Quantum Chaos

Lecture Notes

Sebastian Müller (University of Bristol)

2019/20, teaching block 1

Contents

Chapter 0

Introduction

Quantum chaos

- **times:**
  Wednesday 9am (Fry G.13)
  Thursday 1pm (Fry G.09)
  Friday 10am (Fry G.09)
- **13-14 lectures**
- **4-5 problem/revision classes**
- some material covered in **videos**
- **5 homeworks** (usually Thursday to Thursday) set from booklet
- **assessment:** 100% by exam in January
- lecture material is on **Blackboard**
- **lecturer:** Sebastian Müller (Fry 2A.09, sebastian.muller@bristol.ac.uk)
- **drop in session:** Tuesday 10am

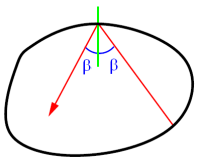
Books

- **Quantum Chaos: An Introduction (Hans-Jürgen Stöckmann)**
  Cambridge University Press (1999)
  an undergraduate textbook on quantum chaos
- **Nonlinear Dynamics and Quantum Chaos (Sandro Wimberger)**
  Springer (2014)
  also an undergraduate textbook
- **Chaos: Classical and Quantum (Pedrag Cvitanovic et al.)**
  ChaosBook.org, Niels Bohr Institute, Copenhagen (2010)
  a webbook about classical and quantum chaos
- **Quantum Signatures of Chaos (Fritz Haake, Sven Gnutzmann, Marek Kus)**
  Springer Verlag (4th edition 2019, previous editions by Haake),
  contains a lot of material on quantum chaos and random matrix theory, parts of chapter 9 and 10 are relevant for this unit

Classical chaos

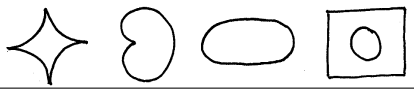
Billiards (classical)

particle moving in domain $B \subset \mathbb{R}^2$



- motion on straight line with constant velocity
- reflection at boundary
  (angle of incidence = angle of reflection)

Billiards with 'irregular' boundary show chaotic behaviour.

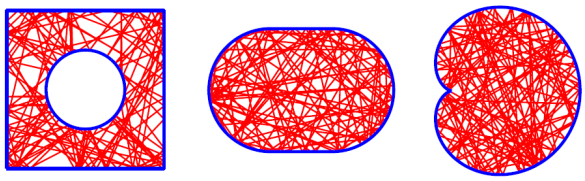


Chaos

1. sensitive dependence on initial conditions

Chaos

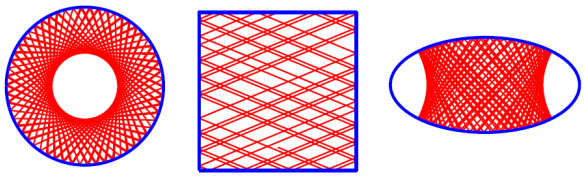
2. long trajectories fill the interior almost uniformly
   + all directions are equally likely
   (**ergodicity**)



Integrable billiards

For billiards with 'regular' boundary:

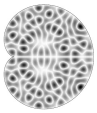
1. no sensitive dependence on initial conditions
2. no ergodicity:
   e.g. long trajectories they fill only part of the interior, prefer certain directions, show regular patterns



Quantum chaos

quantum properties of systems that are classically chaotic

Billiards (quantum)



- **Schrödinger equation** in domain $B \subset \mathbb{R}^2$, no potential

$$-\frac{\hbar^2}{2m}\nabla^2\psi(\boldsymbol{r},t) = i\hbar\frac{\partial}{\partial t}\psi(\boldsymbol{r},t)$$

Dirichlet boundary conditions:

$$\psi(\boldsymbol{r},t) = 0 \quad \text{for} \quad \boldsymbol{r} \in \partial B$$

- **time independent Schrödinger equation**

$$-\frac{\hbar^2}{2m}\nabla^2\psi_n(\boldsymbol{r}) = E_n\psi_n(\boldsymbol{r})$$

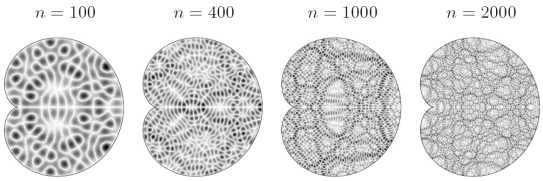
$\psi_n(\boldsymbol{r})$ – energy eigenfunctions

E_n – energy levels

Quantum chaos

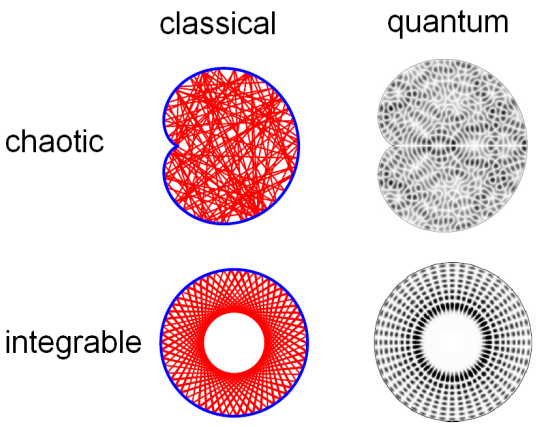
1. **sensitive dependence on initial conditions ?**
   NO, because Schrödinger equation is linear
2. **energy eigenfunctions**

plot $|\psi_n(\boldsymbol{r})|^2$ for eigenfunctions with increasing energy:



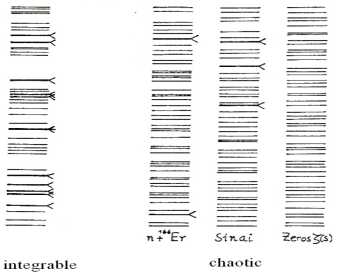
eigenfunctions become more and more equidistributed
(**quantum ergodicity**)

Quantum chaos



Quantum chaos

- **statistics of energy levels**



(large) energy levels of a chaotic system **repel** each other!
statistical properties of energy levels in chaotic systems are universal
want to understand why

Syllabus

- Classical chaos
- Quantum ergodicity
- Semiclassical approximation to connect classical and quantum behaviour
  - The propagator
  - Stationary-phase approximation
  - The trace formula
- Spectral statistics

Examples for chaos

chaotic behaviour can be observed in many

- billiards
- atoms, molecules, nuclei, electric conductors, many-particle systems, systems in quantum field theory . . .

There is also a connection to number theory.

Chapter 1

Classical chaos

We start by discussing some classical properties of chaotic systems, using the example of billiard systems.

1.1 Phase space of billiards

In a billiard, a particle of mass m moves inside a domain $B \subset \mathbb{R}^2$. It moves with constant speed along straight lines. When hitting the boundary the particle is reflected according to the reflection law (angle of incidence = angle of reflection). The state of a billiard can be characterised by

- the position of the particle $\boldsymbol{r} \in B \subset \mathbb{R}^2$
- and its momentum $\boldsymbol{p} = m\dot{\boldsymbol{r}} \in \mathbb{R}^2$.

Here $\boldsymbol{r}$ and $\boldsymbol{p}$ can be collected into a four-dimensional vector $(\boldsymbol{r}, \boldsymbol{p}) \in B \times \mathbb{R}^2$. The space $B \times \mathbb{R}^2$ where these vectors live is called the **phase space** of the billiard.

The speed $|\dot{\boldsymbol{r}}|$ of the particle and hence its energy $E = \frac{1}{2}m|\dot{\boldsymbol{r}}|^2 = \frac{|\boldsymbol{p}|^2}{2m}$ remain constant, they are conserved quantities. As a consequence a particle that is started with a given energy E will never be able to access points in phase space with a different energy. It can at most access the points with the same E. These points form the **energy shell** associated to E. They can be parametrised by only three real numbers:

- the position $\boldsymbol{r} \in B \subset \mathbb{R}^2$
- and the direction angle ϕ with

$$\boldsymbol{p} = \sqrt{2mE} \underbrace{\begin{pmatrix} \cos\phi \\ \sin\phi \end{pmatrix}}_{\equiv \boldsymbol{d}} .$$

1.2 Dependence on initial conditions

We will now investigate how the motion of a billiard ball changes if we slightly change the initial conditions. We will see that for certain billiards with an 'irregular' boundary the motion depends sensitively on initial conditions. These billiards are called chaotic.

Possible changes of initial conditions. To change the initial conditions of a billiard trajectory we have the following possibilities:

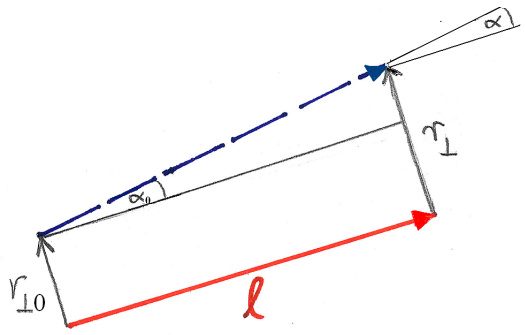


Figure 1.1: Initial deviations $r_{\perp 0}, \alpha_0$ and final deviations $r_\perp, \alpha$ for a straight line.

