Analysis of random matrix actions related to random operators in quantum physics

Der Naturwissenschaftlichen Fakultät
der Friedrich-Alexander-Universität Erlangen-Nürnberg
zur
Erlangung des Doktorgrades

Friedrich-Alexander-Universität
Erlangen-Nürnberg

vorgelegt von
Christian Hermann Sadel
aus Bamberg
Der Wissenschaftler findet seine Belohnung in dem, was Poincaré die Freude am Verstehen nennt, nicht in den Anwendungsmöglichkeiten seiner Erfindung.

*Albert Einstein*

gewidmet meinen Eltern
Preface

This dissertation summarizes my work of the last three years on the analysis on random Schrödinger and also random Dirac operators. The whole subject of considering random Schrödinger operators in mathematical physics started with some heuristic description of what is today called Anderson localization and delocalization by a paper of P. W. Anderson [3] in 1958. In Chapter 1 the main terminologies are introduced by analyzing the simplest model exhibiting Anderson localization, the so called one-dimensional Anderson model. A main mathematical tool will be the study of random products of the transfer matrices, therefore I will give some background on the theory of products of random matrices in Chapter 2. The main quantities for the description of these random products are the so called Lyapunov exponents. As they are very important for the whole work I felt it better to place this chapter at the beginning rather than to put it into the appendix. In Chapter 3 I will relate the transfer matrix in some physics calculation to the scattering matrix. This will give another view on the Lyapunov exponents. These first three chapters should give some background on the subject and work I have done in collaboration with Prof. Schulz-Baldes which follows in the consecutive chapters. Chapter 4 contains the work on the scaling of the Lyapunov exponent at a so called band edge which has already been published [72]. In Chapter 5 I have done similar calculations for a model with strongly mixing potential which is a weaker condition than the i.i.d. potential in the models before. This is also published [73]. In Chapter 6 I generalize some ideas used in Chapter 4 to a more abstract setting which can be used for random operators on strips. Such an example is given in Chapter 7. Finally Chapter 8 contains work on the spectrum of random Dirac operators. The work of the Chapters 6 and 8 are available online as preprints on arXiv, [71] and [70]. They were sent in for publication.

Each chapter is essentially written in a self contained way. However, I recommend to read them in the order they are written down. Chapters 4 to 8 use general facts mentioned in Chapter 2, Chapter 5 generalizes parts of Chapter 4. Chapter 7 depends on Chapter 6 which generalizes parts of Chapter 4. Chapter 8 is mostly independent of the chapters before.

Generally in this work, the meaning of a variable is only fixed within a chapter and some variables may have different meanings in different chapters. However, for important mathematical objects the same symbols are used throughout the thesis.

I assume the reader to be familiar with basic notions in functional analysis, e.g. spectral decomposition of operators as well as pure point, singular and absolutely continuous spectrum. The well established theory can be found in [66].
Before starting with the introduction I give a more mathematical summary in german and english on the next pages.

Acknowledgement

I am thankful to my advisor, Prof. H. Schulz-Baldes who offered me the PhD position right after the Diploma, as well as to Prof. A. Knauf and Prof. A. Bendikov for several discussions. I also like to thank the DFG for funding my work in the past three years and supporting my attendance at several conferences. Furthermore I am grateful to the Cambridge Philosophical Society who supported part of my visit at the Newton Institute in Cambridge. Last but not least I want to thank Prof. J. Bellissard for inviting me to the Georgia Institute of Technology for several months.
Contents

Preface 3

Acknowledgement 4

I Zusammenfassung 9

II Summary 15

1 Introduction on the Anderson model 21
   1.1 The phenomenon of Anderson localization 21
   1.2 Basic quantum mechanics 22
   1.3 Anderson localization in mathematical terms 24
   1.4 Pure point spectrum for the 1D Anderson model 27
      1.4.1 Transfer matrices and Lyapunov exponent 29
      1.4.2 Proof of pure point spectrum 31
   1.5 Integrated Density of States and Rotation number 34
   1.6 Scaling of the Lyapunov exponent 38

2 Theory on products of random matrices 41
   2.1 Lyapunov exponents 41
   2.2 Multiplicative Ergodic Theorem 43
   2.3 Distinct Lyapunov exponents 45
   2.4 The case of symplectic matrices 47

3 Transfer matrix and Scattering matrix 49
   3.1 Introduction and model 49
   3.2 Channels and scattering states 51
   3.3 Normal forms of the transfer matrices 53
   3.4 Reduced transfer matrix 54
   3.5 Scattering operator and scattering matrix 56
   3.6 Divided cable as reference operator 58
   3.7 Relation of scattering matrix and transfer matrix 61
   3.8 Ideal leads 64
4 Scaling diagram for the Lyapunov exponent at a bandedge
  4.1 Main result and short overview ........................................ 69
  4.2 Normal form of transfer matrix at a band edge ..................... 73
  4.3 Definition of anomalies ................................................... 76
  4.4 Phase shift dynamics ...................................................... 77
  4.5 First order anomalies ..................................................... 80
  4.6 Fokker-Planck operator of a second order anomaly ................. 82
  4.7 Birkhoff sums at second order anomalies ............................. 86
  4.8 Application to a band edge .............................................. 89

5 Perturbation Theory with strongly mixing potential
  5.1 Introduction ................................................................. 91
  5.2 Set-up and results ......................................................... 92
  5.3 Anomalies at band center and band edge ................................ 96
  5.4 Phase shift dynamics ...................................................... 98
  5.5 From Birkhoff-like sums to Birkhoff sums ............................ 99
  5.6 Oscillatory sums away from band center and edges .................. 103
  5.7 Fokker-Planck operator for drift-diffusion .......................... 103
  5.8 Application to the band center and band edge ....................... 106
  5.9 Bound on the quantum dynamics ....................................... 107

6 Random Lie Group actions on compact manifolds:
  a perturbative analysis ...................................................... 109
  6.1 Main results and discussion ............................................ 109
  6.2 Fokker-Planck operator and its properties .......................... 115
  6.3 Control of Birkhoff sum in the case $R = 1, P = 0$ ................ 120
  6.4 Extension to lowest order rotations .................................. 121

7 Random phase approximation for randomly coupled wires
  7.1 Main result and short overview ....................................... 129
  7.2 Lyapunov exponents ....................................................... 131
  7.3 Evaluation of Birkhoff sums ............................................ 133
  7.4 Divergence of induced vector fields .................................. 135
  7.5 Volume measure to lowest order ....................................... 136
  7.6 Calculations for the Hamiltonian $H^{R}_{A,w}$ ...................... 138
     7.6.1 Lyapunov exponents .............................................. 138
     7.6.2 Haar measure to lowest order .................................... 139

8 Random Dirac operators with time-reversal symmetry
  8.1 Introduction ................................................................. 141
  8.2 Weyl-Titchmarsh matrices ............................................... 143
  8.3 Green’s function and spectral analysis ................................ 149
  8.4 Stochastic Dirac operators ............................................. 153
  8.5 Kotani theory ............................................................... 157
<table>
<thead>
<tr>
<th>Contents</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.6 The Lyapunov spectrum</td>
<td>163</td>
</tr>
<tr>
<td>8.7 Absence of singular spectrum</td>
<td>170</td>
</tr>
<tr>
<td>A Appendix</td>
<td>177</td>
</tr>
<tr>
<td>A.1 Inhomogeneous singular first order ODE</td>
<td>177</td>
</tr>
<tr>
<td>A.2 Vector valued random variables</td>
<td>180</td>
</tr>
<tr>
<td>A.3 Hörmander operators</td>
<td>181</td>
</tr>
<tr>
<td>A.4 Sobolev spaces on compact manifolds</td>
<td>182</td>
</tr>
<tr>
<td>A.5 Fourier series on compact Lie groups</td>
<td>184</td>
</tr>
<tr>
<td>A.6 Isotropic flag manifold as quotient of compact groups</td>
<td>186</td>
</tr>
<tr>
<td>A.7 Divergence of vector fields on quotients of compact groups</td>
<td>188</td>
</tr>
<tr>
<td>A.8 Measurability of ((\mu, \nu) : \nu \text{ is a.c. w.r.t. } \mu)</td>
<td>189</td>
</tr>
<tr>
<td>Bibliography</td>
<td>193</td>
</tr>
<tr>
<td>Curriculum Vitae</td>
<td>199</td>
</tr>
<tr>
<td>Declaration</td>
<td>200</td>
</tr>
</tbody>
</table>
I Zusammenfassung


Im ersten Kapitel bespreche ich zur Einführung aller wichtigen mathematischen Begriffe das sogenannte eindimensionale Anderson Modell. Hier betrachtet man einen zufälligen Schrödinger Operator auf $\ell^2(\mathbb{Z})$ mit einem zufälligen Potential $V(\omega_n)$ gegeben durch

$$ (H_\omega \Psi)_n = -\Psi_{n-1} - \Psi_{n+1} + \lambda V(\omega_n) \Psi_n . $$

Die reelle Konstante $\lambda$ ist eine sogenannte Kopplungskonstante, die zur störungstheoretischen Analyse später eingeführt wird und den Zufall koppelt. Für ergodische Potentiale (bzgl. Translation auf $\mathbb{Z}$) ist bekannt, dass das Spektrum und seine Bestandteile (Punktspektrum, singulär und absolut stetiges Spektrum) nicht zufällig sind [19, 64]. Wir betrachten als Beispiel eines solchen Potentials den Fall, wenn die Variablen $V(\omega_n)$ unabhängig identisch verteilt sind. Ferner nehmen wir an, dass für den Erwartungswert $E(V) = 0$ gilt. Ohne zufälliges Potential ($\lambda = 0$) hat man absolut stetiges Spektrum im Intervall $[-2, 2]$, sobald man jedoch auch nur ein kleines zufälliges Potential hat ($\lambda > 0$), erhält man sofort reines Punktspektrum mit exponentiell abfallenden Eigenfunktionen [33, 64]. Das ist die mathematische Definition von starker Anderson Lokalisierung. Das wichtigste Werkzeug im Beweis sind die zufälligen Transformatoren $T^E_{\omega_n} = \begin{pmatrix} \lambda V(\omega_n) - E & -1 \\ 1 & 0 \end{pmatrix}$, die man erhält, wenn man die stationäre Schrödinger Gleichung $H_\omega \Psi = E \Psi$ in der Form $\begin{pmatrix} \Psi_{n+1} \\ \Psi_n \end{pmatrix} = T^E_n \begin{pmatrix} \Psi_n \\ \Psi_{n-1} \end{pmatrix}$ schreibt. Das asymptotische Verhalten der zufälligen Produkte $T^E_N(\omega) = T^E_{\omega_n} T^E_{\omega_{n-1}} \cdots T^E_{\omega_1}$ dieser Matrizen wird durch einen Lyapunovexponenten $\gamma = \lim_{N \to \infty} \frac{1}{N} \log(\|T^E_N(\omega)\|)$, der im allgemeinen von der Energie $E$ und der Unordnung $\lambda$ abhängt, beschrieben. Der Lyapunovexponent $\gamma$ hat zugleich physikalische Bedeutung, da er gleich der inversen Lokalisierungsänge ist. Zur Berechnung kann man das zufällige dynamische System auf dem projektiven Raum $\mathbb{R}P(1)$, das durch die Wirkung der Transformatoren gegeben ist, betrachten.
ZUSAMMENFASSUNG

Genauer gesagt stellt sich heraus, dass der Lyapunovexponent als Birkhoff Summe geschrieben werden kann. Gleiches gilt auch für die sogenannte integrierte Zustandsdichte (kurz IDS für 'integrated density of states').

Da eine exakte Berechnung nicht so leicht möglich ist, wird für kleine zufällige Potentiale, d.h. kleines $\lambda$, eine störungstheoretische Rechnung gemacht, die erlaubt, z.B. den Lyapunovexponenten zur niedrigsten Ordnung auszurechnen. Dadurch erhält man Thouless Formel [87, 64] für Energieen im Band und weg vom Bandzentrum, $|E| < 2, E \neq 0$, die aussagt

$$\gamma = \lambda^2 \frac{E(V^2)}{2(4 - E^2)} + \mathcal{O}(\lambda^3).$$

Für $E = 0$ und an den Bandrändern $|E| = 2$ ist die störungstheoretische Analyse etwas komplizierter. Das Bandzentrum wurde in [77] behandelt, der Bandrand war Bestandteil meiner Diplomarbeit und wird im vierten Kapitel in einer allgemeineren Form behandelt.


Im strikt eindimensionalen Modell beschreibt der Lyapunovexponent die Lokalisierungs-länge und hat daher auch mit der Leitfähigkeit zu tun. Um die Bedeutung der Lyapunovexponenten für Modelle auf Streifen zu erkennen, wird im dritten Kapitel die Relation der Transfermatrix mit der sog. Streumatrix aufgezeigt. Dieses Kapitel ist rechentechnisch eher im physikalischen Stil geschrieben und soll als zusätzliches Hintergrundwissen dienen. Dabei wird eine bestimmte Interpretation der Transfermatrizen nahegelegt. Man nehme einen endlichen Block und verbinde ihn mit sogenannten idealen Kabeln in denen die nicht normierbaren 'Eigenzustände' stehende Wellen sind. Betrachtet man nun sogenannte Streuzustände, dann verbindet die Transfermatrix des endlichen Blocks die Koeffizienten für die stehenden Wellen im linken Kabel mit denen im rechten Kabel. Die Streumatrix hingegen bildet die einlaufenden Wellen auf die auslaufenden Wellen ab [5]. Genaugenommen, wenn der Vektor $a^+$ die Koeffizienten der nach rechts laufenden Wellen im idealen Kabel links, $a^-$ die nach links laufenden links, $b^+$ die nach rechts laufenden rechts und $b^-$ die nach links laufenden rechts, beschreibt, dann gilt nach einer entsprechenden Konjugation für die Transfermatrix $T(a^+_{a^-}) = (b^+_{b^-})$. Nun beschreiben $a^+$ und $b^-$ einlaufende Wellen und daher gilt für die Streumatrix $S(a^+_{b^-}) = (a^-_{b^+})$. Es zeigt sich, dass eine eindeutige unitäre Matrix $S$ existiert, die diese Beziehung erfüllt. Aus der Streumatrix erhält man die sogenannten Transmissionskoeffizienten welche für den kohärenrenten Ladungstransport wichtig sind. Für einen Streifen der Breite $L$ gibt es jeweils $L$ nach links und nach rechts laufende Moden und man erhält $L$ Transmissionskoeffizienten $t_1 \geq \ldots \geq t_L$ deren Summe die Landauer Leitfähigkeit ergibt [5, 54]. Ferner sind die Transfermatrizen symplektische $2L \times 2L$ Matrizen und man erhält $2L$ Lyapunovexponenten $\gamma_1 \geq \gamma_2 \geq \ldots \geq \gamma_{2L}$, wobei die zweite
ZUSAMMENFASSUNG

Hälfte die Gegenzahlen der ersten $L$ Lyapunovexponenten sind, d.h. $\gamma_l = -\gamma_{2L+1-l}$. Das asymptotische Verhalten der Transmissionskoeffizienten bzgl. der Blocklänge $N$ ist durch $t_l \sim \exp(-\gamma_l N)$ gegeben, wie es im dritten Kapitel gezeigt wird. Daher werden die Inversen $\gamma_l^{-1}$ gelegentlich auch als Kanal-abhängige Lokalisierungslängen bezeichnet (z.B. in [5]). Gilt $\gamma_L > 0$, dann gehen alle Transmissionskoeffizienten gegen 0 und man erwartet reines Punktspektrum für den unendlich ausgedehnten Block. Für ergodische Operatoren kann dies auf ähnliche Weise bewiesen werden wie für das strikt eindimensionale Anderson Modell im ersten Kapitel. Andererseits entspricht ein Lyapunovexponent der Null ist, einem offenen Kanal in dem Sinn, dass der Transmissionskoeffizient nicht gegen 0 geht. Kotani und Simon [51, 52, 82] haben gezeigt, dass jeder offene Kanal zum a.c. Spektrum mit Multiplizität 2 beiträgt. Das bedeutet das a.c. Spektrum hat eine Multiplizität, die zweimal so groß ist wie die Anzahl offener Kanäle.

Im vierten Kapitel wird das Skalierungsverhalten des Lyapunovexponenten für strikt eindimensionale Modelle an Bandrändern untersucht. Gegeben ist ein periodischer Schrödinger Operator mit nächster-Nachbar-Sprungtermen und Potential. Zu diesem werden zufällige Sprunsterme und zufälliges Potential, beides gekoppelt mit der Kopplungskonstanten $\lambda$, addiert. Um das Skalierungs-Verhalten nahe am Bandrand $E_b$ genauest möglich zu studieren, wird auch die Energie mit $\lambda$ in der Form $E = E_b + \lambda \eta \epsilon$ skaliert. Charakteristisch für eine Energie am Rand eines Bandes, $E_b$, des ungestörten Operators ($\lambda = 0$) ist, dass die entsprechende Transfermatrix zu niedrigster Ordnung äquivalent zu einer nicht diagonalisierbaren Jordan Normalform ist, d.h. nach geeignetem Basiswechsel gilt $T^{E_b} = \pm \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} + O(\lambda^\eta, \lambda)$. Das bedeutet dass die induzierte freie ($\lambda = 0$) Dynamik auf dem projektiven Raum $\mathbb{R}P(1)$ einen parabolischen Fixpunkt hat, was die störungstheoretische Analyse etwas schwierig macht, da der Fixpunkt nur von einer Seite stabil ist. Daher wird der Bereich um den parabolischen Fixpunkt aufgebläht, um die Dynamik genauer zu untersuchen. Genau genommen passiert dies durch eine $\lambda$ abhängige Konjugation mit der Matrix $N_\delta = \begin{pmatrix} \lambda^{\delta} & 0 \\ 0 & 1 \end{pmatrix}$, was zu $N_\delta T^{E_b} N_\delta^{-1} = \begin{pmatrix} 1 & \lambda^{\delta} \\ 0 & 1 \end{pmatrix} + O(\lambda^{\eta-\delta}, \lambda^{1-\delta})$ führt. Der Parameter $\delta$ in der Konjugationsmatrix $N_\delta$ kann nun abhängig von $\eta$ und dem Vorzeichen von $\epsilon$ so geschickt gewählt werden, dass danach die Störungsrechnung auf ähnliche Weise wie im Bandzentrum in [77] durchgeführt werden kann. Es ergeben sich drei Regimes. Zunächst für $\eta < 4/3$ gibt es einen Einfluss vom Vorzeichen von $\epsilon$. Führt es aus dem Band heraus, kommt man in das sog. hyperbolische Regime und $\gamma$ skaliert wie $\gamma \sim \lambda^{\eta/2}$. Führt das Vorzeichen von $\epsilon$ in das Band hinein, ist man im elliptischen Regime und erhält $\gamma \sim \lambda^{2-\eta}$. Für $\eta \geq 4/3$ befindet man sich im parabolischen Regime. In diesem Fall ist die Wahl des Basiswechsels $\delta = 2/3$ und entspricht dem Ansatz in einer formalen Rechnung von Derrida und Gardner [22]. Allerdings wird dies dort nicht auf dem Level von Matrizen eingeführt und erscheint wenig konzeptionell. Die Begründung für den Ansatz ist lediglich, dass er zu einer lösbaren Differentialgleichung führt.

Im folgenden, für Bandränder universellem Phasendiagram wird das Skalierungsverhalten des Lyapunovexponenten und der Änderung $N$ der integrierten Zustandsdichte (IDS)
ZUSAMMENFASSUNG

an Bandrändern nochmals zusammen gefasst.

\[ E = E_b - \varepsilon \lambda^{4/5} \]

\[ E = E_b + \varepsilon \lambda^{4/5} \]

\[ \gamma \sim \lambda^{3/2} \]

\[ N \sim \lambda^{2/3} \]

\[ \gamma \sim \lambda^{2-\eta} \]

\[ N \sim \lambda^{\eta/2} \]

\[ E = E_b + \varepsilon \lambda^{4/5} \]

\[ E = E_b - \varepsilon \lambda^{4/5} \]

\[ \lambda \]

\[ E \]

\[ \text{Band} \]

Abbildung 1: Phasendiagram, Skalierung des Lyapunov exponenten und der IDS


Im weiteren Verlauf der Arbeit wollen wir uns dann ab Kapitel 6 von den rein eindimensionalen Modellen verabschieden und quasi eindimensionale Modelle auf Streifen der Breite \( L \) untersuchen. Dann sind die Transfermatrizen symplektische \( 2L \times 2L \) Matrizen und wie im zweiten und dritten Kapitel gezeigt wird, gibt es mehrere Lyapunovexponenten, deren Skalierungsverhalten interessant ist. Zur störungstheoretischen Berechnung muss dann ein etwas komplizierteres stochastisches dynamisches System auf einem kompakten homogenen Raum \( \mathcal{M} \) untersucht werden, das durch die Wirkung der Transfermatrizen getrieben wird. Da der homogene Raum von der Streifenbreite und davon abhängt, welche Lyapunovexponenten man betrachtet, haben wir uns entschlossen, das Problem abstrakter zu fassen. Daher werden in Kapitel 6 allgemein Markov Prozesse auf kompakten Mannigfaltigkeiten untersucht, die durch eine Liegruppenwirkung und unabhängig, identisch verteilte (u.i.v.) Zufallselemente der Liegruppe induziert werden. Dieser Abschnitt ist der mathematisch abstrakteste Teil meiner Doktorarbeit. Die u.i.v. Liegruppenelemente sind von der Form

\[ T_{\lambda,\sigma} = \mathcal{R} \exp \left( \sum_n \lambda^n \mathcal{P}_{n,\sigma} \right), \]

wobei \( \sigma \in \Sigma \) Element eines abstrakten Wahrscheinlichkeitsraumes ist und \( \sigma \mapsto \mathcal{P}_{n,\sigma} \) Zufallsvariablen mit Werten in der entsprechenden Liealgebra. Von \( \mathcal{R} \) und dem Erwartungswert \( \mathbb{E}(\mathcal{P}_{1,\sigma}) \) nehmen wir an, dass sie eine kompakte, abelsche Gruppe erzeugen. Der homogene Raum \( \mathcal{M} \) sei desweiteren mit einer Riemannschen Metrik versehen und das zugehörige Volumenmaß sei \( \mu \). Der durch u.i.v. Kopien von \( T_{\lambda,\sigma} \) induzierte Markovprozess auf \( \mathcal{M} \) sei
mit \((x_n)_{n \in \mathbb{Z}}\) bezeichnet. Unter bestimmten technischen Voraussetzungen existiert eine glatte, fast sicher positive, \(L^1(\mu)\)-normierte Funktion \(\rho\), so dass für eine hinreichend schöne (weitere technische Voraussetzung), glatte Funktion \(f\) auf \(\mathcal{M}\) folgendes gilt:

\[
\frac{1}{N} \mathbb{E} \sum_{n=1}^{N} f(x_n) = \int d\mu \rho f + \mathcal{O}(\lambda, (N\lambda^2)^{-1}) .
\]

Als Korollar erhält man, dass jede Familie \(\nu_\lambda\) von invarianten Wahrscheinlichkeitsmaßen des Markovprozesses \(x_n\) auf \(\mathcal{M}\) im Limes \(\lambda \to 0\) in der \(\ast\)-schwachen Topologie gegen das Wahrscheinlichkeitsmaß \(\rho \mu\) konvergiert. Im Fall \(\mathcal{R} = 1\) und \(\mathbb{E}(\mathcal{P}_{1,\sigma}) = 0\) können die Birkhoffsummen jeder glatten Funktion zu beliebigen Ordnungen in \(\lambda\) entwickelt werden. Das bedeutet in diesem Fall erhalten wir eine Folge glatter Funktionen \(\rho_m\) auf \(\mathcal{M}\) mit \(\int d\mu \rho_m = \delta_{1,m}\), so dass für jede glatte Funktion \(f\) und jedes feste \(M \in \mathbb{N}\) gilt

\[
\frac{1}{N} \mathbb{E} \sum_{n=1}^{N} f(x_n) = \sum_{m=0}^{M} \lambda^m \int d\mu \rho_m f + \mathcal{O}(\lambda^{M+1}, (N\lambda^2)^{-1}) .
\]

Allerdings hängt die Konstante des Fehlers von \(M\) ab, d.h. der Fehler \(\mathcal{O}(\lambda^M)\) ist genauer gesagt von der Gestalt \(\mathcal{O}(C_M \lambda^{M+1})\) und \(C_M\) wächst sehr schnell, so dass für kleine \(\lambda\) die Eindeutigkeit des invarianten Maßes auf diese Art und Weise nicht gezeigt wird. Im übrigen gibt es auch ein einfaches Beispiel (siehe Remark 6.2, Seite 113), das zeigt, dass es nicht unbedingt ein Intervall \([0, \lambda_0]\) geben muss, so dass für alle \(\lambda \in [0, \lambda_0]\) das invariante Maß des Markovprozesses eindeutig ist. Ersetzt man jedoch ‘alle \(\lambda\) durch ‘Lebesgue fast alle \(\lambda\)’ so habe ich die Vermutung, dass die Aussage stimmt.

Kapitel 7 beschreibt einen zufälligen Schrödinger Operator auf einem Streifen, in dem das Theorem aus Kapitel 6 angewendet werden kann. Genauer handelt es sich um ein Modell von \(L\) zufällig gekoppelten Kabeln, das dem \(L\)-Orbital Modell von Wegner ähnlich ist. Der Operator sieht genauso aus, wie beim eindimensionalen Anderson Modell,

\[
(H_{\lambda,\omega}\Psi)_n = -\Psi_{n-1} - \Psi_{n+1} + \lambda V(\omega_n) \Psi_n , \quad (\Psi_n)_{n \in \mathbb{Z}} \in \ell^2(\mathbb{Z}, \mathbb{C}^L).
\]

Der Unterschied ist, dass jetzt \(\Psi_n \in \mathbb{C}^L\) ein Vektor ist und \(V(\omega_n)\) eine Folge u.i.v. Zufallsvariabler mit Werten in den Hermiteschen \(L \times L\) Matrizen. Um die zugehörigen \(L\) positiven Lyapunovexponenten zu bestimmen, muss ein Markov Prozess auf einer isotropen Fahnennamnigfaltigkeit betrachtet werden, die diffeomorph zu einem Quotienten der Gruppe \(U(L) \times U(L)\) ist und daher ein natürliches Haarmaß besitzt. Unter einem bestimmten Basiswechsel für die Transformatoren ist für Energien \(|E| < 2\), \(E \neq 0\) das invariante Maß zu niedrigster Ordnung (\(\rho \mu\) in obiger Gleichung) durch das Haarmaß gegeben. Dies wird auch als ‘random phase property’ bezeichnet. Desweiteren erlaubt dies eine störungstheoretische Berechnung der Lyapunovexponenten, die zeigt, dass diese für kleine \(\lambda\) näherungsweise equidistant sind.

In Kapitel 8 werden zufälligen Dirac Operatoren analysiert und ein Beispiel gezeigt, in dem ein einziger verschwindender Lyapunovexponent (von den nicht-negativen) ausreicht,
ZUSAMMENFASSUNG

um reines a.c. Spektrum zu erhalten. Genauer ist der zufällige Operatorenpotential auf $L^2(\mathbb{R}, \mathbb{C}^{2L})$ durch

$$H = \mathcal{J} \partial + \mathcal{W} + \sum_{j \in \mathbb{Z}} \mathcal{V}_j \delta_{x_j}, \quad \mathcal{J} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

wobei $\partial$ die Ableitung von $\Psi(x)$ bzgl. $x$ darstellt. Eine physikalischere Art die Multiplikation mit $\mathcal{J}$ im Operator zu schreiben wäre $\mathcal{J} = \sigma_2 \otimes 1_{2L}$, wobei $\sigma_2$ die zweite der drei Pauli-Matrizen ist. $\mathcal{W} \in L^1_{\text{loc}}(\mathbb{R}, \text{Her}(2L, \mathbb{C}))$ soll ein zufälliges, ergodisches Potential sein und die Familie $(\mathcal{V}_j)_j$ ergibt ein u.i.v. zufälliges Dirac Potential auf einem regulären Gitter $(x_j)_{j \in \mathbb{Z}}$. Außerdem soll der Operator die Zeitumkehrsymmetrie erfüllen, das bedeutet

$$\mathcal{J}^* \mathcal{W}(x) \mathcal{J} = \mathcal{W}(x), \quad \mathcal{J}^* \mathcal{V}_j \mathcal{J} = \mathcal{V}_j.$$

Ähnlich wie bei den zufälligen Schrödinger Operatoren kann die stationäre Gleichung $H\Psi = E\Psi$ mit Hilfe von $2L \times 2L$ Transfermatrizen geschrieben werden. Wegen der Zeitumkehrinvarianz sind die Transfermatrizen in der Gruppe $\text{SO}^*(2L)$, welche eine Untergruppe der symplektischen Gruppe $\text{Sp}(2L, \mathbb{C})$ ist. Neben der symplektischen Symmetrie, welche $\gamma_l = -\gamma_{2L+1-l}$ impliziert, erhält man auch Kramer's Degenration, was zu $\gamma_{2l-1} = \gamma_{2l}$ führt. Falls $L$ eine ungerade Zahl ist, hat dies $\gamma_{L+1} = -\gamma_L = \gamma_{L+1}$, also $\gamma_L = 0 = \gamma_{L+1}$ zur Folge. Bestimmte Bedingungen ähnlich denen in Kapitel 2 implizieren, dass die Lyapunovexponenten voneinander verschieden sind, abgesehen von Kramer's Degenration. Damit verschwinden nur die Lyapunovexponenten $\gamma_L$ und $\gamma_{L+1}$. Eine Variante der Kotani Theorie für Dirac Operatoren, die von Sun erarbeitet wurde [86] und hier für Dirac Potentiale erweitert wird, beweist a.c. Spektrum von Multiplizität 2 auf der gesamten reellen Achse. Falls desweiteren die Verteilung der $V_j$ absolut stetig ist, können wir Argumente von Jaksic und Last [41] anpassen um reines a.c. Spektrum zu beweisen.
II Summary

This thesis is about my work on random Schrödinger and Dirac operators in the last three years. The topic started with a paper by the physicist P. W. Anderson [3] about ‘absence of diffusion in certain random lattices’ published in 1958. This phenomenon is today known as Anderson localization. It describes the fact that random impurities in a solid can localize electrons and stop diffusion completely. In 1977 P. W. Anderson got the Nobel prize in physics together with N. F. Mott and J. H. van Vleck for ‘the fundamental theoretical investigations of the electronic structure of magnetic and disordered systems’. The paper mentioned above was part of the work he got the Nobel prize for.

In order to introduce the main relevant mathematical objects, terms and definitions, I will discuss the discrete one-dimensional Anderson model in the first chapter. There one considers a random Schrödinger operator on $l^2(\mathbb{Z})$ with a random potential $V(\omega_n)$, i.e.

$$ (H_{\lambda,\omega}\Psi)_n = -\Psi_{n-1} - \Psi_{n+1} + \lambda V(\omega_n)\Psi_n. $$

The real constant $\lambda$ is called coupling constant which is introduced to do perturbation theory later. It is known that the spectrum and actually all its components (pure point spectrum, singular and absolutely continuous spectrum) are non-random for ergodic potentials w.r.t. translations on $\mathbb{Z}$ [19, 64]. An example is obtained by setting $\omega_n$ and respectively $V(\omega_n)$ to be independent, identically distributed (will be abbreviated by i.i.d.) random variables. We will consider this case and further assume that the potential is centered at zero, i.e. $E(V) = 0$ where $E$ denotes the expectation value. Without the random potential ($\lambda = 0$) the spectrum is absolutely continuous and supported on the interval $[-2, 2]$. But once there is just a little randomness ($\lambda > 0$) one obtains immediately pure point spectrum with exponentially decaying eigenfunctions [33, 64]. This is the mathematical definition of strong Anderson localization. The main tool in the proof are the random transfer matrices

$$ T^E_{\omega_0} = \left( \begin{array}{cc} \lambda V(\omega_0) - E & -1 \\ 1 & 0 \end{array} \right). $$

Using them one can write the stationary Schrödinger equation $H_{\lambda,\omega}\Psi = E\Psi$ in the form

$$ \left( \begin{array}{c} \Psi_{n+1} \\ \Psi_n \end{array} \right) = T^E_n \left( \begin{array}{c} \Psi_n \\ \Psi_{n-1} \end{array} \right). $$

The asymptotic behavior of the random products $T^E_N(\omega) = T^E_{\omega_N}T^E_{\omega_{N-1}} \cdots T^E_{\omega_1}$ can be described by the Lyapunov exponent $\gamma = \lim_{N \to \infty} \frac{1}{N} \log(\|T^E_N(\omega)\|)$ which depends on the energy $E$ and the disorder $\lambda$. The Lyapunov exponent $\gamma$ has also some physical meaning and equals the inverse localization length. In order to calculate it, one may consider the random dynamical system on the projective space $\mathbb{R}P(1)$ induced by the action of the transfer matrices. More precisely the
Lyapunov exponent can be written as a Birkhoff sum of this Markov process. The same is true for the so called integrated density of states (IDS).

As an exact calculation is quite hard, we will do some perturbative calculations and obtain perturbative formulas for the Lyapunov exponent. At energies inside the band, away from the band edges and the band center, $|E| < 2$ and $E \neq 0$, one obtains Thouless formula [87, 64] stating

$$\gamma = \lambda^2 \frac{E(V^2)}{2(4 - E^2)} + O(\lambda^3).$$

For $E = 0$ and at the band edges $|E| = 2$ the perturbative analysis is more involved. The band center was considered in [77], the band edge was part of my Diploma thesis and is considered again in Chapter 4.

In the second chapter I collect some well known facts about products of random matrices and Lyapunov exponents which can be found in [9, 19, 32]. These facts are quite important for understanding the work of the thesis. In the consecutive chapters some models on strips will be considered. There the transfer matrices have a bigger dimension. One has more Lyapunov exponents describing the growth of the singular values of the random products. Chapter 2 shall give some background knowledge on the general theory of products of random matrices.

In the strictly one-dimensional case the Lyapunov exponent describes the localization length and is hence related to the conductivity. In order to obtain a similar meaning for the Lyapunov exponents of models on strips one needs to relate the transfer matrix with the so called scattering matrix which is done in Chapter 3. This chapter is mostly written in physics notation and supposed to give some more background information. There another physical interpretation of the transfer matrix is carried out. Take a finite block and connect it to two infinite ideal leads on both sides which have stationary waves as pseudo-eigenstates. Considering the so called scattering states one realizes that the transfer matrix of the finite block written in the correct basis relates the coefficients of the waves in the left lead with the ones in the right lead. The scattering matrix on the other hand maps incoming waves on outgoing waves [5]. More precisely for one scattering state let the vector $a^+$ denote the coefficients of the waves moving to the right in the left lead, $a^-$ the ones in the left lead moving to the left, $b^+$ the ones in the right lead moving to the right and $b^-$ the one in the right lead moving to the left. Then, after some conjugation, the transfer matrix satisfies $T(a^+) = (b^+)$. Now $a^+$ and $b^-$ describe the incoming wave and hence the scattering matrix satisfies $S(a^+) = (b^-)$. It can be shown that there is a unique unitary matrix satisfying this relation with the transfer matrix $T$. From the scattering matrix one can calculate the transmission coefficients which are important for the coherent transport of charges. On a strip of width $L$ there are $L$ left moving and $L$ right moving modes and hence there are $L$ transmission coefficients $t_1 \geq \ldots \geq t_L$ whose sum gives the Landauer conductance [5, 54]. Moreover the transfer matrices are symplectic $2L \times 2L$ matrices leading to $2L$ Lyapunov exponents $\gamma_1 \geq \ldots \geq \gamma_{2L}$. However, because of the symplectic structure the second $L$ ones are basically the same as the first $L$ ones with opposite sign, i.e. $\gamma_l = -\gamma_{2L+1-l}$. The asymptotic behavior of the transmission coefficients w.r.t. the length $N$ of the block is
given by $t_l \sim \exp(-\gamma_l N)$ as shown in Chapter 3. Therefore the inverses are sometimes called channel-dependent localization lengths (e.g. in [5]). If $\gamma_L > 0$, then all transmission coefficients converge to zero and one expects pure point spectrum for the infinite block $N \to \infty$. For ergodic operators this can be proved in a similar way as for the strictly one-dimensional Anderson model done in Chapter 1. On the other hand, a zero Lyapunov exponent corresponds to an open channel, meaning that the transmission coefficient does not vanish in the limit. Kotani and Simon [51, 52, 82] showed that each open channel on an energy interval adds a.c. spectrum of multiplicity 2. This means the multiplicity of a.c. spectrum equals twice the number of open channels.

In the fourth chapter the scaling of the Lyapunov exponents at band edges for strictly one-dimensional models is investigated. Given a periodic Schrödinger operator with next nearest neighbor hoping terms and periodic potential, centered random hoping terms and a random potential are added, both coupled with a small coupling constant $\lambda$. In order to study the scaling at the band edge $E_b$ in a more detailed way we also scale the energy with $\lambda$ in the form $E = E_b + \lambda^\eta \epsilon$. The characteristic property of an energy $E_b$ at the edge of a band of the unperturbed operator ($\lambda = 0$) is that the transfer matrix is not diagonalizable but equivalent to a Jordan normal form. This means after some basis change $T^{E_b} = \pm \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} + O(\lambda^\eta, \lambda)$. Therefore the free induced dynamics on $\mathbb{RP}(1)$ has a parabolic fix point which is only stable from one side. This makes the perturbative analysis more difficult. In order to study this dynamics more detailed we blow up a small neighborhood around this parabolic fix point using a $\lambda$ dependent basis change. More precisely we take $N_\delta = \begin{pmatrix} \lambda^\delta & 0 \\ 0 & 1 \end{pmatrix}$ and obtain $N_\delta T^{E_b} N_\delta^{-1} = \begin{pmatrix} 1 & \lambda^\delta \\ 0 & 1 \end{pmatrix} + O(\lambda^{\eta-\delta}, \lambda^{1-\delta})$. The parameter $\delta$ can be chosen in a nice way, such that the perturbative calculation gets similar to the one done at the band center in [77]. For this choice, $\delta$ depends on $\eta$ and the sign of $\epsilon$. There are three regimes. For $\eta < \frac{1}{4}$ there is some influence of the sign of $\epsilon$. If it leads to energies outside the band one has the hyperbolic regime and $\gamma$ scales like $\gamma \sim \lambda^{\eta/2}$. If the sign of $\epsilon$ leads to energies inside the band one is in the elliptic regime and obtains $\gamma \sim \lambda^{2-\eta}$. For $\eta \geq \frac{1}{4}$ one obtains the parabolic regime and the choice for $\delta$ is $\delta = \frac{2}{3}$. This corresponds to some Ansatz in a formal calculation done by Derrida and Gardner [22]. However in their work this is not done on the level of matrices and it doesn’t appear in a conceptual way. This Ansatz in the paper is only justified by the fact that it leads to a solvable differential equation.

The universal phase diagram for bandedges below summarizes these scalings of the Lyapunov exponent and the change $N$ of the IDS.

In Chapter 5 the methods of Chapter 4 are generalized to handle strongly mixing potentials. This work complements the perturbative calculations at energies inside the band for strongly mixing potentials done by Chulaevsky and Spencer [17]. A mathematical precise formulation of the mixing condition can be found on page 93.

In the sequel from Chapter 6 on we want to consider models on strips of width $L$ instead of strictly one-dimensional models. In this case the transfer matrices are symplectic $2L \times 2L$ matrices. As shown in the second and third chapter there are more Lyapunov exponents...
whose scaling is interesting. For the perturbative calculation one needs to consider a more complicated stochastical, dynamical system on a compact homogeneous space $\mathcal{M}$ driven by the action of the transfer matrices. As the homogeneous space depends on the width $L$ and the considered Lyapunov exponents one wants to calculate, it is worth to look at this problem in a more abstract way. Therefore Chapter 6 deals with Markov processes on compact manifolds driven by the action of i.i.d. Lie group elements. This part is mathematically the most abstract one of my thesis. The i.i.d. Lie group elements shall be of the form

$$T_{\lambda,\sigma} = \mathcal{R} \exp \left( \sum_n \lambda^n P_{n,\sigma} \right),$$

where $\sigma \in \Sigma$ is an element of an abstract probability space. The maps $\sigma \mapsto P_{n,\sigma}$ are random variables with values in the corresponding Lie algebra. We assume that $\mathcal{R}$ and $P_{1,\sigma}$ create a compact, abelian Lie group. The homogeneous space $\mathcal{M}$ shall have a Riemannian metric for convenience and the corresponding volume measure shall be denoted by $\mu$. The Markov process on $\mathcal{M}$ induced by i.i.d. copies of $T_{\lambda,\sigma}$ shall be denoted by $(x_n)_{n \in \mathbb{Z}}$. If certain technical conditions are satisfied then there is a smooth, $\mu$-almost surely positive and $L^1(\mu)$ normalized function $\rho$, such that for sufficiently nice (technical detail) smooth functions $f$ one has

$$\frac{1}{N} E \sum_{n=1}^N f(x_n) = \int d\mu \rho f + O(\lambda, (N\lambda^2)^{-1}).$$

As a corollary one obtains for any family $(\nu_\lambda)_\lambda$ of invariant probability measures of the Markov process $(x_n)_n$ on $\mathcal{M}$ that $\nu_\lambda$ converges in the weak-* topology to the probability measure $\rho\mu$. In the case $\mathcal{R}^k = 1$ for some natural $k$ and $E(P_{1,\sigma}) = 0$ one can calculate the Birkhoff sums of any smooth function to any order in $\lambda$. This means in this case there is a sequence of smooth functions $\rho_m$ satisfying $\int d\mu \rho_m = \delta_{1,m}$ such that for any smooth $f$ and fixed $M \in \mathbb{N}$ one has

$$\frac{1}{N} E \sum_{n=1}^N f(x_n) = \sum_{m=0}^M \lambda^m \int d\mu \rho_m f + O(\lambda^{M+1}, (N\lambda^2)^{-1}).$$

Figure 1: Phase diagram, scaling of Lyapunov exponent and IDS
However the constant of the error depends on $M$ meaning the error of order $O(\lambda^M)$ is better written as $O(C_M \lambda^M)$ and $C_M$ might grow very fast in $M$. Therefore uniqueness of the invariant measure for small $\lambda$ is not proved. Anyway there is a simple counter example (Remark 6.2 on page 113) showing that one may not find an interval $[0, \lambda_0]$ such that the invariant measure of the Markov process is unique for all $\lambda \in [0, \lambda_0]$. However, replacing 'for all $\lambda$' by 'for Lebesgue almost all $\lambda$' my conjecture is that this is true.

Chapter 7 considers a random Schrödinger operator on a strip for which Theorem 6.2 can be used. More precisely a model of $L$ randomly coupled wires is considered, which is quite similar to the Wegner $L$-orbital model. Formally the Schrödinger operator looks like

$$(H_{\lambda, \omega} \Psi)_n = -\Psi_{n-1} - \Psi_{n+1} + \lambda V(\omega_n) \Psi_n, \quad (\Psi_n)_{n \in \mathbb{Z}} \in \ell^2(\mathbb{Z}, \mathbb{C}^L).$$

The difference is that now $\Psi_n \in \mathbb{C}^L$ is a vector and $V(\omega_n)$ a sequence of i.i.d. $L \times L$ hermitian matrices. To get the $L$ non-negative Lyapunov exponents one has to consider a Markov process on the isotropic flag manifold. This one is diffeomorphic to a quotient of the compact group $U(L) \times U(L)$ and possesses a natural Haar measure. Using a certain basis change for the transfer matrices one obtains that for energies $|E| < 2, E \neq 0$ the invariant measure to lowest order ($\rho\mu$ in equation above) is given by the Haar measure. This is called the 'random phase property'. Furthermore this allows a perturbative calculation of all Lyapunov exponents showing that they are approximately equidistant for small $\lambda$.

In Chapter 8 we analyze random Dirac operators and give an example where just one (of the $L$ non-negative ones) vanishing Lyapunov exponent assures almost surely pure a.c. spectrum. More precisely the random operator is defined on $L^2(\mathbb{R}, \mathbb{C}^{2L})$ and given by

$$H = \mathcal{J} \partial + \mathcal{W} + \sum_{j \in \mathbb{Z}} \mathcal{V}_j \delta_{x_j}, \quad \mathcal{J} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

where $\partial$ denotes the differentiation of $\Psi(x)$ w.r.t. $x \in \mathbb{R}$. A more physical way to write $\mathcal{J}$ in the operator would be of the form $\mathcal{J} = \iota \sigma_2 \otimes 1_L$ where $\sigma_2$ is the second Pauli matrix. $\mathcal{W} \in L^1_{\text{loc}}(\mathbb{R}, \text{Her}(2L, \mathbb{C}))$ shall be some random, ergodic potential and the family $(\mathcal{V}_j)_j$ gives an i.i.d. random Dirac potential supported on a regular lattice $(x_j)_{j \in \mathbb{Z}}$. Furthermore we suppose the operators to satisfy the time reversal symmetry leading to

$$\mathcal{J}^* \mathcal{W}(x) \mathcal{J} = \mathcal{W}(x), \quad \mathcal{J}^* \mathcal{V}_j \mathcal{J} = \mathcal{V}_j.$$

Similar as for the discrete Schrödinger operators a solution of the stationary equation $H \Psi = E \Psi$ can be written using $2L \times 2L$ transfer matrices. Because of the time reversal symmetry this transfer matrices lie in the group $SO^*(2L)$ which is a sub-group of the symplectic group $\text{Sp}(2L, \mathbb{C})$. Besides the symplectic symmetry implying $\gamma_l = -\gamma_{2L+1-l}$ one also has Kramer’s degeneracy leading to $\gamma_{2l-1} = \gamma_{2l}$. Therefore if $L$ is odd then $\gamma_{L+1} = -\gamma_L = \gamma_{L+1}$ and hence $\gamma_L = 0$. Some conditions on the randomness similar to the ones mentioned in Chapter 2 imply distinct Lyapunov exponents besides Kramer’s degeneracy. Therefore only the Lyapunov exponents $\gamma_L$ and $\gamma_{L+1}$ vanish. A version of
Kotani theory for Dirac operators done by Sun [86] which is extended here to the case of Dirac type potentials implies a.c. spectrum of multiplicity 2 on the whole real line. If furthermore the distribution of the $V_j$ is absolutely continuous we can adopt the arguments of Jaksic and Last [41] to show that the a.c. spectrum is pure.
Chapter 1

Introduction on the Anderson model

1.1 The phenomenon of Anderson localization

Let us consider a solid with a perfect crystal structure where the atoms sit on a perfect $d$ dimensional lattice $\mathbb{Z}^d$. For real observable systems one may have very thin cables, surfaces or blocks and hence $d \in \{1, 2, 3\}$. We will just consider the 1-particle Hamiltonian as to first order an description by non-interacting electrons is often sufficient. The possible energies of electrons in such a solid have a band structure (Bloch theory). Electrons with high enough energy can travel through the solid, the others are defracted by the atoms. The energy bands for electrons which can travel through the solid are usually called conducting bands. As the outer electrons of an atom are called valence electrons, the band where they sit is called the valence band. More precisely, as electrons are Fermions, there can not be two different electrons in the same state. Therefore at zero (Kelvin) temperature the lowest energy states are all filled, up to the so called Fermi energy $E_F$. If this energy is in between two bands then the band below is the valence band and all bands above $E_F$ are conducting bands. The material is a semi-conductor if the energy difference of the valence band and the lowest conducting band is very small compared to the temperature, such that there are some electrons in the conducting band by thermal excitation. It is an insulator if the energy gap between the conducting band and the valence band is much greater than the thermal energy an electron gets in average from thermal excitation. In metals the typical situation is that the Fermi energy is inside one of the energy bands and most of the valence electrons form an electron gas able to travel through the whole solid. This leads to a very good conductor.

In 1958 Philip W. Anderson, who got the Nobel prize in 1977 in part for the topic described below, worked out that tiny modifications, such as impurities or defects, can change the behavior of electrons dramatically [3]. Electron waves which may normally travel ballistically through the solid can be stopped completely and appear to be localized in space, provided there is enough disorder (compared to the perfect crystal). This phenomenon might occur even in a situation where the random impurities do not create energy barriers which are higher than the typical energies of the conducting electrons. In such a case a classical particle could not be trapped, on the other hand electrons are not classical particles.
1. INTRODUCTION ON THE ANDERSON MODEL

The heuristic explanation is that the electron waves are scattered at the impurities and one has to consider the interference of all scattered waves. In this situation it might occur that there is enough destructive interference such that the sum of all scattered waves is not extended, but localized. When this happens a conducting material might become an insulator due to the impurities. P.W. Anderson expected a transition from a conductor to an insulator when increasing the disorder.

Today, the phenomenon of absence of diffusion due to a random environment is known as Anderson localization, it occurs also in other situations than the one described above, e.g. light waves. Although this subject seems to be quite old, there are still some open mathematical problems such as proving Anderson localization for weak disorder in the 2D Anderson model or proving delocalization for weak disorder in the 3D Anderson model. Below I will describe these models more precisely in mathematical terms. Even today one can still find recent papers about Anderson localization observed in experiments, e.g. in quantum optics.

1.2 Basic quantum mechanics

This section can be found in any quantum mechanics text-book such as [4]. In mathematical terms a quantum particle is described by its state which in general is a linear functional acting on a unital $*$-algebra or $C^*$-algebra$^1$. In many cases$^2$ the $*$-algebra consists of (possibly unbounded) operators on a separable Hilbert space $\mathcal{H}$. The state of the quantum particle is then identified with a positive trace class operator $\rho$ with unit trace $\text{Tr}(\rho) = 1$ and the corresponding functional is given by $A \mapsto \text{Tr}(\rho A)$. Such operators $\rho$ are also called density matrices. Here one has to be careful with the term matrix as the matrices representing $\rho$ have usually infinite size. Self-adjoint operators $A$ are supposed to model observables for experiments and the value $\text{Tr}(\rho A)$ represents the expectation value for measurements of the observable $A$ when the particle is in the state $\rho$. Using functional calculus one obtains the distribution of $A$ by considering $\text{Tr}(\rho 1_{(-\infty,x]}(A))$ which is the probability that the outcome of the measurement of $A$ is smaller or equal to $x$. Here $1_{(-\infty,x]}$ denotes the characteristic function. As in this interpretation one should get a probability distribution on $\mathbb{R}$ one requires for the state that $\text{Tr}(\rho) = \text{Tr}(\rho 1_{\mathbb{R}}(A)) = 1$ and $\text{Tr}(\rho A) > 0$ whenever $A > 0$ which is equivalent to $\rho > 0$. So from the interpretation in quantum mechanics the restrictions on $\rho$ mentioned above are natural.

For any density matrix $\rho$ the associated probability distribution of an observable $A$ is supported on the spectrum of $A$. Hence the spectrum is an important quantity as it represents the possible values which can be measured in an experiment. For some quantities in physics, the possible values are unbounded and therefore some observables are associated to unbounded, self-adjoint operators. If $A$ is unbounded, then $\rho A$ might not be trace class and hence $A$ may not have an expectation value. However, $1_{(-\infty,x]}(A)$ is always a

$^1$Even though I mention here the general framework of $C^*$ algebras I would like to point out that this work does not include any research in this direction.

$^2$For many models there is a canonical Hilbert space. However in quantum field theory one sometimes has a canonical $C^*$-algebra, but no canonical representation on a Hilbert space.
bounded operator and therefore one still gets a probability distribution on $\mathbb{R}$ associated to the observable $A$ and the state $\rho$.

If $\Psi$ is a unit vector then $\text{Tr}(\langle \Psi \rvert \langle \Psi \rvert \rangle) = 1$ and $\langle \Psi \rvert \langle \Psi \rvert \rangle$ is a density matrix and hence a state. Any state represented by a unit vector in this way is called a pure state. Here $\langle \Psi \rvert \langle \Psi \rvert \rangle$ is the Dirac notation for the map $\varphi \mapsto \langle \Psi, \varphi \rangle$ on $\mathcal{H}$. The expression $\langle \Psi, \varphi \rangle$ denotes the scalar product on $\mathcal{H}$ which in this work will always be assumed to be linear in the second and anti-linear in the first component. The Dirac notation comes from splitting the scalar product and $\langle \Psi \rvert \rangle$ is interpreted as a linear functional on $\mathcal{H}$. The expression $\langle \Psi \rvert \rangle$ denotes the scalar product on $\mathcal{H}$ which in this work will always be assumed to be linear in the second and anti-linear in the first component. The Dirac notation comes from splitting the scalar product and $\langle \Psi \rvert \rangle$ is interpreted as a linear functional on $\mathcal{H}$. By the Riesz-Fischer theorem, any bounded linear functional on $\mathcal{H}$ can be written in that way, reflecting the fact that there is a canonical anti-isomorphism between a Hilbert space and its dual given by $\langle \Psi \rvert \rangle \mapsto \langle \Psi \rvert \rangle$. This also shows that Hilbert spaces are reflexive.

The general framework of quantum mechanics can be seen as a non-commutative generalization of ordinary probability theory. Like the continuous functions on a compact space, the bounded operators form a $C^*$-algebra. If one has a probability measure on the Borel sets of a compact space, the expectation value is a linear functional $\varphi$ acting on the continuous functions. Then one has $\varphi(1) = 1$, $\varphi(f^*) = \varphi(f)^*$ and the convexity condition $\varphi(f^*f) \geq |\varphi(f)|^2$. Indeed, using this rules as axioms for a functional on a $C^*$ algebra gives the general mathematical framework of so called non-commutative probability theory. It can be shown to reduce to the ordinary probability theory whenever the considered $C^*$-algebra is commutative.

One of the most important observables in quantum mechanics is the energy which is represented by the so called self-adjoint Hamilton operator or Hamiltonian $H$. This operator not only represents the energy but also determines the time evolution of a quantum state which is given by $\rho(t) = \exp(-itH)\rho(0)\exp(itH)$. (In this work $i$ represents the imaginary unit, where as $i$ is sometimes used as index variable). As $H$ is self-adjoint the one parameter group of operators $\exp(itH)$ is unitary. In fact it is known that the generators of unitary one-parameter groups are precisely the self-adjoint operators, this is the reason why one has this restriction on $H$.

Therefore $\rho(t)$ is a conjugation of $\rho(0)$ by unitaries and hence again a density matrix. If $\rho(0)$ is a pure state, $\rho(0) = \langle \Psi \rvert \langle \Psi \rvert \rangle$ then also $\rho(t)$ is a pure state represented by the unit vector $\Psi(t) = \exp(-itH)\Psi(0)$. This leads to the so called Schrödinger equation $i\partial_t \Psi(t) = H\Psi(t)$. To solve this equation one has to analyze the spectral decomposition of $H$. If one has for example an eigenvector $H\Psi = E\Psi$ then $\Psi(t) = e^{-itE}\Psi(0)$ is the solution to the equation above and the state $|\Psi(t)\rangle\langle \Psi(t) | = |\Psi(0)\rangle\langle \Psi(0) |$ is fixed. More generally as $H$ commutes with $\exp(\pm itH)$ and the cyclicity of the trace, one obtains $\text{Tr}(\rho(t)H) = \text{Tr}(\rho(0)\exp(itH)H\exp(-itH)) = \text{Tr}(\rho(0)H)$. This basicly means that in average the energy is conserved, a basic principle in physics.

As pure eigenstates of $H$ are fixed under the evolution one defines spectral localization in an energy interval $I$ by $H$ exhibiting pure point spectrum in that interval $I$. For so called strong Anderson localization one also requires the eigenstates to be exponentially localized in space which means the probability to find the particle at a position $x$ should vanish exponentially as $\|x\|$ goes to infinity.
1. INTRODUCTION ON THE ANDERSON MODEL

The distribution of a one-particle state in space is given by an observable \( X \) representing the position. In many cases one uses the coordinate representation for the states. This means that a state \( \Psi \) is a function of \( x \in \mathbb{R}^d \) or \( x \in \mathbb{Z}^d \) with values in some (usually finite) Hilbert space \( \mathcal{H} \) describing other degrees of freedom like spin. If \( | \cdot | \) denotes the norm in \( \mathcal{H} \), then \( |\Psi(x)|^2 \) denotes the probability density (in the continuous case and the probability in the lattice case) to find the particle at position \( x \). The (one-particle) Hilbert space in this case is \( L^2(\mathbb{R}^d) \otimes \mathcal{H} \) or \( \ell^2(\mathbb{Z}^d) \otimes \mathcal{H} \) and the position operator is just the multiplication with the position. This means if \( \Psi = (\Psi(x))_{x \in \mathbb{Z}^d} \in \ell^2(\mathbb{Z}^d) \otimes \mathcal{H} \) then \( X \) is a vector of operators defined by \( (X\Psi)(x) = x\Psi(x) \). Exponentially localization in space of \( \Psi \) is therefore equivalent to \( |\Psi(x)| \) decreasing exponentially as \( \|x\| \to \infty \). The same holds when replacing \( \ell^2(\mathbb{Z}^d) \) with \( L^2(\mathbb{R}^d) \). In the sequel we will ignore the spin and set \( \mathcal{H} = \mathbb{C} \). Without interaction to a magnetic field the spin would just change the multiplicity of the spectrum.

1.3 Anderson localization in mathematical terms

This section is already more specifically related to the topic of the Thesis. Let me come back to the phenomenon of Anderson localization in random media in mathematical terms. Assume an electron which can travel through the perfect solid is described by a Hamiltonian \( H_0 \) acting on \( \ell^2(\mathbb{Z}^d) \) or \( L^2(\mathbb{R}^d) \) which has only absolutely continuous spectrum. Usually \( H_0 \) is given by some kind of discrete or continuous Laplace operator.

Now let us model the random impurities or defects by a random operator \( V_\omega \) (in most cases later this will just be some sort of random potential). Here \( \omega \) is an element of an abstract probability space labeling the randomness. Then consider the random Hamiltonian \( H_\omega = H_0 + V_\omega \). Let us assume that the spectrum of \( H_\omega \) is almost surely non-random. For the Anderson model considered below this is known to be true \([19, 64]\).

**Definition 1.1** One says that the random family \( H_\omega \) exhibits strong Anderson localization in an energy interval \([E_0, E_1]\) if almost surely the spectrum of \( H_\omega \) in this interval is pure point and all eigenstates are exponentially localized.

The one-particle Hilbert space describing one valence electron is made out of the states for the valence electrons of any single atom. As these states are located at the atoms which sit on a lattice one may identify the states with the lattice points and hence the Hilbert space may be identified with \( \ell^2(\mathbb{Z}^d) \). The state at the atom in position \( n \in \mathbb{Z}^d \) is just given by the \( \ell^2 \) function being 1 at \( n \) and 0 elsewhere, in Dirac notation this state may be written as \( |n\rangle \). However by the presence of the other atoms, these single atomic states can not be eigenstates of the Hamilton operator for the solid.

In \( \mathbb{R}^d \) the free motion of a particle is usually described by the negative Laplace operator. A discrete analogue of the Laplacian is obtained when replacing differential operators by difference operators. Hence electrons which can travel ballistically through the solid should be described by

\[
(-\Delta \Psi)_n = - \sum_{m: \|n-m\|=1} (\Psi_m - \Psi_n) = - \sum_{m: \|n-m\|=1} \Psi_m + 2d\Psi_n ,
\]
where $\Psi = (\Psi_n)_{n \in \mathbb{Z}^d} \in \ell^2(\mathbb{Z}^d)$. It can be shown that this is actually a positive operator with spectrum $[0, 4d]$. However for simplicity one often neglects the term $2d\lambda V(n)$ and defines the discrete Laplacian by

$$(H_0 \Psi)_n = -\sum_{m:|n-m|=1} \Psi_m \quad \Leftrightarrow \quad H_0 |n\rangle = -\sum_{m:|m-n|=1} |m\rangle .$$

The spectrum of $H_0$ is now just shifted by $2d$ compared to the one of $-\Delta$. Anyway in physics only energy differences matter. An interpretation of this Hamiltonian is that the motion of the electron consists of hopping between nearest neighbors. This is the so-called ‘tightly bound binding approximation’ and the non-zero matrix elements $\langle n|H_0|m\rangle$ whenever $|n-m|=1$ are called hopping terms.

In some articles the free Hamilton operator may have a different sign, i.e. one has $-H_0$ instead, but $H_0$ and $-H_0$ are unitarily equivalent by the map $|n\rangle \mapsto (-1)^{|n|} |n\rangle$ where $|n|_1$ denotes the 1-norm of $n$. Although in different articles one may find different definitions of the discrete Laplacian on $\mathbb{Z}^d$, for the mathematical analysis it does not really matter which one you take.

Random impurities in a solid such as donors and acceptors lead to recombination which creates electric fields. These electric fields act on a single valence electron like a random potential. Therefore the Anderson model on the lattice $\mathbb{Z}^d$ now basically consists of a random family of Schrödinger operators given by the sum of $H_0$ and an i.i.d. real-valued random potential which may be multiplied by a coupling constant $\lambda$, i.e.

$$(H_{\lambda,\omega} \Psi)_n = ((H_0 + \lambda V_\omega) \Psi)_n = \sum_{|m-n|=1} \Psi_m + \lambda V_{\omega,n} \Psi_n, \quad n \in \mathbb{Z}^d .$$

Let us furthermore assume that the support of $V_{\omega,n}$ is bounded, then the operators $H_{\lambda,\omega}$ are bounded. As $V_{\omega,n}$ is real these operators are also symmetric and hence self-adjoint (for bounded operators the notions of symmetry and self-adjointness are the same).

The variable $\omega$ labels the randomness and the maps $\omega \mapsto V_{\omega,n}$ for $n \in \mathbb{Z}$ shall be real i.i.d. random variables. The usual way to obtain this is using a product probability space $(\Omega, \mathcal{F}) = (\Sigma, \mathbf{P})^\mathbb{Z}^d$. (In this work I may always neglect the $\sigma$-algebra in the terminology of probability spaces. Usually $\Sigma$ can be chosen to be some subset of a real vector space or some other polish space and the Borel $\sigma$-algebra is associated with it.) Then for a measurable function $\sigma \mapsto V(\sigma)$ on $\Sigma$ and $\omega = (\omega_n)_{n \in \mathbb{Z}^d} \in \Sigma^{\mathbb{Z}^d}$ one simply defines $V_{\omega,n} = V(\omega_n)$. The group $\mathbb{Z}^d$ now acts in a natural way on $\Omega$. For $m \in \mathbb{Z}^d$ this action is given by the shift $(T_m \omega)_n = \omega_{n+m}$. It is measure preserving and ergodic and hence the group of automorphisms $T_m$ is called metrically transitive [64, I.1.C]. On the other hand the group $\mathbb{Z}^d$ also acts unitarily on the Hilbert space $\ell^2(\mathbb{Z}^d)$ by the shift operators $(U_m \Psi)_n = \Psi_{n+m}$. Now one obtains $V_\omega(n + m) = V_{T_m \omega}(n)$ which leads to $U_m H_{\lambda,\omega} U_m^* = H_{\lambda, T_m \omega}$. Therefore one says $H_{\lambda,\omega}$ is a random, metrically transitive operator. An important result for such operators is the following. A proof can be found in [64].

**Theorem 1.1** For fixed $\lambda$, the pure point spectrum $\sigma_{pp}(H_{\lambda,\omega})$, the a.c. spectrum $\sigma_{ac}(H_{\lambda,\omega})$ and the s.c. spectrum $\sigma_{sc}(H_{\lambda,\omega})$ are almost surely non-random sets.
Note that the pure point spectrum $\sigma_{pp}$ here denotes the closure of the set of eigenvalues. Typically it consists of some interval in which the set of eigenvalues is dense.

Let me briefly state some known facts and conjectures about this model. Assume that the potential is not almost surely constant to avoid trivialities and let us also assume that the expectation of the potential $V_\sigma$ is zero, i.e. $E(V_\sigma) = \int d\sigma V_\sigma = 0$. For $\lambda = 0$ the free operator $H_0$ has pure a.c. spectrum supported on the interval $[-2d, 2d]$. In any dimension $d$ it is known that for sufficiently large disorder $\lambda$ or sufficiently small energy $E$ one has pure point spectrum with exponentially localized eigenstates [1, 24, 27, 28].

For $d = 1$ it is known that one has immediately pure point spectrum with exponentially localized eigenstates as soon as $\lambda \neq 0$ [64]. For the case of an absolutely continuous distributed potential $V_\sigma$ we will give a proof of the latter fact in the next sections. For singular distributions of the potential, such as a Bernoulli distribution, the analysis is more involved [16, 50].

For dimensions $d \geq 3$ one expects for low disorder to have pure a.c. spectrum in some interval within the band $[-2d, 2d]$ of the free operator but there is no proof so far. It is known that one also has pure point spectrum for low disorder at the edge of the spectrum. Hence for $d \geq 3$ there should be a metal-insulator-transition when increasing the disorder $\lambda$ or in energy. For $d = 2$ one expects pure point spectrum with exponentially localized eigenstates also for low disorder, but the localization length is expected to grow very fast, when $\lambda$ approaches zero.

Related to this different spectral types are so called level spacing statistics. Reducing the Hamilton operator to a large box one obtains a finite dimensional matrix acting on a finite dimensional Hilbert space. Hence one can diagonalize it (as it is self-adjoint) to obtain the eigenvalues. Then one can consider the spacing of these eigenvalues (distance of consecutive eigenvalues) and normalize it, such that the average is 1. There is a limit distribution of the normalized spacings when the large box converges to $\mathbb{Z}^d$. In the regime of strong Anderson localization with pure point spectrum one obtains Poisson statistics ($e^{-x}$) which corresponds to independent eigenvalues. The reason for this is, if the considered box is much larger than the localization length, then all the eigenstates are more or less located in small sub-boxes and the corresponding eigenvalues depend almost only on the potential values in these smaller boxes. Then the non intersecting sub-boxes lead to more or less independent eigenvalues. Poisson statistics was proved first by Molchanov [62] for a one-dimensional continuum random Schrödinger operator. For multidimensional Anderson models on a lattice a proof of Poisson statistics was established by Minami [60]. A crucial ingredient in Minami’s proof is an estimate of the probability of two or more eigenvalues in an interval which is today known as Minami estimate. It complements an estimate by Wegner on the probability to find one eigenvalue in a given interval [88]. Minami’s method and also the more recent proofs [6, 35] do not seem to extend easily to the continuum case. Therefore Minami-type estimates and Poisson statistics for continuum Anderson Hamiltonians is still a challenging question. Concerning the bottom of the spectrum of a continuum Anderson Hamiltonian, Poisson statistics was established quite recently [18].

The level statistics is different in the regime of expected a.c. spectrum. As this is related to extended states also the eigenvectors of $H$ restricted to a large box are not localized in
smaller subboxes. This leads to the fact that the eigenvalues have the tendency not to sit too close to each other, a phenomenon which is called level repulsion. Hence the limit spacing distribution vanishes at zero in this case. In the physics community the idea that the metallic phase of disordered systems is described by random matrix theory seems to be widely accepted. Depending on the symmetry class of the operator the level statistics according to numerical calculations is the same as obtained in random matrix theory for the gaussian unitary (GUE) or gaussian orthogonal ensemble (GOE) [2]. As these ensembles are called Wigner-Dyson ensembles one also speaks about Wigner-Dyson statistics. GUE corresponds to operators without time reversal symmetry, e.g. with magnetic phases in hopping terms, and GOE to operators with time reversal symmetry like the Anderson models I described above. A detailed discussion of random matrix ensembles is given in [58]. When increasing the disorder $\lambda$ in the at least 3 dimensional Anderson model ($d \geq 3$) one expects a metal-insulator-transition as explained above, also called Anderson transition. The level statistics at the Anderson transition and the cross over from Wigner-Dyson to Poisson statistics is quite of some interest [79] and similar to GUE (GOE) statistics above there seems to be a connection to other matrix ensembles [81]. More background on statistics of energy levels and a lot of references can be found in [61].

1.4 Pure point spectrum for the 1D Anderson model

Here I will discuss the methods used to prove Anderson localization in the 1D Anderson model in detail. First proofs were obtained by Goldsheid, Molchanov and Pastur [33], but the techniques used here are closer to Kunz-Soulliard [53]. More details of the proof given here are written out in [64]. The main tool is to rewrite the eigenvalue equation for $H$ in terms of transfer matrices and then study the products of these transfer matrices.

For the remainder of this chapter I will rather drop the coupling constant $\lambda$ as in one dimension the spectral type does not depend on $\lambda$. As above I will assume that the distribution of $V(\sigma)$ is bounded. Then the Hamiltonian for the 1D Anderson model is given by

$$ (H_{\omega} \Psi)_n = -\Psi_{n-1} - \Psi_{n+1} + V(\omega_n) \Psi_n, \quad \Psi = (\Psi_n)_{n \in \mathbb{Z}} \in \ell^2(\mathbb{Z}). \quad (1.1) $$

Before analyzing this model let me briefly show that the free Hamiltonian $H_0$, where the potential is equal to zero, has pure a.c. spectrum (of multiplicity 2) in the interval $[-2,2]$.

**Proposition 1.1** The spectrum of the free Hamiltonian $H_0$ is given by $\sigma(H_0) = [-2,2]$

**Proof.** First note that

$$ \|H_0 \Psi\|^2 = \sum_{n \in \mathbb{Z}} |\Psi_{n-1} + \Psi_{n+1}|^2 \leq \sum_{n \in \mathbb{Z}} (2|\Psi_{n-1}|^2 + 2|\Psi_{n+1}|^2) = 4\|\Psi\|^2 $$

and hence $\|H_0\| \leq 2$ and therefore $\sigma(H_0) \subset [-2,2]$. Now for $E \in [-2,2]$ there exists a $k \in [0,\pi]$ such that $E = -2\cos(k)$. Then the function $\Psi_n = e^{ikn}$ is a formal solution of the eigenvalue equation $H_0 \Psi = E \Psi$ but $\Psi$ is not an $\ell^2$ function. Therefore define $\Psi^m$ by $\Psi^m_n = \frac{1}{2m+1} e^{ikn}$ whenever $|n| \leq m$ and $\Psi^m_n = 0$ elsewhere. Then $\|\Psi^m\| = 1$ and
1. INTRODUCTION ON THE ANDERSON MODEL

\[ \lim_{m \to \infty} \| (H_0 - E) \Psi_m \| = 0. \] Hence \( E \) is an approximate eigenvalue and by the Weyl criterion therefore lies in the spectrum.

The main tool in the proof above was to use the formal solution \( H_0 \Psi = E \Psi \) to create approximate eigenfunctions which show that \( E \) is in the spectrum. This could be done because the behavior of the formal solution \( (\Psi_n)_n \) at infinity is not too bad. In physics literature such formal eigenfunctions which are not in \( \ell^2 \) but might be used to prove that certain energies are in the spectrum, are often called extended states. In fact in physics the terminology of extended states is often used instead of a.c. spectrum. However, mathematically one needs stronger arguments like the spectral decomposition in order to show that there is really a.c. spectrum. Vice versa if there is a.c. spectrum in a set \( A \) then for these types of 1D models it is known that there are Lebesgue almost surely non-\( \ell^2 \) extended states which grow at most polynomial at infinity. This is the so-called Pastur-Ishii theorem and a proof can be found in [19]. It implies that Lyapunov exponent has to be 0 in an interval where one has a.c. spectrum. The opposite is also true and was shown by Kotani [51].

For the free operator \( H_0 \) considered here one can explicitly write down the spectral decomposition with the help of the formal solutions to see that there is only a.c. spectrum. For convenience I will use Dirac notation for this part. Therefore let \( | \Psi, k \rangle \) denote the non-\( \ell^2(\mathbb{Z}) \) vector defined by \( \langle n | \Psi, k \rangle = e^{ikn} \). Then by anti-symmetry of the scalar product one has

\[ \int_{-\pi}^{\pi} \frac{dk}{2\pi} \langle n | \Psi, k \rangle \langle k, \Psi | m \rangle = \delta_{n,m}. \]

Hence in the weak topology one can write \( 1 = \int_{-\pi}^{\pi} | \Psi, k \rangle \langle k, \Psi | \) and one calls the integral a partition of unity. Now for an energy \( E = -2 \cos(k), \) let us define

\[ | \Psi, E, \pm \rangle = \left( \frac{2}{\sin(k)} \right)^{1/2} | \Psi, \pm k \rangle. \]

Then a change of variables shows

\[ 1 = \frac{1}{2\pi} \int_{-2}^{2} dE \left( | \Psi, E, + \rangle \langle \Psi, E, + | + | \Psi, E, - \rangle \langle \Psi, E, - | \right). \]

and furthermore one obtains for continuous functions \( f : \mathbb{R} \to \mathbb{C} \) that

\[ f(H_0) = \frac{1}{2\pi} \int_{-2}^{2} dE \ f(E) \left( | \Psi, E, + \rangle \langle \Psi, E, + | + | \Psi, E, - \rangle \langle \Psi, E, - | \right). \]

Considering \( \langle n | f(H_0) | n \rangle \) one sees that the spectral measure associated to the vector \( | n \rangle \) is given by \( dE \left( \langle n | \Psi, E, + \rangle \langle \Psi, E, + | + \langle n | \Psi, E, - \rangle \langle \Psi, E, - | \right) \) which is clearly absolutely continuous for all \( n \in \mathbb{Z}^d \). Hence the spectrum is absolutely continuous.

In mathematical terms a more rigorous proof of the pure a.c. spectrum can be given by the Fourier transform, which is implicitly used above. Therefore one maps \( \Psi \in \ell^2(\mathbb{Z}) \) on the function \( \hat{\Psi}(k) = \sum_n e^{ikn} \Psi(n) \). It is well known that this map identifies \( \ell^2(\mathbb{Z}) \) with \( L^2([\pi, \pi], \frac{dk}{2\pi}) \), where \( \frac{dk}{2\pi} \) denotes the normalized Lebesgue measure on \([\pi, \pi]\). Now it is not
very hard to see that this map transforms the operator $H_0$ to the multiplication operator $\hat{H}_0$ given by $\hat{H}_0 \Psi(k) = -2\cos(k)\Psi(k)$. For this operator it is obvious that the spectrum consists of the interval $[-2, 2]$ and is a.c. of multiplicity 2.

### 1.4.1 Transfer matrices and Lyapunov exponent

One can write the so-called stationary Schrödinger equation $H_\omega \Psi = E \Psi$ for an energy $E$, which is just an eigenvalue equation, in the following way

$$(H_\omega \Psi)_n = E \Psi_n \iff \begin{pmatrix} \Psi_{n+1} \\ \Psi_n \end{pmatrix} = \begin{pmatrix} V(\omega_n) - E & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \Psi_n \\ \Psi_{n-1} \end{pmatrix}.$$ 

For $\sigma \in \Sigma$ and an energy $E$ the matrix $T_\sigma^E = \begin{pmatrix} V(\sigma) - E & -1 \\ 1 & 0 \end{pmatrix}$ is called the transfer matrix at energy $E$ with potential $V(\sigma)$.

Now the asymptotic behavior for $n \to \pm \infty$ of a formal solution $\Psi_n$ of this equation can be studied by the asymptotic behavior of $N \to \infty$ of the random products $T_{\pm N}(\omega)$ defined by $T_{\pm N}^E(\omega) = \prod_{n=-N}^{N} T_{\omega_n}^E = T_{\omega_N}^E \cdots T_{\omega_1}^E$ and $T_{\pm N}(\omega) = \prod_{n=-N}^{N} (T_{\omega_n}^E)^{-1}$ respectively, where $N \in \mathbb{N}$. For products in a non-commutative group the convention in this work will always be $\prod_{n=1}^{N} T_n = T_N \cdots T_1$. The asymptotic behavior of these random products is described by the Lyapunov exponent.

**Proposition 1.2** The limits

$$\gamma_+^E = \lim_{N \to \infty} \frac{1}{N} \log (\| T_{\pm N}(\omega) \|)$$

exist almost surely and are non-random. Furthermore they are equal and one defines $\gamma_-^E = \gamma_+^E$. The quantity $\gamma^E$ is called the Lyapunov exponent for the sequence of random products $T_N^E$ or $T_{-N}^E$ respectively.

**Proof.** The existence of those limits will be shown in Chapter 2. The most important ingredient is the so-called subadditive ergodic theorem. Given these limits exist a.s. and are a.s. independent of $\omega$, let me show that they are equal. Therefore let $A := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, then by definition of the transfer matrix one gets $AT_\sigma^E A^{-1} = (T_\sigma^E)^{-1}$. Thus the matrix valued random variables $AT_N^E A^{-1}$ and $T_{-N}^E$ have the same distribution. As different norms are equivalent the matrices $T_N$ and $AT_N A^{-1}$ lead to the same Lyapunov exponent and one finds

$$\gamma^- = \mathbb{E}(\gamma^-) = \mathbb{E}(\gamma^+) = \gamma^+$$

almost surely. \hfill $\Box$

Oseledec’s ‘multiplicative ergodic theorem’ connects the asymptotic behavior of $T_N$ with that of $T_N v$ for non-zero vectors $v \in \mathbb{R}^2$. For a proof see e.g. [69].
1. INTRODUCTION ON THE ANDERSON MODEL

**Theorem 1.2** (Multiplicative Ergodic Theorem, Oseledec). Let a sequence of matrices $T_n \in \text{SL}(2, \mathbb{R})$ satisfy $\lim_{N \to \infty} \frac{1}{N} \log \|T_N\| = 0$. If $\gamma := \lim_{N \to \infty} \frac{1}{N} \log \|T_N \ldots T_1\| > 0$, then there exists a one-dimensional subspace $V^- \subset \mathbb{R}^2$ such that

$$\lim_{N \to \infty} \frac{1}{N} \log \|T_N \ldots T_1 v\| = -\gamma \quad \text{for } v \in V^- \quad \text{and}$$

$$\lim_{N \to \infty} \frac{1}{N} \log \|T_N \ldots T_1 v\| = \gamma \quad \text{for } v \notin V^-.$$  

Because of equation (1.3) the Lyapunov exponent will give the rate of decay for the eigenfunctions. Therefore the value $l^E = \frac{1}{\gamma}$ is a measure for the size of the region where the eigenstate and hence an electron in this eigenstate is concentrated and is called ‘localization length’.

Now consider briefly the situation of the free Hamiltonian where the potential is zero. In that case all transfer matrices are equal and one has

$$T_N^E = (T_1^E)^N = \begin{pmatrix} E & -1 \\ 1 & 0 \end{pmatrix}^N.$$

The characteristic polynomial of $T_1^E$ is $\det(T_1^E - \lambda 1) = \lambda^2 - \text{Tr}(T_1^E)\lambda + 1$ as the determinant is equal to 1. Therefore if $|E| = |\text{Tr}(T_1^E)| < 2$, the eigenvalues are $e^{\pm ik}$ for real $k$ with $E = -2 \cos(k)$ and $T_1^E$ is equivalent to diag($e^{ik}$, $e^{-ik}$). Then $T_1^E$ generates a compact group which forces the Lyapunov exponent to be zero. On the other hand if $|E| > 2$, the eigenvalues are of the form $\pm e^{\pm ik}$ (the sign depends on the sign of $E$) for real $k$ and the Lyapunov exponent is equal to $k$ where $2 \cosh(k) = |E|$. Finally if $E = \pm 2$ then $T_1^E$ is equivalent to a non-diagonalizable Jordan normal form and the growth of $\| (T_1^E)^N \|$ is linear which also implies that the Lyapunov exponent is zero.

Hence in this case the pure a.c. spectrum is supported on the set of energies where the Lyapunov exponent is zero. This is no coincidence. For ergodic operators such as the 1D Anderson model, Kotani theory [51, 82, 52] tells us that the support of the a.c. spectrum is precisely the set of energies where the Lyapunov exponent is zero. There is also a generalization to operators on a strip where one has more than just one Lyapunov exponent. In fact the number of zero Lyapunov exponents there coincides with the multiplicity of the a.c. spectrum [52]. In the definition of Lyapunov exponents as in the next chapter the value $-\gamma$ for this 1D model is also a Lyapunov exponent. Therefore if $\gamma = 0 = -\gamma$ in some interval then one has a.c. spectrum of multiplicity two as we obtained above for the discrete Laplacian.

The proof of localization is based on the fact that the Lyapunov exponent for the Anderson model is always positive, provided the potential is really random and not deterministic. For the random Hamiltonian as in (1.1) this simply means that $V(\sigma)$ is not almost surely constant.

The heuristic argument for localization in the case where the Lyapunov exponent is always positive, is the following. Any formal solution of the eigenvalue equation is either exponentially growing or exponentially decaying at $+\infty$ and $-\infty$. The exponential growth
either at $+\infty$ or at $-\infty$ is too bad for the solution to be an extended state. When the solution decays exponentially at both, $\pm \infty$ then it is clearly an element of $\ell^2(\mathbb{Z})$ and hence an eigenvector. This only happens for a countable set of energies though. However there are some technical difficulties to be dealt with.

1.4.2 Proof of pure point spectrum

In this section I will sketch the proof following the arguments of [64]. The most important ingredient is the positivity of the Lyapunov exponent which is a consequence of Furstenberg’s theorem [9, A.II.4].

**Proposition 1.3** If the potential $V_{\sigma}$ is not $p$-almost surely constant, then for any energy $E$, the Lyapunov exponent is positive.

Now let us start with the operator on the semi axis. This means we consider the restriction of $H_\omega$ on $\ell^2(\mathbb{N})$ with left-boundary condition $\Psi_0 = \tan(\alpha)\Psi_1$ and denote this operator by $H_{\omega,\alpha}^+$. This means for $n \geq 1$ the formula for $(H_{\omega,\alpha}^+\Psi)_n$ for $n \geq 2$ is just the same as above and for $n = 1$ one replaces $\Psi_{n-1} = \Psi_0$ by $\tan(\alpha)\Psi_1$. For any boundary condition $\alpha$ the operator $H_{\omega,\alpha}^+$ is self-adjoint. (Note that $V(\sigma)$ was assumed to be bounded leading to a bounded potential.)

Now for each energy $E$ there is exactly one formal solution to the eigenvalue equation $H_{\omega,\alpha}^+ u^E = Eu^E$. It is given by

$$
\begin{pmatrix}
  u^E_{n+1} \\
  u^E_n
\end{pmatrix} = T^E_n(\omega) \begin{pmatrix}
  \cos(\alpha) \\
  \sin(\alpha)
\end{pmatrix}.
$$

Note that these sequences $(u^E_n)_n$ are real. Similar to the calculation above for $H_0$ the extended states can be used to get the resolution of identity or projection valued measure associated to $H_{\omega,\alpha}^+$. This means there exists a positive measure $\rho_{\omega,\alpha}$ on $\mathbb{R}$ such that for any analytic function $f : \mathbb{R} \to \mathbb{C}$ one has [64, eq. (12.19)]

$$
\langle n | f(H_{\omega,\alpha}^+) | m \rangle = \int_\mathbb{R} \rho_{\omega,\alpha}(dE) u^E_n u^E_m f(E).
$$

From this equation it is obvious that $\sigma(H_{\omega,\alpha}^+) = \text{supp}(\rho_{\omega,\alpha})$ and that the type of the spectrum is given by the measure class of $\rho_{\omega,\alpha}$. As we will see in the sequel, only polynomial bounded solutions $(u^E_n)_n$ will contribute. Furthermore note that the spectral measure at $|1\rangle$ is given by $(u^E_1)^2 \rho_{\omega,\alpha} = \cos^2(\alpha) \rho_{\omega,\alpha}$.

The measure $\rho_{\omega,\alpha}$ is obtained in the following way as described in [64, Section 12]. Let $H_{\omega,\alpha}^{N,\beta}$ be the restriction of $H_{\omega,\alpha}^+$ to $\ell^2(\{1, \ldots, N\})$ with the right-boundary condition $\Psi_N = \Psi_N \tan(\beta)$. The eigenvalues of the matrix $H_{\omega,\alpha}^{N,\beta}$ are given by those energies $E_1, \ldots, E_N$ where the solutions $(u^E_n)_n$ satisfy the right-boundary condition parameterized by $\beta$. Set $N_N(E) = \sum_{n=1}^N |u^E_n|^2$ and define the measure $\rho_{\omega,\alpha}^{N,\beta}$ by

$$
\rho_{\omega,\alpha}^{N,\beta}(I) = \sum_{E_i \in I} (N_N(E_i))^{-1},
$$
where $I$ is some interval on $\mathbb{R}$. For $N \to \infty$ these measures converge in the weak-* topology to a limit which is independent of the boundary condition $\beta$ [64, eq. (12.17)] and this measure is precisely $\rho_{\omega,\alpha}$. 

$$w * \lim_{N \to \infty} \rho_{\omega,\alpha}^N = \rho_{\omega,\alpha}.$$ 

As $H^+_{\omega,\alpha}$ is self-adjoint, its spectrum is on the real line. Therefore $\| (H^+_{\omega,\alpha} - i)^{-1} \| \leq 1$ and hence $\| \Im m((H^+_{\omega,\alpha} - i)^{-1}) \| \leq 1$, where the imaginary part of an operator is defined as $\Im m(A) = \frac{1}{2}(A - A^*)$. This leads to

$$1 \geq \langle n | \Im m((H^+_{\omega,\alpha} - i)^{-1})|n \rangle \geq \int \rho_{\omega,\alpha}(dE) \frac{1}{E^2 + 1}(u_n^E)^2 \geq 0$$

Therefore the following sum converges

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \langle n | \Im m((H^+_{\omega,\alpha} - i)^{-1})|n \rangle \geq \int \rho_{\omega,\alpha}(dE) \sum_{n=1}^{\infty} \frac{(u_n^E)^2}{n^2} < \infty,$$

in particular

$$\sum_{n=1}^{\infty} \frac{(u_n^E)^2}{n^2} < \infty, \quad \text{for } \rho_{\omega,\alpha} \text{ a.e. } E.$$ 

(1.4)

This is a version of the Pastur-Ishii theorem. Now we are ready to prove the following.

**Theorem 1.3** For $P$ almost all $\omega$ one has $\sigma_{ac}(H^+_{\omega,\alpha}) = \emptyset$ and $\sigma_{ac}(H_\omega) = \emptyset$.

**Proof.** We already know that for every energy $E$ one has that for $P$ almost every $\omega$ the Lyapunov exponent $\gamma^E(\omega) = \gamma^E$ is positive. But the set of $\omega$ where this is true may heavily depend on $E$. However, Fubini’s theorem tells us, that for $P$ almost all $\omega$ one has that for Lebesgue almost all $E \in \mathbb{R}$ the Lyapunov exponent $\gamma^E(\omega)$ is positive. Now fix such an $\omega$ and call the set of energies $E$ where the Lyapunov exponent is positive $A$. Theorem 1.2 tells us that $u_n^E$ for $E \in A$ either grows or decays exponentially as $n \to \infty$. If it decays exponentially, then $u_n^E$ is an eigenvector and as $\ell^2(\mathbb{N})$ is separable, there are only countably many of them. Therefore the set $\tilde{A}$ of energies $E$ where $u_n^E$ grows exponentially is still a set of full Lebesgue measure. On the other hand by (1.4) one has $\rho_{\omega,\alpha}(\tilde{A}) = 0$. Therefore the measure $\rho_{\omega,\alpha}$ has no absolutely continuous part and hence $\sigma_{ac}(H^+_{\omega,\alpha}) = \emptyset$.

Similar to $H^+_{\omega,\alpha}$ one can restrict $H_\omega$ to $\ell^2(\mathbb{Z}^-)$ and fix a boundary condition $\beta$ to get an operator $H^-_{\omega,\beta}$. Analogue to above one would get that for $P$ almost all $\omega$ one has $\sigma_{ac}(H^-_{\omega,\beta}) = 0$. Now take an $\omega$ where $H^+_{\omega,\alpha}$ and $H^-_{\omega,\beta}$ have no a.c. spectrum which happens for $P$ a.e. $\omega$. Then $H_\omega = H^+_{\omega,\alpha} \oplus H^-_{\omega,\beta} + \Gamma$ where $\Gamma$ is a rank 2 operator. As the a.c. spectrum is stable under finite rank perturbations [83] one obtains

$$\sigma_{ac}(H_\omega) = \sigma_{ac}(H^+_{\omega,\alpha} \oplus H^-_{\omega,\beta}) = \sigma(H^+_{\omega,\alpha}) \cup \sigma_{ac}(H^-_{\omega,\beta}) = \emptyset.$$ 

This finishes the proof. \hfill \Box

The problem for proving pure point spectrum is that one doesn’t know if the Lyapunov exponent exists for $\rho_{\omega,\alpha}$ almost every energy $E$. To connect the Lebesgue measure with
the spectral measures $\rho_{\omega, \alpha}$ one uses spectral averaging. Now the difference of the operator $H_{\omega, 0}^+$ for $\alpha = 0$ and $H_{\omega, \alpha}^+$ is the rank one potential $\tan(\alpha)|1\rangle\langle 1|$. As furthermore the spectral measure at $|1\rangle$ is given by $\cos^2(\alpha)\rho_{\omega, \alpha}$ as mentioned before, the theory of rank one perturbations [83] yields
\[ \int_{-\pi/2}^{\pi/2} d\alpha \rho_{\omega, \alpha}(A) = |A| \]
for any measurable set $A \subset \mathbb{R}$, where $|A|$ denotes the Lebesgue measure of $A$. In fact this equality already holds for the measures $\rho_{\omega, 0}^{N, 0}$. Using this one can prove the following.

