left(\frac{n^x n!}{x(x+1)\dots(x+n)}\right) = \sum_{j=0}^n \frac{1}{(x+j)^2} > 0$$

shows that the limit function $\Gamma(x)$ is logarithmic convex and

$$\frac{d^2}{dx^2} \log \Gamma(x) = \sum_{j=0}^{\infty} \frac{1}{(x+j)^2}. \tag{23}$$

COROLLARY 5 (The Bohr–Mollerup Theorem, 1922 [1, p. 35]). *If $f(x)$ is a positive logarithmic convex function on $x > 0$ satisfying $f(1) = 1$, $f(x+1) = xf(x)$, then $f(x) = \Gamma(x)$.*

It is clear from [3, §17, 19, 20] that Euler continued his attempts at recovering the original Brouncker’s proof for quite a long time. First, using Wallis’ method of infinite products and Theorem 4, Euler obtained the following formula

$$s + \prod_{n=1}^{\infty} \left(\frac{(2n-1)^2}{2s}\right) = (s+1) \frac{\int_0^1 \frac{x^{s+2} dx}{\sqrt{1-x^4}}}{\int_0^1 \frac{x^s dx}{\sqrt{1-x^4}}} \tag{24}$$

The idea is very simple. Integration by parts shows that

$$\int_0^1 x^{m-1} (1-x^4)^{\frac{3}{4}-1} dx = \frac{m+2}{m} \int_0^1 x^{m+4-1} (1-x^4)^{\frac{3}{4}-1} dx.$$

Iterations of this formula give the products shifted by four exactly as zeros of numerators and denominators in Brouncker’s product (17). Repeating Wallis’ proof, one can easily obtain from here that Brouncker’s product equals the quotient of integrals in (24). Then, using calculus, Euler proved that, indeed, formula (24) is correct [3]. Euler’s formula (24) is only one step away from Ramanujan’s formula for $b(s)$:

$$\int_0^1 x^s (1-x^4)^{1/2-1} dx = \frac{1}{4} \int_0^1 x^{\frac{s+1}{4}-1} (1-x)^{\frac{1}{2}-1} dx = \frac{1}{4} B\left(\frac{s+1}{4}, \frac{1}{2}\right),$$

where B is Euler’s Beta function. Combining the well-known expression of the Beta function via Gamma function, we obtain Ramanujan’s formula

$$b(s) = s + \prod_{n=1}^{\infty} \left(\frac{(2n-1)^2}{2s}\right) = 4 \frac{\left[\Gamma\left(\frac{3+s}{4}\right)\right]^2}{\left[\Gamma\left(\frac{1+s}{4}\right)\right]^2} = R(s). \tag{25}$$

As soon as one has Ramanujan’s formula (25), one can easily prove it! Indeed, using the functional equation for the Gamma function we see that

$$R(s)R(s + 2) = 4^2 \left[\frac{\Gamma(\frac{3+s}{4})\Gamma(1 + \frac{1+s}{4})}{\Gamma(\frac{1+s}{4})\Gamma(\frac{3+s}{4})} \right]^2 = (s + 1)^2 ,$$

so R satisfies the same functional equation as b . Now Stirling’s formula

$$\Gamma(x) \sim \sqrt{2\pi}x^{x-1/2}e^{-x} \exp \left\{ \sum_{k=1}^{\infty} \frac{B_{2k}}{2k(2k-1)x^{2k-1}} \right\} \tag{26}$$

where B_n are the Bernoulli numbers defined by

$$\frac{z}{1 - e^{-z}} = \sum_{n=0}^{\infty} \frac{B_n(-1)^n}{n!} z^n , \tag{27}$$

shows that $R(s) > s$ for big s . By Theorem 4 this proves both Ramanujan’s formula (25) and Euler’s formula (24).

5. The Asymptotic Expansion for $b(s)$

Since Brouncker’s function $b(s)$ and Euler’s Gamma function $\Gamma(s)$ are closely related, one may expect that an analogue of (26) holds for $b(s)$. The key role here is played by the continued fraction

$$y(s) = \mathbf{K}_{n=1}^{\infty} \left(\frac{n^2}{s} \right) , \tag{28}$$

partial numerators of which appear in the right-hand side of (5). In his main paper [3] on continued fractions, Euler presented a simple method for evaluation of such continued fractions with differential calculus.

LEMMA 6 (Euler [3]). *Let R and P be two positive functions on $(0, 1)$ such that for $n = 0, 1, 2, \dots$ and some positive α, β, γ ,*

$$(a + n\alpha) \int_0^1 PR^n dx = (b + n\beta) \int_0^1 PR^{n+1} dx + (c + n\gamma) \int_0^1 PR^{n+2} dx ,$$

then

$$\frac{\int_0^1 PR dx}{\int_0^1 P dx} = \frac{a}{b + \beta} + \frac{(a + \alpha)c}{b + \beta} + \frac{(a + 2\alpha)(c + \gamma)}{b + 2\beta} + \frac{(a + 3\alpha)(c + 2\gamma)}{b + 3\beta} + \dots$$

Keeping in mind the integration by parts formula, one can search P and R as functions satisfying the following identity with indefinite integrals

$$(a + n\alpha) \int PR^n dx + R^{n+1}S = (b + n\beta) \int PR^{n+1} dx + (c + n\gamma) \int PR^{n+2} dx .$$

If $R^{n+1}S$ vanishes at 0 and 1, then P and R must satisfy the conditions of Lemma 6. Euler’s formula in differentials looks as follows

$$(a + n\alpha)P dx + R dS + (n + 1)S dR = (b + n\beta)PR dx + (c + n\gamma)PR^2 dx .$$

Considering it as a polynomial in n , one can replace it with a system

$$\begin{aligned} aP dx + R dS + S dR &= bPR dx + cPR^2 dx , \\ \alpha P dx + S dR &= \beta PR dx + \gamma PR^2 dx . \end{aligned}$$

Solving both equations in Pdx , we find that

$$Pdx = \frac{RdS + SdR}{bR + cR^2 - a} = \frac{SdR}{\beta R + \gamma R^2 - \alpha}. \tag{29}$$

It follows from the last equation of (29) that

$$\begin{aligned} \frac{dS}{S} &= \frac{(b - \beta)RdR + (c - \gamma)R^2dR - (a - \alpha)dR}{\beta R^2 + \gamma R^3 - \alpha R} = \\ &= \frac{(a - \alpha)dR}{\alpha R} + \frac{(\alpha b - \beta a)dR + (\alpha c - \gamma a)RdR}{\alpha(\beta R + \gamma R^2 - \alpha)}. \end{aligned} \tag{30}$$

To obtain a formula for continued fraction (28), we apply Euler’s method with

$$\begin{aligned} a = 1, \quad b = s, \quad c = 2, \quad \alpha b - \beta a = s, \\ \alpha = 1, \quad \beta = 0, \quad \gamma = 1, \quad \alpha c - \gamma a = 1. \end{aligned}$$

By Lemma 6,

$$\mathbf{K}_{n=1}^{\infty} \left(\frac{n^2}{s} \right) = \frac{\int_0^1 x \left(\frac{1-x}{1+x} \right)^{s/2} \frac{dx}{\sqrt{1-x^2}}}{\int_0^1 \left(\frac{1-x}{1+x} \right)^{s/2} \frac{dx}{\sqrt{1-x^2}}}. \tag{31}$$

THEOREM 7 (Euler [3, §69]). For $s > 0$,

$$\frac{1}{s + \mathbf{K}_{n=1}^{\infty} \left(\frac{n^2}{s} \right)} = 2 \int_0^1 \frac{x^s dx}{1+x^2} \stackrel{x=e^{-t}}{=} \int_0^{+\infty} \frac{e^{-st} dt}{\cosh(t)}. \tag{32}$$

PROOF. By integration by parts followed by the substitution $x := (1 - x)/(1 + x)$,

$$\begin{aligned} \int_0^1 x \left(\frac{1-x}{1+x} \right)^{s/2} \frac{dx}{\sqrt{1-x^2}} &= \\ &= 1 - s \int_0^1 \left(\frac{1-x}{1+x} \right)^{s/2} \frac{dx}{\sqrt{1-x^2}} = 1 - s \int_0^1 \frac{x^{(s-1)/2} dx}{1+x} \end{aligned}$$

and (31) result in the formula

$$\frac{1}{s + \mathbf{K}_{n=1}^{\infty} \left(\frac{n^2}{s} \right)} = \int_0^1 \frac{x^{(s-1)/2} dx}{1+x} = 2 \int_0^1 \frac{x^s dx}{1+x^2}. \tag{33}$$

□

Expanding $(1 + x^2)^{-1}$ into the power series in (32) and integrating, we obtain

$$\frac{1}{s + \mathbf{K}_{n=1}^{\infty} \left(\frac{n^2}{s} \right)} = 2 \sum_{k=0}^{\infty} \frac{(-1)^k}{s + 2k + 1}, \tag{34}$$

which gives the functional equation for $y(s)$:

$$y(2s) + y(2s + 2) = \frac{2}{2s + 1}. \tag{35}$$

The following elementary lemma describes the solutions to (35) vanishing at $+\infty$. Notice that $\lim_{s \rightarrow +\infty} y(s) = 0$ by (32).

LEMMA 8. Let $g(s)$ be monotonic on $(0, +\infty)$ and $\lim_{s \rightarrow +\infty} g(s) = 0$. Then the unique solution to the functional equation

$$f(s) + f(s + 1) = g(s) \tag{36}$$

satisfying $\lim_{x \rightarrow +\infty} f(s) = 0$ is given by

$$f(s) = \sum_{n=0}^{\infty} (-1)^n g(s + n). \tag{37}$$

On the other hand, an application of the logarithm to (5) shows that

$$\log b(2s) + \log b(2(s + 1)) = 2 \log(2s + 1).$$

The differentiation now implies that $(b'/b)(2s)$ satisfies (35). By Theorem 4,

$$\frac{b'}{b}(s) = \frac{1}{s + 1} + \sum_{n=1}^{\infty} \frac{8}{(s + 4n - 3)(s + 4n - 1)(s + 4n + 1)},$$

implying $\lim_{s \rightarrow +\infty} (b'/b)(2s) = 0$. Hence by Lemma 8,

$$\frac{1}{s + \prod_{n=1}^{\infty} \left(\frac{n^2}{s}\right)} = \frac{b'}{b}(s). \tag{38}$$

THEOREM 9. For $s > 0$,

$$s + \prod_{n=1}^{\infty} \left(\frac{(2n - 1)^2}{2s}\right) = \frac{8\pi^2}{\Gamma^4\left(\frac{1}{4}\right)} \exp\left\{\int_0^s \frac{dt}{t + \prod_{n=1}^{\infty} \left(\frac{n^2}{t}\right)}\right\}. \tag{39}$$

PROOF. Integrating differential equation (38) and observing that

$$y(0) = \frac{1}{y(2)} = \frac{1}{4} \left[\frac{\Gamma\left(\frac{3}{4}\right)}{\Gamma\left(\frac{5}{4}\right)}\right]^2 = 4 \left[\frac{\Gamma\left(\frac{3}{4}\right)\Gamma\left(\frac{1}{4}\right)}{\Gamma^2\left(\frac{1}{4}\right)}\right]^2 = \frac{\Gamma\left(\frac{1}{4}\right)^4}{8\pi^2},$$

we get (39). □

LEMMA 10 (Watson [23]). Let f be a function on $(0, +\infty)$ such that $|f(t)| < M$ for $t > \varepsilon$ and $f(t) = \sum_{k=0}^{\infty} c_k t^k$, $0 < t < 2\varepsilon$. Then

$$\int_0^{+\infty} f(t)e^{-st} dt \sim \sum_{k=0}^{\infty} \frac{k!c_k}{s^{k+1}}, \quad s \rightarrow +\infty \tag{40}$$

is the asymptotic expansion for the Laplace transform of f .

The proof is elementary and can be found in [1, p. 614]. The second formula in (32) and Lemma 10 give the asymptotic expansion for $y(s)$. Indeed,

$$\frac{1}{\cosh(x)} = \sum_{n=0}^{\infty} \frac{E_n}{n!} x^n, \tag{41}$$

where E_n are the Euler numbers: $E_0 = 1$, $E_1 = 0$, $E_2 = -1$, $E_3 = 0$, $E_4 = 5$, $E_5 = 0$, $E_6 = -61$, $E_7 = 0$, $E_8 = 1385$, ... Hence,

$$\frac{1}{s + \prod_{n=1}^{\infty} \left(\frac{n^2}{s}\right)} \sim \sum_{k=0}^{\infty} \frac{E_k}{s^{k+1}}, \quad s \rightarrow +\infty. \tag{42}$$

THEOREM 11. *The following asymptotic relation holds as $s \rightarrow +\infty$:*

$$b(s) = s + \mathbf{K}_{n=1}^{\infty} \left(\frac{(2n-1)^2}{2s} \right) \sim s \cdot \exp \left\{ - \sum_{k=1}^{\infty} \frac{E_{2k}}{2ks^{2k}} \right\} \tag{43}$$

PROOF. By Theorem 4, continued fraction $b(s)$ extends analytically to the left half-plane and vanishes at $s = -1$. Hence, $b(s)$ can be written as

$$b(s) = b(0)(s+1) \exp \left\{ \int_0^{\infty} \gamma(t) dt \right\} \exp \left\{ - \int_s^{\infty} \gamma(t) dt \right\},$$

where

$$\gamma(t) = \frac{1}{t + \mathbf{K}_{n=1}^{\infty} \left(\frac{n^2}{t} \right)} - \frac{1}{1+t}.$$

By (42),

$$\gamma(t) \sim \sum_{k=1}^{\infty} \frac{E_k - (-1)^k}{t^{k+1}}, \quad t \rightarrow +\infty.$$

Integration shows that

$$\int_s^{\infty} \gamma(t) dt \sim \sum_{k=1}^{\infty} \frac{E_k - (-1)^k}{ks^k}, \quad t \rightarrow +\infty.$$

Since $b(s) \sim s$ as $s \rightarrow +\infty$, we see that

$$\int_0^{+\infty} \left(\frac{1}{t + \mathbf{K}_{n=1}^{\infty} \left(\frac{n^2}{t} \right)} - \frac{1}{1+t} \right) dt = \ln \left(\frac{8\pi^2}{\Gamma \left(\frac{1}{4} \right)^4} \right) \tag{44}$$

and

$$s + \mathbf{K}_{n=1}^{\infty} \left(\frac{(2n-1)^2}{2s} \right) \sim (s+1) \exp \left\{ - \sum_{k=1}^{\infty} \frac{E_k - (-1)^k}{ks^k} \right\}. \tag{45}$$

The proof is completed by observing that

$$\sum_{k=1}^{\infty} \frac{(-1)^k}{ks^k} = - \ln \left(1 + \frac{1}{s} \right), \quad s > 1. \quad \square$$

It is useful to compare formula (43) with Stirling's formula (26) for the Gamma function. The coefficients in (26) are the Bernoulli numbers, whereas in (43) they are the Euler numbers. Easy computations with the first Euler numbers show that

$$\begin{aligned} s + \mathbf{K}_{n=1}^{\infty} \left(\frac{(2n-1)^2}{2s} \right) &\sim s \cdot \exp \left\{ \frac{1}{2s^2} - \frac{5}{4s^4} + \frac{61}{6s^6} + O \left(\frac{1}{s^8} \right) \right\} = \\ &= s + \frac{1}{2s} - \frac{9}{8s^3} + \frac{153}{16s^5} + O \left(\frac{1}{s^7} \right). \end{aligned}$$

Compare this with the asymptotic series for Brouncker's continued fraction $b(s)$ obtained in (11) by a different method.

If we apply Euler's method with

$$\begin{aligned} a &= 2, & b &= s, & c &= 1, & \alpha b - \beta a &= s, \\ \alpha &= 1, & \beta &= 0, & \gamma &= 1, & \alpha c - \gamma a &= -1, \end{aligned}$$

where

$$\begin{aligned}
 A_1 &= ap + bq + cr + ds, & A_2 &= ar - bs - cp + dq, \\
 B_1 &= -aq + bp + cs - dr, & B_2 &= as + br + cq + dp, \\
 C_1 &= ar + bs - cp - dq, & C_2 &= -ap + bq - cr + ds, \\
 D_1 &= -as + br - cq + dp, & D_2 &= -aq - bp + cs + dr, \\
 A_3 &= -as - br + cq + dp, & A_4 &= aq - bp + cs - dr, \\
 B_3 &= ar - bs + cp - dq, & B_4 &= ap + bq - cr - ds, \\
 C_3 &= aq + bp + cs + dr, & C_4 &= as - br - cq + dp, \\
 D_3 &= -ap + bq + cr - ds, & D_4 &= ar + bs + cp + dq,
 \end{aligned}$$

As an application Euler presented the following square matrix:

$$\left\| \begin{array}{cccc}
 +68 & -29 & +41 & -37 \\
 -17 & +31 & +79 & +32 \\
 +59 & +28 & -23 & +61 \\
 -11 & -77 & +8 & +49
 \end{array} \right\|$$

with orthogonal columns and rows. The sum of the squares in rows and columns equals 8515. The same value has the sum at the corners of the big and central interior squares:

$$68^2 + 37^2 + 49^2 + 11^2 = 31^2 + 79^2 + 23^2 + 28^2 = 8515.$$

In the conclusion [6, §36], Euler writes:

This solution should be paid more attention to since I arrived at it not with some definite method but rather through some guesses; and since in addition it depends on eight arbitrary parameters, which after the normalization can be reduced to seven, one can hardly doubt that it is universal and lists all possible cases. If somebody will find a direct way to this solution, then it will be admitted that he made an outstanding contribution to Analysis. Whether similar solutions exist for wider squares consisting of 25, 36, etc. numbers, I hardly dare to claim. Here not only Algebra but the Diophantine method seemingly will get a great contribution.

Euler’s Problem. *Find rational (or may be algebraic) parameterizations of the manifold of real orthogonal matrices with independent parameters.*

A solution to Euler’s problem can be found in [21, Ch. IX, §1], where it is proved that any rotation of \mathbb{R}^n is a product of two-dimensional rotations of a very special form. Recall that rotations of \mathbb{R}^n correspond to orthogonal $n \times n$ matrices. Since any rotation of a plane can be rationally parameterized (see (49)), their products can be rationally parameterized too.

In 1855, exactly two hundred years after Brouncker’s great discovery, Chebyshev found a parameterization which gave rise to the Theory of Orthogonal Polynomials. Chebyshev’s solution followed important contributions by Gauss, Jacobi, Legendre, and Sturm.

7. Gauss' Continued Fractions for ${}_2F_1$

In [4] Euler presented three different methods to assign a finite value to Wallis' divergent hypergeometric series

$${}_2F_0 \left(\begin{matrix} 1, 1 \\ - \end{matrix}; -\frac{1}{s} \right) = \sum_{k=0}^{\infty} \frac{(-1)^k k!}{s^k} \tag{50}$$

and even computed this value 0.596347362323 at $s = 1$. Notice that this problem is similar to the evaluation of $b(0)$, which according to Brouncker is $1/b(2)$. The formalism of the formal Laurent series is quite useful in handling divergent asymptotic series and exploits Part 3 of Brouncker's Program; see p. 880. The field $\mathbb{C}([1/s])$ of formal Laurent series at $s = \infty$ consists of formal series

$$f(z) = \sum_{k \in \mathbb{Z}} \frac{c_k}{s^k}, \tag{51}$$

where $c_k \in \mathbb{C}$ and $c_k \neq 0$ only for a finite number of negative indices k . It is equipped with a non-archimedean norm

$$\|f\| = \exp(\deg(f)), \quad \deg(f) = -\inf\{k \in \mathbb{Z} : c_k \neq 0\}. \tag{52}$$

We put

$$[f] = \sum_{k \leq 0} \frac{c_k}{z^k}, \quad \text{Frac}(f) = \sum_{k > 0} \frac{c_k}{z^k}.$$

Similar to real numbers, $[f]$ is called the *integer* part and $\text{Frac}(f)$ the *fractional* part of f . Exploiting Brouncker's analogy between decimal fractions and rational functions, and the introduced notations, one can easily develop any element in $\mathbb{C}([1/s])$ into a continued fraction

$$b_0(s) + \overset{\infty}{\underset{k=1}{\mathbf{K}}} \left(\frac{1}{b_k(s)} \right), \tag{53}$$

converging in $\mathbb{C}([1/s])$. Here $b_k(s)$ are polynomials in s . If $\deg(b_k) \geq 1$ for every $k \geq 1$, then this continued fraction, called the *P-fraction* of $f \in \mathbb{C}([1/s])$, is unique. Therefore, it looks helpful to develop some general methods for construction of such continued fractions, at least for special Laurent series.

THEOREM 12. *Suppose that for infinitely many n a formal Laurent series f can be represented as*

$$f = b_0 + \frac{1}{b_1} + \frac{1}{b_2} + \dots + \frac{1}{b_n} + \frac{1}{g_n}, \tag{54}$$

where b_j are polynomials with $\deg(b_j) \geq 0$, $\deg(b_n) \geq 1$ and $\deg(g_n) \geq 0$. Then the continued fraction $b_0 + \overset{\infty}{\underset{k \geq 1}{\mathbf{K}}} \left(\frac{1}{b_k(z)} \right)$ converges to f in $\mathbb{C}([1/z])$.

PROOF. Notice that $\deg(b_j) \geq 1$ infinitely often, which implies $\deg(Q_n) \rightarrow +\infty$. By (54),

$$f - \frac{P_n}{Q_n} = \frac{(-1)^n}{Q_n^2 \left(g_n + \frac{Q_{n-1}}{Q_n} \right)}.$$

Since $\deg(g_n) \geq 0$, $\deg(Q_{n-1}/Q_n) \leq -1$, and $\deg(Q_n) \rightarrow +\infty$, this proves the theorem. □

Remark. As in the number field case (Lemma 6, in which $g_n \geq 0$), it is essential that $\deg(g_n) \geq 0$. Indeed, let $\{b_k\}_{k \geq 1}$ be any sequence of polynomials such that $\deg(b_k) \geq 1$ and $b_1 = z$. It determines the continued fraction $\mathbf{K}_{k \geq 1} \left(\frac{1}{b_k(z)} \right)$. The formula

$$\frac{Q_n}{Q_{n-1}} = b_n + \frac{1}{b_{n-1} + \frac{1}{b_{n-2} + \dots + \frac{1}{b_1}}}$$

shows that the identity

$$\frac{1}{z} = \frac{1}{2z} + \frac{1}{b_2} + \dots + \frac{1}{b_n} + \frac{1}{g_n}$$

holds with $g_n = -Q_{n-1}/Q_n$, $\deg(g_n) \leq -1$. However, since two different P -fractions cannot be equal, this equation is not valid in the limit.

The Pochhammer symbol

$$(x)_n = \frac{\Gamma(x+n)}{\Gamma(x)} = x(x+1)\dots(x+n-1)$$

considerably simplifies the notation for hypergeometric series. As usual, the product with the empty set of multipliers has the value 1. A direct generalization of Wallis' hypergeometric series is the series in $\mathbb{C}([1/s])$

$${}_2F_0 \left(\begin{matrix} a, b \\ - \end{matrix}; -\frac{1}{s} \right) = \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{n!} \frac{(-1)^n}{s^n}. \tag{55}$$

THEOREM 13 (Euler [4]). *In $\mathbb{C}([1/z])$,*

$$\frac{{}_2F_0 \left(\begin{matrix} a, b \\ - \end{matrix}; -\frac{1}{s} \right)}{{}_2F_0 \left(\begin{matrix} a, b+1 \\ - \end{matrix}; -\frac{1}{s} \right)} = 1 + \frac{a}{s} + \frac{b+1}{1} + \frac{a+1}{s} + \frac{b+2}{1} + \frac{a+2}{s+\dots}. \tag{56}$$

PROOF. Elementary algebra with series shows that

$${}_2F_0 \left(\begin{matrix} a, b+1 \\ - \end{matrix}; -\frac{1}{s} \right) = {}_2F_0 \left(\begin{matrix} a, b \\ - \end{matrix}; -\frac{1}{s} \right) - \frac{a}{s} \cdot {}_2F_0 \left(\begin{matrix} a+1, b+1 \\ - \end{matrix}; -\frac{1}{s} \right).$$

It follows that

$$\frac{{}_2F_0 \left(\begin{matrix} a, b \\ - \end{matrix}; -\frac{1}{s} \right)}{{}_2F_0 \left(\begin{matrix} a, b+1 \\ - \end{matrix}; -\frac{1}{s} \right)} = 1 + \frac{a/s}{\frac{{}_2F_0 \left(\begin{matrix} a, b+1 \\ - \end{matrix}; -\frac{1}{s} \right)}{{}_2F_0 \left(\begin{matrix} a+1, b+1 \\ - \end{matrix}; -\frac{1}{s} \right)}}.$$

Since ${}_2F_0$ is symmetric with respect to a and b , the iterations complete the proof by Theorem 12. □

COROLLARY 14 (Euler [4]). *In $\mathbb{C}([1/s])$,*

$$\frac{1}{s} \cdot {}_2F_0 \left(\begin{matrix} 1, 1 \\ - \end{matrix}; -\frac{1}{s} \right) = \sum_{k=0}^{\infty} \frac{(-1)^k k!}{s^{k+1}} = \frac{1}{s} + \frac{1}{1} + \frac{1}{s} + \frac{2}{1} + \frac{2}{s} + \frac{3}{1} + \frac{3}{s} + \dots \tag{57}$$

PROOF. It follows from Theorem 13 and the obvious identity

$${}_2F_0 \left(\begin{matrix} 1, 2 \\ - \end{matrix}; -\frac{1}{s} \right) = 1 - \frac{2!}{s} + \frac{2! \cdot 3!}{2!s^2} - \frac{3! \cdot 4!}{3!s^3} - \dots = s \left\{ 1 - {}_2F_0 \left(\begin{matrix} 1, 1 \\ - \end{matrix}; -\frac{1}{s} \right) \right\}.$$

Just substitute this in (56) and make equivalence transforms. □

The continued fraction in (57) converges for every positive s and for $s = 1$ Euler found its value 0.596347362323. By formal differentiation of

$$y(s) = \sum_{n=0}^{\infty} \frac{(-1)^n n!}{s^{n+1}}.$$

Euler obtained that y satisfies the differential equation $y' = y - 1/s$. This equation has a solution

$$y(s) = e^s \int_s^{\infty} e^{-t} \frac{dt}{t} = \int_0^{\infty} \frac{e^{-st} dt}{1+t} = \int_0^{\infty} \frac{e^{-t} dt}{t+s}. \tag{58}$$

Then by Lemma 10,

$$\int_0^{\infty} e^{-st} \frac{dt}{1+t} = \int_0^{\infty} \frac{e^{-t} dt}{t+s} \sim \sum_{n=0}^{\infty} \frac{(-1)^n n!}{s^{n+1}}. \tag{59}$$

Euler found that the left integral also equals 0.596347362323 for $s = 1$. Most likely this happens since the integral equals the continued fraction. Using the binomial theorem and Lemma 10, we obtain

$$E(a, b; s) = \frac{1}{\Gamma(a)} \int_0^{\infty} \frac{e^{-x} x^{a-1}}{(1+x/s)^b} dx \sim {}_2F_0 \left(\begin{matrix} a, b \\ - \end{matrix}; -\frac{1}{s} \right). \tag{60}$$

To handle the apparent asymmetry of $E(a, b; s)$ in a and b , we write

$$\frac{1}{(1+x/t)^b} = \frac{1}{\Gamma(b)} \int_0^{\infty} e^{-(1+x/s)y} y^{b-1} dy.$$

It follows that

$$E(a, b; s) = \frac{1}{\Gamma(a)\Gamma(b)} \int_0^{\infty} \int_0^{\infty} e^{-x-y-xy/s} x^{a-1} y^{b-1} dx dy = E(b, a; s).$$

Now integration by parts and the symmetry of $E(a, b; s)$ imply

$$E(a, b; s) = E(a, b+1; s) + \frac{a}{s} E(a+1, b+1; s), \tag{61}$$

$$E(a, b; s) = E(a+1, b; s) + \frac{b}{s} E(a+1, b+1; s). \tag{62}$$

Using (61) and (62) as in Lemma 6, we obtain the following corollary:

COROLLARY 15. *For positive a, b and s ,*

$$\frac{E(a, b, s)}{E(a, b+1; s)} = 1 + \frac{a}{s} + \frac{b+1}{1} + \frac{a+1}{s} + \frac{b+2}{1} + \frac{a+2}{s} + \dots$$

Putting $b = 0$ in (61) and $b = 1$ in (62), we find that

$$X = \frac{E(a+1, 1; s)}{E(a+1, 2; s)} = \frac{1-E}{(a+s)E-s},$$

where $E = E(a, 1; s)$. Resolving this equation in E , we find

$$\begin{aligned} \frac{1}{\Gamma(a)} \int_0^{\infty} \frac{e^{-x} x^{a-1}}{x+s} dx &= \frac{E(a, 1; s)}{s} = \frac{1}{s} + \frac{a}{1} + \frac{1}{sX} = \\ &= \frac{1}{s} + \frac{a}{1} + \frac{1}{s} + \frac{a+1}{1} + \frac{2}{s} + \dots + \frac{n}{s} + \frac{a+n}{1} + \dots, \tag{63} \end{aligned}$$

which explains Euler’s striking calculations. It should be mentioned that these formulas were later proved by Stieltjes to develop the theory of moments [18].