- **Shift the initial position in the direction of the trajectory.** This will leave the form of the trajectory unchanged, the start point will just be earlier or later.
- **Shift the initial position in the direction perpendicular to the trajectory.**
- **Change the absolute value of the initial momentum (and thus the energy).** This will not affect the form of the trajectories (straight lines and reflection according to the reflection law). It will just mean that the particle goes through these trajectories faster or slower.
- **Change the direction of the initial momentum.**

The second and fourth kind of change lead to nontrivial results. Hence we should study them in more detail. When doing so we will stick to small changes of the initial conditions and approximate accordingly. We will also allow for changes of initial conditions that combine a change in position perpendicular to the trajectory and a change of the initial angle. We use the following variables:

- The small shift perpendicular to the direction of the trajectory will be denoted by $r_{\perp 0}$. $r_{\perp 0}$ will be taken positive for a shift to the left and negative for a shift to the right.
- The small change of the initial angle will be denoted by α_0, taken positive for a change to the left and negative for a change to the right. (As the mathematically positive sense of rotation is to the left.)

Assuming an initial change by $r_{\perp 0}$ and α_0, we now want to determine the resulting change of the final position and direction, denoted by $r_\perp$ and α. We will do this first for a straight line and for a reflection, and then combine these results to study trajectories with several straight lines and reflections.

Straight lines. We first consider a straight path of length l as in Fig. 1.1. We see that along this path the perpendicular deviation increases by $l \tan \alpha_0$ which for small α_0 can be arroximated by $l\alpha_0$. Hence $r_\perp$ can be approximated by $r_{\perp 0} + l\alpha_0$. In a straight line the direction remains the same, hence we have $\alpha = \alpha_0$. Writing these equations in matrix form we obtain

$$\begin{pmatrix} r_\perp \\ \alpha \end{pmatrix} \approx \begin{pmatrix} 1 & l \\ 0 & 1 \end{pmatrix} \begin{pmatrix} r_{\perp 0} \\ \alpha_0 \end{pmatrix} . \tag{1.1}$$

This equation, like similar equations to follow, is an approximation to leading order in $r_{\perp 0}$ and α_0.

Reflections. We now study reflections at a part of the boundary with circular shape. Here the centre of the circle is assumed to lie outside the billiard, i.e., the part of the boundary is concave. We

consider a trajectory reflected from this circle with angle of incidence and angle of reflection β. In this case the deviations $r_\perp$ and α after the reflection are related to the deviations $r_{\perp 0}$ and α_0 before the reflection by

$$\begin{pmatrix} r_\perp \\ \alpha \end{pmatrix} \approx \begin{pmatrix} -1 & 0 \\ -\frac{2}{R\cos\beta} & -1 \end{pmatrix} \begin{pmatrix} r_{\perp 0} \\ \alpha_0 \end{pmatrix} . \tag{1.2}$$

The proof of this statement is a little involved. We leave it for appendix A.1 and it is not examinable.

Now what happens for reflections at parts of a boundary that are not circles? In this case we approximate the boundary locally by a circle. The radius of that circle is radius of curvature of the boundary. (Apart from the sign convention this is the inverse of the curvature defined in Calculus 1.) It has to be inserted for R in (1.2). If the boundary is locally concave as in the above example the radius of curvature of curvature is taken positive. For locally convex parts of the boundary it is negative.

Trajectories with several reflections and straight lines. Using the building blocks established above, we can consider a trajectory with $\nu+1$ straight-line segments of length l_μ ($\mu = 1, 2, \ldots, \nu+1$) and ν reflections between these segments with radius of curvature R_μ and angle of reflection β_μ ($\mu = 1, 2, \ldots, \nu$). If we change the initial condition of this trajectory by $r_{\perp 0}, \alpha_0$, the corresponding change after the first straight-line segment is given by (1.1). Then the change after the first reflection is obtained by multiplying with a matrix as in (1.2). Continuing like this will the subsequent straight lines and reflections we see that the final change is given by:

$$\begin{pmatrix} r_\perp \\ \alpha \end{pmatrix} \approx M \begin{pmatrix} r_{\perp 0} \\ \alpha_0 \end{pmatrix}$$
$$M = \begin{pmatrix} 1 & l_{\nu+1} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ -\frac{2}{R_\nu \cos\beta_\nu} & -1 \end{pmatrix} \cdots \begin{pmatrix} 1 & l_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ -\frac{2}{R_1 \cos\beta_1} & -1 \end{pmatrix} \begin{pmatrix} 1 & l_1 \\ 0 & 1 \end{pmatrix}$$

The matrix M is called the **stability matrix** of the trajectory.

Determinant. The stability matrix is a product of matrices with determinant 1.

$$\det \begin{pmatrix} 1 & l_\mu \\ 0 & 1 \end{pmatrix} = 1, \qquad \det \begin{pmatrix} -1 & 0 \\ -\frac{2}{R_\mu \cos\beta_\mu} & -1 \end{pmatrix} = 1 .$$

Hence it has determinant 1 itself.

Eigenvalues. To study the dependence of the billiard dynamics on initial conditions it is helpful to determine the eigenvalues of M. For instance, assume that we have an eigenvalue larger than one and choose the corresponding eigenvector as the initial deviation $(r_{\perp 0}, \alpha_0)$ between two trajectories. Then it is clear that the deviation will be larger in the end of the trajectory, i.e., the motion will depend sensitively on the initial conditions. The eigenvalues of M will be denoted by m_1 and m_2. We know that

$$\det M = m_1 m_2 = 1 .$$

Hence the eigenvalues are mutually inverse,

$$m_2 = \frac{1}{m_1} .$$

Now we can compute the trace

$$\operatorname{tr} M = m_1 + m_2 = m_1 + \frac{1}{m_1}$$

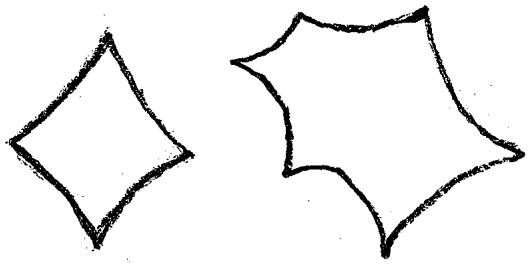


Figure 1.2: Dispersing billiards.

and multiply with m_1 to get a quadratic equation for m_1 (or, equivalently, m_2):

$$m_1^2 - m_1 \mathrm{tr}\, M + 1 = 0 \; .$$

This equation has the solutions

$$m_{1,2} = \frac{1}{2}\left(\mathrm{tr}\, M \pm \sqrt{(\mathrm{tr}\, M)^2 - 4}\right) \; . \tag{1.3}$$

Hence the behaviour of the system depends on $|\mathrm{tr}\, M|$. We have to distinguish between the three cases $|\mathrm{tr}\, M| > 2$, $|\mathrm{tr}\, M| = 2$, and $|\mathrm{tr}\, M| < 2$.

Case I: $|\mathrm{tr}\, M| > 2$ (hyperbolic/unstable case)

For $|\mathrm{tr}\, M| > 2$ the expression below the square root in (1.3) is positive. Hence the eigenvalues m_1 and m_2 are real. As they are also mutually inverse this means that one eigenvalue (say, m_1) has an absolute value larger than one whereas the other eigenvalue (say, m_2) has an absolute value smaller than one. We denote the corresponding eigenvectors by $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$. If the initial deviation between two trajectories is chosen proportional to $\boldsymbol{u}_1$ it will increase with a factor m_1 with $|m_1| > 1$, i.e., the motion depends sensitively on the initial conditions. $\boldsymbol{u}_1$ is called the **unstable direction**. Similarly initial deviations along the direction $\boldsymbol{u}_2$ decrease by a factor m_2 with $|m_2| = \frac{1}{|m_1|} < 1$. Hence $\boldsymbol{u}_2$ is called the **stable direction**. If the initial deviations are chosen as a linear combination of $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ the unstable component will increase and the stable component will decrease. However for large $|m_1|$ the increase of the unstable component dominates.

Example: dispersing billiards. An example for the case $|\mathrm{tr}\, M| > 2$ are dispersing billiards. These are billiards consisting of locally concave segments as in Fig. 1.2.

When studying the dynamics of dispersing billiards two types of trajectories have to be studied separately. These are trajectories hitting corners, and trajectories that pass one of the boundary segments tangentially (being just in between trajectories that are reflected and trajectories that just pass by close the boundary segment without hitting it). For both types of trajectories it is not possible to use stability matrices to study the effects of small changes of the initial conditions. Hence we will disregard these trajectories in the following.

For the other trajectories we have the following result:

Theorem. For all other trajectories with $\nu \geq 1$ reflections

- $|\mathrm{tr}\, M|$ is larger than 2 and is increased by adding further reflections or straight-line segments,
- and all matrix elements of M are nonzero and have the sign $(-1)^\nu$.