**Theorem 1.4** For $P$ a.e. $\omega$ one has that for Lebesgue a.e. $\alpha$ the operator $H_{\omega, \alpha}^+$ has pure point spectrum. An eigenstate at energy $E$ decays exponentially with rate $\gamma^E$, i.e. $u_n^E \sim \exp(-\gamma^E n)$.

**Proof.** For each $\omega$ we denote by $S(\omega)$ the set of energies, where either the Lyapunov exponent does not exist, or where $\gamma^E(\omega) \neq \gamma^E$. Any fixed $E$ is almost surely not in $S(\omega)$. Therefore by Fubini for a.e. $\omega$ the set $S(\omega)$ has zero Lebesgue measure. For any such $\omega$ one therefore has
\[ \int_{-\pi/2}^{\pi/2} d\alpha \rho_{\omega, \alpha}(S(\omega)) = 0. \]
Hence for Lebesgue a.e. $\alpha$, one has $\rho_{\omega, \alpha}(S(\omega)) = 0$. For such an $\alpha$ one has that for $\rho_{\alpha}$ a.e. energies $E$, the function $u_n^E$ either decays or grows exponentially with rate $\gamma^E$. But according to (1.4) it does not grow exponentially for $\rho_{\omega, \alpha}$ almost all $E$. Hence $u_n^E$ decays exponentially with rate $\gamma^E$ for $\rho_{\omega, \alpha}$ a.e. $E$ and forms an eigenstate of $H_{\omega, \alpha}^+$. As there are only countably many of them, $\rho_{\omega, \alpha}$ must be a point measure and therefore the spectrum of $H_{\omega, \alpha}^+$ is pure point.

To obtain pure point spectrum for a.e. operator $H_\omega$, we need some averaging quantity, similar to the boundary condition on the semi-axis. Furthermore we need some representative spectral measure. By the form of $H_\omega$ it is not hard to see that the vector space spanned by all the vectors $H_\omega^m|0\rangle, H_\omega^m|1\rangle$, for $m \in \mathbb{N}$, is dense in $l^2(\mathbb{Z})$. Hence the sum of the spectral measures $\rho_\omega = \rho_\omega^0 + \rho_\omega^1$, associated to the vectors $|0\rangle$ and $|1\rangle$ dominates all spectral measures. For $\omega \in \Omega$ we define $\bar{\omega}$ as the collection of all $\omega_n$ except for $n = 0, 1$. The distribution of $\bar{\omega}$ shall be denoted by $\bar{P}$. Next we define $H_{\bar{\omega}}$ like $H_\omega$ except that the potential at $|0\rangle$ and $|1\rangle$ shall be zero. Furthermore let $H_{\bar{\omega}, V, W} = H_{\bar{\omega}} + V|0\rangle\langle 0| + W|1\rangle\langle 1|$. Then one has $H_\omega = H_{\bar{\omega}, V, W}(\omega_0, V(\omega_1))$. The corresponding spectral measure for $H_{\bar{\omega}, V, W}$ will be denoted by $\rho_{\bar{\omega}, V, W}$. In this notation I do not require $V$ or $W$ to be potentials obtained from $\omega_0$ or $\omega_1$. By the theory of rank one perturbations [83] one obtains for any measurable $A \subset \mathbb{R}$
\[ \int_{\mathbb{R}} dV \rho_{\bar{\omega}, V, W}(A) = |A|, \quad \int_{\mathbb{R}} dW \rho_{\bar{\omega}, V, W}(A) = |A|. \]
Combining both leads to
\[ \int_{B \times \mathbb{R}} dV dW \rho_{\bar{\omega}, V, W}(A) \leq 2|A||B|, \quad (1.5) \]
for any Lebesgue measurable sets $A$ and $B$.

As the measure $\rho_\omega$ dominates all spectral measures, the spectral decomposition can be written in terms of extended states integrated w.r.t. $\rho_\omega$. However as the set of formal solutions $H_\omega u^E = E u^E$ is a 2 dimensional vector space, the spectrum may be in parts twice degenerate. Therefore there are solutions $u^E$ and $v^E$ such that
\[
\langle n | f(H_{\omega,V,W}) | m \rangle = \int \rho_{\omega,V,W}(dE) f(E) (u^E_n u^E_m + v^E_n v^E_m).
\]
At energies where the spectrum is not twice degenerate one may choose $v^E$ to be a multiple of $u^E$. A similar calculation as the one above shows that
\[
\sum_{n \in \mathbb{Z}} \frac{(u^E_n)^2 + (v^E_n)^2}{1 + n^2} < \infty. \tag{1.6}
\]

Now we have all ingredients to prove the pure point spectrum.

**Theorem 1.5** Let the distribution of the single site potential $V_\sigma$ be absolutely continuous. Then for $\mathbf{P}$ a.e. $\omega$, $H_\omega$ has pure point spectrum.

**Proof.** Similar to the case above for $\omega$ define the set $S(\omega)$ such that $\mathbb{R} \setminus S(\omega) = \{ E \in \mathbb{R} : \gamma^E_\pm \text{exist and } \gamma^E = \gamma^E_\pm \}$. By Fubini one has that for $\mathbf{P}$ a.e. $\omega$ the set $S(\omega)$ has Lebesgue measure zero. It is not hard to see that the change of one transfer matrix does not change the limit behavior of the random product of the transfer matrices. Hence the set $S(\omega)$ actually only depends on $\tilde{\omega}$ and we denote it therefore by $S(\tilde{\omega})$ and for $\mathbf{P}$ a.e. $\tilde{\omega}$ this has zero Lebesgue measure. Take such a $\tilde{\omega}$. By (1.5) $\int_{B \times B} dV dW \rho_{\tilde{\omega},V,W}(S(\tilde{\omega})) = 0$ for any compact interval $B$. Hence for Lebesgue a.e. $V, W$ one has $\rho_{\tilde{\omega},V,W}(S(\tilde{\omega})) = 0$. As the distribution of $V(\omega_0)$ and $V(\omega_1)$ are absolutely continuous, one obtains that for fixed $\tilde{\omega}$ chosen like above and $\mathbf{P} \times \mathbf{P}$ a.e. $(\omega_0, \omega_1)$ one has $\rho_\omega(S(\omega)) = 0$, where $\omega = (\tilde{\omega}, \omega_0, \omega_1)$.

This means the extended states $u^E_n, v^E_n$ either grow or decay exponentially with rate $\gamma^E$ for $n \to \pm \infty$ for $\rho_\omega$ a.e. $E$. By (1.6) for $\rho_\omega$ a.e. $E$ they cannot grow exponentially, hence they decay and are eigenstates. Therefore $H_\omega$ has pure point spectrum in this case.

Now $\omega = (\tilde{\omega}, \omega_0, \omega_1)$ and the situation above happens for $\mathbf{P}$ a.e. $\tilde{\omega}$ and $\mathbf{P} \times \mathbf{P}$ a.e. $\omega_0, \omega_1$ dependend on $\tilde{\omega}$. From this one might conclude by Fubini that for $\mathbf{P}$ a.e. $\omega$ one has pure point spectrum but for this one as to assure that this is a measurable condition which is not obvious.

But by the non-randomness of the spectrum according to Theorem 1.1 of Section 1.3 it is clear that the set of $\omega$ where $H_\omega$ has pure point spectrum, is measurable. Furthermore one already knows that this set has either measure zero or one. A proof of this theorem can be found in [19] or [64]. \qed

### 1.5 Integrated Density of States and Rotation number

Next let me introduce the so called 'Integrated Density of States' (short IDS or sometimes IDOS). At an energy $E$ of the family $(H_\omega)_{\omega \in \Omega}$ it can be almost surely defined by [64]
\[
\text{IDS}(E) = \lim_{N \to \infty} \frac{1}{N} \text{Tr}(\Pi_N 1_{(-\infty,E]}(\Pi_N H_\omega \Pi_N)),
\]
where $\Pi_N$ denotes the projection of $\ell^2(\mathbb{Z})$ on $\ell^2(\{1, \ldots, N\}) \subset \ell^2(\mathbb{Z})$, i.e. $\Pi_N u(n) = u(n)$ for $n = 1, \ldots, N$ and $\Pi_N u(n) = 0$ else. $1_{(-\infty, E]}$ denotes the characteristic function of $(-\infty, E]$. Thus the trace counts the number of independent eigenvectors (sum of dimensions of eigenspaces) in $\ell^2(\{1, \ldots, N\})$ of the operator $\Pi_N H_u \Pi_N$ that correspond to eigenvalues less or equal than $E$. (The projection operator before the characteristic function ensures that for $E \geq 0$ the infinitely many eigenvectors not lying in $\ell^2(\{1, \ldots, N\})$ are not counted.) It is obvious that this quantity increases with $E$ and hence associated to it is a measure on $\mathbb{R}$, the so called density of states (short DOS) measure. It is a quantitative description of the asymptotic number of eigenvalues at an energy $E$ when one considers the Hamiltonian restricted to a large box. Let us denote this number by $C^E_N(\omega)$ for the moment and the projected Hamiltonian by $H_N$, i.e.

$$H_{N, \omega} := \Pi_N H_u \Pi_N, \quad C^E_N(\omega) := \text{Tr}(\Pi_N 1_{(-\infty, E]}(H_{N, \omega})).$$

For a more formal definition of the IDS see [19, 9.2].

We will show that the IDS $C^E_N(\omega)$ equals another quantity coming from a dynamical system driven by the transfer matrices. Therefore consider the natural action $s_M$ of an invertible matrix $M \in \text{GL}(2, \mathbb{R})$ on $S^1 \cong \mathbb{R}/2\pi \mathbb{Z}$. Defining $e_\theta = (\cos(\theta), \sin(\theta))$ this action is given by

$$e_{s_M(\theta)} = \frac{M e_\theta}{\|M e_\theta\|}.$$

Actually the matrices even act on the linear subspaces of $\mathbb{R}^2$, thus one can project $s_M$ to an action on $\mathbb{RP}(1) \cong \mathbb{R}/\pi \mathbb{Z}$, for convenience this will be considered later. To get simpler notations let $s^E_\sigma := s_\sigma T^E_\sigma$.

These functions can be lifted to smooth functions $S^E_\sigma$ from $\mathbb{R}$ to $\mathbb{R}$. As $\text{det}(T^E_\sigma) = 1 > 0$ one finds by differentiating that $S^E_\sigma$ is monotonic increasing w.r.t. $\theta$. By the condition

$$\frac{1}{2\pi} < S^E_\sigma(\theta) - \theta < \frac{3}{2\pi}$$

which can be satisfied\(^3\) the lifted functions $S^E_\sigma$ are uniquely determined. Together with an initial value $\theta_0$ they define a Markov process $\theta_n^E(\theta_0)$ on $\Omega = \Sigma^\mathbb{Z}$ by conjunction, i.e. for $\omega \in \Omega$ define iteratively

$$\theta_0^E(\theta_0)(\omega) := \theta_0, \quad \theta_n^E(\lambda, \theta_0)(\omega) := S^E_{\omega_n}(\theta_{n-1}^E(\lambda, \theta_0)(\omega)) \quad (1.7)$$

The angles $\theta_n^E$ or $\theta_n^E$ mod $2\pi$ respectively are called Prüfer phases or Dyson-Schmidt variables. They may be interpreted as a discrete time random dynamical system on $\mathbb{R}$ or on $S^1$ or on $\mathbb{RP}(1)$.

\textbf{Definition 1.2} The $S^1$ rotation number $N$ for the sequence $T^E_{\omega_n}$, if existent, is defined as

$$N(E, \omega) = \frac{1}{\pi} \lim_{N \to \infty} \frac{1}{N} \theta_N^E(\theta_0)(\omega).$$

\(^3\)As $T^E_\sigma \epsilon_{\mathbb{R}} = e_{\mathbb{R}} \forall E, \sigma$ one sets $S^E_\sigma(\mathbb{R}) = \pi$ and as $S^E_\sigma$ projected on $\mathbb{RP}(1)$ is bijective the inequality is satisfied.
According to (1.8) one has

\[ \text{write the rotation number as Birkhoff sum of the Markov process} \]

is almost surely constant. Hence it is almost surely equal to its expectation value

\[ \theta \]

Theorem 1.6

For fixed \( \theta \), the rotation number \( N(E, \omega) \) is defined for almost all \( \omega \) and it is almost surely constant. Hence it is almost surely equal to its expectation value \( N(E) = E N(E, \omega) \).

Proof. Let \( \xi_N(\omega) = \theta_N^E(\pi, \omega) \). \( S \) denotes the ergodic shift operator on \( \Omega \). Furthermore let \( \hat{\theta} \) denote the projection of \( \theta \) modulo \( \pi \), i.e. \( \hat{\theta} \in [0, \pi) \) such that \( \theta - \hat{\theta} \in \pi \mathbb{Z} \). Then one has

\[
\xi_{N+M}(\omega) = \pi + \sum_{n=0}^{N+M-1} \Delta_{\omega_n} \theta_n^E = \xi_N(\omega) + \sum_{n=0}^{M-1} \Delta_{S_n \omega_n} \hat{\theta}_{N+n}^E
\]

\[
\leq \xi_N(\omega) + \hat{\theta}_N^E + \sum_{n=0}^{M-1} \Delta_{S_n \omega_n} \hat{\theta}_{N+n}^E = \xi_N(\omega) + \theta_M(\hat{\theta}_N^E, S^N \omega)
\]

\[
\leq \xi_N(\omega) + \xi_M(S^N \omega).
\]

The last step follows from the monotony of the functions \( S_n^E \), leading to the monotony of \( \theta_N^E(\theta_0, \omega) \) w.r.t. \( \theta_0 \). Thus the process \( \xi_N \) is subadditive. Furthermore we have \( E(\xi_N) \leq \pi + N \frac{\pi}{2} < \infty \) and \( \Gamma(\xi) = \inf (E(\xi_N)/N) \geq -\frac{\pi}{2} > -\infty \), therefore the subadditive ergodic theorem (Theorem 2.1) completes the proof. \[\square\]

Before proving the equality of rotation number and IDS let me point out that one can write the rotation number as Birkhoff sum of the Markov process \( \theta_n^E \) mod \( \pi \) on \( \mathbb{R} \mathbb{P}(1) \). According to (1.8) one has

\[
N(E) = \frac{1}{\pi} \lim_{N \to \infty} \frac{1}{N} E_\omega \sum_{n=1}^{N} \Delta_{\omega_n} \theta_n^E = \frac{1}{\pi} \lim_{N \to \infty} E_\omega \sum_{n=1}^{N} E_{\sigma} \Delta_{\sigma} \theta_n^E(\theta_n^E), \tag{1.9}
\]

where the last equation follows as \( \theta_n^E \) is independent of \( \omega_n \). As the function \( \theta \mapsto E_{\sigma} \Delta_{\sigma} \theta \) is \( \pi \).

The following proof can be found for a bigger class of Hamilton operators, e.g. in [45].
Theorem 1.7 (Sturm-Liouville) The IDS and the $S^1$ rotation number coincide, i.e. one has almost surely

$$N(E) = \lim_{N \to \infty} \frac{1}{N} (\theta^E_N(\omega)) = \lim_{N \to \infty} \frac{1}{N} C^E_N(\omega).$$

More precisely one has $\left| \frac{1}{\pi} \theta^E_N(\omega) - C^E_N(\omega) \right| \leq \frac{1}{2}$ for any $\omega$ if one starts with the initial phase $\theta_0 = 0$.

To prove this we first introduce some more notations. Similar to the subsection above let $u^E$ be a formal solution of the Schrödinger equation at energy $E$ with initial values $u^E_1 = \cos(\theta_0)$ and $u^E_0 = \sin(\theta_0)$, i.e. let

$$\begin{pmatrix} u^E_{n+1} \\ u^E_n \end{pmatrix} = T^E_\omega \begin{pmatrix} u^E_n \\ u^E_{n-1} \end{pmatrix} = T^E_\omega \begin{pmatrix} \cos(\theta_0) \\ \sin(\theta_0) \end{pmatrix}. $$

For convenience in the definition above and in the sequel we drop the dependence on $\omega$ in the notations. Furthermore define $R^E_n > 0$ by

$$R^E_n \begin{pmatrix} \cos(\theta^E_n) \\ \sin(\theta^E_n) \end{pmatrix} = \begin{pmatrix} u^E_{n+1} \\ u^E_n \end{pmatrix}. $$

Lemma 1.1 Starting with the same initial angle $\theta_0$ for all energies $E$ one gets the following:

$$\left(R^E_n\right)^2 \partial_E \theta^E_n = \begin{cases} \sum_{n=1}^N (u^E_n)^2 & \text{if } N > 0 \\ -\sum_{n=N+1}^0 (u^E_n)^2 & \text{if } N < 0 \end{cases} \tag{1.10}$$

Proof. If $u^E_n \neq 0$ one gets from the recurrence relation above

$$\cot(\theta^E_n) = -\frac{\tan(\theta^E_{n-1})}{\Delta^E_n} = \frac{V_\omega_n}{\Delta^E_n} - E .$$

As the condition $-\frac{\pi}{2} < \Delta^E_n < \frac{3\pi}{2}$ implies $\Delta^E_n(\frac{\pi}{2}) = \frac{\pi}{2}$ for all $E$ it follows that the functions $\Delta^E_n$ and therefore $\theta^E_n$ are smooth w.r.t. $E$. Thus one can differentiate the equation above w.r.t. $E$ to get

$$\partial_E \theta^E_n = \frac{\sin^2(\theta^E_n)}{\cos^2(\theta^E_{n-1})} \partial_E \theta^E_{n-1} + \sin^2(\theta^E_n).$$

Multiplying this with $(R^E_n)^2$ and using $u_n = \sin(\theta^E_n) R^E_n = \cos(\theta^E_{n-1}) R^E_{n-1}$ leads to

$$\left(R^E_n\right)^2 \partial_E \theta^E_n = \left(R^E_{n-1}\right)^2 \partial_E \theta^E_{n-1} + (u^E_n)^2. \tag{1.11}$$

If $u^E(n) = 0$ one can consider $\tan(\theta^E_n)$ to get the same result using similar steps. Iterating (1.11) gives (1.10) as $\partial_E \theta^E_0 = 0$ by definition. \[\square\]

Note that $\partial_E \theta^E_n$ is strictly positive for $n \geq 2$. 

Proof of Theorem 1.7. An eigenvector $u$ of $H_N$ in $\ell^2(\{1, \ldots, N\})$ has to fulfill the boundary conditions $u_0 = 0$ leading to an initial phase $\theta_0 = 0$ and $u_{N+1} = 0$ implying
\[ \theta_N \equiv \frac{\pi}{2} \mod \pi. \] For any \( c > 0 \) one can choose \( E \) small enough to get \( V_{\omega_n} - E > c \) for all \( n = 1, \ldots, N \). Using this one finds iteratively that \( u^E_N > 0 \) for \( E \) sufficiently close to \(-\infty\) as well as \( \lim_{E \to -\infty} u^E_N = 0 \) which means \( \lim_{E \to -\infty} \tan(\theta^E_N) = 0 \). From \(-\frac{\pi}{2} < \Delta^E_0 < \frac{\pi}{2}\) one gets \( \lim_{E \to -\infty} \Delta^E_0(0) = 0 \) for all \( \sigma \). Thus we find \( \lim_{E \to -\infty} \theta^E_N = 0 \) for all \( N \geq 0 \). As \( \theta^E_N \) is continuous and monotonic increasing w.r.t. \( E \) it follows that the \( j \)-th eigenvalue of \( H_N \) (counted from below, \( E^j_1 < E^j_2 < \ldots < E^j_N \)) with an eigenvector in \( \ell^2(\{1, \ldots, N\}) \) satisfies
\[
\theta^E_N = \frac{\pi}{2} + \pi(j - 1), \quad \theta_0 = 0.
\]
This oscillation theorem implies immediately
\[
\left| \frac{1}{\pi} \theta^E_N - C^E_N \right| \leq \frac{1}{2}. 
\]

1.6 Scaling of the Lyapunov exponent

As the inverse of the Lyapunov exponent is the localization length, the scaling w.r.t. a coupling constant of the potential is of quite some interest. Therefore we now put back the coupling constant and consider the operator
\[
(H_{\omega, \lambda} \Psi)_n = -\Psi_{n-1} - \Psi_{n+1} + \lambda V(\omega_n) \Psi_n.
\]

\( V(\sigma) \) is a random variable as above and we assume \( \mathbf{E}V(\sigma) = 0 \). Now clearly the transfer matrices, Prüfer phases defined as above, Lyapunov exponent and IDS depend all on \( \lambda \). In order to include this dependence in the notation of the transfer matrix, let us denote them by
\[
T^E_{\lambda, \sigma} = \begin{pmatrix} \lambda V(\sigma) - E & -1 \\ 1 & 0 \end{pmatrix},
\]

furthermore the Lyapunov exponent shall be denoted by \( \gamma^E_{\lambda} \). The scaling behavior of the Lyapunov exponent at energies within a band has been well understood for a long time. Thouless [87] found a perturbative formula for the Lyapunov exponent at an energy \( E = -2 \cos(k) \) in the band spectrum of the discrete Laplacian \( H_0 = H_{\omega,0} \) as defined above.
\[
\gamma^E_{\lambda} = \frac{E_\sigma(V^2(\sigma))}{8 \sin^2(k)} \lambda^2 + \mathcal{O}(\lambda^3) \quad (1.12)
\]

A rigorous proof was provided by Pastur and Figotin using the Prüfer phases \( \theta^E_0 \) [64]. They justified the so-called random phase approximation, stating that the Prüfer phases are distributed according to the Lebesgue measure on the unit circle, at least to lowest order after a suitable basis change. (In this section we consider the Prüfer phases as dynamics on the projective space.) However, as we will see, their argument breaks down at the band
1.6. SCALING OF THE LYAPUNOV EXPONENT

center \( k = \frac{\pi}{2}, \frac{3\pi}{2} \). Actually the random phase approximation is not correct there leading to anomalies in the perturbative formula (4.1) first found by Kappus and Wegner [47]. The scaling \( \lambda^2 \) is still valid but the coefficient is quite different. A perturbative formula as (1.12) for the band center with a control on the error terms was proven in [77].

Furthermore (1.12) has singularities at the band edges \( k = 0, \pi \) of the unperturbed operator \( H_0 \). The analysis there is more complicated and part of my work [72], written out in Chapter 4.

Now to prove (1.12) for \( 0 < |E| < 2 \) similar as in [64] we first need to write \( \gamma_E \) as Birkhoff sum for the Markov process of the Prüfer phases \( \theta_{\omega}^E \). Let us denote the product of the transfer matrices by \( T_{\lambda,\omega}^{E}(\omega) = \prod_{n=1}^{N} T_{\lambda,\omega}^{E} \). Furstenberg’s theorem does not only assure the positivity of the Lyapunov exponent but also gives

\[
\gamma_E = \lim_{N \to \infty} \frac{1}{N} \log \left( \| T_{\lambda,\omega}^{E} \| \right)
\]

\( \mathbf{P} \) almost surely for any initial angle \( \theta_0 \) [9]. Using

\[
\| T_{\lambda}^{E} \| = \| T_{\lambda}^{E} T_{\lambda}^{E-1} \| \| T_{\lambda}^{E-1} \| = \| T_{\lambda}^{E} \| \| T_{\lambda}^{E-1} \|
\]

inductively and taking expectation values, one obtains

\[
\gamma_E = \lim_{N \to \infty} \frac{1}{N} \mathbf{E}_\omega \log \left( \prod_{n=1}^{N} \| T_{\lambda}^{E} \| \right) = \lim_{N \to \infty} \frac{1}{N} \mathbf{E}_\omega \sum_{n=1}^{N} \log \left( \| T_{\lambda}^{E} \| \right)
\]

where the last equation follows from the fact that \( \theta_{\omega}^{E-1} \) is independent of \( \omega_n \) and \( \mathbf{P} \) is a product measure. Hence similar to the IDS (rotation number) the Lyapunov exponent can be obtained as Birkhoff sum of the Markov process \( \theta_{\omega}^{E} \).

By the equivalence of norms the Lyapunov exponent is independent of the used matrix norm. For any invertible matrix \( M \) the map \( T \mapsto ||MTM^{-1}|| \) also defines a norm. Hence the Lyapunov exponent is invariant under conjugation of the transfer matrices. Such a transformation changes the Prüfer phases and the dynamics on \( \mathbb{RP}(1) \). For the corresponding modified Prüfer phases and transfer matrices the same formula for the Lyapunov exponent holds. We want the modified Prüfer phases \( \tilde{\theta}_n \) to be Lebesgue distributed to lowest order on the projective space. Therefore we use a basis change \( M \) such that the free transfer matrix \( (\lambda = 0) \) is transformed to a rotation matrix. For a fixed energy \( E = -2 \cos(k), |E| < 2 \), a possible choice is \( M = \begin{pmatrix} \sin(k) & 0 \\ -\cos(k) & 1 \end{pmatrix} \). The one obtains

\[
MT_{\lambda,\sigma}^{E}M^{-1} = \begin{pmatrix} \cos(k) & -\sin(k) \\ \sin(k) & \cos(k) \end{pmatrix} \begin{pmatrix} 1 & -\frac{\lambda}{\sin(k)} \left( \begin{array}{c} 0 \\ V(\sigma) \end{array} \right) \end{pmatrix}. 
\]
A small calculation shows
\[
E_\sigma \log(\|MT_E^{M^{-1}}e_\theta\|) = \frac{\lambda^2 E(V(\sigma))^2}{2 \sin^2(k)} \left( \cos^2(\theta) - 2 \cos^2(\theta) \sin^2(\theta) \right) + O(\lambda^3)
\]
\[
= \frac{\lambda^2 E(V(\sigma))^2}{8 \sin^2(k)} \left( 1 + e^{2i\theta} + e^{-2i\theta} + \frac{1}{2} e^{4i\theta} + \frac{1}{2} e^{-4i\theta} \right) + O(\lambda^3)
\]
\hspace{2cm} (1.13)

For the dynamics of the corresponding \( M \)-modified Prüfer phases one obtains \( \tilde{\theta}_n = k + \tilde{\theta}_{n-1} + O(\lambda) \) with uniform error in \( \tilde{\theta}_{n-1} \). Therefore one has
\[
\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} e^{2i\tilde{\theta}_n} = \lim_{N \to \infty} \frac{1}{N} e^{2ik} \sum_{n=0}^{N-1} e^{2i\tilde{\theta}_n} + O(\lambda).
\]

implying
\[
(1 - e^{2ik}) \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} e^{2i\tilde{\theta}_n} = 0.
\]

An analogue calculation shows
\[
(1 - e^{4ik}) \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} e^{4i\tilde{\theta}_n} = 0.
\]

Therefore if \( e^{2ik} \neq 1 \) and \( e^{4ik} \neq 1 \) which is true for \( k \in (0, \pi), k \neq \frac{\pi}{2} \), the non constant terms in (1.13) when averaging as Birkhoff sum vanish. Then one obtains
\[
\gamma^E = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} E_\sigma \frac{1}{2} \log \left( \|T_{E}^{E} e_{\theta_{n-1}}\| \right) = \lambda^2 E(V(\sigma))^2 \frac{8 \sin^2(k)}{8 \sin^2(k)} + O(\lambda^3).
\]

This is precisely Thouless formula.

The proof shows that it is enough to apply the random phase hypothesis to the functions \( e^{\pm 2i\theta} \) and \( e^{\pm 4i\theta} \). In the band center \( k = \frac{\pi}{2} \) one has \( e^{4ik} = 1 \) and one can not conclude with the easy argument made above. As stated above, it turns out that Thouless formula is indeed wrong at the band center. At the band edges \( E = \pm 2 \) this technique just fails because the transfer matrix is not equivalent to a rotation matrix. There the analysis is more involved and done in Chapter 4.
Chapter 2

Theory on products of random matrices

2.1 Lyapunov exponents

The asymptotic behavior of products of matrices may be characterized by Lyapunov exponents. Let \((T_n)_{n \in \mathbb{N}} \in \text{GL}(m, \mathbb{R})\) be a sequence of matrices. For \(N \in \mathbb{N}\) let \(T_N = \prod_{n=0}^{N-1} T_{N-n} = T_N \cdots T_1\) and let \(d_{1,N} \geq d_{2,N} \geq \ldots \geq d_{k,N} > 0\) denote the singular values of \(T_N\), i.e. the eigenvalues of \(T_N^* T_N\). Then there are unitary matrices \(U_N\) and \(V_N\) such that \(U_N T_N V_N = \text{diag}(d_{1,N}, \ldots, d_{k,N}) = D_N\).

**Definition 2.1** If existent, the \(k\)-th Lyapunov exponent for the sequence \(T_N\) of products of matrices is defined by

\[
\gamma_k = \lim_{N \to \infty} \frac{1}{N} \log(d_{k,N}) .
\]

We want to see that if the \(T_n\) are random i.i.d. matrices, then all Lyapunov exponents exist almost surely. To obtain this, we need to write the definitions for the Lyapunov exponents in another way following the arguments of [9]. For \(k \leq m\) define the \(k\)-th antisymmetric tensor product of \(T_N\) acting on \(\Lambda^k \mathbb{R}^m\) by \((\Lambda^k T_N)(v_1 \wedge \ldots \wedge v_k) = T_N v_1 \wedge \ldots \wedge T_N v_k\), where \(v_l \in \mathbb{R}^m, l = 1, \ldots, k\). Then \((\Lambda^k U_N)(\Lambda^k T_N)(\Lambda^k V_N) = \Lambda^k D_N\) which is a diagonal matrix and \(\Lambda^k U_N\) and \(\Lambda^k V_N\) are unitary on \(\Lambda^k \mathbb{R}^m\). The norm and greatest eigenvalue of \(\Lambda^k D_N\) can easily be seen to be the product \(\prod_{l=1}^k d_{l,N}\). Hence one has

\[
\log \|\Lambda^k T_N\| = \log \|\Lambda^k D_N\| = \sum_{l=1}^k (\log d_{l,N}) .
\]

This leads to the following.

**Proposition 2.1** If all Lyapunov exponents for \(k = 1, \ldots, m\) exist, then they may be defined iteratively by

\[
\sum_{l=1}^k \gamma_l = \lim_{N \to \infty} \frac{1}{N} \log \|\Lambda^k T_N\| .
\]
In particular, the sum of the $k$ biggest Lyapunov exponents of the sequence $(T_N)_N$ coincides with the biggest Lyapunov exponent of the sequence $(\Lambda^k T_N)_N$.

Now let $(\Sigma, \mathbf{p})$ be some probability space, $\Sigma \ni \sigma \mapsto T_\sigma \in \text{GL}(m, \mathbb{R})$ some measurable function such that $E(\log(\det(T_\sigma))) > -\infty$ and $E(|\log(||T_\sigma||)|) < \infty$ which means $\log(||T_\sigma||)$ and $\log(\det(T_\sigma))$ are integrable w.r.t. $\mathbf{p}$. Here $E$ denotes the expectation value w.r.t. $\mathbf{p}$. Furthermore let $(\Omega, \mathcal{P}) = (\Sigma^Z, \mathbf{P})$ be the product space. Similar to above define for $\omega \in \hat{\Omega} = \Sigma^Z$ and $N \in \mathbb{N}$

$$T_N(\omega) = \prod_{n=0}^{N-1} T_{\omega_{N-n}} = T_{\omega_N} T_{\omega_{N-1}} \ldots T_{\omega_1}$$

We need the following result from probability theory. A nice proof for a more general situation can be found in [46] or any other book about probability theory.

**Theorem 2.1** (Subadditive Ergodic Theorem). Let $(\Omega, \mathcal{A}, \mathbf{P})$ be a probability space and $\mathcal{S}$ be a measure preserving transformation, i.e. $P(S^{-1}(A)) = P(A), \forall A \in \mathcal{A}$. If $\xi_N$ is a real valued, subadditive process, i.e.

$$\xi_{N+M}(\omega) \leq \xi_N(\omega) + \xi_M(\mathcal{S}^N \omega) \quad \text{almost surely},$$

satisfying $E(|\xi_N|) < \infty$ for each $N$ and $\Gamma(\xi) := \inf E(\xi_N)/N > -\infty$, then $\xi_N(\omega)/N$ converges almost surely. If $\mathcal{S}$ is ergodic, then

$$\lim_{N \to \infty} \frac{1}{N} \xi_N(\omega) = \Gamma(\xi) = \inf \frac{E(\xi_N)}{N} \quad \text{almost surely}.$$  

Now we can prove the convergence for the Lyapunov exponents.

**Theorem 2.2** For almost all $\omega$ the Lyapunov exponents $\gamma_k$, $k = 1, \ldots, m$, exist and are independent of $\omega$.

**Proof.** For the space $\Omega = \Sigma^Z$ with product measure $\mathbf{P}^Z$ the shift operator $(S \omega)_n := \omega_{n+1}$ is ergodic. Let $\xi^k_N(\omega) := \log ||\Lambda^k T_N(\omega)||$. Then one has

$$\xi^k_{N+M}(\omega) = \log ||\Lambda^k T_{N+M}(\omega)|| = \log ||(\Lambda^k T_M(S^N \omega))(\Lambda^k T_N(\omega))||$$

$$\leq \log (||\Lambda^k T_M(S^N \omega)|| ||\Lambda^k T_N(\omega)||) = \xi^k_N(\omega) + \xi^k_M(\mathcal{S}^N \omega),$$

thus the process is subadditive. Furthermore we have

$$E_{\omega}(||\xi^k_N(\omega)||) \leq E_{\omega} \sum_{n=1}^N (\log ||\Lambda^k T_{\omega_n}||) \leq k N E_{\sigma} (\log ||T_\sigma||) < \infty$$

by assumption. As $\xi^m_N = \log(\det(T_N)) = \sum_{n=1}^N \log(\det(T_{\omega_n}))$ one has $\inf_N E(\xi^m_N)/N \geq \frac{k}{m} \inf_N E(\xi^m_N)/N = \frac{k}{m} E(\log(\det(T_\sigma))) > -\infty$. Thus one can apply Theorem 2.1 to get the desired result. \hfill \Box
As the Lyapunov exponents are almost surely existent and constant, one may also define them by taking expectation values

$$\sum_{l=1}^{k} \gamma_k = \lim_{N \to \infty} \frac{1}{N} \mathbb{E}_\omega (\log \| \Lambda^k T_N(\omega) \|) .$$

### 2.2 Multiplicative Ergodic Theorem

Now the singular values $d_{k,N}$ as defined above are just the eigenvalues of the positive matrix $(T_N^* T_N)^{\frac{1}{2}}$ and hence the Lyapunov exponents are just the limits of the eigenvalues of $\frac{1}{N} \log(T_N^* T_N)^{\frac{1}{2}}$. Osceledec’s ‘multiplicative ergodic theorem’ shows that for almost all $\omega$, not only the eigenvalues of this matrix converge, but also the eigenspaces converge. Here we just state the theorem, for a proof see [32].

**Theorem 2.3 (Multiplicative Ergodic Theorem, Osceledec)**

(i) For almost all $\omega$ the limit

$$\lim_{N \to \infty} (T_N^*(\omega) T_N(\omega))_N^{\frac{1}{2N}} = \Lambda(\omega)$$

exists and for almost all $\omega$ the eigenvalues of $\Lambda(\omega)$ coincide with the values $e^{\gamma_k}$ for $k = 1, \ldots, m$.

(ii) Let $\exp(\alpha_1(\omega)) < \exp(\alpha_2(\omega)) < \ldots < \exp(\alpha_s(\omega))$ denote the different eigenvalues of $\Lambda(\omega)$ and let $U_1(\omega), U_2(\omega), \ldots, U_s(\omega)$ denote the corresponding eigenspaces. Now set $V_0(\omega) = \{0\}$ and iteratively $V_i(\omega) = U_1(\omega) \oplus \ldots \oplus U_i(\omega)$. Then one obtains for $\mathbb{P}$ almost every $\omega$ that for $v \in V_i(\omega) \setminus V_{i-1}(\omega)$

$$\lim_{N \to \infty} \frac{1}{N} \log \| T_N(\omega) v \| = \alpha_i(\omega) .$$

Note that by (i) the set of values $\alpha_l(\omega)$ for $l = 1, \ldots, s$ coincides almost surely with the set of values $\gamma_k$ for $k = 1, \ldots, m$. For instance one has almost surely $\alpha_1(\omega) = \gamma_m$ and $\alpha_s(\omega) = \gamma_1$.

Part (ii) of the theorem above shows that there is a $\omega$ dependent flag of subspaces $V_i(\omega)$ where the values of the limits $\frac{1}{N} \log \| T_N v \|$ hop from the smallest to the biggest Lyapunov exponents. If one decomposes $v$ in terms of eigenvalues of $\Lambda(\omega)$ then always the part corresponding to the biggest Lyapunov exponent determines the limit. The function mapping $v$ to this limit can be viewed as a map on the projective space $\mathbb{R}P(m-1)$ which has a natural Haar measure $\mu$ defined by the property that it is invariant under the action of the orthogonal group. Then for fixed $\omega$ the set of equivalence classes $[v] \in \mathbb{R}P(m-1)$ for which $v \in V_{s-1}(\omega)$ is a set of measure zero. Hence one obtains the following.
Theorem 2.4 Let \( \mu \) denote the Haar measure on \( \mathbb{R}P(m-1) \). Then one has

\[
\gamma_1 = \lim_{N \to \infty} \frac{1}{N} \int_{\mathbb{R}P(m-1)} E \left( \log \| T_N v \| \right) d\mu([v])
\]

where \( v \) denotes a vector of length 1 corresponding to the equivalence class \([v]\).

One of course gets similar formulas for the sums of the \( k \) biggest Lyapunov exponents using the antisymmetric tensor products. The right hand side of the equation above can be expanded into a sum. Therefore set \( v = v_0 \) and define \( v_n \) by

\[
\| T_N v_0 \| = \sum_{n=0}^{N-1} E \sigma \log (\| T_\sigma v_n \|)
\]

for the last equation note that \( v_{n-1} \) is independent of \( \omega_n \). As the \( v_n \) are unit vectors the values \( \log (\| T_\sigma v_n \|) \) only depend on the equivalence class \([v_n]\) on \( \mathbb{R}P(m-1) \). The sequence \([v_n]\) is just the Markov process on the projective space induced by the action of the family \( T_\omega \). If this Markov process is ergodic then there is a unique invariant probability measure \( \nu \) on \( \mathbb{R}P(m-1) \) for the action of \( T_\sigma \) and for any starting vector \( v = v_0 \) and almost all \( \omega \) one has

\[
\sum_{n=0}^{N-1} E \sigma \log (\| T_\sigma v_n \|) = \int_{\mathbb{R}P(m-1)} E \sigma \log (\| T_\sigma v \|) d\nu([v])
\]

Therefore we finally obtain the following.

Corollary 2.1 If the process \([v_n]\) as defined above is ergodic and \( \nu \) is the unique invariant probability measure on \( \mathbb{R}P(m-1) \) then we get Furstenberg’s formula:

\[
\gamma_1 = \int_{\mathbb{R}P(m-1)} E \sigma (\log \frac{\| T_\sigma v \|}{\| v \|}) d\nu([v])
\]

Similar results as Theorem 2.4 and Corollary 2.1 hold for the sum of the first \( k \) Lyapunov exponents when considering the dynamics on the space \( \mathbb{P}(\Lambda^k \mathbb{R}^m) \). (In this notation \( \mathbb{P}(\mathbb{R}^2) = \mathbb{R}P(1) \).) However, because of the special structure of the matrices \( \Lambda^k T_N \) they leave the set of vectors of the form \( v_1 \wedge \ldots \wedge v_k \) invariant. Therefore one may only consider the dynamics on the submanifold of \( \mathbb{P}(\Lambda^k \mathbb{R}^m) \) given by the equivalence classes of such vectors. By antisymmetry of the wedge product, vectors of that form are not zero only if \( v_1, \ldots, v_k \) are linearly independent. Furthermore two non-zero vectors \( v_1 \wedge \ldots \wedge v_k \) and \( w_1 \wedge \ldots \wedge w_k \) are linear dependent, precisely if the tuples \( (v_1, \ldots, v_k) \) and \( (w_1, \ldots, w_k) \) span the same vectorspace. Therefore this submanifold is diffeomorphic to the Grassmann manifold of \( k \)-dimensional subspaces of \( \mathbb{R}^m \). Hence if the process on this Grassmann manifold is ergodic, one obtains similar formulas as (2.1) and (2.2) for the sum \( \sum_{i=1}^{k} \gamma_k \) as an integral over the Grassmann manifold.
2.3. DISTINCT LYAPUNOV EXPONENTS

Furthermore if the considered products of random matrices are in a sub group of SL($m$, $\mathbb{R}$) which leaves a certain submanifold of the Grassmannian invariant, one may reduce the consideration of the dynamics to this manifold. This is for example the case if all matrices are symplectic which will be the case for the transfer matrices of random Schrödinger operators.

If one wants to calculate all Lyapunov exponents at once with only one invariant measure then one may consider the dynamics on flags of subspaces of $\mathbb{R}^m$ which gives a so called flag manifold. The appropriate flag manifolds for matrices in the Lorentz group $U(L, L)$ or $U(L, L, \mathbb{R})$ is described in Appendix A.6. The Lorentz groups are equivalent to the symplectic groups $Sp(2L, \mathbb{C})$ and $Sp(2L, \mathbb{R})$ respectively.

2.3 Distinct Lyapunov exponents

The proof of the pure point spectrum for the 1D Anderson model in Section 1.4 is based on the positivity of the Lyapunov exponent. A criterion for $\gamma_1 > 0$ was found by Furstenberg [29, 30] and is stated in Theorem 2.5 below. His result was improved by Guivarc’h and Raugi who gave a criterion to find the least $k$ such that $\gamma_1 > \gamma_k$. In particular they gave a criterion implying $\gamma_1 > \gamma_2$ (cf. Theorem 2.7) which can be generalized to obtain a criterion for $\gamma_k > \gamma_{k+1}$. Proofs for these theorems and more background information concerning this topic can be found in [9, A.III]. Before stating the results I have to introduce some notations.

**Definition 2.2** A subset $M \subset GL(m, \mathbb{R})$ is called strongly irreducible, if there is no finite union $W = V_1 \cup \ldots \cup V_l \subset \mathbb{R}^m$ of proper subspaces of $\mathbb{R}^m$ such that $T(W) \subset W$, $\forall T \in M$. A subset $M \subset GL(m, \mathbb{R})$ is called contracting, if there is a sequence $T_n \in M$ of matrices such that $\lim_{n \to \infty} \|T_n\|^2 \|\Lambda^2 T_n\|^{-1} = \infty$.

In the case where all random matrices have determinant one, Furstenberg found the following.

**Theorem 2.5** (Furstenberg) Let $T_\sigma$ be supported in $SL(m, \mathbb{R})$ and let $G$ be the smallest closed group in $SL(m, \mathbb{R})$ which contains the (essential) support of $(T_\sigma)_{\sigma \in \Sigma}$ (w.r.t. $p$). Let us assume that $G$ is strongly irreducible and non compact. Then the upper Lyapunov exponent is positive ($\gamma_1 > 0$).

In the case of $SL(2, \mathbb{R}) = Sp(2, \mathbb{R})$ matrices it is shown in [9, A.II.4] that the irreducible condition can be replaced by a simpler statement giving the following.

**Theorem 2.6** (Furstenberg) Let $G$ be the smallest closed sub group in $SL(2, \mathbb{R})$ which contains the (essential) support of $(T_\sigma)_{\sigma \in \Sigma}$ w.r.t. $p$. Let us assume that

(i) $G$ is not compact,

(ii) for any $\bar{x} \in \mathbb{RP}(1)$ the set $\{T\bar{x} : T \in G\}$ has more than two elements, where $\bar{x}$ denotes the equivalence class of $x \in \mathbb{R}^2$ in $\mathbb{RP}(1)$. 


Then there exists a unique invariant distribution \( \nu \) on \( \mathbb{RP}(-1) \) for the Prüfer phases and \( \nu \) is continuous. Furthermore the (first) Lyapunov exponent is positive.

In order to get \( \gamma_1 > \gamma_2 \) and a unique invariant measure for higher dimensional matrices one also needs the contraction property. To understand this better note that if the sequence \( T_n \) is a typical sequence of the random products which is described by the Lyapunov exponents, then \( \|T_n\|^2 \|\Lambda^2 T_n\|^{-1} \sim \exp(n(2\gamma_1 - (\gamma_1 + \gamma_2))) = \exp(n(\gamma_1 - \gamma_2)) \). Hence if \( \gamma_1 > \gamma_2 \) then \( T \) necessarily fulfills the contracting property. The content of the following theorem is due to Guivarc’h and Raugi (cf. Sections A.III.4 and A.III.6 in [9]).

**Theorem 2.7** (Guivarc’h, Raugi) Let \( T \) be the smallest closed semi group in \( \text{GL}(m, \mathbb{R}) \) which contains the (essential) support of \( (T_\sigma)_{\sigma \in \Sigma} \) (w.r.t. \( p \)). Let us assume that \( T \) is strongly irreducible and contracting. Then there exists a unique invariant distribution \( \nu \) on \( \mathbb{RP}(m - 1) \) for the process \( v_n \) and \( \nu \) is continuous. Furthermore the first two Lyapunov exponents are distinct, i.e. \( \gamma_1 > \gamma_2 \).

Goldsheid and Margulis [32] generalized these theorems by showing that instead of considering the semi group \( T \) it is enough to consider its Zariski or algebraic closure. The Zariski topology on algebraic varieties like \( \text{SL}(m, \mathbb{R}) \) is defined as follows: A set \( M \subset \text{SL}(m, \mathbb{R}) \) of matrices is Zariski closed, if \( M \) is equal to the intersection of the zero sets of all polynomial functions which vanish on \( M \). In other words, the closed sets are precisely the intersections of zero sets of polynomials.

Clearly, if \( T \) is strongly irreducible then its Zariski closure also is. We will see that the other inclusion is true as well. Assume \( \overline{T} \) is strongly irreducible and \( T \) is not. Then there is a finite union of proper subspaces \( \mathcal{W} = V_1 \cup \ldots \cup V_l \) which is invariant under the action of \( T \). The condition \( T(\mathcal{W}) \subset V_i \) is equivalent to \( (v_i, Tw) = 0 \) for \( v \in V_i^\perp \) and \( w \in \mathcal{W} \), where \( (\cdot, \cdot) \) denotes the usual scalar product on \( \mathbb{R}^m \). The condition \( T(\mathcal{W}) \subset \mathcal{W} \) can hence be described by \( \prod_{i=1}^l (v_i, Tw) = 0 \) \( \forall v_i \in V_i^\perp, w \in \mathcal{W} \). The set of all such matrices \( T \) is Zariski closed and therefore contains \( \overline{T} \). Hence \( \overline{T} \) can not be strongly irreducible and we get a contradiction! Hence if \( \overline{T} \) is strongly irreducible then also \( T \) is. Furthermore Goldsheid and Margulis proved the following [32, Theorem 6.3]

**Theorem 2.8** Let \( T \subset \text{GL}(\mathbb{R}^m) \) be a semi group such that its algebraic closure is strongly irreducible and contracting. Then \( T \) is strongly irreducible.

This leads to the following Corollary.

**Corollary 2.2** Let \( T \) be the smallest closed semi group in \( \text{GL}(m, \mathbb{R}) \) which contains the (essential) support of \( (T_\sigma)_{\sigma \in \Sigma} \) (w.r.t. \( p \)) and assume that its algebraic closure \( \overline{T} \) is strongly irreducible and contracting. Then there exists a unique invariant distribution \( \nu \) on \( \mathbb{RP}(m - 1) \) for the process \( v_n \) and \( \nu \) is continuous. Furthermore the first two Lyapunov exponents are distinct, i.e. \( \gamma_1 > \gamma_2 \).

\(^{1}\text{By continuous I mean the measure has no pure point part. The statement does not mean that the measure is absolutely continuous.}\)
These methods can also be used to prove that two consecutive Lyapunov exponents, $\gamma_k$ and $\gamma_{k+1}$ are distinct. Let us consider the sequence $\Lambda^k T_N$ instead of $T_N$ and call the corresponding first two Lyapunov exponents $\hat{\gamma}_1$ and $\hat{\gamma}_2$ (assume that $k < m$). We have already seen that the top Lyapunov exponent is equal to the sum $\hat{\gamma}_1 = \sum_{l=1}^k \gamma_l$. If one considers the singular values one realizes that the second Lyapunov exponent is given by $\hat{\gamma}_2 = \sum_{l=1}^{k-1} \gamma_l + \gamma_{k+1}$. Hence $\gamma_k > \gamma_{k+1}$ if and only if $\hat{\gamma}_1 > \hat{\gamma}_2$. This leads to the following definitions and corollary.

**Definition 2.3** A subset $\mathbb{M} \subset \text{GL}(m, \mathbb{R})$ is called $k$-strongly irreducible, if there is no finite union $\mathbb{W} = V_1 \cup \ldots \cup V_l \subset \Lambda^k \mathbb{R}^m$ of proper subspaces of $\Lambda^k \mathbb{R}^m$ such that $\Lambda^k T(\mathbb{W}) \subset \mathbb{W}$, $\forall T \in \mathbb{M}$. A subset $\mathbb{M} \subset \text{GL}(m, \mathbb{R})$ is called $k$-contracting, if there is a sequence $T_n \in \mathbb{M}$ of matrices such that $\lim_{n \to \infty} \|\Lambda^k T_n\|^2 \|\Lambda^2(\Lambda^{k+1} T_n)\|^{-1} = \infty$.

**Corollary 2.3** Let $\mathbb{T}$ be the smallest closed semi group in $\text{GL}(m, \mathbb{R})$ which contains the (essential) support of $(T_\sigma)_{\sigma \in \Sigma}$ (w.r.t. $p$) and assume that its algebraic closure $\overline{\mathbb{T}}$ is $k$-strongly irreducible and $k$-contracting for $k < m$. Then one has $\gamma_k > \gamma_{k+1}$.

### 2.4 The case of symplectic matrices

As the transfer matrices for Schrödinger or Dirac operators on strips are typically symplectic I would like to add some remarks about this sub-group. More details can be found in [9, A.IV.3]. For $m = 2$ one has $\text{Sp}(2, \mathbb{R}) = \text{SL}(2, \mathbb{R})$ but for bigger, even $m = 2L$, the group $\text{Sp}(2L, \mathbb{R})$ is a proper sub-group of $\text{SL}(2L, \mathbb{R})$, defined by

$$
\text{Sp}(2L, \mathbb{R}) = \{ M \in \text{Mat}(2L, \mathbb{R}) : M^*J M = J \},
$$

where $J$ denotes the so called symplectic form. It is given by the $2L \times 2L$ matrix

$$
J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}
$$

where all entries in this notation are $L \times L$ matrices. The group $\text{Sp}(2L, \mathbb{C})$ is defined by the same relation for complex matrices $M$. First note that this defines a group as for $M, N \in \text{Sp}(2L, \mathbb{R})$ one has

$$(MN)^*JMN = N^*(M^*JM)N = J , \quad (M^{-1})^*J M^{-1} = (M^{-1})^*JM^*M^{-1} = J .$$

Next if $M \in \text{Sp}(2L, \mathbb{R})$ then also $M^*$ is, because as $J^{-1} = -J$ one obtains

$$
MJM^* = [(M^{-1})^*(-J)M^{-1}]^{-1} = J .
$$

Therefore if $M \in \text{Sp}(2L, \mathbb{R})$ then also $M^*M \in \text{Sp}(2L, \mathbb{R})$. Now, the algebraic relation of this group implies some relation on the eigenvalues of $M^*M$ which are the squares of the singular values of $M$. As $J^*J = -J^2 = 1$ one obtains

$$
J^*M^*JM^*M = J^*J = 1 \implies J^*M^*M = (M^*M)^{-1} .
$$
Therefore the eigenvalues of $M^* M$ come in inverse pairs. Hence if $d_1 \geq d_2 \geq \ldots \geq d_{2L}$ are the singular values of $M$, then $d_{L+k} = d_{L+1-k}^{-1}$. Therefore if the random matrices $T_n$ as described in the sections above are supported in the group $\text{Sp}(2L, \mathbb{R})$ then one obtains for the Lyapunov exponents, that $\gamma_{L+k} = -\gamma_{L+1-k}$. Thus it is enough to consider the first $L$ Lyapunov exponents which are non-negative (as $\gamma_L \geq 0 \geq \gamma_{L+1} = -\gamma_L$).

To prove localization on strips one needs that the smallest non-negative Lyapunov exponent is positive, i.e. $\gamma_L > 0$ which is equivalent to $\gamma_L > \gamma_{L+1}$. In fact if $\gamma_L = 0$ for some ergodic operator and some energy interval then Kotani theory tells us, that there is a.c. spectrum [52]. The problem is, however, that one can not apply the theorems above directly as for $L \geq 1$ the whole group $\text{Sp}(2L, \mathbb{R})$ is not $k$-strongly irreducible for any $2L > k > 1$. Hence one has to restrict the action to some subspace of $\Lambda^k \mathbb{R}^{2L}$ as it is described in [9, A.IV.3]. For $k \leq L$ define the vector space $L_k$ by

$$L_k = \text{span} \{ \Lambda^k M(e_1 \wedge \ldots \wedge e_k) : M \in \text{Sp}(2L, \mathbb{R}) \} \subset \Lambda^k \mathbb{R}^{2L}.$$ 

It is clear by construction that the vector space $L_k$ is left invariant by the action of $\text{Sp}(2L, \mathbb{R})$. It is not so clear that this space is a proper sub space of $\Lambda^k \mathbb{R}^{2L}$, but this is shown in [9]. Let us denote the dimension of $L_k$ by $d$. Then any basis of $L_k$ induces a homomorphism $\varphi : \Lambda^k \text{Sp}(2L, \mathbb{R}) \to \text{GL}(d, \mathbb{R})$ and one can show that the first two Lyapunov exponents for the random products $\Lambda^k T_N$ and $\varphi(T_N)$ coincide. Hence one obtains the following.

**Definition 2.4** Let $1 \leq k \leq L$. A subset $\mathbb{M} \subset \text{Sp}(2L, \mathbb{R})$ is called $L_k$-strongly irreducible, if there is no finite union $\mathbb{W} = \mathbb{V}_1 \cup \ldots \mathbb{V}_l \subset \text{GL}(d, \mathbb{R})$ of proper subspaces of $\mathbb{R}^d$ such that $\varphi(\Lambda^k T)(\mathbb{W}) \subset \mathbb{W}$, $\forall T \in \mathbb{M}$.

A subset $\mathbb{M} \subset \text{Sp}(m, \mathbb{R})$ is called $L_k$-contracting, if there is a sequence $T_n \in \mathbb{M}$ of matrices such that $\lim_{n \to \infty} \|\varphi(\Lambda^k T_n)\|^2 \|\Lambda^2 \varphi(\Lambda^{k+1} T_n)\|^{-1} = \infty$.

**Proposition 2.2** Let $\mathbb{T}$ be the smallest closed semi group in $\text{Sp}(2L, \mathbb{R})$ which contains the (essential) support of $(T_\sigma)_{\sigma \in \Sigma}$ (w.r.t. $\mathbf{p}$) and assume that its algebraic closure $\overline{\mathbb{T}}$ is $L_k$-strongly irreducible and $L_k$-contracting for $k < m$. Then one has $\gamma_k > \gamma_{k+1}$.

Another important fact shown in [9] is that the group $\text{Sp}(2L, \mathbb{R})$ itself is $L_k$-contracting and $L_k$ irreducible for any $1 \leq k \leq L$. Therefore as a corollary one obtains the Goldsheid-Margulis criterion [32]:

**Corollary 2.4** (Goldsheid, Margulis) Let $\mathbb{T}$ be the smallest closed semi group in $\text{Sp}(2L, \mathbb{R})$ which contains the (essential) support of $(T_\sigma)_{\sigma \in \Sigma}$ (w.r.t. $\mathbf{p}$). If $\mathbb{T}$ is Zariski dense in $\text{Sp}(2L, \mathbb{R})$ then all $2L$ Lyapunov exponents are distinct. Especially, the smallest non-negative Lyapunov exponent is positive, i.e. $\gamma_L > 0$. 
Chapter 3

Transfer matrix and Scattering matrix

3.1 Introduction and model

In Section 1.4 we discussed the Anderson localization, \textit{i.e.} pure point spectrum with exponential decaying eigenfunctions, for the 1D Anderson model. Let us now consider a 'finite piece' of the Anderson model which means the restriction of the operator to \(l^2(1, \ldots, N)\) for some \(N\). This is supposed to represent a cable of finite length. When increasing \(N\), one should discover the exponential localization of the eigenstates. Therefore \(N\) has to be compared with an important quantity, the localization length \(l^E\) which is given by the inverse of the Lyapunov exponent, \textit{i.e.} \(l^E = \frac{1}{\gamma_E}\). If \(N\) is small compared to \(l^E\), electrons close to the energy \(E\) should be extended over the finite piece and be able to conduct from one end to the other. When \(N\) is big compared to \(l^E\), the probability for an electron to travel from one end to the other one should be exponentially small in \(N\) and scale like \(\exp(-\gamma_E N)\). Hence one expects the conductivity to decrease exponentially.

To measure the conductivity, one connects the finite piece with conducting (perfect) leads and measures the resistivity of the finite piece. This means one treats the finite piece as scatterer and is interested in the transmission and reflection of electron waves of the leads. This is analyzed in terms of scattering theory in this chapter.

Before going into details let me briefly summarize the contents of the following sections. Instead of operators on the line I will consider operators on the strip \(\ell^2(\mathbb{Z}) \otimes \mathbb{C}^L = \ell^2(\mathbb{Z}, \mathbb{C}^L)\). There one compares the Hamiltonian of a finite (randomly disordered) scatterer in between perfect conducting cables with the Hamiltonian of the perfect infinite cable and obtains some scattering operator. This operator describes the relation of incoming and outgoing waves to and from the scatterer. As it commutes with the free Hamiltonian of the perfect cable one can (almost surely) define scattering matrices on energy shells of the free Hamiltonian and one obtains transmission and reflection coefficients. There it turns out that the scattering matrix is related to a reduced transfer matrix for the finite scatterer and the transmission coefficients to the singular values of the reduced transfer matrix. In physics literature, \textit{e.g.} [5, 59], the transfer matrix is defined by the relations to the scattering matrix and
corresponds to what I call reduced transfer matrix here. This reduced transfer matrix does not only depend on the scatterer but also on the cable the scatterer is connected to and may have less dimension than the transfer matrix itself.

However, for so called ideal leads at an energy $E$ the ‘reduced transfer matrix’ is just a conjugation of the transfer matrix for the finite piece and one obtains $L$ transmission coefficients $t_1, \ldots, t_L$ whose sum gives the Landauer conductance at zero temperature [5, 54]. Their asymptotic behavior when increasing the length of the finite scatterer is described by the Lyapunov exponents for the transfer matrices. Now as they are symplectic $2L \times 2L$ matrices, the random products lead to $L$ non-negative Lyapunov exponents $\gamma_1 \geq \ldots \gamma_L \geq 0$ and the other $L$ ones are given by $-\gamma_l$ for $l = 1, \ldots, L$. The transmission coefficient $t_l$ decreases like $e^{-\gamma_l N}$. Analogue to the definition of localization length in the purely 1D Anderson model, one therefore calls the inverse $(\gamma_l)^{-1}$ a channel dependent localization length.

Now let us define the relevant operators on the space $\ell^2(\mathbb{Z}, \mathbb{C}^L) = \{ (\Psi_n)_{n \in \mathbb{Z}} : \Psi_n \in \mathbb{C}^L \land \sum_n \|\Psi_n\|^2 < \infty \}$. The following periodic Hamiltonian $H_0$ represents a ‘perfect’ cable without disorder.

$$(H_0 \Psi)_n = -\Psi_{n+1} - \Psi_{n-1} + \Delta \Psi_n, \quad \Psi = (\Psi_n)_{n \in \mathbb{Z}} \in \ell^2(\mathbb{Z}, \mathbb{C}^L),$$

where $\Delta$ is some self adjoint $L \times L$ matrix describing vertical modes within the cable. Note that the main difference to the Anderson model on $\ell^2(\mathbb{Z})$ is just that $\Psi_n$ is now a vector in $\mathbb{C}^L$ and not a number. Therefore the transfer matrix will be a $2L \times 2L$ matrix instead of a $2 \times 2$ matrix. The operator $H_0$ looks analogue to an operator on the line consisting of the discrete Laplacian and a constant potential. Similar to that case one can prove that this operator has pure a.c. spectrum.

Inserting a finite scatterer is now described by changing this operator on a finite piece. Therefore let $(V_n)_n$ be some sequence of hermitian matrices such that $V_n = \Delta$ for $n < 0$ and $n \geq N$ for some specific $N \in \mathbb{N}$ and define the operator $H$ by

$$(H \Psi)_n = -\Psi_{n+1} - \Psi_{n-1} + V_n \Psi_n.$$  

Solving $H_0 \Psi^0 = E \Psi^0$ formally one obtains

$$(\Psi^0_{n+1} \Psi^0_n) = T^{0,E} (\Psi^0_n \Psi^0_{n-1}), \quad T^{0,E} = \begin{pmatrix} \Delta - E1 & -1 \\ 1 & 0 \end{pmatrix},$$  

where $T^{0,E}$ is called the transfer matrix for the periodic operator $H_0$. If $\Psi$ is a formal solution of $H \Psi = E \Psi$ then one has similarly

$$(\Psi_{n+1} \Psi_n) = T^{E}_n (\Psi_n \Psi_{n-1}), \quad T^{E}_n = \begin{pmatrix} V_n - E1 & -1 \\ 1 & 0 \end{pmatrix}.$$  

Note that $T^{E}_n = T^{0,E}$ for $n < 0$ and $n \geq N$. We define the complete transfer matrix for the finite scatterer by

$$T^{E}_{0,N} = T^{E}_{N-1} \cdot T^{E}_{N-2} \cdots T^{E}_{0}.$$
It is easy to verify that all these transfer matrices are elements of the symplectic group $\text{Sp}(2L, \mathbb{C})$ which means $T^* J_L T = J_L$ where $J_L$ denotes the $2L \times 2L$ symplectic form

$$J_L = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \text{Mat}(2L, \mathbb{R}).$$

Different from the definition in Chapter 2 we indicate the size of the symplectic matrix $J_L$ here by the index $L$, because the size of the reduced transfer matrix is different and another, lower dimensional symplectic form will also be used.

### 3.2 Channels and scattering states

Let $\varphi_\alpha \in \mathbb{C}^L$, $\alpha = 1, \ldots, L$ be an orthonormal basis of eigenvectors of $\Delta$ and denote the corresponding eigenvalue by $\lambda_\alpha$, i.e. $\Delta \varphi_\alpha = \lambda_\alpha \varphi_\alpha$. I will use some Dirac notations, hence expressions like $|\lambda\rangle$ corresponding eigenvalue by $\lambda$. It is easy to verify that all these transfer matrices are elements of the symplectic group $\text{Sp}(2L)$ (not necessarily in $\ell^2$) and for $n \in \mathbb{Z}$ denote the vector $\Psi_n \in \mathbb{C}^L$ by $|n\rangle$. Let $|n, l\rangle$ for $n \in \mathbb{Z}, l \in \{1, \ldots, L\}$ denote the $\ell^2(\mathbb{Z}, \mathbb{C}^L)$ vector defined by $\langle m|n, l\rangle = \delta_{m,n} \delta_{l,1}$, where $e_l$ is the $l$-th canonical basis vector in $\mathbb{C}^L$. Then one can think of $|n\rangle$ as the $L$-tuple $(|n, 1\rangle, \ldots, |n, L\rangle)$. One can interpret $|n\rangle$ as $|n\rangle^*$ and the scalar product $\langle n|\Psi \rangle$ gives in fact a column vector. In these notations it also makes sense to write $\langle m|n\rangle = \delta_{m,n} 1_L$ meaning $\langle m, j|n, l\rangle = \delta_{m,n} \delta_{j,l}$, where $1_L$ is the $L \times L$ unit matrix.

Let us define the pseudo-eigenvectors $|\Psi_\alpha^0, k\rangle$ by $\langle n|\Psi_\alpha^0, k\rangle = \varphi_\alpha e^{ikn}$. They form a partition of unity in the sense that

$$\sum_\alpha \frac{1}{2\pi} \int_{-\pi}^\pi \langle m, j|\Psi_\alpha^0, k\rangle \langle \Psi_\alpha^0, k|n, l\rangle \, dk = \delta_{m,n} \delta_{j,l} = \langle m, j|n, l\rangle$$

which might also be written in the form

$$\sum_\alpha \frac{1}{2\pi} \int_{-\pi}^\pi \langle m|\Psi_\alpha^0, k\rangle \langle \Psi_\alpha^0, k|n\rangle \, dk = \delta_{m,n} 1_L = \langle m|1_L|n\rangle.$$

Therefore in a weak operator topology induced by the functionals $H \mapsto \langle m|H|m\rangle$ (as I am not testing with all $\ell^2$ vectors this topology is actually weaker than the usual weak operator topology) one can write

$$1 = \sum_\alpha \frac{1}{2\pi} \int_{-\pi}^\pi |\Psi_\alpha^0, k\rangle \langle \Psi_\alpha^0, k| \, dk.$$

Using $H_0|\Psi_\alpha^0, k\rangle = (-2\cos(k) + \lambda_\alpha)|\Psi_\alpha^0, k\rangle$ one can obtain the spectral decomposition of $H_0$. An eigenvector $\varphi_\alpha$ is said to be an elliptic channel for the energy $E$ iff $|E - \lambda_\alpha| < 2$. In that case there exists $k_\alpha \in (0, \pi)$ such that $E = -2\cos(k_\alpha) + \lambda_\alpha$. The terminology 'elliptic' comes from the fact, that this corresponds to eigenvalues $e^{\pm ik_\alpha}$ of the transfer matrix and is therefore related to a rotation. Now consider $k_\alpha$ also as a function $k_\alpha(E)$, where the
interval on which this function is defined depends on $\alpha$. To change the normalization of the pseudo-eigenvectors w.r.t. energy define for the elliptic channels

$$|\Psi^0_{\alpha}, E, \pm\rangle = (2 \sin(k_\alpha))^{-1/2} |\Psi^0_{\alpha}, \pm k_\alpha\rangle$$

then a change of variables shows

$$1 = \sum_{\alpha} \frac{1}{2\pi} \int_{-2+\lambda_\alpha}^{2+\lambda_\alpha} dE \left( |\Psi^0_{\alpha}, E, +\rangle \langle \Psi^0_{\alpha}, E, +| + |\Psi^0_{\alpha}, E, -\rangle \langle \Psi^0_{\alpha}, E, -| \right). \quad (3.4)$$

Furthermore we say that $\varphi_\alpha$ is an hyperbolic channel iff $|E - \lambda_\alpha| > 2$ and a parabolic channel iff $|E - \lambda_\alpha| = 2$. Now let $E$ be some energy in the a.c. spectrum of $H_0$ without any parabolic channel\(^1\).

Then there is at least one elliptic channel for $E$. Let us reorder the channels such that $\varphi_1, \ldots, \varphi_s$ are elliptic and $\varphi_{s+1}, \ldots, \varphi_L$ are hyperbolic channels. Furthermore for the hyperbolic channels $\alpha > s$ define $\gamma_\alpha > 0$ and $u_\alpha \in \{-1, 1\}$ such that

$$E = -2u_\alpha \cosh(\gamma_\alpha) + \lambda_\alpha, \quad (\alpha > s).$$

Then the $2L$ vectors

$$\begin{pmatrix} \varphi_\alpha \\  e^{\pm i k_\alpha} \varphi_\alpha \end{pmatrix}, \quad 1 \leq \alpha \leq s, \quad \text{and} \quad \begin{pmatrix} \varphi_\alpha \\ u_\alpha e^{\pm i \gamma_\alpha} \varphi_\alpha \end{pmatrix}, \quad s < \alpha \leq L$$

are eigenvectors of $T^0_E$ and form a basis $\mathbb{C}^{2L}$. Therefore any formal eigenvector $\Psi_0$ of $H_0$ satisfying $H_0 \Psi_0 = E \Psi_0$ is a linear combination of $|\Psi^0_{\alpha}, E, \pm\rangle, \alpha \leq s$, and $|\hat{\Psi}^0_{\alpha}, E, \pm\rangle$ defined by $\langle n| \hat{\Psi}^0_{\alpha}, E, \pm\rangle = \varphi_\alpha \frac{u_\alpha^{\pm 1+1}}{2 \sinh(\gamma_\alpha)} u_\alpha^n e^{\pm i \gamma_\alpha n}, \alpha > s$. The factor in front may seem strange, but it leads to nice relations for the coefficients of formal eigenvectors in the next section. For a formal eigenvector $|\Psi^0, E\rangle$ of $H^0$ define coefficients $c^+_\alpha, c^-_\alpha$ for $\alpha \leq s$ and $\hat{c}^+_\alpha, \hat{c}^-_\alpha$ for $\alpha > s$ such that

$$\langle n| \Psi^0, E\rangle = \sum_{\alpha \leq s, \sigma \in \{+,-\}} c^\sigma_\alpha |\Psi^0_{\alpha}, E, \sigma\rangle + \sum_{\alpha > s, \sigma \in \{+,-\}} \hat{c}^\sigma_\alpha |\hat{\Psi}^0_{\alpha}, E, \sigma\rangle$$

But only the elliptic channels appear as pseudo-eigenvectors in the spectral decomposition and give the multiplicity of $E$ for $H_0$.

Now let $|\Psi, E\rangle$ be some formal eigenvector of $H$ with eigenvalue $E$. Then for $n \leq 0$ and $n \geq N - 1$ it looks like a formal eigenvector of $H_0$. Therefore define the constants $a^+_\alpha, a^-_\alpha, b^+_\alpha, b^-_\alpha$ for $\alpha \leq s$ and $\hat{a}^+_\alpha, \hat{a}^-_\alpha, \hat{b}^+_\alpha, \hat{b}^-_\alpha$ for $\alpha > s$ associated to $|\Psi, E\rangle$ by

$$\langle n| \Psi, E\rangle = \sum_{\alpha \leq s, \sigma \in \{+,-\}} a^\sigma_\alpha \langle n| \Psi^0_{\alpha}, E, \sigma\rangle + \sum_{\alpha > s, \sigma \in \{+,-\}} \hat{a}^\sigma_\alpha \langle n| \hat{\Psi}^0_{\alpha}, E, \sigma\rangle, \quad n \leq 0 \quad (3.5)$$

\(^1\)There are only finitely many energies with parabolic channels. As the perfect cable has a.c. spectrum, these energies can be neglected for scattering theory.
\( \langle n|\Psi, E \rangle = \sum_{\alpha \leq s, \sigma \in \{+,-\}} b^\alpha e^{-\sigma k \alpha_n} \langle n|\Psi^0, E, \sigma \rangle + \sum_{\alpha > s, \sigma \in \{+,-\}} \hat{b}^\alpha e^{-\sigma n} \langle n|\hat{\Psi}^0, E, \sigma \rangle \) (3.6)

for \( n \geq N - 1 \).

\(|\Psi, E\rangle\) is an eigenvector of \( H \) iff there are only exponential decaying parts for the limits \( n \to \pm \infty \), which means that \( a^+ = a^- = b^+ = b^- = 0 \), \( \hat{a}^+ = \hat{a}^- = \hat{b}^+ = \hat{b}^- = 0 \) where \( a^+, \hat{a}^+, a^-, \hat{a}^-, b^+, \hat{b}^+, b^-, \hat{b}^- \) are correspondingly defined. \(|\Psi, E\rangle\) is called a scattering state or pseudo-eigenvector of \( H \) iff it is not an eigenvector and has no exponential growing parts, neither at \( +\infty \) nor at \( -\infty \) which means \( \hat{a}^- = \hat{b}^+ = 0 \). This means for the model considered here a pseudo-eigenvector \(|\hat{\Psi}, E\rangle\) includes at least one elliptic channel at least at one side. Therefore \( \langle n|\hat{\Psi}, E \rangle \) is not going to zero for \( n \to \infty \) or \( n \to -\infty \) but \( \langle n|\hat{\Psi}, E \rangle \) is bounded.

### 3.3 Normal forms of the transfer matrices

The coefficients for \( n \leq 0 \) and \( n \geq N - 1 \) are related by the transfer matrix as one has

\[
\mathcal{T}^E_{0,N} \left( \sum_{\alpha \leq s, \sigma \in \{+,-\}} \frac{a^\alpha}{(2\sin(k^\alpha))^1/2} \left( \begin{array}{c} \varphi_{\alpha} \\ \varphi_{\alpha} e^{-\sigma k^\alpha} \end{array} \right) + \sum_{\alpha > s, \sigma} \frac{\hat{a}^\alpha_{-\sigma}}{2\sqrt{\sinh(\gamma)}} \left( \begin{array}{c} \varphi_{\alpha} \\ \varphi_{\alpha} u^\alpha e^{-\sigma n} \end{array} \right) \right) = \mathcal{T}^E_{0,N} \left( \begin{array}{c} \langle 0|\Psi, E \rangle \\ \langle -1|\Psi, E \rangle \end{array} \right) = \left( \begin{array}{c} \langle N|\hat{\Psi}, E \rangle \\ \langle N - 1|\hat{\Psi}, E \rangle \end{array} \right) = \left[ \sum_{\alpha \leq s, \sigma \in \{+,-\}} \frac{b^\alpha}{(2\sin(k^\alpha))^1/2} \left( \begin{array}{c} \varphi_{\alpha} \\ \varphi_{\alpha} e^{-\sigma k^\alpha} \end{array} \right) + \sum_{\alpha > s, \sigma} \frac{\hat{b}^\alpha_{-\sigma}}{2\sqrt{\sinh(\gamma)}} \left( \varphi_{\alpha} u^\alpha e^{\sigma n} \varphi_{\alpha} \right) \right] \right.
\]

Working in the symplectic group one can diagonalize the hyperbolic channels. In order to do this define \( U = (\varphi_{\alpha})_{1 \leq \alpha \leq L} \), \( k = \text{diag}(k_1, \ldots, k_s) \), \( \gamma = \text{diag}(\gamma_{s+1}, \ldots, \gamma_L) \) and \( u = \text{diag}(u_{s+1}, \ldots, u_L) \). With these matrices define the symplectic 2\( L \times 2\) matrix

\[
M = \begin{pmatrix}
U & 0 \\
0 & U
\end{pmatrix}
\left( \begin{array}{cccc}
\frac{\sin(k)}{2} & 0 & 0 & 0 \\
0 & u(2\sinh(\gamma))^{-1} & 0 & (2\sinh(\gamma))^1/2 \\
\cos(k)(2\sinh(\gamma))^{-1} & 0 & (2\sinh(\gamma))^{1/2} & 0 \\
0 & e^{-\gamma}(2\sinh(\gamma))^{-1} & 0 & u e^{-\gamma}(2\sinh(\gamma))^{-1/2}
\end{array} \right)
\]

Then \( M \) transforms the free transfer matrix to its symplectic normal form

\[
M^{-1} \mathcal{T}^0 \mathcal{M} = \begin{pmatrix}
\cos(k) & 0 & -\sin(k) & 0 \\
0 & u e^{-\gamma} & 0 & 0 \\
\sin(k) & 0 & \cos(k) & 0 \\
0 & 0 & 0 & u e^{-\gamma}
\end{pmatrix}.
\] (3.8)
and one obtains
\[
M^{-1}T_{0,N}E M \frac{1}{\sqrt{2}} \begin{pmatrix}
\begin{array}{c}
a^+ + a^- \\
2\hat{a}^+ \\
(i(a^- - a^+))
\end{array}
\end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix}
\begin{array}{c}
b^+ + b^- \\
2\hat{b}^+ \\
i(b^- - b^+)
\end{array}
\end{pmatrix}.
\] (3.9)

Therefore denote \(T^E = M^{-1}T_{0,N}E \in \text{Sp}(2L, \mathbb{C})\). To diagonalize the elliptic channels define
\[
C_L = \frac{1}{\sqrt{2}} \begin{pmatrix}
1 & 1 \\
1 & -1
\end{pmatrix} \in \text{Mat}(2L, \mathbb{C})
\]

Then one obtains \(C_L \text{Sp}(2L, \mathbb{C})C_L^* = U(L,L)\) and the normal form of the free transfer matrix in the Lorentz group is
\[
C_L M^{-1}T_{0,E}MC_L^* = \begin{pmatrix}
e^{ik} & 0 & 0 & 0 \\
0 & u \cosh(\gamma) & 0 & \sinh(\gamma) \\
0 & 0 & e^{-ik} & 0 \\
0 & \sinh(\gamma) & 0 & u \cosh(\gamma)
\end{pmatrix}.
\]

Furthermore one obtains
\[
C_L T^E C_L = \begin{pmatrix}
a^+ \\
\hat{a}^+ + i\hat{a}^- \\
\hat{a}^- - i\hat{a}^+
\end{pmatrix} = \begin{pmatrix}
b^+ \\
\hat{b}^+ + i\hat{b}^- \\
\hat{b}^- - i\hat{b}^+
\end{pmatrix}
\]

### 3.4 Reduced transfer matrix

If possible we want to define a reduced transfer matrix relating the coefficients for the elliptic channels appearing in scattering states. This means we look for solutions of the equations above where \(\hat{a}^- = \hat{b}^+ = 0\). Given \(a^+\) and \(a^-\) the question is: Is there a unique \(\hat{a}^+\) such that we get \(\hat{b}^+ = 0\)? This is the case if the following \((L-s) \times (L-s)\) matrix
\[
\begin{pmatrix}
0_{(L-s)\times s} & 1_{(L-s)\times (L-s)} & 0_{(L-s)\times s} & 0_{(L-s)\times (L-s)}
\end{pmatrix} T^E \begin{pmatrix}
0_{s\times (L-s)} & 1_{(L-s)\times (L-s)} & 0_{s\times (L-s)} & 0_{(L-s)\times (L-s)}
\end{pmatrix}
\] (3.10)
is invertible. The indices indicate the size of the matrices. Within an energy interval without parabolic channel, the elliptic and hyperbolic channels for the different energies stay the same. Therefore one can consider the above \((L-s) \times (L-s)\) matrix and its determinant as a function of \(E\). As this function is analytic, this matrix is invertible for Lebesgue almost every energy \(E\). In this case any vectors \(a^+, a^-\) define a unique scattering
state characterized by the coefficients $a^+, a^-, b^+, b^-$ and $\hat{a}^+, \hat{b}^-$ as defined in (3.5) and (3.6). Then letting

$$2\hat{a}^+ = -\begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix} T^E \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}^{-1} \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix} T^E \begin{bmatrix} a^+ + a^- \\ 0 \\ 0 \\ \frac{1}{i}(a^- - a^+) \end{bmatrix}$$

one obtains

$$T^E \begin{bmatrix} a^+ + a^- \\ 2\hat{a}^+ \\ \frac{1}{i}(a^- - a^+) \\ 0 \end{bmatrix} = \begin{bmatrix} b^+ + b^- \\ 0 \\ \frac{1}{i}(b^- - b^+) \end{bmatrix}.$$ 

In this case we define the reduced $2s \times 2s$ transfer matrix $\hat{T}^E$ by

$$\hat{T}^E \begin{bmatrix} a^+ + a^- \\ \frac{1}{i}(a^- - a^+) \end{bmatrix} = \begin{bmatrix} b^+ + b^- \\ \frac{1}{i}(b^- - b^+) \end{bmatrix}. \quad (3.11)$$

Another way to write $\hat{T}^E$ would be

$$\hat{T}^E = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} T^E \left\{ 1 - \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \left[ \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right]^{-1} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\} \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}. $$

The sizes of the matrix entries for the matrices inside the brackets $\{\}$ in the equation above are similar as in (3.10). The size of the first matrix on the r.h.s. of the equation is $2s \times 2L$, the columns are divided into two blocks, each of size $s$, and the rows are divided in 4 blocks of sizes $s$, $L - s$, $s$ and $L - s$ in that order. The last matrix is the transpose of the first one.

Note, if $s = L$ then there is no hyperbolic channel and therefore all formal eigenvectors of $H$ are scattering states and $T^E$ already relates the elliptic channels. Therefore on this case one simply defines $\hat{T}^E = T^E$ and the ‘reduced’ transfer matrix is symplectic. This is actually always true, if the reduced transfer matrix exists.

**Proposition 3.1** The reduced transfer matrix is symplectic, i.e. $\hat{T}^E \in \text{Sp}(2s)$.

**Proof.** Let $x_i, y_i \in \mathbb{C}^s$, $i = 1, 2$ and define $\hat{x}_i, \hat{y}_i$ by

$$\hat{T}^E \begin{bmatrix} x_i \\ y_i \end{bmatrix} = \begin{bmatrix} \hat{x}_i \\ \hat{y}_i \end{bmatrix}, \quad \text{for } i = 1, 2 \Rightarrow \exists \hat{a}_i, \hat{b}_i \in \mathbb{C}^{L-s} : T \begin{bmatrix} x_i \\ y_i \end{bmatrix} = \begin{bmatrix} \hat{x}_i \\ \hat{y}_i \\ \hat{b}_i \end{bmatrix}. $$

Then one obtains

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}^* (\hat{T}^E)^* J_s \hat{T}^E \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} \hat{x}_1 \\ \hat{y}_1 \end{bmatrix}^* J_s \begin{bmatrix} \hat{x}_2 \\ \hat{y}_2 \end{bmatrix} = \hat{x}_1 \hat{y}_2 - \hat{y}_1 \hat{x}_2.$$
As this is true for arbitrary \( x, y \) one has \( (\hat{T}^E)^* J_s \hat{T}^E = J_s \) and hence \( \hat{T}^E \) is symplectic. \( \square \)

Thus if the matrix (3.10) is invertible then there is a reduced transfer matrix \( C_s \hat{T}^E C_s^* \in U(s, s) \) such that there is a scattering state with the elliptic channel coefficients \( a^\pm, b^\mp \) if and only if

\[
C_s \hat{T}^E C_s^* \begin{pmatrix} a^+ \\ a^- \end{pmatrix} = \begin{pmatrix} b^+ \\ b^- \end{pmatrix}.
\]  

(3.12)
Thus if \( \alpha > s \), then the solutions for \( \epsilon = 0 \) are \( u_a e^{\pm \gamma_\alpha} \) and \( \xi_\alpha^\epsilon \) converges to \( u_a e^{-\gamma_\alpha} \) for \( \epsilon \to 0 \). For the elliptic channels \( \alpha \leq s \) the solutions for \( \epsilon = 0 \) are \( e^{\pm ik_\alpha} \) being both on the unit circle. As \( |\xi_\alpha^\epsilon| < 1 \) the sign of its imaginary part is different to the sign of the imaginary part of \( \xi_\alpha^\epsilon + (\xi_\alpha^\epsilon)^{-1} = \lambda_\alpha - E - \iota \epsilon \). Therefore and as \( k_\alpha \in (0, \pi) \) one has \( \lim_{\epsilon \downarrow 0} \xi_\alpha^\epsilon = e^{ik_\alpha} \) and \( \lim_{\epsilon \uparrow 0} \xi_\alpha^\epsilon = e^{-ik_\alpha} \). Hence by the calculations above and (3.13), (3.14) we get

\[
\langle n | \Psi, E \rangle = \langle n | \Psi_\text{in}, E \rangle + \sum_{\alpha \leq s} \sum_{m=0}^{N} \frac{\varphi_\alpha e^{\iota km}}{2t \sin(k_\alpha)} (\varphi_\alpha^* V_m \langle m | \Psi, E \rangle)
\]

\[
+ \sum_{\alpha > s} \sum_{m=0}^{N} \frac{\varphi_\alpha u_a e^{-\gamma_\alpha m}}{e^{-\gamma_\alpha} - e^{\gamma_\alpha}} (\varphi_\alpha^* V_m \langle m | \Psi, E \rangle)
\]

\[
\langle n | \Psi, E \rangle = \langle n | \Psi_\text{out}, E \rangle + \sum_{\alpha \leq s} \sum_{m=0}^{N} \frac{\varphi_\alpha e^{-\iota km}}{2t \sin(k_\alpha)} (\varphi_\alpha^* V_m \langle m | \Psi, E \rangle)
\]

\[
+ \sum_{\alpha > s} \sum_{m=0}^{N} \frac{\varphi_\alpha u_a e^{-\gamma_\alpha m}}{e^{-\gamma_\alpha} - e^{\gamma_\alpha}} (\varphi_\alpha^* V_m \langle m | \Psi, E \rangle)
\]

Thus if \( |\Psi, E\rangle \) is the scattering state associated to the coefficients \( a^+, a^-, b^+, b^- \) as in (3.6) and (3.6) then

\[
|\Psi^0_\text{in}, E\rangle = \sum_{\alpha \leq s} \left[ a^+_\alpha |\Psi^0_\alpha, E, +\rangle + e^{ik_\alpha N} b^-_\alpha |\Psi^0_\alpha, E, -\rangle \right] \tag{3.15}
\]

and

\[
|\Psi^0_\text{out}, E\rangle = \sum_{\alpha \leq s} \left[ a^-_\alpha |\Psi^0_\alpha, E, -\rangle + e^{-ik_\alpha N} b^+_\alpha |\Psi^0_\alpha, E, +\rangle \right]. \tag{3.16}
\]

Therefore the scattering operator reduced to the energy shell can be described by the \( 2s \times 2s \) matrix \( S^E \) defined by

\[
S^E \begin{pmatrix} a^+ \\ b^- \end{pmatrix} = \begin{pmatrix} a^- \\ b^+ \end{pmatrix} \tag{3.17}
\]
We will see in Section 3.7 that such a unitary matrix $S^E$ exists and is uniquely determined by the reduced transfer matrix $\tilde{T}^E$.

Putting equations (3.15), (3.16), (3.17) and $S|\Psi_{in}^0\rangle = |\Psi_{out}^0\rangle$ together one can write the operator $S$ as an integral over the energy $E$. The number of elliptic channels is a step function $s(E)$. So far we took just one energy and set the elliptic channels to be the ones for $\alpha = 1, \ldots, s$. But when varying $E$ one should take into account that the channels which are elliptic are different ones for different energy intervals. Therefore let $\alpha(E, 1), \ldots, \alpha(E, s(E))$ denote the elliptic channels for $E$. Correspondingly for pseudo-eigenstates satisfying $S|\Psi_{in}^0, E\rangle = |\Psi_{out}^0, E\rangle$, define the coefficients $a^\pm_{\alpha(E, i)}$ and $b^\pm_{\alpha(E, i)}$. Then the scattering matrix $S^E$ satisfies (3.17) with $a^\pm = (a^\pm_{\alpha(E, 1)}, \ldots, a^\pm_{\alpha(E, s(E))})^t$ and the analogue definitions for $b^\pm$. Furthermore let $e_{E,i,+}$ be the $i$-th and $e_{E,i,-}$ be the $(s(E) + i)$-th canonical basis vector of $\mathbb{C}^{2s(E)}$ for $i = 1, \ldots, s(E)$. Then by (3.15), (3.16) and (3.17) the matrix element $e^*_{E,j,+}S^E e_{E,i,+}$ corresponds to the coefficient $a_{\alpha(E, j)}^-$ if one has $|\Psi_{in}^0, E\rangle = |\Psi_{out}^0_{\alpha(E, i), E}, +\rangle$. The meaning of the other matrix elements can also be read off these equations and one finally obtains the following.

**Proposition 3.2** The scattering operator $S$ is given by

$$S = \int dE \left[ \sum_{i,j=1,\ldots,s(E)} e^{-i\theta(E,i,j,\sigma,\sigma')} |\Psi_{\alpha(E,j)}^0, E, -\sigma'\rangle \left( e^*_{E,j,\sigma'} S^E e_{E,i,\sigma} \right) \langle \Psi_{\alpha(E,i)}^0, E, \sigma| \right],$$

where the correction phase $\theta(E, i, j, \sigma, \sigma')$ is

$$\theta(E, i, j, \sigma, \sigma') = \frac{\sigma_1 + 1}{2} k_{\alpha(E, i)} N + \frac{\sigma'_1 - 1}{2} k_{\alpha(E, j)} N.$$

Hence there are unitary operators $U, U': L^2(\mathbb{R}, \mathbb{C}^L) \to \int \oplus dE \mathbb{C}^{2s(E)}$ such that

$$S = U^* \left( \int \oplus dE S^E \right) U'.$$

The phase $\theta(E, i, j, \sigma, \sigma')$ comes from terms of the form $e^{ik_\alpha N}$ appearing as factors in (3.15) and (3.16).