Similar arguments can be applied to the hypergeometric series

$${}_2F_1 \left(\begin{matrix} a, b \\ c \end{matrix}; z \right) = \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{n! (c)_n} z^n, \tag{64}$$

which Euler introduced in [5], although Euler considered partial cases earlier in relation with Wallis' interpolation method. The ratio test shows that the series (64) converges in the unit disc $|z| < 1$ at least for positive a, b, c .

THEOREM 16 (Gauss [7]). *For $z = 1/s$ in $\mathbb{C} ([1/s])$,*

$$\frac{{}_2F_1 \left(\begin{matrix} a, b+1 \\ c+1 \end{matrix}; z \right)}{{}_2F_1 \left(\begin{matrix} a, b \\ c \end{matrix}; z \right)} = \frac{c}{c} - \frac{a(c-b)z}{c+1} - \frac{(b+1)(c-a+1)z}{c+2} - \dots \\ - \frac{(a+n)(c-b+n)z}{c+2n+1} - \frac{(b+n+1)(c-a+n+1)z}{c+2n+2} - \dots$$

The strategy of the proof in case ${}_2F_1$ is absolutely the same as for ${}_2F_0$ considered above. In his work [7] Gauss obtained the convergence of the continued fraction in $\mathbb{C} ([1/s])$. However, the pointwise convergence for $s > 0$ can be easily derived from Euler's integral representation for ${}_2F_1$, which is obtained similarly to (60).

THEOREM 17 (Euler [5]). *For $c > b > 0$,*

$${}_2F_1 \left(\begin{matrix} a, b \\ c \end{matrix}; z \right) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 \frac{t^{b-1}(1-t)^{c-b-1}}{(1-zt)^a} dt. \tag{65}$$

Theorems 16 and 17 imply that the formula

$$\frac{\int_0^1 \frac{t^b(1-t)^{c-b-1}}{(1+zt)^a} dt}{\int_0^1 \frac{t^{b-1}(1-t)^{c-b-1}}{(1+zt)^a} dt} = \frac{b}{c} + \frac{a(c-b)z}{c+1} + \frac{(b+1)(c-a+1)z}{c+2} + \dots \\ + \frac{(a+n)(c-b+n)z}{c+2n+1} + \frac{(b+n+1)(c-a+n+1)z}{c+2n+2} + \dots \tag{66}$$

is valid for positive z . Let us apply (66) to Euler's formula (24) and put $z = 1$, $b = \frac{s+1}{2}$, $c - b = 1/2$, $a = 1/2$ in (66). Then

$$s + \prod_{n=1}^{\infty} \left(\frac{(2n-1)^2}{2s} \right) = \frac{b(s+1)}{c} \frac{{}_2F_1 \left(\begin{matrix} a, b+1 \\ c+1 \end{matrix}; -1 \right)}{{}_2F_1 \left(\begin{matrix} a, b \\ c \end{matrix}; -1 \right)} = \\ = \frac{(s+1)^2}{s+2 \cdot 1} + \frac{1^2}{s+2 \cdot 2} + \frac{(s+3)^2}{s+2 \cdot 3} + \frac{3^2}{s+2 \cdot 4} + \frac{(s+5)^2}{s+2 \cdot 5} + \frac{5^2}{s+2 \cdot 6} + \dots$$

For $s = 1$, this turns into a nice formula for π :

$$\pi = 3 + \frac{1^2}{5} + \frac{4^2}{7} + \frac{3^2}{9} + \frac{6^2}{11} + \frac{5^2}{13} + \frac{8^2}{15} + \dots \tag{67}$$

Putting formally $b = 0$ in Theorem 16 and observing that $F(a, 0; c; -z) = 1$, we arrive at the following important theorem:

THEOREM 18. For $z \geq 0$,

$${}_2F_1 \left(\begin{matrix} a, 1 \\ c \end{matrix}; -z \right) = \frac{1}{1 + \frac{az}{c} + \frac{1 \cdot (c-a)z}{c+1} + \frac{(a+1)cz}{c+2} + \frac{2(c-a+1)z}{c+3} + \dots + \frac{(a+n)(c+n-1)z}{c+2n} + \frac{(n+1)(c-a+n)z}{c+2n+1} + \dots} \quad (68)$$

Since many known functions can be obtained from ${}_2F_1(a, 1; c; -z)$, Theorem 18 gives useful expansions into continued fractions, which converge at least in $\mathbb{C}([1/s])$. In particular,

$$\ln \frac{1+z}{1-z} = 2zF(1/2, 1; 3/2; z^2) = \frac{2z}{1} - \frac{1^2 z^2}{3} - \frac{2^2 z^2}{5} - \frac{3^2 z^2}{7} - \frac{4^2 z^2}{9} - \dots \quad (69)$$

Replacing z by $1/s$ in (69) and applying equivalence transforms, we obtain

$$\frac{1}{2} \ln \frac{s+1}{s-1} = \frac{1}{s} - \frac{1^2}{s-3s} - \frac{2^2}{5s} - \frac{3^2}{7s} - \frac{4^2}{9s} - \dots \quad (70)$$

8. Gauss' Quadrature Formula

In 1671 Newton discovered the 3/8 rule for definite integrals of continuous functions:

$$\int_{x_1}^{x_4} f(x) dx \approx \frac{3h}{8} \{f(x_1) + 3f(x_2) + 3f(x_3) + f(x_4)\} \quad (71)$$

Here $x_2 - x_1 = x_3 - x_2 = x_4 - x_3 = h > 0$. Newton's quadrature formula (71) has a remarkable property: it is a true identity on any polynomial f , $\deg(f) \leq 4 - 1 = 3$. This property of (71) follows from the formula for Lagrange interpolation polynomials

$$L(x) = \sum_{k=1}^n y_k \frac{Q(x)}{(x-x_k)Q'(x_k)}, \quad Q(x) = (x-x_1) \dots (x-x_n), \quad (72)$$

satisfying $L(x_k) = y_k$ for $k = 1, 2, \dots, n$. Fixing the nodes x_k in (72) and integrating in x , we obtain the Newton-Cotes quadrature

$$\int_{x_1}^{x_n} f(x) dx \approx \sum_{k=1}^n l_k f(x_k), \quad l_k = \int_{x_1}^{x_n} \frac{Q(x)}{(x-x_k)Q'(x_k)} dx, \quad (73)$$

which is a true identity on polynomials of degree $n - 1$.

Since Newton-Cotes quadratures depend on n nodes, a proper choice of $\{x_k\}$ may result in a formula which is a true identity on any polynomial of degree $n - 1 + n = 2n - 1$. In [8] Gauss studied this problem for nodes $-1 \leq x_1 < x_2 < \dots < x_n \leq 1$:

$$\int_{-1}^1 f(x) dx \approx \sum_{k=1}^n l_k f(x_k).$$

Quadrature formula (73) is an identity for any polynomial of degree $2n - 1$ if and only if it is the identity for any monomial $f(x) = x^m, m = 0, 1, \dots, 2n - 1$. It follows that for these m ,

$$\delta_m \stackrel{def}{=} \int_{-1}^1 x^m dx - \sum_{k=1}^n l_k x_k^m = \frac{1 - (-1)^{m+1}}{m+1} - \sum_{k=1}^n l_k x_k^m = 0. \quad (74)$$

Since l_k are defined by (73), $\delta_m = 0$ for $m = 0, \dots, n-1$. Notice also that $\delta_m = O(1)$ as $m \rightarrow \infty$. For $|z| > 1$, Gauss considered a convergent Laurent series

$$G(z) = \sum_{m=0}^{\infty} \frac{1 - (-1)^{m+1}}{(m+1)z^{m+1}} = \sum_{k=1}^n l_k \sum_{m=0}^{\infty} \frac{x_k^m}{z^{m+1}} + \sum_{m=n}^{\infty} \frac{\delta_m}{z^{m+1}}$$

and multiplied it by Q to obtain the formula

$$\begin{aligned} Q(z)G(z) &= \sum_{k=1}^n \frac{l_k Q(z)}{z - x_k} + \sum_{m=n}^{\infty} \frac{\delta_m Q(z)}{z^{m+1}} = \\ &= P(z) + \frac{(\delta_n z^{n-1} + \dots + \delta_{2n-1})Q(z)}{z^{2n}} + O\left(\frac{1}{z^{n+1}}\right), \end{aligned} \tag{75}$$

where P is a polynomial. Notice that $\delta_n = \dots = \delta_{2n-1} = 0$ in (75) if and only if $\deg(QG - P) \leq -n - 1$, where $\deg(Q) = n$. Since

$$\ln\left(\frac{z+1}{z-1}\right) = \sum_{m=0}^{\infty} \frac{1 - (-1)^{m+1}}{(m+1)z^{m+1}} = G(z),$$

as elementary calculations with power series show, by Lagrange’s theorem, P/Q is the n th convergent (see (70)) to

$$\ln \frac{z+1}{z-1} = \frac{2}{z} - \frac{1^2}{3z} - \frac{2^2}{5z} - \frac{3^2}{7z} - \frac{4^2}{9z} - \dots$$

It follows that Q is one of the polynomials defined by the Euler–Wallis formulas

$$Q_{n+1}(x) = (2n+1)xQ_n(x) - n^2Q_{n-1}(x), \quad Q_0(x) = 1, \quad Q_1(x) = x. \tag{76}$$

It remains to check that all zeros of Q_n are simple and are located in $(-1, 1)$.

THEOREM 19. *The zeros $x_{n1} < x_{n-11} < \dots < x_{nn}$ of Q_n are located in $(-1, 1)$ and interlace the zeros of Q_{n-1} .*

PROOF. By inspection the statement is true for $Q_0 = 1, Q_1 = x, Q_2 = 3x^2 - 1$ and $Q_3 = 3x(5x^2 - 3)$. Easy induction with (76) shows that $Q_n(x)$ is even if n is even and is odd if n is odd, $\deg(Q_n(x)) = n$ and the leading coefficient of $Q_n(x)$ is $(2n-1)!!$. The identity $(2n+1)n! - n^2(n-1)! = (n+1)!$ and (76) imply by induction that $Q_n(1) = n! > 0$. Now the proof is completed by induction by counting zeros and applying Bolzano’s theorem. \square

THEOREM 20 (Gauss [8]). *For every positive integer n , there are n nodes $-1 < x_1 < \dots < x_n < 1$ such that*

$$\int_{-1}^1 f(x) dx = \sum_{k=1}^n l_k f(x_k) \tag{77}$$

for every polynomial $f, \deg(f) \leq 2n - 1$. The nodes x_k are the zeros of the denominator of the n th convergent to the continued fraction (70).

Notice that since $\deg(Q_n) = n$, the approximation of $G(z)$ by the convergent of the n th order cannot give the order $2n$ at infinity, which implies that $2n - 1$ is the best possible degree in Gauss’s quadrature.

9. Jacobi's Contribution

In [9] Jacobi noticed that the choice of $f(x) = x^k Q_n(x)$ in (77) with $0 \leq k < n$ gives

$$\int_{-1}^1 Q_n dx = \int_{-1}^1 x Q_n dx = \dots = \int_{-1}^1 x^{n-1} Q_n dx = 0, \tag{78}$$

since $\deg(x^k Q_n(x)) \leq 2n - 1$ and $Q_n(x_k) = 0$. On the other hand, in [16] Legendre introduced orthogonal polynomials $P_n(x)$:

$$\int_{-1}^1 P_n(x) P_m(x) dx = \frac{1}{2n + 1} \delta_{nm}, \delta_{nm} \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } n = m \\ 0 & \text{if } n \neq m, \end{cases} \tag{79}$$

which are now called the Legendre polynomials. One look at (78) and (79) is enough to understand that Q_n in (78) are constant multiples of the Legendre Polynomials. Integration by parts

$$\int_{-1}^1 uv' dx = uv \Big|_{-1}^1 - \int_{-1}^1 vu' dx$$

shows that the polynomial

$$v'(x) = \frac{d^n}{dx^n} (x^2 - 1)^n$$

of degree n satisfies (78) since it has zeros of order n at $x = -1$ and $x = 1$. Taking into account the formulas for the leading coefficients of the polynomials, we obtain that

$$Q_n(x) = \frac{(2n - 1)!! \cdot n!}{(2n)!} \frac{d^n}{dx^n} (x^2 - 1)^n.$$

By Rolle's theorem all roots of Q_n are located in $(-1, 1)$.

Around 1843 Jacobi returned to these results and extended them to more general Jacobi's weights. This paper [10] was published in 1859 by Heine after Jacobi's death.

In [10] Jacobi observes that continued fraction (70), which played a crucial role in Gauss' quadratures, is a partial case of a more general Gauss' continued fraction (68). An important ingredient of Gauss' proof was the fact that continued fraction (70) represents in $\mathbb{C} ([1/z])$ the Cauchy type integral of a constant positive weight on $[-1, 1]$. In fact, Gauss considered the interval $[0, 1]$. Comparing Euler's formula (65) with Gauss' continued fraction (68) and keeping in mind the symmetry of ${}_2F_1$ in a and b (see (64)), we can easily obtain a continued fraction for Cauchy integrals of some special weights on $[0, 1]$. More precisely, let $\alpha > 0, \beta > 0$. Then by Theorem 17,

$${}_2F_1 \left(\begin{matrix} \alpha, 1 \\ \alpha + \beta \end{matrix}; z \right) = {}_2F_1 \left(\begin{matrix} 1, \alpha \\ \alpha + \beta \end{matrix}; z \right) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 \frac{t^{\alpha-1}(1-t)^{\beta-1}}{1-zt} dt.$$

Hence by Theorem 18,

$$\begin{aligned} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 \frac{t^{\alpha-1}(1-t)^{\beta-1}}{z-t} dt &= {}_2F_1\left(\alpha, 1; \frac{1}{z}\right) \frac{1}{z} = \\ &= \frac{1}{z} - \frac{\alpha}{\alpha + \beta} - \frac{1 \cdot \beta}{(\alpha + \beta + 1)z} - \frac{(\alpha + 1) \cdot (\alpha + \beta)}{\alpha + \beta + 2} - \frac{2 \cdot (\beta + 1)}{(\alpha + \beta + 3)z} - \dots \\ &\quad - \frac{(\alpha + n) \cdot (\alpha + \beta + n - 1)}{\alpha + \beta + 2n} - \frac{(n + 1) \cdot (\beta + n)}{(\alpha + \beta + 2n + 1)z} - \dots \end{aligned}$$

This formula allows one to obtain quadrature formulas analogous to Gauss' quadrature formula. It can also be used to obtain the recurrence relation for Jacobi orthogonal polynomials; see [20, §3.4, Ch. IV] for details.

It is interesting that Jacobi was aware of Euler's paper [6] which motivated Chebyshev in his discovery of general orthogonal polynomials. Moreover, Jacobi even wrote a manuscript [11], where he explained Euler's ideas. As in the case of [10], [11] was published only after Jacobi's death.

10. Sturm's Theorem

The proof of Theorem 19 can easily be extended to cover a more general case.

THEOREM 21. *Let $\{b_n(x)\}_{n \geq 1}$ be any sequence of real linear polynomials with positive coefficients at x and $\{a_n\}_{n \geq 1}$ be any nonzero real sequence. Then the zeros of the denominators Q_n of any convergent to*

$$\frac{a_1^2}{b_1(x)} - \frac{a_2^2}{b_2(x)} - \frac{a_3^2}{b_3(x)} - \dots - \frac{a_n^2}{b_n(x)} - \dots \tag{80}$$

are all real and interlace the zeros of Q_{n-1} .

PROOF. The difference compared with Theorem 19 is that here it is not necessary to prove that all zeros lie in $[-1, 1]$. Since $Q_n(x) \sim cx^n$, $c > 0$, if $x \rightarrow \infty$, the location of the smallest and the greatest zeros can be controlled. \square

Let us summarize the properties of Q_n , which were used in Theorem 20.

THEOREM 22. *Let Q_n be the denominator of the n th convergent to the continued fraction (80). Then the sequence*

$$f_0(x) = Q_n(x), f_1(x) = Q_{n-1}(x), \dots, f_n(x) = Q_0(x) \tag{81}$$

satisfies the following properties

- (a) *the product $f_0(x)f_1(x)$ changes the sign from $-$ to $+$ if x passes any zero of $f_0(x)$ in the positive direction;*
- (b) *no polynomials $f_k(x)$, $f_{k+1}(x)$, $k = 0, 1, \dots, m - 1$, may have common zeros;*
- (c) *if $f_k(\alpha) = 0$, $1 \leq k \leq m - 1$, then $f_{k-1}(\alpha)f_{k+1}(\alpha) < 0$;*
- (d) *$f_n(x)$ has no real zeros.*

PROOF. Applying the determinant identity to continued fraction (21), we obtain that

$$P_k Q_{k-1} - P_{k-1} Q_k = a_1^2 a_2^2 \dots a_k^2,$$

which implies (b). The Euler-Wallis formula, $Q_{k+1} = b_{k+1}Q_k - Q_{k-1}$, implies (c). Since $Q_0 = 1$, we have (d). Let us prove (a). Since the zeros $\{x_{(n-1)k}\}$ of Q_{n-1} alternate the zeros $\{x_{nk}\}$ of Q_n , we have $x_{(n-1)k} \in [x_{n1}, x_{nn}]$. If n is even, then

$Q_n(x) > 0, Q_{n-1}(x) < 0$ for $x < x_{n1}$, which proves (a) since the signs alternate. The case of odd n is considered similarly. \square

Definition. A sequence of nonzero polynomials

$$f(x) = f_0(x), f_1(x), \dots, f_m(x) \tag{82}$$

in $\mathbb{R}[X]$ is called a Sturm’s series for a polynomial $f(x)$ on $[a, b], a < b$ if conditions (a)–(d) of Theorem 22 are valid on $[a, b]$.

Given $x \in [a, b]$ let $W_S(x)$ be the number of sign changes in (82).

THEOREM 23 (Sturm [1829]). *Let $f \in \mathbb{R}[x]$ be a polynomial with simple roots on $[a, b], a < b$, satisfying $f(a)f(b) \neq 0$, and (82) be a Sturm’s series for $f(x)$. Then the number of roots of $f(x)$ on (a, b) is $W_S(a) - W_S(b)$.*

PROOF. It is clear that a sign change in (82) may occur only at zeros of polynomials f_j . No sign changes happen in f_m , since by (d) it has no zeros on $[a, b]$. If $0 < j < m$, then by (c) the number of the sign variations in the triple $(f_{j-1}(x), f_j(x), f_{j+1}(x))$ is kept invariant when x passes through any zero of f_j . If x passes through a zero of f_0 , then the product f_0f_1 changes the sign from $-$ to $+$, which implies that $W_S(x)$ decreases by one. \square

The second important ingredient of Sturm’s method is a choice of Sturm’s sequence. In 1829 Sturm [19] proposed to take the first two polynomials f and f' from Budan’s series and then apply the method of continued fractions. Thus the idea was to present the continued fraction of f'/f in the form of Lambert and Gauss:

$$\frac{f'}{f} = \frac{1}{b_1} - \frac{1}{b_2} - \frac{1}{b_3} - \dots \tag{83}$$

11. Chebyshev’s Solution to Euler’s Problem

The logarithmic derivative of any polynomial $f = c(x - x_1)^{n_1} \dots (x - x_k)^{n_k}$ is a rational function with simple poles and positive residues

$$\frac{1}{N}(\log f)' = \frac{1}{N} \frac{f'}{f} = \sum_{j=1}^k \frac{n_j/N}{x - x_j}.$$

Definition. A rational function $r(x)$ is called a *Chebyshev rational function* if it can be represented as

$$r(x) = \sum_{j=1}^n \frac{c_j}{x - x_j}, \tag{84}$$

where $c_j > 0$ and $x_1 < x_2 < \dots < x_n$.

The behavior of rational functions r in (84) with real c_j is controlled by the *Cauchy index*:

$$I_a^b \left(\sum_{j=1}^n \frac{c_j}{x - x_j} \right) = \sum_{x_j \in (a,b)} \text{sign}(c_j).$$

THEOREM 24 (Chebyshev [2]). *A rational function r is a Chebyshev function if and only if*

$$r(z) = \frac{1}{b_1(z)} - \frac{1}{b_2(z)} - \dots - \frac{1}{b_n(z)}, \tag{85}$$

where $n = I_{-\infty}^{+\infty} r$ and b_1, b_2, \dots, b_n are real linear polynomials with positive coefficients at z .

PROOF. Since

$$r(x) = \sum_{j=1}^n \frac{c_j}{x - x_j} = \frac{\sum_{j=1}^n c_j}{x} + O\left(\frac{1}{x^2}\right), \quad x \rightarrow \infty,$$

the real polynomial $\llbracket 1/r \rrbracket$ is linear and has a positive coefficient at x . For any Chebyshev function r and $z \in \mathbb{C}$, $\Im r(z) = -\Im z \sum_{j=1}^n c_j |z - x_j|^{-2}$, which implies that $r(z)$ can vanish only on \mathbb{R} . Let

$$r(x) = \frac{1}{b_1(x) - r_2(x)}. \tag{86}$$

Then the poles of r_2 are located on \mathbb{R} at the zeros of r_1 . Since

$$r'(x) = -\sum_{j=1}^n \frac{c_j}{(x - x_j)^2} < 0, \tag{87}$$

the function r_1 decreases on \mathbb{R} . If $r(c) = 0$, then

$$\lim_{x \rightarrow c} (x - c)r_2(x) = \lim_{x \rightarrow c} (x - c)b_1(x) - \frac{1}{r'(c)} > 0$$

by (87) implying that r_2 is a Chebyshev function. By (87) the poles of r_2 interlace the poles of r . This proves that $I_{-\infty}^{+\infty} r_2 = I_{-\infty}^{+\infty} r_1 - 1$. \square

If f is a separable polynomial with real roots, then f'/f is a Chebyshev function. Hence all real polynomials b_j in (83) have positive coefficients at x . If we compare Gauss' continued fraction with Sturm's sequence and Chebyshev's theorem, we come to the conclusion that most likely the orthogonality of the denominators of the convergents is not a particular property of Gauss's continued fraction but follows from a more general fact. Gauss' proof even hints that Lagrange's criteria for convergents must play the crucial role here.

We present here arguments for positive Borel measures σ with compact support $\text{supp}(\sigma)$ in \mathbb{C} . The Cauchy's integral of σ

$$C_\sigma(z) = \int \frac{d\sigma(t)}{z - t}$$

is holomorphic at $z = \infty$ and therefore determines a unique element in $\mathbb{C}(\llbracket 1/z \rrbracket)$. This element can be developed into a continued fraction (80).

THEOREM 25. *Let P/Q be a convergent to C_σ , $n = \text{deg}(Q)$. Then*

$$\int Q(t)t^k d\sigma = 0, \quad k = 0, 1, \dots, n - 1. \tag{88}$$

PROOF. By Lagrange's theorem,

$$C_\sigma - \frac{P}{Q} = O\left(\frac{1}{z^{2n+1}}\right).$$

Following an idea of Lagrange from Number Theory, we consider a linear form in C_σ with polynomial coefficients

$$QC_\sigma - P = \int \frac{Q(z) - Q(t)}{z - t} d\sigma - P + \int \frac{Q(t)}{z - t} d\sigma = O\left(\frac{1}{z^{n+1}}\right).$$

It follows that

$$P = \int \frac{Q(z) - Q(t)}{z - t} d\sigma; \tag{89}$$

$$\int \frac{Q(t)}{z - t} d\sigma = O\left(\frac{1}{z^{n+1}}\right). \tag{90}$$

Observing that the series converges uniformly in t ,

$$\frac{1}{t - z} = -\frac{1}{z} - \frac{t}{z^2} - \frac{t^2}{z^3} - \dots,$$

if $|z| > 2|t|$, $t \in \text{supp}(\sigma)$, we obtain (88). □

Remark. Compare this proof with Gauss' arguments; see (74) and (75). This proof was explicitly used by Jacobi in [10, §8] to establish the orthogonality of Jacobi polynomials. Chebyshev used similar arguments. However they are not so transparent since he considered the case of discrete masses.

Chebyshev's solution to Euler's problem follows from Theorems 24 and 25. Indeed, let $r = P/Q$ be a Chebyshev function, $\deg(Q) = n$. Then the coefficients of both polynomials P and Q are polynomials in c_1, \dots, c_n and in x_1, \dots, x_n . The long division of polynomials shows that the coefficients of the linear polynomials b_1, \dots, b_n are rational functions of these parameters. By Euler–Wallis recursions, the same is true for the polynomials Q_0, \dots, Q_n .

THEOREM 26 (Chebyshev [2]). *Let $\{Q_n\}$ be the denominators of the convergents for the continued fraction*

$$\int_{\mathbb{R}} \frac{d\sigma(t)}{z - t} = \frac{a_1^2}{b_1(z) - b_2(z)} - \dots - \frac{a_n^2}{b_n(z)} - \dots,$$

where $b_n(z) = k_n z + l_n$. Then

$$\int_{\mathbb{R}} Q_n^2(t) d\sigma(t) = \frac{a_1^2 \cdot \dots \cdot a_{n+1}^2}{k_{n+1}}. \tag{91}$$

PROOF. By the Euler–Wallis formulas, $Q_n(t) = k_n \dots k_1 t^n + \dots + Q(0)$. Observing that $Q_n \perp t^k$ for $k = 0, 1, \dots, n - 1$, we obtain from here that

$$\int_{\mathbb{R}} Q_n^2 d\sigma = k_n \dots k_1 \int_{\mathbb{R}} Q_n t^n d\sigma.$$

Integrating the Euler–Wallis formula multiplied by t^{n-1}

$$Q_{n+1} t^{n-1} = k_{n+1} t^n Q_n + l_{n+1} Q_n t^{n-1} - a_{n+1}^2 Q_{n-1} t^{n-1}$$

and using the orthogonality property, we obtain that

$$k_{n+1} \int_{\mathbb{R}} Q_n t^n d\sigma = a_{n+1}^2 \int_{\mathbb{R}} Q_{n-1} t^{n-1} d\sigma,$$

which implies that

$$\int_{\mathbb{R}} Q_n^2 d\sigma = \frac{(k_n \dots k_1)(a_{n+1}^2 \dots a_2^2)}{k_{n+1} \dots k_2} \int_{\mathbb{R}} Q_0 d\sigma = \frac{k_1(a_{n+1}^2 \dots a_2^2)}{k_{n+1}} \sigma(\mathbb{R});$$

The asymptotic formula for C_σ shows that $\sigma(\mathbb{R}) = a_1^2/k_1$. □

By Theorems 25 and 26, the matrix

$$\begin{pmatrix} \sqrt{k_1 c_1} & \sqrt{k_1 c_2} & \dots & \sqrt{k_1 c_n} \\ \sqrt{k_2 c_1} Q_1(x_1) & \sqrt{k_2 c_2} Q_1(x_2) & \dots & \sqrt{k_2 c_n} Q_1(x_n) \\ \dots & \dots & \dots & \dots \\ \sqrt{k_n c_1} Q_{n-1}(x_1) & \sqrt{k_n c_2} Q_{n-1}(x_2) & \dots & \sqrt{k_n c_n} Q_{n-1}(x_n) \end{pmatrix} \tag{92}$$

is orthogonal.