Proof. We proceed by induction. As the theorem doesn't hold for $\nu = 0$ we have to start with $\nu = 1$. A trajectory with one straight line of length l_1, a reflection (with angle of reflection β_1, radius of curvature $R_1 > 0$, and $a_1 = \frac{2}{R_1 \cos\beta_1} > 0$), and a further straight line of length l_2 has the stability matrix

$$\begin{aligned} M &= \begin{pmatrix} 1 & l_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ -a_1 & -1 \end{pmatrix} \begin{pmatrix} 1 & l_1 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & l_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & -l_1 \\ -a_1 & -a_1 l_1 - 1 \end{pmatrix} \\ &= \begin{pmatrix} -1 - a_1 l_2 & -l_1 - a_1 l_1 l_2 - l_2 \\ -a_1 & -a_1 l_1 - 1 \end{pmatrix} \end{aligned}$$

All symbols in this matrix are positive. The only possible exception is that either l_1 or l_2 could be zero. (However the case $l_1 = l_2 = 0$ is excluded as it would not be considered a proper trajectory.) Given this fact and the minus signs, we see that all entries of the matrix are < 0. Hence as desired they have the sign $(-1)^\nu = -1$. The trace satisfies

$$|\mathrm{tr}\, M| = |-1 - a_1 l_2 - a_1 l_1 - 1| = 2 + a_1 \underbrace{(l_1 + l_2)}_{>0} > 2\ ,$$

also in line with the statement.

We now investigate what happens if we add a reflection or a straight-line segment to the trajectory. We have to show that in this case $|\mathrm{tr}\, M|$ stays above 2 and even increases, and that the signs of all matrix elements change if we add a reflection but stay the same if we add a straight-line segment.

We first **add a reflection to an existing trajectory**. If we denote the stability matrix of the existing trajectory by M and the stability matrix including the reflection by $\tilde{M}$ we obtain

$$\tilde{M} = \begin{pmatrix} -1 & 0 \\ -a & -1 \end{pmatrix} \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} = \begin{pmatrix} -M_{11} & -M_{12} \\ -aM_{21} - M_{22} & -aM_{12} - M_{22} \end{pmatrix}\ .$$

with $a = \frac{2}{R\cos\beta}$ if β is the angle of reflection and R the corresponding radius of curvature. The elements of $\tilde{M}$ involve elements of M and minus signs (and sometimes a factor $a > 0$). Hence all these elements are nonzero and their signs are opposite to the sign of the matrix elements of M. Moreover we have

$$|\mathrm{tr}\, \tilde{M}| = |-M_{11} - aM_{12} - M_{22}| = |\mathrm{tr}\, M| + a|M_{12}| > |\mathrm{tr}M| > 2$$

where we used that M satisfies the theorem and hence $M_{12} \neq 0$ and $|\mathrm{tr}M| > 2$.

Now we **add a straight line of length l to an existing trajectory**. Again denoting the new stability matrix by $\tilde{M}$ we find

$$\tilde{M} = \begin{pmatrix} 1 & l \\ 0 & 1 \end{pmatrix} \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} = \begin{pmatrix} M_{11} + M_{21} l & M_{12} + M_{22} l \\ M_{21} & M_{22} \end{pmatrix}$$

Due to $l > 0$ all elements of $\tilde{M}$ have the same sign as the elements of M. The trace satisfies

$$|\mathrm{tr}\, \tilde{M}| = |\mathrm{tr}\, M + M_{21} l| = |\mathrm{tr}M| + |M_{21}| l > |\mathrm{tr}\, M| > 2\ .$$

Hence all statements of the theorem have been proved.

Lyapunov exponent. In the case $|\mathrm{tr}\, M| > 2$ the eigenvalues of M can be written as

$$m_1 = \pm e^{\lambda_L L}, \quad m_2 = \pm e^{-\lambda_L L}$$

where L is the length of the trajectory and $\lambda_L > 0$ is called the **finite time Lyapunov exponent**. One can obtain λ_L from m_1 via

$$\lambda_L = \frac{1}{L} \ln |m_1|$$

It is interesting to study what happens to λ_L if we follow the trajectory for a longer and longer length L. For dispersing billiards (and many other types of types of chaotic billiards) it turns out that λ_L converges; the limit is called the (infinite time) **Lyapunov exponent** λ,

$$\lambda_L \to \lambda \quad (L \to \infty).$$

For a very long trajectory it is a good approximation to replace λ_L by λ in the formula for m_1,

$$m_1 \sim \pm e^{\lambda L} \, .$$

m_1 is the factor by which an initial deviation between trajectories along the direction $\boldsymbol{u}_1$ increases after length L. We see that the increase is exponential. For dispersing billiards we have already above that $|\mathrm{tr} M|$ increases if further reflections or straight-line segments are added to a trajectory. The exponential increase is an even stronger statement.

Case II: $|\mathrm{tr}\, M| = 2$ (parabolic/marginally stable case)

In the case $|\mathrm{tr}\, M| = 2$ the expression below the square root in (1.3) vanishes and the eigenvalues are

$$m_1 = m_2 = \frac{1}{2} \mathrm{tr}\, M = \pm 1 \, .$$

Hence the deviation between two nearby trajectories either stays constant or flips its sign. This happens for many integrable systems such as the rectangular billiard (see below) or the circle billiard (which will be considered in an exercise).

Example: Rectangular billiard. The rectangular billiard (see Fig. 1.3) has a boundary consisting of straight lines, with a vanishing curvature and thus an infinite radius of curvature. Therefore, the stability matrix for each reflection is the diagonal matrix

$$\begin{pmatrix} -1 & 0 \\ -\frac{2}{R\cos\beta} & -1 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \, .$$

Multiplication with this matrix will just flip the sign, we can thus multiply with the number -1 instead. The stability matrix of a trajectory with ν reflections is hence given by the sign factor $(-1)^\nu$ times the product of the stability matrices associated to the motions along straight lines

$$M = (-1)^\nu \prod_\mu \begin{pmatrix} 1 & l_\mu \\ 0 & 1 \end{pmatrix} \, . \tag{1.4}$$

Here l_μ is the length of the μth straight line. It is easy to show that the matrix product in (1.4) just gives

$$M = (-1)^\nu \begin{pmatrix} 1 & \sum_\mu l_\mu \\ 0 & 1 \end{pmatrix} \, . \tag{1.5}$$

This means up to a sign the stability properties of a trajectory in the rectangular billiard are like those of a single straight line of length $L = \sum_\mu l_\mu$. The eigenvalues are $m_1 = m_2 = \frac{1}{2}\mathrm{tr}\, M = (-1)^\nu$. I.e. we

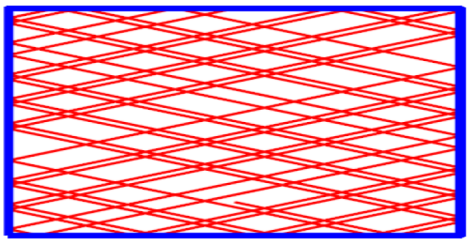


Figure 1.3: Rectangular billiard. (Picture by Arnd Bäcker.)

have one double degenerate eigenvalue, and the size of deviations neither grows nor shrinks but it can flip sign.

Side remark (optional): The eigenvectors $\boldsymbol{u}$ corresponding to this eigenvalue can be determined from

$$\begin{aligned} 0 &= \left[M - m_1 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right] \boldsymbol{u} = \left[(-1)^\nu \begin{pmatrix} 1 & L \\ 0 & 1 \end{pmatrix} - (-1)^\nu \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right] \boldsymbol{u} = (-1)^\nu \begin{pmatrix} 0 & L \\ 0 & 0 \end{pmatrix} \boldsymbol{u} \\ \Rightarrow \boldsymbol{u} &\propto \begin{pmatrix} 1 \\ 0 \end{pmatrix} . \end{aligned}$$

So there is only one eigenvector! This is in line with the theorems of Linear Algebra: In general you are only guaranteed that there is at least one eigenvector per eigenvalue. For some types of matrices (e.g. symmetric or orthogonal ones) there is also a guarantee of n linearly independent eigenvectors for an eigenvalue with degeneracy n but as M is neither symmetric nor orthogonal this statement does not apply in our case.

Case III: $|\mathrm{tr}\, M| < 2$ (elliptic/stable case)

In the case $|\mathrm{tr}\, M| < 2$ the **eigenvalues** $m_{1,2} = \frac{1}{2}\left(\mathrm{tr}\, M \pm \sqrt{(\mathrm{tr}\, M)^2 - 4}\right)$ are complex and complex conjugate with respect to each other, i.e., $m_2 = m_1^*$. One can show that in this case the motion also does not depend sensitively on the initial conditions, but to do so one has to understand the meaning of the complex eigenvalues as the motion is in real space. This will be left for appendix A.2 and it is not examinable. Some examples for elliptic/stable trajectories will be studied in exercises.

Mixed phase space

There are many systems for which some trajectories have $|\mathrm{tr}\, M| > 2$ whereas others have $|\mathrm{tr}\, M| = 2$ or $|\mathrm{tr}\, M| < 2$. Again, examples will be given in problems.