### 3.6 Divided cable as reference operator

The situation may be also analyzed in another way. Instead of having the infinite cable $H_0$ as reference operator, one might argue that it should be the two pieces of semi-infinite cables. Then the finite block scatterer is putted in between. This leads to a different scattering operator defined by Møller operators but as we will see, it can be described using the same scattering matrices $S^E$ as above. Therefore they really include the relevant physics.
3.6. DIVIDED CABLE AS REFERENCE OPERATOR

Hence let us now define a new operator $\tilde{H}_0$ which describes a cable like $H_0$ cut into two pieces, one on the left and one on the right hand side of the finite block from 1 to $N$. This means

$$\langle n|\tilde{H}_0|\Psi\rangle = -\langle n+1|\Psi\rangle - \langle n-1|\Psi\rangle + \Delta\langle n|\Psi\rangle, \quad n \leq -2 \lor n \geq N + 2$$

$$\langle -1|\tilde{H}_0|\Psi\rangle = \langle -2|\Psi\rangle + \Delta\langle n|\Psi\rangle,$$

$$\langle N+1|\tilde{H}_0|\Psi\rangle = -\langle N+2|\Psi\rangle + \Delta\langle N+1|\Psi\rangle$$

$$\langle n|\tilde{H}_0|\Psi\rangle = 0, \quad n \in \{0, 1, \ldots, N\}.$$ 

Let $|n\rangle\langle n|$ denote the projection of $\Psi$ on $\Psi_n$, then $\sum_{n=0}^{N}|n\rangle\langle n|$ is the projection on the pure point part of $\tilde{H}_0$. Let $\tilde{P}_{ac}$ be the projection on the absolute continuous part, i.e.

$$\tilde{P}_{ac} = 1 - \sum_{n=0}^{N}|n\rangle\langle n|.$$ 

Now for large modulus of $n$, the scattering states of $\tilde{H}_0$ can be described by constants $a^\pm, b^\pm$ and $\hat{a}^+, \hat{b}^- \text{ like above. But this time there is no transfer and hence no transfer matrix relating } a \text{'s and } b \text{'s. The scattering states have the property } a^+ = -a^- \text{ and } b^+ = -b^- \text{ and one has the scattering matrix } S^E_{\tilde{H}_0} = -1 \text{ symbolizing that any wave in one part of the cable is reflected at the end. To get the partition of } \tilde{P}_{ac} \text{ with pseudo-eigenstates of } \tilde{H}_0 \text{ define }$$

$$|\tilde{\Psi}_0^0, k, -\rangle \text{ and } |\tilde{\Psi}_0^0, k, +\rangle \text{ by}$$

$$\langle n|\tilde{\Psi}_0^0, k, -\rangle = \begin{cases} \varphi_\alpha(e^{ikn} - e^{-ikn}), & n \leq 0 \\ 0, & n \geq 0 \end{cases}$$

$$\langle n|\tilde{\Psi}_0^0, k, +\rangle = \begin{cases} \varphi_\alpha(e^{-ik(n-N)} - e^{+ik(n-N)}), & n \geq N \\ 0, & n \leq N \end{cases}.$$ 

The $-$ indicates the left side and $+$ the right side of the finite scatterer. Then one has

$$\tilde{P}_{ac} = \sum_\alpha \int_0^{\pi} \frac{dk}{2\pi} \left( |\tilde{\Psi}_0^0, k, -\rangle \langle \tilde{\Psi}_0^0, k, - | + |\tilde{\Psi}_0^0, k, +\rangle \langle \tilde{\Psi}_0^0, k, + | \right)$$

and

$$\tilde{H}_0|\tilde{\Psi}_0^0, k, -\rangle = (-2\cos(k) + \lambda_\alpha)|\tilde{\Psi}_0^0, k, -\rangle,$$

the same holds for $|\tilde{\Psi}_0^0, k, +\rangle$. Therefore we define for an energy $E$ and the elliptic channels $\alpha$ the pseudo-eigenstates

$$|\tilde{\Psi}_0^0, E, -\rangle = \frac{1}{\sqrt{2\sin(k_\alpha)}}|\tilde{\Psi}_0^0, k, -\rangle, \quad |\tilde{\Psi}_0^0, E, +\rangle = \frac{1}{\sqrt{2\sin(k_\alpha)}}|\tilde{\Psi}_0^0, k, +\rangle.$$ 

One has for $n < 0$

$$\langle n|\tilde{\Psi}_0^0, E, -\rangle = \langle n|\tilde{\Psi}_0^0, E+\rangle - \langle n|\tilde{\Psi}_0^0, E, -\rangle \quad (3.18)$$
and for \( n > N \)
\[
\langle n|\tilde{\Psi}_\alpha^0, E, \uparrow \rangle = e^{ik_an} \langle n|\Psi_\alpha^0, E, - \rangle - e^{-ik_an} \langle n|\Psi_\alpha^0, E, + \rangle . \tag{3.19}
\]

From now on let \( E \) and \( \varphi_\alpha \) be chosen as above, meaning \( E \) has no parabolic channels and \( \varphi_1, \ldots, \varphi_s \) are the elliptic channels.

As \( \hat{H}_0 \) has eigenstates, the Møller operators are not defined on the whole Hilbert space, but on the absolute continuous one, this means \( \tilde{\Omega}^\pm = \lim_{t \rightarrow \pm \infty} e^{iHt} e^{-\hat{H}_0t} \hat{P}_{ac} \). Then the scattering operator \( \hat{S} = (\hat{\Omega}^-)^* \hat{\Omega}^+ \) is unitary on the absolute continuous Hilbert space of \( \hat{H}_0 \). Let \( |\tilde{\Psi}_\text{in}, E\rangle \) be some pseudo-eigenvector of \( \hat{H}_0 \) and let \( |\tilde{\Psi}_\text{out}, E\rangle = \hat{S}|\tilde{\Psi}_\alpha^0, E\rangle \) as well as \( |\Psi, E\rangle = \hat{\Omega}^+|\tilde{\Psi}_\text{in}, E\rangle = \hat{\Omega}^-|\tilde{\Psi}_\text{out}, E\rangle \). Then for \( n < 0 \) one has
\[
\langle n|(E - \hat{H}_0 - i\epsilon)^{-1}|\Psi, E\rangle = \sum_\alpha \int_0^\pi \frac{dk}{2\pi} \frac{\varphi_\alpha(e^{ikn} - e^{-ikn})}{E + i\epsilon + e^{ik} + e^{-ik} - \lambda_\alpha} \langle \tilde{\Psi}_\alpha^0, k, \uparrow |(E - \hat{H}_0 - i\epsilon)^{-1}(H - \hat{H}_0)|\Psi, E \rangle
\]
\[
= \sum_\alpha \int_0^\pi \frac{dk}{2\pi} \frac{\varphi_\alpha(e^{ikn} - e^{-ikn})}{E + i\epsilon + e^{ik} + e^{-ik} - \lambda_\alpha} \langle \tilde{\Psi}_\alpha^0, k, \uparrow |H - \hat{H}_0|\Psi, E \rangle
\]
\[
= \sum_\alpha \int_0^\pi \frac{dk}{2\pi} \frac{\varphi_\alpha(e^{ikn} - e^{-ikn})}{E + i\epsilon + e^{ik} + e^{-ik} - \lambda_\alpha} \varphi_\alpha^*(0)|\Psi, E\rangle
\]
\[
= \sum_\alpha \int_0^\pi \frac{dk}{2\pi} \frac{\varphi_\alpha(e^{ikn} - e^{-ikn})}{E + i\epsilon + e^{ik} + e^{-ik} - \lambda_\alpha} \varphi_\alpha^*(0)|\Psi, E\rangle
\]

with \( \xi_\alpha^0 \) defined as above. For \( n > N \) one has to replace \( \tilde{\Psi}_\alpha^0, k, \uparrow \) by \( \tilde{\Psi}_\alpha^0, k, \downarrow \) to obtain
\[
\langle n|(E - \hat{H}_0 - i\epsilon)^{-1}|\Psi, E\rangle = \sum_\alpha \varphi_\alpha (\xi_\alpha^0)^{-n}\varphi_\alpha^*(N)|\Psi, E\rangle .
\]

Using the Lippmann Schwinger equation one therefore obtains for \( n < 0 \)
\[
\langle n|\Psi, E\rangle - \sum_{\alpha > s} \varphi_\alpha e^{\gamma_\alpha n} \varphi_\alpha^*(0)|\Psi, E\rangle = \langle n|\tilde{\Psi}_\text{in}, E\rangle + \sum_{\alpha \leq s} \varphi_\alpha e^{-ik_\alpha n} \varphi_\alpha^*(0)|\Psi, E\rangle \tag{3.20}
\]
and for \( n > N \)
\[
\langle n|\Psi, E\rangle - \sum_{\alpha > s} \varphi_\alpha e^{-\gamma_\alpha n} \varphi_\alpha^*(N)|\Psi, E\rangle = \langle n|\tilde{\Psi}_\text{out}, E\rangle + \sum_{\alpha \leq s} \varphi_\alpha e^{ik_\alpha (n-N)} \varphi_\alpha^*(N)|\Psi, E\rangle \tag{3.21}
\]
Now let $|\Psi, E\rangle$ be the scattering state with coefficients $a^+, a^-, b^+, b^-$ as in (3.5) and (3.6). Then equation (3.17) $S^E(a^+_{b^-}) = (a^-_{b^+})$ holds. Combining (3.5) and (3.20) one obtains

$$
\sum_{\alpha \leq s} (a^+_\alpha e^{ik_n \varphi_\alpha} + a^-_\alpha e^{-ik_n \varphi_\alpha}) = \langle n|\tilde{0}, E\rangle + \sum_{\alpha \leq s} \varphi_\alpha e^{-ik_n \varphi^*_\alpha} \langle 0|\Psi, E\rangle.
$$

For $n < 0$ one can write $\langle n|\tilde{0}, E\rangle$ in the form

$$
\langle n|\tilde{0}, E\rangle = \sum_{\alpha \leq s} c_\alpha |\tilde{0}, E, \alpha\rangle
$$

for certain coefficients $c_\alpha$. Comparison with the formula above and using (3.18) shows $c_\alpha = a^+_\alpha$. Using similar arguments and the equations (3.5), (3.6), (3.18), (3.19), (3.20) and (3.21) one finally obtains

$$
|\tilde{0}, E\rangle = \sum_{\alpha \leq s} \left[ a^+_{\alpha} |\tilde{0}, E, \alpha\rangle + b^-_{\alpha} |\tilde{0}, E, \alpha\rangle \right]
$$

(3.22)

$$
|\tilde{0}, E\rangle = -\sum_{\alpha \leq s} \left[ a^-_{\alpha} |\tilde{0}, E, \alpha\rangle + b^+_{\alpha} |\tilde{0}, E, \alpha\rangle \right].
$$

(3.23)

Similar as in the last section define $e_{E,i,-}$ to be the $i$-th and $e_{E,i,+}$ to be the $(s(E) + i)$-th canonical basis vector of $\mathbb{C}^{2s(E)}$. The equations (3.17), (3.22), (3.23) and $\tilde{S}|\tilde{0}, E\rangle = |\tilde{0}, E\rangle$ now imply the following.

**Proposition 3.3** The scattering operator $\tilde{S}$ is given by

$$
\tilde{S} = \int dE \left[ \sum_{i,j=1, \ldots, s(E)} \sum_{\sigma \in \{-, +\}} -|\tilde{0}_{E,j}, E, \sigma\rangle \langle \hat{\Psi}_{E,i}, E, \sigma'| \right] S^{E_{E,i}} e_{E,i,\sigma} |\tilde{0}_{E,j}, E, \sigma'|.
$$

There are unitary operators $V, V'$ from the a.c. Hilbert space $\tilde{P}_{ac}(\ell^2(\mathbb{Z}, \mathbb{C}^L))$ to $\int dE \mathbb{C}^{2s(E)}$ such that

$$
\tilde{S} = V^* \left( \int dE S^{E} \right) V'.
$$

Therefore the scattering matrices $S^{E}$ also describes the scattering operator $\tilde{S}$.

### 3.7 Relation of scattering matrix and transfer matrix

If the reduced transfer matrix as defined in (3.11) exists, then (3.12) and (3.17) relates the reduced transfer matrix with the scattering matrix by

$$
C_s \hat{T}^E C^*_s \begin{pmatrix} a^+ \cr a^- \end{pmatrix} = \begin{pmatrix} b^+ \cr b^- \end{pmatrix} \iff S^E \begin{pmatrix} a^+ \cr b^- \end{pmatrix} = \begin{pmatrix} a^- \cr b^+ \end{pmatrix}
$$

The following theorem shows the existence of $S^E$ with this property in a constructive way. It is based on a polar decomposition and it can already be found in [57, 59].
Theorem 3.1 For any matrix $\tilde{T} \in U(s,s)$ there is a unique unitary matrix $S \in U(2s)$ with the property that for any $a^+, a^-, b^+, b^- \in \mathbb{C}^s$ one has

$$\tilde{T}\begin{pmatrix} a^+ \\ a^- \end{pmatrix} = \begin{pmatrix} b^+ \\ b^- \end{pmatrix} \iff S\begin{pmatrix} a^+ \\ a^- \end{pmatrix} = \begin{pmatrix} a^- \\ b^+ \end{pmatrix} \quad (3.24)$$

Remark. The inverse is not true. One cannot find a matrix $\tilde{T}$ for all unitary matrices $S \in U(2s)$ such that the relation above is fullfilled. Using the terminology explained below, such a $\tilde{T}$ only exists if the transmission coefficients of $S$ are all non zero and in that case $\tilde{T}$ is unique. The transfer matrices have something to do with the transfer of waves from the left to the right. But if there is one transmission coefficient equal to zero then this means that there is one planar wave which is totally reflected by the scatterer. Hence a transfer does not occur for this wave and the transfer matrix is not defined.

Proof. Let $\tilde{T} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$. We first prove the existence and then uniqueness of $S$. As $\tilde{T} \in U(s,s)$ one has

$$A^*A = 1 + C^*C, \quad AA^* = 1 + BB^*$$

$$D^*D = 1 + B^*B, \quad DD^* = 1 + CC^*$$

$$A^*B = C^*D, \quad AC^* = BD^* \quad (3.25)$$

As $A^*A \geq 1$ there exists a unitary matrix $U$ and a real diagonal matrix $Q \geq 1$ such that $A^*A = U^*QU$. Define $V$ by

$$V = AU^*\sqrt{Q^{-1}} \iff A = V\sqrt{Q}U \quad (3.26)$$

Then $V^*V = \sqrt{Q^{-1}}UA^*AU^*\sqrt{Q^{-1}} = 1$ and hence $V$ is unitary. Furthermore one has $C^*C = A^*A - 1 = U^*(Q - 1)U$. Hence there exists $\tilde{V} \in U(s)$ such that

$$\tilde{V}\sqrt{Q - 1} = CU^* \iff C = \tilde{V}\sqrt{Q - 1}U \quad (3.27)$$

$\tilde{V}$ is uniquely determined if $Q - 1$ is invertible, otherwise it is not. Now define $\tilde{U}$ by

$$\tilde{U} = \sqrt{Q^{-1}}\tilde{V}^*D \iff D = \tilde{V}\sqrt{Q}\tilde{U} \quad (3.28)$$

Then one has using (3.25) and (3.27)

$$\tilde{U}\tilde{U}^* = \sqrt{Q^{-1}}\tilde{V}^*DD^*\sqrt{Q^{-1}} = \sqrt{Q^{-1}}\tilde{V}^*[1 + \tilde{V}(Q - 1)\tilde{V}^*]\tilde{V}\sqrt{Q^{-1}} = 1$$

and hence $\tilde{U}$ is also unitary. Furthermore one obtains using (3.25), (3.26), (3.27) and (3.28)

$$B = (A^*)^{-1}C^*D = (U^*\sqrt{Q}V^*)^{-1}U^*\sqrt{Q - 1}\tilde{V}^*\tilde{V}\sqrt{Q}\tilde{U}$$

$$= V\sqrt{Q^{-1}}\sqrt{Q - 1}\sqrt{Q}\tilde{U} = V\sqrt{Q - 1}\tilde{U} \quad (3.29)$$

Now using (3.26), (3.27), (3.28) and (3.29) one obtains

$$\tilde{T} = \begin{pmatrix} V & 0 \\ 0 & \tilde{V} \end{pmatrix} \begin{pmatrix} \sqrt{Q} & \sqrt{Q - 1} \\ \sqrt{Q - 1} & \sqrt{Q} \end{pmatrix} \begin{pmatrix} U & 0 \\ 0 & \tilde{U} \end{pmatrix} \quad (3.30)$$
As \( Q \geq 1 \), one has \( 1 - Q^{-1} \geq 0 \) and hence \( \sqrt{1 - Q^{-1}} \) is a well-defined, non-negative diagonal matrix. Thus we can define the unitary matrix

\[
S = \begin{pmatrix}
\hat{U}^* & 0 \\
0 & V
\end{pmatrix}
\begin{pmatrix}
-\sqrt{1 - Q^{-1}} & \sqrt{Q^{-1}} \\
\sqrt{Q^{-1}} & \sqrt{1 - Q^{-1}}
\end{pmatrix}
\begin{pmatrix}
U & 0 \\
0 & \hat{V}^*
\end{pmatrix}
\] (3.31)

It is now easy to check, that \( S \) fullfills (3.24).

To prove uniqueness of \( S \) assume \( \hat{S} \) also fullfills (3.24), then for any vectors \( a^+, a^-, b^+, b^- \in \mathbb{C}^s \) one has

\[
S(a^+) = (b^-) \Leftrightarrow T(a^+) = (b^+) \Leftrightarrow \hat{S}(a^+) = (b^-)
\]

which means \( S = \hat{S} \). \( \square \)

Using this theorem for \( \hat{T} = C_s \hat{T} E C_s^* \) one obtains the scattering matrix \( S^E \). Now let us write the scattering matrix in the block structure \( S^E = \begin{pmatrix} R & T' \\ T & R' \end{pmatrix} \). The matrix \( R \) determines the overlap of the incoming waves from the left (\( a^- \)) with the outgoing waves to the left (\( a^+ \)) and represents the reflection to the left cable. Similar \( R' \) represents the reflection to the right cable, \( T \) represents the transmission from left to right and \( T' \) from right to left. As \( S^E \) is unitary, one has \( (S^E)^* S^E = 1 = S^E(S^E)^* \) implying

\[
1 = R^* R + T^* T' = R^* R' + T'^* T = R R^* + T'^* T' = R R^* + T T^*.
\]

These equations reflect the fact that there is no loss of energy. Everything is either reflected or transmitted. Now for any quadratic matrix \( A \) the eigenvalues of \( A^* A \) and \( A A^* \) coincide\(^2\). Therefore the eigenvalues of \( T T^*, T' T'^* \) and \( 1 - R R^* \) and \( 1 - R' R'^* \) coincide and are called transmission coefficients. Let us denote them by \( 0 \leq t_1 \leq \ldots \leq t_s \). Equations (3.30) and (3.31) now hold when replacing \( \hat{T} \) and \( S \) by \( C_s \hat{T} E C_s^* \) and \( S^E \) and comparing them induces the following algebraic relation between transfer and transmission matrices:

\[
\left( 2 + (C_s \hat{T} E C_s^*)^* (C_s \hat{T} E C_s^*)^* + \left( (C_s \hat{T} E C_s^*)^* (C_s \hat{T} E C_s^*)^* \right)^{-1} \right)^{-1} = \frac{1}{4} \begin{pmatrix}
T T^* & 0 \\
0 & T'^* T'^*
\end{pmatrix} \quad (3.32)
\]

For any matrix \( \tilde{T} \) in the group \( U(s, s) \) one has

\[
\tilde{T}^* \begin{pmatrix} 1 & 0 \\
0 & -1
\end{pmatrix} \tilde{T} = \begin{pmatrix} 1 & 0 \\
0 & -1
\end{pmatrix} \Leftrightarrow \tilde{T} = \begin{pmatrix} 1 & 0 \\
0 & 1
\end{pmatrix} \tilde{T}^* \begin{pmatrix} 1 & 0 \\
0 & 1
\end{pmatrix}
\]

and therefore \( \tilde{T} \) and \( \tilde{T}^{-1} \) have the same singular values. In particular the singular values come in inverse pairs. Therefore the \( 2s \) eigenvalues of \( (C_s \hat{T} E C_s^*)^* (C_s \hat{T} E C_s^*)^* \) may be denoted

\(^2\)Write \( A = U D V \) with unitaries \( U, V \) and \( D \) real diagonal, then \( A A^* = U D^2 U^* \) and \( A^* A = V^* D^2 V \).
by $\exp(\pm x_l)$ for $x_1 \geq \ldots \geq x_s \geq 0$. Note as $C_L$ is unitary, they coincide with the eigenvalues of $\hat{T}^E(\hat{T}^E)^*$ and them of $(\hat{T}^E)^*\hat{T}^E$. By (3.32) one obtains

$$t_l = \frac{1}{2 + e^{x_l} + e^{-x_l}} = \frac{1}{(\cosh(x_l/2))^2}.$$  \hfill (3.33)

If $H_0$ is an ideal lead (explained in the next section) the problem of reducing the transfer matrix does not occur (i.e. $s = L$ in that case) and one has $L$ transmission coefficients. The conductance $G$ of the finite piece at zero temperature is then given by the sum of the transmission coefficients [5], $G = G_0 \sum_{l=1}^L t_l$, where $G_0 = \frac{2e^2}{h}$ is a natural unit. This is known as Landauer formula or also Landauer conductance because of his pioneering paper [54].

### 3.8 Ideal leads

The set of elliptic channels and therefore the reduced transfer matrix and especially its dimension depends very much on the transverse modes of the cable (in the model considered here they are described by the matrix $\Delta$). The operator $\Delta$ determines also the channel vectors $\varphi_\alpha$ and hence the conjugation of the transfer matrix that is used. For all that reasons the transmission coefficients are not only a property of the scatterer but very much also of the used cable. In order to get something which is more like a physical property of the scatterer (like the whole transfer matrix itself) one should connect it to some cable where one has only elliptic channels and any vector $\varphi \in \mathbb{C}^L$ is a channel. This means $\Delta = c \mathbf{1}$ where the energy shift $c$ lies in a region such that there are only elliptic channels at the given energy $E$. Such a cable does not have transverse modes and is called an ideal lead as any vector is a channel. The problem of reducing the transfer matrix does not occur for any $H_0$ having only elliptic channels at energy $E$. In the sequel we allow any such $H_0$ as an ideal lead at energy $E$. In this case the transmission coefficients $t_l$ depend only on the singular values of $\hat{T}^E = T^E$. Taking $N \rightarrow \infty$, the behavior is described by the Lyapunov exponents $\gamma_l$ for the sequence of products $T^E_{0,N}$, provided they exist. Then the asymptotic behavior of the singular values $\exp(\pm x_n)$ of $T^E$, which now depend on $N$, is given by $\exp(\pm x_l) \sim \exp(\pm \gamma_l N)$. (Note that increasing $N$ actually means that the operator $H$ might be changed, because the condition $V_n = \Delta$ is assumed to hold for $n \geq N$.)

By (3.33) this implies $t_l(N) \sim [\cosh(\gamma_l N/2)]^{-2} \sim \exp(-\gamma_l N)$. From this one immediately realizes, that the transmission coefficients converge to zero, if the first (non-negative) $L$ Lyapunov exponents are positive. In the situation, when a transmission coefficient converges to zero the conduction through this 'channel' vanishes in the limit, one may speak about a closing channel. However, this is not related to one of the channels of $H_0$ as introduced above, since all channels of $H_0$ are mixed by the scattering matrix. The terminology of 'closing channels' corresponds to conduction 'channels' through the block of size $N$ in $H$ which are given by the scattering states.

In that terminology if all channels close, then there is eventually (for large $N$) no conduction and hence one expects localization. The inverses of the Lyapunov exponents $(\gamma_l)^{-1}$ are sometimes also called channel dependent localization lengths e.g. [5].
Note that the transmission coefficients are always bounded between zero and one. If the Lyapunov exponent $\gamma_l$ is equal to zero, it only implies that $t_l$ does not converge to zero exponentially fast, but it can still converge to zero. However if that is not the case and if furthermore one has a limit above from zero, one says that the channel remains open for conduction to a certain fraction, given by the limit of $t_l(N)$ for $N \to \infty$.

Now let me give another point of view on the ‘reduced transfer matrix’ and the related scattering matrix of lower dimension in case of a ‘real’ lead (with elliptic and hyperbolic channels). We fix the size of the finite scatterer we putted in between the real leads. Then we connect a finite piece of the real lead to the left and the right and consider this block as scatterer for an ideal lead. Finally we will take the limit of the length of the finite real leads going to infinity. Hence the scatterer for the ideal lead consists of 3 blocks, the real lead, the block and again the real lead. As an ideal lead we will take one, which has the same channel vectors $\varphi_\alpha$ as the real lead.

The aim of this section is to prove that this limit process gives the same transmission coefficients and in some sense the same scattering matrix. This shows once more that the reduced transfer matrix and the corresponding scattering matrix already describe the finite scatterer together with the (non-ideal) infinite cables on the right and the left. In fact, the hyperbolic channels are related to closing channels, whereas the elliptic channels remain open. The limit gives the same transmission coefficient as one would obtain for the scattering matrix related to the reduced transfer matrix related to the real lead.

Therefore let us construct an ideal lead which has the same channels $\varphi_\alpha$ with the same eigenvalues $\lambda_\alpha$ for $\alpha = 1, \ldots, s$ and with eigenvalues $E$ instead of $\lambda_\alpha$ for $\alpha > s$. This means that the inverse wavelengths for the additional elliptic channels satisfy $k_\alpha = \frac{\pi}{2} (\alpha > s)$. The hermitian matrix for one slice is given by $\Delta I = U \text{diag}(\lambda_1, \ldots, \lambda_s, E, \ldots, E) U^*$ where $U = (\varphi_1, \ldots, \varphi_l)$ as above. The Hamiltonian for this at $E$ ideal lead would be

$$ (H_I \psi)_n = -\psi_{n+1} - \psi_{n-1} + \Delta_I \psi_n. $$

Now placing a finite block scatterer somewhere one can introduce the vectors $a_I^{\pm}$ and $b_I^{\pm}$ in $\mathbb{C}^L$ describing a formal solution of the eigenvalue equation like above, but this time there are only elliptic channels.

Now let us consider the block described by $H$ surrounded by cables of length $m$ described by $H_0$ as scatterer w.r.t. the ideal lead. The transfer matrix of this three blocks as a function of $m$ is given by $T(m) = (T^{0,E})^m T_{0,N}^{E}(T^{0,E})^m$. $T^{0,E}$ is the transfer matrix for the operator $H_0$ as defined in (3.1). To get the relation between $a_I^{\pm}$ and $b_I^{\pm}$ let us introduce the matrix $M_I$ similar to $M$ in (3.7) by

$$ M_I = \begin{pmatrix} U & 0 \\ 0 & U \end{pmatrix} \begin{pmatrix} (\sin(k))^{-1/2} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \cos(k)(\sin(k))^{-1/2} & 0 & (\sin(k))^{1/2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. $$

Then $M_I^{-1} T(m) M_I \begin{pmatrix} a_I^+ + \hat{a}_I^- \\ i(a_I^+ - a_I^-) \end{pmatrix} = \begin{pmatrix} b_I^+ + \hat{b}_I^- \\ i(b_I^+ - \hat{b}_I^-) \end{pmatrix}$. Let $a^\pm, \hat{a}^\pm, b^\pm, \hat{b}^\pm$ satisfy the relations as in (3.9),
3. Transfer Matrix and Scattering Matrix

i.e.

\[
M^{-1} T_{0,N}^E M \begin{pmatrix} a^+ + a^- \\ 2\hat{a}^+ \\ 2\hat{a}^- \end{pmatrix} = \begin{pmatrix} b^+ + b^- \\ 2\hat{b}^+ \\ 2\hat{b}^- \end{pmatrix}.
\]

Then using (3.8) one obtains

\[
M^{-1} T(m)M \begin{pmatrix} e^{-imk}a^+ + e^{imk}a^- \\ 2u^m e^{-im\gamma} \hat{a}^+ \\ 2u^m e^{-im\gamma} \hat{a}^- \end{pmatrix} = \begin{pmatrix} e^{imk}b^+ + e^{-imk}b^- \\ 2u^m e^{im\gamma} \hat{b}^+ \\ 2u^m e^{im\gamma} \hat{b}^- \end{pmatrix}.
\]

Furthermore by (3.7) and the definition of \( M_I \) one has

\[
M_I^{-1} M = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & u(2\sinh\gamma)^{-\frac{1}{2}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & u\gamma(2\sinh\gamma)^{\frac{1}{2}} \end{pmatrix}.
\]

To simplify notations define \( \hat{a}^+ = u(2\sinh\gamma)^{-\frac{1}{2}} \hat{a}^+ \) and \( \hat{a}^- = u(2\sinh\gamma)^{\frac{1}{2}} \hat{a}^- \), the vectors \( \hat{b}^\pm \) are analogue defined. Then the equation above and (3.34) give

\[
M_I^{-1} T(m) M_I \begin{pmatrix} e^{-imk}a^+ + e^{imk}a^- \\ 2u^m e^{-im\gamma} \hat{a}^+ \\ 2u^m e^{-im\gamma} \hat{a}^- \end{pmatrix} = \begin{pmatrix} e^{imk}b^+ + e^{-imk}b^- \\ 2u^m e^{im\gamma} \hat{b}^+ \\ 2u^m e^{im\gamma} \hat{b}^- \end{pmatrix}.
\]

Therefore the related scattering matrix of the three blocks w.r.t. the ideal lead satisfies

\[
S_I(m) \begin{pmatrix} e^{-imk}a^+ \\ u^m(e^{-im\gamma} \hat{a}^+ + ie^{im\gamma} \hat{a}^-) \\ e^{-imk}b^- \\ u^m(e^{-im\gamma} \hat{b}^+ + ie^{im\gamma} \hat{b}^-) \end{pmatrix} = \begin{pmatrix} e^{imk}a^- \\ u^m(e^{-im\gamma} \hat{a}^- - ie^{im\gamma} \hat{a}^+) \\ e^{imk}b^+ \\ u^m(e^{-im\gamma} \hat{b}^+ - ie^{im\gamma} \hat{b}^-) \end{pmatrix}.
\]

(3.35)

Now introduce a phase normalization and consider the matrices

\[
\hat{S}_I(m) = \begin{pmatrix} e^{-imk} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & e^{-imk} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} S_I(m) \begin{pmatrix} e^{-imk} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & e^{-imk} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.
\]

As the unitary group is compact, there is a limit point of this sequence, lets call it \( \hat{S}_I \).

Dividing (3.35) by \( u^m e^{im} \) and taking the limit \( m \to \infty \) one obtains

\[
\hat{S}_I \begin{pmatrix} 0 \\ -i\hat{a}^+ \\ 0 \\ \hat{b}^+ \end{pmatrix} = \begin{pmatrix} 0 \\ -i\hat{a}^+ \\ 0 \\ \hat{b}^+ \end{pmatrix}.
\]

(3.36)
3.8. IDEAL LEADS

As the reduced transfer matrix exists, one can always choose $\hat{a}^+$ such that $\hat{b}^+ = 0$ which means $\tilde{b}^- = 0$. Therefore and by linearity (3.36) holds for arbitrary $\hat{a}^-, \hat{b}^+ \in \mathbb{C}^{L-s}$. Furthermore one can choose $\hat{a}^- = 0$ and tune $\hat{a}^+$ such that $\hat{b}^+ = 0$. Then $\tilde{a}^- = 0 = \tilde{b}^+$, $a^-, b^+$ are related by the scattering matrix $S^E$ and the limit $m \to \infty$ of (3.35) yields

$$\begin{pmatrix} a^+ \\ 0 \\ b^- \\ 0 \end{pmatrix} = \begin{pmatrix} a^- \\ 0 \\ b^+ \\ 0 \end{pmatrix} \iff S^E \begin{pmatrix} a^+ \\ b^- \end{pmatrix} = \begin{pmatrix} a^- \\ b^+ \end{pmatrix}. \quad (3.37)$$

These two equations (3.36) and (3.37) determine any limit point of $\hat{S}_I(m)$ uniquely, hence we have proved the following.

**Theorem 3.2** The limit $\hat{S}_I = \lim_{m \to \infty} \hat{S}_I(m)$ exists. Furthermore there is a relation between $\hat{S}_I$ and $S^E$ given by

$$S^E = \begin{pmatrix} R & 0 \\ T & T' \end{pmatrix} \quad \Rightarrow \quad \hat{S}_I = \begin{pmatrix} R & 0 & 0 \\ 0 & -1 & 0 \\ T' & 0 & R' \end{pmatrix}.$$ 

This implies that the non-zero transmission coefficients of $\hat{S}_I$ coincide with them of $S^E$.

So we have seen that the scattering matrix and especially the transmission coefficients of a finite scatterer w.r.t. a periodic background operator $H_0$ can be obtained by a limit process of the scattering matrix w.r.t. an ideal lead. The elliptic channels of $H_0$ survive in this limit giving positive transmission coefficients and the hyperbolic channels are closing.

The asymptotic behavior of the transmission coefficients is described by the Lyapunov exponents and therefore the scaling behavior if the Lyapunov exponent is of physical interest.

For large blocks the smallest non-negative Lyapunov exponent describes the biggest transmission coefficient. Positive Lyapunov exponents always correspond to closing channels and the Lyapunov exponents give the exponential decay rate. Therefore for random models $H_0 + \lambda V$ on strips with random matrix potential $V$ one expects pure point spectrum if all Lyapunov exponents are non-zero. The eigenfunctions should decay with the rate of the smallest positive Lyapunov exponent. The proof for this statement is very similar as the one for the 1D Anderson model done in the first chapter. In the terminology of this chapter, the randomness is closing all channels leading to localization.

If one has just one open channel in the limit, i.e. one of the non-negative Lyapunov exponents is zero, one might have pure a.c. spectrum. An example for that case is shown in Chapter 8.
3. TRANSFER MATRIX AND SCATTERING MATRIX
Chapter 4

Scaling diagram for the Lyapunov exponent at a bandedge

This chapter consists of the work done in the publication [72]. In this chapter we turn back to pure one dimensional models, i.e. we are working on $\ell^2(\mathbb{Z})$.

4.1 Main result and short overview

In Section 1.6 we calculated the Lyapunov exponent perturbatively for certain energies in the 1D Anderson model $H_0 + \lambda V$ where $H_0$ denotes the discrete Laplacian. The result was that for energies $E = 2 \cos(k)$, away from the band center and the band edges, i.e. $0 < |E| < 2$, Thouless formula [87] holds:

$$\gamma_\lambda(E) = \frac{E_\sigma(V_\sigma^2)}{8 \sin^2(k)} \lambda^2 + O(\lambda^3), \quad (4.1)$$

where $V_\sigma$ is one of the identically distributed and centered entries of $V$ and $E_\sigma$ denotes the expectation value.

Furthermore we saw that the control on the error terms breaks down at the band center, namely for $k = \frac{\pi}{2}, \frac{3\pi}{2}$; this leads to anomalies in the perturbative formula (4.1) first found by Kappus and Wegner [47], and consecutively analyzed by several authors [22, 11, 15, 80]. A perturbative formula as (4.1) for the band center with a control on the error terms was proven in [77]. In this chapter we will refine the methods used there to handle the band edges.

The method presented in Section 1.6 does not apply to the band edge, because the unperturbed transfer matrix (without disorder) is not diagonalizable and hence only equivalent to a Jordan normal form rather than a rotation matrix. This situation is typical for a bandedge (Proposition 4.1) and therefore we consider a slightly more general situation here. Namely given is some one-dimensional discrete random Schrödinger operator consisting of a periodic background operator and weakly-coupled random hoping terms and potential. The periodic (non-random) operator has a band structure. In the vicinity of a band edge,
we present a rigorous perturbation theory in the coupling constant and the energy for the Lyapunov exponent and the density of states. This leads to a new and rich scaling diagram describing these two self-averaging quantities at a band edge and showing how the $\lambda^2$ is modified. Parts of this diagram were already given by Derrida and Gardner [22]. These authors actually found the correct scaling in the parabolic regime of Theorem 4.1 below, but the wrong prefactor (cf. the comment at the end of Section 4.8). Moreover, they could not give a better justification of their scaling Ansatz than that it leads to a differential equation they could solve. Our more conceptual approach shows why the scaling is natural in the situation considered in [22]. It exhibits a far richer scaling behavior near a band edge and also allows to rigorously control the higher order corrections. Moreover, if one wants to give a perturbative proof of uniform positivity of the Lyapunov exponent in an energy interval around a band edge, all of the scaling behaviors considered below are needed.

Let us now describe the main result in more detail. We consider a one-parameter family of random Jacobi matrices $(H_{\lambda,\omega})_{\lambda \geq 0}$ given as the sum of an $J$-periodic background operator $H_0 = H_{0,\omega}$ and a random perturbation $H_{\lambda,\omega} - H_0$ which is linear in the (small) coupling constant $\lambda$. More precisely, for every fixed configuration $\omega$, the operator $H_{\lambda,\omega}$ acts on $\psi \in l^2(\mathbb{Z})$ as

$$H_{\lambda,\omega}|n\rangle = t_{\lambda,\omega}(n + 1)|n + 1\rangle + v_{\lambda,\omega}(n)|n\rangle + t_{\lambda,\omega}(n)|n - 1\rangle, \quad n \in \mathbb{Z},$$

where the coefficients $t_{\lambda,\omega}(n) > 0$ and $v_{\lambda,\omega}(n) \in \mathbb{R}$ are constructed as described in the following: let $(\hat{t}_1, \ldots, \hat{t}_J, \hat{v}_1, \ldots, \hat{v}_J)$ be given real constants with $\hat{t}_j > 0$ for $j = 1, \ldots, J$; set $\Sigma = [-1,1]^{2J}$ so that each $\sigma \in \Sigma$ is of the form $\sigma = (\hat{t}_1(\sigma), \ldots, \hat{t}_J(\sigma), \hat{v}_1(\sigma), \ldots, \hat{v}_J(\sigma))$; then $\Omega = \Sigma^\mathbb{Z} \times \{1, \ldots, J\}$ is the configuration space and to each $\omega = ((\sigma_m)_{m \in \mathbb{Z}}, k) \in \Omega$ there are associated sequences

$$t_{\lambda,\omega}(k - 1 + Jm + j) = \hat{t}_j + \lambda \hat{v}_j(\sigma_m), \quad v_{\lambda,\omega}(k - 1 + Jm + j) = \hat{v}_j + \lambda \hat{v}_j(\sigma_m),$$

which for $\lambda$ sufficiently small satisfy $t_{\lambda,\omega}(n) > 0$; these sequences define $H_{\lambda,\omega}$ by (4.2). In order to make $(H_{\lambda,\omega})_{\lambda \geq 0}$ into a family of random operators, we equip $\Omega$ with a probability measure $\mathbf{P} = \mathbf{p}^\mathbb{Z} \times \frac{1}{J} \sum_{j=1}^J \delta_{\hat{\sigma}}$ where $\mathbf{p}$ is a probability measure on $\Sigma$. Expectation values w.r.t. $\mathbf{P}$ and $\mathbf{p}$ will be denoted by $\mathbf{E}$ and $\mathbf{E}_\sigma$. We suppose that $\mathbf{E}_\sigma(\hat{t}_j(\sigma)) = \mathbf{E}_\sigma(\hat{v}_j(\sigma)) = 0$.

For every fixed energy $E \in \mathbb{R}$ and coupling parameter $\lambda$, there are two self-averaging quantities of interest, namely the integrated density of states (IDS) $\mathcal{N}_\lambda(E)$ and the Lyapunov exponent (or inverse localization length) $\gamma_\lambda(E)$. The definitions will be recalled in Section 4.4 below. The periodic operator $H_0 = H_{0,\omega}$ has a band structure and we are interested in the scaling of the IDS and Lyapunov exponent at one of its band edges $E_b$ (band touching excluded). In order to state the precise result, we need to introduce two quantities. Let $\mathcal{T}_{\lambda,\sigma}^E$ be the random transfer matrix at energy $E$ over a unit cell of length $J$ (see Section 4.2 for the explicit formula). Then set

$$x = \partial_E \text{Tr}(\mathcal{T}_{0,\sigma}^E)|_{E=E_b}, \quad x_\sigma = \partial_\lambda \text{Tr}(\mathcal{T}_{\lambda,\sigma}^E)|_{\lambda=0}.$$
4.1. MAIN RESULT AND SHORT OVERVIEW

Theorem 4.1 Let $E_b$ be a band edge (band touching excluded) of the periodic background operator $H_0 = H_{0,\omega}$ of a family of random Jacobi matrices $(H_{\lambda,\omega})_{\lambda \geq 0}$. The perturbation is supposed to be non-trivial in the sense that $x_\sigma$ does not vanish $\mathbf{p}$-almost surely. Then the scaling near the band edge is

$$
\lambda (E_b + \epsilon \lambda^\eta) = \lambda_0(E_b) + A \lambda^\alpha + \mathcal{O}(\lambda^{\alpha+\delta}), \quad \gamma_\lambda(E_b + \epsilon \lambda^\eta) = B \lambda^{\beta} + \mathcal{O}(\lambda^{\beta+\delta}), \quad \text{(4.3)}
$$

where $A, B, \alpha, \beta$ and $\delta > 0$ depend on $\epsilon \in \mathbb{R}$ and $\eta > 0$ as described in the following and resumed in Figure 4.1.

(i) (Elliptic regime) Let $\eta < \frac{4}{3}$ and let the sign of $\epsilon \neq 0$ be such that $E_b + \epsilon \lambda^\eta$ is inside the band of $H_0$. For the case of the Lyapunov exponent, we also suppose $\eta > \frac{4}{5}$. Then

$$
\alpha = \frac{\eta}{2}, \quad \beta = 2 - \eta,
$$

and

$$
A = \text{sgn}(\epsilon) \frac{\sqrt{|\epsilon x|}}{J \pi}, \quad B = \frac{1}{8J} \frac{E_\sigma(|x_\sigma|^2)}{|\epsilon x|}.
$$

(ii) (Parabolic regime) Let $\eta = \frac{4}{3}$, then

$$
\alpha = \beta = \frac{2}{3},
$$

and $A = A(\epsilon)$ and $B = B(\epsilon)$ are given by integrals written out explicitly in Section 4.8. For $\eta > \frac{4}{3}$, $A$ and $B$ are independent of $\epsilon$, namely the result is the same as for $\epsilon = 0$.

(iii) (Hyperbolic regime) Let $\frac{4}{5} < \eta < \frac{4}{3}$ and let the sign of $\epsilon \neq 0$ be such that $E_b + \epsilon \lambda^\eta$ is outside the band of $H_0$. Then

$$
\alpha > \frac{\eta}{2}, \quad \beta = \frac{\eta}{2},
$$

and

$$
B = \sqrt{|\epsilon x|}.
$$

The result is summarized in Figure 4.1. All the error estimates, in particular, the value of $\delta$, are controlled more explicitly and given in Section 4.8. As will become more clear below, the terms elliptic, parabolic and hyperbolic regimes reflect the nature of the corresponding averaged transfer matrix to lowest order in $\lambda$. Let us point out that we do not provide the asymptotics for the IDS in the hyperbolic regime, but only give an upper bound on it. Actually the IDS between the Lifshitz tails and the band edge of the periodic operator is exponentially small in $\lambda$, a fact that results from a more delicate large deviation behavior. This will be dealt with elsewhere. The Lifshitz tails themselves were analyzed in [75].

The proof of Theorem 4.1 begins (in Section 4.2) with an adequate basis change on the transfer matrices. A band edge is characterized by the fact that the modulus of the
trace of the transfer matrix is 2. If there is no band touching, the transfer matrix is then not diagonalizable and its normal form is hence a Jordan block (see Proposition 4.1). The action of a Jordan block on projective space (in fact, identified with the Prüfer phases) has exactly one fixed point, which is unstable under random perturbations induced by the randomness in the Hamiltonian. In order to analyze the associated invariant distribution on projective space, a further basis change blowing up the vicinity of the fixed point is necessary. If this is done adequately, a formal perturbative calculation typically leads to a differential equation for the invariant distribution.

Instead of studying the concrete form of the random matrices obtained in Section 4.2 after the various basis changes, we rather choose to single out the more general concept of an anomaly of a family of random matrices (cf. Section 4.3). The term anomaly is chosen by reference to the center of band anomaly in the one-dimensional Anderson model as studied by Kappus and Wegner [47]. The latter is a special case which is analyzed in a prior work [77]. The formalism to calculate the Lyapunov exponents and rotation numbers at a general anomaly is built up in Section 4.4. In particular, it turns out that one needs to evaluate certain Birkhoff sums for the perturbative calculation of these quantities. The perturbative evaluation of these Birkhoff sums directly leads to a differential operator on projective space (Propositions 4.2 and 4.3). In some situations this operator is of first order, and consequently the anomaly is called first order as well (this is dealt with in Section 4.5). However, in the more interesting cases treated in Sections 4.6 and 4.7, the differential operator is of second order and of the Fokker-Planck type. The latter operator was already used in [77] in order to study the center of band anomaly, but here one is confronted with the supplementary difficulty that its ellipticity is destroyed at a band edge. Thus a thorough analysis of the corresponding singularities is needed. Section 4.6 deals with the Fokker-Planck operator and its groundstate, while Section 4.7 shows how it is used to calculate the Birkhoff sums. Both sections heavily depend on Appendix A.1, where inhomogeneous singular first order ordinary differential equations are studied in detail. Sections 4.3 up to 4.7 are kept slightly more general than needed for the proof of Theorem 4.1 as completed in Section 4.8, but this stresses the structural aspects of the analysis. There may possibly
also be further applications.

4.2 Normal form of transfer matrix at a band edge

In this section we motivate and carry out the basis changes on the transfer matrices in the vicinity of a band edge. The final normal forms obtained in the various regimes will then motivate the definition of the term anomaly in the next section.

First, one rewrites the Schrödinger equation \( H_{\lambda, \omega} \psi = E \psi \), with \( E \in \mathbb{R} \), \( \psi \in l^2(\mathbb{Z}) \) and \( H_{\lambda, \omega} \) as given in (4.2), in the standard way \( (t_{\lambda, \omega}(n+1) \Psi(n+1)) = T_{\lambda, \omega, n}(t_{\lambda, \omega}(n) \Psi(n)) \) using transfer matrices (e.g. [45]). Using the notation introduced in (4.4) below, the transfer matrices are given by \( T_{\lambda, \omega, n} = T^E(t_{\lambda, \omega}(n), \sigma(n)) \).

Due to the periodicity of \( H_0 \) it is convenient to introduce the transfer matrix over \( J \) sites associated to one configuration \( \sigma \in \Sigma \) of the disorder on these sites:

\[
T^E_{\lambda, \sigma} = \prod_{l=1}^{L} T^E(\hat{t}_l + \lambda \hat{t}(\sigma), \hat{v}_l + \lambda \hat{v}(\sigma)), \quad T^E(t, v) = \begin{pmatrix} (E - v)t^{-1} & -t \\ t^{-1} & 0 \end{pmatrix}.
\]

Note that \( T^E(t, v) \in \text{SL}(2, \mathbb{R}) \) and hence \( T^E_{\lambda, \sigma} \in \text{SL}(2, \mathbb{R}) \). If \( E_b \) is a band edge of \( H_0 \), then \( \text{Tr}(T^E_{0, \sigma}) = \pm 2 \) and, because there is no band touching, \( \partial_E \text{Tr}(T^E_{0, \sigma}) \neq 0 \). One can show that necessarily \( T^E_{0, \sigma} \) has only one eigenvector.

**Proposition 4.1** The transfer matrix at a band edge has only one eigenvector and is hence equivalent to one of the Jordan normal forms \( \pm \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \).

**Proof.** We may assume \( E_b = 0 \). Then we have a real analytic family \( E \in \mathbb{R} \mapsto T^E \in \text{SL}(2, \mathbb{R}) \) with

\[
\text{Tr}(T^0) = \pm 2, \quad \partial_E \text{Tr}(T^E)|_{E=0} \neq 0.
\]

Let us only consider the case where the trace is equal to 2. Then 1 is an eigenvalue as the characteristic polynomial is equal to \( \det(T^0 - x1) = (x - 1)^2 \). We show that necessarily the eigenvalue 1 of \( T^0 \) has geometric multiplicity 1, so that the Jordan form of \( T^0 \) is non-diagonal. A similar statement holds if \( \text{Tr}(T^0) = -2 \). For this purpose, let us choose the notation

\[
T^E = \begin{pmatrix} a + AE & b + BE \\ c + CE & d + DE \end{pmatrix} + O(E^2).
\]

Then one deduces \( ad - bc = 1, a + d = 2 \) and \( A + D \neq 0 \). Furthermore the order \( E \) of \( \det(T^E) = 1 \) implies that \( aD + dA - cB - bC = 0 \). We need to show that \( T^0 \neq 1 \). If \( a \neq 1 \), this is true. Hence suppose \( a = 1 \) so that also \( d = 1 \). Then \( ad - bc = 1 \) implies that either \( b = 0 \) or \( c = 0 \). Hence \( aD + dA - cB - bC = 0 \) implies that either \( A + D - cB = 0 \) or \( A + D - bC = 0 \). Because \( A + D \neq 0 \) it follows that either \( c \neq 0 \) or \( b \neq 0 \).

\( \square \)
Hence one can find a basis change\footnote{We denote this basis change by $N_J$ as instead of $N$, to distinguish the matrix from the variable $N$ used later as natural number. The index $J$ is chosen, as it leads to a Jordan normal form. In this and the next chapter, $N$ with an index is a matrix.} $N_J \in \text{SL}(2, \mathbb{R})$ such that $N_J T_{0,\sigma}^E N_J^{-1}$ is a Jordan block with eigenvalue either 1 or $-1$. Using this basis change, one can write the full energy rescaled transfer matrix as follows:

$$N_J T_{\lambda,\sigma}^{E_b + \epsilon \lambda} N_J^{-1} = \pm \exp \left( \begin{pmatrix} 0 & \pm 1 \\ 0 & 0 \end{pmatrix} + \sum_{k \geq 1} \lambda^{\eta_k} Q_{\eta_k,\sigma} \right). \tag{4.5}$$

Here the exponents $\eta_k$ are in $\mathbb{N} + \eta \mathbb{N}$. They are put into increasing order $\eta_k < \eta_{k+1}$ and the $Q_{\eta_k,\sigma}$ are in the Lie algebra $\text{sl}(2, \mathbb{R})$. Hence the lowest terms are $\lambda^\eta Q_{\eta,\sigma}$ or $\lambda Q_{1,\sigma}$, pending on the value of $\eta$. Let us note that $Q_{k\eta,\sigma}$ are independent of $\sigma$ unless $k\eta \in \mathbb{N}$. Also, $E_\sigma(Q_{1,\sigma}) = 0$ unless there is a $k$ such that $k\eta = 1$. Each combination of the signs can occur, the one inside the exponential indicates if it is a lower (+ sign) or an upper (- sign) bandedge. For sake of concreteness, we choose both signs positive.

A few further remarks on the $Q_{\eta_k,\sigma}$ will be relevant later on. Because there is no band touching at $E_b$, for $\epsilon \neq 0$ one gets

$$0 = \epsilon \partial_E \left. \text{Tr}(N_J T_{0,\sigma}^E N_J^{-1}) \right|_{E=E_b} = \text{Tr} \left( Q_{\eta,\sigma} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \right) = \epsilon x, \tag{4.6}$$

where the first identity follows from the definition (4.5), Duhamel’s formula and the cyclicity of the trace, and the second one defines $x$. By the same calculation (with $\epsilon = 0$)

$$\partial_\lambda \left. \text{Tr}(N_J T_{\lambda,\sigma}^{E_b} N_J^{-1}) \right|_{\lambda=0} = \text{Tr} \left( Q_{1,\sigma} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \right) = x_\sigma,$n

the latter by definition of the centered random variable $x_\sigma$. Of course this expression is centered because the perturbation is centered. However, the assumption in Theorem 4.1 on the non-invariance of the band edge under perturbation translates into

$$E \left( \left[ \text{Tr} \left( Q_{1,\sigma} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \right) \right]^2 \right) > 0. \tag{4.7}$$

It turns out (see Section 4.4 or [45]) that one needs to study the random dynamical system given by random iteration of the natural action of the transfer matrices $N_J T_{\lambda,\sigma}^{E_b + \epsilon \lambda} N_J^{-1}$ on $S^1_\pi \cong \mathbb{R}/\pi \mathbb{Z}$, given by (see Section 4.4 for details)

$$e_{\theta_n} = \pm \frac{N_J T_{\lambda,\sigma}^{E_b + \epsilon \lambda} N_J^{-1} e_{\theta_{n-1}}}{\| N_J T_{\lambda,\sigma}^{E_b + \epsilon \lambda} N_J^{-1} e_{\theta_{n-1}} \|}, \quad e_\theta = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}. \tag{4.8}$$

At $\lambda = 0$, the action is given by the graph of Figure 4.2. It has an unstable parabolic fixed point at $\theta = 0$, which is being approached from the left under the dynamics without
4.2. NORMAL FORM OF TRANSFER MATRIX AT A BAND EDGE

Figure 4.2: Dynamics at $\lambda = 0$ at lower band edge, i.e. + sign in (4.5).

passing it though. For $\lambda > 0$, the dynamics is the same to lowest order $\lambda^0$, but the random perturbations of higher order may allow the dynamics to cross through $\theta = 0$. Nevertheless, the angles $\theta_n$ are close to $\theta = 0$ for most $n$ and thus their distribution is more and more concentrated in the vicinity of $\theta = 0$ as $\lambda$ tends to 0 and, in fact, converges weakly to a Dirac peak. In order to extract the shape of the non-trivial distribution it is necessary to rescale the neighborhood of $\theta = 0$ with an adequate power of $\lambda$. This is done by conjugation with

$$N_{\lambda,\delta} = \begin{pmatrix} \lambda^\delta & 0 \\ 0 & 1 \end{pmatrix}.$$  

The conjugation has the following effect on $2 \times 2$ matrices:

$$N_{\lambda,\delta} \begin{pmatrix} a & b \\ c & d \end{pmatrix} N_{\lambda,\delta}^{-1} = \lambda^{-\delta} \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix} + \lambda^\delta \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix}.$$  \hspace{1cm} (4.9)

One deduces from (4.5)

$$N_{\lambda,\delta} N_J T_{E_{n+\epsilon \lambda^\eta}} N_{\lambda,\delta}^{-1} = \exp \left( N_{\lambda,\delta} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} N_{\lambda,\delta}^{-1} + \sum_{k \geq 1} \lambda^\eta_k N_{\lambda,\delta} Q_{\eta_k,\sigma} N_{\lambda,\delta}^{-1} \right).$$

Note that the powers of $\lambda$ appearing in the exponential now lie in $\mathbb{N} + \eta \mathbb{N} + \{-\delta,0,\delta\}$, but none is negative for $\delta < \min\{1,\eta\}$. The following three contributions of low order in $\lambda$ turn out to be relevant:

$$X_{\delta,1} = \lambda^\delta \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad X_{\delta,2} = \lambda^{\eta-\delta} \lim_{\lambda' \to 0} \lambda' N_{\lambda',\delta} Q_{\eta,\sigma} N_{\lambda',\delta}^{-1},$$

$$X_{\delta,\sigma,3} = \lambda^{1-\delta} \lim_{\lambda' \to 0} \lambda' N_{\lambda',\delta} Q_{1,\sigma} N_{\lambda',\delta}^{-1},$$

where the notation reflects that $Q_{\eta,\sigma}$ is independent of $\sigma$. By equations (4.9) and (4.6), $X_{\delta,2}$ has a non-vanishing constant entry only in the lower left corner unless $\epsilon = 0$. Similarly
by (4.9) and (4.7), \( X_{\delta,\sigma,3} \) has a centered random entry in the lower left corner with positive variance, while all the other entries vanish. It is helpful to think of \( X_{\delta,1} \) and \( X_{\delta,2} \) as drift terms in the action (4.8), and of \( X_{\delta,\sigma,3} \) as generator of a diffusion. Now we blow up the parabolic fixed point by augmenting \( \delta \) until the drift \( X_{\delta,1} \) is of same order of magnitude as either the drift \( X_{\delta,2} \) (then \( \delta = \frac{\eta}{2} \)) or the diffusive force (variance) of \( X_{\delta,\sigma,3} \) (then \( 2(1-\delta) = \delta \) so that \( \delta = \frac{2}{3} \)). Hence we choose

\[
\delta = \min \left\{ \frac{\eta}{2}, \frac{2}{3} \right\}.
\]

Explicitly, for \( \eta < \frac{4}{3} \), one has with this choice of \( \delta \)

\[
N_{\lambda,\delta} N_J T_{\lambda,\sigma} \frac{E_0 + \epsilon \lambda \eta}{\lambda} N_J^{-1} \lambda^{-1} = \exp \left( \frac{\lambda \eta}{2} \begin{pmatrix} 0 & 1 \\ \epsilon & 0 \end{pmatrix} \right) + \lambda^{-\frac{3}{2}} \frac{Q'_{\eta} + \lambda Q'_{1,\sigma} + O(\lambda^{-\frac{2}{3},\lambda^{1+\frac{2}{3}}})}{\lambda} \right), \tag{4.10}
\]

where the specific form of \( Q'_{\eta} \) and \( Q'_{1,\sigma} \) will not be of importance for the asymptotics of Theorem 4.1. For \( \eta > 1 \), the second term is of lower order than the first, but as it is centered, we shall see that it does not enter into the asymptotics of the IDS and the Lyapunov exponent because its variance is smaller than \( O(\lambda^{\frac{2}{3}}) \). Hence (4.10) is an anomaly of first order in the sense of the following section. In the case \( \eta = \frac{4}{3} \),

\[
N_{\lambda,\delta} N_J T_{\lambda,\sigma} \frac{E_0 + \epsilon \lambda \eta}{\lambda} N_J^{-1} \lambda^{-1} = \exp \left( \frac{\lambda \eta}{2} \begin{pmatrix} 0 & 0 \\ \epsilon & 0 \end{pmatrix} \right) + \lambda^{\frac{2}{3}} \left( \begin{pmatrix} 0 & 1 \\ \epsilon & 0 \end{pmatrix} \right) + O(\lambda) \right). \tag{4.11}
\]

This is actually the same formula as (4.10), but it was written out again in order to clearly exhibit the so-called anomaly of second order (see Definition 4.2 below), namely the variance of the first centered term is precisely of the same order of magnitude as the expectation value of the second. If \( \eta > \frac{4}{3} \), one obtains the same formula with \( \epsilon = 0 \), but with an additional error term \( O(\lambda^{\frac{2}{3}}) \).

### 4.3 Definition of anomalies

Let us consider families \( (T_{\lambda,\sigma})_{\lambda \in \mathbb{R}, \sigma \in \Sigma} \) of matrices in \( \text{SL}(2, \mathbb{R}) \) depending on a random variable \( \sigma \) in some probability space \( (\Sigma, \mathbf{p}) \) as well as a real coupling parameter \( \lambda \). In order to avoid technicalities, we suppose that \( T_{\lambda,\sigma} \) has compact support for small \( \lambda \). Furthermore we assume that the dependence on \( \lambda \) can be expanded in some power series with non-negative exponents (not necessarily integers). Later on, we shall choose \( T_{\lambda,\sigma} \) to be the matrices in (4.10) and (4.11).

**Definition 4.1** The value \( \lambda = 0 \) is an anomaly of the family \( (T_{\lambda,\sigma})_{\lambda \in \mathbb{R}, \sigma \in \Sigma} \) if for all \( \sigma \in \Sigma \) and a sign that may depend on \( \sigma \in \Sigma \):

\[
T_{0,\sigma} = \pm 1. \tag{4.12}
\]
In order to further classify the anomalies, let us suppose that for a fixed \( \lambda \)- and \( \sigma \)-
independent basis change \( M \in \text{SL}(2, \mathbb{R}) \) to be chosen later, the transfer matrix is of the
following form:

\[
M T_{\lambda, \sigma} M^{-1} = \pm \exp \left( \sum_{k \geq 1} \lambda^{\eta_k} P_{\eta_k, \sigma} \right).
\]

(4.13)

Here \( P_{\eta_k, \sigma} \in \text{sl}(2, \mathbb{R}) \), \( \eta_k > 0 \) for \( k \in \mathbb{N} \) and \( K \), such that \( \eta_j < \eta_k \) if \( j < k \), \( \eta_K = 2\eta_1 \), \( \eta_{K+1} \leq \eta_1 + \eta_2 \) and \( p(P_{\eta_k, \sigma} = 0) < 1 \) for \( k = 1, \ldots, K - 1 \) (which means none of these
matrices shall be identically 0 for \( p \)-almost all \( \sigma \)).

**Definition 4.2** Let \( (T_{\lambda, \sigma})_{\lambda \in \mathbb{R}, \sigma \in \Sigma} \) have an anomaly and suppose given the expansion
(4.13) for a fixed basis change \( M \). The anomaly is said to be of first order and \( K \)th kind if \( E(P_{\eta_k, \sigma}) = 0 \)
for \( k = 1, \ldots, K - 1 \) and \( E(P_{\eta_K, \sigma}) \) is non-vanishing. An anomaly of first order
and \( K \)th kind is called elliptic if \( \det(E(P_{\eta_K, \sigma})) > 0 \), hyperbolic if \( \det(E(P_{\eta_K, \sigma})) < 0 \) and
parabolic if \( \det(E(P_{\eta_K, \sigma})) = 0 \). Note that all these notions are independent of the choice of
\( M \).

If \( E(P_{\eta_k, \sigma}) = 0 \), for \( k = 1, \ldots, K - 1 \) (then the variance of \( P_{\eta_k, \sigma} \) is non-vanishing), then
an anomaly is said to be of second order.

It may happen that the transfer matrices have to be modified as follows in order to
obtain an anomaly. If there is a mapping \( \lambda \mapsto M_\lambda \in \text{GL}(2, \mathbb{R}) \) for \( \lambda > 0 \) where \( M_\lambda \)
is independent of \( \sigma \), such that

\[
\lim_{\lambda \to 0} M_\lambda T_{\lambda, \sigma} M_\lambda^{-1} = \pm 1,
\]

(4.14)

then we say that \( T_{\lambda, \sigma} \) is transformed to an anomaly by \( M_\lambda \). Note that the limits \( \lim_{\lambda \downarrow 0} M_\lambda \)
and \( \lim_{\lambda \downarrow 0} M_\lambda^{-1} \) need not exist.

As argued in the last section, the transfer matrices at a band edge can be trans-
formed into anomalies (4.10) and (4.11) of respectively first and second order. The term
anomaly first appeared in the work of Kappus and Wegner [47] on the band center of a
one-dimensional Anderson model. Indeed, as discussed in [77], the square of the transfer
matrix at a band edge is an anomaly in the above sense (of second order, if the random
potential is centered). More generally, higher powers of transfer matrices are transformed
into anomalies for rational quasimomenta, and thus lead to anomalies in higher order per-
turbation theory [22]. A more systematic classification of anomalies as in Definition 4.2
was done in [77]. However, we felt it adequate to change the term degree of an anomaly
used in [77] to the present order of an anomaly because it leads to a differential equation
of corresponding order.

### 4.4 Phase shift dynamics

The bijective action \( S_T \) of a matrix \( T \in \text{SL}(2, \mathbb{R}) \) on \( S^1_\pi = \mathbb{R}/\pi \mathbb{Z} = [0, \pi) \) is given by

\[
e_{S_T(\theta)} = \pm \frac{T e_\theta}{\|T e_\theta\|}, \quad e_\theta = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}, \quad \theta \in S^1_\pi ,
\]

(4.15)
with an adequate choice of the sign. This defines a group action, namely \( S_{T'} = S_T S_{T'} \). In particular the map \( S_T \) is invertible and \( S_{T^{-1}} = S_{T^{-1}} \). Note that this is an action on projective space. In order to shorten notations, we write

\[
S_{\lambda,\sigma} = S_{MT,\sigma} M^{-1}.
\]

At an anomaly one has \( S_{\lambda,\sigma}(\theta) = \theta + \mathcal{O}(\lambda) \).

Given an initial angle \( \theta_0 \) and iterating this dynamics defines a Markov process \( \theta_n(\omega, \theta_0) \) on \( \Omega = \Sigma^k \times \{1, \ldots, L\} \), i.e. for \( \omega = ((\sigma_n)_{n \in \mathbb{Z}}, k) \) one defines iteratively

\[
\theta_0(\omega, \theta_0) = \theta_0, \quad \theta_{n+1}(\omega, \theta_0) = S_{\lambda,\sigma_{n+1}}(\theta_n(\omega, \theta_0)).
\]

In order to shorten notations, we denote \( \theta_n(\omega, \theta_0) \) by \( \theta_n \). Under adequate identifications, this corresponds to the dynamics (4.8). The \( \theta_n \) are also called modified Prüfer variables.

In order to analyze the dynamics in more detail, some further notations are needed. For \( k \in \mathbb{N} \), we define the trigonometric polynomials:

\[
p_{k,\sigma}(\theta) = \Im \left( \frac{\langle v | P_{\eta,\sigma}| e_\theta \rangle}{\langle v | e_\theta \rangle} \right), \quad v = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}.
\]

More explicitly, in terms of the matrix elements one obtains

\[
P_{\eta,\sigma} = \begin{pmatrix} a_{\eta,\sigma} & b_{\eta,\sigma} \\ c_{\eta,\sigma} & -a_{\eta,\sigma} \end{pmatrix} \quad \implies \quad p_{k,\sigma}(\theta) = c_{\eta,\sigma} \cos^2(\theta) - b_{\eta,\sigma} \sin^2(\theta) - a_{\eta,\sigma} \sin(2\theta).
\]

Now starting from the identity

\[
e^{2\pi S_{\lambda,\sigma}(\theta)} = \frac{\langle v | MT_{\lambda,\sigma} M^{-1} | e_\theta \rangle}{\langle v | e_\theta \rangle},
\]

the definition (4.13) and the identity \( \langle v | e_\theta \rangle = \frac{1}{\sqrt{2}} e^{i\theta} \), one can expand \( S_{\lambda,\sigma}(\theta) \) to get

\[
S_{\lambda,\sigma}(\theta) = \theta + \Im \left( \sum_{k=1}^{K} \lambda_{\eta} \frac{\langle v | P_{\eta,\sigma} | e_\theta \rangle}{\langle v | e_\theta \rangle} + \lambda^{2n} \left[ \frac{\langle v | P_{\eta,\sigma} | e_\theta \rangle}{2 \langle v | e_\theta \rangle} - \frac{\langle v | P_{\eta,\sigma} | e_\theta \rangle^2}{4 \langle v | e_\theta \rangle^2} \right] \right)
\]

with an error of order \( \mathcal{O}(\lambda^{\eta+1}) \). As one readily verifies that

\[
P_{\eta,\sigma}^2 = -\det(P_{\eta,\sigma}) 1, \quad \Im \left( \frac{\langle v | P_{\eta,\sigma} | e_\theta \rangle^2}{\langle v | e_\theta \rangle^2} \right) = -p_{1,\sigma} \partial_{\theta} p_{1,\sigma}(\theta),
\]

it follows that

\[
S_{\lambda,\sigma}(\theta) = \theta + \sum_{k=1}^{K} \lambda_{\eta} p_{k,\sigma}(\theta) + \frac{1}{2} \lambda^{2n} p_{1,\sigma} \partial_{\theta} p_{1,\sigma}(\theta) + \mathcal{O}(\lambda^{\eta+1}).
\]
4.4. PHASE SHIFT DYNAMICS

As one has \( \exp \left( \sum \lambda^n P_{\eta,\sigma} \right)^{-1} = \exp \left( - \sum \lambda^n P_{\eta,\sigma} \right) \) the same procedure leads to

\[
S_{\lambda,\sigma}^{-1}(\theta) = \theta - \sum_{k=1}^{K} \lambda^n p_{k,\sigma}(\theta) + \frac{1}{2} \lambda^{2n} \partial_{\theta} p_{\eta,\sigma}(\theta) + \mathcal{O}(\lambda^{3n+1}). \tag{4.19}
\]

Our main interest is the perturbative calculation of the Lyapunov exponent \( \gamma(\lambda) \) and rotation number \( \mathcal{R}(\lambda) \) associated to the random family of matrices \((T_{\lambda,\sigma,n})_{n \geq 1}\) which are defined by

\[
\gamma(\lambda) = \lim_{N \to \infty} \frac{1}{N} \mathbb{E} \sum_{n=0}^{N-1} \mathbb{E}_\sigma \log(\| MT_{\lambda,\sigma} M^{-1} \epsilon_{\theta_n} \|), \tag{4.20}
\]

and

\[
\mathcal{R}(\lambda) = \frac{1}{\pi} \lim_{N \to \infty} \frac{1}{N} \mathbb{E} \sum_{n=1}^{N} \mathbb{E}_\sigma \varphi_{\lambda,\sigma}(\theta_n). \tag{4.21}
\]

where \( \varphi_{\lambda,\sigma}: S_\pi \to \mathbb{R} \) is the phase shift given by \( \varphi_{\lambda,\sigma}(\theta) = S_{\lambda,\sigma}(\theta) - \theta \). If \( T_{\lambda,\sigma} = \tilde{N}T_{\lambda,\sigma}^{E+\epsilon \lambda^n} \tilde{N}^{-1} \) where the latter is the transfer matrix as given in (4.4) associated to a random Jacobi matrix \( H_{\lambda,\omega} \) with an \( L \)-periodic background operator defined in (4.2) and \( \tilde{N} \) is an arbitrary basis change such as the ones used in (4.10) and (4.11), then the Lyapunov exponent \( \gamma_{\lambda}(E) \) (inverse localization length) and IDS \( N_{\lambda}(E) \) of the random Jacobi matrix are

\[
\gamma_{\lambda}(E + \epsilon \lambda^n) = \frac{1}{f} \gamma(\lambda), \quad N_{\lambda}(E + \epsilon \lambda^n) = -\frac{1}{f} \mathcal{R}(\lambda) \mod \frac{1}{f}. \tag{4.22}
\]

The first identity follows immediately from the definition; the second is a consequence of the oscillation theorem and the gap labelling for the periodic operator (see e.g. [45] for details). In both of the expressions (4.20) and (4.21), one can now expand each summand w.r.t. \( \lambda \).

**Lemma 4.1** Set

\[
\alpha_{\eta,\sigma} = \langle v | P_{\eta,\sigma} | v \rangle, \quad \beta_{\eta,\sigma} = \langle \overline{v} | P_{\eta,\sigma} | v \rangle, \quad \gamma_{\eta,\sigma} = \langle \overline{v} | P_{\eta,\sigma} |^2 | v \rangle
\]
as well as

\[
\delta_{\eta,\eta_2,\sigma} = \frac{1}{2} \overline{\langle \overline{v} | (P_{\eta,\sigma} + P_{\eta_1,\sigma}^*) P_{\eta_2,\sigma} + (P_{\eta_2,\sigma} + P_{\eta_1,\sigma}^*) P_{\eta_1,\sigma} | v \rangle}.
\]

Then \( p_{\eta,\sigma}(\theta) = \Im m(\alpha_{\eta,\sigma} - \beta_{\eta,\sigma} e^{2i\theta}) \). Furthermore, with errors of order \( O(\lambda^{3n}) \),

\[
\log(\| MT_{\lambda,\sigma} M^{-1} \epsilon_{\theta} \|) = \Re e \sum_{k \geq 1} \left( \lambda^{n_k} \beta_{\eta,\sigma} e^{2i\theta} + \frac{\lambda^{2n_k}}{2} \left( | \beta_{\eta,\sigma} |^2 + | \gamma_{\eta,\sigma} | e^{2i\theta} - | \beta_{\eta,\sigma} |^2 e^{4i\theta} \right) \right)
\]

\[
+ \Re e \sum_{k_1 < k_2} \left( \lambda^{n_{k_1}+n_{k_2}} \left( \beta_{\eta_{k_1},\sigma} \beta_{\eta_{k_2},\sigma} + \delta_{\eta_{k_1},\eta_{k_2},\sigma} e^{2i\theta} - \beta_{\eta_{k_1},\sigma} \beta_{\eta_{k_2},\sigma} e^{4i\theta} \right) \right), \tag{4.23}
\]
and up to errors of order $\mathcal{O}(\lambda^{n+\eta})$

$$\varphi_{\lambda,\sigma}(\theta) = \Im \sum_{k \geq 1} \left( \lambda^{\eta_k} \left( \alpha_{\eta_k,\sigma} - \beta_{\eta_k,\sigma} e^{2i\theta} \right) + \frac{1}{2} \lambda^{2\eta_k} \left( -2 \alpha_{\eta_k,\sigma} \beta_{\eta_k,\sigma} e^{2i\theta} + \beta_{\eta_k,\sigma}^2 e^{4i\theta} \right) \right).$$

(4.24)

**Proof.** This follows from straightforward algebra using

$$\langle e_\theta | T | e_\theta \rangle = \frac{1}{2 \pi} \text{Tr}(T) + \text{Re} \left( \langle \bar{v} | T | v \rangle e^{2i\theta} \right), \quad \text{Tr}(|P_{\eta,\sigma}|^2 + P_{\eta,\sigma}^2) = 4|\beta_{\eta,\sigma}|^2,$$

as well as the identities $\text{Tr}(P_{\eta,\sigma}) = 0$ and $\langle \bar{v} | P_{\eta,\sigma}^2 | v \rangle = 0$. \hfill \qed

Once these formulas are replaced in (4.20) and (4.21), one hence needs to consider Birkhoff sums of the type

$$I_N(f) = \frac{1}{N} \mathbb{E} \sum_{n=0}^{N-1} f(\theta_n), \quad I(f) = \lim_{N \to \infty} I_N(f),$$

(4.25)

for $\pi$-periodic functions $f$. For $\gamma(\lambda)$ and $\mathcal{R}(\lambda)$ one actually only needs the functions $e^{2i\theta}$ and $e^{4i\theta}$. These Birkhoff sums have to be evaluated perturbatively in $\lambda$ with a rigorous control on the error terms. For this purpose, one needs to know the distribution of the $\theta_n$ as generated by the random dynamics (4.18). This is the main focus of the next sections.

### 4.5 First order anomalies

It turns out that it is easiest to calculate the Birkhoff sums $I_N(f)$ in case of an elliptic first order anomaly of $K$th kind. As then $\det(E(P_{\eta,K,\sigma})) > 0$, one can choose the basis change $M$ in (4.13) such that

$$E(P_{\eta,K,\sigma}) = \begin{pmatrix} 0 & -\mu \\ \mu & 0 \end{pmatrix},$$

(4.26)

with $\mu \neq 0$. Hence $\mathbb{E}(p_{\eta,K,\sigma}(\theta)) = \mu$.

**Proposition 4.2** Let $T_{\lambda,\sigma}$ have an elliptic anomaly of first order and $K$th kind and suppose that $\mathbb{E}(p_{\eta,K,\sigma}(\theta)) = \mu$. Then for any $f \in C^1(S^1_\lambda)$, one has

$$I_N(f) = \int_0^\pi \frac{d\theta}{\pi} f(\theta) + \mathcal{O}\left(\lambda^{K'-\eta}, (\lambda^{\eta} N)^{-1}\right),$$

with $K' = \min\{\tilde{K}, K\} > K$ where $\tilde{K}$ is such that $P_{\eta,K}$ is the first uncentered term in (4.13) after $P_{\eta,K}$. 

4.5. FIRST ORDER ANOMALIES

Proof. Because \( I_N(f) = c + I_N(f - c) \) for \( c = \int_0^\pi d\theta f(\theta)/\pi \), we may assume that \( \int_0^\pi d\theta f(\theta) = 0 \). Then \( f \) has an antiderivative \( F \in C^2(S^2_\pi) \). Using the Taylor expansion and \( 2\eta_1 = \eta_K \),

\[
F(S_{\lambda,\sigma}(\theta)) = F\left(\theta + \sum_{k=1}^K \lambda^{\eta_k} p_{k,\sigma}(\theta) + O(\lambda^{\eta_{K+1}})\right)
= F(\theta) + F'(\theta) \sum_{k=1}^{K-1} \lambda^{\eta_k} p_{k,\sigma}(\theta) + O(\lambda^{\eta_K}).
\]

From this one deduces after taking expectation and summing over \( n \):

\[
I_N(F) = I_N(F) + \mu \lambda^{-\eta_K} I_N(f) + O(\lambda^{-\eta_K}, N^{-1}).
\]

As \( \mu \neq 0 \) this proves the estimate. \( \square \)

Concerning hyperbolic and parabolic first order anomalies, we do not provide an exhaustive treatment, but rather present some procedures on how these anomalies can (possibly) be transformed into a more accessible anomaly such as one of second order. This applies, in particular, to the hyperbolic first order anomaly (4.10) with \( \varepsilon \xi \) negative, corresponding to the hyperbolic regime in Theorem 4.1. For a hyperbolic first order anomaly of \( K \)-th kind, let us first choose \( M \) such that

\[
E(P_{\eta_K,\sigma}) = \begin{pmatrix} -\mu & 0 \\ 0 & \mu \end{pmatrix}, \quad \mu > 0.
\]

(4.27)

Then to lowest order in \( \lambda, \theta = 0 \) is the unstable and \( \theta = \frac{\pi}{2} \) the stable fixed point of (the averaged dynamics of) \( S_{\lambda,\sigma} \). Like in (4.11), we use the \( \lambda \)-dependent basis change \( N_{\lambda,\delta} \). Let \( \chi \) be the smallest exponent such that \( E(c_{\xi,\sigma}) > 0 \). First suppose that \( E(c_{\xi,\sigma}) = 0 \) for all \( \xi < \frac{1}{2} \eta_K + \chi \), then choose \( \delta = \chi - \frac{1}{2} \eta_K > 0 \) so that \( 2(\chi - \delta) = \eta_K \). One obtains

\[
N_{\lambda,\delta} M T_{\lambda,\sigma} M^{-1} N_{\lambda,\delta}^{-1} = \pm \exp \left[ \lambda^{\eta_K} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + \ldots + \lambda^{\eta_K} \begin{pmatrix} a_{\eta_K,\sigma} \\ b_{\eta_K,\sigma} \\ -a_{\eta_K,\sigma} \end{pmatrix} + \ldots \right],
\]

(4.28)

where, moreover, \( E(b_{\eta_K,\sigma}) = 0 \) and \( E(a_{\eta_K,\sigma}) = -\mu < 0 \). This is a second order anomaly for which the Birkhoff sums are analyzed in Theorem 4.4 below. If, on the other hand, there is a \( \xi < \frac{1}{2} \eta_K + \chi \) such that \( E(c_{\xi,\sigma}) \neq 0 \) and \( \xi \) is the smallest such exponent, choose \( \delta = \xi - \eta_K > 0 \) so that

\[
N_{\lambda,\delta} M T M^{-1} N_{\lambda,\delta}^{-1} = \pm \exp \left[ \lambda^{\zeta} \begin{pmatrix} a_{\zeta,\sigma} \\ a_{\zeta,\sigma} \end{pmatrix} + \ldots + \lambda^{\eta_K} \begin{pmatrix} a_{\eta_K,\sigma} \\ b_{\eta_K,\sigma} \\ -a_{\eta_K,\sigma} \end{pmatrix} \right]
\]

where \( \zeta \) is either \( \eta_1 \) or \( \chi - \delta \). For both cases we have \( 2\zeta > \eta_K \). As \( E(b_{\eta_K,\sigma}) = 0 \), one thus has again a hyperbolic first degree anomaly. Repeating the procedure, one may now be in the above advantageous case.

A parabolic first order anomaly of \( K \)-th kind can be treated just as a the band edge (which is parabolic to 0th order) in Section 4.2. One first chooses \( M \) to be a rotation
matrix, such that the parabolic fixed point is \( \theta = 0 \), namely \( \mathbf{E}(P_{\eta K,\sigma})e_0 = 0 \), and then considers the matrices \( N_{\lambda,\delta}MT_{\lambda,\sigma}M^{-1}N_{\lambda,\delta}^{-1} \). Let again \( \chi \) be the smallest exponent such that \( \mathbf{E}(c_{\chi,\sigma}^2) > 0 \) and let \( \xi \) be the smallest exponent such that \( \mathbf{E}(c_{\xi,\sigma}) \neq 0 \). If \( \xi \geq \frac{4}{3}\lambda + \frac{1}{2}\eta_{K} \) and \( \mathbf{E}(a_{\eta,\sigma}) = 0 \) for all \( \eta < \frac{2}{3}(\chi + \eta_{K}) \), then one chooses \( \delta = \frac{1}{3}(2\chi - \eta_{K}) > 0 \) in order to get a second order anomaly

\[
\pm \exp \left[ \lambda^2(\chi + \eta_{K}) \begin{pmatrix} a_{\chi-\delta,\sigma} & 0 \\ c_{\chi,\sigma} & -a_{\chi-\delta,\sigma} \end{pmatrix} + \ldots + \lambda^2(\chi + \eta_{K}) \begin{pmatrix} a_{\eta K + \delta,\sigma} & b_{\eta K,\sigma} \\ c_{\eta K + 2\delta,\sigma} & -a_{\eta K + \delta,\sigma} \end{pmatrix} + \ldots \right].
\]

### 4.6 Fokker-Planck operator of a second order anomaly

In order to calculate \( I_N(f) \) at a second order anomaly, we first refine the proof of Proposition 4.2. This naturally leads to an associated differential operator, as shown in the next proposition. The remainder of this section analyzes the properties of these operators.

**Proposition 4.3** Let the random family \( (T_{\lambda,\sigma})_{\sigma \in \Sigma} \) have a second order anomaly and \( F \in C^3(S^1_n) \). Then, for

\[
f = \mathbf{E}(p_{1,\sigma}^2)F'' + (2 \mathbf{E}(p_{K,\sigma}) + \mathbf{E}(p_{1,\sigma}\partial_\theta p_{1,\sigma}))F',
\]

one has

\[
I_N(f) = \mathcal{O}(\lambda^{\tilde{\eta}-\eta_{K}}, (\lambda^{\eta_{K}} N)^{-1}).
\]

with \( \tilde{\eta} = \min\{\eta_{\tilde{K}}, \eta_{1} + \eta_{2}\} \) where \( \tilde{K} \) is the first uncentered term in (4.13) after \( P_{\eta_{K},\sigma} \).

**Proof.** A Taylor expansion implies

\[
F(S_{\lambda,\sigma}(\theta)) = F \left( \theta + \sum_{k=1}^{\tilde{K}-1} \lambda^{n_k} p_{K,\sigma}(\theta) + \frac{1}{2} \lambda^{2n_1} p_{1,\sigma} \partial_\theta p_{1,\sigma}(\theta) + \mathcal{O}(\lambda^{\tilde{\eta}}) \right) =
\]

\[
F(\theta) + F'(\theta) \sum_{k=1}^{\tilde{K}-1} \lambda^{n_k} p_{K,\sigma}(\theta) + \frac{1}{2} \lambda^{2n_1} \left( p_{1,\sigma} \partial_\theta p_{1,\sigma}(\theta) F'(\theta) + p_{1,\sigma}^2(\theta) F''(\theta) \right) + \mathcal{O}(\lambda^{\tilde{\eta}}).
\]

Taking expectation values and summing over \( n \), one gets the estimate by the same argument as in Proposition 4.2 because \( \eta_{K} = 2\eta_{1}. \)

Let us sketch the further strategy. In view of (4.29), it is natural to introduce:

\[
p = \mathbf{E}(p_{1,\sigma}^2), \quad q = 2 \mathbf{E}(p_{K,\sigma}) + \mathbf{E}(p_{1,\sigma} \partial_\theta p_{1,\sigma}).
\]

(4.30)

Both \( p \) and \( q \) are the trigonometric polynomials of degree 4 on \( S^1_n \), namely linear combinations of \( e^{\pm 4\theta}, e^{\pm 2i\theta} \) and the constant function. Moreover, \( p \geq 0 \). Hence \( p \) can have at most one zero of order 4, or two zeros of order 2. Actually a zero \( \tilde{\theta} \) of order 4 only appears in the rather special case described next. By applying a rotation matrix \( M \) in (4.13), one may assume \( \tilde{\theta} = \frac{\pi}{2} \) and then one readily checks that

\[
p(\theta) = \mathbf{E}(c_{\eta_{1},\sigma}^2) \cos^4(\theta), \quad P_{\eta_{1},\sigma} = \begin{pmatrix} 0 & 0 \\ c_{\eta_{1},\sigma} & 0 \end{pmatrix}.
\]
4.6. **FOKKER-PLANCK OPERATOR OF A SECOND ORDER ANOMALY**

This is precisely the situation in (4.11). Furthermore, let us note that \( q \) can have zeros with orders adding up to at most 4.

Proposition 4.3 now states that one can control Birkhoff sums of functions which are in the range of the differential operator on \( C^3(S^1_\pi) \) given by

\[
L = (p \partial_\theta + q) \partial_\theta.
\]

The formal adjoint on \( L^2(S^1_\pi) \), also considered as an operator defined on \( C^3(S^1_\pi) \), is given by

\[
L^* = \partial_\theta (p \partial_\theta - q).
\]

The operator \( L^* \) is a Fokker-Planck (or forward Kolmogorov) operator on \( S^1_\pi \) describing a drift-diffusion dynamics on \( S^1_\pi \), while \( L \) is the associated backward Kolmogorov operator [67]. The operator \( L^* \) was already used in [77], but only in the situation of a strictly positive \( p \). In the case of the anomaly (4.11) corresponding to the parabolic regime at a band edge, \( p \) has a zero. Hence the ellipticity of \( L^* \) is destroyed. This is the main technical problem that has to be dealt with in the present work.

As \( \text{Ran}(L) \subset \text{Ker}(L^*)^\perp \), a necessary condition for controlling the Birkhoff sum \( I_N(f) \) is hence that \( f \in \text{Ker}(L^*)^\perp \). As will shortly be shown in Theorem 4.2, in many situations \( \text{Ker}(L^*) \) is one-dimensional and spanned by a smooth positive, \( L^1 \)-normalized function \( \rho \). Let us call \( \rho \) the groundstate of the Fokker-Planck operator, even though the spectrum of the Fokker-Planck operator is non-positive. Furthermore, the necessary condition \( f \in \text{Ker}(L^*)^\perp \) combined with \( f \in C^2(S^1_\pi) \) will in most cases turn out (Theorem 4.3 below) to be sufficient for finding a solution \( F \in C^3(S^1_\pi) \) of the inhomogeneous differential equation (4.29). Thus one deduces from Proposition 4.3 that, for \( f \in C^2(S^1_\pi) \),

\[
I_N(f) = \int d\theta \, \rho(\theta) \, f(\theta) + O(\lambda^{\eta K + 1 - \eta K}, (\lambda^{\eta K} N)^{-1}).
\]

In some other situations, relevant for the hyperbolic regime in Theorem 4.1, the kernel of \( L^* \) is given by a Dirac peak \( \rho \) (if \( L^* \) is considered as an operator on the space of distributions), and the same formula holds.

Before proving the main result on the groundstate \( \rho \), let us point out that \( \rho \) can be seen as the lowest order formal approximation to the Furstenberg measure \( \nu_\lambda \) on \( S^1_\pi \) associated to a family \( (T_{\lambda,\sigma})_{\sigma \in \Sigma} \) of random matrices. It is defined by [9]

\[
\int_0^\pi \nu_\lambda(d\theta) \, f(\theta) = \mathbb{E} \int_0^\pi \nu_\lambda(d\theta) \, f(S_{\lambda,\sigma}(\theta)), \quad f \in C(S^1_\pi) .
\]

(4.31)

In fact, supposing \( \nu_\lambda(d\theta) = \rho_\lambda(\theta) \, d\theta \), (4.31) leads to

\[
\mathbb{E}\left( \partial_\theta S_{\lambda,\sigma}^{-1}(\theta) \rho_\lambda(S_{\lambda,\sigma}^{-1}(\theta)) \right) = \rho_\lambda(\theta) .
\]

Furthermore assuming that \( \rho_\lambda \) is twice differentiable and \( \rho_\lambda = \rho + O(\lambda^{\eta K + 1 - \eta K}) \), this gives the equation \( L^* \rho = 0 \) (cf. [77] for further details).
Theorem 4.2 The Fokker-Planck operator $\mathcal{L}^*$ has a unique groundstate $\rho$, which is non-negative, normalized and continuous except possibly in the following cases:

(I) $p$ has a zero $\hat{\theta}$ of order 2 (and possibly another zero) for which, moreover,

$$\partial_\theta p(\hat{\theta}) = q(\hat{\theta}) = 0, \quad \partial^2_\theta p(\hat{\theta}) \geq \partial_\theta q(\hat{\theta}). \tag{4.32}$$

(II) $p$ has a zero $\hat{\theta}$ of order 4 for which, moreover,

$$E(p_{K,\sigma}(\hat{\theta})) = 0, \quad E(\partial_\theta p_{K,\sigma}(\hat{\theta})) \leq 0. \tag{4.33}$$

Furthermore, the groundstate is infinitely differentiable except possibly at one point $\hat{\theta}$, at which $p$ has a zero of order 2 and

$$\partial_\theta p(\hat{\theta}) = q(\hat{\theta}), \quad \partial^2_\theta p(\hat{\theta}) < \partial_\theta q(\hat{\theta}). \tag{4.34}$$

Let us note that (4.32) is equivalent to

$$E(\partial_\theta p_{\tilde{1},\sigma}(\hat{\theta})) = 4E(p_{K,\sigma}(\hat{\theta})), \quad E(\partial^2_\theta p_{\tilde{1},\sigma}(\hat{\theta})) > 4E(\partial_\theta p_{K,\sigma}(\hat{\theta})).$$

Furthermore, let us point out that whenever there is a point $\hat{\theta}$ which is a zero of both $p$ and $q$, then the Dirac peak $\delta_{\hat{\theta}}$ is a groundstate of $\mathcal{L}^*$ if considered as an operator on distributions. This Dirac peak is dynamically stable under the drift-diffusion described by $\mathcal{L}^*$ if, moreover,

$$\partial^l ((\partial_\theta p - q)(\hat{\theta})) > 0, \quad l \in \{1, 3\}.$$ 

This is precisely the case in (I) and (II). A corresponding dual result, namely that the Birkhoff sums $I_N(f)$ are equal to $f(\hat{\theta})$ to lowest order, will be proven in Theorem 4.4 below. Let us also note that there may be two such angles (e.g. if $E(p_{K,\sigma}) = 0$), in which case the stable groundstate is degenerate. Finally, we encourage the reader to draw the curves of $p$ and $q$ in each of the cases of the proof below, and interpret $p$ as a diffusive force, and $q$ as a drift. This allows to better understand the formal proof, as well as the result itself.