12. Back to Brouncker’s Polynomials

Using (13), we can easily list the first Brouncker’s polynomials

$$\begin{aligned} P_{-1}(s) &= 1, \\ P_0(s) &= s, \\ P_1(s) &= 2s^2 + 1, \\ P_2(s) &= 4s^3 + 11s, \\ P_3(s) &= 8s^4 + 72s^2 + 25, \\ P_4(s) &= 16s^5 + 340s^3 + 589s, \\ P_5(s) &= 32s^6 + 1328s^4 + 3410s^2 + 2025. \end{aligned}$$

It is clear that $\deg(P_n) = n + 1$ and $P_n(s) = 2^n s^{n+1} + \dots$. Putting $s = 0$ in (13), we obtain

$$P_n(0) = \begin{cases} 0 & \text{if } n \text{ is even,} \\ (2n - 1)^2(2n - 5)^2(2n - 9)^2 \dots & \text{if } n \text{ is odd.} \end{cases}$$

For instance, $P_5(0) = 9^2 \cdot 5^2 \cdot 1^2 = 2025$. For odd n , (14) implies

$$\begin{aligned} (2n + 1)P_n(0) &= (2n + 1)(2n - 1)(2n - 1)(2n - 5)(2n - 5) \dots > \\ &> (2n + 1)(2n - 1)(2n - 3)(2n - 5)(2n - 7) \dots = (2n + 1)!! = P_n(1). \end{aligned}$$

LEMMA 27. *Brouncker’s continued fraction (4) converges in the right half-plane $\Re s > 0$ to a holomorphic function having a positive real part.*

PROOF. Let $s_0(w) = s + w$, $s_k(w) = (2k - 1)^2 / (2s + w)$. Then if $\Re s > 0$, then each s_k maps $\Re w \geq 0$ into $\Re w > 0$. Hence the same is true for

$$\frac{P_n(s)}{Q_n(s)} = s_0 \circ s_1 \circ \dots \circ s_n(0),$$

which implies that the family of holomorphic functions P_n/Q_n is normal in $\Re s > 0$. But these convergents converge to $b(s)$ for positive s . Hence they converge for $\Re s > 0$ by the uniqueness theorem for holomorphic functions. \square

COROLLARY 28. *Identity (17) holds for $\Re s > 0$ and the infinite product in (17) has positive real part for $\Re s > 0$.*

Stirling’s formula (26) in the classical form

$$\Gamma(z) \sim \sqrt{2\pi} z^{z-1/2} e^{-z}, \quad |z| \rightarrow +\infty \tag{93}$$

is valid in any angular domain $|\arg(z)| < \pi - \delta, \delta > 0$; see [1, Cor. 1.4.3]. Simple calculations using Ramanujan’s formula (25) show that in any such domain,

$$b(z) \sim z, |z| \rightarrow +\infty. \tag{94}$$

THEOREM 29. For $\Re s > 0$,

$$\frac{1}{b(s)} = \int_{\mathbb{R}} \frac{d\mu(t)}{s - it}, \quad d\mu = \frac{1}{8\pi^3} \left| \Gamma\left(\frac{1+it}{4}\right) \right|^4 dt,$$

where μ is a probability measure on \mathbb{R} .

PROOF. A well-known formula for the Gamma function [24, §12.14]

$$\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin \pi s}$$

with $s = 1/4 - it/4$ shows that

$$\Gamma\left(\frac{3+it}{4}\right) = \Gamma(1-s) = \frac{\pi}{\sin \pi s} \frac{1}{\Gamma(s)}$$

and therefore,

$$\frac{1}{4} \left[\frac{\Gamma\left(\frac{1+it}{4}\right)}{\Gamma\left(\frac{3+it}{4}\right)} \right]^2 = \frac{1 - \cos 2\pi s}{8\pi^2} |\Gamma(s)|^4 = \frac{1}{8\pi^2} |\Gamma(s)|^4 \left(1 - i \sinh \frac{\pi t}{2}\right).$$

It follows from (25) that

$$\Re\left(\frac{1}{b(it)}\right) = \frac{1}{8\pi^2} \left| \Gamma\left(\frac{1+it}{4}\right) \right|^4 > 0.$$

By (94), we can apply Cauchy’s integral formula for the holomorphic function $1/b(s)$ to the right half-plane $\Re s > 0$:

$$\begin{aligned} \frac{1}{b(s)} &= \frac{1}{2\pi} \int_{+\infty}^{-\infty} \frac{1}{b(it)} \frac{1}{it-s} d(it) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{1}{b(it)} \frac{1}{s-it} dt \\ 0 &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{1}{b(it)} \frac{1}{it+s} dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{1}{b(it)} \frac{1}{s-it} dt. \end{aligned}$$

Adding these two formulas, we complete the proof of the formula for $1/b(s)$. Since $1/b(s) = 1/s + o(1/s^2)$ as $s \rightarrow +\infty$, we obtain that

$$\frac{1}{8\pi^3} \int_{-\infty}^{+\infty} \left| \Gamma\left(\frac{1+it}{4}\right) \right|^4 dt = 1 \tag{95}$$

and hence $d\mu$ is a probability measure on \mathbb{R} . □

Notice that by (93), one can easily obtain an asymptotic formula for the density of $d\mu$ (see [1, Corollary 1.4.4]):

$$\frac{1}{8\pi^3} \left| \Gamma\left(\frac{1+it}{4}\right) \right|^4 \sim \frac{2}{|t|\pi} \exp\left(-\frac{\pi}{2}|t|\right), \quad t \rightarrow \pm\infty. \tag{96}$$

COROLLARY 30. For $\Im z > 0$,

$$\frac{1}{8\pi^3} \int_{-\infty}^{+\infty} \left| \Gamma\left(\frac{1+it}{4}\right) \right|^4 \frac{dt}{z-t} = \frac{1}{z} - \frac{1^2}{2z} - \frac{3^2}{2z} - \frac{5^2}{2z} - \dots \tag{97}$$

PROOF. If $\Im z > 0$ and $s = z/i$, then $\Re s > 0$ and

$$\frac{1}{b(s)} = \int_{\mathbb{R}} \frac{d\mu(t)}{s - it} = i \int_{\mathbb{R}} \frac{d\mu(t)}{z + t} = i \int_{\mathbb{R}} \frac{d\mu(t)}{z - t},$$

since μ is symmetric. On the other hand, continued fraction (4) can be easily transformed into (80) by the substitution $z = is$:

$$\frac{1}{b(s)} = \frac{i}{is} + \frac{i1^2}{s} + \dots = \frac{i}{z - z} + \frac{i3^2}{s} + \dots = \frac{i}{z - z} - \frac{3^2}{z} - \frac{5^2}{z} - \dots$$

which implies (97). □

The partial denominators U_n of the convergents to the continued fraction in (97) satisfy the recurrence relation

$$U_{n+1}(z) = 2zU_n(z) - (2n - 1)^2U_{n-1}(z), \quad U_0(z) = 1, \quad U_1(z) = z.$$

Observing that $z = is$ and using the recurrence relations for P_n and U_n , we see that these polynomials are related by

$$U_n(z) = i^n P_{n-1}(s). \tag{98}$$

By Theorem 21, all roots of polynomials U_n are real. Hence, all roots of Brouncker’s polynomials are placed on the imaginary axis. These roots taken in their totality make a barrier for the analytic continuation of Brouncker’s continued fraction from the right half-plane to the left half-plane as a continued fraction. However, as Theorem 4 shows $b(s)$ still extends analytically through the imaginary axis by the formula of Wallis’ infinite product.

By Theorem 25, the polynomials U_n are orthogonal with respect to μ . By (98) this implies that $P_n(it)$ are orthogonal polynomials. By (91)

$$\int_{\mathbb{R}} U_n^2 d\mu = \frac{1}{2} (2n - 1)!!^2, \quad n = 1, 2, \dots$$

It follows that the system

$$V_n(s) = \frac{\sqrt{2}}{(2n - 1)!!} U_n(z), \quad n = 1, 2, \dots, \quad V_0(z) \equiv 1$$

is an orthonormal system in $L^2(d\mu)$. The operator of multiplication by z acts on V_n by the formulas

$$zV_n(z) = \left(n + \frac{1}{2}\right) V_{n+1} + \left(n - \frac{1}{2}\right) V_{n-1}, \quad n = 1, 2, \dots, \quad zV_0(z) = \frac{1}{\sqrt{2}} V_1(z).$$

Hence, the Jacobi matrix of this operator is

$$\begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & 0 & 0 & 0 & \dots & \dots & 0 & \dots \\ \frac{1}{\sqrt{2}} & 0 & \frac{3}{2} & 0 & 0 & \dots & \dots & 0 & \dots \\ 0 & \frac{3}{2} & 0 & \frac{5}{2} & 0 & \dots & \dots & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & n - \frac{1}{2} & 0 & n + \frac{1}{2} & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

Let us put $z = iy$, $y \rightarrow +\infty$ in (97). Then

$$\frac{y}{b(y)} \sim \sum_{k=0}^{\infty} (-1)^k \frac{1}{y^{2k}} \int_{\mathbb{R}} t^{2k} d\mu, \quad y \rightarrow +\infty. \tag{99}$$

Comparing this formula with (11), we obtain that

$$\int t^2 d\mu = \frac{1}{2}, \quad \int t^4 d\mu = \frac{11}{8}, \quad \int t^6 d\mu = \frac{173}{16}.$$

Further moments can be evaluated by (43). Since by (96),

$$\int_{\mathbb{R}} \frac{\log(d\mu/dx) dx}{1+x^2} = -\infty,$$

polynomials are complete in $L^2(d\mu)$.

It took 160–200 years after Brouncker's discovery until the relationship between orthogonal polynomials and continued fractions was clarified, and it took another 122 years until Wilson constructed Wilson's polynomials [25], of which Brouncker's polynomials turned out to be an important partial case corresponding to the choice of the parameters $a = 0$, $b = \frac{1}{2}$, $c = d = \frac{1}{4}$. (See [1, p. 152] for the definition of Wilson's polynomials.)

References

- [1] Andrews G. E., Askey R., Roy R.: Special functions, in *Encyclopedia of Mathematics and its Applications* 71. Cambridge University Press, Cambridge, 1999.
- [2] Chebyshev P.: O nepreryvnykh drobyakh, *Uch. Zap. Imp. Acad. Sci.*, III, (1855), 636–664; French translation: Sur les fractions continues, *Journ. de math. pures et appl.*, II, série III (1858), 289–323.
- [3] Euler L.: De fractionibus continuis observationes (Observations on continued fractions), presented on January 22, 1739. Originally published in *Commentarii academiae scientiarum Petropolitanae* 11 (1750), pp. 32–81; Opera Omnia: Series 1, 14, pp. 291–349.
- [4] Euler L.: De seriebus divergentibus (On divergent series), presented on March 12, 1753. Originally published in *Novi Commentarii academiae scientiarum Petropolitanae* 5 (1760), pp. 205–237; Opera Omnia: Series 1, 14, pp. 585–617.
- [5] Euler L.: Institutioium calculi integralis volumen secundum, St. Petersburg (1763), E366 = Opera Omnia: Series 1, Volume 12.; Russian translation: *Integral Calculus*, Leningrad-Moscow, GITTL, 1957.
- [6] Euler L.: Problema algebraicum ob affectiones prorsus singulares memorabile (An algebraic problem that is notable for some quite extraordinary relations), presented on March 5, 1770. Originally published in *Novi Commentarii academiae scientiarum Petropolitanae* 15 (1771), pp. 75–106; Opera Omnia: Series 1, 6, pp. 287–315.
- [7] Gauss C. F.: Disquisitiones generales circa seriem infinitam, *Commentationes societas regiae scientiarum Gottingensis recentioris*, II, 1813; in: *Werke*, 3, Göttingen, 1876, pp. 123–162.
- [8] Gauss C. F.: Methodus nova integralium valores per approximationem inveniendi, *Commentationes societas regiae scientiarum Gottingensis recentioris*, III, 1816; in: *Werke*, 3, Göttingen, 1876, pp. 165–196.
- [9] Jacobi C. G. J.: Ueber Gauss' neue Methode, die werthe der Integrale näherungsweise zu finden, *J. Reine Angew. Math.* 1 (1826), 301–308.
- [10] Jacobi C. G. J.: Untersuchungen über die Differentialgleichung der hypergeometrischen Reihe, *J. Reine Angew. Math.* 56 (1859), 149–165.
- [11] Jacobi C. G. J.: Bemerkungen zu einer Abhandlung Euler's über die orthogonale Substitution, in: *Gesammelte Werke*, 3 (1884), pp. 599–609.
- [12] Khovanskii A. N.: *The Application of Continued Fractions and Their Generalizations to Some Problems in Approximation Theory*, GIITL, Moscow, 1956; English translation by P. Wynn, P. Noordhoff, Groningen, 1963.
- [13] Khrushchev S. V.: A recovery of Brouncker's proof for the quadrature continued fraction, *Publ. Mat.* 50 (2006), 3–42.
- [14] Khrushchev S. V.: On Euler's differential methods for continued fractions, *Electronic Transactions on Numerical Analysis* (<http://etna.mcs.kent.edu>) 25 (2006), Special Volume on Constructive Function Theory.

- [15] Kramar F. D.: Integration methods of J. Wallis, (Russian) *Istor.-Mat. Issled.* **14** (1961) 11–100.
- [16] Legendre A. M.: Sur l'attraction des Sphéroïdes. *Mém. Math. et Phys. présentés à l'Ac. r. des. sc. par divers savants.* **10** (1785).
- [17] Pringsheim A.: Über ein Convergenz-Kriterium für die Kettenbrüche mit positiven Gleidern, *Sitzungsber. der math.-phys Klasse der Kgl. Bayer. Akad. Wiss., München* **29** (1899), 261–268.
- [18] Stieltjes T. J.: Recherches sur les fractions continues, *Ann. Faculté Sci. Toulouse* **8** (1894–1895); J76–J122; *ibid.* **9**, A5–A47.
- [19] Sturm C. F.: Mémoire sur la résolution des équations numériques, *Bull. des sciences de Férussac* **11** (1829), 419–422.
- [20] Szegö G.: *Orthogonal Polynomials*, Amer. Math. Soc. Colloq. Publ., **23**. American Mathematical Society, Providence, R.I., 1939; 3rd edition, 1967.
- [21] Vilenkin N. Ya.: *Special Functions and the Theory of Groups Representations*, 2nd edition. Nauka, Moscow, 1991.
- [22] Wallis J.: *Arithmetica Infinitorum*, Oxford, 1656. English translation: Stedall J. A.: *John Wallis: The Arithmetic of Infinitesimals*. Sources and Studies in the History of Mathematics and Physical Sciences. Springer-Verlag, New York, 2004.
- [23] Watson G. N.: Asymptotic expansions of hypergeometric series, *Trans. Cambridge Phil. Soc.*, **22** (1918), 277–308.
- [24] Whittaker E. T., Watson G. N.: *A Course of Modern Analysis*, 4th edition. Cambridge University Press, Cambridge, 2000.
- [25] Wilson J. A.: Three-term contiguous relations and some new orthogonal polynomials, in *Padé and Rational Approximations* (Proc. Internat. Sympos., Univ. South Florida, Tampa, Fla., 1976), E. B. Saff and R. S. Varga, eds., pp. 227–232. Academic Press, New York, 1977.

ATILIM UNIVERSITY, DEPARTMENT OF MATHEMATICS, 06836 INCEK, ANKARA, TURKEY
E-mail address: svk_49@yahoo.com

Spectral Theory via Sum Rules

Rowan Killip

Happy Birthday, Barry Simon

ABSTRACT. We survey some results in the spectral theory of certain one-dimensional differential and finite-difference operators: Jacobi matrices, Krein systems, Schrödinger operators and CMV matrices. What ties these results together is the use of sum rules relating the coefficients and the spectral data.

CONTENTS

1. Introduction
 2. Background
 3. Trace Formulae for Jacobi Matrices
 4. Trace Formulae for Other Operators
 5. Point Spectrum
 6. A.C. Spectrum
 7. The Step-by-Step Method
 8. Necessary and Sufficient Conditions
- References

1. Introduction

We survey some results in the spectral theory of certain one-dimensional differential and finite-difference operators. What ties these results together is their use of sum rules relating the coefficients and spectral data. Contemporary mathematicians are perhaps most familiar with these identities in the context of the inverse scattering solution of integrable systems; however, as we will explain, the natural precursor is a formula of Szegő and Verblunsky uncovered in the study of orthogonal polynomials.

2000 *Mathematics Subject Classification.* 34L05, 47E05, 47B39, 47A55.

Key words and phrases. Jacobi matrix, Krein system, CMV matrix, Schrödinger operator, sum rule.

The author was supported in part by NSF grant DMS-0401277 and a Sloan Foundation Fellowship.

We will describe results for a quartet of operators: Jacobi matrices, Krein systems, Schrödinger operators, and CMV matrices. However, in the interests of brevity and clarity, we limit the discussion of proofs to the case of Jacobi matrices.

2. Background

We start at the very beginning:

THEOREM 2.1. *The sum of the diagonal entries of an $n \times n$ matrix is equal to the sum of its eigenvalues (counting algebraic multiplicity).*

This innocuous sounding result is surprisingly deep; perhaps more importantly, it and its descendants can be surprisingly useful. The key point is the following: in general, one cannot hope to determine the eigenvalues of an operator; however, computing the trace is easy and says something potentially useful about the eigenvalues. A good example of the power of this little fact is shown by the following ingenious application due to Avron, van Mouche, and Simon, [1]:

THEOREM 2.2. *Consider the almost Mathieu operator*

$$(H_\theta u)(n) = u(n + 1) + u(n - 1) + \lambda \cos(2\pi n\alpha + \theta)u(n)$$

acting on $\ell^2(\mathbb{Z})$ with $\alpha = p/q$ rational and $\lambda \leq 2$. Then $\sigma_- = \cap_\theta \sigma(H_\theta)$ has Lebesgue measure $4 - 2\lambda$.

Let me outline the proof when q is odd. Due to a remarkable formula of Chambers, [11], it is possible to show that σ_- is the union of q bands; moreover, each band edge corresponds to an eigenfunction of either $H_{\theta=0}$ or $H_{\theta=\pi}$ belonging to a specific symmetry class: periodic/anti-periodic (under translation) and even/odd (under reflection). In this way, one is led to the conclusion that

$$|\sigma_-| = \text{tr}(H_{\theta=\pi}^{p,e}) - \text{tr}(H_{\theta=0}^{a,e}) + \text{tr}(H_{\theta=0}^{a,o}) - \text{tr}(H_{\theta=\pi}^{p,o})$$

where the subscripts indicate the restriction of this operator to the subspace with the corresponding symmetry. This equality comes from regarding the trace as the sum of eigenvalues; the traces are easily evaluated as the sum of diagonal entries, which gives the result. (The result is also true for irrational α ; see [43, 58].)

The proof of Theorem 2.1 is easily found once one pauses to remember the definition of algebraic multiplicity: the number of occurrences on the diagonal in Jordan normal form, or the order of the root in the characteristic polynomial. Of course, one may identify all coefficients of the characteristic polynomial in terms of its roots (the eigenvalues) and in terms of the matrix entries. This leads to

THEOREM 2.3. *For any $n \times n$ matrix, A ,*

$$1 + \sum_{k=1}^n z^k \text{tr}(\wedge^k A) = \det[1 + zA] = 1 + \sum_{k=1}^n z^k \sum_{l_1 > \dots > l_k} \lambda_{l_1} \cdots \lambda_{l_k}$$

where $\lambda_1, \dots, \lambda_n$ denote the eigenvalues of A according to algebraic multiplicity.

REMARK. The matrix elements of $\wedge^k A$ —the restriction of $A \otimes \dots \otimes A$ to antisymmetric tensors—are exactly the $k \times k$ minors of A ; thus

$$\text{tr}(\wedge^k A) = \sum_{l_1 > \dots > l_k} A(l_1, \dots, l_k) = \frac{1}{k!} \sum_{l_1, \dots, l_k} A(l_1, \dots, l_k). \tag{1}$$

Our notation for the minors is as follows: the upper list of indices gives the rows used for the minor and the lower, the columns. The second equality comes from summing over permutations of the indices and noting that the (minor) determinant vanishes if two indices coincide.

PROOF. The right-hand equality comes from expanding $\prod(1+z\lambda_l)$. These formulae for the coefficients of a polynomial in terms of its roots are usually attributed to François Viète, a 16th century French lawyer and mathematician.

The determinant is multi-linear in the columns, thus one may expand $\det(1+zA)$ in much the same way as the product in the previous paragraph. A few column operations are all that is required to finish the proof. \square

It is natural to extend this theorem to Banach spaces; it is here that one realizes that things are not so simple after all. An operator on a Banach space E is called nuclear if it can be written as $\sum e_j \langle l_j, \cdot \rangle$ for sequences $e_j \in E$ and $l_j \in E^*$ with $\sum \|e_j\| \|l_j\| < \infty$. The big surprise is that the eigenvalues of nuclear operators are only guaranteed to be absolutely summable if E is isomorphic to a Hilbert space, [44].

In the Hilbert space setting, the space of nuclear operators is better known as trace class, \mathcal{J}_1 , and the more usual definition is as those compact operators, A , whose singular values are summable. (Recall that the singular values are the eigenvalues of $(A^*A)^{1/2}$.) Here, the sum of the moduli of the eigenvalues is finite; indeed it is bounded by the sum of the singular values. A very general and elegant proof of this fact can be found in [109]; [87] contains three further proofs and historical references.

A second obstruction to the extension of Theorem 2.1 to Banach spaces is more devastating: there is a nuclear operator A on ℓ^1 with $\text{tr}(A) = 1$ and $A^2 = 0$. A textbook presentation of this example can be found in [62, §2.d].

THEOREM 2.4. *Let A be a trace class operator on a Hilbert space. For any orthonormal basis, $\{\phi_j\}$, $\sum \langle \phi_j | A \phi_j \rangle$ is equal to the sum of the eigenvalues repeated according to algebraic multiplicity.*

For a compact operator, the algebraic multiplicity of a non-zero eigenvalue λ can be defined as the rank of $\oint_\gamma (z - A)^{-1} dz$ where γ is a small circle around λ excluding the remainder of the spectrum of A . It is not necessary to assign a multiplicity to $\lambda = 0$ as such eigenvalues do not contribute to the sum.

Theorem 2.4 is widely known as Lidskiĭ's theorem, [60]. As pointed out by Pisier, [70], the statement can be found earlier in [40, §4]. Neither paper gives much detail; thorough treatments can be found in several textbooks: [37, 59, 88]. The majority of proofs of this theorem go one step further and treat the determinant:

THEOREM 2.5. *Let A be a trace class operator on a Hilbert space, then*

$$\det(1+zA) := 1 + \sum_{k=1}^{\infty} z^k \text{tr}(\wedge^k A) = \prod(1+z\lambda_l). \quad (2)$$

PROOF. The sum converges to an entire function for any nuclear operator on a Banach space; indeed by applying the Hadamard inequality in (1) one has $\text{tr}(\wedge^k A) = O(C^k k^{-k/2})$. Moreover, by finite-rank approximations, its zeros are $z_l = -\lambda_l^{-1}$ where λ_l are the eigenvalues of A with appropriate multiplicities.

The restriction to Hilbert space has two effects: Firstly, $\sum |z_l|^{-1} < \infty$ and so $\det(1 + zA) = e^{g(z)} \prod (1 + z\lambda_l)$ for some entire function g . Secondly, one can improve the bound on $\text{tr}(\wedge^k A)$ to $O(\epsilon^k/k!)$ for any $\epsilon > 0$. In this way, we obtain $|\det(1 + zA)| \lesssim \exp(\epsilon|z|)$ for every $\epsilon > 0$. With further effort, one can then deduce $\text{Re } g(z) \leq 2\epsilon|z| + C_\epsilon$, which implies that $|g(z)| = o(|z|)$ by the Borel–Carathéodory inequality. Thus g is constant, and taking $z = 0$ shows $g(z) \equiv 0$. \square

The proof outlined above matches the scant details in [40]. It seems to be the simplest approach and was independently discovered by Barry, [87]. Related ideas were used by Carleman, [9], to prove a product representation for the (regularized) determinant of integral operators with Hilbert–Schmidt kernels.

Every entire function has a product representation; the product over the zeros can be made to converge by adding exponential factors. This is a famous idea of Weierstrass. Implementing it in the context of determinants leads to the notion of regularized determinants:

THEOREM 2.6. *Given an integer $p \geq 1$, let $G(x) = \sum_{k=1}^{p-1} x^k/k$ then*

$$\det_p(1 - zA) := \det(1 - zA)e^{\text{tr } G(zA)} = \prod (1 - z\lambda_l)e^{G(z\lambda_l)} = \det[(1 - zA)e^{G(zA)}]$$

extends from trace class to a continuous function on \mathcal{I}_p , the space of operators whose singular values are ℓ^p .

The notion of regularized determinants entered mathematics incrementally, beginning in the early twentieth century; see [87, §6] for a discussion of the history. One approach to the theory (introduced by Seiler, [86]) is to notice that $(1 - x)e^{G(x)}$ can be written as $1 - x^p f(x)$ for some entire function f .

Undoubtedly, the most famous work on infinite determinants is that of Fredholm concerning integral equations, [32]. In addition to constructing the determinant, he obtains a series expansion for the kernel of the inverse operator by analogy with Cramer’s rule. We will use the discrete analogue of these formulae:

THEOREM 2.7. *Suppose $A \in \mathcal{I}_1$, then $1 + A$ is invertible if and only if*

$$\det(1 + A) = 1 + \sum_{k=1}^{\infty} \sum_{l_1 > \dots > l_k} A(l_1, \dots, l_k) \tag{3}$$

is non-vanishing. The inverse can then be written as $1 - B$ where

$$B(n, m) = \frac{1}{\det(1 + A)} \sum_{k=0}^{\infty} \sum_{l_1 > \dots > l_k} A\left(\begin{matrix} n, l_1, \dots, l_k \\ m, l_1, \dots, l_k \end{matrix}\right). \tag{4}$$

PROOF. The first sentence follows from (1) and Theorem 2.5. Note that $1 + A$ is invertible if and only if -1 is not an eigenvalue of A .

The second statement follows by direct computation. For more detail, see [59, Ch. 24] or [88, Ch. 5]. \square

This theorem can be extended to operators in the higher trace ideals, \mathcal{I}_p . Specifically, $\det_p(1 + zA)$ can be written as a sum of modified minors: when calculating these determinants as a sum over the symmetric group, one must omit any permutation containing a cycle of length less than p . The modification to (4) is not so simply explained; see [88].