1.3 Ergodicity

If we plot a long trajectory in a billiard such as the Sinai billiard, the stadium billiard, the cardioid billiard, or a dispersing billiard (see Fig. 1.4) we see that the interior of the billiard is filled almost uniformly. We also see that all angles of direction are equally likely. Recalling that the energy shell of a billiard is parameterised by the positions $\boldsymbol{r} \in B$ and direction angles $\phi \in [0, 2\pi)$ we can thus say that the energy shell is filled almost uniformly. This property is known as **ergodicity**. In the following we want to formulate it in a quantitative way:

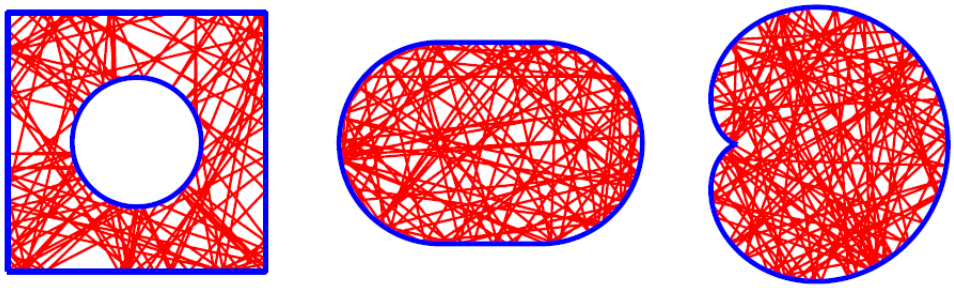


Figure 1.4: Long trajectories in the Sinai, stadium, and cardioid billiard, displaying ergodic behaviour. (Picture by Arnd Bäcker.)

We consider a part of the energy shell with position vectors $\boldsymbol{r}$ in a region $B_0 \subseteq B$ (where B is the billiard domain), and direction angles ϕ in an interval $[\phi_0, \phi_0 + \Delta\phi)$. This part of the energy shell will be denoted by $\Omega = B_0 \times [\phi_0, \phi_0 + \Delta\phi)$. Then for the billiards mentioned above we have (apart from the exceptions below)

$$\frac{\text{time a trajectory spends in } \Omega \text{ up to time } T}{T} \quad \to \quad \frac{|B_0|\Delta\phi}{|B|2\pi} \quad (T \to \infty)\,. \tag{1.6}$$

I.e. for large T the ratio of its time the trajectory spends in Ω converges to the ratio between the size of Ω and the size of the energy shell. Here the size of Ω is obtained by multiplying the area $|B_0|$ of B_0 with the angle range $\Delta\phi$. Similarly the size of the energy shell is obtained by multiplying the area $|B|$ with the range of angles 2π.

Exceptions. However, there are exceptions to (1.6). For example, if a particle moves on a periodic orbit it will move through the same positions and angles again and again without exploring the rest of the energy shell. For a system to be called ergodic these exceptions have to be rare. More precisely, one has to demand that **the exceptions are of "measure zero"**. This term is properly defined in measure theory but roughly it means the following: We consider the three dimensional energy shell parameterised by $\boldsymbol{r}$ and ϕ. Inside this energy shell we then consider those values of $\boldsymbol{r}$ and ϕ that, taken as initial conditions, will lead to exceptional trajectories. We now integrate over these initial conditions. A system is called ergodic if this three dimensional integral is zero. This would be the case for example if the exceptions would form a line in three dimensional space, as a line only has a length but no three dimensional volume. (Similarly an plane in three dimensions has an area but no volume). This quantifies the amount of exceptions permissible in an ergodic system.

Example. As an example we try to find out how often long (non-exceptional) trajectories in an ergodic billiard are reflected from the boundary. We are also interested in the position where the boundary is hit and the reflection angle. The position is indicated by a coordinate (arclength) s ranging from 0 to the circumference of the boundary. Instead of the reflection angle β we use $p = \sin\beta$ as a variable (the reason for this will become clear later). We denote by T the time for which we follow the trajectory. Then $N(T)$ indicates the number reflections occurring during this time at a given boundary segment of length Δs with p in the interval $[p_0, p_0 + \Delta p)$. Then we get

$$\lim_{T\to\infty} \frac{N(T)}{T} = \frac{v}{2\pi|B|}\Delta s \Delta p \tag{1.7}$$

where v is the speed of the billiard ball.

Proof. We only consider the case that Δs and Δp are small. For this case we identify the part of the energy shell Ω where the particle can be found during a small time interval τ before the reflection. As the particle stays in Ω for a time τ per reflection, $N(T)$ is the time spent in Ω divided by τ. Ergodicity then gives

$$\lim_{T\to\infty} \frac{N(T)}{T} = \lim_{T\to\infty} \frac{\text{time spent in } \Omega \text{ up to } T}{\tau\, T} = \frac{|B_0|\Delta\phi}{\tau|B|2\pi}. \tag{1.8}$$

Now how can we describe Ω geometrically? As Δs is small we approximate the segment of the boundary where the particle will be reflected by a straight line. As Δp is small all reflections involve reflection angles β close to β_0 with $p_0 = \sin\beta_0$. In the time interval τ before reflection the particle is thus in a parallelogram with side lengths Δs and $v\tau$. As seen in the figure one of the angles enclosed between these sides is $\frac{\pi}{2} + \beta_0$.

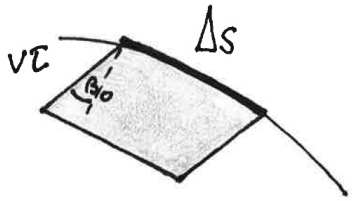


Figure 1.5: Area where particles are in a time interval τ before a reflection with reflection angle β_0.

This leads to an area

$$|B_0| = v\tau\Delta s\cos\beta_0.$$

The range of possible direction angles (enclosed with the x-axis) $\Delta\phi$ coincides with the range reflection angles (enclosed with the normal direction) $\Delta\beta$. We can relate this to the range of p using that $p = \sin\beta$ implies $\Delta p = \cos\beta_0\Delta\beta$ (where $\Delta\beta$ is an infinitesimal change away from β_0). Hence we have

$$\Delta\phi = \frac{1}{\cos\beta_0}\Delta p.$$

If we insert these results into (1.8) we obtain (1.7), as desired.

Numerical experiment (Poincaré plot). To check the above result numerically one can generate numerically a long trajectory in an ergodic billiard, say, the cardioid billiard, and record s and p for every reflection point in a two dimensional plot, see Fig 1.6. One sees that the whole plot is filled with points, and their density is the same everywhere. This is in line with (1.7) which indicates that the number of points (divided by T and taken in the limit $T \to \infty$) should be the same for equally large areas in the plot. (Note that the densities would have been different if we had used β as a variable instead of p.)

For systems that are not ergodic the plot looks different. E.g. for the circle billiard, which is completely integrable, all points fill a line with constant p because for a trajectory in such a billiard all reflection angles coincide (see Fig. 1.7). For systems whose shape is between a circle and a cardioid it depends on the initial conditions: for some initial condition one gets a picture very similar to the one for the cardioid but leaving out some parts of the plot (the "chaotic sea"); for other initial conditions one obtains regular looking structures (the "stability islands") in particular circles. This is typical for systems with mixed phase space. The form of these islands can be understood a little better based on a more careful treatment of elliptic/stable trajectories, see appendix A.3.

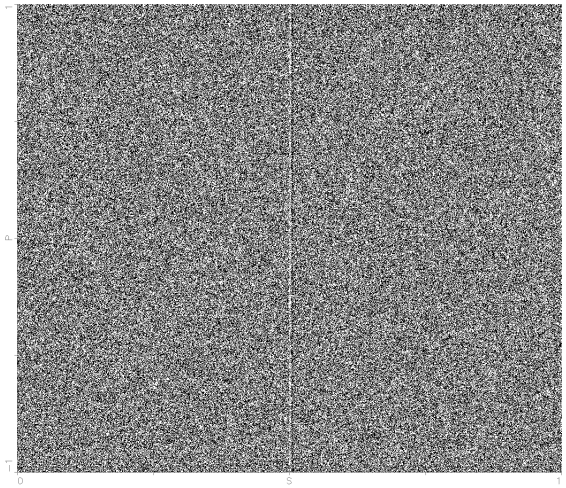


Figure 1.6: Poincaré plot for the cardioid billiard. The points indicate where and with which angle a particle bounces from the wall. The x-coordinate is s, the y-coordinate is $p = \sin\beta$. The line in the middle is just a numerical artefact. (Picture generated with a program from "Chaos: A Program Collection for the PC", by H.J. Korsch, H.-J. Jodl, Springer Verlag, 1998.)

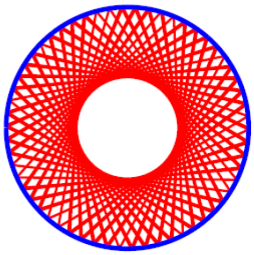


Figure 1.7: Trajectory in a circular billiard. All angles of reflection are the same. (Picture by Arnd Bäcker.)