Proof of Theorem 4.2. The equation $\mathcal{L}^*\rho = 0$ can be integrated once. Hence one needs to solve

$$(p \partial_\theta + \tilde{q}) \rho = C, \quad \tilde{q} = (\partial_\theta p) - q, \tag{4.35}$$

where the real constant $C$ has to be chosen such that (4.35) admits a non-negative, $\pi$-periodic, $L^1$-normalized solution $\rho$. The equation (4.35) is locally precisely the type of singular ODE studied in Appendix A.1, with coefficients which are moreover real analytic. The only supplementary property of (4.35) is that the inhomogeneity is constant. When writing out solutions, we use similar notations as in the appendix, except that we shall put tildes on $w$ and $W$ in order to distinguish them from the corresponding dual objects appearing in the proof of Theorem 4.3 below, and that we do not include the constant inhomogeneity $C$ in the definition of $\tilde{W}$ below as done in (A.5).
4.6. FOKKER-PLANCK OPERATOR OF A SECOND ORDER ANOMALY

If \( p > 0 \) so that there is no singularity and \( \mathcal{L}^* \) is elliptic, the groundstate \( \rho \) can readily be calculated as in [77]. Set, for some \( \hat{\theta} \in S^1_x \),

\[
\tilde{w}(\theta) = \int_\theta^\theta d\xi \frac{\tilde{q}(\xi)}{p(\xi)}, \quad \tilde{W}(\theta) = \int_\theta^\theta d\xi \frac{e^{\tilde{w}(\xi)}}{p(\xi)}.
\]

Then

\[
\rho = C_1 e^{-\tilde{w}} (C_2 \tilde{W} + 1), \quad C_2 = \frac{e^{\tilde{w}(\hat{\theta} + \pi)} - 1}{\tilde{W}(\hat{\theta} + \pi)}, \quad C = C_1 C_2,
\]

where one first determines the normalization constant \( C_1 \) which then also fixes \( C \) in (4.35) by the last identity. Note that \( \tilde{W}(\theta) > 0 \) for \( \theta > \hat{\theta} \) so that, in particular, \( C_2 \) is well-defined. Furthermore \( C_2 = 0 \) if \( w \) is \( \pi \)-periodic. The groundstate \( \rho \) is clearly real analytic in this case.

Now let \( \hat{\theta} \) be the only zero of \( p \). It is either of order 2 or 4. If \( \tilde{q}(\hat{\theta}) \neq 0 \), one is locally near \( \hat{\theta} \) in the situation of Proposition A.1(iv) or (v). For each \( C \) in (4.35), the smooth integrable solution is unique to one side of \( \hat{\theta} \). Continuing it cyclically around \( S^1_x \) shows that it is also unique on the other side of \( \hat{\theta} \). For \( \hat{\theta} \leq \theta < \hat{\theta} + \pi \), this smooth solution is given by

\[
\rho = C e^{-\tilde{w}} \tilde{W},
\]

with \( \tilde{w} \) and \( \tilde{W} \) defined as in (4.36) using \( \hat{\theta} < \tilde{\theta} < \hat{\theta} + \pi \) in the first equation and \( \tilde{\theta} = \hat{\theta} \) or \( \tilde{\theta} = \hat{\theta} + \pi \) in the second one, such that the singularities cancel. It only remains to choose \( C \neq 0 \) (and hence the ODE (4.35) itself) such that \( \rho \) is normalized.

If \( p \) has two zeros \( \hat{\theta}_1 \) and \( \hat{\theta}_2 \) of order 2 for which \( \tilde{q}(\hat{\theta}_1) \neq 0 \) and \( \tilde{q}(\hat{\theta}_2) \neq 0 \), one has to distinguish two cases. If the signs of \( \tilde{q}(\hat{\theta}_1) \) and \( \tilde{q}(\hat{\theta}_2) \) are the same, then one proceeds cyclically twice as in the case with one zero of \( p \) which was just treated. The solution is as in (4.38) and \( C \neq 0 \). If on the other hand \( 0 < \hat{\theta}_1 < \hat{\theta}_2 < \pi \) with \( \tilde{q}(\hat{\theta}_1) < 0 \) and \( \tilde{q}(\hat{\theta}_2) > 0 \), then one chooses \( C = 0 \) and sets

\[
\rho(\theta) = \begin{cases} 
C_3 e^{-\tilde{w}(\theta)} & \hat{\theta}_1 < \theta < \hat{\theta}_2, \\
0 & \text{otherwise,}
\end{cases}
\]

with \( \tilde{w} \) defined as in (4.36) for some \( \tilde{\theta} \) with \( \hat{\theta}_1 < \tilde{\theta} < \hat{\theta}_2 \) and some normalization constant \( C_3 \). Note that this \( \rho \) has no singularities because \( \lim_{\theta \to \hat{\theta}_1} \tilde{w}(\theta) = +\infty \) and \( \lim_{\theta \to \hat{\theta}_2} \tilde{w}(\theta) = +\infty \).

As shown in Proposition A.1(iv) and (v), \( \rho \) is smooth at both \( \hat{\theta}_1 \) and \( \hat{\theta}_2 \).

Now follow the singular cases where \( p \) and \( q \) have common zeros. The situations (I) and (II) are of this type, but there will appear others for which we then show as stated that there is a unique continuous groundstate. As argued in Appendix A.1, there can then only be integrable solutions of (4.35) if the equation is homogeneous, namely \( C = 0 \). One such solution is \( \rho = 0 \), but there may be others. First let \( p \) have only one zero \( \hat{\theta} \) of order 2 and suppose that \( \tilde{q}(\hat{\theta}) = 0 \). Then \( \hat{\theta} \) is a zero of order 2 of \( q \) if and only if \( \partial_\theta \tilde{q}(\hat{\theta}) = 0 \), that is \( E(\partial_\theta^2 p_{1,\sigma}(\hat{\theta})) = 4 E(\partial_\theta p_{K,\sigma}(\hat{\theta})) \). In this situation, there is a one-parameter family
of smooth solutions according to Proposition A.1(i). However, these solutions are typically not periodic and hence this is part of situation (I). If \( \hat{\theta} \) is a zero of order 1 of \( \tilde{q} \), then Proposition A.1(ii) and (iii) can be used for analyzing the local regularity. Case (iii) applies if \( \partial_q \tilde{q}(\hat{\theta}) < 0 \). Thus one has a two-parameter family of continuous solutions near \( \hat{\theta} \); however, one parameter has to assure the periodicity of the solution so that it remains continuous, while the second has to be chosen so that \( \rho \) is normalized. Hence the unique continuous groundstate is \( \rho = C_4 e^{-\tilde{w}} \) with an adequate normalization constant \( C_4 \). In the situation (I) when (4.32) holds with a strict inequality, one has \( \partial_q \tilde{q}(\hat{\theta}) > 0 \) so that Proposition A.1(ii) implies that the solution is unique. As \( \rho = 0 \) is one solution, it is the only one.

Next let us consider the case where \( p \) has two zeros \( \hat{\theta}_1 \) and \( \hat{\theta}_2 \) of order 2, and that at least one of them, say \( \hat{\theta}_1 \), is a zero of \( \tilde{q} \). If \( \partial_q \tilde{q}(\hat{\theta}_1) > 0 \) (namely situation (I) again), there is only the zero solution due to Proposition A.1(ii).

If \( \partial_q \tilde{q}(\hat{\theta}_1) = 0 \) and \( \tilde{q}(\hat{\theta}_2) \neq 0 \), the singularity \( \hat{\theta}_1 \) can due to Proposition A.1(i) only be resolved by choosing \( C = 0 \), but then the zero \( \hat{\theta}_2 \) for which \( \tilde{q}(\hat{\theta}_2) \neq 0 \) leads to a singularity. Thus there is no continuous solution and this is included in situation (I). If \( \partial_q \tilde{q}(\hat{\theta}_1) < 0 \) and say \( \tilde{q}(\hat{\theta}_2) > 0 \), then one can appeal to Proposition A.1(iii) at \( \hat{\theta}_1 \) and Proposition A.1(iv) at \( \hat{\theta}_2 \), and construct the groundstate exactly as in (4.39). However, \( \rho \) might not be differentiable at \( \hat{\theta}_1 \). If \( \hat{\theta}_1 \) and \( \hat{\theta}_2 \) are zeros of both \( p \) and \( \tilde{q} \), then either \( \partial_q \tilde{q}(\hat{\theta}_1) > 0 \) or \( \partial_q \tilde{q}(\hat{\theta}_2) > 0 \). Indeed, \( \tilde{q} = \frac{1}{2} \partial_q p - C(p_{K,\sigma}) \) and \( C(p_{K,\sigma}) \) is a trigonometric polynomial of degree 1 and hence can only compensate either \( \partial_q p(\hat{\theta}_1) > 0 \) or \( \partial_q p(\hat{\theta}_2) > 0 \). Therefore this corresponds again to situation (I) and there is no normalizable solution.

In the last remaining case \( \hat{\theta} \) is a zero of \( p \) of order 4 (hence the only zero of \( p \)) and \( \tilde{q}(\hat{\theta}) = 0 \). If \( \partial_q \tilde{q}(\hat{\theta}) < 0 \), which is equivalent to \( C(\partial_q p_{K,\sigma} (\hat{\theta})) > 0 \), Proposition A.1(vii) can be applied and the groundstate \( \rho = C_5 e^{-\tilde{w}} \) is smooth. If \( \partial_q \tilde{q}(\hat{\theta}) > 0 \), then Proposition A.1(vi) implies that the zero solution is the only solution. If \( \partial_q \tilde{q}(\hat{\theta}) = 0 \) and \( \partial_q p(\hat{\theta}) = 0 \) identically, then \( \tilde{q} = \frac{1}{2} \partial_q p \) has a zero at \( \hat{\theta} \) of order 3 and \( \partial_q \tilde{q}(\hat{\theta}) > 0 \), so that Proposition A.1(ii) implies that the zero solution is the only solution. The same holds if \( \partial_q \tilde{q}(\hat{\theta}) = 0 \) and \( \partial_q p(\hat{\theta}) \neq 0 \) due to Proposition A.1(iv) or (v), applied either to the left or right of \( \hat{\theta} \) pending on the sign of \( \partial_q p_{K,\sigma}(\hat{\theta}) \). The last three cases are regrouped in situation (II).

\[ \square \]

4.7 Birkhoff sums at second order anomalies

In this section we calculate the Birkhoff sums \( I_N(f) \) perturbatively for many second order anomalies. This involves the analysis of the operator \( \mathcal{L} \) dual to the Fokker-Planck operator. We will not present a treatment as exhaustive as Theorem 4.2 for \( \mathcal{L}^* \) and do not treat the singular cases where \( p \) and \( q \) have common zeros, except for the case of Theorem 4.4 below, which is needed for the analysis of a band edge. Comments on the remaining cases are given at the end of the section.

**Theorem 4.3** Let \( (T_{l,\sigma})_{\sigma \in \Sigma} \) have an anomaly of second order. Suppose \( q(\hat{\theta}) \neq 0 \) whenever \( E(p_{1,\sigma}^2(\hat{\theta})) = 0 \). Furthermore let \( \rho \) be the groundstate of the Fokker-Planck operator given
by Theorem 4.2. Then, for \( f \in C^2(S^1_\pi) \) one has
\[
I_N(f) = \int_0^\pi d\theta f(\theta) \rho(\theta) + O\left(\lambda^{\tilde{\eta}-\eta_K},(\lambda^{\eta_K} N)^{-1}\right)
\]
with \( \tilde{\eta} = \min\{\eta_\tilde{K},\eta_1 + \eta_2\} \) where \( \tilde{K} \) is the first uncentered term in (4.13) after \( P_{\eta_K}\sigma \). If \( E(p^2_{1,\sigma}) \) has no zero, it is enough to suppose \( f \in C^1(S^1_\pi) \).

**Proof.** As already argued in Section 4.6, one can control the Birkhoff sums using Proposition 4.3 only if \( f \in \text{Ker}(L^*\perp) \). Under the hypothesis stated, the kernel of \( L^* \) is one-dimensional by Theorem 4.2 and spanned by the groundstate \( \rho \). Replacing \( f \) by \( f - \langle \rho | f \rangle \), we may now assume that \( \int_0^\pi d\theta f(\theta) \rho(\theta) = 0 \) and have to show for such \( f \) that the Birkhoff sum \( I_N(f) \) is of the order stated. This will follow directly from Proposition 4.3 once we have found \( G \in C^2(S^1_\pi) \) satisfying
\[
(p \partial_\theta + q) G = f, \quad \int_0^\pi d\theta G(\theta) = 0 .
\]  
Indeed, the second identity allows to take an antiderivative \( F \) of \( G \) which then satisfies (4.29).

First let us consider the case where \( p \) has no zero and \( C \neq 0 \) in (4.35). Setting, for some \( \tilde{\theta} \in S^1_\pi \),
\[
w(\theta) = \int_\theta^\tilde{\theta} d\xi \frac{q(\xi)}{p(\xi)}, \quad W(\theta) = \int_\theta^\tilde{\theta} d\xi \frac{e^{w(\xi)}}{p(\xi)} f(\xi) ,
\]  
the solution is given by
\[
G = e^{-w}(W + c_1) , \quad c_1 = \frac{e^{-w(\tilde{\theta}+\pi)} W(\tilde{\theta} + \pi)}{1 - e^{-w(\tilde{\theta}+\pi)}},
\]
where \( e^{-w(\tilde{\theta}+\pi)} \neq 1 \) as this happens if and only if \( w \) and \( \tilde{w} \) are \( \pi \)-periodic leading to \( C = 0 \). One readily checks that the first equation in (4.40) is satisfied and that \( G \) is periodic, hence in \( C^2(S^1_\pi) \) because \( f \in C^1(S^1_\pi) \). Furthermore, multiplying the first equation in (4.40) by \( \rho \) and integrating over \( S^1_\pi \) shows that also the second equation holds:
\[
0 = \int \rho f = \int \rho (p \partial_\theta + q) G = - \int G (\partial_\theta p - q) \rho = -C \int G(\theta) .
\]

Next, if \( p \) has no zeros and \( C = 0 \) in (4.35), then \( C_2 = 0 \) in (4.37) and
\[
\rho = C_1 e^{-\tilde{w}} = C_1 \frac{e^{w}}{p} , \quad (p \partial_\theta + q) e^{-w} = 0 ,
\]  
so that \( W \) defined as in (4.41) satisfies \( W(\tilde{\theta} + \pi) = W(\tilde{\theta}) = 0 \) because \( \int \rho f = 0 \). Hence
\[
G = e^{-w}(W + c_2)
\]
is a (periodic) function in \( C^2(S_1^1) \) and, for adequate choice of \( c_2 \), of vanishing integral.

Now let \( p \) have one zero \( \hat{\theta} \). The hypothesis implies \( q(\hat{\theta}) \neq 0 \). According to Proposition A.1(iv) or (v), the first equation in (4.40) has a \( C^2 \)-solution in a neighborhood of \( \hat{\theta} \), which is unique to one side of \( \hat{\theta} \). Continuing this solution around \( S_1^1 \) one deduces that the solution is unique, and actually given by \( G = e^{-w}W \). Due to the argument in the proof of Theorem 4.2, \( C \neq 0 \) in (4.35) in this case, so that by the same argument as above the second equation in (4.40) is satisfied. The case with two zeros \( \hat{\theta}_1 \) and \( \hat{\theta}_2 \) of for which \( q(\hat{\theta}_1) \) and \( q(\hat{\theta}_2) \) have the same sign is treated similarly (cf. the proof of Theorem 4.2).

Finally we deal with the case of two zeros \( 0 \leq \hat{\theta}_1 < \hat{\theta}_2 < \pi \) of for which \( q(\hat{\theta}_1) > 0 \) and \( q(\hat{\theta}_2) < 0 \). By (4.39), the support of the groundstate is \([\hat{\theta}_1, \hat{\theta}_2]\) and hence not all of \( S_1^1 \). For \( \theta \in [\hat{\theta}_1, \hat{\theta}_2] \), we choose \( \hat{\theta}_1 < \theta < \hat{\theta}_2 \) in the first equation of (4.41) and \( \hat{\theta} = \hat{\theta}_1 \) in the second equation of (4.41). Then \( G = e^{-w}W \) is a solution in \((\hat{\theta}_1, \hat{\theta}_2)\). Due to Proposition A.1(iv) this is the unique solution that can be smoothly extended to the left of \( \theta_1 \). Furthermore (4.42) and \( \int \rho f = 0 \) imply \( W(\hat{\theta}_2) = 0 \), so that this solution can also be smoothly extended through \( \hat{\theta}_2 \) by Proposition A.1(v). On both sides one has one free parameter. One is chosen such that \( G \) is \( \pi \)-periodic and the remaining one such that the second equation of (4.40) is also satisfied. This reflects that \( G_0 = e^{-w}w_{[\hat{\theta}_2, \hat{\theta}_1]} \) is a smooth positive solution of the homogeneous equation \((p \partial_\theta + q)G_0 = 0\). \( \square \)

The remaining singular cases for which Theorem 4.2 guarantees existence of a unique groundstate cannot be treated by the techniques of Theorem 4.3 without further hypothesis on \( f \), which seem somewhat unnatural, but reflect the delicate dynamical behavior at such points (e.g. one needs that \( f \) and its derivative have a particular behavior near the common zero of \( p \) and \( q \)). The next theorem deals with an anomaly corresponding to situation (II) of Theorem 4.2, albeit with a strict inequality in (4.33).

**Theorem 4.4** Let \( (T_{\lambda, \sigma})_{\sigma \in \Sigma} \) have an anomaly of second order. Suppose that \( \hat{\theta} \) is a zero of order 4 of \( E(p_{1, \sigma}^2) \) and a zero of order 1 of \( E(p_{K, \sigma}) \) and that, moreover, \( E(\partial_\theta p_{K, \sigma}(\hat{\theta})) < 0 \). Then for \( f \in C^2(S_1^1) \)

\[
I_N(f) = f(\hat{\theta}) + O(\lambda^{\eta_K}, (\lambda^{\eta_K} N)^{-1})
\]

with \( \hat{\eta} = \min\{\eta_K, \eta_1 + \eta_2\} \) where \( K \) is the first uncentered term in (4.13) after \( P_{\eta_K, \sigma} \).

**Proof.** The procedure is exactly as in the proof of Theorem 4.3. Hence we search for a solution of (4.40) for a function \( f \) satisfying \( f(\hat{\theta}) = 0 \). Close to \( \hat{\theta} \) the equation \((p \partial_\theta + q)G = f\) has a two-parameter family of solutions due to Proposition A.1(vii). One parameter is fixed by requiring \( G \) to be periodic. The other reflects that one can, by the same procedure, find a non-vanishing solution \( G_0 \) to \((p \partial_\theta + q)G_0 = 0 \). As \( G_0 \) has definite sign, one can use it to normalize \( G \) so that also the second equation in (4.40) is satisfied. \( \square \)

As to extensions of Theorem 4.4, the case with an equality in the second equation of (4.33) cannot be treated by the presented technique because \((p \partial_\theta + q)G_0 = 0 \) has only the trivial solution by Proposition A.1(iv) or (v). The case with one zero of order 2 corresponding to situation (I) in Theorem 4.2 can be treated similarly as long as \( 3 \| E(p_{1, \sigma}^2(\hat{\theta})) \| < 4 \| E(\partial_\theta p_{K, \sigma}(\hat{\theta})) \| \), a condition which originates in Proposition A.1(iii).
4.8 Application to a band edge

This section contains the proof of Theorem 4.1. Hence let \( T_{\lambda, \sigma} = N_{\lambda, \sigma} N J T_{\lambda, \sigma}^{E_\beta + \epsilon \lambda^\eta} N_{J}^{-1} N_{\lambda, \sigma}^{-1} \) be the anomaly given in (4.10) or (4.11). Then (4.22) allows to calculate the Lyapunov exponent and the IDS, so that one may focus on the calculation of \( \gamma(\lambda) \) and \( R(\lambda) \) based on (4.20), (4.21) and the expansions given in Lemma 4.1.

Let us begin with the elliptic first order regime, hence (4.10). The adequate basis change assuring (4.26) is

\[
M = \begin{pmatrix} 1 & 0 \\ 0 & -|\epsilon x|^{-\frac{1}{2}} \end{pmatrix}.
\]

Then \( (MN_{\lambda, \delta} N J) T_{\lambda, \sigma}^{E_\beta + \epsilon \lambda^\eta} (MN_{\lambda, \delta} N J)^{-1} \) is equal to

\[
\exp \left( \lambda^\eta |\epsilon x|^{\frac{1}{2}} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} - \lambda^{1-\eta} \frac{x_\sigma}{|\epsilon x|^{\frac{1}{2}}} \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} + \lambda^\eta P_\eta + \lambda P_{1, \sigma} + O(\lambda^{\frac{3\eta}{2}}, \lambda^{1+\frac{\eta}{2}}) \right).
\]

Here \( E(P_{1-\frac{\eta}{2}, \sigma}) = 0 \), and also \( E(P_{1, \sigma}) = 0 \). With the notations of Lemma 4.1, this gives \( \beta_\frac{\eta}{2} = 0 \) and \( \beta_{1-\frac{\eta}{2}, \sigma} = i \epsilon x_\sigma |\epsilon x|^{-\frac{1}{2}} / 2 \). Thus Proposition 4.2 implies \( I(\epsilon^{2j+\delta}) = O(\lambda^{\frac{\eta}{2}}, \lambda^{2-\frac{2\eta}{2}}) \) for \( j = 1, 2 \). Hence the first 4 terms with powers \( \lambda^m \) in (4.23) either vanish or are of higher order, the same holds for the mixed terms \( \lambda^{m_1+\eta_2} \), the first of the terms with \( \lambda^{2n} \) is also of higher order, but the second one gives the leading order contribution:

\[
\gamma(\lambda) = \lambda^{2-\eta} \frac{E(x_\sigma^2)}{8|\epsilon x|} + O(\lambda^{\frac{3\eta}{2}}, \lambda^{1+\frac{\eta}{2}}),
\]

which together with (4.22) establishes the first claim. Similarly one verifies \( \alpha_\frac{\eta}{2} = i |\epsilon x|^{\frac{1}{2}} \) so that by (4.24) one finds \( R(\lambda) = \lambda \frac{3\eta}{2} |\epsilon x|^{\frac{1}{2}} + O(\lambda^\eta, \lambda^{2-\eta}) \).

In the hyperbolic regime, the basis change leading to (4.27) is

\[
M = \begin{pmatrix} 1 & -|\epsilon x|^{-\frac{1}{2}} \\ 1 & |\epsilon x|^{-\frac{1}{2}} \end{pmatrix}.
\]

Then \( (MN_{\lambda, \delta} N J) T_{\lambda, \sigma}^{E_\beta + \epsilon \lambda^\eta} (MN_{\lambda, \delta} N J)^{-1} \) calculated from (4.10) is equal to

\[
\exp \left( \lambda^\eta |\epsilon x|^{\frac{1}{2}} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} + \lambda^{1-\eta} \frac{x_\sigma}{2|\epsilon x|^{\frac{1}{2}}} \begin{pmatrix} -1 & -1 \\ 1 & 1 \end{pmatrix} + \lambda^\eta P_\eta + \lambda P_{1, \sigma} + O(\lambda^{\frac{3\eta}{2}}, \lambda^{1+\frac{\eta}{2}}) \right).
\]

Now let us apply a further basis change \( N_{\lambda, \delta} \) as described in (4.28), namely in the present situation \( \eta_\kappa = \frac{\eta}{2}, \chi = 1 - \frac{\eta}{2} \) and \( \delta = 1 - \frac{3\eta}{4} \). Due to (4.9) the transformed anomaly then becomes of second order:

\[
\exp \left( \lambda^\eta |4 \epsilon x|^{-1} \begin{pmatrix} 0 & 0 \\ x_\sigma & 0 \end{pmatrix} + \lambda^\eta |\epsilon x|^{\frac{1}{2}} \begin{pmatrix} -1 & 0 \\ 1 & 1 \end{pmatrix} + \lambda^{1-\eta} P_{1-\frac{\eta}{2}, \sigma} + \lambda^{2-\frac{3\eta}{2}} P_{2-\frac{3\eta}{4}, \sigma} + O \right),
\]
where $\mathcal{O} = \mathcal{O}(\lambda^{\frac{2}{3}}, \lambda^{\frac{2}{3}-1})$. Here $P'_{1-\frac{2}{3},\sigma}$ and $P'_{2-\frac{2}{3},\sigma}$ are possibly of lower order than the second term (but not as the first one as $\eta < \frac{1}{3}$), but they are both centered so that they do not enter into the asymptotics (cf. Definition 4.2). The obtained second order anomaly can now be dealt with by Theorem 4.4, implying that $I(f) = f(\pi) + \mathcal{O}(\lambda^{\frac{5}{3}}, \lambda^{1-\frac{4}{3}})$. Furthermore, one calculates $\beta_{\frac{2}{3},\sigma} = -ix_\sigma|\epsilon x|^{-\frac{3}{4}}/4$, $\gamma_{\frac{2}{3},\sigma} = x_\sigma|\epsilon x|^{-1/8}$ and $\gamma_{\frac{2}{3}} = -|\epsilon x|^2$. Replacing into (4.23) one realizes that the factor of $\lambda^{\frac{2}{3}}$ vanishes so that $\gamma(\lambda) = \lambda^{\frac{2}{3}}|\epsilon x|^{\frac{1}{4}} + \mathcal{O}(\lambda^{\frac{5}{3}}, \lambda^{1-\frac{4}{3}})$.

Concerning the rotation number, one has $\alpha_{\frac{2}{3}} = 0$ and $\alpha_{\frac{2}{3},\sigma} = ix_\sigma|\epsilon x|^{-\frac{3}{4}}/4$. The latter implies that the contribution $\lambda^{\frac{2}{3}}$ vanishes in (4.24) due to the $\mathcal{O}(m)$. Thus $\mathcal{R}(\lambda) = \mathcal{O}(\lambda^{\frac{5}{3}}, \lambda^{1-\frac{4}{3}})$. Note that the condition $\frac{7\eta}{4} - 1 > \frac{5}{4}$ is equivalent to $\eta > \frac{4}{7}$.

Finally let us deal with the parabolic case for which (4.11) is directly a second order anomaly. The polynomials (4.30) are explicitly given by

$$p(\theta) = E(x_\sigma^2) \cos^4(\theta), \quad q(\theta) = \epsilon x - 1 + (\epsilon x + 1) \cos(2\theta) - 2 E(x_\sigma^2) \cos^3(\theta) \sin(\theta).$$

Because $q(\frac{\pi}{2}) = -2 \neq 0$, Theorem 4.2 guarantees the existence of a unique normalized groundstate $\rho \in C^\infty(S^1_{\pi})$. As $\alpha_{\frac{1}{3},\sigma} = ix_\sigma/2$, $\beta_{\frac{1}{3},\sigma} = -ix_\sigma/2$ and $\gamma_{\frac{1}{3},\sigma} = x_\sigma^2/2$, one deduces from Lemma 4.1 and Theorem 4.3 that up to errors of order $\mathcal{O}(\lambda)$, $\gamma(\lambda)$ is given by

$$\lambda^{\frac{2}{3}} \left( \frac{1}{2} (\epsilon x + 1) \int d\theta \rho(\theta) \sin(2\theta) + \frac{E(x_\sigma^2)}{8} \int d\theta \rho(\theta) (1 + 2 \cos(2\theta) + \cos(4\theta)) \right),$$

and, furthermore that $\mathcal{R}(\lambda)$ is, also up to errors of order $\mathcal{O}(\lambda)$, equal to

$$\frac{\lambda^{\frac{2}{3}}}{2\pi} \left( \int \int d\theta \rho(\theta) \cos(2\theta) - \frac{E(x_\sigma^2)}{4} \int d\theta \rho(\theta) (2 \sin(2\theta) + \sin(4\theta)) \right).$$

Let us remark that this formula is not the same as equation (35) of [22], even though the invariant measure of [22] coincides with the above. In fact, their expansion of the expression $\log(1 + \lambda^{2/3}t) \approx \lambda^{2/3}t$ is erroneous for large $t$. 

4. SCALING DIAGRAM AT A BANDEDGE
Chapter 5

Perturbation Theory with strongly mixing potential

So far we considered random Schrödinger operators on $\mathbb{Z}$ with an i.i.d. random potential. In this chapter we want to relax the i.i.d. condition a little bit and show how the methods of the previous chapter can be modified to treat this case. The contents of this chapter is the work of the publication [73].

5.1 Introduction

As before let $\Sigma$ be a topological space and $\Omega = \Sigma^\mathbb{Z}$ the associated Tychonov product space. Furthermore let $\mathbf{P}$ be a probability measure on $\Omega$ which is invariant and ergodic w.r.t. the left shift $S : \Omega \to \Omega$. (A product measure $\mathbf{p}^\mathbb{Z}$ like in the previous sections is of course an example.) Now given a measurable real-valued function $V$ on $\Omega$ and a coupling constant $\lambda > 0$, one can associate an ergodic family of Jacobi matrices $(H_{\lambda, \omega})_{\omega \in \Omega}$ (also called discrete Schrödinger operators) each acting on $\ell^2(\mathbb{Z})$:

$$H_{\lambda, \omega} |n\rangle = |n + 1\rangle + \lambda V(S^n \omega) |n\rangle + |n - 1\rangle,$$

where $|n\rangle$ is the Dirac notation for the state in $\ell^2(\mathbb{Z})$ localized at site $n \in \mathbb{Z}$. If $\mathbf{P} = \mathbf{p}^{\otimes \mathbb{Z}}$ is a product measure of a compactly supported probability measure $\mathbf{p}$ on $\Sigma$ and $V(\omega)$ depends only on one entry of $\omega$, say $\omega_0 \in \Sigma$, then the random variables of the sequence $(V(S^n \omega))_{n \in \mathbb{Z}}$ of potential values are independent and we have the same model as considered before. As we have seen in Section 1.4, this model exhibits so-called Anderson localization, namely the spectrum of $H_{\lambda, \omega}$ is $\mathbf{P}$-almost surely pure-point with exponentially localized eigenstates [64] and the induced quantum dynamics is bounded in time (in the precise sense given below, [44]). The question considered in this chapter (and many others, see the reviews [43, 20] and references therein) concerns the spectral properties as well as the quantum dynamics in situations where the random variables $(V(S^n \omega))_{n \in \mathbb{Z}}$ are correlated. This situation typically arises when the dynamical system $(\Omega, S, \mathbf{P})$ is the symbolic dynamics associated to a (possibly weakly) hyperbolic discrete time dynamics; then $\Sigma$ is the Markov partition. If now the correlations of the potential decay sufficiently fast, then one expects that the model is
still in the regime of Anderson localization. Here we complement on the prior work \[17, 10\] and prove that this holds at least in a weak sense when the correlations satisfy a power law decay.

The proof of localization for these models is based on the positivity of the Lyapunov exponent. This positivity can either be established by Kotani theory \[20\], a version of Furstenberg’s theorem for correlated random matrices (work by Avila and Damanik cited in \[20\]) or by a perturbative calculation (for small \(\lambda\)) of the Lyapunov exponent. This latter calculation was first done by Chulaevski and Spencer \[17\] by carrying over the argument of Thouless \[87\], in a version given by Pastur and Figotin \[64\] (shown in Section 1.6), to the case of correlated potential values. The resulting formula is recalled in Section 5.2. Based on this result, Bourgain and Schlag then proved localization \[10\]. The only flaw left is that in \[17\] not all energies could be dealt with, but the band center and the band edges were spared out. Here we show how the techniques of our prior works \[72, 77\], shown in the last chapter, on anomalies and band edges \[22, 47\] combine with those of \[17\] to rigorously control the perturbation theory for the Lyapunov exponent also at these energies. Instead of repeating the rather complicated proofs of \[10\], we then adapt to the case of correlated potentials the elementary and short argument of \[44\] showing that positivity of the Lyapunov exponents implies at most logarithmic growth of quantum dynamics and hence, by Guarneri’s inequality \[36\], zero Hausdorff dimension of the spectral measures. Even though this is a weaker localization result than pure-point spectrum with exponential localized eigenfunctions, it proves the behavior which is stable under perturbation and we hence consider, as argued in \[44\], that it already captures the physically relevant effect. In the next section, the results and the precise hypothesis are described and discussed in detail. The other sections contain the proofs.

## 5.2 Set-up and results

In order to fix terms and notations, we have to begin by reviewing some basic definitions of symbolic dynamics and strong mixing \[12, 63\]. Let \(\Sigma\) be a countable set furnished with the discrete topology. We designate a reference element \(0 \in \Sigma\). For any subset \(I \subset \mathbb{Z}\) and \(\omega = (\sigma_n)_{n \in \mathbb{Z}} \in \Omega\), let us define

\[
\pi_I(\omega) = (\hat{\sigma}_n)_{n \in \mathbb{Z}}, \quad \hat{\sigma}_n = 0 \text{ for } n \notin I, \quad \hat{\sigma}_n = \sigma_n \text{ for } n \in I.
\]

For a bounded, measurable function \(g : \Omega \to V\) into a real, normed vector space \((V, \| \cdot \|)\), the variation on \(I\) is defined by

\[
\text{Var}_I(g) = \sup_{\pi_I(\omega) = \pi_I(\omega')} \| g(\omega) - g(\omega') \|.
\]

Then \(g\) is called quasi-local with rate \(0 < r < 1\) if and only if there exists a constant \(C = C(g)\) such that, for any \(m, n \geq 1\),

\[
\text{Var}_{[-m,n]}(g) \leq C(g) r^{m \wedge n}, \quad m \wedge n = \min\{m, n\}
\]  (5.2)
The set of all quasi-local functions with rate $r$ is denoted by $\mathcal{Q}_r(\mathcal{V})$.

Next let us state precisely the strong mixing hypothesis used in this work. For $m < n$ and $a_k \in \Sigma$ with $m \leq k \leq n$, the associated cylinder set is denoted by $A_{m,n}(a_m, \ldots, a_n) = \{ \omega = (\sigma_k)_{k \in \mathbb{Z}} | \sigma_k = a_k, m \leq k \leq n \}$. Then the invariant measure $\mathbf{P}$ on the shift space $(\Omega, \mathbb{Z})$ is said to satisfy a power law $\psi$-mixing [13] with exponent $\alpha > 0$ if there is a constant $C > 0$ such that for all $k < l < m < n$ and all $A_{k,l}$, $A_{m,n}$, one has

$$\left| \mathbf{P}(A_{k,l} \cap A_{m,n}) - \mathbf{P}(A_{k,l}) \mathbf{P}(A_{m,n}) \right| \leq C \mathbf{P}(A_{k,l}) \mathbf{P}(A_{m,n}) |m-l|^{-\alpha}.$$  \hspace{1cm} (5.3)

Equivalently, for any $\pi_{[k,l]}$-measurable function $g_1$ and $\pi_{[m,n]}$-measurable function $g_2$ holds

$$\left| \mathbf{E}(g_1 g_2) - \mathbf{E}(g_1) \mathbf{E}(g_2) \right| \leq C \mathbf{E}(|g_1|) \mathbf{E}(|g_2|) |m-l|^{-\alpha},$$  \hspace{1cm} (5.4)

where $k < l < m < n$ and $C$ as above. This also implies ergodicity. Examples when (5.4) holds are given in Remark 5.1 and 5.2 below, after the main results are stated. Averages over $\omega$ w.r.t. $\mathbf{P}$ are denoted by $\mathbf{E}$, or also by $\mathbf{E}_\omega$ if the dependence on $\omega$ is retained in the integrand. Furthermore, the set of centered quasi-local functions will be denoted by $\mathcal{Q}_r^0(\mathcal{V}) = \{ g \in \mathcal{Q}_r(\mathcal{V}) | \mathbf{E}(g) = 0 \}$.

Throughout we suppose that the potential in (5.1) is given by a centered real-valued quasi-local function $V \in \mathcal{Q}_r^0(\mathbb{R})$. It is well-known and verified in Lemma 5.4 that (5.4) implies the decay of correlations $|\mathbf{E}(V(\omega)V(S^m\omega))| \leq C|m|^{-\alpha}$ for some constant $C$. For $\alpha > 1$, one can hence define its (positive) spectral density $D_V(k) at k \in [0, 2\pi)$:

$$D_V(k) = \sum_{n \in \mathbb{Z}} e^{i kn} \mathbf{E}_\omega(V(\omega)V(S^m\omega)) = \lim_{N \to \infty} \frac{1}{N} \mathbf{E}_\omega \left( \left| \sum_{n=0}^{N-1} e^{i kn} V(S^n\omega) \right|^2 \right).$$  \hspace{1cm} (5.5)

As final preparation let us recall the definition of the Lyapunov exponent $\gamma_\lambda(E)$ at energy $E \in \mathbb{C}$ associated to (5.1). If the transfer matrices are defined by

$$T^E_{\lambda,S^n \omega} = \begin{pmatrix} E - \lambda V(\omega) & -1 \\ 1 & 0 \end{pmatrix} \in \mathcal{Q}_r(\text{SL}(2, \mathbb{R})), \hspace{1cm} (5.5)$$

then

$$\gamma_\lambda(E) = \lim_{N \to \infty} \frac{1}{N} \mathbf{E}_\omega \log \left( \left\| \prod_{n=1}^{N} T^E_{\lambda,S^n \omega} \right\| \right).$$

The main result of Chulaevski and Spencer [17] is that for $\alpha > 2$, at an energy $E = 2 \cos(k)$ in the spectrum $[-2, 2]$ of the discrete Laplacian away from the band edges $E = -2, 2$ and the band center $E = 0$, one has

$$\gamma_\lambda(E) = \lambda^2 \frac{D_V(k)}{8 \sin^2(k)} + O\left( \frac{\lambda^{3n+2}}{d(k)} \right),$$  \hspace{1cm} (5.6)

where $d(k)$ denotes the distance of $k$ from 0 mod $\frac{\pi}{2}$. As we need to build up the whole formalism anyway, the main element of the proof of (5.6) is reproduced in Section 5.6.
As indicated, the control of the error terms breaks down at the band edges and the band center. Our first result provides perturbative formulas for the Lyapunov exponent at these energies, generalizing respectively our prior results for independent potential values [77, 72].

**Theorem 5.1** Assume \( \alpha > 2, D_V(0) > 0 \) and \( D_V(\pi) > 0 \) (the latter is only needed for (i)).

(i) The Lyapunov exponent near the band center \( E = 0 \) is given by

\[
\gamma_\lambda(\epsilon\lambda^2) = \lambda^2 \frac{D_V(\pi)}{8} \int_0^\pi d\theta \rho_\epsilon(\theta) (1 + \cos(4\theta)) + \mathcal{O}(\lambda^{\frac{3\alpha+2}{\alpha+2}}), \tag{5.7}
\]

where \( \rho_\epsilon \) is a \( \pi \)-periodic smooth probability density.

(ii) Up to errors of order \( \mathcal{O}(\lambda^{\frac{3\alpha+2}{\alpha+2}}) \), the Lyapunov exponent near the upper band edge \( E = 2 \) is given by

\[
\gamma_\lambda(2 + \epsilon\lambda^\frac{4}{3}) = \lambda^\frac{2}{3} \int d\theta \rho_\epsilon(\theta) \left( \frac{1 - \epsilon}{2} \sin(2\theta) + \frac{D_V(0)}{8}(1 + 2\cos(2\theta) + \cos(4\theta)) \right), \tag{5.8}
\]

where \( \rho_\epsilon \) is a \( \pi \)-periodic smooth probability density written out explicitly in Section 5.8. The same formula holds at the lower band edge \( E = -2 \).

The formulas (5.6), (5.7) and (5.8) combined allow to study the Lyapunov exponents at all energies \([-2, 2]\). In order to assure positivity for \( \lambda > 0 \), one first has to check that the spectral density is positive (cf. Remark 5.3 below) and then prove that the integrals appearing in (5.7) and (5.8) are positive. This is immediate for (5.7). For (5.8) we could not produce an analytic proof, but, given the explicit formula (5.42) for \( \rho_\epsilon \), one can readily do a numerical evaluation.

Nevertheless, the three formulas are not yet sufficient to prove uniform positivity of the Lyapunov exponent on the whole spectrum for some fixed small, but positive value of \( \lambda \). Indeed, for once the non-random spectrum \( \sigma(H_{\lambda,\omega}) \) may (and typically will) fill the whole interval \([-2 - \lambda\|V\|_\infty, 2 + \lambda\|V\|_\infty]\) where \( \|V\|_\infty = P - \text{ess sup}|V(\omega)| \) (use approximate eigenfunctions as Weyl sequences in order to show this). For an energy \( 2 + \epsilon\lambda, \epsilon > 0 \), the asymptotics (5.8) then says nothing. However, one can combine the techniques of this chapter with those of the previous one ([72]) in order to prove, as in the case of independent potential values (Section 4.8)

\[
\gamma_\lambda(2 + \epsilon\lambda^\eta) = \sqrt{\epsilon\lambda^\eta} + \mathcal{O}(\lambda^{1-\frac{\eta}{2}}, \lambda^{\frac{2\eta-1}{2}}, \lambda^{\frac{4\alpha+2}{\alpha+2}}), \quad \epsilon > 0, \tag{5.9}
\]

where \( \frac{4}{3} < \eta < \frac{4}{5} \) is such that the error terms are of lower order than \( \lambda^{\frac{2}{3}} \) (in particular, \( \eta = 1 \) is allowed for \( \alpha \) sufficiently large). Moreover, the formulas (5.6) and (5.8) do not imply positivity of the Lyapunov exponent at a fixed \( \lambda \) for all energies in \([2 - \lambda, 2]\) because the error term in (5.6) explodes as one approaches the band edge. However, once again one can transpose Section 4.8 to the case of a strongly mixing potential:

\[
\gamma_\lambda(2 - \epsilon\lambda^\eta) = \lambda^{2-\eta} \frac{D_V(0)}{8\epsilon} + \mathcal{O}(\lambda^{4-\frac{5\eta}{2}}, \lambda^{\frac{9\eta}{2}}, \lambda^{(1-\frac{\eta}{2})\frac{6\alpha+2}{\alpha+2}}), \quad \epsilon > 0, \tag{5.10}
\]
where again $\frac{4}{5} < \eta < \frac{4}{3}$ has to assure that the error terms are subdominant. A careful analysis now allows to show (modulo the issues discussed above) that for $\lambda$ sufficiently small the Lyapunov exponent is positive on $[2-c, 2+c]$ for $c > 0$. We do not provide the detailed argument here, but do claim to have presented all the essential ingredients in order to complete it. Similarly, by analyzing the Lyapunov exponent $\gamma_{\lambda}(e^{\lambda \eta})$, $1 \leq \eta \leq 2$, using the techniques of [77, Section 5.1] or Section 4.5, one can show that the Lyapunov exponent is positive near the band center for $\lambda$ sufficiently small.

Let us now assume that uniform positivity of the Lyapunov exponent has been verified for all energies in the spectrum, either by the above or some other argument, and then deduce localization estimates from this. One standard way to quantify the spreading (delocalization) of an initially localized wave packet $|0\rangle$ under the quantum mechanical time evolution $e^{-itH_{\lambda,\omega}}$ is to consider the growth of (time and disorder averaged) moments of the position operator $X$ on $l^2(\mathbb{Z})$:

$$M^q_T = \int_0^\infty \frac{dt}{T} e^{-\frac{t}{T}} \mathbb{E}_\omega \langle 0 | e^{iH_{\lambda,\omega}t} | X^q e^{-iH_{\lambda,\omega}t} | 0 \rangle, \quad q > 0.$$ (5.11)

Boundedness of $M^q_T$ uniformly in time is called dynamical localization. Logarithmic growth in time as obtained in the following theorem is quite close to that.

**Theorem 5.2** Consider an ergodic family of Jacobi matrices $(H_{\lambda,\omega})_{\omega \in \Omega}$ of the form (5.1) with a quasi-local potential $V$ and an invariant measure $P$ satisfying (5.4) with $\alpha > 0$. Suppose that the spectrum is included in an open interval $(E_0, E_1)$ on which the Lyapunov exponent is uniformly positive:

$$\gamma_{\lambda}(E) \geq \gamma_0 > 0, \quad E \in (E_0, E_1).$$ (5.12)

Then for any $\beta > 2$ there exists a constant $C(\beta, q)$ such that

$$M^q_T \leq (\log T)^{\beta \gamma} + C(\beta, q).$$ (5.13)

Furthermore, the Hausdorff dimension of the spectral measure of $H_{\lambda,\omega}$ vanishes $P$-almost surely.

The elementary proof (fitting on 4-5 pages) of (5.13) is almost completely contained in [44]. It is therefore not reproduced here, but we discuss in detail in Section 5.9 the only step that has to be modified. As already indicated in the introduction, the last statement then follows directly from Guarneri’s inequality [36].

Now follow remarks on when the hypothesis of the above theorems are satisfied.

**Remark 5.1** The strong mixing condition (5.4) clearly holds if $P$ is the product measure of some probability measure on $\Sigma$, because the functions $g_1$ and $g_2$ are then independent. The mixing condition also holds if $P$ stems from a Markov process given by a stochastic kernel having only one invariant measure on a countable set $\Sigma$. Then the decay on the r.h.s. of (5.4) is actually exponential, with rate given by the Perron-Frobenius gap of the stochastic
kernel. Yet more general, let us consider a hyperbolic dynamical system \((X, T)\) (Axiom A) given by a map \(T : X \to X\). Then one has a finite Markov partition \(\Sigma\), with associated symbolic dynamics \((\Omega, S)\) [12], and there is a wealth of so-called Gibbs measures associated to Hölder continuous (i.e. quasi-local) functions which all satisfy (5.4) with an exponential mixing rate [12, Proposition 2.4]. Two standard examples of this type already cited in [17] are the period doubling map and the Arnold cat maps. Moreover, if the phase space \(X\) is a manifold, then any differentiable real function on this manifold gives rise to a quasi-local potential under the coding map. For all these examples with exponential \(\psi\)-mixing, the error bounds in (5.6), (5.7) and (5.8) are given by the error bounds of the independent case [64, 77, 72] multiplied by \(\log(\lambda)\). The error bounds in the independent case are recovered by sending \(\alpha \to \infty\) in (5.6), (5.7), (5.8), (5.9) and (5.10).

**Remark 5.2** Concrete examples of dynamical systems \((X, T)\) having not an exponential, but only a power law decay in (5.4) have only be analyzed more recently. Necessarily \(T\) is then not uniformly hyperbolic, but it is supposed to have only a few parabolic points. Such examples can be constructed even if \(X\) is an interval, but the invariant measure then has a non-normalizable density w.r.t. the Lebesgue measure. It is, however, possible to construct a symbolic dynamics over a countable alphabet \(\Sigma\) which then has a shift-invariant probability measure \(P\) satisfying the strong mixing estimates (5.3) and (5.4). Instead of producing a long citation list, we refer to the references in [34] which contains a proof of (5.4) for several concrete examples. It is precisely in order to deal with these cases at the verge that we bothered to work with (5.4) instead of exponential mixing.

**Remark 5.3** The positivity of the spectral density \(D_V(k)\) can for some examples be checked by an explicit calculation, but there are also further techniques available in order to verify this [13]. The case of \(D_V(0)\) is particularly well studied because of its importance for central limit theorems [63, 34]. For the Gibbs measures of Remark 5.1 and the examples of Remark 5.2, \(D_V(0) = 0\) holds if and only if \(V = v \circ S - v\) is a cocycle given by another quasi-local function \(v\). By suspension, one can deal similarly with \(k = \pi\) and actually any rational \(\frac{k}{2\pi}\).

**Remark 5.4** The above results transpose if \(\mathbb{Z}\) is replaced by \(\mathbb{N}\), namely for \(\Omega = \Sigma^\mathbb{N}\) furnished with the left shift and \(H_{\lambda, \omega}\) acts on \(\ell^2(\mathbb{N})\). As the inverse \(S^{-1}\) of the left shift operator is not defined in that case, one needs to replace in all proofs functions like \(g \circ S^{-n}\) for \(n > 0\) by \((U^*)^n g\), where \(U^*\) is the \(L^2(\Sigma^\mathbb{N}, P)\)-adjoint operator of \(U : g \mapsto g \circ S\).

### 5.3 Anomalies at band center and band edge

Let us begin by recalling that the transfer matrix \(\mathcal{T}_{\lambda, \omega}^E \in \text{SL}(2, \mathbb{R})\) given in (5.5) is elliptic for an energy \(E = 2 \cos(k) \in (-2, 2)\) and \(\lambda = 0\), and it can hence, to zeroth order in \(\lambda\), be transformed into a rotation. More explicitly,

\[
MT_{\lambda, \omega}^E M^{-1} = R_{\tilde{k}} \left( 1 + \lambda \frac{V(\omega)}{\sin(k)} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right),
\]  
(5.14)
where
\[ R_k = \begin{pmatrix} \cos(k) & -\sin(k) \\ \sin(k) & \cos(k) \end{pmatrix}, \quad M = \frac{1}{\sqrt{\sin(k)}} \begin{pmatrix} \sin(k) & 0 \\ -\cos(k) & 1 \end{pmatrix}. \]

In the next section we will consider the action of the matrix (5.14) on the real projective line, which is identified with a circle. To lowest order \( \lambda^0 \), this action induced by (5.14) is then a rotation on the circle. For irrational \( \frac{k}{2\pi} \), there is a unique invariant measure given by the Lebesgue measure. For rational \( \frac{k}{2\pi} = \frac{p}{q} \), at least Birkhoff sums of harmonics of order lower than \( q \) vanish.

At the band center \( k = \frac{\pi}{2} \), the square of the transfer matrix (5.14) (note that \( M = 1 \) here) is the unit matrix and one can only control the lowest order harmonic, which turns out not to be sufficient for the calculation of the Lyapunov exponent. It is then more convenient to consider directly the square of the transfer matrix
\[
T_{\lambda,\omega}^{\lambda^2} T_{\lambda,\omega}^{\lambda^2} = -\left( 1 - \lambda^2 V(\omega)V(S\omega) \epsilon \lambda^2 - \lambda V(\omega) \right) + O(\lambda^3) \tag{5.15}
\]
and to group the coordinates of \( \omega \) in pairs and consider \( \tilde{\Omega} = \tilde{\Sigma}^2 \) where \( \tilde{\Sigma} = \Sigma \times \Sigma \), and furnish it with a probability \( \tilde{P} \) naturally induced by \( P \). Again the suspension \( (\tilde{\Omega}, \tilde{P}) \) is a shift space with power law mixing. However, the matrix (5.15) is now in the form of an anomaly as discussed at the end of this section.

At a band edge, \( e.g. E = -2 \) and \( k = \pi \), the basis change in (5.14) becomes singular and one has a Krein collision. Nevertheless, the transfer matrix at \( \lambda = 0 \) can be transformed into a non diagonalizable Jordan normal form\(^1\):
\[
N_j T_{\lambda,\omega}^{\lambda^2} T_{\lambda,\omega}^{\lambda^2} N_j^{-1} = -\exp \left( \lambda \begin{pmatrix} 0 & -V(S\omega) \\ V(\omega) & 0 \end{pmatrix} + \lambda \begin{pmatrix} 0 & 2\epsilon \\ -2\epsilon & 0 \end{pmatrix} + O(\lambda^3) \right) \tag{5.16}
\]

Resuming, after adequate basis change and possibly regrouping of terms, one has to study in each of the three situations (5.14), (5.15) and (5.16) families of random matrices \( (T_{\lambda,\omega})_{\lambda \geq 0, \omega \in \Omega} \subset Q_r(\text{SL}(2, \mathbb{R})) \) of the following form:
\[
T_{\lambda,\omega} = \pm R_k \exp \left( \lambda^2 P_{1,\omega} + \lambda^{3\eta} P_{2,\omega} + O(\lambda^{3\eta}) \right) \tag{5.17}
\]
\(^1\)The corresponding basis change is denoted by \( N_j \), the index is used to distinguish this matrix from \( N \) used as a natural number.
where \( \eta > 0 \), \( P_{j,\omega} \in Q_{\tau}(\text{sl}(2, \mathbb{R})) \) for \( j = 1, 2 \), \( \mathbf{E}(P_{1,\omega}) = 0 \) and the error term \( \mathcal{O}(\lambda^{3\eta}) \) is uniformly bounded \( \text{(i.e. the bound is } \omega\text{-independent)} \). If \( k = 0, \pi \), namely at a band center (5.15) and a band edge (5.16), such a family is said to have an anomaly of second order [77, 72] as defined in the previous chapter. In the following sections, we treat general families of the form (5.17), and then go back to the explicit cases in Section 5.8 in order to complete the proof of Theorem 5.1.

### 5.4 Phase shift dynamics

The bijective action \( S_T \) of a matrix \( T \in \text{SL}(2, \mathbb{R}) \) on \( S^1_\pi = \mathbb{R}/\pi\mathbb{Z} = [0, \pi) \) is given by

\[
e_{S_T}(\theta) = \pm \frac{Te_\theta}{\|Te_\theta\|}, \quad e_\theta = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}, \quad \theta \in S^1_\pi,
\]

(5.18)

with an adequate choice of the sign. This defines a group action, namely \( S_{TT'} = S_T S_{T'} \). In order to shorten notations, we write \( S_{\lambda,\omega} = S_{T_{\lambda,\omega}} \) where \( T_{\lambda,\omega} \) is of the form (5.17). One thus has \( S_{\lambda,\omega}(\theta) = \theta + k + \mathcal{O}(\lambda) \).

Given an initial angle \( \theta_0 \) and iterating this dynamics by the left shift on \( \Omega \) defines a stochastic process \( \theta_n(\omega) \), also simply denoted by \( \theta_n \) below:

\[
\theta_0(\omega) = \theta_0, \quad \theta_{n+1}(\omega) = S_{\lambda,S^n}(\theta_n(\omega)).
\]

(5.19)

As in the last chapter, let us introduce for \( j = 1, 2 \) the trigonometric polynomials

\[
p_{j,\omega}(\theta) = \Im \left( \frac{\langle v| P_{j,\omega} |e_\theta \rangle}{\langle v| e_\theta \rangle} \right), \quad v = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}.
\]

(5.20)

By similar calculations as done in Section 4.4 one has

\[
S_{\lambda,\omega}(\theta) = \theta + k + \sum_{j=1}^{2} \lambda^{j\eta} p_{j,\omega}(\theta) + \frac{1}{2} \lambda^{2\eta} p_{1,\omega} \partial_\theta p_{1,\omega}(\theta) + \mathcal{O}(\lambda^{3\eta}).
\]

(5.21)

Due to Lemma 3 of [45] and a telescoping argument, the Lyapunov exponent \( \gamma(\lambda) \) characterizing the exponential growth of the products of matrices in the ergodic family \( (T_{\lambda,S^n})_{n \geq 0} \) is given by

\[
\gamma(\lambda) = \lim_{N \to \infty} \frac{1}{N} \mathbf{E}_{\theta_0} \mathbf{E}_\omega \sum_{n=0}^{N-1} \log(||T_{\lambda,S^n}e^{\theta_n(\omega)}||),
\]

(5.22)

where \( \mathbf{E}_{\theta_0} \) denotes an average over the initial condition \( \theta_0 \) w.r.t. an arbitrary continuous probability measure on \( S^1_\pi \). As our interest is perturbation theory for \( \gamma(\lambda) \) w.r.t. \( \lambda \), we shall need the following expansions for the summands of (5.22) obtained similar to Lemma 4.1.
5.5. FROM BIRKHOFF-LIKE SUMS TO BIRKHOFF SUMS

Lemma 5.1 Set
\[\alpha_{j,\omega} = \langle v | P_{j,\omega} | v \rangle, \quad \beta_{j,\omega} = \langle v | P_{j,\omega}^2 | v \rangle, \quad \gamma_{j,\omega} = \langle v | P_{j,\omega} | v \rangle.\]
Then \(p_{j,\omega}(\theta) = \Im m(\alpha_{j,\omega} - \beta_{j,\omega} e^{2i\theta}).\) Furthermore,
\[
\log(\|T_{\lambda,\omega} e^{\theta}\|) = \Re e \left( 2 \sum_{j=1}^{2} \lambda^j \beta_{j,\omega} e^{2i\theta} + \frac{\lambda^{2n}}{2} (|\beta_{1,\omega}|^2 + \gamma_{1,\omega} e^{2i\theta} - \beta_{1,\omega}^2 e^{4i\theta}) \right) + O(\lambda^{3n}).
\]
(5.23)

Formula (5.22) and also its perturbative evaluation based on (5.23) hence leads us to consider sums of the type
\[
\hat{I}_N(G) = \frac{1}{N} \mathbb{E}_{\omega} \sum_{n=0}^{N-1} G(S_n \omega, \theta_n(\omega)) , \quad \hat{I}(G) = \lim_{N \to \infty} \hat{I}_N(G),
\]
(5.24)
for functions \(G\) on \(\Omega \times S_1^1\) of the type \(G(\omega,\theta) = \sum_j g_j(\omega) f(\theta)\). More explicitly, the above lemma shows that one only needs functions of the form \(g(\omega) e^{2i\theta}\) and \(g(\omega) e^{4i\theta}\) with \(g \in \mathcal{Q}_r(\mathbb{C})\). For a \(\pi\)-periodic function \(f \in C(S_1^1)\), we also introduce
\[
I_N(f) = \frac{1}{N} \mathbb{E} \sum_{n=0}^{N-1} f(\theta_n), \quad I(f) = \lim_{N \to \infty} I_N(f),
\]
This is a Birkhoff sum of the process \(\theta_n = \theta_n(\omega)\). In the sum (5.24) there is, moreover, an explicit dependence of \(G\) on \(\omega\), hence let us use the term Birkhoff-like sums for the sums \(\hat{I}(G)\). The defined limits may not exist but with the methods used one can obtain estimates for certain finite sums and then work with \(\lim \sup\) and \(\lim \inf\). But this only makes the writing more technical and more complicated.

5.5 From Birkhoff-like sums to Birkhoff sums

The aim of this section is, as indicated in the title, to reduce the perturbative evaluation of the Birkhoff-like sums (5.24) to the evaluation of Birkhoff sums by invoking the correlation decay (5.4).

Proposition 5.1 Suppose \(\alpha > 2\) and \(k = 0\). Let \(g \in \mathcal{Q}_r(\mathbb{C})\) and \(f \in C^2(S_1^1)\). Define \(G(\omega,\theta) = g(\omega) f(\theta)\). Then
\[
\hat{I}(G) = \mathbb{E}(g) I(f) + O(\lambda^{\frac{2n}{\alpha}}).
\]
(5.25)
If \(\mathbb{E}(g) = 0\), one has the following convergent expression for the next higher order contribution:
\[
\hat{I}(G) = \lambda^j \sum_{j=1}^{\infty} I(f_j) + O(\lambda^{\frac{2n}{\alpha}}), \quad f_j(\theta) = \mathbb{E}_{\omega}(g(S_j \omega)p_{1,\omega}(\theta)) f'(\theta).
\]
(5.26)
Lemma 5.2

One has for \( m, n \geq 1 \)

\[
\text{Var}_{[-m-n,n+m]}(\theta) \leq \mathcal{O}(r^m \lambda^n) .
\]

Proof. Using equation (5.19),

\[
|\theta_{n+1}(\omega) - \theta_{n+1}(\omega')| \leq |S_{\lambda,S_n^\omega}(\theta_n(\omega)) - S_{\lambda,S_n^\omega}(\theta_n(\omega'))| + |S_{\lambda,S_n^\omega}(\theta_n(\omega')) - S_{\lambda,S_n^\omega}(\theta_n(\omega'))|
\]

one deduces

\[
|\theta_{n+1}(\omega) - \theta_{n+1}(\omega')| \leq \left( \sup_{\omega,\theta} |S_{\lambda,\omega}(\theta)| \right) |\theta_n(\omega) - \theta_n(\omega')| + \sup_{\theta} |S_{\lambda,S_n^\omega}(\theta) - S_{\lambda,S_n^\omega}(\theta)| .
\]

Using the estimate

\[
\left\| \frac{x - x'}{\|x\|} - \frac{x' - x'}{\|x'\|} \right\| = \left\| \frac{x - x'}{\|x\|} + x' \left( \frac{\|x'\| - \|x\|}{\|x\| \|x'\|} \right) \right\| \leq \frac{2}{\|x\|} \|x - x'\|
\]

and the definition of \( S_{\lambda,\omega} \), it follows

\[
\left\| e_{S_{\lambda,\omega}(\theta)} - e_{S_{\lambda,\omega'}(\theta)} \right\| \leq \frac{2}{\|T_{\lambda,\omega} e_\theta\|} \|T_{\lambda,\omega} - T_{\lambda,\omega'}\| \leq 2 \left( \sup_{\omega} \|T_{\lambda,\omega}^{-1}\| \right) \|T_{\lambda,\omega} - T_{\lambda,\omega'}\| .
\]

This implies

\[
\sup_{\theta} |S_{\lambda,S_n^\omega}(\theta) - S_{\lambda,S_n^\omega'}(\theta)| \leq C_1 \lambda^n \|Q_{\lambda,S_n^\omega} - Q_{\lambda,S_n^\omega'}\| \tag{5.28}
\]

where \( C_1 \) is a constant and \( T_{\lambda,\omega} = 1 + \lambda^\theta Q_{\lambda,\omega} \) for a matrix-valued function \( Q_{\lambda,\omega} \) that is analytic in \( \lambda^n \) and uniformly quasi-local for small \( \lambda \) (i.e. the constant and rate is \( \lambda \)-independent). Furthermore, one has

\[
\sup_{\omega,\theta} |S_{\lambda,\omega}'(\theta)| \leq 1 + C_2 \lambda^n
\]

for \( \lambda \) sufficiently small and some constant \( C_2 \). Applying this and (5.28) to (5.27) one gets

\[
\text{Var}_I(\theta_{n+1}) \leq (1 + C_2 \lambda^n) \text{Var}_I(\theta_n) + C_1 \lambda^n \text{Var}_I(Q_{\lambda,S_n^\omega}) .
\]

Iterating this estimate and using \( \text{Var}_I(\theta_0) = 0 \), it follows that

\[
\text{Var}_{[-m-n,n+m]}(\theta_n) \leq \sum_{j=1}^{n} (1 + C_2 \lambda^n)^{j-1} C_1 \lambda^n \text{Var}_{[-m-n,n+m]}(Q_{\lambda,S_{n-j}^\omega})
\]

\[
\leq C_1 \lambda^n C_3 r^{m+1} \sum_{j=0}^{\infty} (1 + C_2 \lambda^n)^j r^j = \mathcal{O}(\lambda^n r^m)
\]
for \( \lambda \) sufficiently small.

In order to state the next two lemmata, we introduce the following notation extending (5.24):

\[
\hat{I}_N^m(G) = \frac{1}{N} E \sum_{n=0}^{N-1} G(S^{m+n}\omega, \theta_n(\omega)), \quad \hat{I}^m(G) = \lim_{N \to \infty} I_N^m(G).
\]

**Lemma 5.3** Let \( g_1, g_2 \in \mathcal{Q}_r(\mathbb{C}) \) and \( f \in \mathcal{C}^1(S^1_+) \). Furthermore let \( k \geq l \geq 0 \) and \( m \geq 1 \). Then

\[
E_\omega(g_1(S^{3m+k+n}\omega) g_2(S^{3m+l+n}\omega) f(\theta_n(\omega))) = E(f(\theta_n)) E(g_1 \circ S^{k-l} g_2) + O(m^{-\alpha}),
\]

(5.29)

uniformly in \( k, l \) and \( n \). This implies, for \( G(\omega, \theta) = g_1(S^k\omega) g_2(S^l\omega)f(\theta) \),

\[
\hat{I}^m(G) = E((g_1 \circ S^{k-l} g_2) I(f)) + O(m^{-\alpha}).
\]

(5.30)

**Proof.** By Lemma 5.2 and because \( f \) is Lipshitz-continuous, one has uniformly in \( n \)

\[
|f(\theta_n(\omega)) - f(\theta_n(\pi[-m-n+m](\omega)))| \leq O(\lambda^n r^m).
\]

As \( g_1 \) and \( g_2 \) are quasi-local and therefore bounded, one deduces uniformly in \( k, n \) and \( l \)

\[
|g_1(S^{k+n+m}\omega) g_2(S^{l+m}\omega) - ((g_1 \circ S^k)(g_2 \circ S^l)) \circ S^{m+n} \circ \pi_{n+k+4m}(\omega)| \leq O(r^m).
\]

Let us denote the two functions inside the modulus by \( g \) and \( \hat{g} \) respectively. Similarly denote \( f \circ \theta_n \circ \pi_{n+k+4m} \) by \( \hat{f} \). Now consider \( E(g f(\theta_n)) \). As the functions \( f \) and \( g \) are bounded, it follows from the estimates above and (5.4) that with errors of order \( O = O(m^{-\alpha}) \geq O(r^m) \geq O(\lambda^n r^m) \) (for big \( m \) and small \( \lambda \)) in each step we get

\[
E(g f(\theta_n)) = E(\hat{g} \hat{f}(\theta_n)) + O = E(\hat{g}) E(\hat{f}) + O = E(g) E(f(\theta_n)) + O.
\]

This finishes the proof. \( \square \)

Replacing \( g_2(S^{3m+n+l}\omega) \) by \( g_2(S^{l+n}\omega) \) for \( 0 \leq l \leq k \), one can modify the argument by grouping \( g_2 \) and \( f \) together. This gives the following

**Lemma 5.4** Let \( g_1, g_2 \in \mathcal{Q}_r(\mathbb{C}) \) and let \( f \in \mathcal{C}^1(S^1_+) \). Then one has for \( 0 \leq l \leq k \) and \( m \geq 1 \)

\[
E_\omega(g_1(S^{3m+k+n}\omega) g_2(S^{l+n}\omega) f(\theta_n(\omega))) = E(g_1) E_\omega(g_2(S^{l+n}\omega)) f(\theta_n(\omega)) + O(m^{-\alpha}),
\]

uniformly in \( l, k \) and \( n \). This implies for \( G(\omega, \theta) = g_1(S^{3m+k}\omega) g_2(S^l(\omega)) f(\theta) \)

\[
\hat{I}(G) = E(g_1) \hat{I}(g_2(S^l(\omega)) f(\theta)) + O(m^{-\alpha}),
\]

(5.31)

and leads, for \( f = 1 \) and \( l = 0 \), to

\[
E(g_1(S^{3m+k}(\omega)) g_2(\omega)) = E(g_1) E(g_2) + O(m^{-\alpha}).
\]

(5.32)
Proof of Proposition 5.1. By Taylor expansions and 
P_1, S^{n+j+\omega}(\theta_n) = P_1, S^{n+j+\omega}(\theta_n) + O(j^\eta),
one finds
\[ f(\theta_n + 6m) = f(\theta_n) + \lambda^\eta \sum_{j=0}^{6m-1} P_1, S^{n+j+\omega}(\theta_n) f'(\theta_n) + O(m^2 \lambda^{2\eta}). \]
Therefore multiplying with \( g \circ S^{6m+n} \) and averaging over \( \omega \) and \( n \) gives
\[ \hat{I}(\mathcal{G}) = \hat{I}_{6m}(\mathcal{G}) + \lambda^\eta \sum_{j=0}^{6m-1} \hat{I}(\mathcal{G}_j) + O(m^2 \lambda^{2\eta}), \]
where \( \mathcal{G}_j(\omega, \theta) = g(S^{6m+j}\omega)p_1, S^j(\theta) f'(\theta) \). As \( p_1, \omega(\theta) \) is a trigonometric polynomial in \( \theta \), Lemma 5.3 can be applied to each summand in order to obtain
\[ \hat{I}_{6m}(\mathcal{G}) = E(g) I(f) + O(m^{-\alpha}). \]
Because the functions \( \mathcal{G}_j \) are uniformly bounded, one has \( \lambda^\eta \sum_{j=0}^{6m-1} \hat{I}(\mathcal{G}_j) = O(m^\lambda) \). Using \( m = \lambda^{-\eta \alpha + 2} \) now proves the first part.

Now let \( E(g) = 0 \). Again because \( p_1, \omega \) is a trigonometric polynomial, Lemma 5.3 gives, for \( j \geq 3m \) and \( f_j \) as defined in (5.26),
\[ \hat{I}(\mathcal{G}_j) = I(E_\omega(g(S^{6m-j}\omega)p_1, \omega(\theta)) f') + O(m^{-\alpha}) = I(f_{6m-j}) + O(m^{-\alpha}). \]
Using Lemma 5.4, one obtains for \( j < 3m \)
\[ \hat{I}(\mathcal{G}_j) = E(g) \hat{I}(p_1, \omega f'(\theta)) + O(m^{-\alpha}) = O(m^{-\alpha}). \]
All together, one has
\[ \hat{I}(\mathcal{G}) = \lambda^\eta \sum_{j=3m}^{6m-1} I(f_{6m-j}) + O(m^2 \lambda^{2\eta}, \lambda^\eta m^{1-\alpha}, m^{-\alpha}). \]
Because (5.32) gives
\[ \sum_{j=3m+1}^{\infty} |f_j(\theta)| = \sum_{j=3m+1}^{\infty} |E_\omega(g(S^j \omega)p_1, \omega(\theta)) f'(\theta)| \leq C \sum_{j=3m+1}^{\infty} j^{-\alpha} = O(m^{1-\alpha}), \]
one therefore deduces
\[ \hat{I}(\mathcal{G}) = \lambda^\eta \sum_{j=1}^{\infty} I(f_j) + O(m^2 \lambda^{2\eta}, \lambda^\eta m^{1-\alpha}, m^{-\alpha}). \]
Finally choosing \( m = \lambda^{-\eta \alpha + 2} \) concludes the proof.
5.6 Oscillatory sums away from band center and edges

As already explained in Section 5.4, for the calculation of the Lyapunov exponent one needs to evaluate the Birkhoff-like sums of functions of the type \( G(\omega, \theta) = g(\omega)e^{2ij\theta} \), \( j = 1, 2 \). This is done in Proposition 5.2 below for energies away from the band center and band edge. By applying it to the terms appearing when (5.23) is replaced in (5.22), this result allows to complete the proof of formula (5.6). As the straightforward algebraic calculations are carried out in detail e.g. in [17, 45] and we present a similar calculation for the band edge in Section 5.8, we skip the details.

Proposition 5.2 Let \( \alpha > 2 \). Suppose that the lowest order rotation phase \( k \) in the dynamics (5.21) satisfies \( d(k) = \text{dist}(k \mod \frac{\pi}{2}, 0) > 0 \). Consider \( G_j(\omega, \theta) = g(\omega)e^{2j\theta} \) with \( j = 1, 2 \) and \( g \in \mathcal{Q}_r(\mathbb{R}) \). Then

\[
\hat{I}(G_j) = \mathcal{O}\left( \frac{\lambda^{\frac{\alpha}{\alpha - 2}}}{d(k)} \right).
\]

If, moreover, \( E(g) = 0 \),

\[
\hat{I}(G_1) = \lambda^{\eta} \sum_{j=1}^{\infty} \mathbb{E}_\omega \left( g(S^j\omega), \beta_{1,\omega} \right) + \mathcal{O}\left( \frac{\lambda^{\frac{2\eta}{\alpha - 2}}}{d(k)} \right).
\]

(5.33)

Proof. [64, 17, 45] The dynamics and the definition of the Birkhoff sums implies that \( I_N(e^{2i\theta}) = e^{2ijk}I_N(e^{2i\theta}) + \mathcal{O}(N^{-1}, \lambda^\eta) \) and hence \( I(e^{2i\theta}) = \mathcal{O}(d(k)^{-1}\lambda^\eta) \). Therefore the modifications of (5.25) and (5.26) in Proposition 5.1 mentioned in the remark are irrelevant. The bound (5.25) thus implies the first statement. The formula (5.33) now follows after a short calculation from (5.26), the identity \( p_{1,\omega}(\theta) = \Im(\alpha_{1,\omega} - \beta_{1,\omega}e^{2i\theta}) \) and the first statement. \( \square \)

5.7 Fokker-Planck operator for drift-diffusion

We now focus on energies for which the rotation angle \( k \) in (5.21) satisfies \( k \mod \frac{\pi}{2} = 0 \) so that the argument of Proposition 5.2 does not apply in order to calculate the Birkhoff sum \( I(e^{2i\theta}) \). For this purpose, let us introduce the bilinear form

\[
\langle g_1, g_2 \rangle_\Omega = \mathbb{E}_\omega (g_1(\omega)g_2(\omega)) + 2\sum_{m=1}^{\infty} \mathbb{E}_\omega (g_1(\omega)g_2(S^m\omega)), \quad g_1, g_2 \in \mathcal{Q}_r^0(\mathbb{R}),
\]

which by (5.32) is well-defined. Note that \( D_V(0) = \langle V, V \rangle_\Omega \). Let us use the notation \( p_j(\omega, \theta) = p_{j,\omega}(\theta) \) and \( p'_j = \partial_\theta p_j \). Then expressions like \( \langle p_1, p'_1 \rangle_\Omega \) are functions of \( \theta \) on \( S^1_\pi \).

Proposition 5.3 Let the family \( T_{\lambda,\omega} \) be as in (5.17) with \( k = 0 \), and \( F \in C^3(S^1_\pi) \). For \( f \in C^1(S^1_\pi) \) given by

\[
f = \langle p_1, p_1 \rangle_\Omega F'' + (\langle p_1, p'_1 \rangle_\Omega + 2 \mathbb{E}(p_2,\omega)) F',
\]

(5.34)

one then has for \( \alpha > 2 \)

\[
I(f) = \mathcal{O}\left( \lambda^{\frac{\alpha}{\alpha - 2}} \right).
\]
Proof. By a Taylor expansion, one has with errors of order $O(\lambda^{3q})$

$$F(S_{\lambda,\omega}(\theta)) = F(\theta) + \sum_{k=1}^{2} \lambda^{kn} p_k(\theta) F'(\theta) + \lambda^{2n} \frac{1}{2} \left[ F'(\theta) p_{1,\omega}(\theta) p'_{1,\omega}(\theta) + p_{1,\omega}^2(\theta) F''(\theta) \right].$$

We now use this for $\theta = \theta_n$ and average over $n$. Because $p_{1,\omega}$ is centered and a polynomial, one can apply equation (5.26) of Proposition 5.1 to the term with power $\lambda^q$ and (5.25) to the other terms. This gives

$$I(F) = I(F) + \frac{1}{2} \lambda^{2n} \left( I(\langle p_1, p'_1 \rangle_{\Omega} F') + I(\langle p_1, p_1 \rangle_{\Omega} F'') + 2 I(\mathbb{E}_{\omega}(p_{2,\omega}) F') \right) + O$$

with errors of order $O(\lambda^{3q+2})$. As the functional $I$ is linear, resolving this equation for $I(f)$ gives the desired result. \hfill \Box

This proposition shows, that we can control error terms on Birkhoff sums for a function $f$, if $f$ is in the image of the operator $\mathcal{L}$ on functions on $S_\pi^1$ given by

$$\mathcal{L} = (p \partial_\theta + q) \partial_\theta, \quad p = \langle p_1, p_1 \rangle_{\Omega}, \quad q = \langle p_1, p'_1 \rangle_{\Omega} + 2 \mathbb{E}(p_{2,\omega}). \quad (5.35)$$

As one needs to calculate Birkhoff sums $I(f)$ perturbatively, we are looking for some class of functions where $\lim_{\lambda \to 0} I(f)$ exists. For $f$ in the image under $\mathcal{L}$ of $C^3(S_\pi^1)$, this limit is 0. Thus, if this map is given by the scalar product with some $L^2$-function $\rho$, one has $\rho \in \text{Ran}(\mathcal{L})^\perp = \text{Ker}(\mathcal{L}^*)$, where the formal adjoint is given by

$$\mathcal{L}^* = \partial_\theta (\partial_{\theta p} - q).$$

$\mathcal{L}^*$ is a forward Kolmogorov or Fokker-Planck operator describing the drift-diffusion dynamics of the process $\theta_n$ on $S_\pi^1$, and $\mathcal{L}$ is the associated backward Kolmogorov operator [67]. It will be shown that in the situations considered here, $\text{Ker}(\mathcal{L}^*)$ is spanned by a smooth, $L^1$-normalized function $\rho$. Furthermore, the following theorem shows that $f \in \text{Ker}(\mathcal{L}^*)^\perp \cap C^2(S_\pi^1)$ turns out to be sufficient for finding a solution $F \in C^3(S_\pi^1)$ of the differential equation (5.34) so that Proposition 5.3 actually applies. Even though contained in the last chapter and [72], let us give the proof for sake of completeness.

Theorem 5.3 Suppose that $p(\hat{\theta}) = 0$ for at most one angle $\hat{\theta} \in S_\pi^1$. Furthermore suppose $q(\hat{\theta}) \neq 0$ in that case. Then the Fokker-Planck operator $\mathcal{L}^*$ has a unique groundstate $\rho \in C^\infty(S_\pi^1)$, which is non-negative and normalized. Furthermore, for $f \in C^2(S_\pi^1)$, one has

$$I(f) = \int_0^\pi d\theta \rho(\theta) f(\theta) + O(\lambda^{3q+2}).$$

Proof. Integrating the equation $\mathcal{L}^* \rho = 0$ once gives

$$(p \partial_\theta + (\partial_\theta p) - q) \rho = C, \quad (5.36)$$
where $C$ is some real constant. As $I(f + c) = c + I(f)$ for $c = \langle \rho, f \rangle$, we may assume \( \int_{\theta}^{\pi} d\theta \rho(\theta) f(\theta) = 0 \) once we found the normalized solution of (5.36). Proposition 5.3 then gives the bound on $I(f)$ if one finds a solution $G \in C^{2}(S^{1}_{\pi})$ of

\[
(p\partial_{\theta} + q)G = f, \quad \int_{0}^{\pi} d\theta G(\theta) = 0.
\]

First let us consider the case $p > 0$. Then there is no singularity and $L^{*}$ is elliptic. The groundstate $\rho$ and the function $G$ can be calculated. For some $\hat{\theta}$ set

\[
w(\theta) = \int_{\hat{\theta}}^{\theta} d\xi \frac{q(\xi)}{p(\xi)}, \quad W(\theta) = \int_{\hat{\theta}}^{\theta} d\xi \frac{e^{w(\xi)}}{p(\xi)} f(\xi), \quad \hat{W}(\theta) = \int_{\hat{\theta}}^{\theta} d\xi e^{-w(\xi)}.
\]

Then

\[
\rho = C_{1} e^{w} \left( C_{2} \hat{W} + 1 \right), \quad G = e^{-w} \left( W + C_{3} \right),
\]

where $C_{2}$ is fixed by the condition that $\rho$ is $\pi$-periodic and $C_{1} > 0$ is a normalization constant. This fixes $C = C_{1}C_{2}$ in (5.36). $G$ is a solution of the first equation of (5.37) and for $C \neq 0$ the constant $C_{3}$ is fixed by the condition that $G$ is $\pi$-periodic. Furthermore one has

\[
0 = \int \rho f = \int \rho \left( p\partial_{\theta} + q \right) G = -\int G \left( \partial_{\theta} p - q \right) \rho = -C \int G(\theta).
\]

Thus $G$ is a solution of (5.37). If $C = 0 \iff C_{2} = 0$, then $w$ is $\pi$-periodic as well as $W$ which follows from $\int \rho f = 0$. Therefore $G$ is $\pi$-periodic and $C_{3}$ is chosen such that the integral in (5.37) vanishes.

Now let $p(\hat{\theta}) = 0$ for exactly one $\hat{\theta} \in S^{1}_{\pi}$ and for sake of concreteness let $q(\hat{\theta}) > 0$ which implies $\hat{q}(\hat{\theta}) > 0$. Then choose $\bar{\theta} \in (\hat{\theta}, \hat{\theta} + \pi)$ in the first equation of (5.38), $\theta = \bar{\theta}$ in the second one and $\theta = \hat{\theta} + \pi$ in the third one. As $\lim_{\theta \to \hat{\theta}} e^{w(\theta)} = 0$ and $\lim_{\theta \to \hat{\theta} + \pi} e^{w(\theta)} = \infty$ in this case, $w$, $W$ and $\hat{W}$ are well-defined for $\theta \in (\hat{\theta}, \hat{\theta} + \pi)$. Using de l’Hospital’s rule, one can prove by induction (see Appendix A.1 for details) that

\[
\rho = C e^{w} \hat{W}, \quad G = e^{-w} W,
\]

can both be continued to a smooth (even at $\hat{\theta}$) and $\pi$-periodic function. $C > 0$ is again a normalization constant and hence equation (5.40) shows that $G$ solves (5.37).

Before applying this result in order to prove Theorem 5.1, let us present another derivation of the equation $L^{*}\rho = 0$, albeit a formal one, which shows that $\rho$ is the lowest order approximation for the asymptotic invariant measure of the process $\theta_{n}$. Expanding the function $S_{\lambda,\omega}^{N} = \mathcal{S}_{\lambda,\omega}^{N} \circ \ldots \circ \mathcal{S}_{\lambda,\omega} \circ \mathcal{S}_{\lambda,\omega}$ shows that the coefficients of

\[
\mathcal{S}_{\lambda,\omega}^{N}(\theta) = \theta + \lambda^{n} \mathcal{P}_{n}^{N}(\theta) + \frac{1}{2} \lambda^{2n} \mathcal{F}_{n}^{N}(\theta) + \mathcal{O}(\lambda^{3n}),
\]
are
\[ \hat{p}_\omega^N = \sum_{n=0}^{N-1} p_{1,s^n\omega}, \quad \hat{q}_\omega^N = \sum_{n=0}^{N-1} \left( p_{1,s^n\omega} + \sum_{j=0}^{n-1} p_{1,s^j\omega} \right) p'_{1,s^n\omega} + 2 \sum_{n=0}^{N-1} p_{2,s^n\omega}. \]

An invariant measure \( \nu_{\lambda,N} \) for \( N \) steps of the dynamics \( \theta_n \) on \( S_1^\pi \) satisfies
\[ \int_0^n \nu_{\lambda,N}(d\theta) f(\theta) = E \int_0^n \nu_{\lambda,N}(d\theta) f(S_{\lambda,S^{-n}\omega}(\theta)), \quad f \in C(S_1^\pi). \] (5.41)

Supposing \( \nu_{\lambda,N}(d\theta) = \rho_{\lambda,N}(\theta) d\theta + o(\lambda^0) \), (5.41) leads to
\[ \mathcal{L}_N^* \rho_{\lambda} = 0, \quad \mathcal{L}_N^* = \partial_\theta \left( \partial_\theta E((\hat{p}_{\omega}^N)^2) - E(\hat{q}_\omega^N) \right). \]

Using the stationarity of \( P \) and the definitions of \( \hat{p}_\omega^N \) and \( \hat{q}_\omega^N \), one deduces
\[ \lim_{N \to \infty} \frac{1}{N} E((\hat{p}_{\omega}^N)^2) = p, \quad \lim_{N \to \infty} \frac{1}{N} E(\hat{q}_\omega^N) = q, \]
where the convergences are uniform in \( \theta \). This shows that \( \frac{1}{N} \mathcal{L}_N^* \to \mathcal{L}^* \) weakly for \( N \to \infty \).

### 5.8 Application to the band center and band edge

This section contains the proof of Theorem 5.1. Let us first consider item (i), that is the band center. As described in Section 5.3 we have to work with the probability space \( \tilde{\Omega} = (\Sigma \times \Sigma)^2 \) which is isomorphic to \( \Omega \) by the pairing isomorphism \( \mathcal{P} \). Using this isomorphism and the potential \( V \), which is defined on \( \Omega \), let us define the two random variables on \( \tilde{\Omega} \)
\[ v_{\tilde{\omega}} = V(\mathcal{P}^{-1}(\bar{\omega})) = V(\omega), \quad u_{\tilde{\omega}} = V(S\mathcal{P}^{-1}(\bar{\omega})) = V(S\omega). \]

Then according to equation (5.15) the family of matrices we have to consider is given by
\[ T_{\lambda,\tilde{\omega}} = -\exp \left[ \lambda \begin{pmatrix} 0 & -u_{\tilde{\omega}} \\ v_{\tilde{\omega}} & 0 \end{pmatrix} + \frac{\lambda^2}{2} \begin{pmatrix} -u_{\tilde{\omega}}v_{\tilde{\omega}} & 2\epsilon \\ -2\epsilon & u_{\tilde{\omega}}v_{\tilde{\omega}} \end{pmatrix} + O(\lambda^3) \right]. \]

In this situation one has \( \alpha_{1,\tilde{\omega}} = \nu(v_{\tilde{\omega}} + u_{\tilde{\omega}})/2 \), \( \beta_{1,\tilde{\omega}} = \nu(u_{\tilde{\omega}} - v_{\tilde{\omega}})/2 \), \( \alpha_{2,\tilde{\omega}} = -\nu \) and \( \beta_{2,\tilde{\omega}} = -\frac{1}{2}u_{\tilde{\omega}}v_{\tilde{\omega}} \). Using Lemma 5.1 and \( \langle v - u, v - u \rangle_\Omega = 2D_V(\pi) \) and \( \langle v + u, v + u \rangle_\Omega = 2D_V(0) \), one obtains that the polynomials (5.35) are explicitly given by
\[ p(\theta) = \frac{1}{2} D_V(0) + \frac{1}{2} D_V(\pi) \cos^2(2\theta), \quad q(\theta) = -\frac{1}{2} D_V(\pi) \sin(4\theta) - \epsilon. \]

By assumption on \( V \), one has \( p > 0 \) uniformly on \( S_1^\pi \). By Theorem 5.3 there is thus a smooth, positive and \( L^1 \)-normalized groundstate \( \rho_{\epsilon} \) for the operator \( \mathcal{L}^* \) (which can readily be written out). Furthermore, one checks \( \gamma_{1,\tilde{\omega}} = (v_{\tilde{\omega}}^2 - u_{\tilde{\omega}}^2)/2 \). Then equation (5.23), Theorem 5.3 and Proposition 5.1 combined with some algebra leads to (5.7) for \( \gamma_\lambda(\epsilon \lambda^2) = \frac{1}{2} \gamma(\lambda) \).
5.9. BOUND ON THE QUANTUM DYNAMICS

Now let us prove Theorem 5.1(ii). Hence let \( T_{\lambda, \omega} = N_\lambda N_j T_{\lambda, \omega}^{-2+\epsilon^2} N_j^{-1} N_\lambda^{-1} \) be the anomaly given in (5.16). As \( \alpha_{1,\omega} = iV(\omega)/2 \), \( \beta_{1,\omega} = -iV(\omega)/2 \), \( \alpha_{2,\omega} = -i(\epsilon-1)/2 \), and \( \beta_{2,\omega} = i(\epsilon-1)/2 \), one deduces, using \( \langle V, V \rangle_\Omega = D_V(0) \),

\[
p(\theta) = D_V(0) \cos^4(\theta), \quad q(\theta) = -\epsilon + 1 + (1 - \epsilon) \cos(2\theta) - 2 D_V(0) \cos^3(\theta) \sin(\theta).
\]

By assumption on \( V \) one has \( p(\theta) > 0 \) for \( \theta \not\in \frac{\pi}{2} \), and as \( q(\frac{\pi}{2}) = -2 \neq 0 \), there is a unique groundstate \( \rho_\epsilon \in C^\infty(S^1_\epsilon) \) by Theorem 5.3. Explicitly, one obtains

\[
\rho_\epsilon(\theta) = C \int_{-\frac{\pi}{2}}^{\theta} d\xi \frac{\cos^2(\xi)}{\cos^6(\theta)} \exp\left( \frac{2}{3 D_V(0)} \left( \tan^3(\xi) - \tan^3(\theta) + 3\epsilon \tan(\xi) - 3\epsilon \tan(\theta) \right) \right),
\]

(5.42)

where \( C \) is some normalization constant. Furthermore, one checks \( \gamma_{1,\omega} = V(\omega)^2/2 \) and hence (5.23), Proposition 5.1 and Theorem 5.3 imply (5.8).