Later, we will use the following consequence of the Fredholm formulae. While I have not seen this precise statement elsewhere, results of this type are well known.

LEMMA 2.8. Let G be a bounded operator with semi-separable kernel, that is,

$$G(n, m) = \begin{cases} f(n)g(m) & : n \geq m, \\ f(m)g(n) & : n \leq m. \end{cases}$$

Suppose K is a finite rank operator with $K(n, m) \neq 0$ only when $m \leq n < N$ for some integer N , then

$$\tilde{f} = [1 + GK]^{-1}f \text{ obeys } \tilde{f}(n) = a^{-1}f(n) \text{ for } n > N$$

where $a = \det(1 + GK)$.

PROOF. We will give the main computation and then justify the steps. Writing A for GK and using the Fredholm formulae from Theorem 2.7, we see that for n sufficiently large,

$$\begin{aligned} \tilde{f}(n) &= f(n) - a^{-1} \sum_{k=0}^{\infty} \sum_{l_1 > \dots > l_k} \sum_m A(\begin{smallmatrix} n, l_1, \dots, l_k \\ m, l_1, \dots, l_k \end{smallmatrix}) f(m) \\ &= f(n) - a^{-1} \sum_{k=0}^{\infty} \sum_{m > l_1 > \dots > l_k} A(\begin{smallmatrix} n, l_1, \dots, l_k \\ m, l_1, \dots, l_k \end{smallmatrix}) f(m) \\ &= f(n) - a^{-1} \sum_{k=0}^{\infty} \sum_{m > l_1 > \dots > l_k} A(\begin{smallmatrix} m, l_1, \dots, l_k \\ m, l_1, \dots, l_k \end{smallmatrix}) f(n) \\ &= f(n)[1 - a^{-1}(a - 1)], \end{aligned}$$

which simplifies to $a^{-1}f(n)$. The second line follows by noting that if $m \leq l_1$ then the top two rows of the minor are linearly dependent; this uses the fact that K is upper triangular and G is semi-separable. For the same reasons, $f(m)A(n, p) = f(n)A(m, p)$ whenever $p \leq \max\{n, m\}$. This justifies the third equality. The last line follows by recognizing the determinant from (3). \square

3. Trace Formulae for Jacobi Matrices

In this section, we will present *a priori* sum rules for Jacobi matrices.

Given two sequences $a_n > 0$ and $b_n \in \mathbb{R}$ indexed over $n = 1, 2, \dots$, the associated Jacobi matrix is the tri-diagonal matrix with these sequences as entries:

$$J = \begin{bmatrix} b_1 & a_1 & & & \\ a_1 & b_2 & a_2 & & \\ & a_2 & b_3 & \ddots & \\ & & \ddots & \ddots & \ddots \end{bmatrix}. \tag{5}$$

This defines a bounded self-adjoint operator if and only if the sequences are bounded. When they are unbounded, the operator may or may not be essentially self-adjoint when defined on finite sequences; see [91] for a discussion of this and its significance for the moment problem.

Given a pair of Jacobi matrices, \tilde{J} and J , that differ at only finitely many entries, we can define the perturbation determinant:

$$a(z) = \det \left[\frac{\tilde{J}-z}{J-z} \right] = \det[1 + G(\tilde{J} - J)] \text{ with } G = G(z) = (J - z)^{-1}, \tag{6}$$

which is an analytic function off $\sigma(J)$ and meromorphic for $z \notin \sigma_{\text{ess}}(J)$. We will derive trace formulae by studying the behaviour of $\log |a(z)|$ at infinity and in a neighbourhood of the spectrum. This is simpler if we assume that J is a bounded operator, which we do henceforth.

The behaviour near infinity is the easiest to describe: for $|z|$ sufficiently large,

$$\begin{aligned} \log[a(z)] &= \log \det \left[\frac{1-z^{-1}\tilde{J}}{1-z^{-1}J} \right] = \text{tr}(\log[1 - z^{-1}\tilde{J}] - \log[1 - z^{-1}J]) \\ &= - \sum_{k=1}^{\infty} \frac{1}{k} z^{-k} \text{tr}(\tilde{J}^k - J^k). \end{aligned} \tag{7}$$

Taking the real part gives the asymptotics of $\log |a(z)|$.

Understanding the behaviour of $\log |a(z)|$ near the spectrum is considerably more involved and will require a number of preliminaries.

The vector $e_1 = [1, 0, \dots]^\dagger$ is cyclic for J ; we will write $d\mu$ for the corresponding spectral measure. Because of the existence of a cyclic vector, all eigenspaces are one-dimensional and hence all zeros and poles of $a(z)$ are simple. Given a concrete Jacobi matrix, the natural way to determine $d\mu$ is via the m -function:

$$m(z) = \langle e_1 | (J - z)^{-1} e_1 \rangle = \int \frac{d\mu(t)}{t - z}. \tag{8}$$

This requires knowledge of the Green function, which we will now describe.

The Green function is constructed from two solutions of the finite difference equation associated to J . Let us define polynomials $p_n(z)$ of degree $n \geq 0$ by the recurrence

$$a_{n+1}p_{n+1}(z) + b_n p_n(z) + a_{n-1}p_{n-1}(z) = z p_n(z) \tag{9}$$

with $p_{-1}(z) \equiv 0$ and $p_0 \equiv 1$. Then $\sum_{n=0}^{\infty} p_n(z)e_{n+1}$ is a formal solution of $Ju = zu$. Moreover, these polynomials form an orthonormal basis for $L^2(d\mu)$ and $p_n(J)e_1 = e_{n+1}$. The second solution is the Weyl solution, which is only defined for $z \notin \sigma(J)$. It is given by

$$\sum_{n=0}^{\infty} \psi_n(z)e_{n+1} = (J - z)^{-1} e_1. \tag{10}$$

In practice, one uses the following equivalent formulation: $\psi_n(z)$ obeys (9) for $n \geq 1$, it is square summable, and $a_1\psi_1 + (b_1 - z)\psi_0 = 1$. It is now easy to check that the Green function is given by

$$G(n + 1, n' + 1; z) = \langle e_{n+1} | (J - z)^{-1} e_{n'+1} \rangle = p_{\min\{n, n'\}}(z) \psi_{\max\{n, n'\}}(z).$$

The point in introducing all this machinery is the following discrete analogue of a theorem of Jost and Pais, [45]:

THEOREM 3.1. *Suppose \tilde{J} and J are bounded Jacobi matrices differing in only finitely many entries, then for n sufficiently large, the Weyl solutions associated to these Jacobi matrices obey*

$$a(z) \tilde{\psi}_n(z) = \psi_n(z) \prod_{k=1}^{\infty} \frac{\tilde{a}_k}{a_k}. \tag{11}$$

In particular, for a.e. x with respect to $d\mu_{ac}$,

$$\frac{d\tilde{\mu}}{d\mu} = \frac{\text{Im } \tilde{m}(x + i0)}{\text{Im } m(x + i0)} = |a(x + i0)|^{-2} \prod_{k=1}^{\infty} \frac{\tilde{a}_k^2}{a_k^2}. \tag{12}$$

PROOF. Let D be the diagonal matrix with $D_{nn} = \prod_{k=n}^\infty \tilde{a}_k/a_k$. By the resolvent identity,

$$D(\tilde{J} - z)^{-1}D^{-1} = (J - z)^{-1} - (J - z)^{-1}[D^{-1}\tilde{J}D - J]D(\tilde{J} - z)^{-1}D^{-1},$$

which implies that the sequence $\xi_n := (D_{nn}/D_{11})\tilde{\psi}_n$ obeys

$$\xi_n = \psi_n - GK\xi_n$$

where G is the Green function for J and K is the matrix $D^{-1}\tilde{J}D - J$, which is lower triangular. As D_{nn} is eventually identically one and

$$\det[1 + GK] = \det[D^{-1}(\tilde{J} - z)D(J - z)^{-1}] = a(z),$$

Lemma 2.8 implies that (11) holds.

To prove (12) we merely combine (11), the fact that

$$\frac{d\mu}{dx} = \lim_{y \downarrow 0} \frac{1}{\pi} \operatorname{Im} m(x + iy) = \lim_{y \downarrow 0} \frac{y}{\pi} \|(J - x - iy)^{-1}e_1\|^2 = \lim_{y \downarrow 0} \frac{y}{\pi} \|\psi_n(x + iy)\|^2$$

Lebesgue almost everywhere, and the corresponding result for \tilde{J} . □

With a little more care, one may use the reasoning above to deduce that $\log |a(z)|$ has a non-tangential limit at $d\mu_{ac}$ -a.e. $x \in \mathbb{R}$. Actually, it is easy to obtain this kind of information about $a(z)$:

LEMMA 3.2. *If $\tilde{J} - J$ has finite rank, then $a(z)$ is a polynomial in z and $m(z)$.*

PROOF. The key observation is that every matrix element of $(J - z)^{-1}(\tilde{J} - J)$ is a polynomial in $m(z)$ and z . This can be justified by noting that

$$\langle e_{n+1} | (J - z)^{-1} e_{m+1} \rangle = \int \frac{p_n(x)p_m(x) d\mu(x)}{x - z}$$

and $x^k(x - z)^{-1} = x^k + zx^{k-1}(x - z)^{-1}$. □

The reader is no doubt familiar with function theory in the unit disk and hence in any simply connected domain. In multiply-connected domains, matters are more complicated, primarily because of the non-existence of Blaschke products. Of course, one may always lift questions to the universal cover and apply results from the disk case, but in general, the covering map can be a horror. We do not wish to get waylaid by these problems and so treat a very simple case.

HYPOTHESIS 3.3. *We assume that J is periodic, that is, the sequences a_n and b_n are periodic.*

Under this hypothesis, $\sigma(J)$ consists of finitely many compact intervals together with finitely many points. (This remains true for finite-rank perturbations.)

Let us write Ω for the complement of $\sigma_{\text{ess}}(J)$ in the Riemann sphere. By applying Joukowski transformations, this region can be transformed to one bounded by finitely many analytic curves; thus we can apply the general results described in [31, 77].

The trace formulae we will derive amount to the relation between $\log |a(z)|$ on $\sigma_{\text{ess}}(J)$ as given in Theorem 3.1 and the asymptotics given in (7). In essence, $\log |a(z)|$ is the Poisson integral of its boundary values; however $a(z)$ may have both

zeros and poles. When Ω is simply connected, the traditional approach has been to remove the problem with Blaschke products. We use Green’s identity:

$$\int_{\Omega} f \Delta g - g \Delta f = \int_{\partial\Omega} f \frac{\partial g}{\partial n} - g \frac{\partial f}{\partial n}. \tag{13}$$

(The Poisson integral representation of harmonic functions follows by choosing g to be the Dirichlet Green function for Ω .)

DEFINITION 3.4. Let $g_0(z)$ denote the (Dirichlet) Green function for Ω with singularity at infinity. That is, g_0 is the unique continuous function on \mathbb{C} that is harmonic on $\mathbb{C} \setminus \sigma_{\text{ess}}(J)$, vanishes on $\sigma_{\text{ess}}(J)$, and has asymptotics

$$g_0(z) = -\frac{1}{2\pi} \log |z| + O(1) \quad \text{as } z \rightarrow \infty.$$

Similarly we introduce functions $g_k, k \geq 1$, that are continuous and harmonic as before, but with asymptotics

$$g_k(z) = -\frac{1}{2\pi} \operatorname{Re} z^k + O(1) \quad \text{as } z \rightarrow \infty.$$

We will also use the analogue of harmonic measure:

$$d\nu_k(x) = -2 \left[\lim_{y \downarrow 0} \frac{\partial}{\partial y} g_k(x + iy) \right] dx,$$

which is supported on $\sigma_{\text{ess}}(J)$.

Note that Green’s identity with $f \equiv 1$ shows that $\int d\nu_k = \delta_{k0}$. These functions are Green functions for infinity in the following sense: if f is smooth and supported in a small neighbourhood of infinity with

$$f(z) \sim \sum_l c_l \operatorname{Re}(z^{-l}) + d_l \operatorname{Im}(z^{-l}), \quad \text{then} \quad \int f \Delta g_k = \begin{cases} kc_k & : k \geq 1, \\ c_0 & : k = 0. \end{cases} \tag{14}$$

THEOREM 3.5. Suppose J is a periodic Jacobi matrix and \tilde{J} is a finite-rank perturbation, then

$$-\sum \log[\tilde{a}_j/a_j] = 2\pi \sum [g_0(\tilde{E}_j) - g_0(E_j)] - \frac{1}{2} \int \log \left[\frac{d\tilde{\mu}}{d\mu} \right] d\nu_0(x), \tag{15}$$

where E_j and \tilde{E}_j enumerate the discrete spectrum of J and \tilde{J} . For $k \geq 1$,

$$-\operatorname{tr}(\tilde{J}^k - J^k) = 2\pi \sum [g_k(\tilde{E}_j) - g_k(E_j)] - \frac{1}{2} \int \log \left[\frac{d\tilde{\mu}}{d\mu} \right] d\nu_k(x). \tag{16}$$

PROOF. The equations follow from Green’s identity, (13), with $f = \log |a(z)|$ and $g = g_k$. As $a(z)$ has simple poles/zeros at the eigenvalues of J and \tilde{J} ,

$$-\Delta f = \sum 2\pi \delta(z - E_j) - \sum 2\pi \delta(z - \tilde{E}_j),$$

while $\int_{\Omega} f \Delta g_k$ can be evaluated with (7) and (14).

We need to show that the integrals over the boundaries can be taken in an almost-everywhere, rather than distributional, sense. Because we have assumed that J is periodic, its m -function is extremely well behaved and so Lemma 3.2 makes this elementary. In more general settings, one needs to use the fact that $m \in H^p(\Omega)$ for any $0 < p < 1$ and hence $a(z) \in H^p(\Omega)$ for p sufficiently small.

From the definition, $a(\bar{z}) = \overline{a(z)}$ and so $f(z) = f(\bar{z})$. This allows us to combine the contributions from upper and lower edges of each slit. The final result follows by re-writing f on the boundary via (12). Note that $\log[\tilde{a}_j/a_j]$ appears with coefficient one in (15) and not at all in (16) because $\int d\nu_k = \delta_{k0}$. □

These formulae may seem rather far removed from the trace formulae we discussed in Section 2. However, the important property has remained: the left-hand side involves the coefficients of the operator, while the the right, the spectral properties.

The simplest periodic operator has constant coefficients; by scaling and shifting, it suffices to consider $a_k \equiv 1$ and $b_k \equiv 0$. The resulting Jacobi matrix has spectrum $[-2, 2]$, which is purely absolutely continuous,

$$d\mu(x) = \frac{1}{2\pi} \chi_{[-2,2]}(x) \sqrt{4 - x^2} dx.$$

For this choice of J , Theorem 3.5 is due to Case, [10], although he was very much inspired by the trace formulae for Schrödinger operators that we will describe in the next section.

Certain linear combinations of the Case formulae turn out to be more useful for applications; the key ingredient is positivity. The following example synthesizes [47, 54]. Let T_n and U_n denote the usual Chebyshev polynomials:

$$T_n(\cos(\theta)) = \cos(n\theta) \quad \text{and} \quad U_n(\cos(\theta)) = \frac{\sin[(n + 1)\theta]}{\sin(\theta)}.$$

Then for each $n \geq 1$,

$$\begin{aligned} & -\frac{n}{\pi} \int_{-2}^2 \log \left[\frac{d\tilde{\mu}}{d\mu} \right] \sqrt{4 - x^2} |U_{n-1}(\frac{x}{2})|^2 dx + \sum G_n(\tilde{E}_j) \\ & = \text{tr} \left\{ [2T_n(\frac{1}{2}\tilde{J}) - 2T_n(\frac{1}{2}J)]^2 \right\} + 4 \sum_{j=1}^{\infty} F(\tilde{a}_j \cdots \tilde{a}_{j+n-1}) + X_n \end{aligned} \tag{17}$$

where $F(x) = x - 1 - \log(x) \geq 0$,

$$G_n(\beta + \beta^{-1}) = \beta^{2n} - \beta^{-2n} - 4n \log |\beta| \quad \text{for } |\beta| > 1, \tag{18}$$

and X_n is a simple function of the first few entries of \tilde{J} . This is most easily deduced by simply repeating the proof of Theorem 3.5 using the harmonic function $g = \frac{1}{2\pi} \text{Re } G_n$ in place of any particular g_k . One further observation is necessary however: $2T_n(\frac{1}{2}J)$ differs from the matrix with ones on the n th sub- and super-diagonals and zeros elsewhere in only a few entries. This implies

$$-8 \text{tr} \left\{ [T_n(\frac{1}{2}\tilde{J}) - T_n(\frac{1}{2}J)] T_n(\frac{1}{2}J) \right\} = X_n - 4 \sum [\tilde{a}_j \cdots \tilde{a}_{j+n-1} - 1],$$

with the proper choice of X_n ; in fact, $X_1 \equiv 0$.

For future reference, let us note that the right-hand side of this equation is finite if and only if $T_n(\frac{1}{2}\tilde{J}) - T_n(\frac{1}{2}J)$ is Hilbert-Schmidt; the sum over $F(\tilde{a}_j \cdots \tilde{a}_{j+n-1})$ is bounded by the sum of the squares of the entries on the n th super-diagonal of this matrix.

The first two terms on the right-hand side of (17) are manifestly positive, as is the sum over the eigenvalues. Strict positivity of the integral is not essential; however, by Lemma 6.2 it is bounded from below. There are several other sum rules for Jacobi matrices that have good positivity properties; see, for example, [47, 53, 81, 100]. A very general (but rather abstract) approach to the positivity problem can be found in [66].

The observation regarding $2T_n(\frac{1}{2}J)$ has an analogue for general periodic Jacobi matrices, which gives rise to similar formulae. Let J be a periodic Jacobi matrix

with period p scaled so that $a_1 a_2 \cdots a_p = 1$ and let Δ denote the corresponding discriminant. For any \tilde{J} differing from J by finite rank,

$$\begin{aligned}
 2\pi \sum [g(\tilde{E}_j) - g(E_j)] - \frac{1}{2} \int \log \left[\frac{d\tilde{\mu}}{d\mu} \right] d\nu(x) \\
 = \frac{1}{4p} \operatorname{tr} \left\{ [\Delta(\tilde{J}) - \Delta(J)]^2 \right\} + \frac{1}{4p} \sum F(\tilde{a}_j \cdots \tilde{a}_{j+p-1}) + X
 \end{aligned}
 \tag{19}$$

where $g = \frac{-1}{2\pi p} \operatorname{Re} \left\{ \log \left[\frac{1}{2} \Delta + \frac{1}{2} \sqrt{\Delta^2 - 4} \right] - \frac{1}{4} \Delta \sqrt{\Delta^2 - 4} \right\}$ and $d\nu$ is the probability measure supported on $\sigma_{\text{ess}}(J)$ with density $\frac{d\nu}{dx} = \frac{1}{2\pi p} |\Delta'(x)| \sqrt{4 - \Delta(x)^2}$. These formulae and related results are the topic of forthcoming joint work, [14]. Note also that (17) follows from this formula by considering the case of constant coefficients as a period- n problem.

4. Trace Formulae for Other Operators

As mentioned in the previous section, the derivation of sum rules for Jacobi matrices follows earlier results for Schrödinger operators. The first goal of this section is to describe these results. After that we will briefly discuss certain older results that fit naturally into the same framework. As in the previous section, we will state *a priori* versions of these sum rules; that is, with far stronger hypotheses than turn out to be necessary.

Consider the whole-line Schrödinger operator associated to a smooth compactly supported potential V ,

$$[\tilde{L}u](x) = -u''(x) + V(x)u(x),$$

and write L for the free operator ($V \equiv 0$). In this setting, the perturbation determinant $a(z) = \det[(\tilde{L} - z)/(L - z)]$ happens to be equal to the reciprocal of the transmission coefficient and most references we quote take this point of view. The analogue of Theorem 3.5 is much better known, primarily because of its role in the inverse scattering solution of the KdV equation. As $\sigma(L)$ is not compact, one studies the behaviour of $a(z)$ as z approaches infinity in a particular direction, specifically, along the negative real axis.

THEOREM 4.1. *If V is C^∞ and of compact support, then for $n \geq 0$,*

$$\int_0^\infty \log |a(E + i0)| E^{n-1/2} dE = \frac{(-1)^n \pi}{2^{2n+1}} \int \xi_{2n+1}(x) dx + \frac{(-1)^n 2\pi}{2n+1} \sum E_m^{n+1/2} \tag{20}$$

where $E_m < 0$ enumerate the discrete spectrum and ξ_{2n+1} is defined by the following recurrence: $\xi_0(x) = 0$, $\xi_1(x) = V(x)$ and $\xi_{n+1} + \xi'_n + \sum_{\ell=1}^n \xi_\ell \xi_{n-\ell} = 0$.

The original paper is [110], which builds upon earlier work [7, 34, 35]. The reader may have noticed that Jacobi matrices are parameterized over a half-line, while now we discuss whole-line Schrödinger operators. The trace formulae for half-line Schrödinger operators, [7], contain values of V (and its derivatives) at the origin; this makes them unsuitable for the applications we have in mind.

The formulae for $\int \xi_k dx$ can be simplified by recognizing complete derivatives. We will primarily discuss the $n = 1$ case of (20):

$$\frac{1}{\pi} \int_0^\infty \log |a(E + i0)| E^{1/2} dE + \frac{2}{3} \sum |E_m|^{3/2} = \frac{1}{8} \int |V(x)|^2 dx. \tag{21}$$

If V is supported on the positive half-axis, then we obtain the following analogue of the Jost–Pais theorem: for $k > 0$,

$$|a(k^2 + i0)|^2 = \frac{|m(k^2 + i0) + ik|^2}{4k \operatorname{Im} m(k^2 + i0)} \geq 1 \tag{22}$$

where m denotes the Weyl m -function associated to the half-line Schrödinger operator with potential V and a Dirichlet boundary condition. In Theorem 3.1, we made a direct link to the spectral measure; that is not quite possible here. While $\frac{1}{\pi} \operatorname{Im} m(E + i0)$ is equal to the Radon–Nikodym derivative of the spectral measure, the formula for a involves $\operatorname{Re} m$ and hence the Hilbert transform of the spectral measure.

Jacobi matrices are naturally associated to the theory of orthogonal polynomials for measures supported on the real line. There is an analogous theory of orthogonal polynomials for measures on the unit circle in the complex plane. While this theory is of considerable vintage, the proper analogue of Jacobi matrices was discovered surprisingly recently. We will now describe these operators and describe how the corresponding sum rules relate to certain classical questions.

Given a probability measure $d\mu$ on $S^1 = \{z \in \mathbb{C} : |z| = 1\}$ (with infinite support), we can construct a system of orthonormal polynomials $\phi_k(z)$ by applying the Gram–Schmidt procedure to $1, z, z^2, \dots$. As in the Jacobi case, these polynomials obey a recurrence relation. As it is simpler in this case, we write the relation for the monic orthogonal polynomials:

$$\Phi_{k+1}(z) = z\Phi_k(z) - \bar{\alpha}_k \Phi_k^*(z), \quad \Phi_{k+1}^*(z) = \Phi_k^*(z) - \alpha_k z \Phi_k(z). \tag{23}$$

Here $\alpha_k \in \mathbb{D}$ are the recurrence coefficients, which we call Verblunsky coefficients, and Φ_k^* denotes the reversed polynomial: $\Phi_k^*(z) = z^k \Phi_k(\bar{z}^{-1})$.

In general, these polynomials need not form a basis for $L^2(d\mu)$, as can be seen when $d\mu = \frac{1}{2\pi} d\theta$. Instead, we may apply the Gram–Schmidt procedure to $1, z, z^{-1}, z^2, z^{-2}, \dots$; in this way we obtain an orthonormal basis $\chi_k(z)$ for $L^2(d\mu)$, which are related to the orthonormal polynomials by

$$\chi_k(z) = \begin{cases} z^{-k/2} \phi_k^*(z) & : k \text{ even} \\ z^{-(k-1)/2} \phi_k(z) & : k \text{ odd.} \end{cases} \tag{24}$$

Let \mathcal{C} be the matrix representing $f(z) \mapsto zf(z)$ in this basis. The resulting class of matrices are known as CMV matrices and comprise a natural unitary analogue of Jacobi matrices. The name is taken from authors of [8]; however, the original discovery predates this paper as discussed in [95] and [108].

Let us write \mathcal{C}_0 for the CMV matrix associated to $d\mu = \frac{1}{2\pi} d\theta$, which corresponds to $\alpha_k \equiv 0$. The analogue of Theorems 3.1 and 3.5 can be combined into one:

THEOREM 4.2. *Suppose $\mathcal{C} - \mathcal{C}_0$ is of finite rank (that is, $\alpha_k = 0$ for all but finitely many k). Then $d(z) := \det[(1 - z\mathcal{C}^\dagger)/(1 - z\mathcal{C}_0^\dagger)]$ is related to the Szegő function,*

$$D(z) = \exp \left\{ \frac{1}{4\pi} \int \frac{e^{i\theta} + z}{e^{i\theta} - z} \log \left[2\pi \frac{d\mu}{d\theta} \right] d\theta \right\}, \quad \text{by} \quad d(z)D(z) = \prod (1 - |\alpha_k|^2)^{1/2}.$$

Notice that $|D(e^{i\theta})|^2 = 2\pi \frac{d\mu}{d\theta}$.

This result is Theorem 4.2.14 in [94]. By comparing the Taylor coefficients of $d(z)$ and $D(z)$ one easily deduces sum rules resembling (15) and (16).

While the interpretation of $D(z)$ as a perturbation determinant of CMV matrices is very recent, the primary content of this theorem is not. Szegő, [102, 103], proved that when the integral defining D is convergent, $\phi_n^*(z) \rightarrow D(z)^{-1}$ uniformly on compact subsets of \mathbb{D} . By using the recurrence relations for ϕ_n^* , one can deduce the same family of sum rules.

Kreĭn, [52], introduced a continuous analogue of the recurrence (23),

$$\frac{d}{dr}P(r, z) = izP(r, z) - \bar{A}(r)P_*(r, z), \quad \frac{d}{dr}P_*(r, z) = -A(r)P(r, z), \quad (25)$$

where $A(r)$ is a complex function on $[0, \infty)$ and $P_*(0, z) = P(0, z) \equiv 1$. These equations are referred to as the Kreĭn system. Note that $P_*(r, z) = e^{irz}\overline{P(r, \bar{z})}$ and

$$P(r, z) = e^{irz} - \int_0^r \gamma_r(s)e^{i(r-s)z} dz$$

for some integrable function γ_r , which explains the relation of r to the degree of the polynomial. The polynomial analogy is further strengthened by the existence of a measure $d\mu$ on \mathbb{R} so that $\int(1+x^2)^{-1}d\mu(x) < \infty$ and $\int P(r, x)\overline{P(s, x)}d\mu = \delta(r-s)$. While of interest in their own right, results for Kreĭn systems also have consequences for Schrödinger operators; the key observation is that if A is real-valued, then

$$\psi(r; z) = e^{-irz} \frac{P(2r, z) - P_*(2r, \bar{z})}{2iz} \quad \text{solves} \quad -\psi'' + V\psi = z^2\psi$$

where $V(r) = 4A(2r)^2 - 4A'(2r)$.