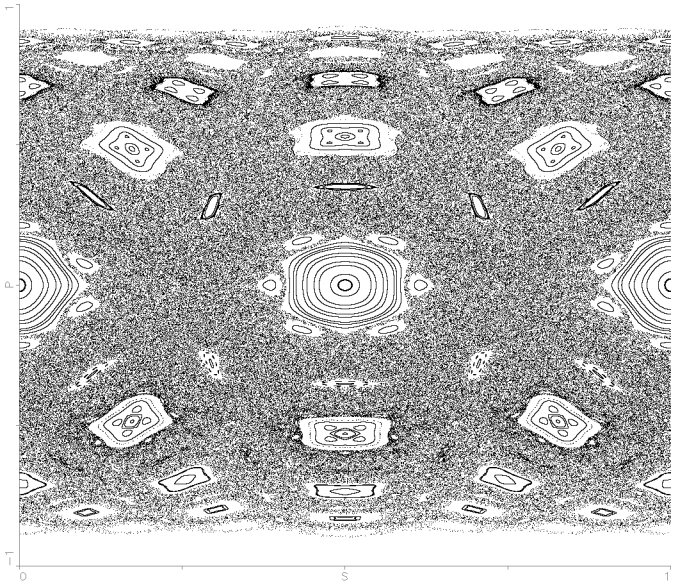


Figure 1.8: As in Fig. 1.6, but for a billiard between cardioid and circular shape. The are filled with dots corresponds to one trajectory, the circles corresponding to other trajectories. (Picture generated with a program from the book by H.J. Korsch and H.-J. Jodl.)

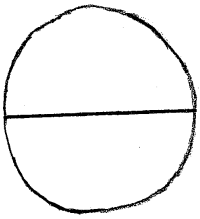


Figure 1.9: A periodic orbit in a billiard with a shape between circle and cardioid.

Chapter 2

Quantum ergodicity

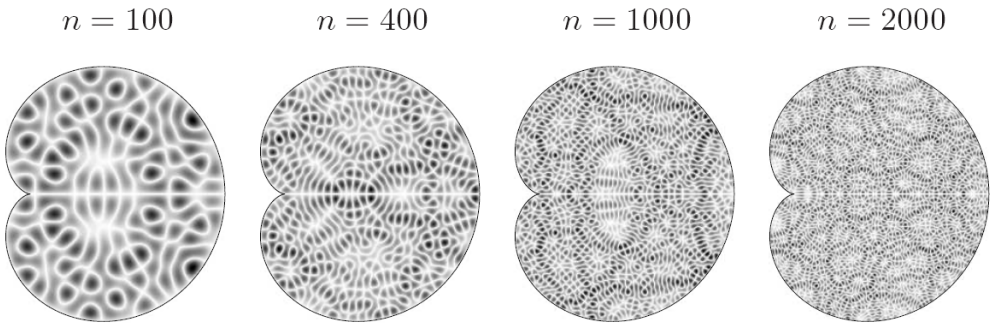


Figure 2.1: Probablity density for three energy eigenfunctions in the cardioid billiard. (Picture by Arnd Bäcker.)

We now consider the energy eigenfunctions $\psi_n(\boldsymbol{r})$ of ergodic systems, and the corresponding probability density $|\psi_n(\boldsymbol{r})|^2$. In Fig. 2.1 we depict the probability density for different eigenfunctions of the cardioid billiard. We see that for larger larger energy levels the plot becomes more and more uniform, and appears to lead to a constant density in the limit of infinite energy.

To formulate this statement quantitatively we consider the quantum mechanical probability that a particle in the n-th eigenstate is found in a part $B_0 \subseteq B$ of the billiard domain. This probability is denoted by $\mu_n(B_0)$. It can be obtained by integrating the probability density over B_0,

$$\mu_n(B_0) = \int_{B_0} |\psi_n(\boldsymbol{r})|^2 d^2r \ .$$

Now we take the limit of large energies. If we order the energy levels according to $E_1 \leq E_2 \leq E_3 \leq \ldots$ this limit can be written as $n \to \infty$. In this limit our above observation suggests that the probability to be found in a part B_0 should just be proportional to the area $|B_0|$ of B_0. Due to the normalisation of the wavefunction it then has to be equal to $\frac{|B_0|}{|B|}$. If all eigenfunctions behaved as in Fig. 2.1 we would thus expect the following limit to hold:

$$\mu_n(B_0) \ \ \to \ \ \frac{|B_0|}{|B|} \ \ (n \to \infty) \ . \tag{2.1}$$

Exceptions. However like for classical ergodicity there are exceptions. Some wavefunctions of ergodic billiards are very nonuniform (even for large energies). These exceptions are called **"scars"**. An example for the cardioid billiard is shown in Fig. 2.2. Here the wavefunction is concentrated close to the periodic orbit that is just bouncing up and down. Similarly to the exceptional classical trajectories

these exceptional eigenfunctions are "rare". We must define properly what rare means and then write down a correct version of Eq. (2.1). The idea is that from the sequence of natural numbers 1 2 3 4 5 6 7 8 9 10 ... we have to take out those for which $\psi_n(\boldsymbol{r})$ is exceptional. To consider a simplified example let us first imagine that instead we take out the squares, keeping the numbers 2 3 5 6 7 8 10 ... . We then define the sequence enumerating the remaining natural numbers $n_1 = 2, n_2 = 3, n_3 = 5, n_4 = 6$, etc. The exceptions are considered rare if this sequence is a sequence of density one.

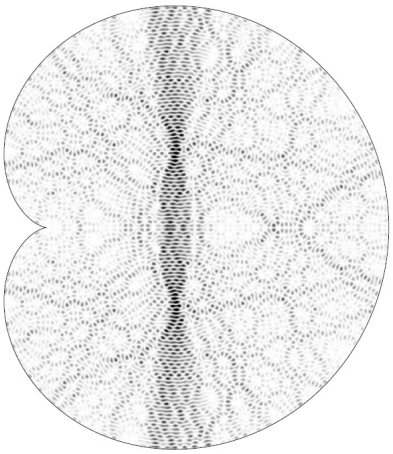


Figure 2.2: A scar in the cardioid billiard. (Picture by Arnd Bäcker.)

Definition. A sequence n_k is called a **sequence of density one** if the number of exceptions below N satisfies

$$\frac{\#\text{exceptions} \leq N}{N} \;\rightarrow\; 0 \;\; (N \rightarrow \infty) \, .$$

Here the exceptions are natural numbers m that do not appear in the sequence n_k. Hence we can write more formally

$$\frac{\#\{m \leq N : m \neq n_k \forall k\}}{N} \;\rightarrow\; 0 \;\; (N \rightarrow \infty) \, .$$

Example: Let us determine the number of exceptions $\leq N$ for the sequence of natural numbers excluding the squares. If N is a square itself there are $\sqrt{N}$ exceptions, the squares of the numbers from 1 to $\sqrt{N}$. If N is not a square we have to take instead $[\sqrt{N}]$, which is defined as the smallest integer $\leq \sqrt{N}$. We thus have

$$\frac{\#\text{exceptions} \leq N}{N} = \frac{[\sqrt{N}]}{N} \;\rightarrow\; 0 \;\; (N \rightarrow \infty) \, ,$$

i.e., our sequence is a sequence of density one.

Now the corrected version of Eq. (2.1) is:

Shnirelman's theorem for billiards. For each classically ergodic billiard there is a sequence n_k of density one such that

$$\mu_{n_k}(B_0) = \int_{B_0} |\psi_{n_k}(\boldsymbol{r})|^2 d^2 r \;\rightarrow\; \frac{|B_0|}{|B|} \;\; (k \rightarrow \infty) \, .$$

This property is called quantum ergodicity.

The proof is tricky, we are not going to do it here.

Remarks. We are considering the quantum probability to be in an area in position space, not in a part of the energy shell. The reason for this is that in quantum mechanics we cannot make exact measurements of position and momentum at the same time.

There are a few systems for which one does not need to restriction to a sequence of density one; for these systems the original statement (2.1) is correct. This property is called quantum unique ergodicity. In 2010, Elon Lindenstrauss received a Fields medal partly for his work on quantum unique ergodicity.

Chapter 3

The propagator

So far we have discussed the sensitive dependence on initial conditions and ergodicity as classical properties of chaotic systems, and quantum ergodicity as a quantum property. We will later consider properties of the energy levels, in particular the effect chaos has on the statistical distribution of energy levels. We then also want to investigate how this quantum mechanical property follows from classical chaos, and to this end we need tools to connect the classical and quantum behaviour of a system. We start to explore these tools in the present chapter, by introducing the propagator of a quantum system. This is simply the wavefunction for a particular initial condition, namely the condition that the wavefunction is localised at a given position $\boldsymbol{r}_0$ at time $t = 0$. We will then see that the propagator can be expressed in terms of classical trajectories of the system. We will first consider simple systems where this expression in terms of trajectories is exact and then dispersing billiards for which the expression in terms of trajectories is an approximation.

3.1 Definition and properties

The wavefunction. We start by recalling how the wavefunction $\psi(\boldsymbol{r}, t)$ of a quantum system is determined.

- The wavefunction has to satisfy the time dependent **Schrödinger equation**

$$\hat{H}\psi(\boldsymbol{r}, t) = i\hbar\frac{\partial}{\partial t}\psi(\boldsymbol{r}, t)$$

  with the Hamiltonian

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + U(\boldsymbol{r})$$

  where $\nabla^2 = \Delta = \sum_i \frac{\partial^2}{\partial r_i^2}$.