5.9 Bound on the quantum dynamics

As already said above, the proof of Theorem 5.2 follows exactly the proof of Theorem 1 in [44] given in Section 3 and 4 therein, except that the proof of Lemma 4 of [44] has to be refined in order to deal with strong mixing (5.4) instead of independent potential values \( V(S^\omega) \). The conclusion of the following lemma is hence exactly the same as of Lemma 4 of [44], and we thereby consider the proof of Theorem 5.2 to be complete.

Let us set \( U = \{ E \in C \mid E_0 \leq \Re(e(E)) \leq E_1, \ |\Im(m(E))| \leq 1 \} \). Furthermore introduce the transfer matrices over several sites:

\[
T_{\lambda, \omega}^E(k, m) = \prod_{n=m}^{k-1} T_{\lambda, \omega}^E(n, n), \quad k > m,
\]

Furthermore, \( T_{\lambda, \omega}^E(k, m) = (T_{\lambda, \omega}^E(m, k))^{-1} \) for \( k < m \) and \( T_{\lambda, \omega}^E(m, m) = 1 \).

**Lemma 5.5** Let \( E \in U \) and \( N \in \mathbb{N} \). Then there is a constant \( \hat{C} \) such that the set

\[
\hat{\Omega}_N(E) = \left\{ \omega \in \Omega \mid \max_{0 \leq n \leq N} \| T_{\lambda, \omega}^E(n, 1) \| \geq e^{\hat{C} N^{3/2}} \right\}
\]

satisfies

\[
P(\hat{\Omega}_N(E)) \geq 1 - e^{-\hat{C} N^{3/2}}.
\]

**Proof.** For sake of notational simplicity, we will drop the index \( \lambda \) on the transfer matrices \( T_{\lambda, \omega}^E \). Let us fix \( E \in U \) and \( N \in \mathbb{N} \) and then split \( N \) into \( N_1 = N_0 + N_1 + 2 N_2 \). For \( j = 0, \ldots, \frac{N_1}{N_3} \), we consider the following events:

\[
\Omega_j^0 = \left\{ \omega \in \Omega \mid \| T_{\omega}^E(j N_3 + N_0, j N_3) \| \leq e^{\hat{C} \gamma_0 N_0} \right\},
\]

\[
\Omega_j^1 = \left\{ \omega \in \Omega \mid \| T_{\pi(j N_3 - N_2, N_3 + N_0 + N_2)}(\omega)(j N_3 + N_0, j N_3) \| \leq e^{\hat{C} \gamma_0 N_0} \right\},
\]

\[
\Omega_j^2 = \left\{ \omega \in \Omega \mid \| T_{\omega}^E(j N_3 + N_0, j N_3) \| \leq e^{\hat{C} \gamma_0 N_0} \right\}.
\]
First we note that uniformly in $\omega$ and for some $\gamma_1 > 0$
\[ \| T^E_\omega (n, m) \| \leq e^{\gamma_1 |n-m|}. \]
Therefore the hypothesis (5.12) implies as in the proof of Lemma 3 of [44] that, for $E \in U$ and $N_0 \in \mathbb{N}$, we have
\[ P(\Omega^2_j) \leq 1 - p_0 < 1, \quad p_0 > 0. \] (5.43)
To shorten notations let us define $\pi_j = \pi_j[N_3 - N_2, j, N_3 + N_0 + N_2]$ and $T^E_\omega j = T^E_\omega (N_3 j + N_0, N_3 j)$.
Using the quasi-locality of $g(\omega) = T^E_\omega$ we get
\[
\| T^E_\omega j - T^E_\pi_j(\omega) \| = \| \sum_{j=N_3}^{j=N_3-1} \left( \prod_{k=j}^{j-l-1} T^E_\omega \right) \left( T^E_\omega - T^E_{\pi_j(\omega)} \right) \left( \prod_{k=l+1}^{j=N_3} T^E_{S^j \pi_j(\omega)} \right) \| \leq N_0 (\sup_\omega (T^E_\omega))^{N_0 - 1} C^{N_2},
\]
where $C = C(g)$ as in (5.2). Now choosing $N_2 = cN_0$ for an adequate constant $c$, it follows that
\[
\| T^E_\omega j - T^E_\pi_j(\omega) \| \leq e^{\gamma_0 N_0}
\]
Therefore for $\omega \in \Omega^0_j$
\[
\| T^E_\pi_j(\omega) \| \leq \| T^E_\omega \| + e^{\gamma_0 N_0} \leq 2 e^{\gamma_0 N_0} \leq e^{\gamma_0 N_0}
\]
for $N_0$ large enough, implying $\Omega^0_j \subset \Omega^1_j$. By a similar calculation, one obtains the second inclusion of
\[
\Omega^0_j \subset \Omega^1_j \subset \Omega^2_j. \] (5.44)
By (5.43) this implies
\[ P(\Omega^1_j) \leq P(\Omega^2_j) \leq 1 - p_0. \]
Now clearly $\Omega^1_j$ is $\pi_j = \pi_j[N_3 - N_2, j, N_3 + N_0 + N_2]$-measurable. Therefore the strong mixing condition (5.3) implies that $P(\Omega^1_j \cap \Omega^1_j) \leq P(\Omega^1_j) (1 + C N_i^{-1}) \leq (1 - p_0)^2 (1 + C N_i^{-1})$. At the next step, one obtains $P(\Omega^1_j \cap \Omega^1_j \cap \Omega^1_j) \leq (1 - p_0)^3 (1 + C N_i^{-1})^2$. Iteration and (5.44) therefore give
\[ P\left( \bigcap_{j=0}^{\infty} \Omega^0_j \right) \leq P(\bigcap_{j=0}^{\infty} \Omega^1_j) \leq (1 - p_0)(1 + C N_i^{-1})^{\frac{N}{N}}. \]
Now let us choose $N_1$ sufficiently large such that $1 - p_1 = (1 - p_0)(1 + C N_i^{-1}) < 1$. Then
\[ P\left( \left\{ \omega \in \bigcap_{j=0}^{\infty} \Omega^0_{j} \bigg| 0 \leq j \leq N/N_3 \sum_{i=N/N_3}^{j=N/N_3} \right\} \right) \leq (1 - p_1)^{\frac{N}{N_3}}. \]
Furthermore $T^E_\omega (j, N_3 + N_0, j, N_3) = T^E_\omega (j, N_3 + N_0, 1) T^E_\omega (j, N_3, 1)^{-1}$. As $A = BC$ implies either $\|B\| \geq \|A\|^{\frac{1}{2}}$ or $\|C\| \geq \|A\|^{\frac{1}{2}}$ for arbitrary matrices, and $\|A^{-1}\| = \|A\|$ for $A \in \text{SL}(2, \mathbb{C})$, it therefore follows that
\[ P\left( \left\{ \omega \in \bigcap_{j=0}^{\infty} \Omega^0_{j} \bigg| 0 \leq j \leq N/N_3 \sum_{i=N/N_3}^{j=N/N_3} \right\} \right) \leq (1 - p_1)^{\frac{N}{N_3}}. \]
Furthermore $T^E_\omega (j, N_3 + N_0, j, N_3) = T^E_\omega (j, N_3 + N_0, 1) T^E_\omega (j, N_3, 1)^{-1}$. As $A = BC$ implies either $\|B\| \geq \|A\|^{\frac{1}{2}}$ or $\|C\| \geq \|A\|^{\frac{1}{2}}$ for arbitrary matrices, and $\|A^{-1}\| = \|A\|$ for $A \in \text{SL}(2, \mathbb{C})$, it therefore follows that
\[ P\left( \left\{ \omega \in \bigcap_{j=0}^{\infty} \Omega^0_{j} \bigg| 0 \leq j \leq N/N_3 \sum_{i=N/N_3}^{j=N/N_3} \right\} \right) \leq (1 - p_1)^{\frac{N}{N_3}}. \]
Choosing $N_0 = cN_3^{\frac{1}{2}}$ with adequate $c$ concludes the proof. \[ \square \]
Chapter 6

Random Lie Group actions on compact manifolds:
a perturbative analysis

The Lyapunov exponent which was perturbatively calculated in the last chapters for certain one dimensional models, determines the decay of the eigenfunctions of the infinite model. Furthermore we obtained for models on strips that all the different Lyapunov exponents are related to the decay of transmission values. Hence it might be interesting to make a perturbation theory for all of them. In Chapter 2 we saw that in general one can calculate all Lyapunov exponents by considering dynamical systems on homogeneous spaces, driven by the transfer matrices. As the homogeneous space depends very much on the size of the strip and the Lyapunov exponents, one wants to calculate, we felt it worth to have a more abstract, conceptual approach which is given in this chapter. It generalizes parts of the work of Chapter 4.

6.1 Main results and discussion

Suppose there are given a Lie group $G \subset \text{GL}(L, \mathbb{C})$, a compact, connected, smooth Riemannian manifold $\mathcal{M}$ without boundary and a smooth, transitive group action $\cdot : \mathcal{G} \times \mathcal{M} \rightarrow \mathcal{M}$. Thus $\mathcal{M}$ is a homogeneous space. Furthermore, let $T_{\lambda,\sigma} \in \mathcal{G}$ be a family of group elements depending on a coupling constant $\lambda \geq 0$ and a parameter $\sigma$ varying in some probability space $(\Sigma, p)$, which is of the following form:

$$
T_{\lambda,\sigma} = \mathcal{R} \exp \left( \sum_{n=1}^{\infty} \lambda^n P_{n,\sigma} \right),
$$

(6.1)

where $\mathcal{R} \in \mathcal{G}$ and $P_{n,\sigma}$ are measurable maps on $\Sigma$ with compact image in the Lie algebra $\mathfrak{g}$ of $\mathcal{G}$ such that

$$
\limsup_{n \to \infty} \sup_{\sigma \in \Sigma} (\|P_{n,\sigma}\|)^{\frac{1}{n}} < \infty
$$

(6.2)
for some norm on $g$. This implies that $T_{\lambda,\sigma}$ is well-defined and analytic in $\lambda$ for $\lambda$ sufficiently small. The expectation value of the first order term $P_{1,\sigma}$ will be denoted by $P = \int p(\sigma) P_{1,\sigma}$.

Let us consider the product probability space $(\Omega, P) = (\Sigma^N, p^N)$. Associated to $\omega = (\sigma_n)_{n \in \mathbb{N}} \in \Omega$ there is a sequence $(T_{\lambda,\sigma_n})_{n \in \mathbb{N}}$ of group elements. An $M$-valued Markov process $x_n(\lambda, \omega)$ with starting point $x_0 \in M$ is defined iteratively by

$$x_n(\lambda, \omega) = T_{\lambda,\sigma_n} \cdot x_{n-1}(\lambda, \omega).$$

(6.3)

The averaged Birkhoff sum of a complex function $f$ on $M$ is

$$I_{\lambda,N}(f) = E_\omega \frac{1}{N} \sum_{n=0}^{N-1} f(x_n(\lambda, \omega)) = \frac{1}{N} \sum_{n=0}^{N-1} (T_{\lambda}^n f)(x_0),$$

(6.4)

where in the second expression we used the Markov transition operator given by $(T_{\lambda} f)(x) = E_\sigma(f(T_{\lambda,\sigma} \cdot x))$. Here and below expectation values w.r.t. $P$ (or $p$) will be denoted by $E$ (or $E_\sigma$ and $E_\omega$). Next recall that an invariant measure $\nu_\lambda$ on $M$ is defined by the property $\int \nu_\lambda(dx) f(x) = \int \nu_\lambda(dx) (T_{\lambda} f)(x)$. The operator ergodic theorem [46, 19.2] then states that $I_{\lambda,N}(f)$ converges almost surely (in $x_0$) w.r.t. any invariant measure $\nu_\lambda$ and for any integrable function $f$. In case that $M$ is a projective space and the action is matrix multiplication, several contributions within the field of products of random matrices have been made. Guivarc’h and Raugi generalized a theorem Furstenberg proved for $2 \times 2$ matrices [9, 29]. If the group generated by $T_{\lambda,\sigma}$, with $\sigma$ varying in the support of $p$, is strongly irreducible and contracting, then there is a unique invariant measure $\nu_\lambda$ which is, moreover, Hölder continuous [9]. To our best knowledge, little seems to be known in more general situations and also concerning the absolute continuity of $\nu_\lambda$ (except for some particular examples [56] and under supplementary hypothesis [15, 84]).

Let $p_1$ be the distribution of the random variable $P_{1,\sigma}$ on the Lie algebra $g$, i.e. for any measurable $b \subset g$ one has $p_1(b) = p\{P_{1,\sigma} \in b\}$. We are interested in a perturbative calculation of $I_{\lambda,N}(f)$ in $\lambda$ for smooth functions $f$ with rigorous control on the error terms. This can be achieved if the support of $p_1$ is large enough in the following sense. First let us focus on the special case $R = 1$ and $P = 0$.

**Theorem 6.1** Let $T_{\lambda,\sigma}$ be of the form (6.1) and assume $R = 1$, $P = E(P_{1,\sigma}) = 0$. Let $x_n$ be the associated Markov process on $M$ as given by (6.3) and let $v = \operatorname{Lie(supp}(p_1))$ be the smallest subspace of $g$ that is closed under the Lie bracket and contains the support of $p_1$. Hypothesis: Suppose that there is a Lie subgroup $\mathcal{U} \subset G$ (not necessarily a submanifold) acting transitively on $M$ such that its Lie algebra $u \subset g$ is contained in $v$.

Then there is a sequence of smooth functions $\rho_m$ with $\int_M d\mu \rho_m = \delta_{m,0}$ and $\rho_0 > 0 \mu$-almost surely, such that for any $M \in \mathbb{N}$ and any function $f \in C^\infty(M)$, one obtains

$$I_{\lambda,N}(f) = \sum_{m=0}^{M} \lambda^m \int_M \mu(dx) \rho_m(x) f(x) + O\left(\frac{1}{N^{\lambda^2}} \lambda^{M+1}\right).$$

(6.5)
6.1. MAIN RESULTS AND DISCUSSION

When $R \neq 1$ or $P \neq 0$ further assumptions are needed in order to control the Birkhoff sums. We assume that $R$ and $P$ generate commuting compact groups, i.e. $RP^{-1} = Ad_{R}(P) = P$ and the abelian groups $\langle R \rangle = \langle RK : k \in \mathbb{Z} \rangle$ and $\langle P \rangle = \{exp(\lambda P) : \lambda \in \mathbb{R} \}$ are compact (the bar denotes the closure). While $\langle P \rangle$ is always connected, $\langle R \rangle$ can possibly be disconnected. However as $\langle R \rangle$ is a closed sub-group of a Lie group it is a Lie group itself and hence a manifold. As it is also compact it can only have finitely many components. 

Similarly, the maps $\theta \mapsto \theta L$ directly leads to the Fourier decomposition of the function $\theta \mapsto f(R(\theta) \cdot x)$, notably

$$f(R(\theta) \cdot x) = \sum_{j \in \mathbb{Z}^{L_{R}}} f_{j}(x) e^{ij \cdot \theta}, \quad (6.6)$$

where

$$f_{j}(x) = \int_{T^{L_{R}}} \frac{d\theta}{(2\pi)^{L_{R}}} e^{-ij \cdot \theta} f\left(\hat{R}(\theta) \cdot x\right), \quad j \cdot \theta = \sum_{l=1}^{L_{R}} j_{l} \theta_{l}.$$ 

Similarly, the maps $\theta \mapsto \theta L_{P}$ and $\theta \mapsto \theta L_{R,P}$ lead to Fourier series.

**Definition 6.1** A function $f \in C^{\infty}(\mathcal{M})$ is said to consist of only low frequencies w.r.t. $\langle R \rangle$ if the Fourier coefficients $f_{j} \in C^{\infty}(\mathcal{M})$ vanish for $j$ with norm $\|j\| = \sum_{l=1}^{L_{R}} |j_{l}|$ larger than some fixed integer $J > 0$. Similarly, $f$ is defined to consist of only low frequencies w.r.t. $\langle P \rangle$ or $\langle R, P \rangle$.

The following definitions are standard (see [49] for references).

**Definition 6.2** Let us define $\hat{\theta}_{R} \in T^{L_{R}}$ by $R_{\hat{\theta}_{R}} = R$ and $\hat{\theta}_{P} \in \mathbb{R}^{L_{P}}$ by $R_{\hat{\theta}_{P}}(\lambda \hat{\theta}_{P}) = \exp(\lambda P)$. Then $R$ is said to be a diophantine rotation or simply diophantine if there is some $s > 1$ and some constant $C$ such that for any non-zero multi-index $j \in \mathbb{Z}^{L_{R}} \setminus \{0\}$ one has

$$|e^{ij \cdot \theta_{R}} - 1| \geq C\|j\|^{-s}.$$
Similar, \( P \) is said to be diophantine, or a diophantine generator of a rotation, if there is some \( s > 1 \) and some constant \( C \), such that for any non-zero multi-index \( j \in \mathbb{Z}^{L_{\gamma}} \setminus \{0\} \) one has
\[
|j : \hat{\theta}_P| \geq C\|j\|^{-s}.
\]

As final preparation before stating the result, let us introduce the measure \( \mathbf{p} \) on the Lie algebra \( \mathfrak{g} \) obtained from averaging the distribution \( \mathbf{p}_1 \) of the lowest order terms \( P_{1,\sigma} \) w.r.t. the Haar measure \( dR \) on the compact group \( \langle R, P \rangle \), namely for any measurable set \( \mathbf{b} \subset \mathfrak{g} \),
\[
\mathbf{p}(\mathbf{b}) = \int_{\langle R, P \rangle} dR \mathbf{p}(\{\sigma \in \Sigma : R P_{1,\sigma} R^{-1} \in \mathbf{b}\}).
\]

**Theorem 6.2** Let \( T_{\lambda,\sigma} \) be of the form (6.1) and \( x_n \) the associated Markov process on \( M \) as given in (6.3). Let \( \mathfrak{u} = \text{Lie} (\text{supp}(\mathbf{p}), P) \) be the Lie subalgebra of \( \mathfrak{g} \) generated by the support of \( \mathbf{p} \) and \( P \). Suppose that \( \mathcal{U} \subset \mathcal{G} \) is a Lie subgroup of \( \mathcal{G} \) acting transitively on \( M \) such that its Lie algebra \( \mathfrak{u} \subset \mathfrak{g} \) is contained in \( \mathfrak{v} \). Suppose that \( f \in C^\infty(M) \) and one of the following conditions hold:

(i) \( R \) and \( P \) are diophantine and \( M = K/H \) where \( K \) and \( H \) are compact Lie groups.

(ii) \( f \) consist of only low frequencies w.r.t. \( \langle R, P \rangle \).

Then there is a \( \mu \)-almost surely positive function \( \rho_0 \in C^\infty(M) \) normalized w.r.t. the Riemannian volume measure \( \mu \) on \( M \), such that
\[
I_{\lambda,N}(f) = \int_M \mu(dx) \rho_0(x) f(x) + O\left(\frac{1}{N\lambda^2}, \lambda\right). \tag{6.7}
\]

Moreover, the probability measure \( \rho_0 \mu \) is invariant under the action of \( \langle R, P \rangle \).

The probability measures \( \sum_{m=0}^{M} \lambda^m \rho_m \mu \) in Theorem 6.1 and \( \rho_0 \mu \) in Theorem 6.2 can be seen as perturbative approximations of the invariant measures \( \nu_\lambda \). In fact, integrating (6.5) over the initial condition \( x_0 \) w.r.t. any invariant measure \( \nu_\lambda \) and then taking the limit \( N \to \infty \), shows that for any smooth function
\[
\int_M \nu_\lambda(dx) f(x) = \sum_{m=0}^{M} \lambda^m \int_M \mu(dx) \rho_m(x) f(x) + O\left(\lambda^{M+1}\right).
\]

This means that the invariant measure is unique in a perturbative sense and, moreover, its unique approximations are absolutely continuous with smooth density. In fact as a corollary one obtains the following.

**Corollary 6.1** Assume all assumptions as in Theorem 6.1 or Theorem 6.2 are fulfilled. Let \( (\nu_\lambda)_\lambda \) be a family of invariant probability measures in \( \lambda \). Then one has for \( \rho_0 \) defined as in the theorems above
\[
\text{w}^* - \lim_{\lambda \to 0} \nu_\lambda = \rho_0 \mu,
\]
where \( \text{w}^* - \lim \) denotes convergence in the weak-* topology on the set of Borel measures.
6.1. MAIN RESULTS AND DISCUSSION

Proof. Approximating smooth functions by their uniform convergent Fourier series shows that the set of smooth functions consisting of only low frequencies w.r.t. \( \langle R, \mathcal{P} \rangle \) is dense in the set of continuous functions w.r.t. \( \| \cdot \|_\infty \)-norm. The set of probability measures is norm bounded by 1 w.r.t. the dual norm. Now let \( g \in C(M) \). For any \( \epsilon > 0 \) there is a smooth function \( g \) consisting of only low frequencies such that \( \| f - g \|_\infty < \epsilon \). Then one has
\[
| \nu_\lambda(f) - \rho_0 \mu(f) | \leq | \nu_\lambda(f - g) | + | \nu_\lambda(g) - \rho_0 \mu(g) | + | \rho \mu(g - f) | \leq 2 \epsilon + | \nu_\lambda(g) - \rho_0 \mu(g) |.
\]
Taking the \( \limsup_{\lambda \to 0} \) one obtains \( \limsup_{\lambda \to 0} | \nu_\lambda(f) - \rho_0 \mu(f) | \leq 2 \epsilon \) for any \( \epsilon > 0 \). Therefore one has
\[
\limsup_{\lambda \to 0} | \nu_\lambda(f) - \rho_0 \mu(f) | = 0
\]
for any continuous function \( f \in C(M) \) which gives the desired result. \( \square \).

Note that Corollary 6.1 does not give any estimate on the error term \( | \nu_\lambda(f) - \rho_0 \mu(f) | \) for continuous functions in terms of \( \lambda \).

Now let us give some further remarks on the theorems stated above.

Remark 6.1 The error bound \( O((N \lambda^2)^{-1}, \lambda^{M+1}) \) in Theorem 6.1 means that the modulus of the error term can be estimated by \( C((N \lambda^2)^{-1} + \lambda^{M+1}) \) for some constant \( C \). But this constant depends on \( f \) and on \( M \). Actually our bound grows fast in \( M \) and hence we have not shown that there is a unique invariant measure for small \( \lambda \).

Remark 6.2 In fact, there is a simple example satisfying the conditions of Theorem 6.1 where one does not find \( \lambda_0 > 0 \) such that for all \( \lambda \in [0, \lambda_0] \) the invariant measure is unique. Let the Lie group and the manifold be given by the one-dimensional torus \( G = M = S^1 = \{ z \in \mathbb{C} : |z| = 1 \} \) and the Lie group action is the ordinary multiplication. Let furthermore \( R = 1 \) and \( P_{1, \sigma} \) be Bernoulli distributed with probability \( \frac{1}{2} \) at \( +\pi \) and \( -\pi \). (The Lie algebra of \( S^1 \) is given by \( i \mathbb{R} \).) The other terms \( P_{n, \sigma} \) shall be zero for all \( n \geq 2 \). Clearly, all conditions of Theorem 6.1 are fulfilled. Now for rational \( \lambda = \frac{p}{q} \), any measure on \( S^1 \) leaving the rotation with angle \( \frac{\pi}{q} \) invariant, is an invariant measure and there are many of them. However, for irrational \( \lambda \) the invariant measure is unique and given by the Haar measure on \( S^1 \). Our conjecture is: Given the conditions of Theorem 6.1 one finds \( \lambda_0 > 0 \) just that for Lebesgue a.e. \( \lambda \in [0, \lambda_0] \) there is a unique invariant measure.

Remark 6.3 The main hypothesis of Theorems 6.1 and 6.2 is that \( u \subset v \). This can roughly be thought of as a Lie algebra equivalent of Furstenberg’s irreducibility condition. Let us note that non-trivial \( \mathcal{R}, \mathcal{P} \) lead to a larger support for \( \mathcal{P} \) and hence weaken this hypothesis. A second hypothesis is that the group \( \langle \mathcal{R} \rangle \) is compact. This excludes many situations appearing in physical models where hyperbolic or parabolic channels appear. In some particular situations this could be dealt with \([72, 76]\), (cf. Chapter 4).

Remark 6.4 As the action is transitive, \( M \) is always a homogeneous space and given as quotient of \( G \) w.r.t. some isotropy group, but hypothesis (i) requires that \( M \) is, moreover, a quotient of a compact group (which in the example of the next chapter is a subgroup of \( G \)). The assumption that \( G \subset \text{GL}(L, \mathbb{C}) \) (or, equivalently \( G \) has a faithful representation) is only needed for the proof of Theorem 6.2 under hypothesis (ii).
Remark 6.5 Let us suppose that $\mathcal{K}$ in Theorem 6.2(i) is a compact subgroup of $G$. As $\mathcal{K}$ acts transitive on $M$, the Haar measure $dk$ on $\mathcal{K}$ induces a unique natural $\mathcal{K}$-invariant measure on $M$ which one may choose to be $\mu$ (which is also the volume measure of the metric $\int dk K_\ast g$). It is interesting to examine whether $\rho_0 = 1_M$, that is the lowest order approximation of the invariant measure is given by the natural measure. The proof below provides a technique to check this. More precisely, in the notations developed below, the criterion is that $\hat{L}^{\ast}1_M = 0$. An example, where this can indeed be checked is developed in the next chapter. Note that, if $\mathcal{K}$ is as above, then any conjugation $N\mathcal{K}N^{-1}$ with an element $N \in G$ has another natural measure, given by $J_N \mu$ where $J_N$ is the Jacobian of the map $x \mapsto N \cdot x$. Unless $\mu$ is invariant under all of $G$, the equality $\rho_0 = 1_M$ is hence linked to a good choice of $\mathcal{K}$. If $\mu$ is invariant under $G$, then it is also an invariant measure for the Markov process and under the hypothesis of Theorem 6.2 one therefore has $\rho_0 = 1_M$.

Remark 6.6 If $\langle R, P \rangle$ acts transitive on $M$, then the measure $\rho_0 \mu$ is uniquely determined by the fact that it is invariant under the action of $\langle R, P \rangle$ and normalized. Moreover, $M$ is isomorphic to the quotient of $\langle R, P \rangle$ and the stabilizer $S_x$ of any point $x \in M$ (which is a compact abelian subgroup of $\langle R, P \rangle$). Hence in this case, $M$ is a torus and the action is simply the translation on the torus. Consequently the measure $\rho_0 \mu$ is the Haar measure. Note that, if $P = 0$, this holds independently of the perturbation and is imposed by the deterministic process for $\lambda = 0$.

Remark 6.7 If the action of $\langle R \rangle$ on $M$ is not transitive, there are many invariant measures $\nu_0$ for the deterministic dynamics (in particular, if $R = 1$ any measure is invariant under $\langle R \rangle$). Under the hypothesis of Theorems 6.1 and 6.2, the random perturbations $P_{1,\sigma}$ and $P_{2,\sigma}$ single out a unique perturbative invariant measure $\rho_0 \mu$.

Remark 6.8 We believe the condition that $\mathcal{R}$ and $P$ commute to be unnecessary. In fact, we expect that conditions on $\mathcal{P}$ can be replaced by conditions on $\mathcal{P} = \int_{\langle R \rangle} dR R P R^{-1}$.

Remark 6.9 Let us relate these results to some prior work on the rigorous perturbative evaluation of the averaged Birkhoff sums (6.4). In the case of $G = \text{SL}(2, \mathbb{R})$, $M = \mathbb{R}P(1)$ and a rotation matrix $R$ in (6.1), Pastur and Figotin [64] showed (6.7) for the lowest two harmonics whenever $\mathcal{R}, \mathcal{R}^2 \neq \pm 1$ (see Section 1.6). The above result combined with the calculations in the next chapter shows that (6.7) holds also for other functions with $\rho_0 = 1_M$. Without the conditions $\mathcal{R}, \mathcal{R}^2 \neq \pm 1$, Theorem 6.2 was proven in Chapter 4 ([77, 72]). Moreover, when $\mathcal{R}^K = 1$ (at so-called anomalies) and for an absolutely continuous distribution on $G$, Theorem 6.1 was proven by Campanino and Klein [15]. Quasi-one-dimensional generalizations (higher dimensional matrices) of [64] in the case where $G$ is a symplectic group were obtained in [76, 78]. An attempt to treat higher dimensional anomalies is [23]. To further generalize the above to quasi-one-dimensional systems was the main motivation for this work.

In order to clearly exhibit the strategy of the proof of the theorems, we first focus on the case $\mathcal{R} = 1$ and $\mathcal{P} = 0$ in Sections 6.2 and 6.3, which corresponds to a higher dimensional
anomaly in the terminology of the prior work in Chapter 4 [77, 72]. The main idea is then to expand $T_{\lambda} f$ into a Taylor expansion in $\lambda$. This directly leads to a second order differential operator $\mathcal{L}$ on $\mathcal{M}$ of the Fokker-Planck type, for which the Birkhoff sums $I_{\lambda,N}(\mathcal{L} f)$ vanish up to order $\lambda$. Under the hypothesis of Theorem 6.1, it can be shown to be a sub-elliptic Hörmander operator on the smooth functions on $\mathcal{M}$ with a one-dimensional cokernel. Then one can deduce that $\mathbb{C} + \mathcal{L}(C^\infty(\mathcal{M})) = C^\infty(\mathcal{M})$ and that the kernel of $\mathcal{L}^*$ is spanned by a smooth positive function $\rho_0$. These are the main elements of the proof of Theorem 6.1 for $M = 1$. Then using the properties of the operators $\mathcal{L}$ and $\mathcal{L}^*$ and a further Taylor expansion of $T_{\lambda} f$ one can prove Theorem 6.1 by induction. The additional difficulties for other $\mathcal{R}, \mathcal{P}$ in Theorem 6.2 are dealt with in the more technical Section 6.4.

### 6.2 Fokker-Planck operator and its properties

In this section we suppose $\mathcal{R} = 1$ and $\mathcal{P} = \mathbf{E}(\mathcal{P}_{1,\sigma}) = 0$ in (6.1) and introduce in this case the backward Kolmogorov operator $\mathcal{L}$ and its adjoint $\mathcal{L}^*$, called forward Kolmogorov or also Fokker-Planck operator. Their use for the calculation of the averaged Birkhoff sum is exhibited and several properties of these operators are studied. One way to define the operator $\mathcal{L} : C^\infty(\mathcal{M}) \to C^\infty(\mathcal{M})$ is

$$
(\mathcal{L} f)(x) = \frac{d^2}{d\lambda^2} \bigg|_{\lambda=0} (T_{\lambda} f)(x). \tag{6.8}
$$

Let us rewrite this using the smooth vector fields $\partial_P$ associated to any element $P \in \mathfrak{g}$ by

$$
\partial_P f(x) = \frac{d}{d\lambda} \bigg|_{\lambda=0} f(e^{\lambda P} \cdot x). \tag{6.9}
$$

Then $\mathcal{L}$ is given by

$$
\mathcal{L} = \mathbf{E}_\sigma \left( \partial_{P_{1,\sigma}}^2 + 2 \partial_{P_{2,\sigma}} \right). \tag{6.10}
$$

**Proposition 6.1** For $F \in C^\infty(\mathcal{M})$ one has

$$
I_{\lambda,N}(\mathcal{L} F) = O\left( \frac{1}{N^2 \lambda^4} \right). \tag{6.11}
$$

**Proof.** For $P \in \mathfrak{g}$, a Taylor expansion with Lagrange remainder gives

$$
F(e^P \cdot x) = F(x) + (\partial_P F)(x) + \frac{1}{2} (\partial_P^2 F)(x) + \frac{1}{6} (\partial_P^3 F)(e^P \cdot x),
$$

for some $\chi \in [0, 1]$. Choose $P = \lambda P_{1,\sigma} + \lambda^2 P_{2,\sigma} + \lambda^3 S_\sigma(\lambda)$, where $S_\sigma(\lambda) = \sum_{n=3}^\infty \lambda^{n-3} P_{n,\sigma}$ and use that $P_{1,\sigma}$ is centered to obtain

$$
\mathbf{E}_\sigma F(T_{\lambda,\sigma} \cdot x) = F(x) + \mathbf{E}_\sigma \left( \lambda^2 \left( \frac{1}{2} \partial_{P_{1,\sigma}}^2 F(x) + 2 \partial_{P_{2,\sigma}} F(x) \right) \right) + O(\lambda^3)
$$

$$
= F(x) + \frac{1}{2} \lambda^2 \mathcal{L} F(x) + O(\lambda^3).
$$


The error terms depend on derivatives of $F$ up to order 3 and are uniform in $x$ because $\mathcal{M}$ is compact and $\mathcal{P}_{1,\sigma}, \mathcal{P}_{2,\sigma}$ and $\mathcal{S}_\sigma(\lambda)$ are compactly supported by (6.2). Due to definition (6.3) this implies

$$E_\omega \frac{1}{N} \sum_{n=1}^{N} F(x_n(\lambda, \omega)) = E_\omega \frac{1}{N} \sum_{n=0}^{N-1} F(x_n(\lambda, \omega)) + \frac{\lambda^2}{2} I_{\lambda, N}(\mathcal{L}F) + O(\lambda^3).$$

As the appearing sums only differ by a boundary term, resolvinng for $I_{\lambda, N}(\mathcal{L}F)$ finishes the proof.

Next let us bring the operator $\mathcal{L}$ into a normal form. According to Appendix A.2, one can decompose $\mathcal{P}_{1,\sigma}$ into a finite linear combination of fixed Lie algebra vectors $\mathcal{P}_i \in \mathfrak{g}, i \in I$, with uncorrelated real random coefficients, namely

$$\mathcal{P}_{1,\sigma} = \sum_{i=1}^{I} v_{i,\sigma} \mathcal{P}_i, \quad v_{i,\sigma} \in \mathbb{R}, \quad E_\sigma(v_{i,\sigma}v_{i',\sigma}) = \delta_{i,i'}. \quad \text{Then (6.10)}$$

implies that $\mathcal{L}$ is in the so-called Hörmander form

$$\mathcal{L} = \sum_{i=1}^{I} \partial^2_{\mathcal{P}_i} + 2 \partial_Q,$$

where $Q = E_\sigma(\mathcal{P}_{2,\sigma})$. Using the main assumption of Theorem 6.1 (that is, $\mathfrak{v} \subset \mathfrak{u}$) one can show that $\mathcal{L}$ satisfies the strong Hörmander property of rank $r \in \mathbb{N}$ [40, 42, 68].

**Proposition 6.2** Under the assumptions of Theorem 6.1, there exists $r \in \mathbb{N}$ such that $\mathcal{L}$ satisfies a strong Hörmander property of rank $r$, i.e. the vector fields $\partial_{\mathcal{P}_i}$ and their $r$-fold commutators span the whole tangent space at every point of $\mathcal{M}$.

In order to check this, one needs to calculate the commutators of vector fields $\partial_P, \partial_Q$ for $P, Q \in \mathfrak{g}$. Let $X_P, X_Q$ denote the left-invariant vector fields on $\mathcal{G}$ and furthermore introduce for each $x \in \mathcal{M}$ a function on $\mathcal{G}$ by $f_x(T) = f(T \cdot x), T \in \mathcal{G}$. Then one obtains

$$\partial_P \partial_Q f(x) = \frac{d}{d\lambda} \bigg|_{\lambda=0} (\partial_Q f)(e^{\lambda P} \cdot x) = \frac{d^2}{d\lambda d\mu} \bigg|_{\lambda,\mu=0} f(e^{\mu Q}e^{\lambda P} \cdot x) = X_Q X_P f_x(1),$$

which implies

$$(\partial_P \partial_Q - \partial_Q \partial_P) f(x) = (X_Q X_P - X_P X_Q) f_x(1) = X_{[Q,P]} f_x(1) = \partial_{[Q,P]} f(x), \quad (6.11)$$

where $[Q, P]$ denotes the Lie bracket (this is well-known, see Theorem II.3.4 in [37]). We also need the following lemma for the proof of Proposition 6.2.

**Lemma 6.1** Let $\mathcal{U} \subset \mathcal{G}$ be a Lie subgroup of $\mathcal{G}$ that acts transitively on $\mathcal{M}$ and denote the Lie algebra of $\mathcal{U}$ by $\mathfrak{u}$. Then the vector fields $\partial_P, P \in \mathfrak{u}$, span the whole tangent space at each point of $\mathcal{M}$.
6.2. FOKKER-PLANCK OPERATOR AND ITS PROPERTIES

Proof. First let us show that there is a dense set of points in $\mathcal{M}$ for which the vector fields $\partial_P$, $P \in \mathfrak{u}$, span the whole tangent space. Indeed, for a fixed $x \in \mathcal{M}$ consider the surjective, smooth map $\varphi_x : U \to \mathcal{M}$, $\varphi_x(U) = U \cdot x$. A point $x' \in \mathcal{M}$ is called regular for $\varphi_x$ if and only if for any point in the pre-image of $x'$ the differential $D\varphi_x$ is surjective. For each point $x'$, the hypothesis implies that there is a $U \in \mathcal{U}$ such that $x' = \varphi_x(U) = U \cdot x$ and the regularity of $x'$ then shows that the paths $\lambda \mapsto \varphi_x(e^{\lambda P} U) = e^{\lambda P} \cdot x'$, $P \in \mathfrak{u}$, span the whole tangent space at $x'$. By Sard’s theorem [39], the set of regular points is dense in $\mathcal{M}$.

Actually the existence of only 1 regular point $x$ implies that all points are regular. In fact, again any other point is of the form $x' = U \cdot x$. As the map $x \mapsto x' = U \cdot x$ is a diffeomorphism, the push-forward of the paths $\lambda \mapsto \exp(\lambda P) \cdot x$, $P \in \mathfrak{u}$, given by the paths $\lambda \mapsto U \exp(\lambda P) \cdot x = e^{MU^{-1}P} \cdot x'$, $P \in \mathfrak{u}$, span the tangent space also at $x'$.

Proof of Proposition 6.2. Define iteratively the subspaces $\mathfrak{v}_r \subset \mathfrak{g}$ by

$$
\mathfrak{v}_1 = \text{span}\{\mathcal{P}_i : 1 \leq i \leq I\}, \quad \mathfrak{v}_r = \text{span}\{\mathfrak{v}_{r-1} \cup [\mathfrak{v}_{r-1}, \mathfrak{v}_1]\}.
$$

(6.12)

By definition one has $\mathfrak{v}_1 = \text{span}(\text{supp}(\mathcal{P}_r))$. The space $\mathfrak{v} \subset \mathfrak{g}$ defined in Theorem 6.1 is equal to $\mathfrak{v} = \text{Lie}(\mathfrak{v}_1)$. Due to (6.11), the strong Hörmander property of rank $r$ is equivalent to the property that $\partial_P$, $P \in \mathfrak{v}_r$, spans the whole tangent space at every point $x \in \mathcal{M}$.

By the Lemma 6.1 and the assumption of Theorem 6.1 this is fulfilled if $\mathfrak{v}_r = \mathfrak{v}$ for some $r$. As the vector spaces $\mathfrak{v}_r$ are nested and $\mathfrak{g}$ is finite-dimensional, the sequence has to become stationary. This means, there is some $r$ such that $\mathfrak{v}_r = \mathfrak{v}_{r+1}$. Using the Jacobi identity, one then checks that $\mathfrak{v}_r$ is closed under the Lie bracket and therefore $\mathfrak{v}_r = \mathfrak{v}$. □

Next we want to recollect the consequences of the strong Hörmander property of rank $r$ as proven in [40, 68, 42] (cf. Appendix A.3). The first basic fact is the subelliptic estimate within any chart

$$
\|f\|_{(\mathfrak{g})} \leq C(\|\mathcal{L}f\|_{(0)} + \|f\|_{(0)}),
$$

(6.13)

where $\| . \|_{(\mathfrak{g})}$ denotes the norm of the Sobolev space $H^s$ (also denoted by $W^{s,2}$) of the chart. Using a finite atlas of $\mathcal{M}$ one can define a global Sobolev space $H^s(\mathcal{M})$ with norm denoted by $\| . \|_{(\mathfrak{g})}$ (see Appendix A.4). Then the estimate (6.13) holds also w.r.t. these global norms. Moreover, the norm $\| . \|_{(0)}$ can be seen to be equivalent to the norm in $L^2(\mathcal{M}, \mu)$ where $\mu$ is the Riemannian volume measure. As usual, the embedding of $H^{s+\epsilon}(\mathcal{M})$ in $H^s(\mathcal{M})$ is compact for any $\epsilon > 0$.

The second basic fact is the hypoellipticity of $\mathcal{L}$. In order to state this property, let us first extend $\mathcal{L}$ in the usual dual way to an operator $\mathcal{L}_{\text{dis}}$ on the space $\mathcal{D}' = (C^\infty(\mathcal{M}))'$ of distributions on $\mathcal{M}$. Then hypoellipticity states that, for any smooth function $g$, the solution $f$ of $\mathcal{L}_{\text{dis}}f = g$ is itself smooth.

The Fokker-Planck operator $\mathcal{L}^*$ is the adjoint of $\mathcal{L}$ in $L^2(\mathcal{M}, \mu)$. Because $\mathcal{M}$ is compact and has no boundary, the domain $\mathcal{D}(\mathcal{L}^*)$ of $\mathcal{L}^*$ contains the smooth functions $C^\infty(\mathcal{M})$. Furthermore $\mathcal{L}^*$ is again a second-order differential operator with the same principal symbol as $\mathcal{L}$. Therefore also $\mathcal{L}^*$ satisfies the strong Hörmander condition of rank $r$. Thus the
subelliptic estimate as well as the hypoellipticity property also holds for $L^*_\text{dis}$. We, moreover, deduce that $L$ is closable with closure $\overline{L} = L^{**} \subset L_{\text{dis}}$.

The following proposition recollects properties of $L$ as a densely defined operator on the Hilbert space $L^2(\mathcal{M}, \mu)$.

**Proposition 6.3** There exists $c_0 > 0$ such that for $c > c_0$ the following holds.

(i) $L - c$ is dissipative.

(ii) $(L - c)(C^\infty(\mathcal{M}))$ is dense in $L^2(\mathcal{M}, \mu)$.

(iii) $\overline{L} - c$ is maximally dissipative.

(iv) $\overline{L} - c$ is the generator of a contraction semigroup on $L^2(\mathcal{M}, \mu)$.

(v) The resolvent $(\overline{L} - c)^{-1}$ exists and is a compact operator on $L^2(\mathcal{M}, \mu)$.

**Proof.** (i) Let us rewrite $L$:

$$Lf = \sum_{i=1}^{I} \left[ \text{div}(\partial_{P_i}(f) \partial_{P_i}) - \text{div}(\partial_{P_i} \partial_{P_i}(f)) \right] + 2 \partial_Q(f).$$

Defining $X$ to be the smooth vector field $2 \partial_Q - \sum_i \text{div}(\partial_{P_i}) \partial_{P_i}$, one has

$$Lf = \sum_{i=1}^{I} \text{div}(\partial_{P_i}(f) \partial_{P_i}) + X(f).$$

For a real, smooth function $f$, the divergence theorem and estimate on the negative quadratic term gives

$$\langle f | Lf \rangle = \int_{\mathcal{M}} d\mu \left[ -\sum_{i=1}^{I} \partial_{P_i}(f) \partial_{P_i}(f) + f X(f) \right] \leq \int_{\mathcal{M}} d\mu \ f X(f).$$

Using $2fX(f) = X(f^2) = \text{div}(f^2X) - f^2 \text{div}(X)$ and again the divergence theorem, it follows that

$$\langle f | Lf \rangle \leq -\frac{1}{2} \int_{\mathcal{M}} d\mu \ \text{div}(X) f^2 \leq \frac{1}{2} \| \text{div}(X) \|_{\infty} \| f \|_2^2. \quad (6.14)$$

As $L$ is real, it follows that $\Re \langle f | (L - c)f \rangle \leq 0$ for $f \in C^\infty(\mathcal{M})$ and $c > c_0$ where $c_0 = \frac{1}{2} \| \text{div}(X) \|_{\infty}$. By definition, this means precisely that $L - c$ is dissipative.

(ii) Let $h \in L^2(\mathcal{M}, \mu)$ such that $\langle h | Lf - cf \rangle = 0$ for all $f \in C^\infty(\mathcal{M}) = \mathcal{D}(L)$. Then $h$ is in the kernel of $L^*_\text{dis}$. By hypoellipticity it follows that $h \in C^\infty(\mathcal{M})$. Therefore $\langle h | Lh \rangle = c\| h \|^2_2$ contradicting (6.14) unless $h = 0$.

The statement (iii) means that there is no dissipative extension, which follows directly from (i) and (ii) by [21, 2.24, 2.25 and 6.4]. Item (iv) follows from the same reference.

Concerning (v), the existence of the resolvent follows directly upon integration of the contraction semigroup. Its compactness follows from the subelliptic estimate (6.13) and the compact embedding of $H^s(\mathcal{M})$ into $L^2(\mathcal{M}, \mu)$. 

The next proposition is based on Bony’s maximum principle for strong Hörmander operators [8], as well as standard Fredholm theory.
Proposition 6.4 (i) The kernel of $\mathcal{L}$ consists of the constant functions on $\mathcal{M}$.

(ii) The kernel of $\mathcal{L}^*$ is one-dimensional and spanned by a smooth function $\rho_0$.

(iii) $\text{Ran} \mathcal{L} = (\ker \mathcal{L}^*)^\perp$ and $\text{Ran} \mathcal{L}^* = (\ker \mathcal{L})^\perp = (\ker \mathcal{L})^\perp$.

(iv) $\rho_0$ is $\mu$-almost surely positive.

Proof. (i) By Corollaire 3.1 of [8] a smooth function $f$ which has a local maximum and for which $\mathcal{L}f=0$ has to be constant on (the pathwise connected compact set) $\mathcal{M}$. If $f$ lies in the kernel of the closure $\overline{\mathcal{L}} = \mathcal{L}^*$, then $\mathcal{L}\mu f = 0$. As $\mathcal{L}$ is hypoelliptic, $f \in C^\infty(\mathcal{M})$ and therefore $f$ is again constant.

(ii) Choose $c > c_0$ as in Proposition 6.3 and let $K = (\overline{\mathcal{L}} + c)^{-1}$. Then one has

$$\mathcal{L}f = g \iff \overline{\mathcal{L}}(\overline{\mathcal{L}} + c) = cf + g \iff f = cKf + Kg \iff (1 - cK)f = Kg,$$

and similarly $\mathcal{L}^*f = g \iff (1 - cK^*)f = Kg$. For $g = 0$ this implies $\ker \mathcal{L} = \ker(1 - cK)$ and $\ker \mathcal{L}^* = \ker(1 - cK^*)$. By the Fredholm alternative (the index of $1 + cK$ is 0) the dimension of these two kernels are equal and by (i) hence both one-dimensional. The smoothness of the function in the kernel follows from the hypoellipticity of $\mathcal{L}^*$.

(iii) For $v \in \ker \mathcal{L}^* = \ker(1 - cK^*)$ and $\langle g \mid v \rangle = 0$, one has $0 = \langle g \mid v \rangle = \langle g \mid cK^*v \rangle = c\langle Kg \mid v \rangle$, therefore $g \in (\ker \mathcal{L}^*)^\perp$ implies $Kg \in \ker(1 - cK)^\perp$ and the Fredholm alternative states that $(1 - cK)f = Kg$ is solvable. Hence by the above, $\overline{\mathcal{L}}f = g$ is solvable. Therefore $\text{Ran} \overline{\mathcal{L}} = (\ker \mathcal{L}^*)^\perp$. The other equality is proven analogously.

(iv) Let $f \geq 0$ be smooth and suppose that $\int d\mu \rho_0 f = 0$. According to (ii), (iii) and hypoellipticity this implies that $f = \mathcal{L}F \geq 0$ for some smooth $F$. Again by Bony’s maximum principle $F$ is constant and therefore $f = 0$. Hence for any non-vanishing positive function $f$ one has $\int d\mu \rho_0 f > 0$.

Even though not relevant for the sequel, let us also prove the following

Proposition 6.5 $\mathcal{L}$ generates a contraction semigroup in $(C(\mathcal{M}), \| \cdot \|_\infty)$, also called a Feller semigroup.

Proof. This will follow directly from the Hille-Yosida theorem [46, 19.11] once we verified that $(\mathcal{L} - c)C^\infty(\mathcal{M})$ is dense in $C(\mathcal{M})$ for some $c > 0$ and that $\mathcal{L}$ satisfies the positive-maximum principle. The first property follows from the existence of the resolvent (Proposition 6.3) and the hypoellipticity. For the second, let a smooth $f$ have a positive local maximum at some $x \in \mathcal{M}$. Then one only has to check $(\mathcal{L}f)(x) \leq 0$, which follows because the first derivatives of $f$ vanish, its second derivative is negative and the principal symbol is positive definite.

One can rewrite (6.8) as $\lim_{\lambda \to 0} \frac{1}{2\lambda^2}(T_\lambda - 1)f = \mathcal{L}f$ in $\| \cdot \|_\infty$ and for $f \in C^\infty(\mathcal{M})$. Hence the above statement and [46, 19.28] implies directly the following approximation result of the Feller process by the discrete time Markov processes.

Corollary 6.2 Let $e^{\mathcal{L}}$ denote the Feller semigroup of Proposition 6.5. Then with convergence in $(C(\mathcal{M}), \| \cdot \|_\infty)$

$$\lim_{\lambda \to 0} T_{\lambda}^{\frac{1}{2\lambda^2}} f = e^{\mathcal{L}} f.$$
6.3 Control of Birkhoff sum in the case \( R = 1, P = 0 \)

The aim of this section is the proof of Theorem 6.1.

**Proposition 6.6** Let \( R = 1 \) and \( P = 0 \). The kernel of \( \mathcal{L}^*_{du} \) is spanned by a non-negative smooth function \( \rho_0 \) that is normalized by \( \int_M \mu \rho_0 = 1 \). For \( f \in C^\infty(M) \)

\[
I_{\lambda,N}(f) = \int_M \mu \rho_0 f + O\left(\frac{1}{N^2 \lambda^2}, \lambda\right).
\]

**Proof.** By hypoellipticity the kernel of \( \mathcal{L}^*_{du} \) coincides with the kernel of \( \mathcal{L}^* \). First we show \( C^\infty(M) = C^1(M) + L^2(C^\infty(M)) \). Indeed, let \( f \in C^\infty(M) \). Set \( C = \int_M \mu f \rho_0 \) and \( \hat{f} = f - C \). Then one has \( \int_M \mu \hat{f} \rho_0 = 0 \) and therefore \( \hat{f} \in (\ker \mathcal{L}^*)^\perp = \text{Ran} \mathcal{L}^* \) by Proposition 6.4. By hypoellipticity, \( \hat{f} \in L^2(C^\infty(M)) \).

Now using Proposition 6.1 and the above decomposition

\[
I_{\lambda,N}(f) = I_{\lambda,N}(\hat{f} + C) = C + I_{\lambda,N}(\mathcal{L}F) = C + O(N^{-1} \lambda^{-2}, \lambda),
\]

one completes the proof. \( \Box \)

In order to prove Theorem 6.1, let us define the operators

\[
\mathcal{L}^{(M)} f(x) = \frac{d^M}{d\lambda^M} \bigg|_{\lambda=0} (T_\lambda f)(x), \quad f \in C^\infty(M).
\]

Then \( \mathcal{L}^{(0)} = 0 \) as \( P_{1,\sigma} \) is centered and \( \mathcal{L}^{(2)} = \mathcal{L} \). Using (6.1) these operators can be written as

\[
\mathcal{L}^{(M)} f = E_\sigma \left( \sum_{m=0}^{M} \sum_{n_1 + \ldots + n_m = M} \frac{M!}{m!} \partial_{P_{n_1, \sigma}} \ldots \partial_{P_{n_m, \sigma}} f \right).
\]

Hence \( \mathcal{L}^{(M)} \) is a differential operator of order \( M \). As \( 1_M \in \ker \mathcal{L}^{(m)} \) and hence \( \ker \mathcal{L} \subset \ker \mathcal{L}^{(m)} \) for all positive \( m \), one obtains using Proposition 6.4 (iii)

\[
\text{Ran} \mathcal{L}^{(m)*} \subset (\ker \mathcal{L}^{(m)})^\perp \subset (\ker \mathcal{L})^\perp = \text{Ran} \mathcal{L}^*.
\]

Therefore, and as \( \ker \mathcal{L}^* \) is one dimensional, the functions \( \rho_M \) for \( M \in \mathbb{N} \) are iteratively and uniquely defined by

\[
\mathcal{L}^* \rho_M = \sum_{m=1}^{M} \frac{2}{(m + 2)!} \mathcal{L}^{(m+2)*} \rho_{M-m}, \quad \int_M \mu \rho_M = 0 , \quad (6.15)
\]

with \( \rho_0 \) given by Proposition 6.6. By induction and hypoellipticity of \( \mathcal{L}^* \), it follows that \( \rho_M \) is a smooth function for all \( M \), therefore the r.h.s. of (6.15) always exists. Now we can complete the

**Proof of Theorem 6.1.** The proof will be done by induction. The case \( M = 0 \) is contained in Proposition 6.6. For the step from \( M - 1 \) to \( M \), we first need a Taylor expansion of
higher order than done so far, As \( P_{1,\sigma} \) is centered and due to the compact support of \( P_{n,\sigma} \) and \( \sum_{m \geq n} \lambda^{m-n} P_{n,\sigma} \) (uniform for small \( \lambda \) by (6.2)), one obtains with uniform error bound
\[
T_\lambda F(x) = F(x) + \frac{1}{2} \lambda^2 \mathcal{L} F(x) + \sum_{m=3}^{M+2} \frac{\lambda^m}{m!} \mathcal{L}^{(m)} F(x) + \mathcal{O}(\lambda^{M+3}) ,
\]
which using the induction hypothesis implies for Birkhoff sums
\[
I_{\lambda,N}(\mathcal{L} F) = \sum_{m=1}^{M} \frac{2\lambda^m}{(m+2)!} I_{\lambda,N}(\mathcal{L}^{(m+2)} F) + \mathcal{O} \left( \lambda^{M+1}, \frac{1}{\lambda^2 N} \right)
= \sum_{m=1}^{M} \lambda^{m} \int d\mu \rho_{m} \mathcal{L}^{(m+2)} F + \mathcal{O} \left( \lambda^{M+1}, \frac{1}{\lambda^2 N} \right)
= \sum_{m=1}^{M} \lambda^{m} \int d\mu \rho_{m} (\mathcal{L} F) + \mathcal{O} \left( \lambda^{M+1}, \frac{1}{\lambda^2 N} \right).
\]
The last step follows from the definition (6.15) of \( \rho_{m} \). Now given any smooth function \( f \), we can write it as \( f = \int d\mu \rho_{0} f + \mathcal{L} F \) and obtain
\[
I_{\lambda,N}(f) = \int d\mu \rho_{0} f + \sum_{m=1}^{M} \lambda^{m} \int d\mu \rho_{m} \mathcal{L} F + \mathcal{O} \left( \lambda^{M+1}, \frac{1}{\lambda^2 N} \right)
= \sum_{m=0}^{M} \lambda^{m} \int d\mu \rho_{m} f + \mathcal{O} \left( \lambda^{M+1}, \frac{1}{\lambda^2 N} \right)
\]
where the last step follows from \( \int d\mu \rho_{m} = 0 \) for \( m \geq 1 \). \( \square \)

### 6.4 Extension to lowest order rotations

In this section the lowest order matrix \( R \) is an arbitrary rotation and \( E(P_{1,\sigma}) = P \) commutes with \( R \) and generates a rotation. For any \( R \in \mathcal{G} \) let us consider the associated diffeomorphism \( x \in \mathcal{M} \mapsto R \cdot x \) and its differential \( DR \). Then the push-forward of functions \( f : \mathcal{M} \to \mathbb{C} \) and vector fields \( X = (X_{x})_{x \in \mathcal{M}} \) are defined by
\[
(R_{*} f)(x) = f(R^{-1} \cdot x) , \quad (R_{*} X)_{R \cdot x} = DR_{x}(X_{x}) .
\]
The pull-back is then \( R^{*} = (R_{*})^{-1} \). With these notations, \( R_{*}(X f) = (R_{*} X)(R_{*} f) \) and
\[
R_{*}(\partial_{P}(R^{*} f)) = (R_{*} \partial_{P}) f = \partial_{R^{PR-1} f} .
\]
Furthermore we set $R_*(XY) = (R_*X)(R_*Y)$ for the composition of two vector fields $X$ and $Y$.

Now let $L$ be defined as in (6.10) (note that this is not equal to the r.h.s. of (6.8)). As $\mathcal{R}$ is a zeroth order term in $\lambda$, the Birkhoff sums are to lowest order given by averages along the orbits of $\mathcal{R}$. Furthermore the expectation of the first order term, $\lambda \mathcal{P}$, then leads to averages over the group $\langle \mathcal{P} \rangle$ to order $\lambda$. It is hence reasonable to expect that an averaged Kolmogorov operator has to be considered. In order to define it recall that there are unique, normalized Haar measures on the compact groups $\langle R \rangle$, $\langle P \rangle$ and $\langle R, P \rangle$. Averages with respect to these measures will be denoted by $E_{\langle R \rangle}$, $E_{\langle P \rangle}$ and $E_{\langle R, P \rangle}$; the integration variable will be $R$. As the Haar measure is defined by left invariance and the groups $\langle R \rangle$ and $\langle P \rangle$ commute by hypothesis, one has

$$E_{\langle R, P \rangle}(g(R)) = E_{\langle P \rangle}(\hat{g}(R))$$

for $\hat{g}(\tilde{R}) = E_{\langle R \rangle}(g(\tilde{R}R))$ and any function $g$ on $\langle R, P \rangle$. Then set

$$\hat{L} = E_{\langle R, P \rangle}(R_*L) = E_{\langle R, P \rangle}\left(\sum_{i=1}^{I} \partial_{\mathcal{P}_i R}^2 + \partial_{\mathcal{P}}^2 + 2\partial_{R\mathcal{P}_i R}^2\right)$$

(6.16)

where $\mathcal{P}_i$ are obtained by decomposing the centered random variable $\mathcal{P}_1, \sigma - \mathcal{P}$ into a sum $\sum_i v_i, \sigma \mathcal{P}_i$ such that the real coefficients satisfy $E(v_i, \sigma v_i, \sigma') = \delta_{i,i'}$ (cf. Appendix A.2). With this definition we are able to prove a result similar to Proposition 6.1.

**Proposition 6.7** Let $f \in C^\infty(\mathcal{M})$ and assume one of the following conditions to hold:

(i) $\mathcal{R}$ and $\mathcal{P}$ are diophantine and $\mathcal{M} = \mathcal{K}/\mathcal{H}$ for compact Lie groups $\mathcal{K}$ and $\mathcal{H} \subset \mathcal{K}$.

(ii) $f$ consists of only low frequencies w.r.t. $\langle \mathcal{R}, \mathcal{P} \rangle$.

Then one has

$$I_{\lambda,N}(\hat{L}f) = O\left(\frac{1}{N\lambda^2}, \lambda\right).$$

For the proof we first need the following lemma.

**Lemma 6.2** Let $f \in C^\infty(\mathcal{M})$, $f_0 = E_{\langle \mathcal{R} \rangle}(R^*f)$ and $\hat{f}_0 = E_{\langle \mathcal{P} \rangle}(R^*f_0) = E_{\langle \mathcal{R}, \mathcal{P} \rangle}(R^*f)$. If either (i) or (ii) as in the Proposition 6.7 holds, then

$$f - f_0 = g - \mathcal{R}^*g, \quad f_0 - \hat{f}_0 = \partial_\mathcal{P} E_{\langle \mathcal{R} \rangle}(R^*\hat{g}),$$

(6.17)

for smooth functions $g, \hat{g} \in C^\infty(\mathcal{M})$.

**Proof.** The group $\langle \mathcal{R} \rangle$ is isomorphic to a torus $\mathbb{T}^{L_{\mathcal{R}}}$ with isomorphism $R_{\mathcal{R}}(\theta) \in \langle \mathcal{R} \rangle$. Furthermore we define $\hat{\theta}_{\mathcal{R}}$ by $\mathcal{R} = R_{\mathcal{R}}(\hat{\theta}_{\mathcal{R}})$. If $f$ consists of only low frequencies w.r.t. $\langle \mathcal{R}, \mathcal{P} \rangle$, it can be written as finite sum of its Fourier coefficients

$$f = \sum_{\|j\| < J} f_j, \quad \text{where} \quad f_j(R_{\mathcal{R}}(\theta) \cdot x) = e^{ij^j \theta} f_j(x),$$

where $j^j$ denotes the $j$-th component of $j$. Then

$$f_0 - \hat{f}_0 = \partial_\mathcal{P} E_{\langle \mathcal{R} \rangle}(R^*\hat{g}),$$

where $\hat{g}(\tilde{R}) = E_{\langle \mathcal{R} \rangle}(g(\tilde{R}R))$ and $g$ is the function given in the lemma. 

Finally, we have

$$I_{\lambda,N}(\hat{L}f) = O\left(\frac{1}{N\lambda^2}, \lambda\right).$$

This completes the proof.
where the Fourier coefficients are calculated as in (6.6). Now set
\[ g = \sum_{0<|j|<J} \frac{f_j}{1 - e^{ij \cdot \theta_P}}. \]
This is well-defined because \( \dot{\theta}_R \) is irrational as it generates the whole torus. Then \( g - R^*g = \sum_{0<|j|<J} f_j = f - f_0. \)

As \( \langle P \rangle \) is an embedded subtorus in \( \langle R, P \rangle \), \( f_0 \) consists of only low frequencies w.r.t. \( \langle P \rangle \). Let \( R_P(\theta) \) denote the isomorphism of \( T^{L_R} \) with \( \langle P \rangle \) such that \( e^{\lambda_p} = R_P(\lambda \dot{\theta}_P) \). One can decompose \( f_0 = E_{\langle R \rangle}((R^*)f) \) into a Fourier sum w.r.t. the group \( \langle P \rangle \):
\[ f_0 = \sum_{|j|<J} \tilde{f}_j, \quad \text{where} \quad \tilde{f}_j(R_P(\theta) \cdot x) = e^{ij \cdot \theta} \tilde{f}_j(x). \]

Then
\[ \tilde{g} = \sum_{0<|j|<J} \frac{\tilde{f}_j}{ij \cdot \theta_P} \]
satisfies \( \partial_P \tilde{g} = f_0 - \tilde{f}_0 \). Furthermore \( f_0 - \tilde{f}_0 \) is invariant under \( R \) which commutes with \( P \), thus
\[ f_0 - \tilde{f}_0 = E_{\langle R \rangle}(R^* \partial_P \tilde{g}) = \partial_P E_{\langle R \rangle}(R^* \tilde{g}). \]

In case (i), \( g \) and \( \tilde{g} \) will be defined by the same formulas, but with infinite sums. Thus we have to show that these sums are well defined and that they define smooth functions on \( M \). Let \( p : \tilde{R} \to M \) be the projection identifying \( M \) with \( \tilde{R} / \tilde{g} \) and define the smooth class function \( F(K, \theta) = f(R_K(\theta) \cdot p(K)) \) on the compact Lie group \( \tilde{R} \times T^{L_R} \). We want to compare the Fourier series (6.6) of \( f \) w.r.t. \( R \) with the Fourier series of \( F \) as given by the Peter-Weyl theorem. By Theorem A.6 in Appendix A.5, this Fourier series of \( F \) is given by
\[ f(R_K(\theta) \cdot p(K)) = F(K, \theta) = \sum_{a \in W_+} \sum_{j \in \mathbb{Z}^{L_R}} d(a) \operatorname{Tr}(\mathcal{F}F(a, j) \pi_a(K)) e^{ij \cdot \theta} \]
where \( W_+ \) denotes the set of highest weight vectors of \( \tilde{R} \), \( \pi_a : \tilde{R} \to U(d(a)) \) is the \( d(a) \)-dimensional, unitary representation of \( \tilde{R} \) parameterized by \( a \), and \( \mathcal{F}F(a, j) \) is a \( d(a) \times d(a) \) matrix given by
\[ \mathcal{F}F(a, j) = \int_{\tilde{R}} dK \int_{T^{L_R}} d\theta F(K, \theta) \pi_a(K^{-1}) e^{-ij \cdot \theta}. \]

Here \( d\theta \) and \( dK \) denote the normalized Haar measures. Comparing this equation with (6.6), one obtains that the Fourier coefficients w.r.t. \( \langle R \rangle \) satisfy
\[ f_j(p(K)) = \sum_{a \in W_+} d(a) \operatorname{Tr}(\mathcal{F}F(a, j) \pi_a(K)). \]
Let \( g_j(x) = (1 - e^{ij \cdot \theta_P})^{-1} f_j \) for \( \|j\| > 0 \). The next aim is to verify that the infinite sum \( g = \sum_{\|j\|>0} g_j \) defines a smooth function on \( M \).
As $F$ is smooth, the Fourier coefficients $\mathcal{F}(a,j)$ are rapidly decreasing by [85] or Theorem A.5 in Appendix A.5, meaning that $\lim_{\|(a,j)\| \to \infty} \|(a,j)\|^h \|F(a,j)\| = 0$ for any natural $h$. Here one may choose some norm for which $\|(a,j)\| \geq \|j\|$ and $\|F(a,j)\|$ denotes the Hilbert-Schmidt norm. As $\mathcal{R}$ is diophantine, $|e^{ij \partial \mathcal{R}} - 1| \geq C \|j\|^{-s} \geq C \|(a,j)\|^{-s}$ for some natural $s$ and the coefficients $\mathcal{G}(a,j) = \left(1 - e^{ij \partial \mathcal{R}}\right)^{-1} \mathcal{F}(a,j)$ defined for $\|j\| > 0$ are still rapidly decreasing. Therefore

$$G(K, \theta) = \sum_{\|j\| > 0} \sum_{a \in \mathcal{W}_+} d(a) \text{Tr}(\mathcal{F}(a,j)\pi_a(K)) \ e^{ij \theta} = \sum_{\|j\| > 0} g_j(p(K)) \ e^{ij \theta}.$$ 

is a smooth function and the series converges absolutely and uniformly by Theorem A.5. Setting $\theta = 0$, this implies that $\sum_{\|j\| > 0} g_j$ converges uniformly to a smooth function $g$ on $\mathcal{M}$ satisfying $g - \mathcal{R}^* g = \sum_{\|j\| > 0} f_j = f - f_0$.

As before we write $f_0 = \mathbf{E}_{\mathcal{R}}(\mathcal{R}^* f)$ as sum of Fourier coefficients w.r.t. $\langle \mathcal{P} \rangle$, so $f_0 = \sum_j \tilde{f}_j$, and let $\tilde{g}_j = (ij \cdot \theta \mathcal{P})^{-1} \tilde{f}_j$ for $\|j\| > 0$. Consider the function $\tilde{G}(K, \theta) = f_0(\mathcal{R}\mathcal{P}(\theta) \cdot p(K))$ on $\mathcal{R} \times \mathbb{T}^{L \mathcal{P}}$, just as above define the Fourier coefficients $\mathcal{F} \tilde{F}(a,j)$ for $a \in \mathcal{W}_+$, $j \in \mathbb{Z}^{L \mathcal{P}}$ and let $\mathcal{F} \tilde{G}(a,j) = (ij \cdot \theta \mathcal{P})^{-1} \mathcal{F} \tilde{F}(a,j)$. As $|j \cdot \theta \mathcal{P}| \geq C \|j\|^{-s} \geq C \|(a,j)\|^{-s}$ the coefficients $\mathcal{F} \tilde{G}(a,j)$ are rapidly decreasing, the series

$$\tilde{G}(K, \theta) = \sum_{a \in \mathcal{W}_+} \sum_{j \in \mathbb{Z}^{L \mathcal{P}}} d(a) \text{Tr}\left(\mathcal{F}(a,j)\pi_a(K)\right) e^{ij \theta} = \sum_{\|j\| > 0} \tilde{g}_j(p(K)) \ e^{ij \theta}$$

converges absolutely and $\tilde{G}$ is smooth. Thus $\tilde{g} = \sum_{\|j\| > 0} \tilde{g}_j$ exists, is smooth and

$$\partial_\mathcal{P} \tilde{g} = \frac{d}{d\lambda} \bigg|_{\lambda = 0} \sum_{\|j\| > 0} \tilde{g}_j e^{\lambda ij \partial \mathcal{P}} = \sum_{\|j\| > 0} \tilde{f}_j = f_0 - \tilde{f}_0.$$ 

As $f_0 - \tilde{f}_0$ is $\mathcal{R}$-invariant one obtains also $\partial_\mathcal{P} \mathbf{E}_{\mathcal{R}}(\mathcal{R}^* \tilde{g}) = f_0 - \tilde{f}_0$. \hfill $\square$

**Lemma 6.3** If either (i) or (ii) as in the Proposition 6.7 holds, one has

$$I_{\lambda,N}(f) = I_{\lambda,N}(\mathbf{E}_{\mathcal{R},\mathcal{P}}(\mathcal{R}^* f)) + \mathcal{O}\left(\frac{1}{\lambda N}\right).$$

**Proof.** Similar as in the proof of Proposition 6.1 a Taylor expansion gives

$$\mathbf{E}_\sigma F(T_{\lambda,\sigma} \cdot x) = \mathcal{R}^* F(x) + \lambda \partial_\mathcal{P} \mathcal{R}^* F(x) + \frac{\lambda^2}{2} \mathcal{L} \mathcal{R}^* F(x) + \mathcal{O}(\lambda^3),$$

where the error term is uniform in $x$. For Birkhoff sums this implies

$$I_{\lambda,N}(F - \mathcal{R}^* F) = \lambda I_{\lambda,N}(\partial_\mathcal{P} \mathcal{R}^* F) + \frac{\lambda^2}{2} I_{\lambda,N}(\mathcal{L} \mathcal{R}^* F) + \mathcal{O}\left(\lambda^3, \frac{1}{N}\right). \quad (6.18)$$
6.4. EXTENSION TO LOWEST ORDER ROTATIONS

Using this for \( F = g \), it therefore follows that \( I_{\lambda,N}(f - f_0) = I_{\lambda,N}(g - R^*g) = \mathcal{O}(\lambda, N^{-1}) \).

The function \( F = E_{(\mathcal{R})}(R^*g) \) is \( \mathcal{R} \)-invariant, so that the l.h.s. of (6.18) vanishes, and it follows that

\[
I_{\lambda,N}(f_0 - \hat{f}_0) = I_{\lambda,N}(\partial_{\mathcal{P}} E_{(\mathcal{R})}(R^*g)) = \mathcal{O}
\left( \lambda; \frac{1}{\lambda N} \right).
\]

Combining both estimates completes the proof. \( \square \)

The following lemma is only needed for the proof of Theorem 6.2 under the hypothesis (ii).

**Lemma 6.4** For any Lie algebra element \( P \in \mathfrak{g} \), smooth function \( f \) on \( \mathcal{M} \) and any \( x \in \mathcal{M} \), the map \( \langle \mathcal{R}, \mathcal{P} \rangle \rightarrow \mathbb{C}, R \mapsto \partial_{RPR^{-1}}^{(i)} f(x) \), \( i \in \mathbb{N} \), is a trigonometric polynomial on \( \langle \mathcal{R}, \mathcal{P} \rangle \) with uniformly bounded coefficients and uniform degree in \( x \in \mathcal{M} \) (depending on \( i \) though). This implies that the function \( \mathcal{L}(E_{(\mathcal{R},\mathcal{P})}(R^*f)) \) consists of only low frequencies w.r.t. \( \langle \mathcal{R}, \mathcal{P} \rangle \).

**Proof.** As stated above, \( \langle \mathcal{R}, \mathcal{P} \rangle \subset \mathcal{G} \subset \text{GL}(L, \mathbb{C}) \) is isomorphic to \( \mathbb{T}^{L_{\mathcal{R},\mathcal{P}}} \) and the isomorphism is denoted by \( R_{\mathcal{R},\mathcal{P}}(\theta) \in \langle \mathcal{R}, \mathcal{P} \rangle \). Furthermore this group lies in some maximal torus of \( \text{GL}(L, \mathbb{C}) \). As all maximal tori are conjugate to each other, so that by exchanging \( \mathcal{G} \) with some conjugate subgroup in \( \text{GL}(L, \mathbb{C}) \) one may assume \( \langle \mathcal{R}, \mathcal{P} \rangle \) to be diagonal, i.e. it consists of diagonal matrices \( R(\theta) = \text{diag}(e^{i\varphi_1(\theta)}, \ldots, e^{i\varphi_L(\theta)}) \). Beneath the \( \varphi_1(\theta), \ldots, \varphi_L(\theta) \) there are maximally \( L_{\mathcal{R},\mathcal{P}} \) rationally independent, and each is a linear combination with integer coefficients of \( \theta_1, \ldots, \theta_{L_{\mathcal{R},\mathcal{P}}} \). Hence any trigonometric polynomial in \( \varphi(\theta) \) is a trigonometric polynomial in \( \theta \) (possibly of higher degree), that is a trigonometric polynomial on \( \langle \mathcal{R}, \mathcal{P} \rangle \).

On \( \mathfrak{g} \subset \text{gl}(L, \mathbb{C}) \) consider the usual real scalar product \( \mathbb{R} \text{e} \text{Tr}(P^*Q) = \mathbb{R} \text{e} \sum_{a,b} P_{ab} Q_{ab} \), where \( P_{ab} \) denotes the entries of the matrix \( P \). Let \( M = \dim_{\mathbb{R}}(\mathfrak{g}) \) and \( B^1, \ldots, B^M \in \mathfrak{g} \) be some orthonormal basis for \( \mathfrak{g} \) w.r.t. this scalar product. If \( R = \text{diag}(e^{i\varphi_1}, \ldots, e^{i\varphi_L}) \in \langle \mathcal{R}, \mathcal{P} \rangle \) and \( P \in \mathfrak{g} \), then one has

\[
RPR^{-1} = \sum_{m=1}^M \sum_{a,b=1}^L \mathbb{R} \text{e} \left( B^m_{ab} (RPR^{-1})_{ab} \right) B^m = \sum_{m=1}^M \sum_{a,b=1}^L \mathbb{R} \text{e} \left( B^m_{ab} P_{ab} e^{i(\varphi_a - \varphi_b)} \right) B^m,
\]

and therefore

\[
\partial_{RPR^{-1}}^{(i)} f = \left( \sum_{m=1}^M \sum_{a,b=1}^L \mathbb{R} \text{e} \left( B^m_{ab} P_{ab} e^{i(\varphi_a - \varphi_b)} \right) \partial_{B^m} \right)^i f
\]

is a trigonometric polynomial in \( \varphi \). Thus by definition of \( \mathcal{L} \), the map \( R \mapsto R_\ast(\mathcal{L}(R^*f)) = (R_\ast \mathcal{L})f \) is a trigonometric polynomial on \( \langle \mathcal{R}, \mathcal{P} \rangle \), and therefore also \( R \mapsto R_\ast(\mathcal{L} f) \) for \( f = E_{(\mathcal{R},\mathcal{P})}(R^*f) \). But this means precisely that \( \mathcal{L} \hat{f} \) consists of only low frequencies w.r.t. \( \langle \mathcal{R}, \mathcal{P} \rangle \). \( \square \)

**Proof of Proposition 6.7.** As \( \hat{\mathcal{L}} = E_{(\mathcal{R},\mathcal{P})}(R_\ast \mathcal{L}) \), it follows for \( R \in \langle \mathcal{R}, \mathcal{P} \rangle \) that \( (R_\ast \hat{\mathcal{L}})f = \hat{\mathcal{L}} f = (R^* \hat{\mathcal{L}}) f \). This implies \( R^*(\hat{\mathcal{L}} f) = \hat{\mathcal{L}} (R^* f) \) and \( E_{(\mathcal{R},\mathcal{P})}(R^* f) = \hat{\mathcal{L}} (E_{(\mathcal{R},\mathcal{P})}(R^* f)) \). Hence the Fourier coefficients of \( \hat{\mathcal{L}} f \) are given by

\[
(\hat{\mathcal{L}} f)_j = \hat{\mathcal{L}} (f_j).
\]  (6.19)
Therefore $\hat{L}f$ consists of only low frequencies w.r.t. $(R, P)$ whenever $f$ does. Furthermore one obtains for $\hat{f} = E_{(R, P)}(R^*f)$ the following equalities

$$E_{(R, P)}(R^*(\hat{L}f)) = \hat{L}\hat{f} = E_{(R, P)}(R_*(\hat{L}(R^*\hat{f}))) = E_{(R, P)}(R_*(\hat{L}\hat{f}))$$

Now $\hat{L}\hat{f}$ consists of only low frequencies by Lemma 6.4.

Thus applying Lemma 6.3 twice (the hypothesis are given either by hypothesis (i) of Proposition 6.7 or by (ii) and Lemma 6.4). One obtains

$$I_{\lambda, N}(\hat{L}f) = I_{\lambda, N}(E_{(R, P)}(R^*(\hat{L}f))) + O\left(\frac{1}{\lambda N}\right) = I_{\lambda, N}(\hat{L}\hat{f}) + O\left(\frac{1}{\lambda N}\right).$$

As $R^*\hat{f} = \hat{f}$ and $\partial_p \hat{f} = 0$, equation (6.18) for $F = \hat{f}$ implies

$$I_{\lambda, N}(\hat{L}\hat{f}) = O\left(\frac{1}{\lambda^2 N}, \lambda\right),$$

which combined with the above finishes the proof.

After these preparations the proof of Theorem 6.2 is analogous to the case $R = 1$.

**Proof of Theorem 6.2.** Consider the Markov process on $\mathcal{M}$ induced by the random family

$$T_{\lambda, \sigma} = \exp(\lambda P_{1, \sigma} + \lambda^2 P_{2, \sigma})$$

where $\sigma = (\sigma, R, \alpha) \in \hat{\Sigma} = \Sigma \times (\mathcal{R}, \mathcal{P}) \times \{-1, 1\}$ and $P_{1, \sigma} = (R P_{1, \sigma} R^{-1} - \mathcal{P}) + \alpha \mathcal{P}$, $P_{2, \sigma} = R P_{2, \sigma} R^{-1}$, and $\hat{\Sigma}$ is equipped with the probability measure $p \times dR \times \frac{1}{2}(\delta_{-1} + \delta_1)$ where $dR$ denotes the Haar measure on $(\mathcal{R}, \mathcal{P})$. As $E_{\sigma}(P_{1, \sigma}) = 0$,

$$\hat{L} = E_{(R, P)} E_{\sigma} \left(\partial^2_{P_{1, \sigma} R^{-1} - \mathcal{P}} + \partial^2_{\mathcal{P}} + 2 \partial_{R P_{2, \sigma} R^{-1}}\right) = E_{\sigma} \left(\partial^2_{P_{1, \sigma}} + 2 \partial_{P_{2, \sigma}}\right)$$

and span($\text{supp}(P_{1, \sigma})$) = span($\mathcal{P}$, span($\mathcal{P}$)), this new process leads to the operator $\hat{L}$ instead of $L$ and the whole analysis done for $L$ in the case $R = 1$, $\mathcal{P} = 0$ is applicable to $\hat{L}$ now due to the hypothesis of Theorem 6.2. In particular, $\hat{L}$ and $\hat{L}^*$ are hypoelliptic operators, the kernel of $\hat{L}$ consists of the constant functions and the kernel of $\hat{L}^*$ is one-dimensional and spanned by a normalized, smooth function $\rho_0 \geq 0$. Furthermore $C^{\infty}(\mathcal{M}) = C1_\mathcal{M} + \hat{L}C^{\infty}(\mathcal{M})$ and hence for any smooth function $f$ and $C = \int_\mathcal{M} d\mu \rho_0 f$, there is a smooth function $g$ such that $f = C + \hat{L}g$.

Assume $f$ consists of only low frequencies, i.e. $f_j = 0$ for $\|j\| > J$. Then by equation (6.19) one obtains for frequencies $\|j\| > 0$ that $f_j = (f - C)_j = \hat{L}g_j$ and hence $\hat{L}g_j = 0$ for $\|j\| \geq J$. Therefore $g_j$ is constant, which means $g_j = 0$ as $\|j\| > J > 0$ and $g$ consists of only low frequencies if $f$ does. Hence Proposition 6.7 implies for both cases (i) and (ii) the first statement of Theorem 6.2:

$$I_{\lambda, N}(f) = C + I_{\lambda, N}(\hat{L}g) = C + O\left(\frac{1}{\lambda^2 N}\right).$$
To see that the measure \( \rho_0 \mu \) is \((\mathcal{R}, \mathcal{P})\)-invariant, let again \( f \) be any smooth function. As mentioned above, there exists \( g \in C^\infty(\mathcal{M}) \) and \( C \in \mathbb{C} \) such that \( \hat{\mathcal{L}} g = f - C \). For all \( R \in (\mathcal{R}, \mathcal{P}) \), this implies \( \hat{\mathcal{L}} R^* g = R^* \hat{\mathcal{L}} g = R^* f - C \) and hence \( f - R^* f = \hat{\mathcal{L}} (g - R^* g) \in (\ker \hat{\mathcal{L}}^*)^\perp \) which gives
\[
\int_{\mathcal{M}} d\mu_0 (f - R^* f) = 0.
\]
This is precisely the stated invariance property of the measure \( \rho_0 \mu \). \( \square \)
6. RANDOM LIE GROUP ACTIONS ON COMPACT MANIFOLDS
Chapter 7

Random phase approximation for randomly coupled wires

7.1 Main result and short overview

This chapter presents an application of Theorem 6.2 of the previous chapter. We consider disordered wires, consisting of \( L \) identical subwires (all described by a one-dimensional discrete Laplacian) which are pairwise coupled by random hopping elements. In one model these hopping elements have random magnetic phases, in another one they do not. Moreover, within each wire there is a random potential like in the Anderson model. Hence the Hilbert space is \( \ell^2(\mathbb{Z}, \mathbb{C}^L) \) and a state \( \Psi \) is given by a sequence \( (\Psi_n)_{n \in \mathbb{Z}} \) of vectors \( \Psi_n \in \mathbb{C}^L \).

As before let \((\Sigma, \mathbf{p})\) be some probability space and \( W_\sigma \) be a random variable mapping from \( \Sigma \) into the hermitian \( L \times L \) matrices. Suppose that the entries \( W_{i,j} \in \mathbb{C} \), \( 1 \leq i < j \leq L \), and \( W_{k,k} \in \mathbb{R} \), \( 1 \leq k \leq L \), are independent and centered random variables with variances satisfying \( E(W_{i,j}^2) = 0 \), \( E(|W_{i,j}|^2) = 1 \), \( E(W_{k,k}^2) = 1 \). Using \( \bar{W} = W^* \) one realizes that these relations can be summarized with the equation

\[
E(W_{i,j}W_{k,l}) = \delta_{i,l} \delta_{j,k} \quad (7.1)
\]

Let \((\Omega, \mathbf{P}) = (\Sigma^\mathbb{Z}, \mathbf{p}^\mathbb{Z})\) be the product space and for \( \omega = (\sigma_n)_{n \in \mathbb{Z}} \in \Omega \) and a coupling constant \( \lambda > 0 \) define the random Hamiltonian \( H_{\lambda,\omega} \) on \( \ell^2(\mathbb{Z}, \mathbb{C}^L) \) by

\[
(H_{\lambda,\omega}\psi)_n = -\psi_{n+1} - \psi_{n-1} + \lambda W_{\sigma_n}\psi_n, \quad \psi = (\psi_n)_{n \in \mathbb{Z}} \in \ell^2(\mathbb{Z}, \mathbb{C}^L),
\]

where \( \lambda \in \mathbb{R} \) is the coupling constant and supposed to be small. This describes the disordered wire with random magnetic phases in the coupling terms. To obtain the model without magnetic phases define the random variable \( W_{\sigma}^R = \Re(W_{\sigma}) \) and the random Hamiltonian \( H_{\lambda,\omega}^R \) by

\[
(H_{\lambda,\omega}^R\psi)_n = -\psi_{n+1} - \psi_{n-1} + \lambda W_{\sigma_n}^R\psi_n, \quad \psi = (\psi_n)_{n \in \mathbb{Z}} \in \ell^2(\mathbb{Z}, \mathbb{C}^L).
\]

For the entries \( W_{i,j}^R \) of the random Matrix \( W_{\sigma}^R \) one obtains the following covariances

\[
E(W_{i,j}^RW_{k,l}^R) = \frac{1}{2} (\delta_{i,k}\delta_{j,l} + \delta_{i,l}\delta_{j,k}).
\]
We are interested in the weak coupling limit of $\lambda$. Associated to each energy $E$ there are transfer matrices and their random products are characterized by $L$ positive Lyapunov exponents, which are self averaging quantities and $\mathbb{P}$-almost surely constant. For $L = 1$ both Hamiltonians coincide with the one-dimensional Anderson model where the random potential has variance one.

As already described in previous chapters, associated to the operator $H_{\lambda,\omega}$ and an energy $E$ are transfer matrices ([9, 64], see Chapter 3) given by

$$
\hat{T}_{\lambda,\sigma}^E = \begin{pmatrix} \lambda W_{\sigma} - E & 1 \\ 1 & 0 \end{pmatrix} \in \text{Sp}(2L, \mathbb{C}),
$$

where $\text{Sp}(2L, \mathbb{C})$ denotes the hermitian symplectic group. As a first step let us do some diagonalization. Setting $E = -2 \cos(k)$ and

$$
\mathcal{C} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i1 \\ 1 & i1 \end{pmatrix}, \quad \mathcal{N} = \frac{1}{\sqrt{\sin(k)}} \begin{pmatrix} \sin(k) & 1 \\ -\cos(k) & 1 \end{pmatrix} \in \text{Sp}(2L, \mathbb{C}),
$$

(7.2)

where $|E| < 2$, $\sin(k) \neq 0$. Then one obtains

$$
\mathcal{T}_{\lambda,\sigma} = \mathcal{CN} \hat{T}_{\lambda,\sigma}^E \mathcal{N}^{-1} \mathcal{C}^* = R_k e^{\lambda P_{\sigma}} \in \text{U}(L, L) = \mathcal{C}\text{Sp}(2L, \mathbb{C}) \mathcal{C}^*,
$$

where the first equation is the definition of $\mathcal{T}_{\lambda,\sigma}$ and

$$
R_k = \begin{pmatrix} e^{-ik} & 0 \\ 0 & e^{ik} \end{pmatrix}, \quad P_{\sigma} = \frac{i}{2\sin(k)} \begin{pmatrix} W_{\sigma} & W_{\sigma} \\ -W_{\sigma} & -W_{\sigma} \end{pmatrix}.
$$

(7.3)

In the notation $\mathcal{T}_{\lambda,\sigma}$ we suppress the dependence on $E$ (resp. $k$) as this will be fixed in the considerations later. $\text{U}(L, L)$ denotes the Lorentz group consisting of all complex $2L \times 2L$ matrices $T$ satisfying

$$
T^*GT = G, \quad G = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.
$$

Note that the group generated by $R_k$ is a subgroup of the group consisting of all $R_\theta$ for $\theta \in \mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$.

The group $\text{U}(L, L)$ naturally acts on the Grassmanian flag-manifold of subspaces of $\mathbb{C}^{2L}$ on which the form $\mathcal{G} = i\mathcal{C}\mathcal{J}\mathcal{C}^*$ vanishes. This flag-manifold is diffeomorphic to a quotient $\mathcal{F} = \text{U}(L) \times \text{U}(L) / \mathcal{F}$, where $\mathcal{F}$ is isomorphic to the torus $\mathbb{T}^L$ and given by all elements $(\exp(i\Phi), \exp(i\Phi))$ where $\Phi$ is a real, diagonal matrix (see Appendix A.6). The Haar measure of $\text{U}(L) \times \text{U}(L)$ induces a Haar measure $\mu$ on the flag-manifold $\mathcal{F}$. The group action of $\text{U}(L, L)$ on $\mathcal{F}$ can be lifted to an action on $\text{U}(L) \times \text{U}(L)$ which is given by

$$
\begin{pmatrix} A & B \\ C & D \end{pmatrix} \cdot \begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} (AU + BV)S \\ (CU + DV)S \end{pmatrix}
$$

where $S$ is the unique upper triangular matrix with positive entries on the diagonal, such that $(AU + BV)S$ and $(CU + DV)S$ are unitary. If $T \cdot (\psi) = (\psi')$ in these notations then
7.2. LYAPUNOV EXPONENTS

we may also write \( T \cdot (U, V) = (U', V') \). Furthermore we will always identify functions on \( \mathbb{F} \) with class functions on \( U(L) \times U(L) \).

The corresponding object analyzed in the previous chapter is now the induced Markov process \((U_n, V_n)\) on \( \mathbb{F} \) defined by some starting point \((U_0, V_0)\) and \((U_n, V_n) = T_{\lambda, \sigma_n} \cdot (U_{n-1}, V_{n-1})\).