Kreĭn does not give sum rules *per se*, but under suitable hypotheses, $P_*(r, z) \rightarrow \Pi(z)$ as $r \rightarrow \infty$ where $\Pi(z)$ is the outer function on the upper half-plane that obeys $|\Pi(x+i0)|^{-2} = 2\pi \frac{d\mu}{dx}$. This is essentially equivalent as discussed above. Lastly, the reader should be warned that Kreĭn’s paper contains no proofs; fortunately, details can be found in [84, 105].

5. Point Spectrum

As first noted in [33, p. 115], it follows from (21) that the bound-state energies, E_m , of a whole-line Schrödinger operator with potential $V \in L^2$ obey

$$\frac{2}{3} \sum |E_m|^{3/2} \leq \frac{1}{8} \int |V(x)|^2 dx. \quad (26)$$

This can be justified as follows: Choose $V_n \in C_c^\infty$ converging to V in L^2 . Then $L + V_n$ converges to \tilde{L} in strong resolvent sense, which implies (individual) convergence of the eigenvalues. Applying Fatou’s lemma to the sum over eigenvalues and using the fact that $|a(E+i0)| \geq 1$ for any potential gives (26). The existence of non-trivial reflectionless potentials shows that the constant in this inequality is optimal.

Inequalities of this kind are known as Lieb–Thirring inequalities and hold in considerable generality, including higher dimensions; see [61]. Considerable attention has been paid to the question of the optimal constants. In [57], Laptev and Weidl made a major breakthrough:

THEOREM 5.1. *The negative eigenvalues of $-\Delta + V$ acting in $L^2(\mathbb{R}^d)$ obey*

$$\sum |E_m|^\gamma \leq \frac{\Gamma(\gamma + 1)}{2^d \pi^{d/2} \Gamma(\gamma + \frac{1}{2}d + 1)} \int |V(x)|^{\gamma + \frac{1}{2}d} dx$$

for any $d \geq 1$ and $\gamma \geq \frac{3}{2}$. Moreover, the constant is optimal.

This is proved by extending the inequality (26) to operator-valued potentials (using trace formula methods) and then employing induction in dimension. An alternate proof of the trace formula portion of the argument appears in [5].

It is elementary to apply the reasoning described above to (17); this leads to the following result [47, 54]:

$$[T_n(\frac{1}{2}\tilde{J}) - T_n(\frac{1}{2}J)] \in \mathfrak{J}_2 \Rightarrow \sum [|\tilde{E}_j| - 2]^{3/2} < \infty. \quad (27)$$

(Analogous results can be found in [53, 81].) The particular case $n = 1$, treated in [47], is the natural Jacobi-matrix analogue of (26). For further inequalities of this type, see [41].

For CMV matrices and Kreĭn systems with decaying coefficients, the essential spectrum fills S^1 and \mathbb{R} , respectively. Thus there is no discrete spectrum.

6. A.C. Spectrum

It is well known that a one-dimensional Schrödinger operator with sufficiently rapidly decreasing potential has a.c. spectrum on $[0, \infty)$ —with sufficient decay it will even be purely absolutely continuous. But how quickly is sufficiently quickly?

On the basis of sparse, [50, 68, 72], and random, [18, 19, 51, 89], examples, it was known that there are potentials just outside L^2 which produce no a.c. spectrum whatsoever. Indeed, Simon has shown that this is generic, [90]. Eventually, the weight of this and other evidence led Kiselev, Last, and Simon, [50], to conjecture that L^2 was the correct borderline.

In his thesis, Kiselev made a significant step toward verifying this conjecture. This approach was later refined in [12], while an alternate approach was developed by Remling, [73]. The central conclusion of this work was: If $|V(x)| \lesssim (1+x^2)^{-\epsilon-1/4}$ then for almost every positive energy, all generalized eigenfunctions are bounded. In particular, the essential support of the a.c. spectrum fills $[0, \infty)$. It would be extremely interesting to know whether eigenfunctions are bounded at almost every positive energy when $V \in L^2$; in the regime of infinitesimal coupling, this reduces to Carleson's theorem on a.e. convergence of Fourier integrals. See [104] for more on this perspective.

The spectral question for $V \in L^2$ has been resolved using sum-rule methods, [16]:

THEOREM 6.1. *The absolutely continuous spectrum of a half-line Schrödinger operator with potential $V \in L^2$ is essentially supported by $[0, \infty)$.*

PROOF. Keeping only the imaginary part of m in (22) leads to

$$|a(k^2 + i0)|^2 \geq \frac{[\text{Im } m(k^2 + i0) + k]^2}{4k \text{Im } m(k^2 + i0)} \geq 1$$

Notice that $|a|$ is large wherever $\frac{d\mu}{dx} = \frac{1}{\pi} \text{Im } m$ is small, but by (21), we know that the integral of $\log |a|$ is controlled by the L^2 norm of the potential. The only obstacle is that we only know (21) for compactly supported potentials; this is resolved by choosing a sequence $V_n \rightarrow V$ and applying a simple semi-continuity argument. \square

There are now many results proved by similar means; we will give a brief overview of these and then turn to the Jacobi matrix case, where we offer a more detailed presentation. After that we will describe the analogous results for CMV matrices and Kreĭn systems, which are actually the oldest of all.

Using higher-order sum rules, [65] proves full a.c. spectrum under the hypotheses $V^{(p-1)} \in L^2$ and $V \in L^{p+1}$ for any integer $p \geq 1$. By using the connection to Kreĭn systems, Denisov obtained the same conclusion under the following hypotheses: V is uniformly L^2_{loc} and $V = A'$ with $A \in L^2$, see [20]; or $\limsup V(x) = 0$ and $V' \in L^2$, see [21]. See also [78].

In [46], a modification of the trace formula method was developed that works locally in energy. Specifically, one studies the perturbation determinant in a small region which touches the boundary along an interval, which allows one to consider perturbations of operators with non-zero potentials. In this way, it was shown that the a.c. spectrum of periodic Schrödinger operators is invariant under L^2 perturbation. The Stark operator was also studied; see [69] for further developments in this direction and for references to work on this operator that is not based on sum rules.

Another result from [46] is the following: if $V \in L^3$ and (the distribution) \hat{V} agrees with an L^2 function on an interval $[a^2, b^2]$, then $-\partial_x^2 + V$ has a.c. spectrum throughout the interval $[2a, 2b]$. See also [81], which treats Jacobi matrices. By combining the problems for V and $-V$ as in [83], one can see that the condition $V \in L^3$ can be replaced by $V \in L^4$. This was pointed out to me by O. Safronov.

The most interesting recent development of the trace formula method has been its extension to higher dimensions. For Dirac operators, there are the impressive results of Denisov, [25, 26]. Progress for Schrödinger operators has been slower for two reasons: bound states are especially problematic in the multi-dimensional case and there is no satisfactory WKB theory without smoothness assumptions on the potential. For the state of the art, see [24, 27, 55, 56, 82, 83] and the Denisov–Kiselev contribution to this Festschrift.

We will now present a Jacobi-matrix analogue of Theorem 6.1. The case of discrete Schrödinger operators was discussed in [16]; however, our treatment follows [47] with additional input from [54, 66]. The final result is from [66]. As suggested in the last proof, the main ingredient is a semi-continuity statement:

LEMMA 6.2. *Given probability measures $d\nu$ and $d\sigma$ on \mathbb{R} ,*

$$S(d\nu|d\sigma) := \inf \int e^g d\sigma - \int (g + 1) d\nu = \begin{cases} - \int \log[w] d\nu & : d\nu = wd\sigma \\ -\infty & : \text{otherwise} \end{cases}$$

where the infimum is over bounded continuous functions g . As a consequence, if $d\sigma_n$ converges weak-* to $d\sigma$, then $S(d\nu|d\sigma) \geq \limsup S(d\nu|d\sigma_n)$.

PROOF. The case where $d\nu$ is not $d\sigma$ -a.c. is easily dealt with; we suppose $d\nu = wd\sigma$. Let us write $g = c + h$ where $c = \int g d\nu$. By Jensen's inequality,

$$S(d\nu|d\sigma) \leq e^c \int w^{-1} e^h d\nu - c - 1 \leq \exp\{c - \int \log[w] d\nu\} - c - 1.$$

The minimizing value of c is $\int \log[w] d\nu$, which proves $S(d\nu|d\sigma) \leq - \int \log[w] d\nu$.

The fact that this inequality can be saturated follows by choosing g to approximate $\log[w]$, which corresponds to the case of equality in Jensen's inequality. \square

REMARK. By choosing $g \equiv 0$, it follows that $S(d\nu|d\sigma) \leq 0$. Consequently,

$$\int \log \left[\frac{d\nu}{d\mu} \right] d\nu \leq -S(d\nu|d\mu). \tag{28}$$

THEOREM 6.3. *Let J be the Jacobi matrix with $a_k \equiv 1$ and $b_k \equiv 0$. If \tilde{J} is a Jacobi matrix with $[T_n(\frac{1}{2}\tilde{J}) - T_n(\frac{1}{2}J)] \in \mathfrak{I}_2$ for some integer $n \geq 1$, then*

$$\int_{-2}^2 \log \left[\frac{d\tilde{\mu}}{d\mu} \right] \sqrt{4 - x^2} |U_{n-1}(\frac{x}{2})|^2 dx < \infty$$

where T_n and U_n represent Chebyshev polynomials as in (17).

PROOF. The result follows by combining (17) and Lemma 6.2 once we know that there are a sequence of operators J_k each differing from J by finite rank such that $J_k \rightarrow \tilde{J}$ strongly (which implies weak- $*$ convergence of the spectral measures) and for which $T_n(\frac{1}{2}J_k) - T_n(\frac{1}{2}J)$ is bounded in Hilbert–Schmidt norm. Such a sequence does exist because $\tilde{a}_n \rightarrow 1$ and $\tilde{b}_n \rightarrow 0$. This can be shown by examining the top three diagonals in $T_n(\frac{1}{2}\tilde{J})$; for details see Lemma 6.6 of [66]. \square

Jacobi matrix results have developed along lines parallel to the Schrödinger case—though the proper analogue of [65] remains particularly stubborn; see [53] for the latest on this problem.

That the a.c. spectrum fills S^1 for CMV matrices with $\alpha_k \in \ell^2$ follows from early work of Szegő, [102, 103]. Indeed much more is true; see Theorem 8.1.

With regard to higher-order sum rules for CMV matrices, see [22, 28, 39, 94, 101].

For Kreĭn systems, we have the following [52]:

THEOREM 6.4. *When $A \in L^2(dr)$, the spectral measure obeys*

$$- \int \log \left[\frac{d\mu}{dx} \right] \frac{dx}{1+x^2} < \infty.$$

In particular, the essential support of the a.c. spectrum is \mathbb{R} .

7. The Step-by-Step Method

As we have seen, the *a priori* sum rules presented in Section 3 are ample for applications in (forward) spectral theory. In the next section, we will be presenting results that incorporate inverse spectral theory and for this purpose, we need to discuss a second kind of *a priori* sum rule. The main idea can be found in [47, §4], but was first emphasized in [100]. The function-theoretic essence of the argument was distilled in [93]. We will present only the simplest case; it is not difficult to extend the results to the generality presented in Section 3.

HYPOTHESIS 7.1. *We assume $d\tilde{\mu}$ is a probability measure with support $[-2, 2] \cup \{\tilde{E}_j\}$ where \tilde{E}_j obeys $\sum[|\tilde{E}_j| - 2]^{3/2} < \infty$ and $\frac{d\tilde{\mu}}{dx} > 0$ almost everywhere in $[-2, 2]$.*

As previously, we write $\tilde{m}(z) = \langle e_1 | (\tilde{J} - z)^{-1} e_1 \rangle = \int \frac{1}{t-z} d\tilde{\mu}(t)$, which is a meromorphic function on Ω , the complement of $[-2, 2]$ in the Riemann sphere. We also enumerate the point spectrum $\{\tilde{E}_j\}$ so that $|\tilde{E}_j|$ is non-increasing.

A single step consists of removing the first row and column from J . We will denote the resulting Jacobi matrix by $\tilde{J}^{(1)}$, its spectral measure by $d\tilde{\mu}^{(1)}$, and m -function, $\tilde{m}^{(1)}(z)$.

LEMMA 7.2. *If $d\tilde{\mu}$ obeys Hypothesis 7.1, then so does $d\tilde{\mu}^{(1)}$.*

PROOF. By the min-max characterization of eigenvalues,

$$|\tilde{E}_j^{(1)}| \leq |\tilde{E}_j|. \tag{29}$$

Indeed, by the theory of rank-one perturbations, the eigenvalues of J and $J^{(1)}$ interlace. By the well-known formulae for inverting block matrices,

$$\tilde{m}(z) = [b_1 - z - a_1^2 \tilde{m}^{(1)}(z)]^{-1}. \tag{30}$$

In particular, taking the imaginary part we find

$$\frac{d\tilde{\mu}^{(1)}}{dx} \div \frac{d\tilde{\mu}}{dx} = \frac{\text{Im } \tilde{m}^{(1)}(x + i0)}{\text{Im } \tilde{m}(x + i0)} = \tilde{a}_1^{-2} |\tilde{m}(x + i0)|^{-2} \tag{31}$$

for a.e. $x \in [-2, 2]$. This completes the proof; Herglotz functions have non-zero boundary values almost everywhere. \square

The step-by-step approach studies $\log |\tilde{m}(z)|$ in very much the same manner as we studied $\log |a(z)|$ in Section 3; its boundary values can be read off (31) while the behaviour at infinity is governed by

LEMMA 7.3. *If $d\tilde{\mu}$ has compact support,*

$$\log[-z\tilde{m}(z)] = \sum_{k=1}^{\infty} -\frac{1}{k} z^{-k} \text{tr}\{\tilde{J}^k - (0 \oplus \tilde{J}^{(1)})^k\} \tag{32}$$

for z sufficiently large. Note, $0 \oplus \tilde{J}^{(1)}$ differs from \tilde{J} by having $a_1 = b_1 = 0$.

PROOF. By writing $m(z) = \text{tr}\{P(\tilde{J} - z)^{-1}\}$ with $P = |e_1\rangle\langle e_1|$ and expanding,

$$\log[-zm(z)] = \sum_{k=1}^{\infty} z^{-k} \sum_{p=1}^k \frac{(-1)^p}{p} \sum_{t_1 + \dots + t_p = k} \text{tr}\{P\tilde{J}^{t_1}P\tilde{J}^{t_2}\dots P\tilde{J}^{t_k}\}$$

where t_1, \dots, t_p are positive integers. Writing out the matrix product, we can regard the trace here as a sum over m -tuples (i_1, \dots, i_m) where $i_s = 1$ whenever s belongs to the set $\{1, 1 + t_1, \dots, 1 + t_1 + \dots + t_{p-1}\}$. Similarly,

$$\text{tr}\{\tilde{J}^k - (0 \oplus \tilde{J}^{(1)})^k\} = - \sum \tilde{J}(j_1, j_2)\tilde{J}(j_2, j_3)\dots\tilde{J}(j_k, j_1)$$

where the sum is taken over k -tuples with $j_s = 1$ for at least one s .

To connect the two, one should perform inclusion/exclusion on the number of times a k -tuple visits the value 1; the role of p is to restrict to k -tuples visiting 1 at least p times. \square

I have not seen (32) in the literature. This is not the simplest proof; however having typed all those indices, I am loath to delete them. A simpler proof was suggested to me by Barry: By (8) and Cramer’s rule, $m(z)$ can be written as a ratio of determinants and thus $\log[-zm(z)]$ can be written as the differences of traces. To make this fully rigorous, one first treats finite Jacobi matrices and then observes that this suffices.

THEOREM 7.4. *Let J denote the Jacobi matrix with $a_j \equiv 1$ and $b_j \equiv 0$. Let us fix $n \geq 1$ and suppose $d\tilde{\mu}$ obeys Hypothesis 7.1. Then for each $k \geq 1$,*

$$\begin{aligned}
 &-\frac{n}{\pi} \int_{-2}^2 \log \left[\frac{d\tilde{\mu}}{dx} \div \frac{d\tilde{\mu}^{(k)}}{dx} \right] \sqrt{4-x^2} |U_{n-1}(\frac{x}{2})|^2 dx + \sum G_n(E_j) - G_n(E_j^{(k)}) \\
 &= \text{tr} \left\{ [2T_n(\frac{1}{2}\tilde{J}) - 2T_n(\frac{1}{2}J)]^2 - [2T_n(\frac{1}{2}\mathbf{0} \oplus \tilde{J}^{(k)}) - 2T_n(\frac{1}{2}\mathbf{0} \oplus J^{(k)})]^2 \right\} \quad (33) \\
 &\quad + 4 \sum_{j=1}^k F(\tilde{a}_j \cdots \tilde{a}_{j+n-1}) + X_n - X_n^{(k)}
 \end{aligned}$$

where $\mathbf{0}$ represents the $k \times k$ zero matrix, $F(x) = x - 1 - \log(x) \geq 0$, G_n is given by (18), and X_n and $X_n^{(k)}$ are simple functions of the first few entries of J and $J^{(k)}$, respectively.

PROOF. It suffices to prove the case $k = 1$ since the general case follows by applying this successively. This case corresponds to Green’s identity with $f(z) = \log|m(z)|$ and $g(z) = G_n(z)$. Note that $m(z)$ has a pole at every eigenvalue of J and a zero at those of $J^{(1)}$. Also, $0 \leq G(x) \lesssim [|x| - 2]^{3/2}$ for $x \in \mathbb{R}$ and so by Lemma 7.2, the sum over each set of eigenvalues is absolutely convergent. \square

Because of the interlacing property of the discrete spectrum and the monotonicity of G_n , it is not necessary to assume that $d\tilde{\mu}$ has the Lieb–Thirring property.

8. Necessary and Sufficient Conditions

In this section, our presentation will most closely resemble the historical development; though as previously, we will restrict detailed discussions to the Jacobi case. The primary topic is the optimal versions of the sum rules we have described—the versions with no hypotheses; the left-hand side equals the right, be they finite or infinite.

The first sum rule to reach this stage of development is that of Verblunsky [107]:

THEOREM 8.1. *The coefficients of a CMV matrix, α_k , and its spectral measure, $d\mu$, always obey*

$$\prod_{k=0}^{\infty} (1 - |\alpha_k|^2) = \exp \left\{ \int \log \left[2\pi \frac{d\mu}{d\theta} \right] \frac{d\theta}{2\pi} \right\}.$$

In particular, the right-hand side is finite if and only if $\alpha_k \in \ell^2$.

This result admits several ‘higher order’ analogues where $\frac{d\theta}{2\pi}$ is replaced by $|P(\theta)|^2 \frac{d\theta}{2\pi}$ with P a trigonometric polynomial; see [100] and [94, §2.8].

Theorem 8.1 is often referred to as Szegő’s theorem in deference to [102, 103]; see [94] for a thorough historical discussion. There is a related sum rule which goes under the name ‘strong Szegő theorem’. The definitive version of this is due to Golinskii and Ibragimov, [38, 42]:

THEOREM 8.2. *If $d\mu = \frac{1}{2\pi} e^{h(\theta)} d\theta$, then*

$$\prod_{k=0}^{\infty} (1 - |\alpha_k|^2)^{-k-1} = \exp \left\{ \sum_{n=1}^{\infty} n |\hat{h}(n)|^2 \right\}$$

and the left-hand side is infinite if $d\mu$ cannot be written in this form.

There are two more results of similar nature, although neither has a corresponding trace formula. The first is due to Baxter, [4]:

THEOREM 8.3. $\alpha_k \in \ell^1$ if and only if $d\mu = \frac{1}{2\pi} e^{h(\theta)} d\theta$ with $\hat{h} \in \ell^1$.

This can be interpreted as a statement about the Wiener algebra. As discussed by Baxter, the result extends to other algebras; see also [94].

The second result is from [67]:

THEOREM 8.4. $\limsup |\alpha_k|^{1/k} \leq R^{-1} < 1$ if and only if $d\mu = |f(e^{i\theta})|^{-2} d\theta$ with $f(z)$ an analytic function on $|z| < R$.

This result has recently been the subject of much study, including several extensions in the circle case, [2, 3, 17, 96], and also to Jacobi matrices, [15, 97]. See also the review article [98].

In the remainder of this section, we will discuss analogues of Theorem 8.1 for Jacobi matrices and Schrödinger operators; I am not aware of a corresponding result for Kreĭn systems. It would be interesting to find analogues of Theorems 8.2 and 8.3. As far as I know, the only work on this question is [79, 80], which treats Jacobi matrices. Note that as the rate of decay improves, the analysis becomes more tractable; for instance, the classical theorems of forward and inverse scattering (as used to solve KdV and the Toda lattice), [63, 106], have weighted L^1 hypotheses.

The following result is from [66]; it extends earlier results from [47] and [54]. We give a slightly different proof.

THEOREM 8.5. Fix $n \geq 1$ and write J for the Jacobi matrix with $a_j \equiv 1$ and $b_j \equiv 0$. Then $T_n(\frac{1}{2}\tilde{J}) - T_n(\frac{1}{2}J)$ is Hilbert–Schmidt if and only if the spectral measure $d\tilde{\mu}$ obeys

- (i) (Blumenthal–Weyl) $\text{supp}(d\tilde{\mu})$ is compact and $\text{ess-supp}(d\tilde{\mu}) = [-2, 2]$.
- (ii) (Normalization) $d\tilde{\mu}$ is a probability measure.
- (iii) (Lieb–Thirring Bound)

$$\sum (|\tilde{E}_j| - 2)^{3/2} < \infty \tag{34}$$

- (iv) (Quasi-Szegő Condition) Let $d\tilde{\mu}_{ac}(E) = w(E) dE$. Then

$$\int_{-2}^2 \log[w(E)] |U_{n-1}(\frac{1}{2}E)|^2 \sqrt{4 - E^2} dE > -\infty. \tag{35}$$

REMARK. When n is small, the Hilbert–Schmidt condition can be reduced to simple explicit hypotheses on the coefficients by brute force; the general case was treated in [66] by using the recurrence relation for Chebyshev polynomials. The reformed condition is

$$\begin{aligned} (u_j + u_{j+1} + \dots + u_{j+n-1}) &\in \ell^2, \quad u_j \in \ell^4, \\ (\tilde{b}_j + \tilde{b}_{j+1} + \dots + \tilde{b}_{j+n-1}) &\in \ell^2, \quad \text{and } \tilde{b}_j \in \ell^4, \end{aligned}$$

where $u_j = \tilde{a}_j^2 - 1$.

PROOF. The forward implication follows from Weyl’s theorem (on relatively compact perturbations), (27), and Theorem 6.3.

For the other direction, we use Theorem 7.4. The first observation is that LHS(33) is bounded from above as $k \rightarrow \infty$; naively, it may happen that

$$\frac{n}{\pi} \int_{-2}^2 \log \left[\frac{d\tilde{\mu}^{(k)}}{d\mu} \right] \sqrt{4 - x^2} |U_{n-1}(\frac{x}{2})|^2 dx \rightarrow -\infty$$

but by (28), this sequence cannot diverge to $+\infty$. Therefore, RHS(33) must also be bounded above as $k \rightarrow \infty$.

As $d\tilde{\mu}$ has compact support, the coefficients of \tilde{J} are uniformly bounded. The sequence \tilde{a}_j is bounded from below, for if it were not, trace-class perturbation theory would imply that $d\mu$ is purely singular, [30, 99]. In this way, the bound on RHS(33) translates into

$$\limsup_{k \rightarrow \infty} \sum_{j=1}^k \left\langle e_j \left| [T_n(\tfrac{1}{2}\tilde{J}) - T_n(\tfrac{1}{2}J)]^2 e_j \right\rangle < \infty,$$

which completes the proof. \square

An analogous result for perturbations of periodic operators can be found in [14].

The proof of Theorem 8.5 given above avoids a very interesting idea that was employed in [54, 66], namely Denisov's extension of Rakhmanov's theorem, [23]:

THEOREM 8.6. *Let J be a bounded Jacobi matrix, and $d\mu$ its spectral measure. If $\sigma_{\text{ess}}(J) = [-2, 2]$ and $\frac{d\mu}{dx} > 0$ a.e. there, then $a_n \rightarrow 1$ and $b_n \rightarrow 0$.*

The original theorem of Rakhmanov, [64, 71], says the following: if the spectral measure of a CMV matrix obeys $\frac{d\mu}{d\theta} > 0$ a.e. on the unit circle, then $\alpha_k \rightarrow 0$.

To obtain the Schrödinger analogue of Theorem 8.5, one must confront two new difficulties.

First, every probability measure is the spectral measure for some Jacobi matrix, but not every positive measure on \mathbb{R} is the spectral measure of a Schrödinger operator. Necessary and sufficient conditions are known, [63]; they involve the large-energy asymptotics of the spectral measure. In addition, for technical reasons, one would like a statement that guarantees the existence of an L^2_{loc} potential.

The second problem is the occurrence of the real part of m in the natural trace formula. By analogy with Theorem 8.5, one would like to have a condition on the logarithmic integrability of the Radon–Nikodym derivative of the spectral measure.

The theorem below is from [48]. But first, a few remarks about how these difficulties are overcome.

Let $d\rho$ denote the spectral measure for a half-line Schrödinger operator (or a candidate for this role) and let $d\rho_0$ denote the measure for the free ($V \equiv 0$) case. We define a signed measure $d\nu$ on $(1, \infty)$ by

$$\frac{2}{\pi} \int f(k^2)k d\nu(k) = \int f(E)[d\rho(E) - d\rho_0(E)], \quad \forall f \in C_c^\infty((1, \infty)). \quad (36)$$

Notice that $d\nu$ is parameterized by momentum, k , rather than energy, E . Using Barry's A -function approach to the inverse problem, [36, 76, 92], it is possible to show that if $\sum [|\nu|(n, n+1)]^2$ is finite, then $d\rho$ is the spectral measure of a potential $V \in L^2_{\text{loc}}$. Using trace-formula methods, it is possible to show that this sum is finite for any $V \in L^2$.

Following the work of Burkholder, Gundy, and Silverstein, [6], it is understood that L^p bounds on the maximal function are equivalent to such bounds on the conjugate function. This is progress in our setting because it removes the spectre of cancellation. It also unifies the way one measures the size of the singular and absolutely continuous parts of $d\rho$. The specific hypothesis below makes use of a

short-range modification of the usual Hardy–Littlewood maximal function:

$$(M_s\nu)(x) = \sup_{0 < L \leq 1} \frac{|\nu|([x - L, x + L])}{2L}.$$

THEOREM 8.7. *A positive measure $d\rho$ on \mathbb{R} is the spectral measure associated to a (Dirichlet) half-line Schrödinger operator with potential $V \in L^2(\mathbb{R}^+)$ if and only if*

- (i) (Weyl) $\text{supp}(d\rho)$ is bounded from below and $\text{ess-supp}(d\rho) = [0, \infty)$.
- (ii) (Normalization)

$$\int \log \left[1 + \left(\frac{M_s\nu(k)}{k} \right)^2 \right] k^2 dk < \infty \tag{37}$$

- (iii) (Lieb–Thirring)

$$\sum_j |E_j|^{3/2} < \infty \tag{38}$$

- (iv) (Quasi-Szegő)

$$\int_0^\infty \log \left[\frac{1}{4} \frac{d\rho}{d\rho_0} + \frac{1}{2} + \frac{1}{4} \frac{d\rho_0}{d\rho} \right] \sqrt{E} dE < \infty \tag{39}$$

One consequence of this theorem is that L^2 perturbations can give rise to more or less arbitrary embedded singular spectrum. A related result was proved in [29]; indeed, this paper was a major stimulus for [47, 48]. Other results on the nature of embedded singular spectrum (not using trace formula methods) can be found in [13, 49, 74, 75] and the Denisov–Kiselev contribution to this Festschrift.