- We need to specify an initial condition, or more generally the values $\psi(\boldsymbol{r}, t_0)$ of the **wavefunction at a particular time** t_0.

- If the system has a boundary we also have to specify **boundary conditions**. In this course we usually consider billiards with Dirichlet boundary conditions, meaning that $\psi(\boldsymbol{r}, t) = 0$ if $\boldsymbol{r}$ is on the boundary of the system, denoted by ∂B.

The propagator is the wavefunction with an initial wavefunction localised at $\boldsymbol{r}_0$. To express this initial condition we use a delta function:

"Definition". The **delta function** $\delta(\boldsymbol{r})$ in $\mathbb{R}^n$ is defined such that

- $\delta(\boldsymbol{r}) = 0$ for $\boldsymbol{r} \neq 0$
- $\delta(\boldsymbol{r})$ diverges at $\boldsymbol{r} = 0$
- $\int_{\mathbb{R}^n} f(\boldsymbol{r})\delta(\boldsymbol{r})d^n r = f(0)$ for functions $f(\boldsymbol{r})$.

The last line implies that $\int_{\mathbb{R}^n} \delta(\boldsymbol{r})d^n r = 1$.

Proper definition. (Not examinable.) The above definition is not rigorous; in particular it appears to redefine the notion of an integral. The rigourous definition is based on the theory of distributions. Distributions are mappings from a space of test functions $\mathcal{F}$ to $\mathbb{R}$. One often uses

$$\mathcal{F} = \{f \in C^\infty(\mathbb{R}^n) : (1 + |\boldsymbol{r}|^2)^k f(\boldsymbol{r}) \text{ is bounded for all } k \in \mathbb{N}\}.$$

Functions $g : \mathbb{R}^n \to \mathbb{R}$ give rise to distributions

$$\begin{aligned} &D_g : \mathcal{F} \to \mathbb{R} \\ &D_g[f] = \int_{\mathbb{R}^n} f(\boldsymbol{r})g(\boldsymbol{r})d^n r \in \mathbb{R} \end{aligned} \tag{3.1}$$

where $f \in \mathcal{F}$. The δ-distribution maps each function $f \in \mathcal{F}$ to its value at zero

$$\begin{aligned} &D_\delta : \mathcal{F} \to \mathbb{R} \\ &D_\delta[f] = f(0) \ . \end{aligned}$$

We now use the following notation motivated by (3.1)

$$D_\delta[f] = \int_{\mathbb{R}^n} f(\boldsymbol{r})\delta(\boldsymbol{r})d^n r \ .$$

We thus work with the δ-distribution as we would work with a function.

The delta function as a limit. $\delta(\boldsymbol{r})$ can be represented as the limit of a Gaussian with a peak at $\boldsymbol{r} = 0$. We normalise the Gaussian such that its integral is one. If we then increase the height and reduce the width the delta function is obtained as a limit. In the one dimensional case we have

$$\lim_{\epsilon \to 0} \frac{1}{(\epsilon\pi)^{1/2}} e^{-x^2/\epsilon} = \delta(x) \ ,$$

for $\boldsymbol{r} \in \mathbb{R}^n$ this generalises to

$$\lim_{\epsilon \to 0} \frac{1}{(\epsilon\pi)^{n/2}} e^{-\boldsymbol{r}^2/\epsilon} = \delta(\boldsymbol{r}) \ .$$

There is also a version of the previous formula where an i has been sneaked into the exponent and prefactor

$$\lim_{\epsilon \to 0} \frac{1}{(i\epsilon\pi)^{n/2}} e^{i\boldsymbol{r}^2/\epsilon} = \delta(\boldsymbol{r}) \ . \tag{3.2}$$

In this formula the limit procedure is a bit tricky. It involves first averaging over an interval of values for ϵ, and then taking the limit $\epsilon \to 0$ in the usual way. The formula obviously diverges for $\boldsymbol{r} = 0$, $\epsilon \to 0$ as in this case the prefactor diverges and the exponential is 1. For $\boldsymbol{r} \neq 0$, $\epsilon \to 0$ the exponent involves the imaginary unit times a factor going to infinity. Hence the exponential oscillates rapidly between -1 and 1. This wouldn't lead to a convergence in the usual sense. However if we first average over a small interval of ϵ the positive and negative contributions have a chance to cancel. If one takes this average first and then the limit $\epsilon \to 0$ one can show that the exponential goes to zero faster than the prefactor

diverges, and the expression converges to zero for $\boldsymbol{r} \neq 0$. (There is also a more formal definition of convergence in the theory of distributions that leads to the same result.)

Finally we have the following representation of the delta function as the limit of a Lorentzian. This formula will be needed only in one dimension.

$$\lim_{\epsilon \to 0} \frac{1}{\pi} \frac{\epsilon}{x^2 + \epsilon^2} = \delta(x). \tag{3.3}$$

We are now ready to define the propagator of a quantum system as follows.

Definition. The propagator $K(\boldsymbol{r}, \boldsymbol{r}_0, t)$ of a system with Hamiltonian $\hat{H}$ is defined such that

- it obeys the Schrödinger equation for $t \neq 0$

$$\hat{H} K(\boldsymbol{r}, \boldsymbol{r}_0, t) = i\hbar \frac{\partial}{\partial t} K(\boldsymbol{r}, \boldsymbol{r}_0, t)\,,$$

- it satisfies the initial condition

$$K(\boldsymbol{r}, \boldsymbol{r}_0, 0) = \delta(\boldsymbol{r} - \boldsymbol{r}_0)\,,$$

- if the system has a boundary $K(\boldsymbol{r}, \boldsymbol{r}_0, t)$ obeys the same boundary conditions as the wavefunction.

Remarks. This definition is in line with our characterisation of the wavefunction and the initial localisation at $\boldsymbol{r}_0$. The parameters $\boldsymbol{r}$ and t are the same as for the wavefunction, and the gradients in $\hat{H}$ act on $\boldsymbol{r}$. The parameter $\boldsymbol{r}_0$ just takes account of the initial condition. In the first property we excluded the case $t = 0$ to avoid technical complications associated to applying $\hat{H}$ to a highly singular function.

We also note that the propagator can equivalently be defined as

$$K(\boldsymbol{r}, \boldsymbol{r}_0, t) = \langle \boldsymbol{r} | e^{-\frac{i}{\hbar}\hat{H}t} | \boldsymbol{r}_0 \rangle$$

where $|\boldsymbol{r}_0\rangle$ is Dirac notation for the delta function located at $\boldsymbol{r}_0$; $e^{-\frac{i}{\hbar}\hat{H}t}$ is the time evolution operator defined in third year Quantum Mechanics. The propagator thus defined is the state of a system started with the wavefunction localised at $\boldsymbol{r}_0$ and evolved over a time t according to the Schr ödinger equation. Then the multiplication with $\langle \boldsymbol{r}|$ means that we look at its wavefunction at position $\boldsymbol{r}$. In this course we will not use Dirac notation and hence stick to the definition given earlier, however it is important to note that both are equivalent.

A nice property of the propagator is that it helps us to get the wavefunction for arbitrary initial conditions or conditions at time t_0. In this case the wafunction is given as an integral involving the propagator and the initial condition:

Theorem. The wavefunction $\psi(\boldsymbol{r}, t)$ is obtained from the wavefunction at time t_0 via

$$\psi(\boldsymbol{r}, t) = \int K(\boldsymbol{r}, \boldsymbol{r}_0, t - t_0) \psi(\boldsymbol{r}_0, t_0) d^n r_0\ . \tag{3.4}$$

Here and in the following the integrals go over $\mathbb{R}^n$ or, if the system has a boundary, just over the part of $\mathbb{R}^n$ enclosed by this boundary.

Proof. We show that the $\psi(\boldsymbol{r}, t)$ given in the theorem satisfies all the properties of the wavefunction.

- The Schrödinger equation is satisfied due to

$$\begin{aligned}\hat{H}\psi(\boldsymbol{r},t) &= \hat{H}\int K(\boldsymbol{r},\boldsymbol{r}_0,t-t_0)\psi(\boldsymbol{r}_0,t_0)d^n r_0\\ &= \int(\hat{H}K(\boldsymbol{r},\boldsymbol{r}_0,t-t_0))\psi(\boldsymbol{r}_0,t_0)\\ &= \int(i\hbar\frac{\partial}{\partial t}K(\boldsymbol{r},\boldsymbol{r}_0,t-t_0))\psi(\boldsymbol{r}_0,t_0)\\ &= i\hbar\frac{\partial}{\partial t}\int K(\boldsymbol{r},\boldsymbol{r}_0,t-t_0)\psi(\boldsymbol{r}_0,t_0)d^n r_0\\ &= i\hbar\frac{\partial}{\partial t}\psi(\boldsymbol{r},t)\ .\end{aligned}$$

  Here we used that derivatives w.r.t. $\boldsymbol{r}$ and t act only on the propagator but not on $\psi(\boldsymbol{r}_0,t_0)$, and that the propagator satisfies the Schrödinger equation.