The random phase property now states, that the invariant measure for this Markov process should be approximately given by the Haar measure \( \mu \). For this model this is perturbatively true and can be shown by the methods of the last chapter. The methods developed there also allow to calculate the Lyapunov exponents perturbatively.

The operator \( H_{\lambda, \omega}^R \) can be treated similarly. There we get slightly different transfer matrices \( T_{\lambda, \sigma}^R \) satisfying an additional symmetry. Therefore one also has a different flag-manifold \( \mathbb{F}_R \) which is a quotient of \( U(L) \) and hence also adopted with a Haar measure \( \mu^R \) (see Appendix A.6). Thus there is also a random phase property. Now the main theorem to be proved in this chapter is the following.

**Theorem 7.1** Let \(|E| < 2\), \(E = 2 \cos(k)\) and \(E \neq 0\). If \((\nu_\lambda)\), (resp. \((\nu_\lambda^R)\)) is the family of invariant probability measures of the Markov processes for different \(\lambda\) on \(\mathbb{F}\) as defined above (or the corresponding Markov processes on \(\mathbb{F}_R\)), then one has

\[
\lim_{\lambda \to 0} \nu_\lambda = \mu, \quad \lim_{\lambda \to 0} \nu_\lambda^R = \mu^R.
\]

The \(p\)-th greatest Lyapunov exponent \(\gamma_p(E)\) for the complex Hamiltonian \(H_{\lambda, \omega}\) is given by

\[
\gamma_p(E) = \lambda^2 \frac{1 + 2(L - p)}{8 \sin^2(k)} + \mathcal{O}(\lambda^3).
\]

The \(p\)-th greatest Lyapunov exponent \(\gamma_p^R(E)\) for the real Hamiltonian \(H_{\lambda, \omega}^R\) is given by

\[
\gamma_p^R(E) = \lambda^2 \frac{1 + L - p}{8 \sin^2(k)} + \mathcal{O}(\lambda^3).
\]

Although Theorem 6.2 does not state uniqueness of the invariant measures we can speak about ‘the’ invariant measure \(\nu_\lambda\) in the theorem above as the theory of products of random matrices assures uniqueness for this model see e.g. [9].

In the next sections we will first focus on the calculations for the complex Hamiltonian \(H_{\lambda, \omega}\).

### 7.2 Lyapunov exponents

To obtain the Lyapunov exponents we first define the functions \(h_p : U(L, L) \times \mathbb{F} \to \mathbb{R}\) by

\[
h_p(T, (U, V)) = 2^{-p} \det_p \left( 1_{p \times L} U^* V^* T^* T \left( \begin{array}{c} U \\ V \end{array} \right) 1_{L \times p} \right)
\]
for \(1 \leq p \leq L\), where \(1_{p \times L} = (1, 0)\) is a \(p \times L\) matrix and \(1_{L \times p} = 1_{p \times L}^\ast\). Clearly the definition of \(h_p\) only depends on the equivalence class of \((U, V)\) modulo \(\mathcal{O}\). The functions \(h_p\) are multiplicative cocycles. In fact let \(S\) be the triangular matrix such that \(T_1(U^v) = (U^v)\) for unitary matrices \(U^v, V^v\). As \(S\) is triangular, there is an invertible upper triangular \(p \times p\) matrix \(S_p\) such that \(SL = 1_{L \times p}S_p\).

\[
h_p(T_2, (U, V)) = \frac{\det(S_p^\ast h_p(T_2, (U, V)) \det(S_p)}{\det(S_p^\ast h_p(T_1, (U, V)) \det(S_p)} h_p(T_1, (U, V)) = \frac{\det \left(1_{p \times L}(U^v, V^v)T_1^v[T_2\{U^v\}]S\right) 1_{L \times p}}{\det \left(1_{p \times L}(U^v, V^v)\right) 1_{L \times p}} h_p(T_1, (U, V)) = h_p(T_2, (U, V)) h_p(T_1, (U, V)).
\]

To get another view on this possibly strange looking function set \((U^v) = (u_1, \ldots, u_L)\) where the \(u_k\) are the corresponding column vectors. Then as the inner product on \(\Lambda^p \mathbb{C}^2L\) is given by a certain determinant one actually obtains

\[
h_p(T, (U, V)) = \|Tu_1 \wedge \ldots \wedge Tu_p\|_{\Lambda^p \mathbb{C}^2L}^2.
\]

(7.4)

Let \((U_0, V_0)\) be some starting point on the flag manifold and let \(T_{\lambda, \sigma_n}\) be the random family of i.i.d. variables on \(U(L, L)\) as defined above. Furthermore let \((U_n, V_n)\) denote the Markov process on \(\mathbb{F}\) defined by \((U_n, V_n) = T_{\lambda, \sigma_n} \cdot (U_{n-1}, V_{n-1})\). Using (7.4) the sum of the biggest \(p\) Lyapunov exponents is given by

\[
\sum_{l=1}^p \gamma_l = \lim_{N \rightarrow \infty} E_\omega \frac{1}{2N} \log h_p(T_{\lambda, \sigma_n} \cdot T_{\lambda, \sigma_1}, (U_0, V_0))
\]

\[
= \lim_{N \rightarrow \infty} E_\omega \frac{1}{2N} \sum_{n=1}^N \log h_p(T_{\lambda, \sigma_n}, (U_{n-1}, V_{n-1})) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N E_\omega \hat{h}_p(U_n, V_n),
\]

(7.5)

where \(\hat{h}_p(U, V) = \frac{1}{2} E_\sigma \log h_p(T_{\lambda, \sigma}, (U, V))\) and \(E_\omega\) denotes the integration over \(\omega = (\sigma_n)_n\) w.r.t. to \(P\). The last equation follows as \((U_{n-1}, V_{n-1})\) does not depend on \(\sigma_n\). By definition of \(T_{\lambda, \sigma}\) and because \(P_\sigma^2 = 0\), one now has

\[
\hat{h}_p(U, V) = \frac{1}{2} E_\sigma \log \det \left(1_{p \times p} + \frac{1}{2} 1_{p \times L}(U^v, V^v)\left[\lambda(P_\sigma^* + P_\sigma) + \lambda^2 P_\sigma^* P_\sigma\right] \right) 1_{L \times p}.
\]

A Taylor expansion gives

\[
\log(\det(1 + \lambda A + \lambda^2 B)) = \text{Tr}(\lambda A + \lambda^2 (B - A^2/2)) + O(\lambda^3).
\]

Furthermore one has \(E_\sigma(P_\sigma) = 0\),

\[
(U^v, V^v)(P_\sigma^* + P_\sigma) \left(\begin{array}{c} U \\ V \end{array}\right) = \frac{i}{\sin(k)} (U^v W_\sigma V - V^v W_\sigma U)
\]
7.3. EVALUATION OF BIRKHOFF SUMS

as well as
\[
E_\sigma(U^*, V^*)P_\sigma^*P_\sigma \left( \begin{array}{c} U \\ V \end{array} \right) = \frac{L}{2 \sin^2(k)} (21 + V^*U + U^*V),
\]
where we used \( E_\sigma(W^2_\sigma) = L \mathbf{1} \). Putting these things together one obtains the following:
\[
\hat{h}_p(U, V) = \frac{\lambda^2}{8 \sin^2(k)} \left[ L \text{Tr}(\mathbf{1}_{p \times L} (21 + (V^*U + U^*V)) \mathbf{1}_{L \times p}) \right. \\
+ \left. \frac{1}{2} \text{Tr} \left( E_\sigma(\mathbf{1}_{p \times L} [U^*W_\sigma V - V^*W_\sigma U] \mathbf{1}_{L \times p})^2 \right) \right] + \mathcal{O}(\lambda^3). \quad (7.6)
\]

By compactness arguments the error term is uniform in \( U, V \) and \( \sigma \). Therefore combining this with (7.5) one has
\[
\sum_{i=1}^p \gamma_p = \lambda^2 \lim_{N \to \infty} \sum_{n=1}^N E \hat{h}_p(U_n, V_n) + \mathcal{O}(\lambda^3), \quad (7.7)
\]
where \( \hat{h}_p = \frac{1}{2 \sin^2(k)} \hat{h}_p \) is a \( \lambda \)-independent function, that can be read off from equation (7.6). Furthermore \( h_p \) induces a function on \( F \). As we will see in the next sections for \( E \neq 0 \) the Birkhoff sum in (7.7) is to lowest order given by the integral of this function over the Haar measure of \( F \) which is the same as the integral of \( \hat{h}_p \) over the Haar measure on \( U(L) \times U(L) \). Using \( E_\sigma(W_\sigma P W_\sigma) = \text{Tr}(P) \mathbf{1} \) this gives
\[
\int \hat{h}_p(U, V) dU dV = \frac{1}{8 \sin^2(k)} \text{Tr} (2L \mathbf{1}_{p \times p} - p \mathbf{1}_{p \times p}) = \frac{2Lp - p^2}{8 \sin^2(k)}.
\]

7.3 Evaluation of Birkhoff sums

To calculate the Lyapunov exponents, one needs to evaluate the Birkhoff sums for the smooth functions \( \hat{h}_p, 1 \leq p \leq L \) perturbatively. As these are class functions that induce functions on \( F \), which we also denote by \( \hat{h}_p \), it is enough to consider the projected random process \((U_n, V_n) \mod \mathcal{H}) on \( F \).

The aim is to show that one can use Theorem 6.1 and to show that the corresponding lowest order invariant measure is given by the Haar measure. Then Corollary 6.1 also gives the perturbative random phase property. Therefore we need to check that the requirements of Theorem 6.2 are fulfilled. First of all \( E(\mathcal{P}_\sigma) = 0 \) and the closed group \( \langle \mathcal{R}_k \rangle \) generated by \( \mathcal{R}_k \) is compact. As \( \hat{h}_p \) is a polynomial in the coefficients of \( U \) and \( V \), the Fourier series of the function \( \theta \to \hat{h}_p(\mathcal{R}_\theta \cdot (U, V)) \) is finite, uniformly in \( (U, V) \). This implies in the terminology of the previous chapter, that the functions \( \hat{h}_p \) consist of only low frequencies w.r.t. \( \langle \mathcal{R}_k \rangle \).

Next we have to find some subgroup \( \mathcal{U} \) of \( U(L, L) \) which acts transitive on \( F \) such that its Lie algebra \( \mathfrak{u} \) is contained in the Lie algebra generated by the elements \( \mathcal{R} \mathcal{P} \mathcal{R}^{-1} \) for \( \mathcal{R} \in \langle \mathcal{R}_k \rangle \) and \( \mathcal{P} \) in the support of the distribution of the random variable \( \mathcal{P}_\sigma \). Therefore define
\[
\mathcal{U} = \{ \text{diag}(U, V) : U, V \in U(L) \text{ and } UV \in SU(L) \} \subset U(L, L),
\]
where here and below expressions like $\text{diag}(U, V)$ denote $2L \times 2L$ matrices $\begin{pmatrix} U & 0 \\ 0 & V \end{pmatrix}$. Its Lie algebra is given by

$$u = \{ \text{diag}(u, v) : u, v \in u(L) \text{ and } \text{Tr}(u + v) = 0 \} \, .$$

Identifying the pair $(U, V)$ with the matrix $\text{diag}(U, V)$ one can consider $U$ as subgroup of $U(L) \times U(L)$. As $\text{diag}(U, V) \cdot (1, 1) = (U, V)$ and as each element of $U(L) \times U(L)$ is equivalent to an element of $\mathcal{U}$ modulo $\mathcal{H}$ one obtains that the action of $\mathcal{U}$ on $\mathbb{F}$ is transitive.

Now let us check the main hypothesis of Theorem 6.2.

**Proposition 7.1** The Lie algebra $u$ is contained in the Lie algebra $v$ generated by the set

$$\{ RPR^{-1} : R \in \langle R_k \rangle, P \in \text{supp}(\sigma) \} \, .$$

**Proof.** We obtain

$$R_k P_{\sigma} R_k^{-1} = \frac{i}{2 \sin(k)} \begin{pmatrix} W_{\sigma} & e^{-2ik} W_{\sigma} \\ -e^{2ik} W_{\sigma} & -W_{\sigma} \end{pmatrix} \, .$$

Hence

$$-2 \cos(2k) P_{\sigma} + R_k P_{\sigma} R_k^{-1} + R_k^{-1} P_{\sigma} R_k = \frac{1 - \cos(2k)}{\sin(k)} \begin{pmatrix} iW_{\sigma} & 0 \\ 0 & -iW_{\sigma} \end{pmatrix} \, .$$

Therefore the space $v$ contains all matrices $\begin{pmatrix} iV & 0 \\ 0 & -iV \end{pmatrix}$ where $V = V^*$. The commutator of such two matrices is $\begin{pmatrix} [iV, \imath W] & 0 \\ 0 & [iV, \imath W] \end{pmatrix}$, hence also obtained in $v$. As $su(L)$ is a simple Lie-algebra and $iV$ and $\imath W$ are arbitrary elements of $u(L)$, the commutators $[iV, \imath W]$ contain any element of $su(L)$. Therefore, taking linear combinations of these terms shows that $u \subset v$. \hfill $\square$

The Killing form on the Lie algebra $u(2L)$ induces a bi-invariant metric on $U(2L)$ and hence on the subgroup $U(L) \times U(L)$. This metric induces a natural metric $g$ on $\mathbb{F}$, such that the projection of $U(L) \times U(L)$ on $\mathbb{F}$ is a Riemannian submersion. Let $\mu$ denote the associated volume measure on $\mathbb{F}$. Theorem 6.2 states that there is a $\mu$-almost surely positive smooth function $\rho$ on $\mathbb{F}$ such that $\rho \mu$ is a probability measure and

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{E} \hat{h}_p(U_n, V_n) = \int_{U(L) \times U(L)} \hat{h}_p(U, V) \rho(U, V) \, d\mu(U, V) + O(\lambda) \, .$$

(7.9)

Furthermore the proof in the previous chapter gives a technique to check if $\rho$ is a constant function. This is the case iff the constant function lies in the kernel of the operator that was called $\hat{L}^*$. In the situation here this means

$$\mathbb{E}_R \mathbb{E}_\sigma \left( \text{div} \left( \text{div}(\partial_{R\mathcal{P}R^{-1}_L}) \partial_{R\mathcal{P}R^{-1}_L} \right) \right) = 0 \, ,$$

where $\mathbb{E}_R$ denotes the integration of $\mathcal{R}$ over the group $\langle R_k \rangle$ w.r.t. to the Haar measure. The vectorfield $\partial_{\mathcal{P}}$ on $\mathbb{F}$ for $\mathcal{P} \in u(L, L)$ is defined by

$$\partial_{\mathcal{P}} f(U, V) = \left. \frac{d}{dt} \right|_{t=0} f(e^{i\mathcal{P} \cdot (U, V)}) \, .$$
7.4 Divergence of induced vector fields

This section is a technical preparation for the calculation of $\rho$ carried out in Section 7.5 below. The Lie algebra $u(L, L)$ is given by

$$u(L, L) = \left\{ \begin{pmatrix} A & B \\ B^* & D \end{pmatrix} : A, D \in u(L) \right\}.$$  

The paths on $U(L) \times U(L)$ which define the induced vector field for $\mathcal{P} = \begin{pmatrix} A & B \\ B^* & D \end{pmatrix} \in u(L, L)$ are given by

$$\exp(t\mathcal{P}) \cdot (U, V) = (U(1 + t[U^* AU + U^* BV + S]), V(1 + t[V^* DU + V^* B^* U + S])) + O(t^2).$$

The upper triangular matrix $S$ is determined by the fact that it has reals on the diagonal such that $U^* AU + U^* BV + S$ is in the Lie algebra $u(L)$. This leads to $S - S^* = -U^* BV - V^* B^* U$. In order to calculate $S - S^*$ let us define the following $\mathbb{R}$-linear function on $\text{Mat}(L, \mathbb{C})$,

$$w(A) = \sum_{j < k} \left[ E_{j,k} (A + A^*) E_{j,k} - E_{k,j} (A + A^*) E_{k,j} \right],$$  \hspace{1cm} (7.10)

where $E_{j,k}$ is the matrix with a one at position $j, k$ and a zero elsewhere. One obtains $S - S^* = w(U^* BV) = -w(U^* BV) \in u(L)$. This leads to

$$\exp(t\mathcal{P}) \cdot (U, V) = \begin{pmatrix} U(1 + t[U^* AU + \frac{1}{2}(U^* BV - V^* B^* U) - \frac{1}{2}w(U^* BV)]) \\ V(1 + t[V^* DU + \frac{1}{2}(V^* B^* U - U^* BV) - \frac{1}{2}w(U^* BV)]) \end{pmatrix} + O(t^2).$$

Hence we associate to the induced (lifted) vector field the function $P(U, V) = (U^* AU + \frac{1}{2}(U^* BV - V^* B^* U) - \frac{1}{2}w(U^* BV), U^* DU + \frac{1}{2}(V^* B^* U - U^* BV) - \frac{1}{2}w(U^* BV))$.

The lifted vector field on $U(L) \times U(L)$ induces a projected vector field $\partial_P$ on $\mathbb{F}$ and we want to calculate its divergence on $\mathbb{F}$. The natural metric on $u(L) \times u(L)$ induced by the Killing form on $u(2L)$ is given by $\langle (u, v)(\hat{u}, \hat{v}) \rangle = \text{Tr}(u^* \hat{u} + v^* \hat{v})$. The Lie algebra $\mathfrak{h}$ of $\mathfrak{g}$ consists of the elements $(\imath \Phi, \imath \Phi)$ for diagonal, real matrices $\Phi$. An orthonormal basis $(u_i, v_i)$ for $\mathfrak{h}^\perp$ in $u(L) \times u(L)$ is given by the matrices $\frac{1}{\sqrt{2}}(E_{j,k} - E_{k,j}, 0), \frac{1}{\sqrt{2}}(E_{j,k} + E_{k,j}, 0), \frac{1}{\sqrt{2}}(0, E_{j,k} - E_{k,j}), \frac{1}{\sqrt{2}}(0, E_{j,k} + E_{k,j})$ and $\frac{1}{\sqrt{2}}(E_{j,j} - E_{j,j})$ for $1 \leq j < k \leq L$. The derivative w.r.t. to the left-invariant vector field on $U(L) \times U(L)$ defined by $(u_i, v_i)$ will be denoted by $\delta_{(u_i, v_i)}$. According to equation (A.12) in Appendix A.7, the divergence $\text{div}(\partial_P)$ on $\mathbb{F}$ is given by

$$\sum_i \delta_{(u_i, v_i)} \langle (u_i^*, v_i^*) | P(U, V) \rangle = \begin{pmatrix} \text{Tr}(u_i^* [U^* AU + \frac{1}{2}(U^* BV - V^* B^* U) - \frac{1}{2}w(U^* BV)]) \\ \text{Tr}(v_i^* [V^* DU + \frac{1}{2}(V^* B^* U - U^* BV) - \frac{1}{2}w(U^* BV)] \end{pmatrix}.$$
Thus one has
\[ \sum_{i} \delta_{(u_i,v_i)} \text{Tr}(u_i^* U^* A U) = \text{Tr}(u_i^* (u_i^* U^* A U + U^* A U u_i)) = \text{Tr}(U^* A U (u_i^2 - u_i^2)) = 0. \]

Thus one has \( \sum_{i} \delta_{(u_i,v_i)} \text{Tr}(u_i^* U^* A U) = 0 \) and analogously \( \sum_{i} \delta_{(u_i,v_i)} \text{Tr}(v_i^* V^* D V) = 0 \). Next consider \( \sum_{i} \delta_{(u_i,v_i)} \text{Tr}((u_i + v_i)w(U^* B V)) \). It is easy to check that for \( j \neq k \) one has \( \sum_{i} u_i E_{j,k} \bar{v}_i = \sum_{i} \bar{u}_i E_{j,k} v_i = 0 \) and \( \sum_{i} u_i E_{j,k} u_i = \sum_{i} u_i E_{j,k} \bar{u}_i = E_{k,j} \). The same holds with \( v_i \) and \( u_i \) exchanged. From these equations, the cyclicity of the trace and the definition of \( w \) one obtains after some calculatory effort:

\[ \frac{1}{2} \sum_{i} \delta_{(u_i,v_i)} \text{Tr}((u_i + v_i)w(U^* B V)) = \sum_{j < k} \text{Tr}((E_{k,j} - E_{j,k})(U^* B V + V^* B^* U)) \]

The remaining term in \( \text{div}\partial_P \) is given by

\[ \frac{1}{2} \sum_{i} \delta_{(u_i,v_i)} \text{Tr}((u_i^* - v_i^*)(U^* B V - V^* B^* U)) = \frac{1}{2} \sum_{i} \text{Tr}((v_i - u_i)^2(U^* B V + V^* B^* U)) . \]

As \( \sum_{i}(v_i - u_i)^2 = -2L1 \) it follows that

\[ \text{div}(\partial_P) = 2 \Re \text{Tr}(CU^* B V), \quad (7.11) \]

where \( C = -L1 + \sum_{j < k} (E_{k,j} - E_{j,k}) = \sum_{j=1}^{L} (2j - 1 - 2L) E_{j,j} \). Note that \( \text{div}(\partial_P) \) is in fact a function on \( \mathbb{F} \), i.e. it is independent on the choice of the pre-image \( (U, V) \) because \( C \) is a diagonal matrix.

### 7.5 Volume measure to lowest order

The aim of this section is to prove \( \mathbf{E}_R \mathbf{E}_\sigma(\text{div}\text{div}(\partial_{\mathbb{R} \sigma,\mathbb{R}^{-1}}) \partial_{\mathbb{R} \sigma,\mathbb{R}^{-1}}) = 0 \) for \( E \neq 0 \). By the previous chapter this implies that the lowest order invariant measure on \( \mathbb{F} \) is given by the normalized metric measure. By construction of the metric this measure is the same as the Haar measure \( \mu \) on \( \mathbb{F} \) which is the projection of the Haar measure of \( U(L) \times U(L) \). As the group \( \langle \mathcal{R}_k \rangle \) is a closed subgroup of the torus consisting of all \( \mathcal{R}_\theta \) for \( \theta \in \mathbb{T} = \mathbb{R}/2\pi \mathbb{Z} \) the Haar measure of \( \langle \mathcal{R}_k \rangle \) can be considered as a probability measure on \( \mathbb{T} \). Expectations w.r.t. to this measure with integration variable \( \theta \in \mathbb{T} \) will be denoted by \( \mathbf{E}_\theta \). Then for any function \( f \) on \( \langle \mathcal{R}_k \rangle \), one has \( \mathbf{E}_R(f(\mathcal{R})) = \mathbf{E}_\theta(f(\mathcal{R}_\theta)) \).

**Lemma 7.1** Away from the band center for \( E \neq 0 \) one has

\[ \mathbf{E}_\theta(e^{\pm 2i\theta}) = 0, \quad \mathbf{E}_\theta(e^{\pm 4i\theta}) = 0. \]

**Proof.** If \( k \) is an irrational angle, i.e. \( \frac{1}{k} \) is irrational, then the closed group generated by \( \mathcal{R}_k \) is just the set of all \( \mathcal{R}_\theta \) and the measure \( \mathbf{E}_\theta \) is the Haar measure of the torus \( \mathbb{T} \) implying \( \mathbf{E}_\theta(e^{\pm 2i\theta}) = \mathbf{E}_\theta(e^{\pm 4i\theta}) = 0 \). If \( k \) is a rational angle, then the closed group generated by \( \mathcal{R}_k \)
is finite and has, say $s$ elements. The group consists of all $R_\theta$ such that $e^{i\theta}$ is a $s$-th root of 1. As $\sin(k) \neq 0$ we get $s > 2$. The averages over $\theta$ to be made are

$$E_\theta(e^{i2\theta}) = \frac{1}{s} \sum_{\theta=0}^{s-1} e^{i2\theta 2\pi/s}, \quad \text{and} \quad E_\theta(e^{i4\theta}) = \frac{1}{s} \sum_{\theta=0}^{s-1} e^{i4\theta 2\pi/s}.$$ 

As $s > 2$ we have $e^{i2\pi/s} \neq 1$ and the first average is zero. If $s \neq 4$ then also the second average is zero. Note, $s = 4$ which means $k = \pi/2$ and hence $E = 0$ then $E_\theta(e^{i4\theta}) = 1$. This is the origin of the Kappus-Wegner anomaly at the band center and the reason why we have to exclude $E = 0$ in the theorem.

Furthermore we will need the following identities which follow by direct computation using (7.1).

**Lemma 7.2** One has

$$E_\sigma(\text{Tr}(PW_\sigma)\text{Tr}(QW_\sigma)) = \text{Tr}(PQ), \quad E_\sigma(W_\sigma^2) = L1, \quad E_\sigma(W_\sigma Q W_\sigma) = Q^2.$$ 

Define $A_\sigma = B_\sigma = \frac{1}{2 \sin^2(k)} W_\sigma$ and $D_\sigma = -A_\sigma$. Then $R_\theta P_\sigma R_\theta^{-1} = \begin{pmatrix} A_\sigma & e^{-i\theta} B_\sigma \\ e^{i\theta} B_\sigma^* & D_\sigma \end{pmatrix}$. From now on assume $E \neq 0$. First note that \( \text{div}(\text{div}(\partial_\theta \partial_\phi)) = \text{div}(\text{div}(\partial_\theta))^2 + \partial_\theta \text{div}(\text{div}(\partial_\theta)) \) and consider the first summand using (7.11):

$$[\text{div}(\partial_\theta P_\sigma R_\theta^{-1})]^2 = e^{-i4\theta_\sigma} \text{Tr}(CU^* B_\sigma V)^2 + e^{i4\theta_\sigma} \text{Tr}(CV^* B_\sigma^* U)^2 + 2\text{Tr}(CU^* B_\sigma V)\text{Tr}(CV^* B_\sigma^* U).$$

By Lemma 7.1 and Lemma 7.2 one obtains

$$E_R E_\sigma(\text{div}(\partial_\theta P_\sigma R_\theta^{-1})(U, V))^2 = \frac{1}{2 \sin^2(k)} \text{Tr}(VCU^* UC V^*) = \frac{\text{Tr}(C^2)}{2 \sin^2(k)}.$$ 

Next we need to calculate the average of

$$\partial_\theta P_\sigma R_\theta^{-1} \text{div}(\partial_\theta P_\sigma R_\theta^{-1}) = \Re e \text{Tr}
\left[
\frac{1}{e^{2i\theta_\sigma}} 2CU^*(A_\sigma^* B_\sigma + B_\sigma D_\sigma)V + CV^* B_\sigma^* B_\sigma V

+ CU^* B_\sigma B_\sigma^* U - e^{-i4\theta_\sigma} 2U^* B_\sigma VU^* B_\sigma V - e^{-i2\theta_\sigma} U^* B_\sigma V(Cw_{\theta_\sigma}^* + w_{\theta_\sigma} C)
\right],$$

where $w_{\theta_\sigma} = w(e^{-2i\theta_\sigma} U^* B_\sigma V)$ with $w$ as defined in (7.10). Averaging over $\langle R_k \rangle$ and $\sigma$ one gets by Lemma 7.1 and Lemma 7.2

$$E_R E_\sigma(\partial_\theta P_\sigma R_\theta^{-1} \text{div}(\partial_\theta P_\sigma R_\theta^{-1})) = \frac{L \text{Tr}(C)}{2 \sin^2(k)} - E_\theta E_\sigma \Re e \left(e^{-i2\theta_\sigma} \text{Tr}(U^* B_\sigma V(Cw_{\theta_\sigma}^* + w_{\theta_\sigma} C)) \right).$$

The last term with $w_{\theta_\sigma}$ consists of terms of the form $e^{-i4\theta_\sigma} \text{Tr}(U^* B_\sigma V E_{k,j} (U^* B^l V)^t E_{k,j} C)$ and $\text{Tr}(U^* B_\sigma V E_{k,j} (U^* B^l V)^t E_{k,j} C)$. Averaging the latter one over $\sigma$ gives

$$\frac{1}{4 \sin^2(k)} \text{Tr}(U^* U E_{k,j} V^t V E_{j,k} C) = \frac{1}{4 \sin^2(k)} \text{Tr}(E_{k,j} C).$$

Therefore and by a similar result for the term with $w_{\theta_\sigma}^*$ as well as the definition of $C$, one obtains

$$E_\theta E_\sigma \text{Tr}
\left[
\sum_{j<k} \text{Tr}((E_{k,j} - E_{j,k}) C)
\right] = \frac{\text{Tr}((C + L1)C)}{2 \sin^2(k)}.$$
Putting everything together one has
\[
E_R E_\sigma \left( \text{div} \left( \text{div} (\partial_{RP_\sigma R^{-1}}) \partial_{RP_\sigma R^{-1}} \right) \right) = \frac{1}{2 \sin^2(k)} (\text{Tr}(C^2) + L \text{Tr}(C) - \text{Tr}((C + L1)C)) = 0.
\]

Therefore the lowest order invariant measure \( \rho_\mu \) on \( \mathbb{F} \) is given by the Haar measure and the equations (7.7), (7.8) and (7.9) imply for \( E \neq 0 \) and \( |E| < 2 \)
\[
\sum_{l=1}^{p} \gamma_l(E) = \lambda^2 \frac{2Lp - p^2}{8 \sin^2(k)} + \mathcal{O}(\lambda^3) \Rightarrow \gamma_p(E) = \lambda^2 \frac{2L + 1 - 2p}{8 \sin^2(k)} + \mathcal{O}(\lambda^3).
\]

This finishes the proof of Theorem 7.1 concerning the complex Hamiltonian \( H_{\lambda,\omega} \).

### 7.6 Calculations for the Hamiltonian \( H^R_{\lambda,\omega} \)

To calculate the Lyapunov exponents for the real random Hamiltonian \( H^R_{\lambda,\omega} \), one has to consider the corresponding transfer matrices
\[
T_{\lambda,\sigma}^R = \mathcal{R}_k e^{\lambda P_\sigma} \in U(L, L, \mathbb{R}) = C \text{Sp}(2L, \mathbb{R}) C^*.
\]
where \( P_\sigma^R \) is defined as \( P_\sigma \) if one replaces \( W_\sigma \) by \( W_\sigma^R \). The subgroup \( U(L, L, \mathbb{R}) \subset U(L, L) \) consists of matrices of the form \( \begin{pmatrix} A & B \\ B & A \end{pmatrix} \). They act on the real flag manifold \( \mathbb{F}_R \) which is embedded in \( \mathbb{F} \) by the equivalence classes of pairs of unitaries of the form \((U, \overline{U})\). \( \mathbb{F}_R \) is equivalent to the quotient \( U(L) / \mathcal{H}_R \) where \( \mathcal{H}_R \) is the finite group consisting of diagonal matrices with entries \( \pm 1 \) which isomorphic to \((\mathbb{Z}_2)^L\). The induced group action of \( U(L, L, \mathbb{R}) \) on \( \mathbb{F}_R \) is lifted to an action on \( U(L) \) given by
\[
\begin{pmatrix} A & B \\ B & A \end{pmatrix} \cdot U = (AU + B\overline{U}) S
\]
where \( S \) is the unique, real upper triangular matrix such that \((AU + B\overline{U})S \) is unitary.

#### 7.6.1 Lyapunov exponents

Similar as above define the Markov process \( U_n = T_{\lambda,\sigma_n}^R \cdot U_{n-1} \) on \( U(L) \). Then the Lyapunov exponents are given by
\[
\sum_{l=1}^{p} \gamma_l^R = \lim_{N \to \infty} E_\omega 1/2N \sum_{n=1}^{N} \log h_p(T_{\lambda,\sigma_n}^R, (U_{n-1}, \overline{U}_{n-1})) = \lim_{N \to \infty} 1/2N \sum_{n=1}^{N} E_\omega \hat{h}_p^R(U_n),
\]
where \( \tilde{h}^R_p(U) = \frac{1}{p} \mathbf{E}_\sigma \log h_p(\mathcal{T}^R_{\lambda}, (U, \overline{U})) \). By analogue calculations as in Section 7.2 (replacing \( V \) by \( U \)) using \( \mathbf{E}_\sigma((W^R_{\sigma})^2) = \frac{L+1}{2} \mathbf{1} \) one obtains

\[
\frac{d^2}{d\lambda^2} \bigg|_{\lambda=0} \tilde{h}^R_p(U) = \frac{1}{8 \sin^4 k} \left[ \frac{L + 1}{2} \text{Tr}(1_{p \times L}(21 + U^tU + U^*U)\mathbf{1}_{L \times p}) + \frac{1}{2} \text{Tr}(\mathbf{E}_\sigma(1_{p \times L}[U^*W^R_{\sigma}U - U^tW^R_{\sigma}]\mathbf{1}_{L \times p})^2) \right]
\]

and

\[
\sum_{l=1}^{p} \hat{n}^R_l = \lambda^2 \lim_{N \to \infty} \sum_{n=1}^{N} \mathbf{E}_n \hat{h}^R_p(U_n) + \mathcal{O}(\lambda^3).
\]

Furthermore using \( \mathbf{E}(W^R_{\sigma}PW^R_{\sigma}) = \frac{1}{2}(\text{Tr}(P)\mathbf{1} + P^t) \) one obtains

\[
\int_{U(L)} \tilde{h}^R_p(U) \, dU = \frac{2(L + 1)p - p(p + 1)}{16 \sin^2(k)}
\]

7.6.2 Haar measure to lowest order

As above we have to evaluate Birkhoff sums of the Markov process on \( \mathbb{F}_R \) or on \( U(L) \) and we want to use Theorem 6.2. Similar to the above case \( \mathbf{E}(P^R_{\sigma}) = 0 \), the group \( \langle \mathcal{R}_k \rangle \) is compact and the functions \( \tilde{h}^R_p \) consist of only low frequencies w.r.t. \( \langle \mathcal{R}_k \rangle \).

The group \( \mathcal{U}_R = \{ \text{diag}(U, \overline{U}) : U \in U(L) \} \) which is isomorphic to \( U(L) \) acts transitive on \( U(L) \) and its Lie algebra \( u_R = \{ \text{diag}(u, \overline{u}) : u \in u(L) \} \) is contained in the Lie algebra \( \mathfrak{v}_R \) generated by the set \( \{ \mathcal{R}^\mathcal{P} \mathcal{R}^{-1} : \mathcal{R} \in \langle \mathcal{R}_k \rangle, \mathcal{P} \in \text{supp}(P^R_{\sigma}) \} \). To see this consider as above \( -2 \cos(k)\mathcal{P}^R_{\sigma} + \mathcal{R}_k \mathcal{P}^R_{\sigma} \mathcal{R}_k^{-1} + \mathcal{R}_k^{-1} \mathcal{P}^R_{\sigma} \mathcal{R}_k \) to obtain that \( \mathfrak{v}_R \) contains all matrices diag(\( uV, -iV \)) where \( V = V^t \) is a real, symmetric matrix. Then \( \mathfrak{v}_R \) also contains the commutators diag(\( [iV, V], [iV, iW] \)) of such matrices. As the commutators \( [V, W] \) of real symmetric matrices span the antisymmetric matrices one obtains \( u_R \subset \mathfrak{v}_R \).

For \( \mathcal{P} = \begin{pmatrix} A & B \\ B^* & \overline{A} \end{pmatrix} \in u(L, L, \mathbb{R}) \) one has \( B^* = \overline{B} \Leftrightarrow B^t = B \) and the paths on \( U(L) \) defining the induced vectorfield \( \partial_{\mathcal{P}}^R \) are given by

\[
\exp(t\mathcal{P}) \cdot U = U(1 + t[U^*AU + \frac{1}{2}(U^*B\overline{U} - U^tB^*U) - \frac{1}{2}w(U^*B\overline{U}))
\]

Using the natural metric on \( u(L) \), the Killing form, one can calculate the divergence of this vectorfield on \( u(L) \) by the same methods as above and obtains

\[
\text{div}(\partial_{\mathcal{P}}^R) = 2\text{Re Tr}(\tilde{C}U^*B\overline{U})
\]

for \( \tilde{C} = \frac{1}{2}(- (L + 1)\mathbf{1} + \sum_{j<k}(E_{k,k} - E_{j,j})) \). Lemma 7.1 is still valid and one gets the identities

\[
\mathbf{E}(\text{Tr}(PW^R_{\sigma})\text{Tr}(QW^R_{\sigma}) = \frac{1}{2} (\text{Tr}(PQ) + \text{Tr}(PQ^t)) \quad \mathbf{E}_\sigma(W^R_{\sigma}PW^R_{\sigma}) = \frac{1}{2} (\text{Tr}(P)\mathbf{1} + P^t) \]

This leads to
\[ E_R E_\sigma \left( \text{div} (\partial_{R^{\mathbb{R} \mathbb{P}^N \mathbb{R}^{-1}}}) (U) \right)^2 = \frac{\text{Tr}(\tilde{C}^2)}{2 \sin^2(k)} \]
and
\[ E_R E_\sigma (\partial_{R^{\mathbb{R} \mathbb{P}^N \mathbb{R}^{-1}}} \text{div} (\partial_{R^{\mathbb{R} \mathbb{P}^N \mathbb{R}^{-1}}})) = E_\sigma \text{Tr}(\tilde{C}(U^* B_\sigma^R B_\sigma U U^* B_\sigma^R B_\sigma^R U)) - E_\theta E_\sigma \text{Tr}(e^{-2i\theta} U^* B_\sigma \overline{U}(C w_{\theta,\sigma}^* + w_{\theta,\sigma} C)) = \frac{L+1}{2} \frac{\text{Tr}(\tilde{C}) - \text{Tr}((\tilde{C} + \frac{L+1}{2} I)\tilde{C})}{2 \sin^2(k)} = \frac{-\text{Tr}(\tilde{C}^2)}{2 \sin^2(k)}. \]

All together gives
\[ E_R E_\sigma \left( \text{div} (\partial_{R^{\mathbb{R} \mathbb{P}^N \mathbb{R}^{-1}}}) \partial_{R^{\mathbb{R} \mathbb{P}^N \mathbb{R}^{-1}}} \right) = 0 \]
and hence the lowest order invariant measure is given by the Haar measure. By equations (7.12) and (7.13) and Theorem 2 of [71] one has
\[ \sum_{l=1}^p \gamma_l^R = \lambda^2 \frac{2(L + 1)p - p(p + 1)}{16 \sin^2(k)} + O(\lambda^3) \Rightarrow \gamma_p^R = \lambda^2 \frac{1 + L - p}{8 \sin^2(k)} + O(\lambda^3) \]
which proves the second part of Theorem 7.1.
Chapter 8

Random Dirac operators with time-reversal symmetry

8.1 Introduction

In this chapter we consider a random family of Dirac operators on \( L^2(\mathbb{R}, \mathbb{C}^{2L}) \), \( L \in \mathbb{N} \), of the form

\[
H = \mathcal{J} \partial + W + \sum_{j \in \mathbb{Z}} V_j \delta_{x_j}, \quad \mathcal{J} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},
\]

(8.1)

where the potential \( W \in L^1_{\text{loc}}(\mathbb{R}, \text{Her}(2L, \mathbb{C})) \) is a particular space-homogeneous random process described in detail below, and the \( V_j \in \text{Her}(2L, \mathbb{C}) \) are independent and identically distributed singular potentials at \( x_j = j \in \mathbb{R} \). Both potentials are supposed to satisfy time reversal symmetry

\[
\mathcal{J}^* W(x) \mathcal{J} = W(x), \quad \mathcal{J}^* V_j \mathcal{J} = V_j.
\]

(8.2)

This means that \( \mathcal{J} W(x) \) and \( \mathcal{J} V_j \) are elements of the Lie algebra \( \text{so}^*(2L) \) of the classical Lie group \( \text{SO}^*(2L) \) given by those complex \( 2L \times 2L \) matrices \( T \) satisfying \( T^* \mathcal{J} T = \mathcal{J} \) and \( T^t T = 1 \). Hence the Hamiltonian \( H \) is self-dual or in the so-called symplectic symmetry class describing time-reversal invariant particles with odd spin. Apart from this symmetry, we suppose the coupling of the potential to be efficient. This is guaranteed if the distribution of \( \mathcal{J} V_j \) has an absolutely continuous component w.r.t. the volume measure on \( \text{so}^*(2L) \), but can also be satisfied by adequate choice of \( W \) if the \( V_j \)'s vanish. A more technical formulation of the (actually much weaker) coupling hypothesis is given below in Section 8.7.

Our main new result is now:

**Theorem 8.1** Consider the random Dirac operator (8.1) with time reversal invariance (8.2) satisfying the Coupling Hypothesis on the randomness stated in Section 8.4. Let the channel number \( L \) be odd. Then almost surely \( H \) has absolutely continuous spectrum of multiplicity 2 on all of \( \mathbb{R} \) and the Landauer conductance is 1 in natural units. If the distribution of the \( \mathcal{J} V_j \) is absolutely continuous on \( \text{so}^*(2L) \), the absolutely continuous spectrum of \( H \) is almost surely pure.
Theorem 8.1 does not say anything about the singular spectrum in general (i.e. without the supplementary assumption on the distribution of the $\mathcal{V}_j$’s), but we believe it to be always empty. It is crucial that $L$ is odd, as discussed by several authors in the physics literature (please consult [25] for a long list of relevant references). We believe that for even $L$ the spectrum is almost surely pure-point, but did not try to prove this in detail (it should be possible by adapting the techniques of [50]). The main difference between the odd and even case is that there are two vanishing Lyapunov exponents in the odd case and no vanishing Lyapunov exponent in the even case. This is related to Kramer’s degeneracy of the Lyapunov spectrum and is proven in Section 8.6. Based on this fact, the proof of Theorem 8.1 goes on by applying Kotani theory for Dirac operators as developed by Sun [86] along the lines of the work by Kotani and Simon [52]. Even though most of the main identities in [86] are correct, it contains some errors which we felt necessary to correct here. Section 8.5 also generalizes the works [52, 86] to singular and complex-valued potentials. The last claim of the theorem is proven in Section 8.7 by adapting the argument of Jaksic and Last [41] once it is realized that the Dirac peaks lead to similar formulas as in rank one perturbation theory. Sections 8.2 to 8.4 contain preparatory material some of which doesn’t seem to have appeared in the literature and makes this work essentially self-contained.

Let us put Theorem 8.1 in some perspective, both from a mathematical point of view and a physical one. Most quasi-one-dimensional discrete and continuous random Schrödinger operators exhibit Anderson localization, even though some peculiarities such as in the random polymer model may lead to non-trivial quantum diffusion [45]. The situation is different for first order differential operators. For example, consider $h = 1 \otimes i\partial + v$ on $L^2(\mathbb{R}, \mathbb{C}L)$ where $v \in L^\infty(\mathbb{R}, \text{Her}(L, \mathbb{C}))$ is a bounded hermitian potential (which may be thought of as random). Then the initial value problem $\partial u = vu, u(0) = 1$, has a unique solution $u = u(x)$, which lies in the unitary group $U(L)$. Let us use it to define a unitary $U$ on $L^2(\mathbb{R}, \mathbb{C}L)$ by $(U\psi)(x) = u(x)\psi(x)$. Then $U^*hU = 1 \otimes i\partial$ showing that $h$ has absolutely continuous spectrum of multiplicity $L$ for any potential $v$. In physical terms, the operator $h$ can be thought off as an effective model for the chiral edge states of a quantum Hall system with edge conductivity $L$, and the above shows that the nature of the spectrum is conserved under perturbation by a potential, as is the Landauer conductivity which is equal to $L$ (because $U$ commutes with the position operator $X$ on $L^2(\mathbb{R}, \mathbb{C}L)$). Note that the stability of the nature of the spectrum could also be deduced from Mourre theory because $i[h, X] = 1$. For true edge states of a disordered magnetic operator on a half-plane, the proof of conservation of absolutely continuous spectrum [7, 26] and the edge conductivity [48] is much more involved, but possible.

Next let us comment on the physical relevance of the Dirac Hamiltonian (8.1) with time reversal invariance (8.2). It is believed to be an effective model for so-called helical edge states in graphene sheets with a gap at the Dirac point (opened by different hopping strengths between the two sublattices [25]). In such graphene sheets the number of edge channels with spin up and spin down is odd and hence these edge states are protected against localization. This is reflected by Theorem 8.1.
8.2 Weyl-Titchmarsh matrices

This section introduces and analyzes Weyl-Titchmarsh matrices for a fixed non-random Dirac operator with point interactions. In part this is review (compare e.g. [38]) and therefore proofs are kept short, but results need to be written out if only to fix notations. Let \( S = (x_j)_{j \in \mathbb{Z}} \) be a discrete subset of \( \mathbb{R} \) with no accumulation point and associate to each so-called singular point \( x_j \) a singular potential \( V_j \in \text{Her}(2L, \mathbb{C}) \). Furthermore let \( W \in L^1_{\text{loc}}(\mathbb{R}, \text{Her}(2L, \mathbb{C})) \). All this data is encoded in \( \omega = (W, (x_j, V_j))_{j \in \mathbb{Z}} \), but in this and the next section \( \omega \) is fixed and hence suppressed in all notations. The time-reversal symmetry (8.2) is implemented only at the end of the section. The first aim is to make mathematical sense out of \( H \) given in (8.1) as a self-adjoint operator on \( L^2(\mathbb{R}, \mathbb{C}^{2L}) \). As usual, the singular potential is dealt with as a certain self-adjoint extension. Before going on, let us point out that most results of this paper also hold for the self-adjoint operator \( R\partial + W \) where \( x \mapsto R(x) \) is bounded, invertible, and satisfies \( R^* = -R \) as well as \( \partial R = W^* - W \). In order to focus on the essential difficulties, we stick to the case \( R = J \).

Let \( W^{1,2}(\mathbb{R}/S, \mathbb{C}^{2L}) \) be the Sobolev space of functions \( L^2(\mathbb{R}/S, \mathbb{C}^{2L}) \) with square-integrable first distributional derivative. Note that these functions \( \psi \) are continuous away from \( S \) and have left and right limit values \( \psi(x \pm) = \lim_{\epsilon \downarrow 0} \psi(x \pm \epsilon) \) for all \( x \in \mathbb{R} \). First we consider the restriction \( H_0 = H|_{D(H_0)} \) to the domain

\[
D(H_0) = \left\{ \psi \in W^{1,2}(\mathbb{R}/S, \mathbb{C}^{2L}) \mid \psi(x+) = \psi(x-) = 0 \text{ for } x \in S \right\} .
\]

Then the domain of the adjoint is \( D(H_0^*) = W^{1,2}(\mathbb{R}/S, \mathbb{C}^{2L}) \). The proof of the following result is adapted from [55].

**Proposition 8.1** For \( \psi, \phi \in D(H_0^*) \), one has

\[
\langle H_0^* \psi \mid \phi \rangle - \langle \psi \mid H_0^* \phi \rangle = \sum_{x \in S} \left( \psi(x+)^* J \phi(x+) - \psi(x-)^* J \phi(x-) \right) ,
\]

where the scalar product on the l.h.s. is in \( L^2(\mathbb{R}, \mathbb{C}^{2L}) \) and those on the r.h.s. in \( \mathbb{C}^{2L} \).

**Proof.** Let \( \chi_n \in C^\infty(\mathbb{R}, [0, 1]) \) with \( \chi_n[-n,n] = 1 \), \( \chi_n[-2n,2n] = 0 \) and \( \chi'_n = \partial \chi_n \leq C \) for some constant \( C \). For any \( \phi \in D(H_0^*) \) set \( \phi_n = \chi_n \phi \). Then \( \phi_n \to \phi \) and \( H_0^* \phi_n \to H_0^* \phi \) in \( L^2(\mathbb{R}, \mathbb{C}^{2L}) \). Therefore one can calculate as follows:

\[
\langle H_0^* \psi \mid \phi \rangle - \langle \psi \mid H_0^* \phi \rangle = \lim_{n,m \to \infty} \langle H_0^* \psi_n \mid \phi_m \rangle - \langle \psi_n \mid H_0^* \phi_m \rangle
\]

\[
= \lim_{n,m \to \infty} \sum_{j \in \mathbb{Z}} \int_{x_{j-1}}^{x_j} dx \partial (\chi_n(x) \chi_m(x) (J \psi(x))^* \phi(x)) .
\]

This directly implies the proposition. \( \square \)

If \( S \) is empty, then the r.h.s. of (8.3) vanishes and this shows that \( H_0 \) is self-adjoint with domain \( W^{1,2}(\mathbb{R}, \mathbb{C}^{2L}) \). In the terminology of Weyl theory described below, this means that \( H \) is in the limit point case for any locally integrable potential \( W \). This fact also
follows from Weyl theory (more precisely, the bound (8.12) below) without reference to Proposition 8.1. If $\mathcal{S}$ is not empty, then $H_0^*$ has non-trivial deficiency spaces (which are infinite dimensional if and only if $\mathcal{S}$ is infinite). Beneath all the self-adjoint extensions of $H_0$ we are interested in those given by local boundary conditions, namely those not mixing the deficiency spaces corresponding to each of the terms on the r.h.s. of (8.3). Within the class of local boundary conditions we will choose the ones obtained by formally approximating the singular potential $\mathcal{V}_j\delta_{x_j}$ (this will be explained below), namely we consider the domain

$$D(H) = \{ \psi \in W^{1,2}(\mathbb{R}/\mathbb{S}, \mathbb{C}^{2L}) \mid \psi(x_j+) = e^{\mathcal{V}_j}\psi(x_j-) \text{ for } j \in \mathbb{N} \}. \quad (8.4)$$

Then $H = H_0^*|_{D(H)}$ clearly is an extension of $H_0$ and the identity $(e^{\mathcal{V}_j})^* \mathcal{J} e^{\mathcal{V}_j} = \mathcal{J}$ replaced in (8.3) shows that it is self-adjoint.

Now that the operator $H$ is well-defined, let us introduce the transfer matrices (or fundamental solutions) $T^\zeta(x,y) \in \text{Mat}(2L \times 2L, \mathbb{C})$, $x \geq y \in \mathbb{R}$, at a complex energy $z \in \mathbb{C}$ as the unique solutions of

$$(H - z)T^\zeta(.,y) = 0 , \quad T^\zeta(y,y) = 1_{2L} , \quad (8.5)$$

which are right-continuous in $x$ and in $y$ (for $x \geq y$) and for which $x \mapsto T^\zeta(x,y)$ is in $D(H)$. (Recall that a function is left-continuous if $f(x-) = f(x)$ for all $x$ and right-continuous if $f(x+) = f(x)$ for all $x$.) For $x < y$, we set $T^\zeta(x,y) = T^\zeta(y,x)^{-1}$. At $x_j \in \mathcal{S}$ the transfer matrices then satisfy $T^\zeta(x_j,x_j-) = e^{\mathcal{V}_j}$. The general composition rule reads for $x,u,y \in \mathbb{R}$

$$T^\zeta(x,y) = T^\zeta(x,u)T^\zeta(u,y) . \quad (8.6)$$

For later convenience we also set $T^\zeta(x) = T^\zeta(x,0)$. Now let us briefly sketch in which sense the boundary conditions in (8.4) are natural. Indeed, if $\chi_n \in C_0^\infty(\mathbb{R}, \mathbb{R})$ converges weakly to $\delta_{x_j}$ and $T_n^\zeta(x,x')$ is the transfer matrix with potential $\mathcal{V}_j\chi_n$, then taking the limit $n \to \infty$ first, one formally verifies $T_\zeta^\infty(x_j,x_j-) = e^{\mathcal{V}_j}$ which is precisely the jump condition above. Next comes the basic but crucial Wronskian identity for the transfer matrices.

**Lemma 8.1** For $a < b$ and $z, \zeta \in \mathbb{C}$,

$$T^\zeta(b-)^*\mathcal{J}T^\zeta(b-) - T^\zeta(a)^*\mathcal{J}T^\zeta(a) = (\zeta - \bar{z}) \int_a^b dx \ T^\zeta(x)^*T^\zeta(x) . \quad (8.7)$$

**Proof.** Denote the points in $\mathcal{S} \cup (a,b)$ by $x_1, \ldots, x_N$ and set $x_0 = a$ and $x_{N+1} = b$. Then $x \mapsto T^\zeta(x)$ is differentiable away from these points. Thus

$$(\zeta - \bar{z}) \int_a^b dx \ T^\zeta(x)^*T^\zeta(x) =$$

$$= \sum_{j=0}^{N} \int_{x_j}^{x_{j+1}} dx \left[ T^\zeta(x)^*((\zeta - \mathcal{W})T^\zeta(x)) - ((z - \mathcal{W})T^\zeta(x))^*T^\zeta(x) \right]$$

$$= \sum_{j=0}^{N} \left[ T^\zeta(x_{j+1}-)^*\mathcal{J}T^\zeta(x_{j+1}-) - T^\zeta(x_j+)^*\mathcal{J}T^\zeta(x_j+) \right] ,$$
where the second equality follows from the differential equation (8.5) and the fundamental theorem. Replacing $T^c(\psi(x)) = e^{\mathcal{J}V}T^c(\psi(x))$ and using $(e^{\mathcal{J}V})^* \mathcal{J} e^{\mathcal{J}V} = \mathcal{J}$, one sees that only the boundary terms remain and thus the lemma follows.

Next let us consider the restrictions of $H$ to $\mathbb{R}_+ = (0, \infty)$ and $\mathbb{R}_- = (-\infty, 0)$ given by $H_\pm = H|_{L^2(\mathbb{R}_\pm, \mathbb{C}^{2L})}$. These operators are not self-adjoint because the same calculation as above shows

$$\langle H^*_\pm \psi \mid \phi \rangle - \langle \psi \mid H^*_\pm \phi \rangle = \pm \psi(0\pm)^* \mathcal{J} \phi(\pm)$$

for $\psi, \phi \in \mathcal{D}(H^*_\pm) = \{ \psi \in W^{1,2}(\mathbb{R}_\pm/\mathbb{R}, \mathbb{C}^{2L}) \mid \psi(x_\pm) = e^{\mathcal{J}V_j}(\psi(x_\pm)) \text{ for } j \in \mathbb{N} \}$. This shows that the self-adjoint boundary conditions for $H_\pm$ are precisely given by the set $\mathbb{L}_L$ of hermitian Lagrangian planes, namely $L$-dimensional subspaces of $\mathbb{C}^{2L}$ on which the symplectic form $\mathcal{J}$ vanishes. For one such plane $\Phi \in \mathbb{L}_L$, the associated self-adjoint operator will be denoted by $H_{\pm, \Phi}$. It is well-known (see e.g. [78] for a short proof) that $\mathbb{L}_L$ is isomorphic to the unitary group $U(L)$. Thus the deficiency spaces $N^\pm_\Phi = \ker(H^*_\pm - z)$ of $H_\pm$ are $L$-dimensional.

For any analytic function $g$ we denote its complex derivative by $\partial_z g = \dot{g}$.

**Theorem 8.2** For $\Im m(z) \neq 0$ there exist unique so-called Weyl-Titchmarsh matrices $M^\pm_\Phi \in \text{Mat}(L \times L, \mathbb{C})$ such that $\ker(H^*_\pm - z)$ is spanned by the column vectors of

$$\Phi^\pm_\Phi(x) = T^c(x) \left( \begin{array}{c} 1 \\ \pm M^\pm_\Phi \end{array} \right).$$

(Here the column vectors of $\Phi^\pm_\Phi$ are considered as elements of $L^2(\mathbb{R}_\pm, \mathbb{C}^{2L})$, but below $\Phi^\pm_\Phi(x)$ is also used for all $x \in \mathbb{R}$.) They are analytic in $\mathbb{C}/\mathbb{R}$ and satisfy the Herglotz property

$$\frac{\Im m(M^\pm_\Phi)}{\Im m(z)} = \int_{\mathbb{R}_\pm} dx \Phi^*_\Phi(x) \Phi^\pm_\Phi(x) > 0,$$

where $\Im m(Z) = \frac{1}{2}(Z^* - Z)$ for any operator $Z$, as well as

$$(M^\pm_\Phi)^* = M^\mp_\Phi, \quad \dot{M}^\pm_\Phi = \int_{\mathbb{R}_\pm} dx \Phi^*_\Phi(x) \Phi^\pm_\Phi(x).$$

**Proof.** Let us consider the case of the sign $+$ and $\Im m(z) > 0$. It was argued above that the dimension of $\ker(H^*_+ - z)$ is $L$. As every solution of $H_+ \psi = z\psi$ is of the form $\psi(x) = T^c(x)v$ for some vector $v \in \mathbb{C}^{2L}$, it follows that there are $L \times L$ matrices $\alpha$ and $\beta$ such that the column vectors of

$$\begin{pmatrix} \alpha(x) \\ \beta(x) \end{pmatrix} = T^c(x) \begin{pmatrix} \alpha \\ \beta \end{pmatrix},$$

span $\ker(H^*_+ - z)$. As these vectors are, in particular, square integrable, replacing them twice in the Wronskian identity (8.7) with $b = \infty$ and $a = 0$ shows that

$$\nu(\beta^* \alpha - \alpha^* \beta) = 2 \Im m(z) \int_0^\infty dx (\alpha(x)^* \alpha(x) + \beta(x)^* \beta(x)) > 0.$$
From this it follows that both \( \alpha \) and \( \beta \) are invertible because for a vector \( v \) in the kernel of \( \alpha \) or \( \beta \) one would have \( v^*(\beta^* \alpha - \alpha^* \beta) v = 0 \). Therefore one can set \( M_+^z = \beta \alpha^{-1} \) and this also leads to the formula (8.10). The identity \( (M_+^z)^* = M_+^\bar{z} \) follows by replacing \( \zeta = \bar{\zeta} \) and \( a = 0, b = \infty \) in the Wronski identity (8.7). Finally, let us check the analyticity of \( M_+^z \) and derive the formula for its derivative. Again the Wronski identity with \( a = 0 \) and \( b = \infty \) shows for \( z \neq \zeta \) that

\[
\int_0^\infty dx \Phi_+^z(x)^* \Phi_+^z(x) = \frac{M_+^z - M_+^{\bar{z}}}{\zeta - z} .
\]

Note that the integrand on the l.h.s. is square integrable also in the limit \( \zeta \to z \) (at least for \( z \in \mathbb{C}/\mathbb{R} \), so that \( M_+^z \) is indeed holomorphic and the formula for the derivative follows. The proofs for \( M_-^z \) are similar. Let us point out though that due to our definitions the jump \( e^{i\mathcal{V}_0} \) is relevant for \( M_-^z \) if \( x_0 = 0 \in \mathbb{S} \). This is of some importance below.

As a short aside, let us sketch how the modeling of the singular potential in (8.1) by the jump conditions in (8.4) fits with the theory of extensions by von Neumann. For this purpose, let us add the singular potential \( \mathcal{V} = \mathcal{V}_0 \) at \( x_0 = 0 \) to the operator \( H \). Let \( \tilde{H}_0 \) be the restriction of \( H \) to \( \mathcal{D}(\tilde{H}_0) = \{ \psi \in \mathcal{D}(H) \mid \psi(0+) = \psi(0-) = 0 \} \). Due to Theorem 8.2 the deficiency spaces are both 2\( L \)-dimensional and given by

\[
\ker(\tilde{H}_0 - z) = \Psi_+^z \mathbb{C}^L \oplus \Psi_-^z \mathbb{C}^L ,
\]

where

\[
\Psi_\pm^z(x) = \chi(\pm x > 0) \mathcal{T}_\pm^z(x) \left( 1 \pm (\pm M_\pm^z)^2 \right)^{-\frac{1}{2}},
\]

and \( \chi \) is the indicator function. These are partial isometries \( \Psi_\pm^z : \mathbb{C}^L \to N_\pm^L \), namely \( \Psi_\pm^z(\Psi_\pm^z)^* \) is the projection on \( N_\pm^L \) and \( (\Psi_\pm^z)^* \Psi_\pm^z = 1_L \). Now the unitaries from \( \ker(\tilde{H}_0 - z) \) to \( \ker(\tilde{H}_0 - \bar{z}) \) parameterize the self-adjoint extensions of \( \tilde{H}_0 \). Using the partial isometries, these unitaries are precisely given by \( (\Psi_+^z, \Psi_-^z) U (\Psi_+^z, \Psi_-^z)^* \) where \( U \) runs through the unitary group \( U(2L) \). It is now a matter of calculation to check that

\[
U = \begin{bmatrix}
(\Psi_+^z(0+), 0) - e^{i\mathcal{V}_0}(0, \Psi_-^z(0-))
\end{bmatrix}^{-1} \begin{bmatrix}
(\Psi_+^z(0+), 0) - e^{i\mathcal{V}_0}(0, \Psi_-^z(0-))
\end{bmatrix} ,
\]

(8.11)
is well-defined (i.e. the inverse exists), is unitary and gives exactly the self-adjoint extension given by the jump condition \( \psi(0+) = e^{i\mathcal{V}_0}\psi(0-) \). Hence every local boundary condition in (8.4) is an extension within the local 2\( L \)-dimensional deficiency spaces in the sense of von Neumann. On the other hand, there are local von Neumann extensions which are not given by jump conditions (for example, those which do not couple left and right).

Even though it was already shown above that \( H \) is always self-adjoint (so that one is always in the limit point case), we now describe the Weyl theory because it gives quantitative estimates for the Weyl-Titchmarsh matrices needed below. We closely stick to the notations of [74] along the lines of which also the proofs of the results below can be given (even though there are definitely older references such as [38]). The basic idea is to study the restriction of the operator \( \tilde{H}_0 \) to \( L^2((0, x), \mathbb{C}^{2L}) \) and to analyze which initial conditions at 0 lead
to solutions satisfying any self-adjoint boundary conditions at \( x \) (there is an analogous treatment for \( H_- \)). If an adequate chart for these initial conditions is used they have the geometric structure of a matrix circle in the upper half-plane, called the Weyl surface. As \( x \) increases, this circle shrinks in a nested manner. In the so-called limit point case that one always encounters for the Dirac operators, it shrinks to a single point in the limit \( x \to \infty \) identified with the initial condition of (8.9) specified by the Weyl-Titchmarsh matrix \( M_\perp \).

This fact reflects that there is no need to fix a boundary condition at infinity in this case (the \( L^2 \)-condition takes care of it) because \( H \) is already self-adjoint.

Now comes the more technical description. Let \( \mathbb{G}_L \) be the Grassmannian of \( L \)-dimensional planes in \( \mathbb{C}^{2L} \). The chart on \( \mathbb{G}_L \) used is the stereographic projection \( \pi \) sending an \( 2L \times L \) matrix \( \begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix} \) representing the plane to \( \alpha \beta^{-1} \in \text{Mat}(L, \mathbb{C}) \). It is defined on full measure subset \( \mathbb{G}_L^\circ \subset \mathbb{G}_L \) on which the inverse of \( \beta \) exists. Then the Weyl surface at \( x \neq 0 \) is defined by

\[
\partial \mathbb{W}^x(x) = -\pi (\{ \Phi \in \mathbb{G}_L \mid T^x(x) \Phi \in \mathbb{L}_L \}) = \left\{ -M^{-1} \begin{pmatrix} 1 \\ M \end{pmatrix} \in \mathbb{L}_L \right\},
\]

where the equality follows by showing that every plane \( \Phi \) in the first set is of the form in the second one [74, Prop. 7]. Now it is useful to rewrite the condition \( T^x(x) \Phi \in \mathbb{L}_L \) in terms of the quadratic form

\[
Q^x(x) = \frac{1}{i} T^x(x)^* J T^x(x),
\]

namely \( \partial \mathbb{W}^x(x) = -\pi (\{ \Phi \in \mathbb{G}_L \mid \Phi \text{ isotropic for } Q^x(x) \}) \). The definition of \( Q^x(x) \) shows that \( Q^x(x+) = Q^x(x-) \) also for \( x \in \mathbb{S} \) so that \( Q^x(x) \) is continuous and thus \( \partial \mathbb{W}^x(x+) = \partial \mathbb{W}^x(x-) \). Item (i) and (ii) of the following properties of \( Q^x(x) \) follow from the definition and the Wronskian identity, while (iii) can be checked as in [74] once one has verified that \( T^x(x)^{-1} = J^* \mathcal{T}_x^\circ (x) J \).

**Proposition 8.2** The quadratic form \( Q^x(x) \) satisfies:

(i) \( Q^x(x) = \frac{1}{i} J + 2 \mathcal{M}(z) \left\langle T^x(\cdot) | T^x(\cdot) \right\rangle_{L^2(0, \infty, \mathbb{C}^{2L})} \)

(ii) \( \mathcal{M}(z) \partial Q^x(x) \geq 0 \)

(iii) \( Q^x(x)^{-1} = J^* \mathcal{T}_x^\circ (x) J \)

Now the radial and center operator are defined by

\[
R^x(x) = \left[ \begin{pmatrix} 1 \\ 0 \end{pmatrix}^* Q^x(x) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right]^{-1}, \quad S^x(x) = R^x(x) \begin{pmatrix} 1 \\ 0 \end{pmatrix}^* Q^x(x) \begin{pmatrix} 0 \\ 1 \end{pmatrix}.
\]

Both \( R^x(x) \) and \( S^x(x) \) are continuous in \( x \) (apart from the singularity at \( x = 0 \)). It follows from item (i) of Proposition 8.2 that \( R^x(x) > 0 \) and \(-R^\circ(x) > 0 \) for \( \mathcal{M}(z) > 0 \), and item (ii) implies \( \partial R^x(x) \leq 0 \). The terms radial and center operator are justified by the following result which can be checked by the same short calculation as in [74]. It is the basic fact of Weyl theory. Let the matrix upper half-plane \( \mathbb{U}_L \) be defined as the set of matrices \( Z \in \text{Mat}(L, \mathbb{C}) \) satisfying \( \mathcal{M}(Z) > 0 \).
Theorem 8.3 Let $\Im m(z) > 0$. Then
\[
\partial \mathbb{W}^z(x) = \left\{ S^z(x) + R^z(x)^\frac{1}{2} U(-R^\mathbb{C}(x))^\frac{1}{2} \left| U^* U = 1 \right\} \subset \mathbb{U}_L .
\]
If now the open and closed Weyl disc $\mathbb{W}^z(x)$ and $\overline{\mathbb{W}^z(x)}$ are defined by this formula with $U$ running through the set defined by $U^* U < 1$ and $U^* U \leq 1$ instead of the unitary group $U(L)$, then the Weyl surfaces are strictly nested in the sense that for $x > x' > 0$ or $x < x' < 0$
\[
\mathbb{W}^z(x) \subset \mathbb{W}^z(x'), \quad \partial \mathbb{W}^z(x') \cap \overline{\mathbb{W}^z(x)} = \emptyset .
\]

This theorem can also be used to proof the uniqueness of $M^\pm_z$ instead of the above argument based on (8.8), that is, basically the calculation in the proof of Proposition 8.1. Indeed, along the lines of Proposition 11 of [74] one can prove that there exists a constant $c$ such that
\[
\| R^z(x) \| \leq \frac{c}{|x| \Im m(z)^2} .
\]
(8.12)
This implies that $H^\pm_z$ is in the limit point case in the literal sense and that one furthermore has $-(M^\pm_z)^{-1} = \lim_{x \to \pm \infty} S^z(x)$. We need the following consequence for our purposes below. It replaces perturbative arguments in [52, 86] and hence the bounds below hold under the more natural assumptions that $\mathbb{W}$ is locally integrable. For Schrödinger operators a similar reasoning applies if they are supposed to be in the limit point case.

Corollary 8.1 There are constants $c_1, c_2$ depending only on $z$ and the $L^1_{\text{loc}}$-norm of $\mathbb{W}$ such that
\[
\| M^\pm_z \| \leq c_1 , \quad \frac{1}{c_2} \leq \frac{\Im m(M^\pm_z)}{\Im m(z)} \leq c_2 .
\]

Proof. At $x = 0$ the radial operator is infinite in the sense that $R^z(0)^{-1} = 0$. As
\[
\partial (R^z(x)^{-1}) = \Im m(z) \left( \begin{array}{c} 1 \\ 0 \end{array} \right)^* T^z(x)^* T^z(x) \left( \begin{array}{c} 1 \\ 0 \end{array} \right)
\]
is equal to $\Im m(z)1 > 0$ for $x = 0$ and is continuous in $x$ (even differentiable), it follows that $R^z(x)^{-1} > 0$ for some $x > 0$. Hence $\| R^z(x) \| < \infty$ and the Weyl disc $\overline{\mathbb{W}^z(x)}$ is compact and strictly contained in the upper half-plane $\mathbb{U}_L$. Furthermore by Theorem 8.3 the limit point $-(M^\pm_z)^{-1}$ is an element of $\overline{\mathbb{W}^z(x)}$. As $Z \mapsto -Z^{-1}$ maps compact sets of $\mathbb{U}_L$ to compact sets of $\mathbb{U}_L$ the proof is complete. 

The proof of the following result about the implementation of the symmetry is immediate.

Proposition 8.3 Suppose that $H$ is time-reversal invariant, namely satisfies (8.2). Then
\[
\mathcal{J}^* T^z(x) \mathcal{J} = T^{\mathbb{C}}(x) , \quad \overline{M^\pm_z} = -(M^\pm_z)^{-1} .
\]
8.3 Green’s function and spectral analysis

This section deals with the Green function and spectral theory of the self-adjoint operator (8.1) defined by (8.4). We always assume that $x_0 = 0 \in \mathbb{S}$, set $\mathcal{V} = \mathcal{V}_0$ and denote the operator with singular potential $\mathcal{V}$ by $H_{\mathcal{V}}$ (hoping that the reader can distinguish $H_0$ with $\mathcal{V} = 0$ from the $H_0$ in the last section).

**Proposition 8.4** Let $\Im m(z) \neq 0$ and $M^z_\pm$, $T^z(x)$ and $\Phi^z_\pm$ be associated to $H_0$ (this only leads to changes for $x < 0$ and the sign $-$). The resolvent $(H_0 - z)^{-1}$ is an integral operator with kernel

$$G^*_0(x, y) = \Phi^*_\pm(x) (-M^z_\pm - M^z_\pm)^{-1} \Phi^z_\pm(y)^* ,$$  

where the upper and lower signs are taken if $x < y$ and $x > y$ respectively. Furthermore, for a Lagrangian plane $\Phi = (1, \gamma)^*$, the resolvent $(H_{+, \Phi} - z)^{-1}$ is an integral operator with kernel

$$G^*_{+, \Phi}(x, y) = \begin{cases} 
T^z(x) \Phi (-M^z_+ + \gamma)^{-1} \Phi^z_+(y)^* , & x < y , \\
\Phi^z_-(x) (-M^z_- + \gamma)^{-1} \Phi^*_- T^z(y)^* , & x > y . 
\end{cases}$$

**Proof.** Let $G^*_0$ be defined by the formula in the theorem. Using $(M^z_\pm)^* = M^z_\pm$ one readily verifies that for all $x \in \mathbb{R}$,

$$\lim_{\epsilon \downarrow 0} \left[ G^*_0(x + \epsilon, x) - G^*_0(x - \epsilon, x) \right] = T^z(x) \mathcal{J} T^z(x)^* = \mathcal{J} ,$$

where the last equality follows by taking the inverse of $T^z(x)^* \mathcal{J} T^z(x) = \mathcal{J}$, which is the Wronskian identity (8.7) with $\zeta = \tau$, $a = 0$ and $b = x$. Therefore setting $\psi(x) = \int dy G^*_0(x, y) \phi(y)$ for a smooth function $\phi \in L^2(\mathbb{R}, \mathbb{C}^{2L})$, the definition (8.5) of the transfer matrices implies that $(H_0 - z)\psi = \phi$ because $\dd c \text{sgn} = 2\delta_{x_0}$ if $\text{sgn}$ is the sign function and $\delta_x$ is a Dirac peak at $x$. Hence $G^*_0$ is indeed the desired integral kernel. The formula for the half-sided operator is verified in a similar manner. \hfill \square

From Proposition 8.4, (8.11) and the general Krein formula for resolvents of self-adjoint extensions one could now deduce an explicit formula for the integral kernel $G_{\mathcal{V}}(x, y)$ of $H_{\mathcal{V}}$. Then lengthy algebraic calculations lead to Proposition 8.5 below, but we can also deduce it more directly based on the following idea. Both functions $x \mapsto G^*_{\mathcal{V}}(x, y)$ and $y \mapsto G^*_{\mathcal{V}}(x, y)^* = G^*_{\mathcal{V}}(y, x)$ are in the domain $\mathcal{D}(H_{\mathcal{V}})$ and satisfy respectively $(H_{\mathcal{V}} - z)G^*_{\mathcal{V}}(., y) = \delta_y$ and $(H_{\mathcal{V}} - \tau)G^*_{\mathcal{V}}(x, .) = \delta_x$. Away from $x_0 = 0$, the domain of $\mathcal{D}(H_0)$ and the identities for $H_0$ are the same. Thus a good Ansatz is

$$G^*_{\mathcal{V}}(x, y) = G^*_0(x, y) + G^*_0(x, 0+) \mathcal{K} G^*_0(0-, y) ,$$

with a matrix $\mathcal{K}$ to be determined. The jump condition $G^*_{\mathcal{V}}(0+, y) = e^{\mathcal{V} y} G^*_0(0-, y)$ gives for $y \neq 0$

$$G^*_0(0, y) + G^*(0+, 0) \mathcal{K} G^*(0, y) = e^{\mathcal{V} y} \left[ G^*_0(0, y) + G^*(0-, 0) \mathcal{K} G^*(0, y) \right] .$$
Now let us take the difference of this equation for \( y = 0+ \) and \( y = 0- \). Because \( G_\partial^\pi(0+, 0) - G^\pi(0-, 0) = \mathcal{J} \) by (8.14), one obtains

\[
\mathcal{J} + G_\partial^\pi(0+, 0) \mathcal{K} \mathcal{J} = e^{\mathcal{J}V} [\mathcal{J} + G_\partial^\pi(0+, 0) \mathcal{K} \mathcal{J}]
\]

This equation can be formally be solved for \( \mathcal{K} \), leading to the following formula.

**Proposition 8.5** Let \( \Im m(z) \neq 0 \). The resolvent \( (H_V - z)^{-1} \) is an integral operator with kernel

\[
G_V^\pi(x, y) = G_\partial^\pi(x, y) + G_\partial^\pi(x, 0) [e^{\mathcal{J}V} G_\partial^\pi(0-, 0) - G_\partial^\pi(0+, 0)]^{-1} (1 - e^{\mathcal{J}V}) G_\partial^\pi(0, y). \tag{8.15}
\]

**Proof.** It remains to check that the appearing inverse is indeed well-defined. Due to (8.13), there exist two \( L \)-dimensional planes \( \Phi_\pm \) with \( \pm \pi(\Phi_\pm) \in \mathbb{U}_L \) such that \( G_\partial^\pi(0-, 0) = \Phi_+ \Phi_-^* \) and \( G_\partial^\pi(0+, 0) = \Phi_- \Phi_+^* \). Now we claim that for any \( T \in \text{Sp}(2L, \mathbb{C}) \), in particular \( T = e^{\mathcal{J}V} \), one has \( T \Phi_+ \mathbb{C}^L \cap T \Phi_- \mathbb{C}^L = \{0\} \). This implies as desired that \( T \Phi_+ \Phi_-^* - \Phi_- \Phi_+^* \) is invertible. To prove the claim we first note that \( \pi(T \Phi_+) \in \mathbb{U}_L \) (as the Möbius transformation with a matrix in \( \text{Sp}(2L, \mathbb{C}) \) maps \( \mathbb{U}_L \) to \( \mathbb{U}_L \)) so that it is sufficient to consider the case \( T = 1 \).

Now let \( \Phi_\pm w = \Phi\pm w \) for some \( v, w \in \mathbb{C}^L \). Set \( \alpha_\pm = (1 \, 0) \Phi_\pm \) and \( \beta_\pm = (0 \, 1) \Phi_\pm \), both of which are known to be invertible. Then \( \alpha_\pm v = \alpha_\pm w \) and \( \beta_\pm v = \beta_\pm w \). Thus \( v = \beta_\pm^{-1} \beta_\pm w \) so that \( \alpha_\pm \beta_\pm^{-1} \beta_\pm w = \alpha_\pm w \). Therefore \( u = \beta_\pm w \) satisfies \( \alpha_\pm \beta_\pm^{-1} u = \alpha_\pm \beta_\pm^{-1} u \) and thus \( u^\ast \pi(\Phi_\pm) u = u^\ast \pi(\Phi_\pm) u \). By hypothesis this implies \( u = 0 \) and consequently \( w = v = 0 \). \( \square \)

Before going on let us discuss the discontinuities of the \( G_V^\pi(x, y) \) in the vicinity of the point \( (x, y) = (0, 0) \) (any other singular point can be analyzed similarly). Because \( x \mapsto G_V^\pi(x, y) \) and \( y \mapsto G_V^\pi(x, y)^\ast = G_V^\pi(y, x) \) are in the domain \( \mathcal{D}(H_V) \), the singular potential leads to jumps on the lines \( x = 0 \) and \( y = 0 \). According to (8.14) there is furthermore a jump by \( \mathcal{J} \) on the diagonal \( x = y \). Away from these 3 lines crossing at the origin, \( G_V^\pi(x, y) \) is continuous. Hence there are 6 directional limits as \((x, y) \to (0, 0)\). Enumerate them by \( G_1, \ldots, G_6 \) in a clockwise direction starting with \( G_1 = \lim_{\epsilon \downarrow 0} G_V^\pi(\epsilon, 2\epsilon) \).

Setting \( T = e^{\mathcal{J}V} \) one then has

\[
G_2 = G_1 + \mathcal{J}, \quad G_3 = G_2(T^{-1})^\ast, \quad G_4 = T^{-1} G_3, \quad G_5 = G_4 - \mathcal{J}, \quad G_6 = G_5 T^\ast, \quad G_1 = T G_6.
\]

Note that these relations are indeed cyclic because \( T \in \text{Sp}(2L, \mathbb{C}) \). By (8.13) each of the \( G_j \) has rank \( L \). The following proposition shows that, however, an adequate linear combination is a Herglotz function and, in particular, of full rank \( 2L \).

**Proposition 8.6** Let us define the averaged Green matrix

\[
\hat{G}_V^\pi(x) = \lim_{\epsilon \downarrow 0} \left[ \frac{1}{4} G_V^\pi(x + \epsilon, x - \epsilon) + \frac{1}{4} G_V^\pi(x - \epsilon, x + \epsilon) + \frac{1}{8} G_V^\pi(x, x + 2\epsilon) + \frac{1}{8} G_V^\pi(x + 2\epsilon, x + \epsilon) + \frac{1}{8} G_V^\pi(x - \epsilon, x - 2\epsilon) + \frac{1}{8} G_V^\pi(x - 2\epsilon, x - \epsilon) \right].
\]

Then \( z \in \mathbb{U}_1 \mapsto \hat{G}_V^\pi(x) \in \text{Mat}(2L, \mathbb{C}) \) is a Herglotz function for any \( x \in \mathbb{R}/\mathbb{S} \) and has non-negative imaginary part for \( x \in \mathbb{S} \). It satisfies

\[
\Im m(\hat{G}_V^\pi(0)) = \frac{1}{4} (1 + e^{\mathcal{J}V}) \Im m(\hat{G}_V^\pi(0^+)) (1 + e^{\mathcal{J}V})^\ast. \tag{8.16}
\]
8.3. GREEN’S FUNCTION AND SPECTRAL ANALYSIS

Proof. Let us note that for \( x \notin S \) the definition of the averaged Green matrix reduces to \( \hat{G}_\nu(x) = \frac{1}{2}(G_\nu^+(x, x) + G_\nu^-(x, x)) \). For sake of notational simplicity, let us focus on the case \( x = 0 \) with \( V \neq 0 \) modeling \( x \in S \). With the above notations, then by definition \( \hat{G}_\nu^-(0) = \frac{1}{2}(G_1 + G_2 + 2G_3 + G_4 + G_5 + 2G_6) \) which is a weighting of the \( G_j \) according to the area of the corresponding octant or quadrant. Now let \( z = E + i\epsilon \) with \( \epsilon > 0 \) and consider the positive operator \( \Im((H_V - z)^{-1}) = \epsilon((H_V - E)^2 + \epsilon^2)^{-1} \). For any \( \varphi \in L^2(\mathbb{R}, \mathbb{C}^{2L}) \), one thus has

\[
0 < \langle \varphi | \Im((H_V - z)^{-1}) | \varphi \rangle = \frac{1}{2\pi} \int dx \int dy \varphi(x)^* (G_\nu^+(x, y) - G_\nu^-(x, y)) \varphi(y) .
\]

Now let \( \chi_k \in C^K(\mathbb{R}) \) be a positive approximate unit, that is \( w\text{-}\lim_{k \to \infty} \chi_k = \delta_0 \). For any function \( f : \mathbb{R}^2 \cong \mathbb{C} \to \mathbb{C} \) having the directional limits \( f(\theta) = \lim_{r \to 0} f(re^{i\theta}) \), it follows that \( \int dx \int dy \chi_k(x)\chi_k(y) f(x, y) \) converges to \( \hat{f} = \int_0^{2\pi} \frac{d\theta}{2\pi} f(\theta) \). Hence, for \( \varphi_k = \chi_k v \) with \( v \in \mathbb{C}^{2L} \),

\[
0 \leq \lim_{k \to \infty} \langle \varphi_k | \Im((H_V - z)^{-1}) | \varphi_k \rangle = \frac{1}{2\pi} v^* \left( \hat{G}_\nu^+(0) - \hat{G}_\nu^-(0)^* \right) v .
\]

This proves that the imaginary part is non-negative. The Herglotz property for \( 0 \notin S \), namely that the imaginary part is positive, follows from the concrete formula

\[
\hat{G}_0^-(0) = \begin{pmatrix}
(-M_+ - M_-)^{-1} & (-M_+ - M_-)^{-1}(M_+ - M_-) \\
(M_+ - M_-)(-M_+ - M_-)^{-1} & (M_+^{-1} + M_-^{-1})^{-1}
\end{pmatrix}
\]

following from Proposition 8.4, and the Herglotz property of \( M_\pm \) by the Liouville theorem. As the singular points are discrete, there is interval \((0, \epsilon)\) not containing any. Hence \( \hat{G}_\nu(0+) = \frac{1}{2}(G_1 + G_2) \). It is now a matter of an algebraic calculation to verify the second formula. \( \square \)

As for any Herglotz function with sufficient decay properties such as \( \hat{G}_\nu(x) \), there is associated a matrix valued measure \( \mu_x \) on \( \mathbb{R} \) and a self-adjoint matrix \( A_x = A_x^* \) independent of \( z \) (see [31] for a review and properties) such that

\[
\hat{G}_\nu(x) = A_x + \int \mu_x(dE) \left( \frac{1}{E - z} - \frac{1}{1 + E^2} \right) .
\]

Because

\[
\hat{G}_\nu(x) = T^*(x, y) \hat{G}_\nu^-(y) T^*(x, y)^* \]

for \( x, y \notin S \) and \( T^*(x, y) \) is analytic and invertible, the measures \( \mu_x \), \( x \notin S \), all define the same measure class. According to (8.16), the measure \( \mu_0 \) associated to \( \hat{G}_\nu^-(0) \) is also in the same measure class as long as \(-1\) is not in the spectrum of \( e^{J_V} \). We skip the proof of the following result, showing in which sense \( \mu_x \) can rightfully be called a spectral measure of \( H_V \) (see [52]).
Proposition 8.7 Let $\psi, \phi \in L^2(\mathbb{R}, \mathbb{C}^{2L})$ and $f \in C_0(\mathbb{R})$. Then, whenever $\mu_x$ is in the almost sure measure class,
\[
\langle \psi | f(H_V) | \phi \rangle = \int_{\mathbb{R}} f(E) \left( \int dy T^E(y, x)^* \psi(y) \right)^* \mu_x(dE) \left( \int dy T^E(y, x)^* \phi(y) \right),
\]
and the functions of $E$ in the parenthesis are in $L^2(\mathbb{R}, \mu_x)$.

The arguments in Section 8.7 will be based on the following perturbative formula for the averaged Green matrix w.r.t. the finite rank perturbation given by the singular potential $V_0$. For notational convenience let us set $G^z_V = G^z_V(0)$. Furthermore let us introduce the Cayley transform of $V$ by
\[
\hat{V} = 2J(e^{iV} + 1)^{-1}(e^{iV} - 1),
\]
whenever the inverse is well-defined. One readily checks that $\hat{V}^* = \hat{V}^*$ and that $J^*\hat{V}^*J = \hat{V}$ if $J^*VJ = V$.

Proposition 8.8 The averaged Green matrix satisfies (even if $\hat{V}$ is not well-defined)
\[
\hat{G}^z_V = \left( (\hat{G}^z_0)^{-1} + \hat{V} \right)^{-1},
\]
and
\[
\Im m(\hat{G}^z_V) = \left( \left[ 1 + \hat{V} \hat{G}^z_0 \right]^{-1} \right)^* \Im m(\hat{G}^z_0) \left[ 1 + \hat{V} \hat{G}^z_0 \right]^{-1}.
\]

Proof. Let us apply the averaging procedure of Proposition 8.6 to (8.15). This gives
\[
\hat{G}^z_V = \hat{G}^z_0 + \hat{G}^z_0 K \hat{G}^z_0 = \hat{G}^z_0 \left( 1 + K \hat{G}^z_0 \right),
\]
where $K = \left[ e^{iV} G^z_0(0-, 0) - G^z_0(0+, 0) \right]^{-1} (1 - e^{iV})$ as before. Because both $\hat{G}^z_V$ and $\hat{G}^z_0$ are invertible, it follows that also $\left( 1 + K \hat{G}^z_0 \right)$ is invertible. Hence
\[
\hat{G}^z_V = \hat{G}^z_0 + \hat{G}^z_0 \left( 1 + K \hat{G}^z_0 \right)^{-1} K \hat{G}^z_0 = \left[ (\hat{G}^z_0)^{-1} - (1 + K \hat{G}^z_0)^{-1} K \right]^{-1}. \tag{8.21}
\]
Using $G^z_0(0\pm, 0) = \hat{G}^z_0 \pm \frac{i}{2}J$, one readily checks $\left( 1 + K \hat{G}^z_0 \right)^{-1} K = -\hat{V}$ completing the proof of (8.19). That of (8.20) is straightforward. \qed

The average Green matrix always satisfies $(\hat{G}^z_V)^* = \hat{G}^z_V$. Furthermore one has the following.

Proposition 8.9 If $H_V$ has time-reversal symmetry, the averaged Green matrix satisfies
\[
J^* \hat{G}^z_V J = (\hat{G}^z_V)^t, \quad J^* \Im m(\hat{G}^z_V) J = \Im m(\hat{G}^z_V). \tag{8.22}
\]
If furthermore $\phi = (v, J\overline{v})$ for some $v \in \mathbb{C}^{2L}$ satisfying $v^*J\overline{v} = 0$, then the $2 \times 2$ matrix $\phi^* \hat{G}^z_V \phi$ is a constant multiple of the identity.
8.4. STOCHASTIC DIRAC OPERATORS

Proof. The Hamiltonian satisfies \( \mathcal{J}^* H \mathcal{J} = \overline{H} \mathcal{J} \) so that \( \mathcal{J}^*(H \mathcal{J} - z)^{-1} \mathcal{J} = (\overline{H} \mathcal{J} - z)^{-1} \).
This implies that for any vectors \( v, w \in \mathbb{C}^{2L} \), \( v^* \mathcal{J}^* G_{\mathcal{J}}(w) = \overline{w}^* \mathcal{J}^* G_{\mathcal{J}}(v) = v^*(\mathcal{J}^* G_{\mathcal{J}} \mathcal{J})^w \) which implies the first identity in (8.22), from which the second one can be directly deduced.
As to the last point, for any vector \( w \) one has \( w^* G_{\mathcal{J}}^w w = w^*(\mathcal{J}^* G_{\mathcal{J}}^w \mathcal{J})^w = (\mathcal{J}^* G_{\mathcal{J}}^w \mathcal{J})^w. \) Moreover, for any \( w = \lambda v + \lambda' \mathcal{J} v \in \text{Ran}(\phi) \), one checks the orthogonality \( w^* \mathcal{J} \overline{w} = 0. \) These facts imply \( w^* G_{\mathcal{J}}^w w = \frac{1}{2} \text{Tr}(\phi^* G_{\mathcal{J}}^w \phi) \|w\|^2. \) \( \square \)

8.4 Stochastic Dirac operators

In this section we introduce stochastic Dirac operators and state a few of their elementary properties, then introduce the random Dirac operators and give a precise statement of the main coupling hypothesis needed in Theorem 8.1. Let be given a compact dynamical system \((\Omega, \mathcal{P}, T)\) where \( T \) is a continuous \( \mathbb{R} \)-action on the compact space \( \Omega \) w.r.t. which the probability measure \( \mathcal{P} \) is supposed to be ergodic. Then \((H(\omega))_{\omega \in \Omega}\) is called a family of stochastic Dirac operators if each \( H(\omega) \) is of the form (8.1) and the map \( \omega \in \Omega \mapsto H(\omega) \) is strongly continuous in the resolvent sense and covariant, that is, if \( U_x \) denotes the right shift by \( x \) on \( L^2(\mathbb{R}, \mathbb{C}^{2L}) \), then \( U_x(H(\omega) - z)^{-1} U_x^* = (H(T_x \omega) - z)^{-1} \). Each point \( \omega \in \Omega \) is thought of as a configuration, incorporating the positions \( S \) and values \((\mathcal{V}_x)_{x \in S}\) of the singular potential as well as the potential \( \mathcal{W} \). Thus \( S \) is an \( \mathbb{R} \)-ergodic point process. Its density is denoted by \( \rho_S \). The locally integrable potential associated to a given configuration \( \omega \) is then \( \mathcal{W}_\omega(x) = \mathcal{W}(T_x \omega), \ x \in \mathbb{R}, \) where the \( \mathcal{W} \) is a matrix-valued function on \( \Omega \). Hence we suppose this function \( \mathcal{W} \) to be locally integrable along orbits with a uniform bound on the \( L^1 \text{loc} \)-norm.