References

- [1] J. Avron, P. H. M. van Mouche, and B. Simon, On the measure of the spectrum for the almost Mathieu operator. *Comm. Math. Phys.* **132** (1990), 103–118.
- [2] D. Barrios Rolanía, G. López Lagomasino, and E. B. Saff, Asymptotics of orthogonal polynomials inside the unit circle and Szegő–Padé approximants. *J. Comput. Appl. Math.* **133** (2001), 171–181.
- [3] D. Barrios Rolanía, G. López Lagomasino, and E. B. Saff, Determining radii of meromorphy via orthogonal polynomials on the unit circle. *J. Approx. Theory* **124** (2003), 263–281.
- [4] G. Baxter, A convergence equivalence related to polynomials orthogonal on the unit circle. *Trans. Amer. Math. Soc.* **99** (1961), 471–487.
- [5] R. Benguria and M. Loss, A simple proof of a theorem of Laptev and Weidl. *Math. Res. Lett.* **7** (2000), 195–203.
- [6] D. L. Burkholder, R. F. Gundy, and M. L. Silverstein, A maximal function characterization of the class HP . *Trans. Amer. Math. Soc.* **157** (1971), 137–153.
- [7] V. S. Buslaev and L. D. Faddeev, Formulas for traces for a singular Sturm–Liouville differential operator. *Dokl. Akad. Nauk SSSR* **132** (1960), 13–16.
- [8] M. J. Cantero, L. Moral, and L. Velázquez, Five-diagonal matrices and zeros of orthogonal polynomials on the unit circle. *Linear Algebra Appl.* **362** (2003), 29–56.
- [9] T. Carleman, Zur Theorie der linearen Integralgleichungen. *Math. Zeit.* **9** (1921), 196–217.
- [10] K. M. Case, Orthogonal polynomials. II. *J. Math. Phys.* **16** (1975), 1435–1440.
- [11] W. G. Chambers, Linear-network model for magnetic breakdown in two dimensions. *Phys. Rev.* **140** (1965), A135–A143.
- [12] 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** (1998), 771–797.
- [13] M. Christ and A. Kiselev, WKB and spectral analysis of one-dimensional Schrödinger operators with slowly varying potentials. *Comm. Math. Phys.* **218** (2001), 245–262.

- [14] D. Damanik, R. Killip, and B. Simon, in preparation.
- [15] D. Damanik and B. Simon, Jost functions and Jost solutions for Jacobi matrices, II. Decay and analyticity. *Int. Math. Res. Not.* **2006**, Article ID 19396, 32 pages, 2006.
- [16] P. Deift and R. Killip, On the absolutely continuous spectrum of one-dimensional Schrödinger operators with square summable potentials. *Comm. Math. Phys.* **203** (1999), 341–347.
- [17] P. Deift and J. Ostensson, A Riemann-Hilbert approach to some theorems on Toeplitz operators and orthogonal polynomials. *J. Approx. Theory* **139** 2006, 144–171.
- [18] F. Delyon, Appearance of purely singular continuous spectrum in a class of random Schrödinger operators. *J. Statist. Phys.* **40** (1985), 621–630.
- [19] F. Delyon, B. Simon, and B. Souillard, From power pure point to continuous spectrum in disordered systems. *Ann. Inst. H. Poincaré* **42** (1985), 283–309.
- [20] S. A. Denisov, On the application of some of M. G. Krein's results to the spectral analysis of Sturm-Liouville operators. *J. Math. Anal. Appl.* **261** (2001), 177–191.
- [21] S. A. Denisov, On the existence of the absolutely continuous component for the measure associated with some orthogonal systems. *Comm. Math. Phys.* **226** (2002), 205–220.
- [22] S. A. Denisov, Probability measures with reflection coefficients $\{a_n\} \in l^4$ and $\{a_{n+1} - a_n\} \in l^2$ are Erdős measures. *J. Approx. Theory* **117** (2002), 42–54.
- [23] S. Denisov, On Rakhmanov's theorem for Jacobi matrices. *Proc. Amer. Math. Soc.* **132** (2004), 847–852.
- [24] S. Denisov, Absolutely continuous spectrum of multidimensional Schrödinger operator. *Int. Math. Res. Not.* **2004**, 3963–3982.
- [25] S. Denisov, On the existence of wave operators for some Dirac operators with square summable potential. *Geom. Funct. Anal.* **14** (2004), 529–534.
- [26] S. A. Denisov, On the absolutely continuous spectrum of Dirac operator. *Comm. Partial Differential Equations* **29** (2004), 1403–1428.
- [27] S. A. Denisov, On the preservation of absolutely continuous spectrum for Schrödinger operators. *J. Func. Anal.* **231** (2006) 143–156.
- [28] S. Denisov and S. Kupin, Orthogonal polynomials and a generalized Szegő condition. *C. R. Math. Acad. Sci. Paris* **339** (2004), 241–244.
- [29] S. Denisov and S. Kupin, On the singular spectrum of Schrödinger operators with decaying potential. *Trans. Amer. Math. Soc.* **357** (2005), 1525–1544.
- [30] J. Dombrowski, Quasitriangular matrices. *Proc. Amer. Math. Soc.* **69** (1978), 95–96.
- [31] S. D. Fisher, *Function Theory on Planar Domains*. John Wiley and Sons, New York, 1983.
- [32] I. Fredholm, Sur une classe d'équations fonctionnelles. *Acta Math.* **27** (1903), 365–390.
- [33] C. S. Gardner, J. M. Greene, M. D. Kruskal, and R. M. Miura, Korteweg-deVries equation and generalization. VI. Methods for exact solution. *Comm. Pure Appl. Math.* **27** (1974), 97–133.
- [34] I. M. Gelfand, On identities for eigenvalues of a differential operator of second order. (Russian) *Uspehi Mat. Nauk* **11** (1956), 191–198.
- [35] I. M. Gelfand and B. M. Levitan, On a simple identity for the characteristic values of a differential operator of the second order. *Doklady Akad. Nauk SSSR* **88** (1953), 593–596.
- [36] F. Gesztesy and B. Simon, A new approach to inverse spectral theory. II. General real potentials and the connection to the spectral measure. *Ann. of Math.* **152** (2000), 593–643.
- [37] I. C. Gohberg and M. G. Kreĭn, *Introduction to the theory of linear nonselfadjoint operators*. Translations of Mathematical Monographs, **18**, American Mathematical Society, Providence, R.I., 1969.
- [38] B. L. Golinskii and I. A. Ibragimov, On Szegő's limit theorem. *Math. USSR Izv.* **5** (1971), 421–444.
- [39] L. Golinskii and A. Zlatoš, Coefficients of orthogonal polynomials on the unit circle and higher order trace Szegő theorems. Preprint.
- [40] A. Grothendieck, La théorie de Fredholm. *Séminaire Bourbaki*, Mars 1954, exp. no. 91, 53/54. Paris.
- [41] D. Hundertmark and B. Simon, Lieb-Thirring inequalities for Jacobi matrices. *J. Approx. Theory* **118** (2002), 106–130.
- [42] I. A. Ibragimov, A theorem of Gabor Szegő. *Mat. Zametki* **3** (1968), 693–702.
- [43] S. Ya. Jitomirskaya and I. V. Krasovsky, Continuity of the measure of the spectrum for discrete quasiperiodic operators. *Math. Res. Lett.* **9** (2002), 413–421.

- [44] W. B. Johnson, H. König, B. Maurey, and J. R. Retherford, Eigenvalues of p -summing and l_p -type operators in Banach spaces. *J. Funct. Anal.* **32** (1979), 353–380.
- [45] R. Jost and A. Pais, On the scattering of a particle by a static potential. *Physical Rev.* **82** (1951), 840–851.
- [46] R. Killip, Perturbations of one-dimensional Schrödinger operators preserving the absolutely continuous spectrum. *Int. Math. Res. Not.* (2002) 2029–2061.
- [47] R. Killip and B. Simon, Sum rules for Jacobi matrices and their applications to spectral theory. *Ann. of Math.* **158** (2003), 253–321.
- [48] R. Killip and B. Simon, Sum rules and spectral measures of Schrödinger operators with L^2 potentials. To appear in *Ann. of Math.*
- [49] A. Kiselev, Imbedded singular continuous spectrum for Schrödinger operators. *J. Amer. Math. Soc.* **18** (2005), 571–603.
- [50] A. Kiselev, Y. Last, and B. Simon, Modified Prüfer and EFGP transforms and the spectral analysis of one-dimensional Schrödinger operators, *Comm. Math. Phys.* **194** (1998), 1–45.
- [51] S. Kotani and N. Ushiroya, One-dimensional Schrödinger operators with random decaying potentials. *Comm. Math. Phys.* **115** (1988), 247–266.
- [52] M. G. Kreĭn, Continuous analogues of propositions on polynomials orthogonal on the unit circle. *Dokl. Akad. Nauk SSSR* **105** (1955), 637–640.
- [53] S. Kupin, On a spectral property of Jacobi matrices. *Proc. Amer. Math. Soc.* **132** (2004), 1377–1383.
- [54] A. Laptev, S. Naboko, and O. Safronov, On new relations between spectral properties of Jacobi matrices and their coefficients. *Comm. Math. Phys.* **241** (2003), 91–110.
- [55] A. Laptev, S. Naboko, and O. Safronov, A Szegő condition for a multidimensional Schrödinger operator. *J. Funct. Anal.* **219** (2005), 285–305.
- [56] A. Laptev, S. Naboko, and O. Safronov, Absolutely continuous spectrum of Schrödinger operators with slowly decaying and oscillating potentials. *Comm. Math. Phys.* **253** (2005), 611–631.
- [57] A. Laptev and T. Weidl, Sharp Lieb–Thirring inequalities in high dimensions. *Acta Math.* **184** (2000), 87–111.
- [58] Y. Last, A relation between a.c. spectrum of ergodic Jacobi matrices and the spectra of periodic approximants. *Comm. Math. Phys.* **151** (1993), 183–192.
- [59] P. D. Lax, *Functional Analysis*. John Wiley & Sons, New York, 2002.
- [60] V. B. Lidskiĭ, Non-selfadjoint operators with a trace. *Dokl. Akad. Nauk SSSR* **125** 1959, 485–487. (Translated in *Thirteen papers on functional analysis and partial differential equations.*) AMS Transl. Ser. 2, **47**. American Mathematical Society, Providence, R.I., 1965.)
- [61] E. H. Lieb and W. Thirring, Inequalities for the moments of the eigenvalues of the Schrödinger Hamiltonian and their relation to Sobolev inequalities. *Studies in Mathematical Physics. Essays in Honor of Valentine Bargmann*. pp. 269–303, Princeton University Press, Princeton, N.J., 1976.
- [62] J. Lindenstrauss and L. Tzafriri, *Classical Banach Spaces. I. Sequence Spaces*. Ergebnisse der Mathematik und ihrer Grenzgebiete, Vol. 92. Springer-Verlag, Berlin–New York, 1977.
- [63] V. A. Marchenko, *Sturm–Liouville Operators and Applications*. Operator Theory: Advances and Applications, 22. Birkhäuser Verlag, Basel, 1986.
- [64] A. Máté, P. Nevai, and V. Totik, Asymptotics for the ratio of leading coefficients of orthogonal polynomials on the unit circle. *Constr. Approx.* **1** (1985), 63–69.
- [65] S. Molchanov, M. Novitskii, and B. Vainberg, First KdV integrals and absolutely continuous spectrum for 1-D Schrödinger operator. *Comm. Math. Phys.* **216** (2001), 195–213.
- [66] F. Nazarov, F. Peherstorfer, A. Volberg, and P. Yuditskii, On generalized sum rules for Jacobi matrices. *Int. Math. Res. Not.* **2005**, 155–186.
- [67] P. Nevai and V. Totik, Orthogonal polynomials and their zeros. *Acta Sci. Math. (Szeged)* **53** (1989), 99–104.
- [68] D. Pearson, Singular continuous measures in scattering theory. *Comm. Math. Phys.* **60** (1978), 13–36.
- [69] G. Perelman, On the absolutely continuous spectrum of Stark operators. *Comm. Math. Phys.* **234** (2003), 359–381.
- [70] G. Pisier, *The Volume of Convex Bodies and Banach Space Geometry*. Cambridge Tracts in Mathematics, **94**. Cambridge University Press, Cambridge, 1989.

- [71] E. A. Rakhmanov, On the asymptotics of the ratio of orthogonal polynomials. II. *Math. USSR Sb.* **46** (1983), 105–117.
- [72] C. Remling, A probabilistic approach to one-dimensional Schrödinger operators with sparse potentials. *Comm. Math. Phys.* **185** (1997), 313–323.
- [73] C. Remling, The absolutely continuous spectrum of one-dimensional Schrödinger operators with decaying potentials. *Comm. Math. Phys.* **193** (1998), 151–170.
- [74] C. Remling, Bounds on the embedded singular spectrum for one-dimensional Schrödinger operator with decaying potentials. *Proc. Amer. Math. Soc.* **128** (2000), 161–171.
- [75] C. Remling, Schrödinger operators with decaying potentials: some counterexamples. *Duke Math. J.* **105** (2000), 463–496.
- [76] C. Remling, Inverse spectral theory for one-dimensional Schrödinger operators: The A function. *Math. Z.* **245** (2003), 597–617.
- [77] W. Rudin, Analytic functions of class H_p . *Trans. Amer. Math. Soc.* **78**, (1955), 46–66.
- [78] A. Rybkin, On the absolutely continuous and negative discrete spectra of Schrödinger operators on the line with locally integrable globally square summable potentials. *J. Math. Phys.* **45**(2004), 1418–1425.
- [79] E. Ryckman, A spectral equivalence for Jacobi operators. Preprint.
- [80] E. Ryckman, A strong Szegő theorem for Jacobi matrices. To appear in *Comm. Math. Phys.*
- [81] O. Safronov, The spectral measure of a Jacobi matrix in terms of the Fourier transform of the perturbation. *Ark. Mat.* **42** (2004), 363–377.
- [82] O. Safronov, On the absolutely continuous spectrum of multi-dimensional Schrödinger operators with slowly decaying potentials. *Comm. Math. Phys.* **254** (2005), 361–366.
- [83] O. Safronov, Multi-dimensional Schrödinger operators with no negative spectrum. Preprint.
- [84] L. A. Sakhnovich, On the spectral theory of a class of canonical differential systems. *Funktional. Anal. i Prilozhen.* **34** (2000), 50–62.
- [85] D. Sarason, The H^p spaces of an annulus. *Mem. Amer. Math. Soc.* **56** (1965).
- [86] E. Seiler, Schwinger functions for the Yukawa model in two dimensions with space-time cutoff. *Comm. Math. Phys.* **42** (1975), 163–182.
- [87] B. Simon, Notes on infinite determinants of Hilbert space operators. *Adv. in Math.* **24** (1977), 244–273.
- [88] B. Simon, *Trace Ideals and Their Applications*. London Mathematical Society Lecture Note Series, **35**. Cambridge University Press, Cambridge–New York, 1979; second edition, *Mathematical Surveys and Monographs*, **120**, American Mathematical Society, Providence, R.I., 2005.
- [89] B. Simon, Some Jacobi matrices with decaying potential and dense point spectrum. *Comm. Math. Phys.* **87** (1982), 253–258.
- [90] B. Simon, Operators with singular continuous spectrum: I. General operators. *Ann. of Math.* **141** (1995), 131–145.
- [91] B. Simon, The classical moment problem as a self-adjoint finite difference operator. *Adv. Math.* **137** (1998), 82–203.
- [92] B. Simon, A new approach to inverse spectral theory, I. Fundamental formalism. *Ann. of Math.* **150** (1999), 1029–1057.
- [93] B. Simon, A canonical factorization for meromorphic Herglotz functions on the unit disk and sum rules for Jacobi matrices. *J. Funct. Anal.* **214** (2004), 396–409.
- [94] B. Simon, *Orthogonal Polynomials on the Unit Circle, Part 1: Classical Theory*. AMS Colloquium Series, American Mathematical Society, Providence, R.I., 2005.
- [95] B. Simon, CMV matrices: Five years after. To appear in Proc. of the W. D. Evans 65th Birthday Conference.
- [96] B. Simon, Meromorphic Szegő functions and asymptotic series for Verblunsky coefficients. *Acta Math.* **195** (2005), 267–285.
- [97] B. Simon, Meromorphic Jost functions and asymptotic series for Jacobi parameters, to appear in *Funct. Anal. Appl.*
- [98] B. Simon, Orthogonal polynomials with exponentially decaying recursion coefficients. Preprint.
- [99] B. Simon and T. Spencer, Trace class perturbations and the absence of absolutely continuous spectra. *Comm. Math. Phys.* **125** (1989), 113–125.
- [100] B. Simon and A. Zlatoš, Sum rules and the Szegő condition for orthogonal polynomials on the real line. *Comm. Math. Phys.* **242** (2003), 393–423.

- [101] B. Simon and A. Zlatoš, Higher-order Szegő theorems with two singular points. *J. Approx. Theory* **134** (2005), 114–129.
- [102] G. Szegő, Beiträge zur Theorie der Toeplitzschen Formen. *Math. Z.* **6** (1920), 167–202.
- [103] G. Szegő, Beiträge zur Theorie der Toeplitzschen Formen, II. *Math. Z.* **9** (1921), 167–190.
- [104] T. Tao and C. Thiele, *The non-linear Fourier transform*. To appear in “Park City Mathematics Series,” published by the American Mathematical Society.
- [105] A. Teplyaev, A note on the theorems of M. G. Krein and L. A. Sakhnovich on continuous analogues of orthogonal polynomials on the circle. *J. Funct. Anal.* **226** (2005), 257–280.
- [106] G. Teschl, *Jacobi Operators and Completely Integrable Nonlinear Lattices*. Mathematical Surveys and Monographs, **72**, American Mathematical Society, Providence, R.I., 2000.
- [107] S. Verblunsky, On positive harmonic functions. *Proc. London Math. Soc.* **40** (1936), 290–320.
- [108] D. Watkins, Some perspectives on the eigenvalue problem. *SIAM Rev.* **35** (1993), 430–471.
- [109] H. Weyl, Inequalities between the two kinds of eigenvalues of a linear transformation. *Proc. Nat. Acad. Sci. U.S.A.* **35** (1949), 408–411.
- [110] V. E. Zakharov and L. D. Faddeev, The Korteweg–de Vries equation is a completely integrable Hamiltonian system. *Funkcional. Anal. i Priložen.* **5** (1971), 18–27.

UCLA MATHEMATICS DEPARTMENT, BOX 951555, LOS ANGELES, CA 90095, U.S.A.
E-mail address: killip@math.ucla.edu

Barry Simon's List of Publications

Here is Barry's publication list as of October 2006. It is kept updated at <http://math.caltech.edu/people/biblio.html>. It is in groups: first, 309 research papers, then 14 scientific books, and then 48 review articles. Barry's computer books and his reviews in mass market computer magazines are not included.

Papers

- [1] Convergence of regularized, renormalized perturbation series for super-renormalizable field theories, *Nuovo Cimento* **59A** (1969), 199–214.
- [2] Some pictorial compactifications of the real line, *Amer. Math. Monthly* **76** (1969), 536–538.
- [3] On the growth of the number of bound state with increase in potential strength, *J. Math. Phys.* **10** (1969), 1123–1126.
- [4] On the growth of the ground state binding energy with increase in potential strength, *J. Math. Phys.* **10** (1969), 1415–1421.
- [5] On positive eigenvalues of one-body Schrödinger operators, *Commun. Pure Appl. Math.* **12** (1969), 531–538.
- [6] (with J. J. Loeffel, A. Martin and A. S. Wightman) Padé approximants and the anharmonic oscillator, *Phys. Lett.* **30B** (1969), 656–658.
- [7] Coupling constant analyticity for the anharmonic oscillator (with an appendix by A. Dicke), *Ann. Phys.* **58** (1970), 76–136.
- [8] Some comments on the Jin–Martin lower bound, *Phys. Rev.* **D1** (1970), 1240–1241.
- [9] On the infinitude or finiteness of the number of bound states of an N -body quantum system, I, *Helv. Phys. Acta* **43** (1970), 607–630.
- [10] (with S. Graffi and V. Grecchi) Borel summability: Application to the anharmonic oscillator, *Phys. Lett.* **32B** (1970), 631–634.
- [11] Borel summability of the ground state energy in spatially cutoff $(\phi^4)_2$, *Phys. Rev. Lett.* **25** (1970), 1583–1586.
- [12] Distributions and their Hermite expansions, *J. Math. Phys.* **12** (1970), 140–148.
- [13] Hamiltonians defined as quadratic forms, *Commun. Math. Phys.* **21** (1971), 192–210.
- [14] The theory of semi-analytic vectors: A new proof of a theorem of Masson and McClary, *Indiana Univ. Math. J.* **20** (1971), 1145–1151.
- [15] Wave operators for classical particle scattering, *Commun. Math. Phys.* **23** (1971), 37–48.
- [16] (with R. Høegh-Krohn) Hypercontractive semigroups and two-dimensional self-coupled Bose fields, *J. Funct. Anal.* **9** (1972), 121–180.

- [17] Determination of eigenvalues by divergent perturbation series, *Adv. in Math.* **7** (1971), 240–253.
- [18] (with L. Rosen) The $(\phi^{2n})_2$ Hamiltonian for complex coupling constant, *Trans. Amer. Math. Soc.* **165** (1972), 365–379.
- [19] Convergence of time-dependent perturbation theory for autoionizing states of atoms, *Phys. Lett.* **A36** (1971), 23–24.
- [20] Resonances in n -body quantum systems with dilation analytic potentials and the foundations of time-dependent perturbation theory, *Ann. of Math.* **97** (1973), 247–274.
- [21] On the Glimm–Jaffe linear lower bound in $P(\phi)_2$ field theories, *J. Funct. Anal.* **10** (1972), 251–258.
- [22] A remark on groups with the fixed point property, *Proc. Amer. Math. Soc.* **32** (1972), 623–624.
- [23] Continuum embedded eigenvalues in $P(\phi)_2$ field theory, *Proc. Amer. Math. Soc.* **35** (1972), 223–226.
- [24] Essential self-adjointness of Schrödinger operators with positive potentials, *Math. Ann.* **201** (1973), 211–220.
- [25] Quadratic form techniques and the Balslev–Combes theorem, *Commun. Math. Phys.* **27** (1972), 1–9.
- [26] (with M. Reed) A spectral mapping theorem for tensor products of unbounded operators, *Bull. Amer. Math. Soc.* **78** (1972), 730–733.
- [27] (with F. Guerra and L. Rosen) Nelson’s symmetry and the infinite volume behavior of the vacuum in $P(\phi)_2$, *Commun. Math. Phys.* **27** (1972), 10–22.
- [28] Summability methods, the strong asymptotic condition, and unitarity in quantum field theory, *Phys. Rev. Lett.* **28** (1972), 1145–1146.
- [29] Uniform cross norms, *Pacific J. Math.* **46** (1973), 555–560.
- [30] (with M. Reed) Tensor products of closed operators on Banach spaces, *J. Funct. Anal.* **13** (1973), 107–124.
- [31] (with F. Guerra and L. Rosen) The vacuum energy for $P(\phi)_2$: Infinite volume limit and coupling constant dependence, *Commun. Math. Phys.* **29** (1973), 233–247.
- [32] (with F. Guerra and L. Rosen) Statistical mechanical results in the $P(\phi)_2$ quantum field theory, *Phys. Lett.* **44B** (1973), 102–104.
- [33] (with F. Guerra and L. Rosen) The $P(\phi)_2$ Euclidean quantum field theory as classical statistical mechanics, *Ann. of Math.* **101** (1975), 111–259.
- [34] (with F. Guerra and L. Rosen) Boundary conditions for the $P(\phi)_2$ Euclidean field theory, *Ann. Inst. H. Poincaré* **25A** (1976), 231–334.
- [35] Essential self-adjointness of Schrödinger operators with singular potentials *Arch. Rat. Math. Anal.* **52** (1973), 44–48.
- [36] Ergodic semigroups of positivity preserving self-adjoint operators, *J. Funct. Anal.* **12** (1973), 335–339.
- [37] Correlation inequalities and the mass gap in $P(\phi)_2$, I. Domination by the two point function, *Commun. Math. Phys.* **31** (1973), 127–136.
- [38] Schrödinger operators with singular magnetic vector potentials, *Math. Z.* **131** (1973), 361–370.
- [39] (with E. Lieb) Thomas–Fermi theory revisited, *Phys. Rev. Lett.* **31** (1973), 681–683.

- [40] (with R. Griffiths) Griffiths–Hurst–Sherman inequalities and a Lee–Yang theorem for the $(\phi^4)_2$ field theory, *Phys. Rev. Lett.* **30** (1973), 931–933.
- [41] Quadratic forms and Klauder's phenomenon: A remark on very singular perturbations, *J. Funct. Anal.* **14** (1973), 295–298.
- [42] Positivity of the Hamiltonian semigroup and the construction of Euclidean region fields, *Helv. Phys. Acta* **46** (1973), 686–696.
- [43] Pointwise bounds on eigenfunctions and wave packets in N -body quantum systems, I, *Proc. Amer. Math. Soc.* **42** (1974), 395–401.
- [44] Absence of positive eigenvalues in a class of multiparticle quantum systems, *Math. Ann.* **207** (1974), 133–138.
- [45] (with E. Lieb) On solutions to the Hartree–Fock problem for atoms and molecules, *J. Chem. Phys.* **61** (1974), 735–736.
- [46] Pointwise bounds on eigenfunctions and wave packets in N -body quantum systems, II, *Proc. Amer. Math. Soc.* **45** (1974), 454–456.
- [47] (with R. Griffiths) The $(\phi^4)_2$ field theory as a classical Ising model, *Commun. Math. Phys.* **33** (1973), 145–164.
- [48] Correlation inequalities and the mass gap in $P(\phi)_2$, II. Uniqueness of the vacuum for a class of strongly coupled theories, *Ann. of Math.* **101** (1975), 260–267.
- [49] (with F. Guerra and L. Rosen) The pressure is independent of the boundary conditions in $P(\phi)_2$ field theories, *Bull. Amer. Math. Soc.* **80** (1974), 1205–1209.
- [50] (with F. Guerra and L. Rosen) Correlation inequalities and the mass gap in $P(\phi)_2$, III. Mass gap for a class of strongly coupled theories with nonzero external field, *Commun. Math. Phys.* **41** (1975), 19–32.
- [51] Pointwise bounds on eigenfunctions and wave packets in N -body quantum systems, III, *Trans. Amer. Math. Soc.* **208** (1975), 317–329.
- [52] (with J. Rosen) Global support properties of stationary ergodic processes, *Duke Math. J.* **42** (1975), 51–55.
- [53] (with E. Lieb) The Thomas–Fermi theory of atoms, molecules and solids, *Adv. in Math.* **23** (1977), 22–116.
- [54] (with J. Rosen) Fluctuations in $P(\phi)_1$ processes, *Ann. Prob.* **4** (1976), 155–174.
- [55] (with E. Lieb) The Hartree–Fock theory for Coulomb systems, *Commun. Math. Phys.* **53** (1977), 185–194.
- [56] Existence of the scattering matrix for the linearized Boltzmann equation, *Commun. Math. Phys.* **41** (1975), 99–108.
- [57] Operator theory needed in quantum statistical mechanics in boxes, pp. 389–398, Appendix B, in E. Lieb and J. Lebowitz, “The constitution of matter,” *Adv. in Math.* **9** (1972), 316–398.
- [58] Convergence theorems for entropy, Appendix to E. Lieb and M. B. Ruskai, “Proof of the strong subadditivity of quantum mechanical entropy, *J. Math. Phys.* **14** (1973), 1938–1941.
- [59] (with W. Faris) Degenerate and non-degenerate ground states for Schrödinger operators, *Duke Math. J.* **42** (1975), 559–567.
- [60] (with E. Seiler) An inequality among determinants, *Proc. Natl. Acad. Sci. USA* **72** (1975), 3277–3278.