- The condition at time t_0 is satisfied due to

$$\begin{aligned}\psi(\boldsymbol{r},t_0) &= \int K(\boldsymbol{r},\boldsymbol{r}_0,0)\psi(\boldsymbol{r}_0,t_0)d^n r_0\\ &= \int\delta(\boldsymbol{r}-\boldsymbol{r}_0)\psi(\boldsymbol{r}_0,t_0)d^n r_0\\ &= \psi(\boldsymbol{r},t_0)\ .\end{aligned}$$

  Here we used the initial condition for the propagator as well as the property of the delta function $\int\delta(\boldsymbol{r}-\boldsymbol{r}_0)f(\boldsymbol{r})d^n r = f(\boldsymbol{r}_0)$.

- For systems with a boundary we also have to show that $\psi(\boldsymbol{r},t)$ obeys the boundary condition. We will just consider the case of Dirichlet boundary conditions for all points $\boldsymbol{r}$ on the boundary ∂B. In this case the propagator satisfies

$$K(\boldsymbol{r},\boldsymbol{r}_0,t)=0 \ \text{ for all } \ \boldsymbol{r}\in\partial B\ .$$

  For the wavefunction at $\boldsymbol{r}\in\partial B$ this implies

$$\psi(\boldsymbol{r},t)=\int\underbrace{K(\boldsymbol{r},\boldsymbol{r}_0,0)}_{=0}\psi(\boldsymbol{r}_0,t_0)d^n r_0=0\ .$$

  as desired.

Remark. This theorem can be interpreted as follows: The propagator with time argument $t-t_0$ describes the evolution of the system over time $t-t_0$. This is what is needed to turn the wavefunction at time t_0 into the wavefunction at time t. If we ignore the time arguments we face a continuous integral over the second argument of K and the argument of ψ. This integral is analogous to the discrete sum one has to evaluate when multiplying a matrix with a vector.

Corollary. As the propagator $K(\boldsymbol{r},\boldsymbol{r}_I,t)$ is just the wavefunction with initial condition $\delta(\boldsymbol{r}-\boldsymbol{r}_I)$ we can substitute in (3.4) $\psi(\boldsymbol{r})=K(\boldsymbol{r},\boldsymbol{r}_I,t)$. This gives

$$K(\boldsymbol{r},\boldsymbol{r}_I,t)=\int K(\boldsymbol{r},\boldsymbol{r}_0,t-t_0)K(\boldsymbol{r}_0,\boldsymbol{r}_I,t_0)d^n r_0\ .$$

Remark. This result means that combination of the propagator for a time difference t_0 and a time difference $t - t_0$ gives the propagator over the total time t. The integration over $\boldsymbol{r}_0$ is analogous to the summation over indices when evaluating a product of two matrices.

3.2 Free propagator

We now want to determine the propagator of a particle moving in $\mathbb{R}^n$ without potential or boundaries. Quantum mechanically its motion is described by the Hamiltonian

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2.$$

The propagator is

$$K_f(\boldsymbol{r}, \boldsymbol{r}_0, t) = \frac{1}{(i\pi\epsilon)^{n/2}} e^{\frac{i}{\epsilon}(\boldsymbol{r}-\boldsymbol{r}_0)^2}$$

where $\epsilon = \frac{2\hbar t}{m}$.

Proof. The limit representation of the delta function (3.2) implies

$$K_f(\boldsymbol{r}, \boldsymbol{r}_0, t) \to \delta(\boldsymbol{r} - \boldsymbol{r}_0)$$

in the limit where $t \to 0$ and thus $\epsilon \to 0$. Hence the initial condition for the propagator is satisfied.

As there is no boundary there is no boundary condition to satisfy.

All we still have to show is that the Schrödinger equation is satisfied. The Schrödinger equation for the free propagator is

$$\hat{H} K_f(\boldsymbol{r}, \boldsymbol{r}_0, t) = i\hbar \frac{\partial}{\partial t} K_f(\boldsymbol{r}, \boldsymbol{r}_0, t) \ .$$

If we insert $\hat{H}$ and $t = \frac{m}{2\hbar}\epsilon$, $\frac{\partial}{\partial t} = \frac{2\hbar}{m}\frac{\partial}{\partial \epsilon}$ this gives

$$-\frac{\hbar^2}{2m}\nabla^2 K_f(\boldsymbol{r}, \boldsymbol{r}_0, t) = i\hbar \frac{2\hbar}{m}\frac{\partial}{\partial \epsilon} K_f(\boldsymbol{r}, \boldsymbol{r}_0, t)$$

or equivalently

$$\nabla^2 K_f(\boldsymbol{r}, \boldsymbol{r}_0, t) = -4i \frac{\partial}{\partial \epsilon} K_f(\boldsymbol{r}, \boldsymbol{r}_0, t). \tag{3.5}$$

To show that this equation is satisfied we first evaluate the r.h.s. We get

$$\begin{aligned} -4i\frac{\partial}{\partial \epsilon} K_f(\boldsymbol{r}, \boldsymbol{r}_0, t) &= (-4i)\left(-\frac{n}{2}\right)\frac{1}{(i\pi)^{n/2}\epsilon^{n/2+1}} e^{\frac{i}{\epsilon}(\boldsymbol{r}-\boldsymbol{r}_0)^2} \\ &\quad + (-4i)\frac{1}{(i\pi\epsilon)^{n/2}} e^{\frac{i}{\epsilon}(\boldsymbol{r}-\boldsymbol{r}_0)^2}\left(-\frac{i}{\epsilon^2}(\boldsymbol{r}-\boldsymbol{r}_0)^2\right) \\ &= K_f(\boldsymbol{r}, \boldsymbol{r}_0, t)\left(\frac{2in}{\epsilon} + \left(\frac{2i}{\epsilon}(\boldsymbol{r}-\boldsymbol{r}_0)\right)^2\right) \ . \end{aligned} \tag{3.6}$$

Here the first term originates from the derivative acting on the prefactor, and the second term originates from the derivative acting on the exponential.

To evaluate the l.h.s. we use that $\nabla^2 = \nabla \cdot \nabla = \operatorname{div}\operatorname{grad}$ where the gradient satisfies

$$\operatorname{grad}(\boldsymbol{r} - \boldsymbol{r}_0)^2 = 2(\boldsymbol{r} - \boldsymbol{r}_0)$$

and the divergence satisfies

$$\operatorname{div}(\boldsymbol{r}-\boldsymbol{r}_0) = \operatorname{div}\boldsymbol{r} = \sum_{i=1}^{n}\frac{\partial r_i}{\partial r_i} = n\ .$$

Application of these operation to the free propagator gives

$$\begin{aligned}\nabla^2 K_f(\boldsymbol{r},\boldsymbol{r}_0,t) &= \operatorname{div}\operatorname{grad} K_f(\boldsymbol{r},\boldsymbol{r}_0,t)\\ &= \operatorname{div}\frac{1}{(i\pi\epsilon)^{n/2}}e^{\frac{i}{\epsilon}(\boldsymbol{r}-\boldsymbol{r}_0)^2}\frac{2i}{\epsilon}(\boldsymbol{r}-\boldsymbol{r}_0)\\ &= \underbrace{\frac{1}{(i\pi\epsilon)^{n/2}}e^{\frac{i}{\epsilon}(\boldsymbol{r}-\boldsymbol{r}_0)^2}}_{=K_f(\boldsymbol{r},\boldsymbol{r}_0,t)}\left(\left(\frac{2i}{\epsilon}(\boldsymbol{r}-\boldsymbol{r}_0)\right)^2+\frac{2in}{\epsilon}\right) \qquad (3.7)\end{aligned}$$

where in the first summand both derivative acted on the exponent. In the second summand one derivative acted on the exponent yielding a factor $\frac{2i}{\epsilon}(\boldsymbol{r}-\boldsymbol{r}_0)$ whereas the second derivative acted on this factor. We see that (3.7) and (3.6) coincide, implying the Schrödinger equation (3.5).

3.3 System with a single "wall"

To get some practice with evaluating propagators for systems with boundary conditions we now consider a particle moving between along the positive half line without potential. At $x = 0$ we assume Dirichlet boundary conditions, i.e., the wavefunctions satisfy $\psi(0,t) = 0$ and the propagator satisfies $K(0,x_0,t) = 0$. This situation will arise if the potential for negative x is infinite. Then the particle can't access negative x meaning that the wavefunction and propagator for negative x become zero. Due to continuity the values for $x = 0$ thus also have to be zero. We want to show that the system in question has the propagator

$$K(x,x_0,t) = K_f(x,x_0,t) - K_f(x,-x_0,t) = \frac{1}{(i\epsilon\pi)^{1/2}}\left(e^{\frac{i}{\epsilon}(x-x_0)^2} - e^{\frac{i}{\epsilon}(x+x_0)^2}\right).$$

Proof. We have to check that the claimed $K(x,x_0,t)$ satisfies all the properties of the propagator.