Now all objects such as transfer matrices, Weyl-Titchmarsh matrices and Green matrices analyzed in the sections above depend on \( \omega \); however, in the notations this will be made explicit by a supplementary argument only if necessary. Let us introduce some notations for the \( L \times L \) matrix entries of the potential:

\[
\mathcal{W} = \begin{pmatrix} P & R \\ R^* & Q \end{pmatrix}, \quad e^{3\mathcal{V}} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}.
\]

All these objects are random and for \( \mathcal{V} = \mathcal{V}_x, \ x \in S, \) the entries are also denoted \( A_x, B_x, C_x, D_x \). As the matrix \( e^{3\mathcal{V}} \) is in \( \text{Sp}(2L, \mathbb{C}) \), it is well-known that the inverse in the definition of the Möbius transformation

\[
\begin{pmatrix} A & B \\ C & D \end{pmatrix} \cdot Z = (AZ + B)(CZ + D)^{-1},
\]

exists whenever \( Z \) is in the upper or lower half-plane, i.e. \( \pm 3m(Z) > 0 \). If then \( W = e^{3\mathcal{V}} \cdot Z = (AZ + B)(CZ + D)^{-1} \), also \( W \) is in the upper or lower half-plane respectively and one has \( Z = (e^{3\mathcal{V}})^{-1}, W = (D^* - B^* W)(-C^* + A^* Z)^{-1} \). Now we can collect a few first properties of the transfer matrices and the Weyl-Titchmarsh matrices.
Lemma 8.2 Let $\Im m(z) \neq 0$, set

\[
\begin{pmatrix}
\alpha_\pm^z(x, \omega) \\
\beta_\pm^z(x, \omega)
\end{pmatrix} = T^z(x, \omega) \begin{pmatrix}
1 \\
\pm M^\pm_\pm(\omega)
\end{pmatrix} = \Phi^\pm_\pm(x, \omega). \tag{8.23}
\]

(i) The transfer matrices satisfy the cocycle equation

\[
T^z(x + y, \omega) = T^z(x, T_{-y}\omega) T^z(y, \omega), \quad T^z(0, \omega) = 1.
\]

(ii) One has

\[
\begin{pmatrix}
\alpha_\pm(x + y, \omega) \\
\beta_\pm(x + y, \omega)
\end{pmatrix} = \begin{pmatrix}
\alpha_\pm(x, T_{-y}\omega) \\
\beta_\pm(x, T_{-y}\omega)
\end{pmatrix} \alpha_\pm(y, \omega).
\]

In particular, $\alpha_\pm^z(x, \omega)$ is a cocycle:

\[
\alpha_\pm^z(x + y, \omega) = \alpha_\pm^z(x, T_{-y}\omega) \alpha_\pm^z(y, \omega), \quad \alpha_\pm^z(0, \omega) = 1.
\]

(iii) $M^\pm_\pm(T_{-x}\omega) = \pm \beta_\pm(x, \omega) \alpha_\pm(x, \omega)^{-1}$.

(iv) The map $x \mapsto M^\pm_\pm(T_x\omega)$ is differentiable away from $\mathbb{S}$. It is left-continuous and for $-x \in \mathbb{S}$,

\[
\pm M^\pm_\pm(T_{x+}\omega)^{-1} = (e^{J V_\alpha})^{-1} : (\pm M^\pm_\pm(T_x\omega)^{-1}) \cdot
\]

(v) The maps $y \mapsto \alpha_\pm^z(x, T_y\omega)$ and $y \mapsto \beta_\pm^z(x, T_y\omega)$ are left-continuous. For $-y \in \mathbb{S}$,

\[
\begin{pmatrix}
\alpha_\pm(x, T_{y+}\omega) \\
\beta_\pm(x, T_{y+}\omega)
\end{pmatrix} = (e^{J V_\alpha})^{-1} \begin{pmatrix}
\alpha_\pm(x, T_{y}\omega) \\
\beta_\pm(x, T_{y}\omega)
\end{pmatrix} (D^\pm_{-y} - B^\pm_{-y} (\pm M^\pm_\pm(T_y\omega)))^{-1}.
\]

(vi) The map $x \in \mathbb{R}_+ \mapsto \alpha_\pm^z(x, \omega)$ is right-continuous. If $x \in \mathbb{S}$,

\[
\alpha_\pm^z(x, \omega) = (A_x \pm B_x M^\pm_\pm(T_{-x+}\omega)) \alpha_\pm(x-, \omega) = (D^\pm_x \mp B_x^\pm M^\pm_\pm(T_{-x}\omega))^{-1} \alpha_\pm(x-, \omega).
\]

(vii) $\partial_x \alpha_\pm^z(x, \omega) = [-R(T_{-x}\omega)^\pm \mp (Q(T_{-x}\omega) - z) M^\pm_\pm(T_{-x}\omega)] \alpha_\pm^z(x, \omega)$ for $x \notin \mathbb{S}$

(viii) The following Riccati equation holds for $x \notin \mathbb{S}$

\[
\pm \partial_x M^\pm_\pm(T_{-x}\omega) = \begin{pmatrix}
1 \\
\pm M^\pm_\pm(T_{-x}\omega)
\end{pmatrix}^* (\mathcal{W}(T_{-x}\omega) - z) \begin{pmatrix}
1 \\
\pm M^\pm_\pm(T_{-x}\omega)
\end{pmatrix}.
\]

Proof. (i), (ii) and (iii) follow immediately from (8.6) and (8.23). It is clearly sufficient to analyze the directional continuity in (iv) and (v) for the case $x = 0 \in \mathbb{S}$. Let $\epsilon > 0$. Using the composition rule for transfer matrices and their translation property

\[
T^z(x + \epsilon, y + \epsilon, T_\epsilon\omega) = T^z(x, y, \omega),
\]

one deduces

\[
T^z(x, \omega) = T^z(x + \epsilon, x, \omega)^{-1} T^z(x, T_{-\epsilon}\omega) T^z(\epsilon, 0, \omega).
\]
8.4. STOCHASTIC DIRAC OPERATORS

Taking the limit $\epsilon \downarrow 0$ gives $T^\pm(x, \omega) = T^\pm(x, T_{0-\omega})$ which implies $M^\pm_\pm(T_{0-\omega}) = M^\pm_\pm(\omega)$. Similarly, the limit $\epsilon \downarrow 0$ of

$$T^\pm(x, \omega) = T^\pm(x, x - \epsilon, \omega) T^\pm(x, T_\epsilon \omega) T^\pm(0, -\epsilon, \omega)^{-1},$$

leads to

$$T^\pm(x, \omega) = e^{\mathcal{J}V_\epsilon} T^\pm(x, T_{0+\omega}) (e^{\mathcal{J}V_0})^{-1}.$$ 

As the jump at $x$ does not effect the square-integrability in (8.9), this implies that

$$(e^{\mathcal{J}V_0})^{-1} \begin{pmatrix} 1 \\ \pm M^\pm_\pm(\omega) \end{pmatrix} N = \begin{pmatrix} 1 \\ \pm M^\pm_\pm(T_{0+\omega}) \end{pmatrix},$$

for some invertible $L \times L$ matrix $N$. The upper entry implies $N = (D^*_0 - B^*_0(\pm M^\pm_\pm(\omega)))^{-1}$, the lower one

$$\pm M^\pm_\pm(T_{0+\omega}) = (-C^*_0 \pm A^*_0 M^\pm_\pm(\omega))(D^*_0 \mp B^*_0 M^\pm_\pm(\omega))^{-1}. \quad (8.24)$$

This is precisely the equation claimed in (iv) in the case $x = 0$. (v) follows from (8.23) and the last 4 identities. For (vi) we use $T^\pm(x, \omega) = e^{\mathcal{J}V_\epsilon} T^\pm(x, -\omega)$ for $x > 0$, giving

$$\Phi^\pm(x, \omega) = e^{\mathcal{J}V_\epsilon} \Phi^\pm(x, -\omega) = e^{\mathcal{J}V_\epsilon} \begin{pmatrix} 1 \\ \pm M^\pm_\pm(T_{(x-\epsilon)\omega}) \end{pmatrix} \alpha^\pm_\pm(x, \omega),$$

where (iii) was used in the second equality. The upper entry of this identity gives the first equality of (vi). The second one follows by replacing (8.24) and using $A_x D^*_x - B_x C^*_x = 1$ and $A_x B^*_x = B_x A^*_x$. The following calculation gives (vii):

$$\partial_x \alpha^\pm_\pm(x, \omega) = (1 \ 0) \partial_x T^\pm(x, \omega) \begin{pmatrix} 1 \\ \pm M^\pm_\pm(\omega) \end{pmatrix}$$

$$= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mathcal{J}^*(z - \mathcal{W}(T_{-\omega})) T^\pm(x, \omega) \begin{pmatrix} 1 \\ \pm M^\pm_\pm(\omega) \end{pmatrix}$$

$$= \begin{pmatrix} 0 \\ 1 \end{pmatrix} (z - \mathcal{W}(T_{-\omega})) \begin{pmatrix} 1 \\ \pm M^\pm_\pm(T_{-\omega}) \end{pmatrix} \alpha^\pm_\pm(x, \omega).$$

Finally,

$$\partial_x M^\pm_\pm(T_{-\omega}) = (0 \ 1) \partial_x \left[ T^\pm(x, \omega) \begin{pmatrix} 1 \\ \pm M^\pm_\pm(\omega) \end{pmatrix} \alpha^\pm_\pm(x, \omega)^{-1} \right]$$

$$= \begin{pmatrix} 0 \\ 1 \end{pmatrix} (\mathcal{W}(T_{-\omega}) - z) \begin{pmatrix} 1 \\ \pm M^\pm_\pm(T_{-\omega}) \end{pmatrix} \mp M^\pm_\pm(T_{-\omega}) \partial_x \alpha^\pm_\pm(x, \omega) \alpha^\pm_\pm(x, \omega)^{-1},$$

so taking (vii) into account gives (viii).  

The fact that (8.24) is a Möbius transformation with a matrix out of $\text{Sp}(2L, \mathbb{C})$ has a number of consequences which we regroup for later use.
Corollary 8.2 Let $x \in \mathbb{S}$ and set $M^\pm_\pm = M^\pm_\pm(T_{-x}\omega)$, $M^\pm_\pm(+) = M^\pm_\pm(T_{-x+}\omega)$ and $\mathcal{V} = \mathcal{V}_x$.

Then

(i) $\pm M^\pm_\pm(+) = (-C^* \pm A^*M^\pm_\pm) (D^* \mp B^*M^\pm_\pm)^{-1}$

(ii) $\pm M^\pm_\pm(+) = (\pm M^\pm_\pm B - D)^{-1} (C \mp M^\pm_\pm A)$

(iii) $M^\pm_\pm(+) + M^\pm_\pm(-) = (D - M^\pm_\pm B)^{-1}(M^\pm_\pm + M^\pm_\pm) (D^* + B^*M^\pm_\pm)^{-1}$

\[ = (D + M^\pm_\pm B)^{-1}(M^\pm_\pm + M^\pm_\pm) (D^* - B^*M^\pm_\pm)^{-1} \]

(iv) $\pm M^\pm_\pm = (A \pm BM^\pm_\pm(+) (C \pm DM^\pm_\pm(+)^{-1})$

(v) $A \pm BM^\pm_\pm(+) = (D^* \mp B^*M^\pm_\pm)^{-1}$

(vi) $\Im(M^\pm_\pm(+) = (D \mp M^\pm_\pm B)^{-1} \Im(M^\pm_\pm)((D \mp M^\pm_\pm B)^{-1})$

(vii) $\Im(M^\pm_\pm(+) = ((D^* \mp B^*M^\pm_\pm)^{-1}) \Im(M^\pm_\pm) (D^* \mp B^*M^\pm_\pm)^{-1}$

(viii) $\tilde{M}^\pm_\pm(+) - \tilde{M}^\pm_\pm(-) = (M^\pm_\pm B - D)^{-1}\tilde{M}^\pm_\pm(B^*M^\pm_\pm - D^*)^{-1} - (M^\pm_\pm B + D)^{-1}\tilde{M}^\pm_\pm(D^* + B^*M^\pm_\pm)^{-1}$

Proof. This follows by short calculations using e.g. the Appendix of [74] and the identities $AB^* = BA^*$, $CD^* = DC^*$ and $AD^* - BC^* = 1$.

Now let us recall the definition of the Lyapunov exponents and state some of their properties. Because $T_\gamma(x, \omega)$ is a cocycle by Lemma 8.2, Osceledec’s theorem (see [52] for a concise statement) associates $2L$ Lyapunov exponents at $+\infty$ and $-\infty$ which will respectively be denoted by

\[ \gamma^\pm_1 \geq \ldots \geq \gamma^\pm_{2L} \] .

Similarly, $\alpha^\pm_\pm(x, \omega)$ are other cocycles of $L \times L$ matrices, so again each has $L$ Lyapunov exponents at $+\infty$ and $-\infty$ denoted by

\[ \gamma^{\pm}_1 \geq \ldots \geq \gamma^{\pm}_L \] ,

\[ \hat{\gamma}^{\pm}_1 \geq \ldots \geq \hat{\gamma}^{\pm}_L \] .

Part of the following proposition is copied from [52] (even though the definition of $\gamma^{\pm}_l$ differs by a sign).

Proposition 8.10 The various Lyapunov exponents satisfy:

(i) $\gamma_l^\pm = -\hat{\gamma}_{2L-l+1}^\pm$ for $l = 1, \ldots, 2L$

(ii) $\gamma_l^{\pm} = -\hat{\gamma}_{L-l+1}^{\pm}$ for $l = 1, \ldots, L$

(iii) $\gamma_l^\pm = \gamma_l^{\mp \pm}$ for $l = 1, \ldots, L$ and $z \in \mathbb{C}/\mathbb{R}$

(iv) $\gamma_l^\pm = \gamma_{L-l}^{\pm \pm}$ for $l = L+1, \ldots, 2L$ and $z \in \mathbb{C}/\mathbb{R}$

(v) $\hat{\gamma}_l\pm = -\hat{\gamma}_{2L-l+1}^\pm$ for $l = 1, \ldots, 2L$

(vi) $\hat{\gamma}_l^{\pm} = -\hat{\gamma}_{L-l+1}^{\pm \pm}$ for $l = 1, \ldots, L$ and $z \in \mathbb{C}/\mathbb{R}$

Proof. Items (i) and (ii) follow immediately from Lemma 5.2 of [52]. The other items can be proven as in Lemma 5.3 of [52] if one, moreover, uses the identity $\mathcal{T}_\gamma(x, \omega)^{-1} = \mathcal{J}^* \mathcal{T}_\gamma(x, \omega) \mathcal{J}$ following from Wronskian identity (8.7) and invokes Corollary 8.1 to show that $M^\pm_\pm(\omega)$ is uniformly bounded in $\omega$ for every fixed $x$. \qed
Next let us come to the construction of the stochastic Dirac operators of Theorem 8.1 and of the associated dynamical system. Let \( s \in [0, 1) = \mathbb{R}/\mathbb{Z} \). Each operator \( H(\omega) \) is of the form (8.1) with singular potentials at \( S = \mathbb{Z} + s \), hence \( x_j = j + s \). The \( V_j \) are drawn independently and identically out of \( \mathcal{J} so^*(2L) \) with some probability law \( p_V \) with compact support. Furthermore the potential \( \mathcal{W} \in L^1_{\text{loc}}(\mathbb{R}, \mathcal{J} so^*(2L)) \) is of the form

\[
\mathcal{W}(x) = \sum_{j \in \mathbb{Z}} \sum_{k=1}^{K} \lambda_{j,k} \mathcal{W}_k(x + s - j + 1),
\]

(8.25)

where \( K \in \mathbb{N} \), each \( \mathcal{W}_k \in L^1_{\text{loc}}(\mathbb{R}, \mathcal{J} so^*(2L)) \) has support \([0, 1] \) and the vectors \((\lambda_{j,k})_{k=1,...,K} \in \mathbb{R}^K \) are also drawn independently and identically according to a probability distribution \( p_W \) with compact support. Then \( \Omega \) is a compact subset of \((\mathcal{J} so^*(2L) \times \mathbb{R}^K) \times \mathbb{Z} \times \mathbb{R}/\mathbb{Z} \) and \( \mathcal{P} = (p_W \times p_V)^{\times \mathbb{Z}} \times ds \). The \( \mathbb{R} \)-action \( T \) is the natural right shift on \( \Omega \) and \( \mathcal{P} \) is indeed ergodic and even mixing w.r.t. \( T \).

In order to state the main hypothesis on the randomness, it is convenient to introduce the transfer matrix \( T^z(\mathcal{W}, V) \) as the solution \( T^z(1, 0) \) of (8.5) with potential \( \mathcal{W} \) and jump \( e^{\mathcal{J}V} \) at 1. Setting \( \lambda_j = (\lambda_{j,k})_{k=1,...,K} \) (which determines the potential bump \( \mathcal{W}_j = \sum_{k=1}^{K} \lambda_{j,k} \mathcal{W}_k \) between \( j - 1 \) and \( j \)), this notation implies

\[
T^z(\lambda_j, V_j) = T^z(j + s, j + s - 1, \omega).
\]

where the transfer matrix on the r.h.s. is defined by (8.5) with the Hamiltonian \( H(\omega) \).

**Coupling Hypothesis:** The semi-group generated by \( \{T^z(\lambda, V) \mid (\lambda, \mathcal{J}V) \in \text{supp}(p_W \times p_V) \} \) is Zariski dense in \( \text{SO}^*(2L) \) for all \( E \in \mathbb{R} \).

Let us stress that this hypothesis can be verified if \( p_W \times p_V \) is supported on a finite set of points, and also if either \( p_W \) or \( p_V \) is concentrated on a single point, notably the disorder is given only by a random potential \( \mathcal{W} \) or the random Dirac peaks \( V_j \delta_j \). Furthermore this hypothesis is satisfied whenever the set of \( T^E(\lambda, V) \) contains an open set (this property does not depend on \( E \)). This is e.g. the case if \( p_V \) contains an absolutely continuous part w.r.t. to the Haar measure.

### 8.5 Kotani theory

Kotani theory links the absolutely continuous spectrum of stochastic quasi-one-dimensional operators to the set of energies with vanishing Lyapunov exponents, by using analyticity arguments based on a few crucial identities. In all this section it is not needed that the stochastic Dirac operator has time-reversal symmetry or is of the particular random form given in (8.25). Kotani theory for stochastic Dirac operators with bounded potentials was developed in [86] by providing the relevant identities and then following closely the arguments of [52]. As already mentioned, the paper by Sun has some obvious errors which are corrected below. Moreover, we extend the theory in order to include singular potentials and potentials which may be complex-valued matrices. The singular potentials model a
Theorem 8.4 Let be given a stochastic family of Dirac operators with integrable and singular potentials. Then, for \( k = 1, \ldots, L \), the disjoint sets

\[ S_k = \{ E \in \mathbb{R} \mid \text{exactly } 2k \text{ Lyapunov exponents vanish at } E \} \]

are an essential support of the absolutely continuous spectrum of multiplicity \( 2k \).

Just as the crucial identities are different for discrete and continuous Schrödinger operators (compare [52]), there are some variations in the formulas in [86] for stochastic Dirac operators with singular potentials as well. We need to introduce further notations in order to state them. Averaging over \( \omega \) w.r.t. \( P \) is denoted by \( \mathbb{E} \). Another average along the orbit of singular points is

\[
\mathbb{E}_S(f) = \mathbb{E} \left( \lim_{x \to -\infty} \frac{1}{x} \sum_{y \in \mathbb{R} \cap [0, x]} f(T^{-y} \omega) \right) = \rho_S \mathbb{E} \left( \lim_{j \to -\infty} \frac{1}{j} \sum_{j=1}^{\infty} f(T^{-x_j} \omega) \right),
\]

namely one first averages over the random sites of the singular potential. Note that \( \mathbb{E}_S(1) = \rho_S \) and that the average \( \mathbb{E} \) can be dropped \( P \)-almost surely. Furthermore, if \( x_S \in S \) is the point closest to the origin, then \( T^{-x_S} \omega \) has a singular point at the origin and \( \mathbb{E}_S(f) = \int P(d\omega) f(T^{-x_S} \omega) \). Hence \( \mathbb{E}_S \) is closely linked to the Palm measure. Further the sum of the Lyapunov exponents is denoted by \( \gamma^\pm = \sum_{l=1}^{L} \gamma^+_l \) and we introduce two functions on \( \mathbb{C}/\mathbb{R} \) by

\[
w^\pm_S = -\mathbb{E}_S \ln \left( \det(D - M^\pm zB) \right) - \mathbb{E} \text{Tr} \left( R + M^\pm_0(Q - z) \right),
\]

\[
w^\pm_S = \mathbb{E}_S \ln \left( \det(D^* + B^* M^\pm_z) \right) - \mathbb{E} \text{Tr} \left( -R^* + M^\pm_z(Q - z) \right).
\]

By Corollary 8.1 the imaginary part of \( M^\pm_0 \) is uniformly bounded away from 0 so that \( w^\pm_S \) are well-defined. The branch of the logarithm is chosen in a continuous way in \( z \) (for each \( \omega \) separately) so that Theorem 8.2 then shows that \( w^\pm_S \) is analytic. The choice of the branch is of no importance below. Finally for any smooth function \( f \) on \( \Omega \) we define \( \partial f(\omega) = \partial_x f(T_x \omega) \) if \( 0 \notin S \).

Theorem 8.5 Let \( \Im m(z) \neq 0 \).

(i) There is a constant \( c \in \mathbb{R} \) such that \( w^+_S = w^-_S + i c \)

(ii) \( \gamma^\pm = -\Re e(w^\pm_S) \)

(iii) \( \partial_x w^\pm_S = \mathbb{E} \text{Tr}(\hat{G}^\pm) \)

(iv) \( 2 \gamma^\pm = \Im m(z) \mathbb{E} \text{Tr} \left( (1 + |M^\pm|^2) (\Im m(M^\pm_0))^{-1} \right) \)

Items (ii) and (iii) combined provide a Thouless formula for stochastic Dirac operators. The proof is based on a series of algebraic identities which we check first.
Lemma 8.3 Let $\Im m(z) \neq 0$. Away from singular points, the following identities hold.

(i) $\partial \text{Tr}(\ln(M_+^x + M_-^x)) = \text{Tr}(R^* + R + (Q - z)(M_+^x - M_-^x))$

(ii) $\partial \text{Tr}((M_+^x + M_-^x)^{-1}(\partial_x M_+^x - \partial_x M_-^x)) = 2 \text{Tr}(\hat{G}_0^x) + \partial_x \text{Tr}((Q - z)(M_+^x + M_-^x))$

(iii) $\pm \partial \text{Tr}(\ln(\Im m(M_+^x))) = 2 \Re \text{Tr}(\text{Tr}(W^x_+)) - \Im m(z) \text{Tr}((1 + |M_+^x|^2)\Im m(M_+^x))^{-1}$

where $W^x_+ = R + (Q - z)M_+^x$ and $W^x_- = -R^* + (Q - z)M_-^x$

(iv) $\partial_x [\alpha^x_+(x,\omega)^*\Im m(M_+^x(T_{-x}\omega))\alpha^x_+(x,\omega)] = \mp \Im m(z) \alpha^x_+(x,\omega)^* (1 + |M_+^x(T_{-x}\omega)|^2) \alpha^x_+(x,\omega)$

**Proof.** In the formulas below all functions have the argument $T_{-x}\omega$, and one may then set $x = 0$. Using Lemma 8.2(viii), a short calculation shows

$$\partial (M_+^x + M_-^x) = (M_+^x + M_-^x)(R^* - (Q - z)M_-^x) + (R + M_+^x(Q - z))(M_+^x + M_-^x). \quad (8.26)$$

Multiplying this by $(M_+^x + M_-^x)^{-1}$ and then using the cyclicity of the trace shows the formula of (i). For (ii), let us take the derivative $\partial_x$ of the Riccati equation of Lemma 8.2(viii):

$$\partial (M_+^x - \hat{M}_+^x) = -(2 + (M_+^x)^2 + (M_-^x)^2) + (\hat{M}_+^x - \hat{M}_-^x)R^* + R(\hat{M}_+^x - \hat{M}_-^x)$$

$$+ \hat{M}_+^x(Q - z)M_+^x + M_+^x(Q - z)M_+^x + \hat{M}_-^x(Q - z)M_-^x + M_-^x(Q - z)\hat{M}_-^x.$$

Using this and (8.26), some algebra directly leads to (ii) if one also uses the identity

$$\text{Tr}(\hat{G}_0^x) = \text{Tr}(((M_+^x)^{-1} + (M_-^x)^{-1})^{-1} - (M_+^x + M_-^x)^{-1}),$$

following from Propositions 8.4 and 8.6.

Next we turn to the proof of (iii). Let us set $M_+^x = X_+^x + iY_+^x$ with $Y_+^x = \Im m(M_+^x)$.

From $M_+^x = (M_+^x)^*$ follows $X_+^x = X_+^x$ and $Y_+^x = -Y_+^x = (Y_+^x)^*$. Straightforward calculation then shows

$$\partial Y_+^x = RY_+^x + Y_+^x R^* \pm X_+^x(Q - \Re(z))Y_+^x \pm Y_+^x(Q - \Re(z))X_+^x \mp \Im m(z)(1 + (X_+^x)^2 - (Y_+^x)^2).$$

Thus

$$\partial \text{Tr}(\ln(Y_+^x)) = \text{Tr}(R + R^* \pm 2X_+^x(Q - \Re(z))) \mp \Im m(z)\text{Tr}((Y_+^x)^{-1}(1 + (X_+^x)^2 - (Y_+^x)^2))$$

$$= -2 \Re \text{Tr}(\text{Tr}(W^x_+)) \mp \Im m(z)\text{Tr}((Y_+^x)^{-1}(1 + (X_+^x)^2 - (Y_+^x)^2)),$$

where in the last step we used $\text{Tr}(Y^{-1}[X,Y]) = 0$. Finally let us consider (iv). When calculating the derivative on the l.h.s. the product rule leads to three terms. The term containing $\partial Y_+^x$ is given by the above formula, those involving derivatives of $\alpha^x_+(x,\omega)$ by Lemma 8.2(vii). Hence it is sufficient to check

$$(-R^* \mp (Q - z)M_+^x)^*Y_+^x + \partial Y_+^x + Y_+^x(-R^* \mp (Q - z)M_+^x) = \mp \Im m(z)\left(1 + |M_+^x|^2\right).$$

Again this follows from some algebra. □
Proof of Theorem 8.5. (i) Set $I^z = \mathbf{E} \text{Tr}(R^* + R + (Q - z)(M_+^z - M_-^z))$. By the ergodic theorem and Lemma 8.3(i), $\mathbf{P}$-almost surely

$$I^z = \lim_{y \to \infty} \frac{1}{y} \int_0^y dx \text{Tr}\left( R^*(T_x \omega) + R(T_x \omega) + (Q(T_x \omega) - z)(M_+^z(T_x \omega) - M_-^z(T_x \omega)) \right)$$

$$= \lim_{y \to \infty} \frac{1}{y} \int_{-y}^y dx \partial_x \text{Tr}(M_+^z(T_{-x} \omega) + M_-^z(T_{-x} \omega)))$$

$$= \lim_{y \to \infty} \frac{1}{y} \sum_{-y \leq x_j \leq 0} \left[ \ln\left( \det(M_+^z(T_{-x_j} \omega) + M_-^z(T_{-x_j} \omega)) \right) \right]_{x_{j-1}}^{x_j} + 2\pi i n_j,$$

where $S = (x_j)_{j \in \mathbb{Z}}$ with $x_{j-1} \leq x_j$ and $n_j \in \mathbb{Z}$ denotes the number of branches of the logarithm needed in the integral from $x_{j-1}$ to $x_j$ minus 1. Now by Lemma 8.2(iv), $M_+^z(T_{-x_{j+1}} \omega) = M_+^z(T_{-x_j} \omega)$. On the other hand $M_+^z(T_{-x_j} \omega) + M_-^z(T_{-x_j} \omega)$ can be calculated by Corollary 8.2(iii). Thus regrouping the terms shows that

$$I^z = \lim_{y \to \infty} \frac{1}{y} \sum_{-y \leq x_j \leq 0} \left[ \ln\left( \det(D_j - M_+^z(T_{-x_j} \omega)B_j) + \ln\left( \det(D_j^* + B_j^*M_-^z(T_{-x_j} \omega)) \right) - 2\pi i n_j \right].$$

Hence if $c$ is the average of $2\pi i n_j$ over $S$, we have shown

$$I^z = - \mathbf{E}_S \ln(\det(D - M_+^z B)) - \mathbf{E}_S \ln(\det(D^* + B^*M_-^z)) + i c,$$

and thus (i). For (ii) let us start from a formula for $\gamma^z$ which follows from the identities stated in Proposition 8.10:

$$\gamma^z = \lim_{y \to \infty} \frac{1}{y} \ln\left( |\det(\alpha_+^z(y, \omega))| \right),$$

where the convergence holds $\mathbf{P}$-almost surely. Telescoping and regrouping gives

$$\gamma^z = \lim_{y \to \infty} \frac{1}{y} \sum_{0 < x_j < y} \left[ \ln\left( |\det(\alpha_+^z(x_j +, \omega))| \right) - \ln\left( |\det(\alpha_+^z(x_{j-1} +, \omega))| \right) \right]$$

$$= \lim_{y \to \infty} \frac{1}{y} \sum_{0 < x_j < y} \left[ \ln\left( |\det(\alpha_+^z(x_j +, \omega)\alpha_+^z(x_j -, \omega)^{-1})| \right) + \int_{x_{j-1}}^{x_j} dx \partial_x \ln\left( |\det(\alpha_+^z(x, \omega))| \right) \right].$$

The first contribution can be evaluated with Lemma 8.2(vi) and the definition of $\mathbf{E}_S$, the second contribution be summed up and the integrand evaluated:

$$\gamma^z = - \mathbf{E}_S \ln(|\det(D^* + B^*M_-^z)|) + \Re \lim_{y \to \infty} \frac{1}{y} \int_0^y dx \text{Tr}(\alpha_+^z(x, \omega)^{-1}\partial_x \alpha_+^z(x, \omega)).$$

Finally the last expression can be calculated using Lemma 8.2(vii) and then the ergodic theorem completes the proof of (ii). Let us point out that one could have started from

$$\gamma^z = - \lim_{y \to \infty} \frac{1}{y} \ln\left( |\det(\alpha_+^z(y, \omega))| \right).$$
by

Then a similar calculation leads to \( \gamma^z = -\Re(w_+^z) \). Because \( w_\pm^z \) are analytic, this also provides an alternative proof of (i).

(iii) Let us set \( J^z = 2\mathbb{E}\text{Tr}(G^z) + \partial_z\mathbb{E}\text{Tr}(R - R^* + (Q - z)(M_+^z + M_-^z)) \). Because the probability of having a singular potential at 0 vanishes, \( \mathbb{E}\text{Tr}(G^z) \) can be replaced by \( \mathbb{E}\text{Tr}(G_0^z) \). Furthermore the term \( R - R^* \) drops out due to the derivative \( \partial_z \). Hence Lemma 8.3(ii), the ergodic theorem and reordering of the terms imply as above that \( \mathbb{P} \)-almost surely

\[
J^z = \lim_{y \to \infty} \frac{1}{y} \sum_{-y \leq x_j \leq 0} \text{Tr} \left( (M_+^z(T_x \omega) + M_-^z(T_x \omega))^{-1}(\hat{M}_+^z(T_x \omega) - \hat{M}_-^z(T_x \omega)) \right)^{x_j},
\]

where we also used the left-continuity of \( x \in \mathbb{R} \mapsto M_\pm^z(T_x \omega) \). The terms with \( x_j \) now have to be evaluated using Lemma 8.2(iv) or its equivalent formulations. The factor \( (M_+^z(T_{x_j+\omega}) + M_-^z(T_{x_j+\omega}))^{-1} \) is given by the inverse of Corollary 8.2(iii). Corollary 8.2(viii) moreover allows to calculate \( M_+^z(T_{x_j+\omega}) - M_-^z(T_{x_j+\omega}) \). Replacing both identities then shows

\[
J^z = \mathbb{E}_\theta \text{Tr} \left( (M_+^z + M_-^z)^{-1} \left[ (D + M_-^z B)(D - M_+^z B)^{-1} \hat{M}_+^z \\
+ \hat{M}_-^z (D^* + B^* M_-^z)^{-1} (B^* M_+^z - D^*) - \hat{M}_+^z + \hat{M}_-^z \right] \right).
\]

Due to the definitions of \( J^z \) and \( w_\pm^z \) this concludes the proof of (iii).

(iv) We set \( K^z_\pm = \mathbb{E}(2\Re(\text{Tr}(W^z_\pm)) - \Im m(z) \text{Tr}((1 + |M^z_\pm|^2)\Im m(M^z_\pm)^{-1})) \). Using the ergodic theorem and Lemma 8.3(iii) one has \( \mathbb{P} \)-almost surely

\[
\pm K^z_\pm = \lim_{y \to \infty} \frac{1}{y} \sum_{-y \leq x_j \leq 0} \ln \left( \det(\Im m(M^z_\pm(T_x \omega))) \right)^{x_j}.
\]

Now evaluate \( \Im m(M^z_\pm(T_{x_j+\omega})) \) by Corollary 8.2(vi). This implies

\[
K^z_\pm = \pm \mathbb{E}_\theta \ln(\det(|D \mp M^z_\pm B|^2)) = \pm 2\Re \mathbb{E}_\theta \ln(\det(D \mp M^z_\pm B)).
\]

Similarly, using Corollary 8.2(vii), \( K^z_\pm = \mp 2\Re \mathbb{E}_\theta \ln(\det(D^* \mp B^* M^z_\pm)) \). From these identities one readily completes the proof.

The second part of the following theorem establishes Theorem 6.6 of [52] also for complex valued potentials.

**Theorem 8.6** Consider the positive operator \( U^z_\pm = (\Im m(M^z_\pm))^{\frac{1}{2}}(1 + |M^z_\pm|^2)^{-1}(\Im m(M^z_\pm)^2)^{\frac{1}{2}} \) and denote its eigenvalues by \( u^{z}_{1,\pm} \geq \ldots \geq u^{z}_{L,\pm} \geq 0 \). Further let \( E \in \mathbb{R} \), \( \epsilon > 0 \) and \( k = 1, \ldots, L \). Then

\[
\mathbb{E} \sum_{l=1}^{k} \frac{1}{u^{z}_{L+1,l}} \leq \frac{2}{\epsilon} \sum_{l=1}^{k} \gamma^{E\pm\epsilon}_{L+1-l}.
\]
If furthermore $E$ is such that $\gamma^E_l = \lim_{\epsilon \to 0} \gamma^E_{l+\epsilon}$ exists for $l = 1, \ldots, L$, then

$$ E \sum_{l=1}^k \frac{1}{u_{l+\epsilon}^E} \leq \frac{2}{\epsilon} \left[ \sum_{l=1}^k \gamma^E_{L+1-l} + \sum_{l=k+1}^L (\gamma^E_{L+1-l} - \gamma^E_{L+1-l}) \right]. \tag{8.28} $$

**Proof.** This is an adaption and slight generalization of the proof of Theorems 6.5 and 6.6 of [52] (the reasoning in [86] is erroneous at several points). For any $L \times L$ matrix $F$ let $\Lambda^k F$ and $d\Lambda^k F$ the second quantizations on the fermionic tensor product $\Lambda^k \mathbb{C}^L$, such that $e^{d\Lambda^k F} = \Lambda^k e^F$. Let $z = E + i\epsilon$ and $Y^z_\pm = \Im m(M^z_{\pm})$. Define $F^z_\pm(x, \omega) = Y^z_\pm(T_{-x}\omega)^\frac{1}{2} \alpha^z_\pm(x, \omega)$. Then

$$ \partial_x \Lambda^k |F^z_\pm(x)|^2 = \Lambda^k F^z_\pm(x)^* (d\Lambda^k (F^z_\pm(x)^{-1})^* \partial_x |F^z_\pm(x)|^2 F^z_\pm(x)^{-1}) \Lambda^k F^z_\pm(x). $$

Thus by Lemma 8.3(iv)

$$ \partial_x \Lambda^k |F^z_\pm(x, \omega)|^2 = \mp \Im m(z) \Lambda^k F^z_\pm(x, \omega)^* (d\Lambda^k U^z_\pm(T_{-x}\omega)^{-1}) \Lambda^k F^z_\pm(x, \omega), $$

so that for $\Im m(z) > 0$

$$ \partial_x \Lambda^k |F^z_\pm(x, \omega)|^2 \geq - \Im m(z) \|d\Lambda^k U^z_\pm(T_{-x}\omega)^{-1}\| \Lambda^k |F^z_\pm(x, \omega)|^2, $$

$$ \partial_x \Lambda^k |F^z_\pm(x, \omega)|^2 \leq \Im m(z) \|d\Lambda^k U^z_\pm(T_{-x}\omega)^{-1}\| \Lambda^k |F^z_\pm(x, \omega)|^2. $$

Integrating hence gives

$$ \Lambda^k |F^z_\pm(x, \omega)|^2 \geq \exp \left( - \Im m(z) \int_0^x dy \|d\Lambda^k U^z_\pm(T_y\omega)^{-1}\| \right) \Lambda^k |F^z_\pm(0, \omega)|^2, $$

$$ \Lambda^k |F^z_\pm(x, \omega)|^2 \leq \exp \left( \Im m(z) \int_0^x dy \|d\Lambda^k U^z_\pm(T_y\omega)^{-1}\| \right) \Lambda^k |F^z_\pm(0, \omega)|^2. $$

Note that by Lemma 8.2(vi) and Corollary 8.2(vii) the functions $|F^z_\pm(x, \omega)|$ are actually smooth also for $x \in \mathbb{S}$. We combine this with the inequalities

$$ \|\Lambda^k \alpha^z_\pm(x, \omega)\|^2 \|\Lambda^k Y^z_\pm(T_{-x}\omega)^{-1}\|^{-1} \leq \|\Lambda^k |F^z_\pm(x, \omega)|^2\| \leq \|\Lambda^k \alpha^z_\pm(x, \omega)\|^2 \|\Lambda^k Y^z_\pm(T_{-x}\omega)\|. $$

Taking logarithms thus shows

$$ \ln \left( \frac{\|\Lambda^k \alpha^z_\pm(x, \omega)\|^2}{\|\Lambda^k Y^z_\pm(T_{-x}\omega)^{-1}\|} \right) \leq \Im m(z) \int_0^x dy \|d\Lambda^k U^z_\pm(T_y\omega)^{-1}\| + \ln \left( \|\Lambda^k |F^z_\pm(0, \omega)|^2\| \right). $$

Now by Corollary 8.1, $Y^z_\pm(\omega)^{-1}$ is uniformly bounded in $\omega$. Thus dividing by $x$ and then taking the limit $x \to \infty$ shows by Proposition 8.10(iii) and the ergodic theorem

$$ 2 \sum_{l=1}^k \gamma_l^z \leq \Im m(z) E \|d\Lambda^k (U^z_\pm)^{-1}\| = \Im m(z) E \sum_{l=L-k+1}^L \frac{1}{u^z_{l-l}}. \tag{8.29} $$
8.6. THE LYAPUNOV SPECTRUM

Combining this with \( k \) replaced by \( L-k \) together with Theorem 8.5(iv) stating

\[
2 \sum_{l=1}^{L} \gamma_l^z = \Im m(z) \sum_{l=1}^{L} \frac{1}{u_{z,l}^+},
\]

proves inequality (8.27) for the sign \(-\). Similarly one has

\[
\ln \left( \| \Lambda^k \alpha_+^z(x,\omega) \| ^2 \right) \geq - \Im m(z) \int_{0}^{\pi} dy \| d\Lambda^k U_+^z(T_{-y} \omega)^{-1} \| + \ln \left( \frac{\| \Lambda^k F_+^z(0,\omega) \| ^2}{\| \Lambda^k Y_+^z(T_{-\omega}) \| } \right).
\]

As the last term is bounded along the orbit, Proposition 8.10(iii) now implies

\[
2 \sum_{l=1}^{k} \gamma_l^z \leq \Im m(z) \sum_{l=L-k+1}^{L} \frac{1}{u_{z,l}^+},
\]

which again combined with Theorem 8.5(iv) proves (8.27) for the sign \(+\).

For the proof of (8.28) we need the following general fact. If \( T, S > 0 \) are two positive matrices, then the positive operators \( T^+ S T^+ \) and \( S^+ T S^+ \) have the same spectrum (this follows from \( \text{Tr}((T^+ S T^+)^n) = \text{Tr}((S^+ T S^+)^n) \) for all \( n \in \mathbb{N} \)). Hence \( u^z_{\pm,k} \) are also the eigenvalues of the imaginary part of the Herglotz function \((1 + |M_{\pm}^z|^2)^{-\frac{1}{2}} M_{\pm}^z(1 + |M_{\pm}^z|^2)^{-\frac{1}{2}}\) and by the Herglotz representation theorem it follows as in [52] that

\[
\frac{\epsilon}{u^z_{\pm,k}} \geq \frac{\delta}{u^z_{\pm,k}} \quad \text{for } \epsilon \geq \delta > 0.
\]

Combining this fact with Theorem 8.5(iv) and the bounds (8.29) and (8.30) gives

\[
E \sum_{l=1}^{k} \frac{\epsilon}{u^z_{\pm,k}} \leq E \sum_{l=1}^{L} \frac{\epsilon}{u^z_{\pm,k}} - E \sum_{l=k+1}^{L} \frac{\delta}{u^z_{\pm,k}} \leq 2 \sum_{l=1}^{L} \gamma_{E+\epsilon}^z - 2 \sum_{l=1}^{L-k} \gamma_{E+\delta}^z.
\]

Now taking the limit \( \delta \to 0 \) leads to (8.28).

From this point on the proof of Theorem 8.4 is line by line the same as in [52].

8.6 The Lyapunov spectrum

This section proves a criterion for the distinctness (apart from Kramers degeneracy) of the Lyapunov exponents for random products of matrices in \( \text{SO}^*(2L) \). It can be immediately applied to the transfer matrices if the Coupling Hypothesis holds. On the other hand, we believe it to be of somewhat independent interest and thus took care to make it readable without reference to the rest of the paper.
Instead of the group $SO^*(2L)$ as defined in the introduction it will be more convenient to work with an isomorphic group $\mathbb{G}$ for which the polar decomposition takes a more simple form. Thus we define in case of even $L = 2d$ and odd $L = 2d + 1$ respectively

$$A = \frac{1}{\sqrt{2}} \begin{pmatrix} 1_d & 1_d \\ i1_d & -i1_d \end{pmatrix}, \quad A = \frac{1}{\sqrt{2}} \begin{pmatrix} 1_d & 0 & 1_d \\ 0 & \sqrt{2} & 0 \\ i1_d & 0 & -i1_d \end{pmatrix},$$

where $d \times d$ square matrices carry the index $d$. Then set

$$A = \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix},$$

which satisfies $A^* = A^{-1}$. Then we set $\mathbb{G} = A^*SO^*(2L)A$. This group consists of all $2L \times 2L$ matrices $M$ satisfying

$$M^*JM = J, \quad M^*SM = S,$$

where

$$S = \begin{pmatrix} A^tA & 0 \\ 0 & A^tA \end{pmatrix}, \quad A^tA = \begin{pmatrix} 0_d & 0 & 1_d \\ 0 & 1 & 0 \\ 1_d & 0 & 0_d \end{pmatrix} \text{ if } L = 2d + 1,$$

while in the case $L = 2d$ the middle line and column in the formula for $A^tA$ are absent. Note that the matrices $J$ and $S$ commute, $J^* = -J = J^{-1}$ and $S^* = S = S^{-1}$. Further we remark that the group $Sp(2L,\mathbb{H})$ defined in [78] for even $L$ is also isomorphic to $SO^*(2L)$.

**Lemma 8.4** Let $M \in \mathbb{G}$ and $v \in \mathbb{C}^{2L}$.

(i) $M^* \in \mathbb{G}$

(ii) If $Mv = \lambda v$, then $M^*Jv = \lambda^{-1}Jv$, $MJ\bar{S}v = \bar{\lambda}J\bar{S}v$ and $M^*S\bar{v} = \bar{\lambda}^{-1}S\bar{v}$.

(iii) The vectors $v$ and $J\bar{S}v$ are linearly independent for $v \neq 0$.

(iv) For $M > 0, M \in \mathbb{G}$, there exists $U \in \mathbb{G} \cap SU(2L)$ such that $UMU^* = D$, where

$$D = \text{diag}(a_1, \ldots, a_d, 1, a_1^{-1}, \ldots, a_d^{-1}, a_1^{-1}, \ldots, a_d^{-1}, 1, a_1, \ldots, a_d) \text{ if } L = 2d + 1 \text{ and}$$

$$D = \text{diag}(a_1, \ldots, a_d, a_1^{-1}, \ldots, a_d^{-1}, a_1^{-1}, \ldots, a_d^{-1}, a_1, \ldots, a_d) \text{ in case } L = 2d, \text{ with real constants}$$

$$a_1 \geq a_2 \geq \ldots a_d \geq 1. \text{ Note that } D \in \mathbb{G}.$$  

(v) There are unitary matrices $K, U \in \mathbb{G} \cap SU(2L)$ and a diagonal matrix $D$ as in (iv) such that

$$M = KDU.$$  

(vi) One has $\det(M) = 1$ and the group $\mathbb{G}$ is connected.

**Proof.** (i) follows by inverting the relations in (8.31). For (ii) note that $M^*JM = J$ implies $J^*M^*J = M^{-1}$. Hence $J^*M^*Jv = \lambda^{-1}v$ implies $M^*Jv = \lambda^{-1}v$. From $M^*SM = S$ it follows that $S^*M^*S = M^{-1} = J^*M^*J$. Taking the transpose one obtains
\( S M S = J^* \mathcal{M} J \) and hence \( S M S J \pi = -\lambda J^* \pi \) and therefore \( M S J \pi = \lambda S J \pi \). Now using the same calculation as above yields the last equation.

(iii) Writing \( v = (a, b) \) and \( J S v = \lambda v \) gives \( \lambda a = A^t A b \) and \( \lambda b = -A^t A \pi \). As \( A^t A \) is real and \( (A^t A)^2 = 1 \), this implies \( |\lambda|^2 a = A^t A \pi = -a \) and therefore \( (1 + |\lambda|^2) a = 0 \) implying \( a = 0 \) and \( b = 0 \) and hence \( v = 0 \). Therefore these vectors are linearly dependent if and only if \( v = 0 \).

(iv) First we need some basic facts. We say that a subspace \( V \) of the form \( \mathcal{M} > 0 \). One proceeds by induction to complete the proof of the claim.

8.6. THE LYAPUNOV SPECTRUM

For \( M > 0 \) the eigenspaces are orthogonal. Let \( V_1 \) be the eigenspace for the value \( 1 \) (possibly only the zero vector) and \( V_0 \) be the orthogonal complement. By (ii) and the consideration above, these spaces are \( G \)-like and they are invariant under \( M^* M \). By (ii) and (iii) the dimension of \( V_0 \) is divisible by 4, say \( \dim V_0 = 4 \).

**First claim:** \( V_0 \) has an orthonormal basis of eigenvectors of \( M^* M \) of the form \( v_1, v_2, \ldots, v_r, S v_1, \ldots, S v_r, J v_1, \ldots, J v_r, J S v_1, \ldots, J S v_r \).

Indeed, if \( \dim(V_0) = 0 \), there is nothing to prove. Otherwise let \( a_1^2 > 1 \) be the biggest eigenvalue of \( M^* M \) which is also the biggest eigenvalue of \( M^* M \) restricted to \( V_0 \) and let \( v \in V_0 \) be some corresponding eigenvector. Then \( J S \pi \) is another eigenvector for the same eigenvalue. Take \( w = v + \mu J S \pi \), where \( \mu \in \mathbb{C} \) can be chosen in such a way that \( w \) and \( J S \pi \) are orthogonal. Then also \( J w \) and \( S w \) which are eigenvectors to the eigenvalue \( a_1^{-1} \) are orthogonal. As \( a_1 > a_1^{-1} \), the space spanned by \( w \) and \( J S \pi \) is orthogonal to the space spanned by \( J W \) and \( S W \). Therefore normalizing \( w \) to \( v_1 = w / ||w|| \) the vectors \( v_1, S v_1, J v_1, J S v_1 \) are orthonormal. Denote the space spanned by these vectors by \( V_{0,1} \subset V_0 \) and its orthogonal complement in \( V_0 \) by \( V_{0,2} \) which is again a \( G \)-like, \( M \)-invariant subspace. One proceeds by induction to complete the proof of the claim.

**Second claim:** If \( L = 2d \), then \( \dim(V_1) \) is divisible by 4 and there is an orthonormal basis of the form \( v_{d+1}, \ldots, v_d, S v_{d+1}, \ldots, S v_d, J v_{d+1}, \ldots, J v_d, J S v_{d+1}, \ldots, J S v_d \). If \( L = 2d + 1 \), then \( \dim(V_1) \) is congruent to 2 mod 4 and one has an orthonormal basis which is of the form \( v_{d+1}, \ldots, v_d, v_{d+1}, \ldots, v_d, S v_{d+1}, \ldots, S v_d, J v_{d+1}, \ldots, J v_d, J S v_{d+1}, \ldots, J S v_d \) with \( S v_{d+1} = v_{d+1} \).

Indeed, as \( J \) is unitary and operates on \( V_1 \), there is an orthonormal basis of \( V_1 \) of eigenvectors of \( J \). The eigenvalues of \( J \) are \( \pm 1 \). If \( J v = \pm w \), then \( J S \pi = S J \pi = \mp i S \pi \). Hence the dimensions of the eigenspaces of \( J \) in \( V_1 \) are equal. If \( \dim(V_1) \geq 4 \), there are two orthonormal vectors \( w_1, w_2 \) satisfying \( J w_j = i w_j \). As \( J S w_j = -i S w_j \) the vectors \( w_1, w_2, S w_1, S w_2 \) are orthonormal. Set \( v_{d+1} = \frac{1}{\sqrt{2}} (w_1 + S w_2) \). Then the vectors \( v_{d+1}, J v_{d+1}, S v_{d+1}, J S v_{d+1} \) span a 4-dimensional \( G \)-like subspace of \( V_1 \). Denote its orthonormal complement in \( V_1 \) by \( V_2 \) and proceed by induction to obtain the vectors \( v_{d+2}, \ldots, v_d \). In
case $L = 2d$ this shows the above claim; if $L = 2d + 1$, one is left with some $2$-dimensional, $G$-like subspace $V_{d-r+1}$. This space is spanned by the orthonormal vectors $w$ and $\overline{Sv}$ where $Jw = \overline{w}$. Set $v_{d+1} = \sqrt{2}(w + \overline{Sv})$, then $v_{d+1}$ and $Jv_{d+1}$ form an orthonormal basis of $V_{d-r+1}$ and $Sv_{d+1} = \overline{v_{d+1}}$.

**Construction of $U$:** From the first two steps we obtain an orthonormal basis of eigenvectors of $M^*M$ of the form $(v_1, \ldots, v_d, v_{d+1}, S\overline{v_1}, \ldots, S\overline{v_d}, Jv_1, \ldots, Jv_{d+1}, JS\overline{v_1}, \ldots, JS\overline{v_d})$ in case $L = 2d + 1$ and the same without the entries containing $v_{d+1}$ if $L = 2d$. The corresponding eigenvalues of $v_1, \ldots, v_d$ shall be denoted by $a_1^2 \geq a_2^2 \geq \ldots a_d^2 \geq 1$. The eigenvalue corresponding to $v_{d+1}$ if $L = 2d + 1$ is 1. Denote the canonical basis of $\mathbb{C}^{2L}$ by $e_i$, $i = 1, \ldots, 2L$. Let us define the unitary matrix $U$ by

$$
L = 2d \\
Uv_i = e_i \quad i = 1, \ldots, d \\
USv_i = e_{i+d} \quad i = 1, \ldots, d \\
UFv_i = -e_{i+2d} \quad i = 1, \ldots, d \\
USFv_i = -e_{i+3d} \quad i = 1, \ldots, d \\

L = 2d + 1 \\
Uv_i = e_i \quad i = 1, \ldots, d + 1 \\
USv_i = e_{i+d+1} \quad i = 1, \ldots, d \\
UFv_i = -e_{i+2d+1} \quad i = 1, \ldots, d + 1 \\
USFv_i = -e_{i+3d+2} \quad i = 1, \ldots, d.
$$

Then defining the diagonal matrix $D$ as in the statement of the proposition, one has $U^*MU^*$. For $i = 1, \ldots, d$, one has

$$
(U^*JU)v_i = U^*Jv_i = -U^*e_{i+L} = Jv_i, \\
(U^*JU)Fv_i = -U^*Fv_i = -v_i = J(Jv_i),
$$

similar calculations hold for $S\overline{v_i}$, $J\overline{Sv_i}$ and also $v_{d+1}, Jv_{d+1}$ in the case $L = 2d + 1$. Thus one obtains $U^*JU = J$. In the following calculations, we have again $1 \leq i \leq d$ and, whenever the equation is split up into two, then the upper one holds for $L = 2d$ and the lower one for $L = 2d + 1$.

$$
(U^*SU)v_i = U^*Sv_i = \begin{cases} 
U^*e_{i+d} & (L = 2d) \\
U^*e_{i+d+1} & (L = 2d + 1)
\end{cases} = U^*USv_i = Sv_i, \\
(U^*SU)\overline{v_i} = \begin{cases} 
U^*Sv_i & (L = 2d) \\
U^*Sv_i & (L = 2d + 1)
\end{cases} = U^*\overline{USv_i} = \overline{Sv_i},
$$

$$
(U^*SU)e_{i+L} = \begin{cases} 
-U^*e_{i+3d} & (L \neq 2d + 1) \\
-U^*e_{i+3d+2} & (L = 2d + 1)
\end{cases} = JsSv_i = JSv_i, \\
(U^*SU)\overline{v_i} = \begin{cases} 
-U^*\overline{v_i} & (L \neq 2d + 1) \\
-U^*\overline{v_i} & (L = 2d + 1)
\end{cases} = JS\overline{v_i} = JS^2\overline{v_i} = S(JSv_i).
$$

If $L = 2d + 1$, one observes additionally

$$
(U^*SU)v_{d+1} = U^*Sv_{d+1} = \overline{Jv_{d+1}} = S\overline{v_{d+1}} = Sv_{d+1}, \\
(U^*SU)Jv_{d+1} = -U^*Sv_{d+1} = -JuSv_{d+1} = JSv_{d+1} = S\overline{v_{d+1}}.
$$

Therefore one also obtains $U^*SU = S$ and hence $U \in G \cap U(2L)$. Finally, as $U \in G$ we have $AA^* \in SO^*(2L) \cap U(2L) = Sp(2L, \mathbb{R}) \cap O(2L)$ and hence $\det(U) = \det(M^*M) = 1$ and therefore $U \in SU(2L)$.

(v) As $M^*M \in G$ and $M^*M > 0$, by (iv) we find $U \in G \cap SU(2L)$ and a diagonal matrix
8.6. THE LYAPUNOV SPECTRUM

\( \mathcal{D} \) as above, such that \( \mathcal{U} \mathcal{M}^* \mathcal{M}^* = \mathcal{D}^2 \). Set \( \mathcal{K} = \mathcal{M} \mathcal{U}^* \mathcal{D}^{-1} \subset \mathbb{G} \), then \( \mathcal{M} = \mathcal{K} \mathcal{D} \mathcal{U} \) and \( \mathcal{K}^* \mathcal{K} = \mathcal{D}^{-1} \mathcal{U} \mathcal{M}^* \mathcal{M}^* \mathcal{D}^{-1} = 1 \). Hence \( \mathcal{K} \in \mathbb{G} \cap \mathbb{U}(2L) = \mathbb{G} \cap \mathbb{SU}(2L) \).

(vi) By (v), \( \det(\mathcal{M}) = \det(\mathcal{K}) \det(\mathcal{D}) \det(\mathcal{U}) = 1 \). Furthermore as the group \( \text{Sp}(2L, \mathbb{R}) \cap \mathcal{O}(2L) \) is connected, also \( \text{SO}^*(2L) \cap \mathbb{U}(2L) \) is. Using the decomposition in (iv) one easily obtains that \( \mathcal{G} \) is connected.

Now let \( (\mathcal{Y}_n)_{n \geq 1} \) be an i.i.d. sequence in \( \mathcal{G} \). Then by Lemma 8.4 the whole associated Lyapunov spectrum has at least multiplicity two. So let \( \gamma_1, \gamma_1, \gamma_2, \gamma_2, \ldots, \gamma_L, \gamma_L \) be the 2L Lyapunov exponents with \( \gamma_1 \geq \gamma_2 \geq \ldots \gamma_L \). Lemma 8.4 also shows \( \gamma_p = -\gamma_{L+1-p} \) and in the case \( L = 2d+1 \), one has \( \gamma_{d+1} = 0 \). Therefore it is always enough to consider \( \gamma_1, \ldots, \gamma_d \).

Set \( v^{(p)} = e_1 \wedge \ldots \wedge e_p \wedge e_{2L-d+1} \wedge \ldots e_{2L-d+p} \) and define \( \mathbb{L}_p := \text{span}_\mathbb{R}(\{ \Lambda^{2p} \mathcal{M} v^{(p)} \mid \mathcal{M} \in \mathcal{G} \}) \) which is a real linear subspace of \( \Lambda^{2p} \mathcal{C}^{2L} \). Note that \( \mathbb{L}_p \) does not have to be a complex vector space. Taking the real part of the scalar product on \( \Lambda^{2p} \mathbb{C}^{2L} \) induces a scalar product on \( \mathbb{L}_p \) but actually one does not need to take the real part as the following lemma shows.

**Lemma 8.5** The scalar product in \( \Lambda^{2p} \mathbb{C}^{2L} \) two vectors in \( \mathbb{L}_p \) is real. Let \( f_1, f_2, f_3, f_4 \in \mathbb{L}_p \) and consider \( f_1 \wedge f_2, f_3 \wedge f_4 \) on one hand as elements in \( \Lambda^2(\Lambda^{2p} \mathbb{C}^{2L}) \) and on the other hand as elements in \( \Lambda^2 \mathbb{L}_p \) considered as tensor product over the field \( \mathbb{R} \). Then the scalar products coincide, i.e. \( \langle f_1 \wedge f_2, f_3 \wedge f_4 \rangle_{\Lambda^2(\Lambda^{2p} \mathbb{C}^{2L})} = \langle f_1 \wedge f_2, f_3 \wedge f_4 \rangle_{\Lambda^2 \mathbb{L}_p} \).

**Proof.** One finds \( \mathcal{J} \mathcal{S} e_{i} = -e_{2L-d+i} \) and \( \mathcal{J} \mathcal{S} e_{2L-d+i} = e_i \) for \( i = 1, \ldots, d \) which implies \( \Lambda^{2p}(\mathcal{J} \mathcal{S}) v^{(p)} = (1)^{2p} v^{(p)} = v^{(p)} \). For \( \mathcal{M} \in \mathcal{G} \) one has \( \mathcal{S} \mathcal{M} \mathcal{S} = \mathcal{J}^* \mathcal{M} \mathcal{J} \) and hence

\[
\langle v^{(p)}, \Lambda^{2p} \mathcal{M} v^{(p)} \rangle = \langle \Lambda^{2p} \mathcal{M} v^{(p)}, \Lambda^{2p} (\mathcal{S} \mathcal{M} \mathcal{S}^2) v^{(p)} \rangle = \langle \Lambda^{2p} (\mathcal{J} \mathcal{S}) v^{(p)}, \Lambda^{2p} (\mathcal{J}^* \mathcal{M} \mathcal{J}) v^{(p)} \rangle = \langle v^{(p)}, \Lambda^{2p} \mathcal{M} v^{(p)} \rangle.
\]

Therefore \( \langle \Lambda^{2p} \mathcal{M} v^{(p)}, \Lambda^{2p} \mathcal{N} v^{(p)} \rangle = \langle v^{(p)}, \Lambda^{2p} (\mathcal{M}^* \mathcal{N}) v^{(p)} \rangle \) is real for all \( \mathcal{M}, \mathcal{N} \in \mathcal{G} \) and by linearity the \( \Lambda^{2p} \mathcal{C}^{2L} \) scalar product for two vectors in \( \mathbb{L}_p \) is real. The second statement follows from the first one using \( \langle f_1 \wedge f_2, f_3 \wedge f_4 \rangle = \langle f_1, f_3 \rangle \langle f_2, f_4 \rangle - \langle f_1, f_4 \rangle \langle f_2, f_3 \rangle \). \( \square \)

Considering \( f_1 \wedge f_2 \) as element in \( \Lambda^2 \mathbb{L}_p \) on one hand and as an element of \( \Lambda^2(\Lambda^{2p} \mathbb{C}^{2L}) \) on the other hand induces an \( \mathbb{R} \)-linear map \( \Lambda^2 \mathbb{L}_p \to \Lambda^2(\Lambda^{2p} \mathbb{C}^{2L}) \). By Lemma 8.5 this map preserves the inner product and is hence injective. Therefore \( \Lambda^2 \mathbb{L}_p \) can be viewed as real subspace of \( \Lambda^2(\Lambda^{2p} \mathbb{C}^{2L}) \).

**Definition 8.1** A subset \( \mathcal{T} \) of \( \mathcal{G} \) is \( \mathbb{L}_p \)-strongly irreducible if there does not exist a finite union \( \mathcal{W} \) of proper linear subspaces of \( \mathbb{L}_p \) such that \( \langle \Lambda^{2p} \mathcal{M} \rangle(\mathcal{W}) = \mathcal{W} \) for any \( \mathcal{M} \) in \( \mathcal{T} \).

**Proposition 8.11** Let \( (\mathcal{Y}_n)_{n \geq 1} \) be a sequence of i.i.d. random matrices in \( \mathcal{G} \) for \( L = 2d \) or \( L = 2d+1 \) and let \( p \) be an integer \( 1 \leq p \leq d \). Let \( \mathcal{T} \) be the semi-group generated by the support of \( \mathcal{Y}_n \). Suppose that \( \mathcal{T} \) is 2p-contracting and \( \mathbb{L}_p \)-strongly irreducible and that \( \mathbb{E}(\log_+ ||\mathcal{Y}_1||) < \infty \). Then \( \gamma_p > \gamma_{p+1} \).

**Proof.** Let \( k \) be the dimension of \( \mathbb{L}_p \) and \( (f_1, \ldots, f_k) \) an orthonormal basis to be chosen later on. For any \( \mathcal{M} \in \mathcal{G} \) let \( \hat{\mathcal{M}} \) denote the matrix in \( \mathbb{G}(k, \mathbb{R}) \) with the entries

\[
\hat{\mathcal{M}}_{i,j} = \langle f_i, \Lambda^{2p} \mathcal{M} f_j \rangle, \quad 1 \leq i, j \leq k.
\]
If \( \mathcal{U} \in \mathcal{G} \cap \mathcal{U}(2L) \), then \( \Lambda^{2p}\mathcal{U} \in \Lambda^{2p}\mathcal{G} \cap \mathcal{U}(\Lambda^{2p}\mathbb{C}^{2L}) \) and hence the restriction of \( \Lambda^{2p}\mathcal{U} \) to \( \mathbb{L}_p \) is orthogonal, i.e. \( \mathcal{U} \in \mathcal{O}(\mathbb{L}_p) \). Let us use the notation \( \Lambda^{2p}\mathcal{M} = \Phi(\mathcal{M}) \). One has \( \|\hat{\mathcal{M}}\| \leq \|\Phi(\mathcal{M})\| \) as \( \mathbb{L}_p \) is a subspace of \( \Lambda^{2p}\mathbb{C}^{2L} \) and by Lemma 8.5 one also obtains \( \|\Lambda^2\hat{\mathcal{M}}\| \leq \|\Lambda^2\Phi(\mathcal{M})\| \).

Claim: Let \( a_1 \geq a_2 \geq \ldots \geq a_d \geq 1 \) be the singular values of \( \mathcal{M} \) as occurring in the decomposition in Lemma 8.4(v), then \( \|\Phi(\mathcal{M})\| = a_1^2 \cdots a_d^2 = \|\hat{\mathcal{M}}\| \) and \( \|\Lambda^2\Phi(\mathcal{M})\| \geq \|\Lambda^2\hat{\mathcal{M}}\| \geq \|\hat{\mathcal{M}}\| \cdot a_1^2 \cdots a_{d-1}^2 a_{d+1}^2 \). In the case \( p = d \), we define \( a_{d+1} = a_d^{-1} \).

Indeed, set \( f_1 = v(p) = e_1 \wedge \ldots \wedge e_p \wedge e_{2L-d-1} \wedge \ldots \wedge e_{2L-d+p} \) and if \( p < d \) set \( f_2 = e_1 \wedge \ldots \wedge e_{p-1} \wedge e_{p+1} \wedge e_{2L-d-1} \wedge \ldots \wedge e_{2L-d+p-1} \wedge e_{2L-d+p+1} \). In the case \( p = d \), set \( f_2 = e_1 \wedge \ldots \wedge e_{d-1} \wedge e_{L+d} \wedge e_{2L-d-1} \wedge \ldots \wedge e_{2L-1} \wedge e_L \). Further, for any \( d \times d \) invertible matrix \( B \) and any matrix \( C \) with \( B^*C = C^*B \), one can construct the following element of \( \mathcal{G} \):

\[
N = \begin{pmatrix}
B & 0 & 0 & 0 & 0 & 0 \\
0 & \cos(\varphi) & 0 & \sin(\varphi) & 0 \\
0 & 0 & (B^*)^{-1} & 0 & 0 \\
C & 0 & 0 & (B^*)^{-1} & 0 \\
0 & -\sin(\varphi) & 0 & \cos(\varphi) & 0 \\
0 & 0 & 0 & 0 & \overline{B}
\end{pmatrix}, \quad \text{if } L = 2d \text{ pencil out the rows and columns containing } \varphi.
\]

(8.32)

Thus for \( p < d \), one readily finds \( N \in \mathcal{G} \) with \( f_2 = \Lambda^{2p}N f_1 \in \mathbb{L}_p \). In the case \( p = d \) define \( N_1 \) by setting \( B = 1 \) and \( C_{i,j} = 0 \) except \( C_{d,d} = 1 \) and define \( N_2 \) by setting \( B = 2 \cdot 1 \), \( C = 0 \). Then one obtains \( (2^{2(d-1)} - 2^{d-2})) f_2 = (2^{2(d-1)} - \Lambda^{2p}N_2)(\Lambda^{2p}N_1 f_1 - f_1) \in L_d \). In conclusion, \( f_1, f_2 \in \mathbb{L}_p \) can be completed to an orthonormal basis of \( \mathbb{L}_p \). Now let us write \( \mathcal{M} = \mathcal{K}\mathcal{D}\mathcal{U} \) as in Lemma 8.4(v), then

\[
\|\Phi(\mathcal{M})\| = a_1^2 \cdots a_p^2 = \|\Lambda^{2p}\mathcal{D} f_1\| = \|\hat{\mathcal{D}} f_1\| \leq \|\hat{\mathcal{D}}\| \leq \|\Lambda^{2p}\mathcal{D}\| = \|\Phi(\mathcal{M})\|
\]

where the last inequality holds as \( \mathbb{L}_p \) is a subspace of \( \Lambda^{2p}\mathbb{C}^{2L} \). Hence \( \|\Phi(\mathcal{M})\| = \|\hat{\mathcal{D}}\| \), but \( \|\hat{\mathcal{D}}\| = \|\hat{\mathcal{K}}\hat{\mathcal{D}}\hat{\mathcal{U}}\| = \|\hat{\mathcal{M}}\| \). As mentioned above, \( \|\Lambda^2\Phi(\mathcal{M})\| \geq \|\Lambda^2\hat{\mathcal{M}}\| \). Furthermore one has

\[
\|\Lambda^2\hat{\mathcal{M}}\| = \|\Lambda^2\hat{\mathcal{D}}\| \geq \|\Lambda^2\hat{\mathcal{D}}(f_1 \wedge f_2)\| = \|\hat{\mathcal{M}}\| \cdot a_1^2 \cdots a_{p-1}^2 a_{p+1}^2
\]

Hence the claim is proved.

Let \( \hat{T} \) be the semi group induced by the distribution of \( \hat{\mathcal{J}}_1 \). As \( T \in \mathbb{L}_p \)-strongly irreducible, clearly \( \hat{T} \) is a strongly irreducible subset of \( \text{Gl}(k, \mathbb{R}) \). As \( T \) is also \( 2p \)-contracting, there exists a sequence \( (\mathcal{M}_n)_{n \geq 1} \) in \( T \) such that \( \lim_{n \to \infty} \|\Phi(\mathcal{M}_n)\|^2 \|\Lambda^2\Phi(\mathcal{M}_n)\|^{-1} = \infty \). As \( \|\hat{\mathcal{M}}_n\| = \|\Phi(\mathcal{M}_n)\| \) and \( \|\Lambda^2\Phi(\mathcal{M}_n)\| \geq \|\Lambda^2\hat{\mathcal{M}}_n\| \) by the above claim, one obtains

\[
\lim_{n \to \infty} \|\hat{\mathcal{M}}_n\|^2 \|\Lambda^2\hat{\mathcal{M}}_n\|^{-1} \geq \lim_{n \to \infty} \|\Phi(\mathcal{M}_n)\|^2 \|\Lambda^2\Phi(\mathcal{M}_n)\|^{-1} = \infty.
\]

Hence \( \hat{T} \) is contracting.
The two biggest Lyapunov exponents associated to the sequence \( (\mathcal{Y}_n)_{n \geq 1} \) shall be denoted by \( \hat{\gamma}_1 \) and \( \hat{\gamma}_2 \). Then by the claim, the definition of Lyapunov exponents and [9, A.III.6.1] one has
\[
2 \sum_{i=1}^{p} \gamma_i = \hat{\gamma}_1 > \hat{\gamma}_2 \geq 2 \sum_{i=1}^{p-1} \gamma_i + 2\gamma_{p+1},
\]
implying \( \gamma_p > \gamma_{p+1} \). By definition of \( a_{p+1} \) one actually would have to replace \( \gamma_{p+1} \) by \( \gamma_{p+2} = \gamma_{d+2} \) in the case \( L = 2d + 1, p = d \). Then one gets \( \gamma_d > \gamma_{d+2} = -\gamma_d \) and therefore \( \gamma_d > 0 = \gamma_{d+1} \).

\[\textbf{Theorem 8.7} \] Let \( (\mathcal{Y}_n)_{n \geq 1} \) be a sequence of i.i.d. random matrices in \( \mathbb{G} \) for \( L = 2d \) or \( L = 2d + 1 \). Let \( \mathbb{T} \) be the semi-group induced by the support of \( \mathcal{Y}_1 \) and let \( \mathbb{E}(\log_+ ||\mathcal{Y}_1||) < \infty \). Suppose that \( \mathbb{T} \) is Zariski dense in \( \mathbb{G} \), then all Lyapunov exponents are distinct.

\[\textbf{Proof.} \] According to the proof of Proposition 8.11 the inequality \( \gamma_p > \gamma_{p+1} \) follows from the fact that the semi-group \( \hat{T} = \{ \hat{\mathcal{M}} | \mathcal{M} \in \mathbb{T} \} \) is strongly irreducible and contracting in \( \text{GL}(k, \mathbb{R}) \) as defined above. Now \( \hat{T} \) is Zariski dense in \( \hat{\mathbb{G}} = \{ \hat{\mathcal{M}} | \mathcal{M} \in \mathbb{G} \} \). Otherwise there would be a polynomial \( \hat{P} \) on \( \text{GL}(k, \mathbb{R}) \) such that \( \hat{P}(\hat{T}) = 0 \) and \( \hat{P}(\hat{\mathcal{M}}) \neq 0 \) for some \( \mathcal{M} \in \mathbb{G} \). As the entries in \( \hat{\mathcal{M}} \) are polynomials of the entries in \( \mathcal{M} \), this leads to a polynomial \( P \) on \( \text{GL}(2L, \mathbb{C}) \) such that \( P(\mathbb{T}) = 0 \) and \( P(\mathcal{M}) \neq 0 \) for some \( \mathcal{M} \in \mathbb{G} \), contradicting the fact that \( \mathbb{T} \) is Zariski dense in \( \mathbb{G} \).

Now suppose \( \hat{T} \) is not strongly irreducible. Then there would be a finite union of proper subspaces \( \mathcal{W} = \mathcal{V}_1 \cup \ldots \cup \mathcal{V}_n \) such that \( \mathcal{M}(\mathcal{W}) \subseteq \mathcal{W} \) for all \( \mathcal{M} \in \mathbb{T} \). The property \( \mathcal{M}(\mathcal{V}_i) \subseteq \mathcal{V}_k \) can be written as \( \langle w, \mathcal{M}v \rangle = 0 \) for all \( w \in \mathcal{V}_k, v \in \mathcal{V}_i \). Hence the set of all such matrices \( \mathcal{M} \) is Zariski closed. The property \( \mathcal{M}(\mathcal{W}) \subseteq \mathcal{W} \) is therefore a finite intersection of finite unions of Zariski closed sets and hence Zariski closed. As \( \hat{T} \) is Zariski dense in \( \hat{\mathbb{G}} \), this then implies \( \hat{\mathbb{G}}(\mathcal{W}) \subseteq \mathcal{W} \). Therefore, if \( \mathbb{G} \) is strongly irreducible, then also \( \mathbb{T} \) is.

To show that \( \hat{T} \) is contracting we want to use Theorem 6.3 of [32] which states that if the algebraic closure of \( \hat{T} \) is strongly irreducible and contracting, then also \( \mathbb{T} \) is contracting. Hence it is only left to show that \( \mathbb{G} \) is strongly irreducible and contracting.

The property of \( \hat{\mathbb{G}} \) to be strongly irreducible is equivalent to \( \mathbb{G} \) being \( \mathbb{L}_p \)-strongly irreducible. As \( \mathbb{G} \) is connected we have to show that there is a proper subspace \( \mathcal{V} \subseteq \mathbb{L}_p \) such that \( (\Lambda^{2p} \mathcal{M})(\mathcal{V}) \subseteq \mathcal{V} \) for all \( \mathcal{M} \in \mathbb{G} \). Suppose such a \( \mathcal{V} \) exists. For \( a_1 > a_2 > \ldots > a_d > 1 \) take \( \mathcal{D} = \text{diag}(a_1, \ldots, a_d, 1, a_1^{-1}, \ldots, a_d^{-1}, 1, a_1, \ldots, a_d) \). The relation \( \langle \Lambda^{2p} \mathcal{D}^n \mathcal{V}, \mathcal{V} \rangle \subseteq \mathcal{V} \) implies that either \( v^{(p)} \in \mathcal{V} \), but then \( \mathbb{L}_p = \mathcal{V} \) or that \( v^{(p)} \) is in the orthogonal complement \( \mathcal{V}^\perp \). But then by Lemma 8.4(ii) one has for \( v \in \mathcal{V} \)
\[
\langle \Lambda^{2p} \mathcal{M}v^{(p)}, v \rangle = \langle v^{(p)}, \Lambda^{2p} \mathcal{M}^* v \rangle = 0
\]
for any \( \mathcal{M} \in \mathbb{G} \) and hence \( \mathbb{L}_p = \mathcal{V}^\perp \). Therefore \( \mathcal{V} \) is not proper.

Now it is only left to show that \( \hat{\mathbb{G}} \) is contracting. By the proof of Proposition 8.11 this follows if \( \mathbb{G} \) is \( 2p \)-contracting. Therefore take a matrix \( \mathcal{M} \) of the form (8.32) with \( C = 0 \).
and $B = \text{diag}(\lambda_1, \ldots, \lambda_d)$. such that all moduli of the eigenvalues are distinct except for
the fact that always two eigenvalues have the same modulus. The sequence $\mathcal{M}^k$ then shows
that $\mathbb{G}$ is $2p$-contracting. \hfill \Box.

8.7 Absence of singular spectrum

In this section we only consider the random model described at the end of Section 8.4. For any configuration $\omega = ((\lambda_{j,k})_{k=1,\ldots,K,j\in\mathbb{Z}}, (V_j)_{j\in\mathbb{Z}}, s) \in \Omega$ let $\tilde{\omega}$ denote $\omega$ excluded the
singular potential $\mathcal{V} = \mathcal{V}_0$ at $s$, i.e. $\tilde{\omega} = ((\lambda_{j,k})_{k=1,\ldots,K,j\in\mathbb{Z}}, (V_j)_{j\in\mathbb{Z},j\neq 0}, s)$. The distribution
of $\tilde{\omega}$ shall be denoted by $\tilde{\mathbf{P}}$ and that of $\mathcal{V}$ by $\mathbf{P}_\mathcal{V}$. With these notations $\mathbf{P} = \tilde{\mathbf{P}} \times \mathbf{P}_\mathcal{V}$. We only
consider the case where $L$ is odd and $\mathbf{P}_\mathcal{V}$ is absolutely continuous w.r.t. to the Lebesgue
measure. Next recall the definition (8.18) of $\tilde{\mathbf{P}}$. Note that $\tilde{\mathcal{V}}$ is only defined
for almost every $\mathcal{V}$ and for almost every $\tilde{\mathcal{V}}$ there is a pre-image $\mathcal{V}$, which is not necessarily
unique. Furthermore the pre-images of zero sets are zero sets and hence the distribution $\mathbf{P}_\mathcal{V}$
of $\tilde{\mathcal{V}}$, i.e. the image measure of $\mathbf{P}_\mathcal{V}$, is absolutely continuous w.r.t. the Lebesgue measure
on the vector space $\mathcal{J}^{so^+}(2L)$.

As $\mathcal{V}$ denotes the singular potential at $x_0 = s$, let $\tilde{G}_{\mathcal{V}}$ denote the averaged Green matrix
at the point $x_0 = s$, that is, $\tilde{G}_{\mathcal{V}} = \tilde{G}_{\mathcal{V}}(s)$ with the notations of Proposition 8.6. Note
that this matrix actually depends on $\omega = (\tilde{\omega}, \mathcal{V})$, but in most of the arguments below $\tilde{\omega}$
will be fixed. Furthermore, Proposition 8.8 shows that $\tilde{G}_{\mathcal{V}}$ actually only depends on $\mathcal{V}$
(which is a real statement statement since the map $\mathcal{V} \mapsto \tilde{\mathcal{V}}$ is not injective). Hence it is
sufficient to prove almost sure statements w.r.t. the distribution $\mathbf{P}_\mathcal{V}$ of $\tilde{\mathcal{V}}$ instead of w.r.t.
the distribution $\mathbf{P}_\mathcal{V}$ of $\mathcal{V}$.

Let $\mu_\omega = \mu_{\tilde{\omega},\mathcal{V}}$ denote the associated positive matrix valued measure. The function $E \mapsto \frac{1}{1 + E^2}$ is in $L^1(\mu_\omega)$ for all $\omega$. On the set of such measures one may introduce the weak-*
topology induced by the functions $E \mapsto \Im m((E - z)^{-1})$ for $z$ in the upper half plane. As
the pairing of this function with the measure $\mu_\omega$ is just $\Im m(G^2)$ with a $z$ independent
self-adjoint matrix $A$, it follows that the map $\omega \mapsto \mu_\omega$ is Borelian. Finally let $\mu_{\omega,k} = \mu_{\tilde{\omega},\mathcal{V},k}$
denote the measure corresponding to $e_k^\ast \tilde{G}_{\mathcal{V}} e_k$ where $e_k$ is the $k$-th canonical basis vector of $\mathbb{C}^{2L}$.

The aim of this section is to prove that almost surely in $\omega$ the measure $\mu_\omega$ is absolutely
continuous or equivalently, that its singular part vanishes, i.e. $\mu_{\omega,\text{sing}}(\mathbb{R}) = 0$. Therefore we
will first show that almost surely one only needs to consider $\mu_{\omega,1}$ and then we show that
$\mu_{\omega,1,\text{sing}}(\mathbb{R}) = 0$ almost surely. To obtain the first part we compare the measures $\mu_{\tilde{\omega},\mathcal{V},1}$ and $\mu_{\omega,\mathcal{V},k}$ for fixed $\tilde{\omega}$ and show that they are almost surely equivalent. Once cyclicity issues
are settled (Proposition 8.12) and matrix analogues of rank one perturbation results are
proven (Proposition 8.13), the proofs are basically modifications of the arguments of [41]. Our starting point are the following observations linked to Kramers degeneracy.

Lemma 8.6 For $1 \leq k, l \leq L$ let us introduce the $2L \times 2$ matrix $\Psi_k = (e_k, e_{k+L})$.
(i) Let $j$ denote the $2 \times 2$ symplectic form, then $J \Psi_k^\ast = \Psi_k j$.
Furthermore one has $\Psi_k \Psi_k^\ast \in \mathcal{J}^{so^+}(2L)$ and $\Psi_k j \Psi_l^\ast + \Psi_l j^\ast \Psi_k \in \mathcal{J}^{so^+}(2L)$. 

Lemma 8.7 For $1 \leq k, l \leq L$ let us introduce the $2L \times 2$ matrix $\Psi_k = (e_k, e_{k+L})$.
(i) Let $j$ denote the $2 \times 2$ symplectic form, then $J \Psi_k^\ast = \Psi_k j$.
Furthermore one has $\Psi_k \Psi_k^\ast \in \mathcal{J}^{so^+}(2L)$ and $\Psi_k j \Psi_l^\ast + \Psi_l j^\ast \Psi_k \in \mathcal{J}^{so^+}(2L)$.
(ii) For $\mathcal{Y}_1, \mathcal{Y}_2 \in J\text{so}^s(2L)$ one has $\mathcal{Y}_1 \mathcal{Y}_2 \mathcal{Y}_1 \in J\text{so}^s(2L)$.
(iii) $\Psi^*_k \hat{G}_V^* \Psi_k$ is a multiple of the unity matrix, which means $\Psi^*_k \hat{G}_V^* \Psi_k = e^*_k \hat{G}_V^* e_k 1$.

**Proof.** The identity $\mathcal{J} \Psi_k = \Psi_k j$ is readily verified. Furthermore $(\Psi_k \Psi^*_k)^* = \Psi_k \Psi^*_k$ and one has $\mathcal{J}^* \Psi_k \Psi^*_k \mathcal{J} = \Psi_k j^* j \Psi^*_k = \Psi_k \Psi^*_k = (\Psi_k \Psi^*_k)^t$ showing $\Psi_k \Psi^*_k \in J\text{so}^s(2L)$. Similar calculations show $\Psi_k j \Psi^*_j + \Psi_j j^* \Psi_k \in J\text{so}^s(2L)$ and (i) is proved. To obtain (ii), first note that $\mathcal{Y}_1, \mathcal{Y}_2$ are self-adjoint and hence $\mathcal{Y}_1 \mathcal{Y}_2 \mathcal{Y}_1$ is self-adjoint. Furthermore one has $\mathcal{J}^* \mathcal{Y}_1 \mathcal{Y}_2 \mathcal{Y}_1 \mathcal{J} = \mathcal{J}^* \mathcal{Y}_1 \mathcal{J} \mathcal{J}^* \mathcal{Y}_2 \mathcal{J} \mathcal{J}^* \mathcal{Y}_1 \mathcal{J} = \mathcal{Y}_1 \mathcal{Y}_2 \mathcal{Y}_1 \mathcal{Y}_1 = (\mathcal{Y}_1 \mathcal{Y}_2 \mathcal{Y}_1 \mathcal{Y}_1)^t$ and also (ii) is proved. (iii) is just a special case of Proposition 8.9. □

The measure class of $\mu_\omega$ is given by the trace, i.e. by the sum $\sum_{k=1}^{2L} \mu_{\omega,k} = 2 \sum_{k=1}^{L} \mu_{\omega,k}$, where the last identity follows from Lemma 8.6(iii).

**Proposition 8.12** For fixed $\hat{\omega}$, one has that for Lebesgue almost all $\hat{\mathcal{V}} \in J\text{so}^s(2L)$ the set of energies $\{E \in \mathbb{R} | \hat{G}_V^{\hat{\omega}+i0} \text{ exists and } \Psi^*_i \hat{G}^*_V \Psi_k \text{ is invertible} \}$ has full Lebesgue measure.

**Proof.** We first claim that for fixed $z$ in the upper half plane $\mathbb{U}_1$, there is a $\hat{\mathcal{V}} \in J\text{so}^s(2L)$ such that $\Psi^*_i \hat{G}^*_V \Psi_k$ is invertible. Recall that $\hat{G}_V^* = ((\hat{G}_0^*)^{-1} + \hat{\mathcal{V}})^{-1}$. Set $(\hat{G}_0^*)^{-1} = \mathcal{X} + i \mathcal{Y}^{-1}$ with $\mathcal{Y}^{-1} = -\Im m((\hat{G}_0^*)^{-1}) > 0$. As $\mathcal{J}^* \hat{G}_0^* \mathcal{J} = (\hat{G}_0^*)^t$, one has $\mathcal{X}, \mathcal{Y}^{-1}, \mathcal{Y} \in J\text{so}^s(2L)$. Then consider $\hat{\mathcal{V}} = -\Re e((\hat{G}_0^*)^{-1}) + \lambda \mathcal{P}$ with a perturbation $\mathcal{P} \in J\text{so}^s(2L)$. Then

$$\hat{G}_V^* = (\lambda \mathcal{Y})^{-1} = i \mathcal{Y} + \lambda \mathcal{YPY} - i \lambda^2 \mathcal{YPY} \mathcal{YPY} + O(\lambda^3).$$

Note that $\mathcal{V}$ now depends on $\lambda$ and $\mathcal{P}$, furthermore $\mathcal{YPY} \in J\text{so}^s(2L)$ as well as $\mathcal{YPY} \mathcal{YPY} \in J\text{so}^s(2L)$ by Lemma 8.6. For any $2 \times 2$ matrices $A, B, C$ one has $\det(A + \lambda B + \lambda^2 C) = \det(A) + \lambda \text{Tr}(A(j^* Bj)^t) + \lambda^2 \left(\det(B) + \text{Tr}(A(j^* C j)^t)\right) + O(\lambda^3)$. Furthermore for $\mathcal{W} \in J\text{so}^s(2L)$, one has $(j^* \Psi^*_i \mathcal{W} \Psi_k j)^t = j^* \Psi^*_k \mathcal{W} \Psi_i j = \Psi^*_k \mathcal{J}^* \mathcal{W} \mathcal{J} \Psi_i = \Psi^*_k \mathcal{W} \Psi_i$. Thus from the above

$$\det(\Psi^*_i \hat{G}_V^* \Psi_k) = i \det(\Psi^*_i \mathcal{Y} \Psi_k) + i \lambda \text{Tr}(\Psi^*_i \mathcal{Y} \Psi_k \Psi^*_k \mathcal{YPY} \Psi_i) + \lambda^2 \left(\det(\Psi^*_i \mathcal{YPY} \Psi_k) - i \text{Tr}(\Psi^*_i \mathcal{YPY} \Psi_k \Psi^*_k \mathcal{YPY} \Psi_i)\right) + O(\lambda^3) \quad (8.33)$$

If $\det(\Psi^*_i \mathcal{Y} \Psi_k) \neq 0$, then the claim is true (just take $\lambda = 0$). If $\det(\Psi^*_i \mathcal{Y} \Psi_k) = 0$, but $\Psi^*_i \mathcal{Y} \Psi_k \neq 0$, then set $\mathcal{P} = \mathcal{Y}^{-1} \in J\text{so}^s(2L)$ and (8.33) reduces to

$$\det(\Psi^*_i \hat{G}_V^* \Psi_k) = i \lambda \text{Tr}(\Psi^*_i \mathcal{Y} \Psi_k) + O(\lambda^2).$$

Since the coefficient before $\lambda$ only vanishes if $\Psi^*_i \mathcal{Y} \Psi_k = 0$, this is not equal to zero for small $\lambda$ and the claim holds again. Finally, if $\Psi^*_i \mathcal{Y} \Psi_k = 0$, then set $\mathcal{P} = \mathcal{Y}^{-1} \Psi^*_k + \Psi^*_k j^* \Psi_i + \Psi^*_i j^* \Psi_k$ which lies in $J\text{so}^s(2L)$ by Lemma 8.6 part (i) and (ii). Then (8.33) reduces to

$$\det(\Psi^*_i \mathcal{Y} \Psi_k) = \lambda^2 \det(\Psi^*_i \Psi^*_j \Psi^*_k) + \lambda^3(\lambda^3) = \lambda^2 + O(\lambda^3),$$

where we used $\Psi^*_i \Psi_k = \delta_{i,k}$. Hence this determinant is again not zero for small $\lambda$. Thus for all cases we find some $\hat{\mathcal{V}}$ such that $\Psi^*_i \hat{G}^*_V \Psi_k$ is invertible and the claim is proved.
By definition of the determinant and Cramer’s rule the function \( \hat{V} \mapsto \det(\Psi I G V \Psi_k) = \det(\Psi I G V (G_v^\ep)^{-1} + \hat{V})^{-1} \Psi_k) \) is a rational function on the vector space \( J^* \) which does not vanish completely by the claim above, therefore it does not vanish for Lebesgue almost every \( \hat{V} \in J^* \) w.r.t. the Lebesgue measure on \( J^* \).