- [61] (with E. Seiler) On finite mass renormalizations in the two-dimensional Yukawa model, *J. Math. Phys.* **16** (1975), 2289–2293.
- [62] (with E. Seiler) Bounds in the Yukawa₂ quantum field theory: Upper bound on the pressure, Hamiltonian bound and linear lower bound, *Commun. Math. Phys.* **45** (1975), 99–114.
- [63] (with E. Seiler) Nelson symmetry and all that in the Yukawa₂ and $(\varphi^4)_3$ field theories, *Ann. Phys.* **97** (1976), 470–518.
- [64] (with J. Fröhlich and T. Spencer) Phase transitions and continuous symmetry breaking, *Phys. Rev. Lett.* **36** (1976), 804–806.
- [65] (with J. Fröhlich and T. Spencer) Infrared bounds, phase transitions and continuous symmetry breaking, *Commun. Math. Phys.* **50** (1976), 79–85.
- [66] Universal diamagnetism of spinless Bose systems, *Phys. Rev. Lett.* **36** (1976), 1083–1084.
- [67] (with F. J. Dyson and E. Lieb) Phase transitions in the quantum Heisenberg model, *Phys. Rev. Lett.* **37** (1976), 120–123.
- [68] (with F. J. Dyson and E. Lieb) Phase transitions in quantum spin systems with isotropic and non-isotropic interactions, *J. Statist. Phys.* **18** (1978), 335–383.
- [69] A remark on Nelson's best hypercontractive estimates, *Proc. Amer. Math. Soc.* **55** (1976), 376–378.
- [70] The bound state of weakly coupled Schrödinger operators in one and two dimensions, *Ann. Phys.* **97** (1976), 279–288.
- [71] (with P. Deift) On the decoupling of the finite singularities from the question of asymptotic completeness in two-body quantum systems, *J. Funct. Anal.* **23** (1976), 218–238.
- [72] On the absorption of eigenvalues by continuous spectrum in regular perturbation problems, *J. Funct. Anal.* **25** (1977), 338–344.
- [73] Analysis with weak trace ideals and the number of bound states of Schrödinger operators, *Trans. Amer. Math. Soc.* **224** (1976), 367–380.
- [74] Notes on infinite determinants of Hilbert space operators, *Adv. in Math.* **24** (1977), 244–273.
- [75] On the genericity of nonvanishing instability intervals in Hill's equation, *Ann. Inst. H. Poincaré* **A24** (1976), 91–93.
- [76] An abstract Kato's inequality for generators of positivity preserving semigroups, *Indiana Univ. Math. J.* **26** (1977), 1067–1073.
- [77] (with J. Fröhlich) Pure states for general $P(\phi)_2$ theories: Construction, regularity and variational equality, *Ann. of Math.* **105** (1977), 493–526.
- [78] (with M. Reed) The scattering of classical waves from inhomogeneous media, *Math. Z.* **155** (1977), 163–168.
- [79] (with J. Avron) Analytic properties of band functions, *Ann. Phys.* **110** (1978), 85–110.
- [80] (with R. Blankenbecler and M. L. Goldberger) The bound states of weakly coupled long-range one-dimensional quantum Hamiltonians, *Ann. Phys.* **108** (1977), 69–78.
- [81] A canonical decomposition for quadratic forms with applications to monotone convergence theorems, *J. Funct. Anal.* **28** (1978), 377–385.
- [82] Lower semicontinuity of positive quadratic forms, *Proc. Roy. Soc. Edinburgh* **29** (1977), 267–273.

- [83] (with P. Deift) A time-dependent approach to the completeness of multiparticle quantum systems, *Commun. Pure Appl. Math.* **30** (1977), 573–583.
- [84] Geometric methods in multiparticle quantum systems, *Commun. Math. Phys.* **55** (1977), 259–274.
- [85] N -body scattering in the two-cluster region, *Commun. Math. Phys.* **58** (1978), 205–210.
- [86] Kato's inequality and the comparison of semigroups, *J. Funct. Anal.* **32** (1979), 97–101.
- [87] (with J. Avron and I. Herbst) The Zeeman effect revisited, *Phys. Lett.* **62A** (1977), 214–216.
- [88] (with J. Avron and I. Herbst) Formation of negative ions in magnetic fields, *Phys. Rev. Lett.* **39** (1977), 1068–1070.
- [89] (with J. Avron and I. Herbst) Schrödinger operators with magnetic fields, I. General interactions, *Duke Math. J.* **45** (1978), 847–883.
- [90] (with J. Avron and I. Herbst) Schrödinger operators with magnetic fields, II. Separation of center of mass in homogeneous magnetic fields, *Ann. Phys.* **114** (1978), 431–451.
- [91] (with J. Avron and I. Herbst) Schrödinger operators in magnetic fields, III. Atoms in homogeneous magnetic field, *Commun. Math. Phys.* **79** (1981), 529–572.
- [92] (with J. Avron and I. Herbst) Schrödinger operators in magnetic fields, IV. Strongly bound states of hydrogen in intense magnetic field, *Phys. Rev.* **A20** (1979), 2287–2296.
- [93] (with J. Fröhlich, R. Israel, and E. Lieb) Phase transitions and reflection positivity, I. General theory and long range lattice models, *Commun. Math. Phys.* **62** (1978), 1–34.
- [94] (with J. Fröhlich, R. Israel, and E. Lieb) Phase transitions and reflection positivity, II. Lattice systems with short-range and Coulomb interactions, *J. Statist. Phys.* **22** (1980), 297–347.
- [95] (with P. Deift, W. Hunziker, and E. Vock) Pointwise bounds on eigenfunctions and wave packets in N -body quantum systems, IV, *Commun. Math. Phys.* **64** (1978), 1–34.
- [96] Scattering theory and quadratic forms: On a theorem of Schechter, *Commun. Math. Phys.* **53**(1977), 151–153.
- [97] (with E. Lieb) Monotonicity of the electronic contribution to the Born–Oppenheimer energy, *J. Phys.* **B11** (1978), L537–L542.
- [98] Maximal and minimal Schrödinger forms, *J. Oper. Th.* **1** (1979), 37–47.
- [99] (with E. B. Davies) Scattering theory for systems with different spatial asymptotics on the left and right, *Commun. Math. Phys.* **63** (1978), 277–301.
- [100] (with I. Herbst) Stark effect revisited, *Phys. Rev. Lett.* **41** (1978), 67–69.
- [101] (with I. Herbst) Some remarkable examples in eigenvalue perturbation theory, *Phys. Lett.* **78B** (1978), 304–306.
- [102] (with I. Herbst) Dilation analyticity in constant electric field, II. The N -body problem, Borel summability, *Commun. Math. Phys.* **80** (1981), 181–216.
- [103] (with C. Radin) Invariant domains for the time-dependent Schrödinger equation, *J. Diff. Eqn.* **29** (1978), 289–296.
- [104] (with L. Benassi, V. Grecchi, and E. Harrell) The Bender–Wu formula and the Stark effect in hydrogen, *Phys. Rev. Lett.* **42** (1979), 704–707.

- [105] (with E. Harrell) The mathematical theory of resonances which have exponentially small widths, *Duke Math. J.* **47** (1980), 845–902.
- [106] Phase space analysis of simple scattering systems. Extensions of some work of Enns, *Duke Math. J.* **46** (1979), 119–168.
- [107] The definition of molecular resonance curves by the method of exterior complex scaling, *Phys. Lett.* **71A** (1979), 211–214.
- [108] (with J. Morgan) On the asymptotics of Born–Oppenheimer curves for large nuclear separations, *Intl. J. Quan. Chem.* **17** (1980), 1143–1166.
- [109] (with M. Schechter) Unique continuation for Schrödinger operators with unbounded potential, *J. Math. Anal. Appl.* **77** (1980), 482–492.
- [110] Brownian motion, L^p properties of Schrödinger operators and the localization of binding, *J. Funct. Anal.* **35** (1980), 215–229.
- [111] (with S. Graffi and V. Grecchi) Complete separability of the Stark effect in hydrogen, *J. Phys.* **A12** (1979), 2193–2197.
- [112] A remark on Dobrushin’s uniqueness theorem, *Commun. Math. Phys.* **68** (1979), 183–185.
- [113] The classical limit of quantum partition functions, *Commun. Math. Phys.* **71** (1980), 247–276.
- [114] (with M. Klaus) Binding of Schrödinger particles through conspiracy of potential wells, *Ann. Inst. H. Poincaré* **A30** (1979), 83–87.
- [115] (with M. Klaus) Coupling constant thresholds in nonrelativistic quantum mechanics, I. Short range two-body case, *Ann. Phys.* **130** (1980), 251–281.
- [116] (with A. Alonso) The Birman–Krein–Vishik theory of self-adjoint extensions of semibounded operators, *J. Oper. Th.* **4** (1980), 251–270.
- [117] (with J. Avron) A counterexample to the paramagnetic conjecture, *Phys. Lett.* **75A** (1979), 41–42.
- [118] Mean field upper bound on the transition temperature of multi-component spin systems, *J. Statist. Phys.* **22** (1980), 481–493.
- [119] (with M. Hoffmann-Ostenhof and T. Hoffmann-Ostenhof) On the nodal structure of atomic eigenfunctions, *J. Phys.* **A13** (1980), 1131–1133.
- [120] (with M. Hoffmann-Ostenhof and T. Hoffmann-Ostenhof) Brownian motion and a consequence of Harnack’s inequality: Nodes of quantum wave functions, *Proc. Amer. Math. Soc.* **80** (1980), 301–305.
- [121] (with J. Morgan) The calculation of molecular resonances, *J. Phys.* **B14** (1981), L167–L171.
- [122] (with M. Klaus) Coupling constant threshold in non-relativistic quantum mechanics, II. Two-body thresholds in N -body systems, *Commun. Math. Phys.* **78** (1980), 153–168.
- [123] (with M. Klaus) Coupling constant thresholds in non-relativistic quantum mechanics, III. Long range potentials.
- [124] (with V. Enns) Bounds on total cross-sections in atom-atom and atom-ion collisions by geometric methods, *Phys. Rev. Lett.* **44** (1980), 319–321.
- [125] (with V. Enns) Finite total cross-sections in non-relativistic quantum mechanics, *Commun. Math. Phys.* **76** (1980), 177–210.
- [126] Decay of correlations in ferromagnets, *Phys. Rev. Lett.* **44** (1980), 547–549.
- [127] Correlation inequalities and the decay of correlations in ferromagnets, *Commun. Math. Phys.* **77** (1980), 111–126.

- [128] (with M. Aizenman) Local Ward identities and the decay of correlations in ferromagnets, *Commun. Math. Phys.* **77** (1980), 137–143.
- [129] (with M. Aizenman) A comparison of plane rotor Ising models, *Phys. Rev. Lett.* **76A** (1980), 281–282.
- [130] (with K. Miller) Quantum magnetic Hamiltonians with remarkable spectral properties, *Phys. Rev. Lett.* **44** (1980), 1706–1707.
- [131] (with P. Perry and I. Sigal) Absence of singular continuous spectrum in N -body quantum systems, *Bull. Amer. Math. Soc.* **3** (1980), 1019–1024.
- [132] (with P. Perry and I. Sigal) Spectral analysis of multiparticle Schrödinger operators, *Ann. of Math.* **114** (1981), 519–567.
- [133] (with R. Carmona) Pointwise bounds on eigenfunctions and wave packets in N -body quantum systems, V. Lower bounds and path integrals, *Commun. Math. Phys.* **80** (1981), 59–98.
- [134] (with E. Lieb) Pointwise bounds on eigenfunctions and wave packets in N -body quantum systems, VI. Asymptotics in the two cluster region, *Adv. Appl. Math.* **1** (1980), 324–343.
- [135] (with M. Aizenman) Brownian motion and Harnack's inequality for Schrödinger operators, *Commun. Pure Appl. Math.* **35** (1982), 209–273.
- [136] Large time behavior of the L^p norm of Schrödinger semigroups, *J. Funct. Anal.* **40** (1981), 66–83.
- [137] Spectrum and continuum eigenfunctions of Schrödinger operators, *J. Funct. Anal.* **42** (1981), 347–355.
- [138] Convergence in trace ideals, *Proc. Amer. Math. Soc.* **83** (1981), 39–43.
- [139] (with A. Sokal) Making entropy-energy arguments rigorous, *J. Statist. Phys.* **25** (1981), 679–694.
- [140] Pointwise domination of matrices and comparison of \mathcal{L}_p norms, *Pacific J. Math.* **97** (1981), 471–475.
- [141] (with J. Avron) Asymptotics of the gap in the Mathieu equation, *Ann. Phys.* **134** (1981), 76–84.
- [142] (with J. Avron) Transient and recurrent spectrum, *J. Funct. Anal.* **43** (1981), 1–31.
- [143] The rate of falloff of Ising model correlations at large temperatures, *J. Statist. Phys.* **26** (1981), 53–58.
- [144] Absence of continuous symmetry breaking in a one-dimensional n^{-2} model, *J. Statist. Phys.* **26** (1981), 307–311.
- [145] (with S. Friedland) The codimension of degenerate pencils, *Linear Alg. Appl.* **44** (1982), 41–53.
- [146] (with J. Avron) Almost periodic Hill's equation and the rings of Saturn, *Phys. Rev. Lett.* **46** (1981), 1166–1168.
- [147] (with J. Avron) Almost periodic Schrödinger operators, I. Limit periodic potentials, *Commun. Math. Phys.* **82** (1982), 101–120.
- [148] (with J. Avron) Singular continuous spectrum for a class of almost periodic Jacobi matrices, *Bull. Amer. Math. Soc.* **6** (1982), 81–85.
- [149] (with J. Avron) Almost periodic Schrödinger operators, II. The integrated density of states, *Duke Math. J.* **50** (1983), 369–391.
- [150] Hardy and Rellich inequalities in fractional dimension, *J. Oper. Th.* **9** (1983), 143–146.

- [151] Continuity of the density of states in magnetic field, *J. Phys.* **A15** (1982), 2981–2983.
- [152] (with J. Bellissard) Cantor spectrum for the almost Mathieu equation, *J. Funct. Anal.* **48** (1982), 408–419.
- [153] (with L. Yaffe) Rigorous perimeter law upper bound on Wilson loops, *Phys. Lett.* **115B** (1982), 145–147.
- [154] (with E. Harrell) Schrödinger operator methods in the study of a certain non-linear PDE, *Proc. Amer. Math. Soc.* **88** (1983), 376–377.
- [155] (with W. Craig) Subharmonicity of the Lyapunov index, *Duke Math. J.* **50** (1983), 551–560.
- [156] (with N. Corngold and E. Harrell) The mathematical theory of resonances which have exponentially small widths, II, *J. Math. Appl.* **99** (1984), 447–457.
- [157] Some Jacobi matrices with decaying potential and dense point spectrum, *Commun. Math. Phys.* **87** (1982), 253–258.
- [158] Some quantum operators with discrete spectrum but classically continuous spectrum, *Ann. Phys.* **146** (1983), 209–220.
- [159] Nonclassical eigenvalue asymptotics, *J. Funct. Anal.* **53** (1983), 84–98.
- [160] (with M. Hoffmann-Ostenhof and T. Hoffmann-Ostenhof) A multiparticle Coulomb system with bound state at threshold, *J. Phys.* **A16** (1983), 1125–1131.
- [161] Semiclassical analysis of low lying eigenvalues, I. Non-degenerate minima: Asymptotic expansions, *Ann. Inst. H. Poincaré* **38** (1983), 295–307.
- [162] Semiclassical analysis of low lying eigenvalues, II. Tunneling, *Ann. of Math.* **120** (1984), 89–118.
- [163] Instantons, double wells and large deviations, *Bull. Amer. Math. Soc.* **8** (1983), 323–326.
- [164] Equality of the density of states in a wide class of tight binding Lorentzian models, *Phys. Rev.* **B27** (1983), 3859–3860.
- [165] (with F. Bentosela, R. Carmona, P. Duclos, B. Souillard, and R. Weder) Schrödinger operators with electric field and random or deterministic potential, *Commun. Math. Phys.* **88** (1983), 387–397.
- [166] (with W. Craig) Log Hölder continuity of the integrated density of states for stochastic Jacobi matrices, *Commun. Math. Phys.* **90** (1983), 207–218.
- [167] (with J. Avron and W. Craig) Large coupling behavior of the Lyapunov exponent for tight binding one-dimensional random systems, *J. Phys.* **A16** (1983), L209–L211.
- [168] Kotani theory for one-dimensional stochastic Jacobi matrices, *Commun. Math. Phys.* **89** (1983), 227–234.
- [169] (with P. Deift) Almost periodic Schrödinger operators, III. The absolutely continuous spectrum in one dimension, *Commun. Math. Phys.* **90** (1983), 389–411.
- [170] Almost periodic Schrödinger operators, IV. The Maryland model, *Ann. Phys.* **159** (1985), 157–183.
- [171] (with J. Avron and R. Seiler) Homotopy and quantization in condensed matter physics, *Phys. Rev. Lett.* **51** (1983), 51–53.
- [172] Holonomy, the quantum adiabatic theorem and Berry's phase, *Phys. Rev. Lett.* **51** (1983), 2167–2170.
- [173] (with E. B. Davies) Ultracontractivity and the heat kernel for Schrödinger operators and Dirichlet Laplacians, *J. Funct. Anal.* **59** (1984), 335–395.

- [174] Semiclassical analysis of low lying eigenvalues, III. Width of the ground state band in strongly coupled solids, *Ann. Phys.* **158** (1984), 415–420.
- [175] (with E. Lieb, I. Sigal, and W. Thirring) Asymptotic neutrality of large Z ions, *Phys. Rev. Lett.* **52** (1984), 994–996.
- [176] (with F. Delyon and B. Souillard) From power law localized to extended states in a class of one-dimensional disordered systems, *Phys. Rev. Lett.* **52** (1984), 2187–2189.
- [177] Semiclassical analysis of low lying eigenvalues, IV. The flea on the elephant, *J. Funct. Anal.* **63** (1988), 123–136.
- [178] (with F. Delyon and B. Souillard) From power pure point to continuous spectrum in disordered systems, *Ann. Inst. H. Poincaré* **42** (1985), 283–309.
- [179] (with W. Kirsch) Universal lower bounds on eigenvalue splittings for one-dimensional Schrödinger operators, *Commun. Math. Phys.* **97** (1985), 453–460.
- [180] (with W. Kirsch) Lifshitz tails for periodic plus random potentials, *J. Statist. Phys.* **42** (1986), 799–808.
- [181] (with W. Kirsch) Comparison theorems for the gap of Schrödinger operators, *J. Funct. Anal.* **75** (1987), 396–410.
- [182] (with W. Kirsch and S. Kotani) Absence of absolutely continuous spectrum for some one-dimensional random but deterministic Schrödinger operators, *Ann. Inst. H. Poincaré* **42** (1985), 383–406.
- [183] (with J. Avron) Stability of gaps for periodic potentials under variation of the magnetic field, *J. Phys.* **A18** (1985), 2199–2205.
- [184] (with F. Gesztesy and B. Thaller) On the self-adjointness of Dirac operators with anomalous magnetic moment, *Proc. Amer. Math. Soc.* **94** (1985), 115–118.
- [185] Schrödinger semigroups on the scale of Sobolev spaces, *Pacific J. Math.* **122** (1986), 475–480.
- [186] (with E. B. Davies) L^1 -properties of intrinsic Schrödinger semigroups, *J. Funct. Anal.* **65** (1986), 126–146.
- [187] (with M. Taylor and T. Wolff) Some rigorous results for the Anderson model, *Phys. Rev. Lett.* **54** (1985), 1589–1592.
- [188] (with M. Taylor) Harmonic analysis on $SL(2, \mathbb{R})$ and smoothness of the density of states in the one-dimensional Anderson model, *Commun. Math. Phys.* **101** (1985), 1–19.
- [189] (with T. Wolff) Singular continuous spectrum under rank one perturbations and localization for random Hamiltonians, *Commun. Pure Appl. Math.* **39** (1986), 75–90.
- [190] Localization in general one-dimensional random systems, I. Jacobi matrices, *Commun. Math. Phys.* **102** (1985), 327–336.
- [191] Internal Lifshitz tails, *J. Statist. Phys.* **46** (1987), 911–918.
- [192] (with S. Kotani) Localization in general one-dimensional random systems, II. Continuum Schrödinger operators, *Commun. Math. Phys.* **112** (1987), 103–119.
- [193] (with D. Bolle, F. Gesztesy, and H. Grosse) Krein's spectral shift function and Fredholm determinants as efficient methods to study supersymmetric quantum mechanics, *Lett. Math. Phys.* **13** (1987), 127–133.

- [194] (with F. Delyon and B. Souillard) Localization for off-diagonal disorder and for continuous Schrödinger operators, *Commun. Math. Phys.* **109** (1987), 157–165.
- [195] (with D. Bolle, F. Gesztesy, and H. Grosse) Witten index, axial anomaly and Krein's spectral shift function in supersymmetric quantum mechanics, *J. Math. Phys.* **28** (1987), 1512–1525.
- [196] (with E. B. Davies and M. Taylor) L^p spectral theory of Kleinian groups, *J. Funct. Anal.* **78** (1988), 116–136.
- [197] (with F. Gesztesy) Topological invariance of the Witten index, *J. Funct. Anal.* **79** (1988), 91–103.
- [198] (with S. Kotani) Stochastic Schrödinger operators and Jacobi matrices on the strip, *Commun. Math. Phys.* **119** (1988), 403–429.
- [199] (with E. Lieb, I. Sigal, and W. Thirring) Approximate neutrality of large Z ions, *Commun. Math. Phys.* **116** (1988), 635–644.
- [200] (with F. Gesztesy) On a theorem of Deift and Hempel, *Commun. Math. Phys.* **116** (1988), 503–505.
- [201] (with W. Kirsch) Corrections to the classical behavior of the number of bound states of Schrödinger operators, *Ann. Phys.* **183** (1988), 122–130.
- [202] (with H. Englisch, M. Schroder, and W. Kirsch) Density of surface states in discrete models, *Phys. Rev. Lett.* **61** (1988), 1261–1262,
- [203] (with F. Gesztesy, D. Gurarie, H. Holden, M. Klaus, L. Sadun, and P. Vogl) Trapping and cascading of eigenvalues in the large coupling limit, *Commun. Math. Phys.* **118** (1988), 597–634.
- [204] (with J. Avron, L. Sadun, and J. Segert) Topological invariants in Fermi systems with time-reversal invariance, *Phys. Rev. Lett.* **61** (1988), 1329–1332.
- [205] (with J. Avron, L. Sadun, and J. Segert) Chern numbers and Berry's phases in Fermi systems, *Commun. Math. Phys.* **124** (1989), 595–627.
- [206] (with F. Gesztesy) Constructing solutions of the mKdV-equation, *J. Funct. Anal.* **89** (1990) 53–60.
- [207] (with H. Englisch, M. Schroder, and W. Kirsch) Random Hamiltonians ergodic in all but one direction, *Commun. Math. Phys.* **128** (1990), 613–625.
- [208] (with T. Spencer) Trace class perturbations and the absence of absolutely continuous spectrum, *Commun. Math. Phys.* **125** (1989), 113–126.
- [209] (with J. Avron and J. Howland) Adiabatic theorems for dense point spectrum, *Commun. Math. Phys.* **128** (1990), 497–507.
- [210] (with R. Carmona and W. Masters) Relativistic Schrödinger operators: Asymptotic behavior of the eigenfunctions, *J. Funct. Anal.* **91** (1990), 117–142.
- [211] (with J. Avron and P. van Mouche) On the measure of the spectrum for the almost Mathieu operator, *Commun. Math. Phys.* **132** (1990), 103–118.
- [212] (with E. B. Davies) Spectral properties of the Neumann Laplacian of horns, *Geom. Funct. Anal.* **2** (1992), 105–117.
- [213] (with F. Gesztesy and W. Schweiger) Commutation methods applied to the mKdV-equation, *Trans. Amer. Math. Soc.* **324** (1991), 465–525.
- [214] (with F. Gesztesy, H. Holden, and E. Saab) Explicit construction of solutions of the modified Kadomtsev–Petviashvili equation, *J. Funct. Anal.* **98** (1991), 211–228.
- [215] (with F. Gesztesy) A short proof of Zheludev's theorem, *Trans. Amer. Math. Soc.* **335** (1993), 329–340.

- [216] (with J. Avron and R. Seiler) The quantum Hall effect and the relative index for projections, *Phys. Rev. Lett.* **65** (1990), 2185–2188.
- [217] (with R. Hempel and L. Seco) The essential spectrum of Neumann Laplacians on some bounded singular domains, *J. Funct. Anal.* **102** (1991), 448–483.
- [218] Absence of ballistic motion, *Commun. Math. Phys.* **134** (1990), 209–212.
- [219] (with E. B. Davies) L^p -norms of non-critical Schrödinger semigroups, *J. Funct. Anal.* **102** (1991), 95–115.
- [220] Best constants to some operator smoothness estimates, *J. Funct. Anal.* **107** (1992), 66–71.
- [221] The Neumann Laplacian of a jelly roll, *Proc. Amer. Math. Soc.* **114** (1992), 783–785.
- [222] The Weyl transform and L^p functions on phase space, *Proc. Amer. Math. Soc.* **116** (1992), 1045–1047.
- [223] Large time behavior of the heat kernel: On a theorem of Chavel and Karp, *Proc. Amer. Math. Soc.* **118** (1993), 513–514.
- [224] (with F. Gesztesy, H. Holden, and Z. Zhao) On the Toda and Kac–van Moerbeke systems, *Trans. Amer. Math. Soc.* **339** (1993), 849–868.
- [225] (with V. Jaksic and S. Molchanov) Eigenvalue asymptotics of the Neumann Laplacian of regions and manifolds with cusps, *J. Funct. Anal.* **106** (1992), 59–79.
- [226] (with F. Gesztesy and G. M. Graf) The ground state energy of Schrödinger operators, *Commun. Math. Phys.* **150** (1992), 375–384.
- [227] (with G. M. Graf) Asymptotic series for the ground state energy of Schrödinger operators, *J. Funct. Anal.* **112** (1993), 442–446.
- [228] (with J. Avron and R. Seiler) Charge deficiency, charge transport and comparison of dimensions, *Commun. Math. Phys.* **159** (1994), 399–422.
- [229] (with J. Avron and R. Seiler) The index of a pair of projections, *J. Funct. Anal.* **120** (1994), 220–237.
- [230] (with F. Gesztesy, H. Holden, and Z. Zhao) Trace formulae and inverse spectral theory for Schrödinger operators, *Bull. Amer. Math. Soc.* **29** (1993), 250–255.
- [231] (with A. Gordon, V. Jaksic, and S. Molchanov) Spectral properties of random Schrödinger operators with unbounded potentials, *Commun. Math. Phys.* **157** (1993), 23–50.
- [232] Cyclic vectors in the Anderson model, *Rev. Math. Phys.* **6** (1994), 1183–1185.
- [233] (with R. del Rio, S. Jitomirskaya, and N. Makarov) Singular continuous spectrum is generic, *Bull. Amer. Math. Soc.* **31** (1994), 208–212.
- [234] Operators with singular continuous spectrum: I. General operators, *Ann. of Math.* **141** (1995), 131–145.
- [235] (with R. del Rio and N. Makarov) Operators with singular continuous spectrum: II. Rank one operators, *Commun. Math. Phys.* **165** (1994), 59–67.
- [236] (with S. Jitomirskaya) Operators with singular continuous spectrum: III. Almost periodic Schrödinger operators, *Commun. Math. Phys.* **165** (1994), 201–205.
- [237] (with F. Gesztesy and H. Holden) Absolute summability of the trace relation for certain Schrödinger operators, *Commun. Math. Phys.* **168** (1995), 137–161.
- [238] (with F. Gesztesy) Rank one perturbations at infinite coupling, *J. Funct. Anal.* **128** (1995), 245–252.