Next recall that the boundary values \( G_v^{E+\ep} \) exist almost surely in \( E \) by analyticity. For \( \hat{V} \) as described above, the map \( z \mapsto \det(\Psi I G V \Psi_k) \) is analytic in the upper half plane and does not vanish identically. Therefore for Lebesgue almost every \( E \), \( G_v^{E+\ep} \) exists and one has \( \det(\Psi I G V (G_v^\ep)^{-1} \Psi_k) \neq 0 \).

\[ \square \]

**Proposition 8.13** Let \( \hat{\omega} \) and \( \hat{V} \in J^* \) be fixed and define \( \hat{V}_\lambda = \hat{V} + \lambda \Psi_k \Psi_k^* \).

1. The set \( A_{V,k} = \{ E \in \mathbb{R} | \Psi_k G V \Psi_k \text{ exists and } \Im(\Psi_k G V \Psi_k) > 0 \} \) is independent of \( \lambda \) and it is an essential support of the absolutely continuous part of \( \mu_{\hat{\omega},V,k} \).
2. The singular part of \( \mu_{\hat{\omega},V,k} \) is supported on the set \( \{ E \in \mathbb{R} | \Psi_k G V \Psi_k = -\lambda^{-1} \} \).
3. For any \( B \subset \mathbb{R} \) of zero Lebesgue measure, we have \( \mu_{\hat{\omega},V,k}(B) = 0 \) for Lebesgue a.e. \( \lambda \in \mathbb{R} \).

**Proof.** (i) We prove that \( A_{V,k} = A_{V_0,k} \subset A_{V,k} \) for all \( \lambda \); the other inclusion can be obtained analogously. Hence let \( E \in A_{V,k} \). We first claim that \( 1 + \lambda \Psi_k \Psi_k^* G V^{E+\ep} \) is invertible. Suppose \( (1 + \lambda \Psi_k \Psi_k^* G V^{E+\ep}) v = 0 \). Then \( v \) is in the range of \( \Psi_k \) and there are \( \alpha, \eta \in \mathbb{C} \) such that \( v = \alpha e_k + \beta e_{k+L} \). We use \( e_k^* G V^{E+\ep} e_{k+L} = 0 = e_{k+L}^* G V^{E+\ep} e_k \) following from \( J^* G V^* J = (G_v^*)^t \). Thus

\[
\alpha = -\lambda \alpha e_k^* G V^{E+\ep} e_k, \quad \beta = -\lambda \beta e_{k+L}^* G V^{E+\ep} e_{k+L}.
\]

But as \( \Im(e_k^* G V^{E+\ep} e_k) = \Im(e_{k+L}^* G V^{E+\ep} e_{k+L}) > 0 \) for \( E \in A_{V,k} \) this implies \( \alpha = 0 = \beta \) and hence \( v = 0 \). Therefore the kernel of \( 1 + \lambda \Psi_k \Psi_k^* G V^{E+\ep} \) is indeed trivial. Hence by Proposition 8.8, \( G V_{\lambda}^{E+\ep} = (1 + \lambda \Psi_k \Psi_k^* G V^{E+\ep})^{-1} \) exists. Furthermore, also by Proposition 8.8,

\[
\Im(G V_{\lambda}^{E+\ep}) = \left[(1 + \lambda \Psi_k \Psi_k^* G V^{E+\ep})^{-1}\right]^* \Im(G V^{E+\ep})(1 + \lambda \Psi_k \Psi_k^* G V^{E+\ep})^{-1},
\]

and \( (1 + \lambda \Psi_k \Psi_k^* G V^{E+\ep})^{-1} \) leaves the space spanned by \( e_k \) and \( e_{k+L} \) invariant. Therefore one also obtains \( \Im(\Psi_k G V^{E+\ep} \Psi_k) > 0 \) showing \( E \in A_{V,k} \).

(ii) From (8.19),

\[
\hat{G}_V^\ep = \hat{G}_V^\ep + G V[(G_v^\ep)^{-1} - (G_v^\ep)^{-1}]G V^\ep = \hat{G}_V^\ep - \lambda G V^\ep \Psi_k \Psi_k^* G V_{\lambda}^\ep,
\]

and hence \( \Psi_k \hat{G}_V^\ep \Psi_k = (1 + \lambda \Psi_k^* G V \Psi_k)^{-1} \Psi_k^* G V \Psi_k \). Thus Lemma 8.6(iii) implies

\[
e_k^* G V_{\lambda}^\ep e_k = (1 + \lambda e_k^* G V e_k)^{-1} e_k^* G V e_k.
\]

Thus in the limit \( \epsilon \downarrow 0 \), \( e_k^* G V^{E+\ep} e_k \to 0 \) if and only if \( \Psi_k G V^{E+\ep} \Psi_k \to -\lambda^{-1} \).
(iii) From (8.35) one deduces that the map \( \lambda \mapsto \mu_{\omega,\psi_{\lambda},k} \) is integrable in the \(*\)-weak topology over intervals \([a,b]\). Taking imaginary parts of (8.35), one obtains

\[
\Im m(e^*_k \hat{G}_{\psi_{\lambda},k}^* e_k) = \frac{\Im m(e^*_k \hat{G}_{\psi_{\lambda},k}^* e_k)}{(1 + \lambda \Re (e^*_k \hat{G}_{\psi_{\lambda},k}^* e_k))^2 + (\lambda \Im m(e^*_k \hat{G}_{\psi_{\lambda},k}^* e_k))^2}.
\]

Let \( x = \Re (e^*_k \hat{G}_{\psi_{\lambda},k}^* e_k) \) and \( y = \Im m(e^*_k \hat{G}_{\psi_{\lambda},k}^* e_k) \). Then \( \arctan(x/y) \) is an anti-derivative of the function \( \lambda \mapsto \Im m(e^*_k \hat{G}_{\psi_{\lambda},k}^* e_k) \). Therefore \( \int_a^b d\lambda \Im m(e^*_k \hat{G}_{\psi_{\lambda},k}^* e_k) \) is bounded by \( \pi \) and the integral over the whole real line exists and is equal to \( \pi \). This means that the integral \( \int_{-\infty}^{\infty} d\lambda \mu_{\omega,\psi_{\lambda},k} \) actually converges to the Lebesgue measure which has no singular part.

Now let \( B \) be a set of Lebesgue measure zero. Then \( \int_{-\infty}^{\infty} d\lambda \mu_{\omega,\psi_{\lambda},k}(B) = 0 \). As the measures are positive this means that for Lebesgue a.e. \( \lambda \in \mathbb{R} \) one has \( \mu_{\omega,\psi_{\lambda},k}(B) = 0 \). \( \square \)

Note that the equation proved in part (iii) above, \( dE = \int_{-\infty}^{\infty} d\lambda \mu_{\omega,\psi_{\lambda},k}(dE) \), is well-known from the theory of rank one perturbations.

**Theorem 8.8** Let \( \omega = (\hat{\omega}, \mathcal{V}) \) be fixed such that the matrices \( \Psi_1^* \hat{G}_{\psi_{\lambda}}^{E+\delta} \Psi_k \), \( \Psi_1^* \hat{G}_{\psi_{\lambda}}^{E+\delta} \Psi_1 \) as well as \( \Psi_1^* \hat{G}_{\psi_{\lambda}}^{E+\delta} \Psi_k \) exist and are invertible for Lebesgue almost all \( E \). Set \( \psi_{\lambda} = \hat{\psi} + \lambda \Psi_k \Psi_k^* \). Then for Lebesgue almost all \( \lambda \in \mathbb{R} \), the measure \( \mu_{\omega,\psi_{\lambda},k} \) is absolutely continuous w.r.t. \( \mu_{\omega,\psi_{\lambda},1} \).

**Proof.** By the Radon-Nikodym theorem we can write \( \mu_{\omega,\psi_{\lambda},k} = f_k \mu_{\omega,\psi_{\lambda},1} + \tilde{\mu}_k \) where \( f_k \) is a function and \( \tilde{\mu}_k \) is the part of \( \mu_{\omega,\psi_{\lambda},k} \) which is singular to \( \mu_{\omega,\psi_{\lambda},1} \). The statement of the theorem is that \( \tilde{\mu}_k = 0 \) for Lebesgue almost all \( \lambda \).

In order to show this, we first need to verify a few identities. By multiplying (8.34) with \( \Psi_k^* \) from the left and \( \Psi_1 \) from the right, one obtains

\[
\Psi_k^* \hat{G}_{\psi_{\lambda}} \Psi_1 = (1 + \lambda \Psi_k^* \hat{G}_{\psi_{\lambda}} \Psi_k)^{-1} \Psi_k^* \hat{G}_{\psi_{\lambda}} \Psi_1 = \frac{\Psi_k^* \hat{G}_{\psi_{\lambda}} \Psi_1}{1 + \lambda \Psi_k^* \hat{G}_{\psi_{\lambda}} \Psi_k} \quad (8.36)
\]

where the last identity follows from Lemma 8.6(iii). From (8.34), one also obtains

\[
\Psi_1^* \hat{G}_{\psi_{\lambda}} \Psi_1 = \Psi_1^* \hat{G}_{\psi_{\lambda}} \Psi_1 - \lambda \Psi_1^* \hat{G}_{\psi_{\lambda}} \Psi_k \Psi_k^* \hat{G}_{\psi_{\lambda}} \Psi_1. \quad (8.37)
\]

Inserting (8.36) in (8.37) gives

\[
\Psi_1^* \hat{G}_{\psi_{\lambda}} \Psi_1 = \Psi_1^* \hat{G}_{\psi_{\lambda}} \Psi_1 - \lambda \frac{\Psi_1^* \hat{G}_{\psi_{\lambda}} \Psi_k \Psi_k^* \hat{G}_{\psi_{\lambda}} \Psi_1}{1 + \lambda \Psi_k^* \hat{G}_{\psi_{\lambda}} \Psi_k}. \quad (8.38)
\]

Furthermore, it follows from (8.35) that

\[
1 + \lambda \Psi_k^* \hat{G}_{\psi_{\lambda}} \Psi_k = \frac{\Psi_k^* \hat{G}_{\psi_{\lambda}} \Psi_1}{\Psi_k^* \hat{G}_{\psi_{\lambda}} \Psi_k}. \quad (8.39)
\]
Now let $A \subset \mathbb{R}$ be the set of all $k$ where the limit $\hat{G}_V^{E+i\theta}$ exists and all four matrices $\Psi_k^* \hat{G}_V^{E+i\theta} k, \Psi_k^* \hat{G}_V^{E+i\theta} k, \Psi_k^* \hat{G}_V^{E+i\theta} k, \Psi_k^* \hat{G}_V^{E+i\theta} k$ are invertible. By assumption, the set $A$ has full Lebesgue measure and thus by Proposition 8.13(iii) we have $\mu_{\infty, \lambda, k} = \mu_{\infty, \lambda, k}|A$ for Lebesgue a.e. $\lambda \in \mathbb{R}$. Thus we can restrict the measures to the set $A$. We consider the absolutely continuous and singular part of $\mu_{\infty, \lambda, k}$ (w.r.t. the Lebesgue measure) separately and begin with the singular part. Inserting (8.39) into (8.38) and dividing by $e_k^\lambda \hat{G}_V^{E+i\theta} k$ gives

$$\frac{\Psi_k^* \hat{G}_V^{E+i\theta} k \Psi_k^* \hat{G}_V^{E+i\theta} k}{e_k^\lambda \hat{G}_V^{E+i\theta} k} = \frac{\Psi_k^* \hat{G}_V^{E+i\theta} k \Psi_k^* \hat{G}_V^{E+i\theta} k}{e_k^\lambda \hat{G}_V^{E+i\theta} k} - \lambda \frac{\Psi_k^* \hat{G}_V^{E+i\theta} k \Psi_k^* \hat{G}_V^{E+i\theta} k}{e_k^\lambda \hat{G}_V^{E+i\theta} k}.$$ 

Let $E \in A$. Then taking $z = E + i\epsilon$ and the limit $\epsilon \downarrow 0$, it follows that

$$\lim_{\epsilon \downarrow 0} \frac{\Psi_k^* \hat{G}_V^{E+i\epsilon} k \Psi_k^* \hat{G}_V^{E+i\epsilon} k}{e_k^\lambda \hat{G}_V^{E+i\epsilon} k} = \lim_{\epsilon \downarrow 0} \frac{\Psi_k^* \hat{G}_V^{E+i\epsilon} k \Psi_k^* \hat{G}_V^{E+i\epsilon} k}{e_k^\lambda \hat{G}_V^{E+i\epsilon} k} - \lambda \frac{\Psi_k^* \hat{G}_V^{E+i\epsilon} k \Psi_k^* \hat{G}_V^{E+i\epsilon} k}{e_k^\lambda \hat{G}_V^{E+i\epsilon} k},$$

where the last term exists and is not zero (except for $\lambda = 0$) by the invertibility assumptions for $E \in A$. Since $|e_k^\lambda \hat{G}_V^{E+i\epsilon} k| \rightarrow \infty$ as $\epsilon \downarrow 0$ for a.e. $E$ w.r.t. the singular part of $\mu_{\infty, \lambda, k}$ and since, by Lemma 8.6(iii), the matrix on the l.h.s. is a multiple of $1$, one obtains

$$\lim_{\epsilon \downarrow 0} \frac{e_k^\lambda \hat{G}_V^{E+i\epsilon} k}{e_k^\lambda \hat{G}_V^{E+i\epsilon} k} \neq 0$$

for every $\lambda \neq 0$ and a.e. $E \in A$ w.r.t. the singular part of $\mu_{\infty, \lambda, k}|A$. This implies that the singular part of $\hat{\mu}_\lambda|A$ vanishes for every $\lambda \neq 0$ and thus the singular part of $\hat{\mu}_\lambda$ vanishes also for Lebesgue a.e. $\lambda \in \mathbb{R}$.

It remains to consider the absolutely continuous part of $\hat{\mu}_\lambda$. Multiplying both sides of (8.38) with $|1 + \lambda e_k^\lambda \hat{G}_V^{E+i\epsilon} k|^2$ and taking imaginary parts gives

$$|1 + \lambda e_k^\lambda \hat{G}_V^{E+i\epsilon} k|^2 \Im(\Psi_k^* \hat{G}_V^{E+i\epsilon} k) =$$

$$|1 + \lambda e_k^\lambda \hat{G}_V^{E+i\epsilon} k|^2 \Im(\Psi_k^* \hat{G}_V^{E+i\epsilon} k) - \lambda \Im(\Psi_k^* \hat{G}_V^{E+i\epsilon} k \Psi_k^* \hat{G}_V^{E+i\epsilon} k)$$

$$+ \lambda^2 \left[ \Im(e_k^\lambda \hat{G}_V^{E+i\epsilon} k) \Re(\Psi_k^* \hat{G}_V^{E+i\epsilon} k \Psi_k^* \hat{G}_V^{E+i\epsilon} k) - \Re(e_k^\lambda \hat{G}_V^{E+i\epsilon} k) \Im(\Psi_k^* \hat{G}_V^{E+i\epsilon} k \Psi_k^* \hat{G}_V^{E+i\epsilon} k) \right].$$

(8.40)

For $z \in U_1$, the r.h.s. of (8.40) is a second order polynomial in $\lambda$ which we denote by $P(z, \lambda)$. For $z = E + i\epsilon$ and $E \in A$, it converges as $\epsilon \downarrow 0$ to a limiting polynomial $P(E + i0, \lambda)$. As above consider

$$A_{V, k} = \left\{ E \in \mathbb{R} \left| \hat{G}_V^{E+i0} \text{ exists and } \Im(e_k^\lambda \hat{G}_V^{E+i0} k) > 0 \right. \right\}.$$ 

Claim: For $E \in A \cap A_{V, k}$, $P(E + i0, \lambda)$ cannot vanish identically as polynomial in $\lambda$.

Suppose the contrary. Then by considering the constant and the linear term one deduces $
\Im(\Psi_k^* \hat{G}_V^{E+i0} k) = 0$ and

$$\Im(\Psi_k^* \hat{G}_V^{E+i0} k \Psi_k^* \hat{G}_V^{E+i0} k) = 0.$$ 

Finally the quadratic term then gives

$$\Im(e_k^\lambda \hat{G}_V^{E+i0} k) \Re(\Psi_k^* \hat{G}_V^{E+i0} k \Psi_k^* \hat{G}_V^{E+i0} k) = 0.$$ 

As $E \in A_{V, k}$, this now implies
that one also has \( \Re \left( \Psi^*_k \hat{G}^E_{\lambda} + 0 \Psi_k \Psi^*_k \hat{G}^E_{\lambda} + 0 \Psi_1 \right) = 0 \) so that \( \Psi^*_k \hat{G}^E_{\lambda} + 0 \Psi_k \Psi^*_k \hat{G}^E_{\lambda} + 0 \Psi_1 = 0 \).

This is not the case for \( E \in A \) and hence the claim holds.

Hence for \( E \in A \cap A_{V,k} \), \( P(E + i0, \lambda) \neq 0 \) for Lebesgue a.e. \( \lambda \in \mathbb{R} \). As the set of \( (E, \lambda) \) where this happens is clearly measurable, Fubini’s theorem implies that for Lebesgue a.e. \( \lambda \) one has \( P(E + i0, \lambda) \neq 0 \) for Lebesgue a.e. \( E \in A \cap A_{V,k} \). Since \( |1 + \lambda e^*_k \hat{G}^E_{\lambda} \Psi_k |^2 \) exists and is strictly positive for any \( \lambda \in \mathbb{R} \) and \( E \in A \cap A_{V,k} \), it follows from (8.40) that for a.e. \( \lambda \in \mathbb{R} \), Lebesgue a.e. \( E \in A \cap A_{V,k} \), \( \Im(\hat{G}^E_{\lambda} e_k) \) exists and is strictly positive. Therefore for a.e. \( \lambda \in \mathbb{R} \), the absolutely continuous part of \( \mu_{\omega_1,1} \) has almost surely a positive density on \( A \cap A_{V,k} \). By Proposition 8.13(i) the set \( A_{V,k} \) coincides with \( A_{V,k} \) and, as \( A \) has full Lebesgue measure, one obtains that \( A \cap A_{V,k} \) is an essential support of \( \mu_{\omega_1,1,m} \). Therefore for a.e. \( \lambda \in \mathbb{R} \), \( \mu_{\omega_1,1,m} \) is absolutely continuous w.r.t. \( \mu_{\omega_1,1,m} \).

This means that also the absolutely continuous part of \( \mu_\lambda \) must vanish for a.e. \( \lambda \in \mathbb{R} \). □

Corollary 8.3 For fixed \( \bar{\omega} \) and Lebesgue a.e. \( \hat{V} \in \mathcal{J} so^*(2L) \), the matrix valued measure \( \mu_{\omega} \) is absolutely continuous w.r.t. \( \mu_{\omega_1} \). Hence for \( P \) almost all \( \omega = (\bar{\omega}, \mathcal{V}) \) the measure \( \mu_{\omega} \) is absolutely continuous w.r.t. \( \mu_{\omega_1} \).

Proof. Let \( \omega \) be fixed. By Proposition 8.12, the assumptions of Theorem 8.8 are fulfilled for a.e. \( \hat{V} \in \mathcal{J} so^*(2L) \). Therefore for a.e. \( \hat{V} \in \mathbb{R} \Psi_k \Psi_k^* \), the orthogonal complement of \( \mathbb{R} \Psi_k \Psi_k^* \) in \( \mathcal{J} so^*(2L) \), there is some \( \lambda \) such that \( \hat{V}_1 = \hat{V} + \lambda \Psi_k \Psi_k^* \) fulfills the assumptions of Theorem 8.8. Therefore the measure \( \mu_{\omega_1,1,m} \) is absolutely continuous w.r.t. \( \mu_{\omega_1,1} \). For fixed \( \bar{\omega} \), the map \( \hat{V} \mapsto (\mu_{\omega_1,1,k}, \mu_{\omega_1,1,k}) \) is Borelian as well as the set of pairs of measures \( \{ (\mu, \nu) : \nu \text{ is a.c. w.r.t.} \mu \} \) (cf. Appendix A.8). Hence the set of \( \hat{V} \) where \( \mu_{\omega_1,1,k} \) is absolutely continuous is measurable. Therefore Fubini’s theorem now implies that this set has full Lebesgue measure on \( \mathcal{J} so^*(2L) \). This holds for any \( k = 2, \ldots, L \). As a finite intersection of sets of full measure is still a set of full measure we obtain for a.e. \( \hat{V} \in \mathcal{J} so^*(2L) \) the measure \( \sum_{k=1}^L \mu_{\omega_1,1,k} \) is a.c. w.r.t. \( \mu_{\omega_1,1} \), namely \( \mu_{\omega_1} \) is a.e. w.r.t. \( \mu_{\omega_1,1} \).

The maps \( \omega \mapsto \mu_\omega \) and \( \omega \mapsto \mu_{\omega_1} \) are Borelian. By the same arguments as above the set of \( \omega = (\bar{\omega}, \mathcal{V}) \) where \( \mu_{\omega_1} \) is absolutely continuous is measurable. As the distribution \( p_\mathcal{V} \) of \( \hat{V} \) is absolutely continuous, we obtain that for any fixed \( \bar{\omega} \), for \( p_\mathcal{V} \) almost every \( \mathcal{V} \), \( \mu_{\omega_1} \) is a.e. w.r.t. \( \mu_{\omega_1,1} \). By Fubini’s theorem, we obtain that this is true for \( P \) almost all \( \omega \). □

Theorem 8.9 For \( P \) almost every \( \omega \) one has \( \mu_{\omega,1,\text{sing}}(\mathbb{R}) = 0 \). Together with Corollary 8.3 this implies that for \( P \) almost all \( \omega \), one has \( \mu_{\omega,\text{sing}}(\mathbb{R}) = 0 \).

Proof. Let us define \( A_\omega = \{ E | \hat{G}^E \text{ exists and } \operatorname{Tr}(3m(\hat{G}^E + 0)) > 0 \} \) as well as \( A_{\omega,k} = \{ E | \hat{G}^E + 0 \text{ exists and } 3m(e_k \hat{G}^E + 0 e_k) > 0 \} \). By Lemma 8.6(iii), one has \( A_\omega = \bigcup_{k=1}^L A_{\omega,k} \).

Clearly \( A_\omega \) is an essential support of the a.c. part of \( \mu_\omega \) and \( A_{\omega,k} \) is an essential support of the a.c. part of \( \mu_{\omega,k} \).

By Kotani theory and Corollary 8.3 for \( P \) almost all \( \omega \) the set \( A_{\omega,k} \) has full Lebesgue measure and \( \mu_{\omega} \) is a.c. w.r.t. \( \mu_{\omega_1} \). Take such an \( \omega = (\bar{\omega}, \mathcal{V}) \). Then as \( \mu_\omega \) is a.e. w.r.t.
the sets $A_\omega$ and $A_{\omega, 1}$ differ only by a set of measure zero and hence $\mathbb{R} \setminus A_{\omega, 1}$ is a set of zero Lebesgue measure. Let $\hat{\mathcal{V}}$ be the projection of $\mathcal{V}$ orthogonal to $\Psi_1 \Psi_1^*$ and $\mathbf{p}_\omega$ be the distribution of $\hat{\mathcal{V}}$, namely the push forward of $\mathbf{p}_\omega$. Now set $\hat{\mathcal{V}}_\lambda = \hat{\mathcal{V}} + \lambda \Psi_1 \Psi_1^*$ and let $\mathcal{V}_\lambda$ be a pre-image of $\hat{\mathcal{V}}_\lambda$ under the Cayley transformation. Then by Proposition 8.13 one has for Lebesgue a.e. $\lambda \in \mathbb{R}$, $\mu_{\omega, 1}(\mathbb{R} \setminus A_{\omega, 1}) = \mu_{\omega, 1}(\mathbb{R} \setminus A_{\omega, 1}) = 0$, where $\omega_\lambda = (\hat{\omega}, \mathcal{V}_\lambda)$. As $\mu_{\omega, 1, \text{sing}}(A_{\omega, 1}) = 0$ by the definition of $A_{\omega, 1}$, this implies $\mu_{\omega, 1, \text{sing}}(\mathbb{R}) = 0$. Now by Fubini’s theorem for $\mathbf{P}$ a.e. $\hat{\omega}$ the situation described above happens for $\mathbf{p}_\omega$ a.e. $\mathcal{V}$. Then for $\mathbf{p}_\omega$ a.e. $\hat{\mathcal{V}}$ we have $\mu_{\hat{\omega}, \mathcal{V}, 1, \text{sing}}(\mathbb{R}) = 0$ for Lebesgue a.e. $\lambda$. Note that $\mathbf{p}_\omega$ is absolutely continuous and for fixed $\hat{\omega}$ the set of $\mathcal{V}$ where $\mu_{\hat{\omega}, \mathcal{V}, 1, \text{sing}}(\mathbb{R}) = 0$ is measurable, because of Appendix A.8 and the fact that the map $\mathcal{V} \mapsto \mu_{\hat{\omega}, \mathcal{V}, 1}$ is Borelian. Fubini’s theorem thus implies that for Lebesgue almost every $\hat{\mathcal{V}}$ in the strip $\text{supp}(\mathbf{p}_\omega) + \mathbb{R} \Psi_1 \Psi_1^*$ one has $\mu_{\hat{\omega}, \mathcal{V}, 1, \text{sing}}(\mathbb{R}) = 0$. As the distribution of $\hat{\mathcal{V}}$ is supported in this strip, this also holds for $\mathbf{p}_\omega$ a.e. $\mathcal{V}$.

As mentioned, this situation happens to be true for $\mathbf{P}$ a.e. $\hat{\omega}$. By the same arguments as above the set of $\omega$ where $\mu_{\omega, 1, \text{sing}}(\mathbb{R}) = 0$ is measurable. Fubini’s theorem now implies that $\mu_{\omega, 1, \text{sing}}(\mathbb{R}) = 0$ for $\mathbf{P}$ a.e. $\omega$. Since for $\mathbf{P}$ a.e. $\omega$ one also has that $\mu_\omega$ is a.c. w.r.t. $\mu_{\omega, 1}$, we finally obtain that $\mu_{\omega, \text{sing}}(\mathbb{R}) = 0$ for $\mathbf{P}$ a.e. $\omega$. $\square$
Chapter A

Appendix

A.1 Inhomogeneous singular first order ODE

This appendix recollects some results about the solutions of the inhomogeneous first order ordinary differential equation (ODE) for a function $y = y(x)$ on the interval $U = (a, b)$

$$py' + qy = r,$$  \hspace{1cm} (A.1)

where $p, q, r \in C^m(U)$ and $m \in \mathbb{N} \cup \{\infty\}$. In particular, we are interested in smooth solutions of the singular case where $p$, $q$ and $r$ have zeros at $\hat{x} \in U$ of order $0 \leq l_p, l_q < m$ and $l_r$ respectively. We suppose that $p$ and $q$ have no further zeros in $U$. If $l_r < \min\{l_p, l_q\}$, then there cannot be any solution $y \in C^1(U)$ because then

$$\frac{p}{r} y' + \frac{q}{r} y = 1$$

leads to the contradiction $0 = 1$ in the limit $x \to \hat{x}$. Hence let us suppose that $l_r \geq \min\{l_p, l_q\}$. In the regular case $l_p \leq l_q$, the homogeneous equation $py' + qy = 0$ has a one parameter family of solutions $e^{-w}$ where $w(x) = \int_\hat{x}^x \frac{q}{p} \in C^{m-l_p+1}(U)$ is any antiderivative. Then a one-parameter family of solutions of (A.1) is obtained by the method of variation of constants:

$$y(x) = e^{-w(x)} \int_\hat{x}^x ds \ e^{w(s)} \frac{r(s)}{p(s)},$$  \hspace{1cm} (A.2)

where the integral is well-defined because $l_r \geq l_p$. In the singular case, one has $l_q < l_p$ and $l_q \leq l_r$ and the zero of $p$ effectively decouples the intervals $(a, \hat{x})$ and $(\hat{x}, b)$. Any antiderivative $w(x) = \int_\hat{x}^x \frac{q}{p}$ diverges as $x \to \hat{x}$. Nevertheless, we shall show below that these divergences may cancel out in (A.2). Let us note right away though that any solution $y \in C^1(U)$ has to satisfy

$$y(\hat{x}) = \lim_{x \to \hat{x}} \frac{r(x)}{q(x)}.$$  \hspace{1cm} (A.3)

This is a boundary condition for both intervals $(a, \hat{x})$ and $(\hat{x}, b)$. The question is whether there are differentiable solutions (at the point $\hat{x}$, in particular) and whether such a solution is unique, or there are one-parameter or two-parameter families of such solutions.
Proposition A.1 Let \( l_r \geq \min\{l_p, l_q\} \) and set \( l = l_p - l_q \). Then the ODE (A.1) has continuous solutions \( y \) with a number of free parameters as described in the following:

(i) If \( l \leq 0 \), there is a one-parameter family of solutions \( y \in C^{m-l_p+1}(U) \).

(ii) If \( l = 1 \) and \( \partial_q^l(\hat{x}) > 0 \), the solution \( y \in C^{m-l_q}(U) \) is unique.

(iii) If \( l = 1 \) and \( \partial_q^l(\hat{x}) < 0 \), then there is a two-parameter family of solutions \( y \in C^n(U) \) as

\[
\text{long as } n \leq m - l_q \text{ and } n < |\partial_q^l(\hat{x})|^{-1}.
\]

(iv) If \( l \geq 2 \) is even and \( \partial_q^l(\hat{x}) > 0 \), the solution \( y \in C^{m-l_q}(U) \) is unique for \( x > \hat{x} \), but there is one free parameter for \( x < \hat{x} \).

(v) If \( l \geq 2 \) is even and \( \partial_q^l(\hat{x}) < 0 \), the solution \( y \in C^{m-l_q}(U) \) is unique for \( x < \hat{x} \), but there is one free parameter for \( x > \hat{x} \).

(vi) If \( l \geq 3 \) is odd and \( \partial_q^l(\hat{x}) > 0 \), the solution \( y \in C^{m-l_q}(U) \) is unique.

(vii) If \( l \geq 3 \) is odd and \( \partial_q^l(\hat{x}) < 0 \), there is a two-parameter family of solutions \( y \in C^{m-l_q}(U) \).

Any other solution \( y \) has a non-integrable singularity at \( \hat{x} \).

Proof. In case (i) the solutions are given in (A.2). Next let us consider the case (iv). Dividing (A.1) by \( q \) gives \( \frac{d}{dx}y' + y = \frac{r}{q} \). Hence it is sufficient to consider the case \( py' + y = r \) with \( p, r \in C^{m-l_q}(U) \) and \( l_p = l \). Let \( w \) be any antiderivative of \( \frac{1}{p} \) on \( U \setminus \{\hat{x}\} \). Because of \( \partial_q^l(\hat{x}) > 0 \) and \( l \) is even, one has \( \lim_{x \searrow \hat{x}} w(x) = +\infty \) and \( \lim_{x \nearrow \hat{x}} w(x) = -\infty \). Moreover, the growth of \( |w(x)| \) is at least as \( |x - \hat{x}|^{-1} \) because \( l \geq 2 \). Thus for all \( n \in \mathbb{N} \)

\[
\lim_{x \searrow \hat{x}} p^n(x) e^{w(x)} = +\infty, \quad \lim_{x \nearrow \hat{x}} \frac{e^{w(x)}}{p(x)} = 0. \tag{A.4}
\]

Now \( e^{-w} \) is a solution of the homogeneous equation \( py' + y = 0 \). Because \( \lim_{x \searrow \hat{x}} w(x) = -\infty \), the only possible continuous solution of \( py' + y = r \) for \( x > \hat{x} \) is

\[
y(x) = e^{-w(x)} W(x), \quad W(x) = \int_{\hat{x}}^x \frac{r(s)}{p(s)} \, ds, \tag{A.5}
\]

where the last integral exists due to (A.4). For \( x < \hat{x} \), let \( W \) be any antiderivative of \( e^w \frac{r}{p} \) and also set \( y = e^{-w} W \). Hence \( y \) is now a solution of \( py' + y = r \) for \( x \neq \hat{x} \). It remains to be shown that \( y \in C^{m-l_q}(U) \). First of all, \( y \) is continuous because

\[
\lim_{x \searrow \hat{x}} y(x) = \lim_{x \searrow \hat{x}} \frac{W(x)}{e^{w(x)}} = \lim_{x \searrow \hat{x}} \frac{e^{w(x)} r(x)/p(x)}{p(x)} = r(\hat{x}) = y(\hat{x}),
\]

the latter by (A.3). In the second equality, de l’Hospital’s rule could be applied because for \( x \searrow \hat{x} \) the numerator and denominator both converge to 0 by (A.4) and (A.5), while for \( x \nearrow \hat{x} \) as the denominator converges to \( \infty \) by (A.4).
Next follows an inductive argument in order to check the continuity of the higher derivatives of \( y \) at \( \hat{x} \). Let us set:

\[
q_n = np' + 1, \quad r_n = r'_{n-1} - q'_{n-1} y^{(n-1)} , \quad r_0 = r.
\]

Note that \( q_n \) has no zero at \( \hat{x} \). One can check by induction

\[
p y^{(n+1)} + q_n y^{(n)} = r_n . \tag{A.6}
\]

Let \( w_n = w + n \log(|p|) \). Now \( e^{-w_n} \) is a solution of the homogeneous equation \( p y^{(n+1)} + q_n y^{(n)} = 0 \) and satisfies due to (A.4)

\[
\lim_{x \rightarrow \hat{x}} p(x) e^{w_n(x)} = +\infty, \quad \lim_{x \rightarrow \hat{x}} e^{w_n(x)} = 0 . \tag{A.7}
\]

Now let \( y^{(n)} \), \( n < m - l_g \), be continuous by induction hypothesis. Then \( r_n \in C^1(U) \). Set \( W_n = y^{(n)} e^{w_n} \). Due to (A.7), one has \( \lim_{x \rightarrow \hat{x}} W_n(x) = 0 \). Using the identities

\[
y^{(n+1)} = \frac{r_n e^{w_n} - q_n W_n}{e^{w_n} p} , \quad q_n W_n' = r_n w'_n e^{w_n}
\]

which follow from (A.6) and \( w'_n = \frac{q_n}{p} \), one obtains

\[
\lim_{x \rightarrow \hat{x}} y^{(n+1)}(x) = \lim_{x \rightarrow \hat{x}} \frac{-q_n(x) W_n(x) + r_n(x) e^{w_n(x)} - q_n'(\hat{x}) y^{(n)}(\hat{x}) + r'_n(\hat{x})}{q_{n+1}(\hat{x})} ,
\]

where de l’Hospital’s rule can be used for \( x \downarrow \hat{x} \), because the numerator and denominator converge both to 0 according to (A.7) and for \( x \uparrow \hat{x} \) as the denominator converges to \( \pm \infty \). Therefore \( y^{(n)} \) is continuously differentiable in \( \hat{x} \).

The case (v) is identical, once the sides \( x > \hat{x} \) and \( x < \hat{x} \) are exchanged in the above argument. In case (vi), one has \( \lim_{x \rightarrow \hat{x}} \hat{x} = -\infty \) (namely, for both \( x \downarrow \hat{x} \) and \( x \uparrow \hat{x} \)), so that one has to construct the solution as in (A.5) both for \( x > \hat{x} \) and \( x < \hat{x} \); hence the solution is unique. In case (vii), one has \( \lim_{x \rightarrow \hat{x}} w(x) = +\infty \) and the homogeneous solutions on both sides vanish at \( \hat{x} \) faster than any power, so that one has a two-parameter family of solutions.

Finally let us consider \( l = 1 \) and first the case (ii). Hence \( \lim_{x \rightarrow \hat{x}} w(x) = -\infty \) so that \( \lim_{x \rightarrow \hat{x}} e^{w(x)} = 0 \). Now the second limit in (A.4) may diverge to \( +\infty \), but one readily checks that \( \frac{p}{\nu} \) is integrable so that \( W \) can be defined as in (A.5) for \( x > \hat{x} \) and \( x < \hat{x} \). Now the continuity is proven as above, and higher derivatives can also be treated as above because

\[
\lim_{x \rightarrow \hat{x}} e^{w_n(x)} = 0 ,
\]

so that both denominator and numerator of \( y^{(n+1)} \) converge to 0 and de l’Hopital’s rule may be applied as before.
In case (iii), \( \lim_{x \to \hat{x}} w(x) = \infty \) so that \( \lim_{x \to \hat{x}} e^{-w(x)} = 0 \). Furthermore \( \frac{e^w}{p} \) is never integrable and the two-parameter family of solutions \( y \) is defined as in (A.2). The main point to note is that \( w_n = (p'(\hat{x})^{-1} + n) \log(|x - \hat{x}|) + \mathcal{O}(1) \), hence

\[
\lim_{x \to \hat{x}} p(x) e^{w_n(x)} = \infty , \quad \lim_{x \to \hat{x}} p(x) e^{w_{n-1}(x)} = -\infty ,
\]

under the hypothesis stated in (iii). Thus de l'Hopital's rule can be invoked to calculate \( \lim_{x \to \hat{x}} y^{(n)} \) as above (with the index \( n \) shifted by 1). Let us remark that in case \(|p'(\hat{x})|^{-1} \in \mathbb{N}_1 \), one can prove better differentiability, but only for a one-parameter family of solutions (just as for the ODE \( xy = N \)), one can prove better differentiability, but only for a one-parameter family of solutions

\( \xi \). The singularity of all other solutions is at least of the type \( |x - \hat{x}|^{-1} \) in cases (ii) to (vii).

\[ \square \]

### A.2 Vector valued random variables

**Lemma A.1** Let \( a = (a_1, \ldots, a_n)^t : \Sigma \to \mathbb{R}^n \) be a centered, vector-valued random variable on a probability space \((\Sigma, \mathcal{P})\), and each \( a_k \in L^2(\Sigma, \mathcal{P}) \). Then there exist a linear decomposition \( a = \sum_i v_i b_i \) over finitely many fixed vectors \( b_i \in \mathbb{R}^n \) with coefficient \( v_i \) which are centered random variables \( v_i \in L^2(\Sigma, \mathcal{P}) \) that are uncorrelated \( \mathbb{E}(v_i v_{i'}) = \mathbb{E}(v_i^2) \delta_{i,i'} \).

**Proof.** One can assume that the random variables \( a_k \) as elements on \( L^2(\Sigma, \mathcal{P}) \) are linearly independent (otherwise one takes a basis for the vector space span \((\text{supp}(a))\) and rewrites the random variable \( a \) as vector using this basis). Let us introduce \( \lambda_{k,j} \) for \( k > j \) and write the ansatz \( v_k = a_k + \sum_{i=1}^{k-1} \lambda_{k,i} a_i \). Inverting the matrix form of these equations gives

\[
\begin{pmatrix}
(a_1) \\
(\vdots) \\
(a_n)
\end{pmatrix} =
\begin{pmatrix}
1 & 0 & \ldots & 0 \\
\lambda_{2,1} & 1 & \ddots & \vdots \\
\vdots & \ddots & \ddots & 0 \\
\lambda_{n,1} & \ldots & \lambda_{n,n-1} & 1
\end{pmatrix}^{-1}
\begin{pmatrix}
v_1 \\
(\vdots) \\
v_n
\end{pmatrix}.
\]

Hence one can write \( a \) as a sum \( \sum_k v_k b_k \) where the \( b_k \)'s are the vectors of the inverted matrix. The \( v_k \)'s are pairwise uncorrelated, if \( \mathbb{E}(v_k a_i) = 0 \) for all \( i < k \), as this implies \( \mathbb{E}(v_k v_i) = 0 \) for all \( i < k \). Now \( \mathbb{E}(v_k a_i) = 0 \) for \( i = 1, \ldots, k-1 \) is guaranteed if

\[
- \begin{pmatrix}
\mathbb{E}(a_k a_1) \\
\vdots \\
\mathbb{E}(a_k a_{k-1})
\end{pmatrix} =
\begin{pmatrix}
\mathbb{E}(a_1 a_1) & \ldots & \mathbb{E}(a_1 a_{k-1}) \\
\vdots & \ddots & \vdots \\
\mathbb{E}(a_{k-1} a_1) & \ldots & \mathbb{E}(a_{k-1} a_{k-1})
\end{pmatrix}^{-1}
\begin{pmatrix}
\lambda_{k,1} \\
\vdots \\
\lambda_{k,k-1}
\end{pmatrix}.
\]

If the appearing matrix is invertible, one can resolve this equation to get \( \lambda_{k,i} \) for all \( i < k \). So it remains to show that this matrix is invertible which is equivalent to the property that the columns are linearly independent. Now let \( \xi_i \in \mathbb{R} \) such that

\[
\sum_{i=1}^{k-1} \xi_i \mathbb{E}(a_j a_i) = \mathbb{E} \left( a_j \sum_{i=1}^{k-1} \xi_i a_i \right) = 0
\]
for all \( j = 1, \ldots, k \). The vector \( \sum_{1 \leq i \leq k-1} \xi_i a_i \) is then orthogonal in \( L^2(\Sigma, \mu) \) to any vector in the subspace spanned by \( a_1, \ldots, a_{k-1} \) and it therefore has to be zero. As the random variables \( a_i \) are linearly independent, one gets \( \xi_i = 0 \) for all \( i = 1, \ldots, k - 1 \).

### A.3 Hörmander operators

Let \( X_0, X_1, \ldots, X_k \) be smooth vector fields in a domain, i.e., open and connected set, \( D \subset \mathbb{R}^n \). We want to consider the differential operator \( L = \sum_{j=1}^k X_j^2 + X_0 \). As the vector fields are smooth, the operator \( L \) can be extended to act on the distributions on the domain \( D \). For simplicity let us denote this extension also by \( L \). Hörmander was the first one to find the hypoellipticity property (defined below) for such operators under certain conditions that are more general than ellipticity [40].

**Definition A.1** The differential operator \( L \) is called hypoelliptic if for any distribution \( f \) on \( D \) such that \( L(f) \in C^\infty(D) \) one has \( f \in C^\infty(D) \).

This property holds if the operator \( L \) is elliptic which means that the principal symbol of \( L \) is positive. In the written Hörmander form this is equivalent to the vector fields \( X_0, \ldots, X_k \) spanning the whole tangent space of \( D \subset \mathbb{R}^n \) at each point. In particular it implies \( k \geq n \). Hörmander obtained a weaker condition.

**Definition A.2** The operator \( L \) is said to satisfy the (weak) Hörmander condition of rank \( r \) if the vector fields \( X_0, \ldots, X_k \) and its \((r-1)\)th commutators span the whole tangent space at each point of \( D \). \( L \) satisfies the strong Hörmander condition of rank \( r \) if the vector fields \( X_1, \ldots, X_k \) and its \((r-1)\)th fold commutators span the whole tangent space at each point.

Note that an elliptic operator \( L \) is a strong Hörmander operator of rank 1.

**Theorem A.1** (Hörmander) If \( L \) satisfies the weak Hörmander condition, then \( L \) is hypoelliptic.

An essential estimate for second order elliptic operators which can be found in any book on partial differential equations is \( \| f \|_{(2)} \leq C \left( \| Lf \|_{(0)} + \| f \|_{(0)} \right) \). Here \( \| \cdot \|_{(s)} \) denotes the norm in the Sobolev space \( H^s(D) \). Note that the norm \( \| \cdot \|_{(0)} \) is equal to the \( L^2(D) \) norm. The following generalization of this estimate can be found in [42].

**Theorem A.2** Let \( L \) satisfy the Hörmander condition of rank \( r \) then there is a constant \( C \) such that if \( f, Lf \in L^2(D) \) then one has

\[
\| f \|_{(2)} \leq C \left( \| Lf \|_{(0)} + \| f \|_{(0)} \right)
\]

Using the fact that \( H^{\frac{1}{2}}(D) \) is compact embedded in \( L^2(D) \) one obtains the following

**Corollary A.1** Any resolvent of an Hörmander operator in \( L^2(D) \) is compact.
Proof. Let \( R = (\mathcal{L} + c)^{-1} \) be bounded in \( L^2(D) \). Then \( \mathcal{L} R = 1 - cR \) is also a bounded operator on \( L^2(D) \) and one obtains
\[
\| R f \| (\mathcal{L}^2) \leq C (\| \mathcal{L} R f \| + \| R f \| ) \leq C (\| \mathcal{L} R \| + \| R \| ) \| f \| .
\]
The compact embedding of \( H^{\frac{1}{2}} \) in \( L^2(D) \) shows that \( R \) is a compact operator. \( \square \)

Another important result for second order elliptic operators is the maximum principle.

It states that solutions to \( \mathcal{L} f \geq 0 \) have their extreme values always at the boundary of \( D \).
Bony [8] showed that this also holds for strong Hörmander operators.

Theorem A.3 (Bony) Let \( \mathcal{L} \) satisfy the strong Hörmander condition and let \( \mathcal{L} f \geq 0 \) on the domain \( D \). If \( f \) has a local maximum inside \( D \) then \( f \) is constant on \( D \).

Corollary A.2 Let \( \mathcal{L} \) be a strong Hörmander operator on a connected, compact manifold \( \mathcal{M} \). If \( \mathcal{L} f = 0 \) then \( f \) is constant on \( \mathcal{M} \).

Proof. First by hypoellipticity in each chart, \( f \) is smooth. As \( \mathcal{M} \) is compact, there is a point \( x \in \mathcal{M} \) where \( f \) takes its maximum. Let \( y \in \mathcal{M} \) be arbitrary, then there is a path from \( x \) to \( y \). Cover this path with a sequence of connected charts and corresponding open sets \( O_1, \ldots, O_k \) such that \( x \in O_1, y \in O_k \) and \( O_j \cap O_{j+1} \neq \emptyset \). By Bony’s maximum principle \( f \) is constant on \( O_1 \). As \( O_2 \) intersects \( O_1 \), \( f \) takes its maximum also in \( O_2 \) and is constant on \( O_2 \). By induction \( f \) is constant on \( O_1 \cup O_2 \cup \ldots \cup O_k \) and hence \( f(x) = f(y) \). As \( y \) was arbitrary we obtain that \( f \) is constant. \( \square \)

A.4 Sobolev spaces on compact manifolds

Let \( (\mathcal{M}, g) \) be a compact Riemannian manifold without boundary. We want to define the Sobolev spaces using charts. For any point \( x \in \mathcal{M} \) choose a chart with domain \( V_x \subset \mathbb{R}^n \) for some open, connected neighborhood \( V_x \subset \mathcal{M} \) of \( x \) and let \( \Psi_x : V_x \rightarrow \tilde{V}_x \) denote the corresponding diffeomorphism. Furthermore choose some simple connected neighborhood \( \tilde{U}_x \subset \tilde{V}_x \) of \( x \) which is pre-compact in \( \tilde{V}_x \) and let \( U_x = \Psi_x(V_x) \) be the chart domain. As \( \mathcal{M} \) is compact, we can cover it by finitely many charts \( \Psi_i(U_i) = \Psi_{x_i}(U_{x_i}) = U_i \). As the metric tensor \( g_{ab} \) is invertible, \( \sqrt{|g|} = \sqrt{\det g_{ab}} > 0 \), and as \( U_i \) is pre-compact in \( V_x \), we get
\[
0 < C_1 = \min_i \inf_{y \in U_i} \sqrt{|g|} \leq \max_i \sup_{x \in U_i} \sqrt{|g|} = C_2 < \infty . \tag{A.8}
\]
For a smooth function \( f \) on \( \mathcal{M} \) define \( f_{(i)} = f \circ \Psi_i|_{U_i} \), i.e. \( f_{(i)} \) is the part of \( f \) on \( \tilde{U}_i \) represented in the chart \( (U_i, \Psi_i) \). Now let \( \| f_{(i)} \|_{(s)} \) denote the Sobolev norm of \( f_{(i)} \) corresponding to \( H^s(U_i) \) and define \( \| f \|_{(s)}^2 = \sum_i \| f_{(i)} \|_{(s)}^2 \). The set of all functions, where the norm \( \| f \|_{(s)} \) is finite, shall be denoted by \( H^s(\mathcal{M}) \), i.e. \( H^s(\mathcal{M}) = \{ f : f_{(i)} \in H^s(U_i) \} \). Note that although the norm depends on the choice of the atlas \( (U_i) \), the set \( H^s(\mathcal{M}) \) does not. Because as the atlas is finite, a function \( f \) belongs to \( H^s(\mathcal{M}) \) precisely if it is locally in the corresponding Sobolev space which is independent of the used charts. One can also show that the norm for a different finite atlas obtained in the same way is equivalent.
Proposition A.2 For $s > 0$ the space $H^s(\mathcal{M})$ is compactly embedded in $L^2(\mathcal{M})$.

Proof. Clearly by (A.8) one has

$$\|f\|_2^2 \leq \sum_i \int_{U_i} |f \circ \Psi_i|^2 \sqrt{|g|} \, d\lambda \leq C_2 \sum_i \|f_i(0)\|_2^2 \leq C_2 \sum_i \|f_i(0)\|_{L^2(\mathcal{M})}^2 = C_2 \|f\|_{L^2(\mathcal{M})}^2.$$  \hspace{1cm} (A.9)

Now let $f_m$ be some bounded sequence in $H^s(\mathcal{M})$. As $H^s(\mathcal{U}_i)$ is compactly embedded in $L^2(\mathcal{U}_i)$ one can construct inductively some subsequence $f_k$, such that $f_{k(i)} = f_k \circ \Psi_i$ converges to a function $h_i$ in $L^2(\mathcal{U}_i)$ for all $i$.

Now let $U_{ij} = U_i \cap \Psi_i^{-1}(U_j)$ and let $\Psi_{ij}$ denote the diffeomorphism $\Psi_i^{-1} \circ \Psi_j$ from $U_{ji}$ to $U_{ij}$. As $U_i, U_j$ are pre-compact in $V_i, V_j$ the determinant of this diffeomorphism is bounded away from zero and infinity. Then one has

$$\|h_{i} - h_{j} \circ \Psi_{ji}\|_{L^2(U_{ij})} \leq \lim_{k \to \infty} \left\{ \|h_{i} - f_{k(i)}\|_{L^2(U_{i})}^2 + \int_{U_{ij}} |f_{k(i)} - h_{j} \circ \Psi_{ji}|^2 d\lambda \right\}$$

$$\leq \lim_{k \to \infty} \int_{U_{ij}} |f_{k(i)} - h_{j}|^2 |\det \Psi_{ij}| \, d\lambda \leq \sup_{U_{ij}} |\det \Psi_{ij}| \lim_{k \to \infty} \|f_{k(i)} - f_j\|_{L^2(U_{ij})}^2 = 0,$$

hence we get $h_{i} = h_{j} \circ \Psi_{ji}$ almost surely on $U_{ij}$. Thus there is a function $h$, such that $h_{(i)} = h \circ \Psi_i = h_i$ almost surely. Furthermore one has

$$\|f_k - h\|_2^2 \leq C_2 \sum_i \|f_{k(i)} - h_{(i)}\|_{L^2(U_{i})}^2 = C_2 \sum_i \|f_{k(i)} - h_i\|_{L^2(U_{i})}^2 \to 0 \text{ as } k \to \infty,$$

hence $f_k$ converges to $h$ in $L^2(\mathcal{M})$. Thus any bounded sequence in $H^s(\mathcal{M})$ has an $L^2$-convergent subsequence. This finishes the proof. \hfill $\square$

Corollary A.3 $L^2$-resolvents of Hörmander operators are compact.

Proof. Let $\mathcal{L} = \sum_{j=0}^{k} X_j^2 + X_0$ be a smooth, second order differential operator satisfying the Hörmander condition and let $R = (\mathcal{L} + c)^{-1}$ be a bounded operator on $L^2(\mathcal{M})$. The representation of $\mathcal{L}$ in the charts $U_i$ are then also Hörmander type operators. Hence in each chart one has subelliptic estimates of the type

$$\|R f_{i(0)}\|_{L^2(U_i)} \leq C \left( \|LR f_{i(0)}\|_{L^2(U_i)} + \|R f_{i(0)}\|_{L^2(U_i)} \right) \leq \hat{C} \|f_{i(0)}\|_{L^2(U_i)}$$  \hspace{1cm} (A.10)

for adequate constants $C$ and $\hat{C}$. From (A.8) one deduces

$$\|f_{i(0)}\|_2^2 = \int_{U_i} |f \circ \Psi_i|^2 d\lambda \leq \frac{1}{C_1} \int_{U_i} |f \circ \Psi_i|^2 \sqrt{|g|} d\lambda \leq \frac{1}{C_1} \|f\|_2^2.$$  \hspace{1cm} (A.11)

As there are only finitely many sets $U_i$, summing over $i$ in equation (A.10) and using (A.11) one gets

$$\|R f\|_{L^2(\mathcal{M})} \leq \hat{C} (\|f\|_2 + \|Rf\|_2)$$

for an adequate constant $\hat{C}$. As $R$ is bounded on $L^2(\mathcal{M})$ and as $H^s(\mathcal{M})$ is compactly embedded in $L^2(\mathcal{M})$ one can read off immediately, that $R$ is a compact operator. \hfill $\square$
A.5 Fourier series on compact Lie groups

First let us summarize some facts about the representation theory of compact Lie groups. All this is well known and proofs can be found in the literature, e.g. [14], but we need to introduce the notations for the proof of Theorem A.6.

Let $\mathfrak{K}$ be a compact Lie group equipped with its normalized Haar measure and let $\mathfrak{T} \subset \mathfrak{K}$ be some maximal torus $\mathfrak{T} \cong \mathbb{T}^r$, where $r$ is called the rank of $\mathfrak{K}$. The irreducible representations of the torus $\mathfrak{T}$ are given by the characters, i.e. the homomorphisms into the group $S^1 = U(1) \subset \mathbb{C}$. Let us denote them by $X^*(\mathfrak{T})$. They form a $\mathbb{Z}$-module isomorphic to the lattice $\mathbb{Z}^r$ and hence $X^*(\mathfrak{T})$ is a lattice in the vector space $\mathcal{V} = \mathbb{R} \otimes_{\mathbb{Z}} X^*(\mathfrak{T})$, the tensor product over the ring $\mathbb{Z}$. This is an abstract description of the fact, that the characters of the torus $\mathfrak{T}^r$ are given by the maps $\theta \in \mathfrak{T}^r \mapsto e^{ij\theta}$ for a fixed $j \in \mathbb{Z}^r$. In this case, $\mathcal{V} = \mathbb{R}^r$.

Define some $\text{Ad}_{\mathfrak{K}}$-invariant scalar product on the Lie algebra $\mathfrak{k}$ of $\mathfrak{K}$, where $\text{Ad}_{\mathfrak{K}}$ denotes the adjoint representation, and adopt $\mathcal{V}$ with an scalar product $\langle \cdot, \cdot \rangle$ such that the norm of $a \in X^*(\mathfrak{T})$ coincides with the operator norm of the derivative $da$ acting on $\mathfrak{k}$, the Lie algebra of $\mathfrak{T}$.

Let $\mathfrak{p}$ be the orthogonal complement in $\mathfrak{k}$ of $\mathfrak{t}$, the Lie algebra of $\mathfrak{T}$. Then the group $\mathfrak{T}$ acts on the complexification $\mathfrak{p}_\mathbb{C} = \mathbb{C} \otimes_{\mathbb{R}} \mathfrak{p}$ by the adjoint representation and linearity. This representation of $\mathfrak{T}$ can be decomposed into irreducible representations, which means $\mathfrak{p}_\mathbb{C} = \bigoplus_{a \in \Phi} \mathfrak{p}_a$ where $\mathfrak{p}_a$ is the set of $P \in \mathfrak{p}_\mathbb{C}$ such that $\text{Ad}_T(P) = a(T)P$ for all $T \in \mathfrak{T}$. One can show that the spaces $\mathfrak{p}_a$ are one-dimensional complex vector spaces. The appearing characters $a \in \Phi \subset X^*(\mathfrak{T})$ are called roots of $\mathfrak{K}$. If $a \in \Phi$ is a root, then also $-a \in \Phi$. Note that the character $-a$ as a map on $\mathfrak{T}$ is given by $(a)(T) = (a(T))^{-1}$.

One can divide the vector space $\mathcal{V}$ in an upper half space and a lower half space in such a way that there is no root on the boundary. A root in the upper half space is then called a positive root. The set of vectors $v \in \mathcal{V}$ that satisfy $\langle v, a \rangle \geq 0$ for all positive roots $a$ is a so-called positive Weyl chamber $\mathcal{C}_+$. An element of the lattice $X^*(\mathfrak{T})$ lying in the positive Weyl chamber is called a highest weight. The set of highest weights will be denoted by $\mathcal{W}_+$. There is a one-to-one correspondence between the irreducible representations and the highest weight vectors.

**Theorem A.4** Any irreducible (unitary) representation of $\mathfrak{K}$ induces (by restriction) a representation of $\mathfrak{T}$, which when decomposed into irreducible representations of $\mathfrak{T}$ contains exactly one highest weight $a \in \mathcal{W}_+$. For any highest weight vector $a \in \mathcal{W}_+$ there is exactly one irreducible representation of $\mathfrak{K}$ containing $a$.

Let $\pi_a : \mathfrak{K} \rightarrow U(d(a))$ for $a \in \mathcal{W}_+$ be the corresponding irreducible unitary representation of dimension $d(a)$. By Schur orthogonality and the Peter-Weyl theorem the matrix coefficients $\pi_a(K)_{k,l}$, where $1 \leq k, l \leq d(a)$, of these representations, considered as functions on $\mathfrak{K}$, form an orthogonal basis for $L^2(\mathfrak{K})$. The $L^2$ norm of such a matrix coefficient is $d(a)^{-1/2}$. Therefore the orthogonal projection of $f$ onto the space spanned by the matrix
coefficients of the irreducible representation $\pi_a$ is given by

$$d(a) \sum_{k,l=1}^{d(a)} \int dK \left( f(K) \pi_a(K)_{k,l} \right) \pi_a(K)_{k,l} = d(a) \sum_{k,l=1}^{d(a)} \int dK \left( f(K) \pi_a(K^{-1})_{l,k} \right) \pi_a(K)_{k,l} \pi_a(K^{-1})_{l,k} = d(a) \text{Tr} \left( \mathcal{F} f(a) \pi_a(K) \right),$$

where

$$\mathcal{F} f(a) = \int dK f(K) \pi_a(K^{-1}).$$

Hence Schur orthogonality and the Peter-Weyl theorem imply the following.

**Corollary A.4** Let $f \in L^2(\mathfrak{K})$, then one obtains with convergence in $L^2(\mathcal{K})$

$$f(K) = \sum_{a \in \mathcal{W}_+} d(a) \text{Tr} \left( \mathcal{F} f(a) \pi_a(K) \right).$$

As shown in [85] one can characterize the smooth functions on $\mathfrak{K}$ by their Fourier series.

**Theorem A.5** A function $f$ on $\mathfrak{K}$ is smooth if and only if its Fourier coefficients are rapidly decreasing, which means that

$$\forall h > 0 : \lim_{\|a\| \to \infty} \|a\|^h \|\mathcal{F} f(a)\| = 0$$

Here $\|\mathcal{F} f(a)\|$ denotes the Hilbert-Schmidt norm. If this is fulfilled, then the Fourier series converges absolutely in the supremum norm on $\mathfrak{K}$.

Note that the definition of $\mathcal{F} f(a)$ to be rapidly decreasing is independent of the chosen norm on $\mathcal{W}_+ \subset \mathcal{V}$.

Now let us consider the compact group $\mathfrak{K} \times T^L$ with the maximal torus $\mathfrak{T} \times T^L$ and its Lie algebra $\mathfrak{t} \times \mathbb{R}^L$. The characters of this torus also factorize by $X^*(\mathfrak{T} \times T^L) = X^*(\mathfrak{T}) \times \mathbb{Z}^L$. As $\{1\} \times T^L$ lies in the center, the direct product of the scalar product on $\mathfrak{t}$ and the canonical scalar product on $\mathbb{R}^L$ give a scalar product on $\mathfrak{t} \times \mathbb{R}^L$ that is invariant under the adjoint representation of the group $\mathfrak{K} \times T^L$. Therefore the induced scalar product on the vector space $\mathcal{V} \times \mathbb{R}^L$ spanned by the characters also factorizes.

As the adjoint representation of $\{1\} \times T^L$ is trivial, the roots of $\mathfrak{K} \times T^L$ consist of elements $(a,0)$ where $a$ is a root of $\mathfrak{K}$. Therefore the positive roots of $\mathfrak{K} \times T^L$ are simply the positive roots of $\mathfrak{K}$ and, as the scalar product on $\mathcal{V} \times \mathbb{R}^L$ factorizes, the positive Weyl chamber for $\mathfrak{K} \times T^L$ is given by $C_+ \times \mathbb{R}^L$. Hence the highest weight vectors are given by $\mathcal{W}_+ \times \mathbb{Z}^L$.

Now for $a \in \mathcal{W}_+$ the mapping $(K,\theta) \mapsto \pi_a(K)e^{ij\theta}$ is an irreducible representation of $\mathfrak{K} \times T^L$ which contains the highest weight vector $(a,j)$ and by Theorem A.4, it is the unique one containing this weight. Thus we have shown the following.
Theorem A.6 The highest weight vectors of $\mathfrak{R} \times \mathbb{T}^L$ are given by $W_+ \times \mathbb{Z}^L$, where $W_+$ are the highest weight vectors of $\mathfrak{R}$. The irreducible representation parameterized by $(a, j) \in W_+ \times \mathbb{Z}^L$ is given by
\[ \pi(a,j)(K, \theta) = \pi_a(K) e^{ij \theta}. \]

Hence the Fourier series of $F$ is given by
\[ F(K, \theta) = \sum_{a \in W_+} \sum_{j \in \mathbb{Z}^L} d(a) \text{Tr}(\mathcal{F}F(a, j) \pi_a(K)) e^{ij \theta}. \]

with convergence in $L^2(\mathfrak{R} \times \mathbb{T}^L)$, where
\[ \mathcal{F}F(a, j) = \int_{\mathfrak{R}} dK \int_{\mathbb{T}^L} d\theta F(K, \theta) \pi_a(K^{-1}) e^{-ij \theta}. \]

A.6 Isotropic flag manifold as quotient of compact groups

As in the previous chapters define the $2L \times 2L$ complex matrices
\[ C = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i1 \\ 1 & i1 \end{pmatrix}, \quad G = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \]

Then the group $U(L, L) = \text{CSp}(2L, \mathbb{C})C^*$ is given by all complex matrices $T$ such that $T^*GT = G$. The subgroup $U(L, L, \mathbb{R}) = \text{CSp}(2L, \mathbb{R})C^*$ consists of all matrices $T \in U(L, L)$ which additionally satisfy $CTC^* \in \text{Mat}(2L, \mathbb{R})$.

Let $I$ be the set of complex $2L \times L$ matrices $\Phi$ of rank $L$ which satisfy $\Phi^*G\Phi = 0$. Then $I$ is the set of all matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, where $a, b$ are complex $L \times L$ matrices such that $a^*a = b^*b$ is invertible. Furthermore let $I_{\mathbb{R}}$ be the subset of all $\Phi \in I$ where $C\Phi$ is a real matrix. This means $I_{\mathbb{R}}$ consists of all matrices of the form $\begin{pmatrix} a \end{pmatrix} \in I$. The Lorentz group $U(L, L)$ acts naturally on $I$ and $U(L, L, \mathbb{R})$ on $I_{\mathbb{R}}$ by left multiplication. Two elements $\Phi_1, \Phi_2 \in I$ represent the same flag of isotropic subspaces, if $\Phi_2 = \Phi_1 S$ for an invertible upper triangular $L \times L$ matrix $S$. In that case, let us write $\Phi_1 \equiv \Phi_2$. Note that if both, $\Phi_1$ and $\Phi_2$ are elements of $I_{\mathbb{R}}$, then $S$ is a real matrix. It is easy to check that $\equiv$ is an equivalence relation. The flag manifolds are defined by $\mathbb{F} = \mathbb{I}/\equiv$ and $\mathbb{F}_{\mathbb{R}} = I_{\mathbb{R}}/\equiv$. Let us cover this manifolds by other quotients. Let $\Phi_1 \sim \Phi_2$ iff $\Phi_2 = \Phi_1 S$ for an invertible upper triangular $L \times L$ matrix $S$ which has positive reals on the diagonal. Then $\sim$ is also an equivalence relation and $\Phi_1 \sim \Phi_2 \Rightarrow \Phi_1 \equiv \Phi_2$. Hence there is a natural projection of $\mathbb{I}/\sim$ on $\mathbb{F}$ and of $I_{\mathbb{R}}/\sim$ on $\mathbb{F}_{\mathbb{R}}$. As the mentioned group actions are transitive, all these manifolds are homogeneous spaces and equipped with a natural differentiable structure.

Proposition A.3 The manifold $\mathbb{I}/\sim$ is diffeomorphic to $U(L) \times U(L)$ and $I_{\mathbb{R}}/\sim$ is diffeomorphic to $U(L)$. The action of $U(L, L)$ on $U(L) \times U(L)$ is given by
\[ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \cdot \begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} (AU + BV)S \\ (CU + DV)S \end{pmatrix}. \]
and the action of $U(L, L, \mathbb{R})$ on $U(L)$ is given by

$$
\begin{pmatrix} A & B \\ C & D \end{pmatrix} \cdot U = (AU + BV)S
$$

where $S$ is in both cases an upper triangular matrix with positive reals on the diagonal, such that $(AU + BV)S$ is unitary. (Note that in the first case then also $(CU + DV)S$ is unitary and in the second case $S$ is real).

The projections of $\mathbb{I}/ \sim$ on $\mathbb{F}$ and of $\mathbb{R}_1/ \sim$ on $\mathbb{F}_1$ induce diffeomorphisms of $\mathbb{F}$ to $(U(L) \times U(L))/\tilde{\mathcal{H}}$, and of $\mathbb{R}_1$ to $U(L)/\tilde{\mathcal{H}}_R$, where $\tilde{\mathcal{H}} = \{(S, S) : S = \text{diag}(u_1, \ldots, u_L) \text{ and } |u_i| = 1\}$ is isomorphic to the torus $\mathbb{T}^L$ and $\tilde{\mathcal{H}}_R = \{\text{diag}(u_1, \ldots, u_L) : u_i \in \{-1, +1\}\}$ is isomorphic to the finite group $(\mathbb{Z}_2)^L$.

**Proof.** Let $\Phi = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{I}$, where $a, b$ are $L \times L$ matrices. Then $a^*a$ is invertible and defines a scalar product on $\mathbb{C}^L$. By the Gram-Schmidt procedure there is a unique upper triangular matrix $S$ with positive reals on the diagonal such that the column vectors of $S$ form an orthonormal basis w.r.t. the scalar product defined by $a^*a = b^*b$. For this $S$ one has $S^*a^*aS = 1 = S^*b^*bS$. This means that for $(\nu) = \Phi S \sim \Phi$ one obtains $U, V \in U(L)$. As $S$ is uniquely determined and as $(\nu_{U, V}) \in \mathbb{I}$, each equivalence class contains exactly one pair of unitaries and the map $[\Phi]_\sim \mapsto (U, V)$ is bijective.

In case $\Phi_R = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{I}_R$ one has $b = \overline{a}$ and hence $a^*a = \overline{a}\overline{a}$ is a real matrix. Therefore the Gram-Schmidt procedure leads to a real matrix $S$ and $V = \overline{U}$. As $(\nu_{U, V}) \in \mathbb{I}_R$ for any unitary $U$, each element of $\Phi_R \in \mathbb{I}_R$ is equivalent to exactly one element of the form $(\nu_{\overline{U}, V})$ with $U \in U(L)$ and the map $[\Phi_R]_\sim \mapsto U$ is a bijection.

The induced action of $U(L, L)$ on $U(L) \times U(L)$, resp. of $U(L, L, \mathbb{R})$ on $U(L)$, can be seen to be differentiable, as the Gram Schmidt procedure obtaining $S$ as a function of $a$ is differentiable and hence the differentiable structure of $U(L) \times U(L)$, resp. $U(L)$, coincides with the induced one from the group action of $U(L, L)$, resp. $U(L, L, \mathbb{R})$. Therefore the maps $[\Phi]_\sim \mapsto (U, V)$ and $[\Phi_R]_\sim \mapsto U$ are even diffeomorphisms.

Clearly the action of $U(L, L)$ on $U(L) \times U(L)$ is given by

$$
\begin{pmatrix} A & B \\ C & D \end{pmatrix} \cdot \begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} (AU + BV)S \\ (CU + DV)S \end{pmatrix}
$$

where $S$ is in both cases an upper triangular matrix with positive reals on the diagonal, such that $(AU + BV)S$ is unitary. (Note that in the first case then also $(CU + DV)S$ is unitary). Similarly the action of $U(L, L, \mathbb{R})$ on $U(L)$ is also given as stated in the proposition.

For two pairs of unitaries $(U, V)$ and $(U', V')$ write $(U, V) \equiv (U', V')$ iff $(\nu_U) \equiv (\nu_{U'})$. This relation holds iff there is an invertible upper triangular matrix $S$ such that $US = U'$ and $VS = V'$. In that case $S$ has also to be unitary. Hence $S^{-1} = S^*$ is an upper- and lower triangular matrix and therefore diagonal. Thus $(U, V) \equiv (U', V')$ iff $(U, V)^{-1}(U', V') = (U^{-1}U', V^{-1}V') = (S, S) \in \tilde{\mathcal{H}}$ where $\tilde{\mathcal{H}}$ is the subgroup $\tilde{\mathcal{H}} = \{(S, S) : S = \text{diag}(u_1, \ldots, u_L) \text{ and } |u_i| = 1\}$. Therefore $\mathbb{F}$ is diffeomorphic to $(U(L) \times U(L))/\tilde{\mathcal{H}}$.

If one has $(U, \overline{U}) \equiv (U', \overline{U'})$ then this is equivalent to the fact that $S = U^{-1}U' = \overline{U'}S$ is a real, upper triangular and unitary matrix. Therefore as above $S$ is diagonal.
and all entries are equal to ±1. Hence $S \in \mathfrak{h}_\mathbb{R} = \{ \text{diag}(u_1, \ldots, u_L) : u_i \in \{-1,+1\}\}$ and $\mathbb{F}_\mathbb{R}$ is diffeomorphic to $U(L) / \mathfrak{h}_\mathbb{R}$.

\[ \Box \]

### A.7 Divergence of vector fields on quotients of compact groups

Let $\mathfrak{h} \subset \mathfrak{g}$ be compact subgroups of the unitary group $U(L)$. Let $\mathfrak{M} = \mathfrak{g} / \mathfrak{h}$ be the homogeneous quotient and denote the canonical projection by $\pi : \mathfrak{g} \to \mathfrak{M}$. On the Lie algebra $u(L)$ and hence on the Lie algebra $\mathfrak{k}$ of $\mathfrak{g}$, the Killing Form $(u, v) = \text{Tr}(u^* v)$ defines a bi-invariant metric. At each point $K \in \mathfrak{g}$, the Lie algebra $\mathfrak{h}$ of $\mathfrak{h}$ form the vertical vectors, i.e. the kernel of the differential of $\pi$. Hence the tangent space at $\pi(K)$ can be identified with the horizontal vectors, $\mathfrak{h}^\perp$, the orthogonal complement of $\mathfrak{h}$ in $\mathfrak{k}$. This identification depends on the choice of $K$. Two horizontal lifts of some tangent vector on $\mathfrak{M}$ to two different pre-images differ by a conjugation and therefore have the same length due to the invariance of the metric. Thus there is a unique metric on $\mathfrak{M}$ such that the projection $\pi : \mathfrak{g} \to \mathfrak{M}$ is a Riemannian submersion. This metric is invariant under the action of $\mathfrak{g}$.

Let $S_i$ be some orthonormal basis for $\mathfrak{h}^\perp$, then the push forward, $\pi_*(S_i)$ forms an orthonormal basis at $\pi(K)$. (These basis vectors may differ for two different pre-images.)

Let $X$ be some smooth vector field on $\mathfrak{M}$ and denote the horizontal lift to $\mathfrak{g}$ by $\hat{X}$ which then is also smooth. As $\pi$ is a Riemannian submersion, the covariant derivative of $X$ with respect to $\pi_*(S_i)$ is given by $\pi_*(\nabla_{S_i} \hat{X})$. Let $(B_j)$ denote some orthonormal basis of $\mathfrak{k}$ and identify $B_j$ with the left invariant vector field. Furthermore we identify any vector field $Y$ on $\mathfrak{g}$ with a function $Y : \mathfrak{g} \to \mathfrak{k}$ such that the vector at $K$ is given by the path $K \exp(tY(K))$. With $\nabla_S \hat{X}$ we denote the covariant derivative of the vector field $\hat{X}$ and with $\delta_S \hat{X}$ the derivative of the associated function as described above, i.e. if the vector field $\hat{X}$ at $K \in \mathfrak{g}$ is given by the path $K \exp(t \hat{X}(K))$ then $\delta_S \hat{X}(K) = \frac{\partial}{\partial t} \hat{X}(K \exp(tS))$ at $t = 0$. Then using the well known fact that the covariant derivative of a left-invariant vector field $B$ w.r.t. to a left-invariant vector field $S$ is just the left-invariant vectorfield $\frac{1}{2}[S, B]$ one obtains

$$
\nabla_S \hat{X} = \sum_j \nabla_S \text{Tr}(B_j^* \hat{X}) B_j = \sum_j \left[ \text{Tr}(B_j^* \hat{X}) \frac{1}{2} [S, B_j] + \delta_S \hat{X} \right].
$$

If $g$ denotes the metric on $\mathfrak{M}$, then the divergence of $X$ at $\pi(K)$ is given by

$$
\text{div}(X) \circ \pi = \sum_i g(\pi_*(S_i), \nabla_{\pi_*(S_i)} X) \circ \pi = \sum_i \text{Tr}(S_i^\ast \nabla_{S_i} \hat{X})
$$

where we used that $S_i$ is horizontal and that $g(\pi_*(S_i), \pi_*(Y)) = \text{Tr}(S_i^\ast Y)$ for all $Y$.

Using the identity above and the fact that $S_i^\ast = -S_i$ which implies $\text{Tr}(S_i^\ast [S_i, B_j]) = 0$ the expression reduces to

$$
\text{div}(X) \circ \pi = \sum_i \delta_{S_i} \text{Tr}(S_i^\ast \hat{X})
$$

(A.12)
As $\text{Tr}(S^*_1 Y) = 0$ for any vertical vector $Y \in \mathfrak{h}$, the lifted vector field $\hat{X}$ does not need to be horizontal for the last equation to hold.

### A.8 Measurability of $\{(\mu, \nu) : \nu \text{ is a.c. w.r.t. } \mu \}$

Let us consider for some convenience a compact interval $I = [a, b]$. Let $\mathcal{M}_+$ denote the set of non-negative, bounded measures on the Borel $\sigma$-algebra of $[-\pi, \pi]$. Consider $\mathcal{M}_+^2 = \mathcal{M}_+ \times \mathcal{M}_+$ with the product topology of the $*$-weak topology on $\mathcal{M}_+$ induced by the continuous functions $C([a, b])$. We want to consider the set $\mathcal{A} = \{(\mu, \nu) \in \mathcal{M}_+^2 : \nu \text{ is a.c. w.r.t. } \mu \}$.

The main result of this section is the following.

**Theorem A.7** The set $\mathcal{A} = \{(\mu, \nu) \in \mathcal{M}_+^2 : \nu \text{ is a.c. w.r.t. } \mu \}$ is Borelian.

Before proving this let us state that this theorem implies the measurability of the set of pairs $\{(\mu, \nu) : \nu \text{ is a.c. w.r.t. } \mu \}$ of measures as needed in Section 8.7.

**Corollary A.5** Let $\hat{\mathcal{M}}_+$ denote the set of non-negative measures on the Borel $\sigma$-algebra of $\mathbb{R}$ where $x \mapsto \frac{1}{x+1}$ is integrable. Furnish $\hat{\mathcal{M}}_+^2 = \hat{\mathcal{M}}_+ \times \hat{\mathcal{M}}_+$ with the product topology of the weak topology on $\hat{\mathcal{M}}_+$ induced by the functions $f_z : x \mapsto \Im(\frac{1}{x-z})$ for $z$ in the complex upper half plane. Then the set $\hat{\mathcal{A}} = \{ (\hat{\mu}, \hat{\nu}) \in \hat{\mathcal{M}}_+^2 : \hat{\nu} \text{ is a.c. w.r.t. } \hat{\mu} \}$ is Borelian.

**Proof.** Let $\hat{C}(\mathbb{R})$ be the set of continuous functions $g$ on $\mathbb{R}$ where $\lim_{x \to -\infty} g(x)(1+x^2)$ and $\lim_{x \to \infty} g(x)(1+x^2)$ exist. Furthermore let $\mathcal{F}$ be the map defined on $C([-\frac{\pi}{2}, \frac{\pi}{2}])$ with values in $\hat{C}(\mathbb{R})$ given by $(\mathcal{F}f)(x) = \frac{f(\arctan(x))}{1+x^2}$. Note that $\mathcal{F}$ is actually invertible and for $g \in \hat{\mathcal{C}}(\mathbb{R})$ one has $(\mathcal{F}^{-1}g)(y) = g(\tan(y))(1+y^2)$. The dual operator $\mathcal{F}^*$ can be seen as an injective map from $\hat{\mathcal{M}}_+$ to $\hat{\mathcal{M}}_+$, (choosing the compact interval for the measures $\mathcal{M}_+$ to be $I = [-\frac{\pi}{2}, \frac{\pi}{2}]$.)

In order to prove that $\mathcal{F}^*$ is a Borel measurable map (w.r.t. the considered weak topologies), one has to prove measurability of $(\mathcal{F}^*)^{-1}(\{ \hat{\mu} \in \hat{\mathcal{M}}_+ : \hat{\mu}(f) \in O \}) = \{ \hat{\mu} \in \hat{\mathcal{M}}_+ : \hat{\mu}(\mathcal{F}f) \in O \}$, for any $f \in C([-\frac{\pi}{2}, \frac{\pi}{2}])$ and open set $O \subset \mathbb{C}$. This follows immediately, if the map $\hat{\mu} \mapsto \hat{\mu}(\mathcal{F}f)$ is measurable w.r.t. the Borel sets of the weak topology we defined above on $\hat{\mathcal{M}}_+$. Clearly, the functions $\hat{\mu} \mapsto \hat{\mu}(f_z)$ are continuous and hence measurable w.r.t. this topology. Using the fact that the functions $f_z$ approximate a Dirac peak at $x_0$ for $z \to x_0$ (coming from the upper half plane), one finds that the family $f_z$ spans a weakly dense subset in $\hat{\mathcal{C}}(\mathbb{R})$ (w.r.t. the weak topology induced by $\hat{\mathcal{M}}_+$). Approximating $\mathcal{F}f$ weakly shows that the map $\hat{\mu} \mapsto \hat{\mu}(\mathcal{F}f)$ is measurable for any $f \in C([-\frac{\pi}{2}, \frac{\pi}{2}])$.

Therefore $\mathcal{F}^*$ is a Borel measurable map (w.r.t. the considered weak topologies). Furthermore one has $\hat{\mathcal{A}} = (\mathcal{F}^* \times \mathcal{F}^*)^{-1}(\mathcal{A})$ and hence $\hat{\mathcal{A}}$ is Borelian by Theorem A.7.

In order to prove Theorem A.7 let me introduce some more notations. Let $\mathcal{M}_{+\cdot}^2$ denote the set of pairs of measures in $\mathcal{M}_+^2$ where each measure has total mass at most $n$, i.e. $\mathcal{M}_{+\cdot}^2 = \{ (\mu, \nu) : \nu([a, b]) \leq n \land \mu([a, b]) \leq n \}$. Furthermore let $\mathcal{A}^n = \mathcal{A} \cap \mathcal{M}_{+\cdot}^2$, then $\mathcal{A} = \bigcup_n \mathcal{A}^n$. For $(\mu, \nu) \in \mathcal{M}_+^2$ let $\nu_\mu$ be the a.c. part of $\nu$ and $\nu_{\perp \mu}$ the singular part of $\nu$. 

w.r.t. \( \mu \). Hence \( \nu = \nu_\mu + \nu_\perp \mu \) is the Lebesgue decomposition. For any \( n \in \mathbb{N}, \epsilon > 0 \) and positive continuous function \( f \in C([a,b]) \) define
\[
\mathcal{A}_{\epsilon, f}^n = \{(\mu, \nu) \in \mathcal{M}_{+, n}^2 \mid \nu_\perp \mu(f) < \epsilon \}.
\]
Clearly \( \mathcal{A}^n \subset \mathcal{A}_{\epsilon, f}^n \) and \( \mathcal{A}^n = \bigcap_{\epsilon > 0, f > 0, f \in C([a,b])} \mathcal{A}_{\epsilon, f}^n \). By the Weierstrass approximation theorem, the set of continuous functions furnished with the norm of uniform convergence is a separable space. Therefore there is countable family of non-negative, continuous functions \( f_m \in C([a,b]) \) which is dense in the set of all non-negative continuous functions on \([a,b]\). Hence one can actually write \( \mathcal{A}^n \) as a countable intersection
\[
\mathcal{A}^n = \bigcap_{l,n \in \mathbb{N}} \mathcal{A}_{\epsilon, f_m}^n.
\]
(A.13)

Furthermore, for each \( k, n \in \mathbb{N} \) and \( \epsilon > 0 \) let \( \mathcal{N}_{k, \epsilon}^n = \{(\mu, \nu) \in \mathcal{M}_{+, n}^2 \mid \nu \geq (f_k - \epsilon)\mu \} \). Note that \( \mathcal{N}_{k, \epsilon}^n \) is closed in \( \mathcal{M}_{+, n}^2 \) and hence a Borel set.

First we need some important lemma.

**Lemma A.2** For any \( (\mu, \nu) \in \mathcal{A}^n \) and \( g \in C([a,b]), g \geq 0 \), there exists a \( k \in \mathbb{N} \) and an open neighborhood \( U(\mu, \nu) \) such that the intersection \( U(\mu, \nu) \cap \mathcal{N}_{k, \epsilon/5g}^n \) is contained in \( \mathcal{A}^n \).

Proof. Take a pair \( (\mu, \nu) \in \mathcal{A}^n \), then \( \nu = h\mu \) for an integrable, positive function \( h \in L^1(\mu) \). As \([a,b]\) is compact and the function \( g \) is bounded one has \( h \in L^1(\mu) \). The set of continuous functions is dense in \( L^1(\mu) \), therefore there exists a non-negative function \( f \in C([a,b]) \), such that \( f \leq h \) and
\[
0 \leq (h\mu - f\mu)(g) = g\mu(h - f) \leq \epsilon/5.
\]
Now choose \( f_k \) such that \( \|f - f_k\| < \frac{\epsilon}{5n\|g\|} \), then one obtains
\[
|\mu(gh - gf_k)| \leq |\mu(gh - gf)| + |\mu(gf - gf_k)| < \epsilon/5 + \frac{\mu(\|g\|\epsilon)}{5n\|g\|} \leq 2\epsilon/5.
\]
Define
\[
U(\mu, \nu) = \{(\hat{\mu}, \hat{\nu}) : |(\hat{\nu} - \nu)(g)| < \epsilon/5 \text{ and } |(\hat{\mu} - \mu)(f_k g)| < \epsilon/5 \}.
\]
It is not hard to see that \( U(\mu, \nu) \) is an open set in \( \mathcal{M}_{+, n}^2 \) and \( (\mu, \nu) \in U(\mu, \nu) \). Furthermore \( \nu = h\mu \geq f\mu = (f_k - \epsilon/5n\|g\|)h\mu \) and hence also \((\mu, \nu) \in \mathcal{N}_{k, \epsilon/5g}^n \). We want to show that \( U(\mu, \nu) \cap \mathcal{N}_{k, \epsilon/5g}^n \subset \mathcal{A}^n \). Therefore let \( (\hat{\mu}, \hat{\nu}) \) be an element of the intersection on the left hand side. As \( \hat{\nu} \geq (f_k - \epsilon/5n\|g\|)\hat{\mu} \), one has \( \hat{\nu}_1 \geq (f_k - \epsilon/5n\|g\|)\hat{\mu} \), and hence
\[
\hat{\nu}_1(g) = \left( \hat{\nu} - \left( f_k - \frac{\epsilon}{5n\|g\|} \right) \hat{\mu} \right)(g) \leq \left| \frac{\epsilon}{5n\|g\|} \hat{\mu}(g) \right| + \left| (\hat{\nu} - f_k\hat{\mu})(g) \right| \leq \frac{\epsilon}{5} + \left| (\hat{\nu} - \nu)(g) \right| + \left| (h\mu - f_k\mu)(g) \right| + \left| (\mu - \hat{\mu})(gf_k) \right| < \epsilon.
\]
A.8. MEASURABILITY OF \{ (\mu, \nu) : \nu \text{ IS A.C. W.R.T. } \mu \} \\

This means \((\hat{\mu}, \hat{\nu}) \in \mathcal{A}^n_{l,m} \) because one already knows \((\hat{\mu}, \hat{\nu}) \in \mathcal{H}^n_{k,\frac{\|f_m\|}{\|\mu\|}} \subset \mathcal{M}_+^n \).

\[ \text{Proof of Theorem A.7.} \]

Consider the sets \(\mathcal{A}^n_{l,m} \) as in equation (A.13) and fix \(l, m, n\) for the moment. Then let \(\delta = \frac{1}{5dn\|f_m\|} \). By the previous lemma, for each \((\mu, \nu) \in \mathcal{A}^n\) there is an open set \(U(\mu, \nu)\) and a \(k(\mu, \nu) \in \mathbb{N}\) such that

\[ (\mu, \nu) \in U(\mu, \nu) \cap \mathcal{H}^n_{k(\mu, \nu), \delta} \subset \mathcal{A}^n_{l,m}. \quad (A.14) \]

Now collect all pairs \((\mu, \nu)\) having the same \(k\) in the statement above and let \(O_k\) be the union of the corresponding open sets \(U(\nu, \mu)\), i.e. \(O_k = \bigcup_{(\mu, \nu) : k(\mu, \nu) = k} U(\mu, \nu)\). Then \(O_k\) is open and hence measurable and by (A.14) one has

\[ \mathcal{A}^n \subset \mathcal{B}^n_{l,m} = \bigcup_{k \in \mathbb{N}} (O_k \cap \mathcal{H}^n_{k, \delta}) \subset \mathcal{A}^n_{l,m} \]

where the equation above defines the measurable set \(\mathcal{B}^n_{l,m}\). We do the same procedure for any \(l, m, n\) to get the measurable sets \(\mathcal{B}^n_{l,m}\). By (A.13) one then obtains

\[ \mathcal{A}^n = \bigcap_{l,m \in \mathbb{N}} \mathcal{B}^n_{l,m} \Rightarrow \mathcal{A} = \bigcup_{n \in \mathbb{N}} \mathcal{A}^n = \bigcup_{n \in \mathbb{N}} \bigcap_{l,m \in \mathbb{N}} \mathcal{B}^n_{l,m} \quad (A.15) \]

and therefore \(\mathcal{A}\) is Borel measurable. \(\square\)
Bibliography


Curriculum Vitae

name Christian Hermann Sadel
date of birth 22nd March, 1982
place of birth Bamberg

Education


July 1999 - Aug. 1999 Participant of a pre-college program at the Johns Hopkins University of Baltimore, MD, USA.


Oct. 2003 - March 2004 Study at the University of York, GB, part of the Erasmus exchange program.

since May 2006 Ph.D. study in the Math Department of the F.A. University of Erlangen-Nürnberg; advisor: H. Schulz-Baldes.

Jan. 2009 - May 2009 Research assistant of Prof. J. Bellissard at the Georgia Institute of Technology, Atlanta GA, USA
Declaration
I hereby declare that this document has been composed by me and is based on my own work, unless otherwise acknowledged in the text.

Selbständigkeitserklärung
Hiermit erkläre ich, diese Arbeit selbständig und nur unter der Verwendung der angegebenen Literatur verfasst zu haben.