- [239] (with F. Gesztesy, H. Holden, and Z. Zhao) Higher order trace relations for Schrödinger operators, *Rev. Math. Phys.* **7** (1995), 893–922.
- [240] (with F. Gesztesy, H. Holden, and Z. Zhao) A trace formula for multi-dimensional Schrödinger operators, *J. Funct. Anal.* **141** (1996), 449–465.
- [241] (with F. Gesztesy) The xi function, *Acta Math.* **176** (1996), 49–71.
- [242] L^p norms of the Borel transform and the decomposition of measures, *Proc. Amer. Math. Soc.* **123** (1995), 3749–3755.
- [243] (with R. del Rio and G. Stolz) Stability of spectral types for Sturm–Liouville operators, *Math. Research Lett.* **1** (1994), 437–450.
- [244] (with A. Kiselev) Rank one perturbations with infinitesimal coupling, *J. Funct. Anal.* **130** (1995), 345–356.
- [245] (with A. Hof and O. Knill) Singular continuous spectrum for palindromic Schrödinger operators, *Commun. Math. Phys.* **174** (1995), 149–159.
- [246] Operators with singular continuous spectrum, VI. Graph Laplacians and Laplace–Beltrami operators, *Proc. Amer. Math. Soc.* **124** (1996), 1177–1182.
- [247] (with G. Stolz) Operators with singular continuous spectrum, V. Sparse potentials, *Proc. Amer. Math. Soc.* **124** (1996), 2073–2080.
- [248] Operators with singular continuous spectrum, VII. Examples with borderline time decay, *Commun. Math. Phys.* **176** (1996), 713–722.
- [249] (with F. Gesztesy) Uniqueness theorems in inverse spectral theory for one-dimensional Schrödinger operators, *Trans. Amer. Math. Soc.* **348** (1996), 349–373.
- [250] (with R. del Rio, S. Jitomirskaya, and Y. Last) Operators with singular continuous spectrum, IV. Hausdorff dimensions, rank one perturbations, and localization, *J. d'Analyse Math.* **69** (1996), 153–200.
- [251] (with R. del Rio, S. Jitomirskaya, and Y. Last) What is localization?, *Phys. Rev. Lett.* **75** (1995), 117–119.
- [252] (with F. Gesztesy and G. Teschl) Zeros of the Wronskian and renormalized oscillation theory, *Am. J. Math.* **118** (1996), 571–594.
- [253] Bounded eigenfunctions and absolutely continuous spectra for one-dimensional Schrödinger operators, *Proc. Amer. Math. Soc.* **124** (1996), 3361–3369.
- [254] Some Schrödinger operators with dense point spectrum, *Proc. Amer. Math. Soc.* **125** (1997), 203–208.
- [255] (with W. Bulla, F. Gesztesy, and W. Renger) Weakly coupled bound states in quantum waveguides, *Proc. Amer. Math. Soc.* **125** (1997), 1487–1495.
- [256] (with Y. F. Zhu) The Lyapunov exponents for Schrödinger operators with slowly oscillating potentials, *J. Funct. Anal.* **140** (1996), 541–556.
- [257] (with A. Gordon, S. Jitomirskaya, and Y. Last) Duality and singular continuous spectrum in the almost Mathieu equation, *Acta Math.* **178** (1997), 169–183.
- [258] (with F. Gesztesy and G. Teschl) Spectral deformations of one-dimensional Schrödinger operators, *J. d'Analyse Math.* **70** (1996), 267–324.
- [259] (with R. del Rio) Point spectrum and mixed spectral types for rank one perturbations, *Proc. Amer. Math. Soc.* **125** (1997), 3593–3599.
- [260] (with F. Gesztesy) Inverse spectral analysis with partial information on the potential, I. The case of an a.c. component in the spectrum, *Helv. Phys. Acta* **70** (1997), 66–71.

- [261] (with F. Gesztesy) m -functions and inverse spectral analysis for finite and semi-infinite Jacobi matrices, *J. d'Analyse Math.* **73** (1997), 267–297.
- [262] Spectral averaging and the Krein spectral shift, *Proc. Amer. Math. Soc.* **126** (1998), 1409–1413.
- [263] (with Y. Last) Eigenfunctions, transfer matrices, and absolutely continuous spectrum of one-dimensional Schrödinger operators, *Invent. Math.* **135** (1999), 329–367.
- [264] (with F. Gesztesy) Inverse spectral analysis with partial information on the potential, II. The case of discrete spectrum, *Trans. Amer. Math. Soc.* **352** (2000), 2765–2787.
- [265] (with A. Kiselev and Y. Last) Modified Prüfer and EFGP transforms and the spectral analysis of one-dimensional Schrödinger operators, *Commun. Math. Phys.* **194** (1998), 1–45.
- [266] (with R. del Rio and F. Gesztesy) Inverse spectral analysis with partial information on the potential, III. Updating boundary conditions, *Int. Math. Res. Not.* (1997) **no. 15**, 751–758.
- [267] (with Y. Last) Modified Prüfer and EFGP transforms and deterministic models with dense point spectrum, *J. Funct. Anal.* **154** (1998), 513–530.
- [268] (with F. Gesztesy) On the determination of a potential from three spectra, *Amer. Math. Soc. Transl. (2)* **189** (1999), 85–92.
- [269] (with A. Kiselev and C. Remling) Effective perturbation methods for one-dimensional Schrödinger operators, *J. Diff. Eq.* **151** (1999), 290–312.
- [270] The classical moment problem as a self-adjoint finite difference operator, *Adv. in Math.* **137** (1998), 82–203.
- [271] A new approach to inverse spectral theory, I. Fundamental formalism, *Ann. of Math.* **150** (1999), 1029–1057.
- [272] (with F. Gesztesy) A new approach to inverse spectral theory, II. General real potentials and the connection to the spectral measure, *Ann. of Math.* **152** (2000), 593–643.
- [273] (with A. Ramm) A new approach to inverse spectral theory, III. Short-range potentials *J. d'Analyse Math.* **80** (2000), 319–334.
- [274] A Feynman–Kac formula for unbounded semigroups, *Proc. Intl. Conf. Infinite Dimensional (Stochastic) Analysis and Quantum Physics, Leipzig 1999*, Canadian Math. Soc. Conf. Proc. **28** (2000), 317–321.
- [275] (with F. Gesztesy) On local Borg–Marchenko uniqueness results *Commun. Math. Phys.* **211** (2000), 273–287.
- [276] Resonances in one dimension and Fredholm determinants, *J. Funct. Anal.* **178** (2000), 396–420.
- [277] (with D. Hundertmark) An optimal L^p -bound on the Krein spectral shift function, *J. d'Analyse Math.* **87** (2002), 199–208.
- [278] (with W. Kirsch) Approach to equilibrium for a forced Burgers equation, *J. Evol. Eqns.* **1** (2001), 411–419.
- [279] (with A. Kiselev and Y. Last) Stability of singular spectral types under decaying perturbations, *J. Funct. Anal.* **198** (2003), 1–27.
- [280] (with D. Hundertmark) Lieb–Thirring inequalities for Jacobi matrices, *J. Approx. Theory* **118** (2002), 106–130.
- [281] (with R. Killip) Sum rules for Jacobi matrices and their applications to spectral theory, *Ann. of Math.* **158** (2003), 253–321.

- [282] (with A. Zlatoš) Sum rules and the Szegő condition for orthogonal polynomials on the real line, *Commun. Math. Phys.* **242** (2003), 393–423.
- [283] (with S. Denisov) Zeros of orthogonal polynomials on the real line, *J. Approx. Theory* **121** (2003), 357–364.
- [284] (with D. Damanik and D. Hundertmark) Bound states and the Szegő condition for Jacobi matrices and Schrödinger operators, *J. Funct. Anal.* **205** (2003), 357–379.
- [285] (with D. Damanik, D. Hundertmark, and R. Killip) Variational estimates for discrete Schrödinger operators with potentials of indefinite sign, *Commun. Math. Phys.* **238** (2003), 545–562.
- [286] (with D. Hundertmark) A diamagnetic inequality for semigroup differences, *J. Reine Angew. Math.* **571** (2004), 107–130.
- [287] The Golinskii–Ibragimov method and a theorem of Damanik–Killip, *Int. Math. Res. Not.* (2003), **no. 36**, 1973–1986.
- [288] A canonical factorization for meromorphic Herglotz functions on the unit disk and sum rules for Jacobi matrices, *J. Funct. Anal.* **214** (2004), 396–409.
- [289] (with F. Gesztesy) Connectedness of the isospectral manifold for one-dimensional half-line Schrödinger operators, *J. Statist. Phys.* **116** (2004), 361–365.
- [290] Ratio asymptotics and weak asymptotic measures for orthogonal polynomials on the real line, *J. Approx. Theory* **126** (2004), 198–217.
- [291] (with D. Damanik and R. Killip) Necessary and sufficient conditions in the spectral theory of Jacobi matrices and Schrödinger operators, *Int. Math. Res. Not.* (2004), **no. 22**, 1087–1097.
- [292] (with V. Totik) Limits of zeros of orthogonal polynomials on the circle, *Math. Nachr.* **278** (2005), 1615–1620.
- [293] On a theorem of Kac and Gilbert, *J. Funct. Anal.* **223** (2005), 109–115.
- [294] Orthogonal polynomials on the unit circle: New results, *Int. Math. Res. Not.* (2004), **no. 53**, 2837–2880.
- [295] (with D. Damanik and R. Killip) Schrödinger operators with few bound states, *Commun. Math. Phys.* **258** (2005), 741–750.
- [296] (with A. Zlatoš) Higher-order Szegő theorems with two singular points, *J. Approx. Theory* **134** (2005), 114–129.
- [297] Aizenman's theorem for orthogonal polynomials on the unit circle, *Const. Approx.* **23** (2006), 229–240.
- [298] Fine structure of the zeros of orthogonal polynomials, I. A tale of two pictures, to appear in *Proc. Constructive Functions Tech-04*.
- [299] Fine structure of the zeros of orthogonal polynomials, II. OPUC with competing exponential decay, *J. Approx. Theory* **135** (2005), 125–139.
- [300] Fine structure of the zeros of orthogonal polynomials, III. Periodic recursion coefficients, *Commun. Pure Appl. Math.* **59** (2005), 1042–1062.
- [301] (with D. Damanik) Jost functions and Jost solutions for Jacobi matrices, I. A necessary and sufficient condition for Szegő asymptotics, *Invent. Math.* **165** (2006), 1–50.
- [302] (with D. Damanik) Jost functions and Jost solutions for Jacobi matrices, II. Decay and analyticity, *Int. Math. Res. Not.* **2006**, Article ID 19396, 32 pages, 2006.
- [303] Meromorphic Szegő functions and asymptotic series for Verblunsky coefficients, *Acta Math.* **195** (2005), 267–285.

- [304] (with Y. Last) The essential spectrum of Schrödinger, Jacobi, and CMV operators, *J. d'Analyse Math.* **98** (2006), 183–220.
- [305] Meromorphic Jost functions and asymptotic series for Jacobi parameters, to appear in *Funct. Anal. Appl.*
- [306] (with R. Killip) Sum rules and spectral measures of Schrödinger operators with L^2 potentials, to appear in *Ann. of Math.*
- [307] (with E. B. Davies) Eigenvalue estimates for non-normal matrices and the zeros of random orthogonal polynomials on the unit circle, *J. Approx. Theory* **141** (2006), 189–213.
- [308] Rank one perturbations and the zeros of paraorthogonal polynomials on the unit circle, to appear in *J. Math. Anal. Appl.*
- [309] (with Y. Last) Fine structure of the zeros of orthogonal polynomials, IV. A priori bounds and clock behavior, submitted.

Books and Monographs

- [1] *Quantum Mechanics for Hamiltonians Defined by Quadratic Forms*, Princeton Series in Physics, Princeton University Press, 1971.
- [2] (with M. Reed) *Methods of Modern Mathematical Physics, Vol. I: Functional Analysis*, Academic Press, 1972.
- [3] (with M. Reed) *Methods of Modern Mathematical Physics, Vol. II: Fourier Analysis, Self-Adjointness*, Academic Press, 1975.
- [4] *The $P(\phi)_2$ Euclidean (Quantum) Field Theory*, Princeton Series in Physics, Princeton University Press, 1974.
- [5] (with M. Reed) *Methods of Modern Mathematical Physics, Vol. III: Scattering Theory*, Academic Press, 1978.
- [6] (editor, with E. Lieb) *Studies in Mathematical Physics, Essays in Honor of Valentine Bargmann*, Princeton University Press, 1976.
- [7] (with M. Reed) *Methods of Modern Mathematical Physics, Vol. IV: Analysis of Operators*, Academic Press, 1977.
- [8] *Trace Ideals and Their Applications*, Cambridge University Press, 1979; second edition, Mathematical Surveys and Monographs, Vol. 120, American Mathematical Society, 2005.
- [9] *Functional Integration and Quantum Physics*, Academic Press, 1979; second edition, AMS Chelsea Publishing, 2005.
- [10] (with H. Cycon, R. Froese and W. Kirsch) *Schrödinger Operators with Application to Quantum Mechanics and Global Geometry*, Springer, 1987.
- [11] *The Statistical Mechanics of Lattice Gases, Vol. I*, Princeton University Press, 1993.
- [12] *Representations of Finite and Compact Groups*, Graduate Studies in Mathematics, Vol. 10, American Mathematical Society, 1996.
- [13] *Orthogonal Polynomials on the Unit Circle, Part 1: Classical Theory*, AMS Colloquium Publications, American Mathematical Society, 2005.
- [14] *Orthogonal Polynomials on the Unit Circle, Part 2: Spectral Theory*, AMS Colloquium Publications, American Mathematical Society, 2005.

Conference Proceedings, Summer School Notes,
Review Articles

- [1] Analyticity in the coupling constant and the Padé approximation, *Proc. 8th East. Theo. Phys. Conf.* (1969), 167–195.
- [2] The anharmonic oscillator: A singular perturbation theory, *Cargese Lectures in Theoretical Physics 5* (ed. D. Bessis), Gordon and Breach, 1972, 383–414.
- [3] Studying spatially cutoff two-dimensional Bose field theories, *Field Theory and Statistical Mechanics* (eds. R. N. Sen and C. Weil), Keter Press, 1972, 197–224.
- [4] Topics in functional analysis, *Math. of Contemporary Physics* (ed. R. Streater), Academic Press, 1972, 18–76.
- [5] Perturbation theory and coupling constant analyticity in two-dimensional field theories, *Fundamental Interactions in Physics and Astronomy*, Plenum, 1973, 120–136.
- [6] The Glimm–Jaffe ϕ -bound: A Markov proof, *Constructive Quantum Theory* (eds. G. Velo and A. S. Wightman), Springer, 1973, 125–131.
- [7] Bose field theory as statistical mechanics, III. The classical Ising approximation, *Constructive Quantum Field Theory*, (eds. G. Velo and A. S. Wightman), Springer, 1973, 290–297.
- [8] Approximation of path integrals and Markov fields by spin systems, *Proc. 1974 Intl. Cong. Math.*, 1975, 399–402.
- [9] Bose quantum field theory as an Ising ferromagnet: Recent developments, *Intl. Symp. on Math. Problems in Theoretical Physics* (ed. H. Araki), Springer, 1975, 543–553.
- [10] An introduction to the self-adjointness and spectral analysis of Schrödinger operators, *The Schrödinger Equation* (eds. W. Thirring and P. Urban), Springer, 1977 (*Acta Phys. Aus. Suppl.* **17** Vienna, 19–42).
- [11] On the number of bound states of two-body Schrödinger operators: A review, *from* [6], pp. 305–326.
- [12] Quantum dynamics: From automorphism to Hamiltonian, *from* [6], pp. 327–349.
- [13] Classical boundary conditions as a tool in quantum physics, *Adv. in Math.* **30** (1978), 268–281.
- [14] New rigorous existence theorems for phase transitions in model systems, *Proc. IUPAP Stat. Mech. Meeting*, Haifa, 1977.
- [15] Resonances and complex scaling: A rigorous overview, *Intl. J. Quant. Chem.* **14** (1978), 529–542.
- [16] An overview of rigorous scattering theory, *Atomic Scattering Theory: Mathematical and Computational Aspects* (ed. J. Nuttall), Univ. of Western Ontario, 1978, 1–24.
- [17] Identifying the classical limit of a quantum spin system, *Colloq. Math. Soc. Bolyai* **27** (1979), 989–1001.
- [18] Lattice systems, *Encyclopedia of Statistical Sciences 4* (eds. S. Kotz and N. Johnson), Wiley, 1983, 519–522.
- [19] Quantum physics and functional integration, *Encyclopedia of Statistical Sciences 7* (eds. Kotz and Johnson), Wiley, 1986, 453–456.
- [20] Feynman integrals, *McGraw–Hill Encyclopedia of Science and Technology*, fifth edition, 1982, 391–392.
- [21] Schrödinger semigroups, *Bull. Amer. Math. Soc.* **7** (1982), 447–526.

- [22] (with V. Enss) Total cross sections in non-relativistic scattering theory, *Quantum Mechanics in Mathematics, Chemistry and Physics* (eds. K. Gustafson and W. Reinhardt), Plenum, 1981, 1–26.
- [23] Large orders and summability of eigenvalue perturbation theory: A mathematical overview, *Intl. J. Quant. Chem.* **21** (1982), 3–25.
- [24] Spectral analysis of multiparticle Schrödinger operators, *Spectral Theory of Differential Operators* (eds. I. Knowles and R. T. Lewis), North Holland, 1981, 369–370.
- [25] m -functions and the absolutely continuous spectrum of one-dimensional almost periodic Schrödinger operators, *Differential Equations* (eds. I. Knowles and R. T. Lewis), North Holland, 1984, 519.
- [26] Almost periodic Schrödinger operators: A review, *Adv. Appl. Math.* **3** (1982), 463–490.
- [27] Fifteen problems in mathematical physics, *Perspectives in Mathematics*, Oberwolfach Anniversary Volume, Birkhäuser, 1984, 423–454.
- [28] Boundedness of continuum eigenfunctions and their relation to spectral problems, *Linear and Complex Analysis Problem Book* (eds. V. Havin, S. Khrushchev, and N. Nikol'skii), Lecture Notes in Math., **1043**, Springer, 1984.
- [29] (with E. B. Davies) Ultracontractive semigroups and some problems in analysis, *Aspects of Mathematics and its Applications* (ed. J. Barroso), Elsevier, 1986.
- [30] (with B. Souillard) Franco-American meeting on the mathematics of random and almost periodic potentials, *J. Statist. Phys.* **36** (1984), 273–288.
- [31] Some aspects of the theory of Schrödinger operators, *Schrödinger Operators, Como, 1984* (ed. S. Graffi), Lecture Notes in Math. **159**, Springer, 1985.
- [32] Lifshitz tails for the Anderson model, *J. Statist. Phys.* **38** (1985), 65–76.
- [33] Regularity of the density of states for stochastic Jacobi matrices: A review, *IMA Conf. Proc., Random Media* **7** (1987), 245–266.
- [34] Schrödinger operators with random and almost periodic potentials, *Recent Developments in Mathematical Physics* (eds. H. Mitter and L. Pittner), Springer, 1987.
- [35] (with E. B. Davies and L. Gross) Hypercontractivity: A bibliographic review, *Ideas and Methods in Quantum and Statistical Mechanics*, **2** (eds. S. Albeverio et al.), Cambridge University Press, 1992, 370–289.
- [36] Fifty years of eigenvalue perturbation theory, *Bull. Amer. Math. Soc.* **24** (1991), 303–319.
- [37] The Scott correction and the quasi-classical limit, *Asterisque* **210** (1992), 295–302.
- [38] Spectral analysis of rank one perturbations and applications, *Proc. Mathematical Quantum Theory, II: Schrödinger Operators* (eds. J. Feldman, R. Froese and L. Rosen) CRM Proc. Lecture Notes **8** (1995), 109–149.
- [39] Schrödinger operators in the twentieth century, *J. Math. Phys.* **41** (2000), 3523–3555.
- [40] Schrödinger operators in the twenty-first century, *Mathematical Physics 2000* (eds. A. Fokas et al.), Imperial College Press, London, 283–288.
- [41] Sturm oscillation and comparison theorems, *Sturm–Liouville Theory, Past and Present* (eds. W. O. Amrein, A. M. Hinz, and D. B. Pearson), Birkhäuser, 2005, 29–44.

- [42] Analogs of the m -function in the theory of orthogonal polynomials on the unit circle, *J. Comp. Appl. Math.* **171** (2004), 411–424.
- [43] The sharp form of the strong Szegő theorem, *Contemp. Math.* **387** (2005), 253–275.
- [44] Ed Nelson's work in quantum theory, *Diffusion, Quantum Theory, and Radically Elementary Mathematics* (ed. W. G. Faris) Mathematical Notes **47**, Princeton University Press, 2006, 75–93.
- [45] OPUC on one foot, *Bull. Amer. Math. Soc.* **42** (2005), 431–460.
- [46] CMV matrices: Five years after, to appear in *Proc. W. D. Evans 65th Birthday Conference*.
- [47] Orthogonal polynomials with exponentially decaying recursion coefficients, submitted.
- [48] Fine structure of the zeros of orthogonal polynomials: A review, submitted.

Titles in This Series

- 76 **Fritz Gesztesy (Managing editor), Percy Deift, Cherie Galvez, Peter Perry, and Wilhelm Schlag, Editors**, Spectral theory and mathematical physics: A Festschrift in honor of Barry Simon's 60th birthday, Parts 1 and 2 (California Institute of Technology, Pasadena, CA, March 27–31, 2006)
- 75 **Solomon Friedberg (Managing editor), Daniel Bump, Dorian Goldfeld, and Jeffrey Hoffstein, Editors**, Multiple Dirichlet series, automorphic forms, and analytic number theory (Bretton Woods, New Hampshire, July 11–14, 2005)
- 74 **Benson Farb, Editor**, Problems on mapping class groups and related topics, 2006
- 73 **Mikhail Lyubich and Leon Takhtajan, Editors**, Graphs and patterns in mathematics and theoretical physics (Stony Brook University, Stony Brook, NY, June 14–21, 2001)
- 72 **Michel L. Lapidus and Machiel van Frankenhuysen, Editors**, Fractal geometry and applications: A jubilee of Benoît Mandelbrot, Parts 1 and 2 (San Diego, California, 2002 and École Normale Supérieure de Lyon, 2001)
- 71 **Gordana Matić and Clint McCrory, Editors**, Topology and Geometry of Manifolds (University of Georgia, Athens, Georgia, 2001)
- 70 **Michael D. Fried and Yasutaka Ihara, Editors**, Arithmetic fundamental groups and noncommutative algebra (Mathematical Sciences Research Institute, Berkeley, California, 1999)
- 69 **Anatole Katok, Rafael de la Llave, Yakov Pesin, and Howard Weiss, Editors**, Smooth ergodic theory and its applications (University of Washington, Seattle, 1999)
- 68 **Robert S. Doran and V. S. Varadarajan, Editors**, The mathematical legacy of Harish-Chandra: A celebration of representation theory and harmonic analysis (Baltimore, Maryland, 1998)
- 67 **Wayne Raskind and Charles Weibel, Editors**, Algebraic K -theory (University of Washington, Seattle, 1997)
- 66 **Robert S. Doran, Ze-Li Dou, and George T. Gilbert, Editors**, Automorphic forms, automorphic representations, and arithmetic (Texas Christian University, Fort Worth, 1996)
- 65 **M. Giaquinta, J. Shatah, and S. R. S. Varadhan, Editors**, Differential equations: La Pietra 1996 (Villa La Pietra, Florence, Italy, 1996)
- 64 **G. Ferreyra, R. Gardner, H. Hermes, and H. Sussmann, Editors**, Differential geometry and control (University of Colorado, Boulder, 1997)
- 63 **Alejandro Adem, Jon Carlson, Stewart Priddy, and Peter Webb, Editors**, Group representations: Cohomology, group actions and topology (University of Washington, Seattle, 1996)
- 62 **János Kollár, Robert Lazarsfeld, and David R. Morrison, Editors**, Algebraic geometry—Santa Cruz 1995 (University of California, Santa Cruz, July 1995)
- 61 **T. N. Bailey and A. W. Knap, Editors**, Representation theory and automorphic forms (International Centre for Mathematical Sciences, Edinburgh, Scotland, March 1996)
- 60 **David Jerison, I. M. Singer, and Daniel W. Stroock, Editors**, The legacy of Norbert Wiener: A centennial symposium (Massachusetts Institute of Technology, Cambridge, October 1994)
- 59 **William Arveson, Thomas Branson, and Irving Segal, Editors**, Quantization, nonlinear partial differential equations, and operator algebra (Massachusetts Institute of Technology, Cambridge, June 1994)
- 58 **Bill Jacob and Alex Rosenberg, Editors**, K -theory and algebraic geometry: Connections with quadratic forms and division algebras (University of California, Santa Barbara, July 1992)
- 57 **Michael C. Cranston and Mark A. Pinsky, Editors**, Stochastic analysis (Cornell University, Ithaca, July 1993)
- 56 **William J. Haboush and Brian J. Parshall, Editors**, Algebraic groups and their generalizations (Pennsylvania State University, University Park, July 1991)

TITLES IN THIS SERIES

- 55 **Uwe Jannsen, Steven L. Kleiman, and Jean-Pierre Serre, Editors**, *Motives* (University of Washington, Seattle, July/August 1991)
- 54 **Robert Greene and S. T. Yau, Editors**, *Differential geometry* (University of California, Los Angeles, July 1990)
- 53 **James A. Carlson, C. Herbert Clemens, and David R. Morrison, Editors**, *Complex geometry and Lie theory* (Sundance, Utah, May 1989)
- 52 **Eric Bedford, John P. D'Angelo, Robert E. Greene, and Steven G. Krantz, Editors**, *Several complex variables and complex geometry* (University of California, Santa Cruz, July 1989)
- 51 **William B. Arveson and Ronald G. Douglas, Editors**, *Operator theory/operator algebras and applications* (University of New Hampshire, July 1988)
- 50 **James Glimm, John Impagliazzo, and Isadore Singer, Editors**, *The legacy of John von Neumann* (Hofstra University, Hempstead, New York, May/June 1988)
- 49 **Robert C. Gunning and Leon Ehrenpreis, Editors**, *Theta functions – Bowdoin 1987* (Bowdoin College, Brunswick, Maine, July 1987)
- 48 **R. O. Wells, Jr., Editor**, *The mathematical heritage of Hermann Weyl* (Duke University, Durham, May 1987)
- 47 **Paul Fong, Editor**, *The Arcata conference on representations of finite groups* (Humboldt State University, Arcata, California, July 1986)
- 46 **Spencer J. Bloch, Editor**, *Algebraic geometry – Bowdoin 1985* (Bowdoin College, Brunswick, Maine, July 1985)
- 45 **Felix E. Browder, Editor**, *Nonlinear functional analysis and its applications* (University of California, Berkeley, July 1983)
- 44 **William K. Allard and Frederick J. Almgren, Jr., Editors**, *Geometric measure theory and the calculus of variations* (Humboldt State University, Arcata, California, July/August 1984)

For a complete list of titles in this series, visit the
AMS Bookstore at www.ams.org/bookstore/.

ISBN 0-8218-4249-8



9 780821 842492

PSPUM/76.2