- The Schrödinger equation is satisfied because it is satisfied by the two summands:

$$\begin{aligned}\hat{H}K(x,x_0,t) &= \hat{H}K_f(x,x_0,t) - \hat{H}K_f(x,-x_0,t)\\ &= i\hbar\frac{\partial}{\partial t}K_f(x,x_0,t) - i\hbar\frac{\partial}{\partial t}K_f(x,-x_0,t)\\ &= i\hbar\frac{\partial}{\partial t}K(x,x_0,t)\ .\end{aligned}$$

- To check that K has the correct initial condition we use the initial condition for K_f:

$$\begin{aligned}K(x,x_0,0) &= K_f(x,x_0,0) - K_f(x,-x_0,0)\\ &= \delta(x-x_0) - \delta(x+x_0)\end{aligned}$$

  Here the second summand is zero as we are only considering positions x and x_0 that are positive. Hence the argument of the second delta function is positive and the delta function vanishes.[1] (More abstractly speaking, the initial condition was left unchanged by subtracting $K_f(x,-x_0,t)$ because the second parameter was taken outside the range of positions in the problem.)

[1] To be careful, we still have to check whether the case $x = x_0 = 0$ is relevant as in this case the argument of the second delta function would be zero. However we should not have $x_0 = 0$ as then our initial condition would be a delta function localised at 0, and this would disagree with the boundary conditions.

- The subtraction of $K_f(x, -x_0, t)$ makes sure that the boundary condition is satisfied. We have

$$K(0, x_0, t) = K_f(0, x_0, t) - K_f(0, -x_0, t) = 0$$

because $K_f(0, x_0, t) = \frac{1}{(i\epsilon\pi)^{1/2}} e^{\frac{i}{\epsilon}x_0^2}$ is even in x_0.

Interpretation.

- The first summand $K_f(x, x_0, t)$ is the same as without the wall. The second term looks like a wave starting from $-x_0$. Alternatively it could be interpreted as a wave that starts from x_0 and is reflected at the wall.

- We can also relate the terms to classical trajectories (see Fig. 3.1). The first term is associated to a trajectory going directly from x_0 to x. Hence it depends on the length $|x - x_0|$. As the the second term depends on $x + x_0$ it could be associated to a trajectory starting at $-x_0$ and ending at x. More intuitively, it can be associated to a trajectory that starts at x_0, then goes to 0, is reflected, and ends at x. This trajectory also has the length $x + x_0$.

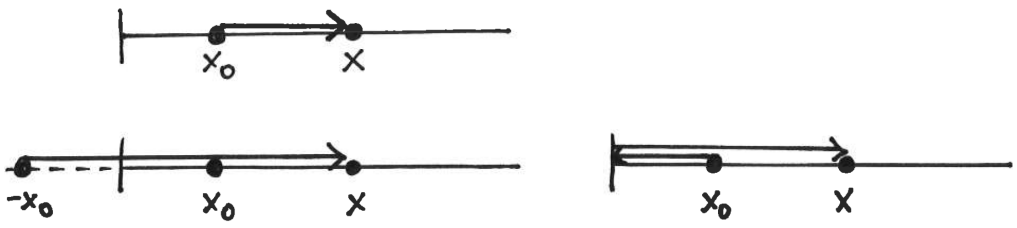


Figure 3.1: Trajectories contributing to the propagator of a one dimensional system with Dirichlet boundary conditions at the origin.

Generalisation. The idea to write the propagator as a combination of terms associated to classical trajectories carries over to many other situations. E.g. we could consider a particle in one dimension that can be reflected at two walls at $x = 0$ and $x = 1$, or in a rectangular box. In these cases (which will be investigated in problems) the correspondence to classical trajectories is exact and the propagator can be written as a linear combination of free propagators related to these trajectories. Then there are more complicated cases such as chaotic billiards. For these systems the propagator can still be *approximated* by a sum over classical trajectories. As an example for this situation we will study dispersing billiards.

3.4 van Vleck propagator

We want to prove an approximate formula for the propagator of dispersing billiards. This is a special case of a more general result derived by John Hasbrouck van Vleck.

Theorem. In dispersing billiards the propagator $K(\boldsymbol{r}, \boldsymbol{r}_0, t)$ can be approximated by by the following sum over classical trajectories γ going from $\boldsymbol{r}_0$ to $\boldsymbol{r}$,

$$K_V(\boldsymbol{r}, \boldsymbol{r}_0, t) = \sum_\gamma \frac{1}{i\epsilon\pi}\sqrt{\frac{L_\gamma}{|M_{\gamma,12}|}}(-1)^{n_\gamma} e^{\frac{i}{\epsilon}L_\gamma^2}\,. \tag{3.8}$$

Here we have again $\epsilon = \frac{2\hbar t}{m}$, n_γ is the number of reflections along the trajectory, L_γ is its length, and M_γ the stability matrix. The formula is an approximation in the following sense:

1. We have
$$\hat{H}K_V(\boldsymbol{r}, \boldsymbol{r}_0, t) = i\hbar\frac{\partial}{\partial t}K_V(\boldsymbol{r}, \boldsymbol{r}_0, t) + O(\epsilon^{-1}). \tag{3.9}$$
Here the two sides agree apart from differences of order ϵ^{-1}. As the expressions on the l.h.s and r.h.s. involve terms of order ϵ^{-2} and ϵ^{-3} this means that the error is small in comparison if ϵ is small.

2. The initial condition
$$K(\boldsymbol{r}, \boldsymbol{r}_0, 0) = \delta(\boldsymbol{r} - \boldsymbol{r}_0) \tag{3.10}$$
holds exactly.

3. The boundary condition
$$K(\boldsymbol{r}, \boldsymbol{r}_0, t) = 0 \quad \text{for } r \in \partial B \tag{3.11}$$
holds exactly.

Remarks.

- The contributions of trajectories to the claimed formula look almost like free propagators in two dimensions. However $|\boldsymbol{r} - \boldsymbol{r}_0|$ is replaced by the length and similarly as in the previous section reflections with Dirichlet boundary condition give rise to a sign factor. The only term one couldn't have guessed from our previous results is $\sqrt{\frac{L_\gamma}{|M_{\gamma,12}|}}$.
- When is our approximation good[2]? We need small $\epsilon = \frac{2\hbar t}{m}$. ϵ is of the dimension of a squared length as the exponent $\frac{i}{\epsilon}L^2$ has to be dimensionless. Hence ϵ has to be small compared to the squares of typical lengths in the problem such as lengths of trajectories or the size of the billiard. One can interpret this limit as follows:
  - As a limit of small times t (since $\epsilon \propto t$).
  - As a limit where the size of the billiard is increased whereas the other parameters such as t remain fixed. Indeed this is what we would expect. The van Vleck formula relates the quantum mechanical propagator to classical trajectories. Quantum mechanics is needed to describe small systems whereas classical mechanics describes the macroscopic world. A relation between the two should be useful for systems that are small enough to bother considering quantum mechanical quantities but large enough such that classical mechanics starts to become relevant. I.e. the systems should be large compared to other quantum systems. (A more quantitative way of saying this is that the size of the system should be much larger than the wavelength.)
  - In the research literature one often imagines that the limit is reached by taking $\hbar \to 0$. More precisely (as $\hbar$ has the dimension of an action, i.e., momentum $\times$ length) $\hbar$ is taken much smaller than typical actions in the system. This is also referred to as the semiclassical limit.

[2]Disclaimer: With "good" I mean that the van Vleck propagator approximately solves the Schödinger equation. Other notions of the quality of the approximation have been studied as well.

- As mentioned the van Vleck propagator given here is an example for a more general result that extends much beyond dispersing billiards. For general chaotic billiards this looks like (3.8) but with an additional phase factor that does not arise for dispersing billiards. For general Hamiltonian or Lagrangian systems the length is replaced as a variable by the classical action or the duration of the trajectory, and there are some further changes.

Proof.

2. For the second part we will give an informal proof. We first consider the trajectory leading on a straight line from $\boldsymbol{r}_0$ to $\boldsymbol{r}$, without reflections. This trajectory has the length

$$L_\gamma = |\boldsymbol{r} - \boldsymbol{r}_0| \ .$$

Its stability matrix is simply that of a straight line,

$$M_\gamma = \begin{pmatrix} 1 & |\boldsymbol{r} - \boldsymbol{r}_0| \\ 0 & 1 \end{pmatrix}$$

implying

$$M_{\gamma,12} = |\boldsymbol{r} - \boldsymbol{r}_0| \ .$$

Hence for this term the square root in (3.8) is 1. This leaves

$$\frac{1}{i\epsilon\pi} e^{\frac{i}{\epsilon}(\boldsymbol{r}-\boldsymbol{r}_0)^2}$$

which is just the free propagator. For the initial condition we need its limit for $t \to 0$, $\epsilon \to 0$. We have already seen that this limit is $\delta(\boldsymbol{r} - \boldsymbol{r}_0)$ which is the desired r.h.s. in (3.10).

Now we need to show that the remaining trajectories don't contribute to the initial condition. These trajectories involve reflections. Hence their length is positive[3]. This means that for $\epsilon \to 0$ the exponential $e^{\frac{i}{\epsilon}L_\gamma^2}$ oscillates rapidly. As in the beginning of the chapter we consider a limit process in which we first average over an interval and then take ϵ to zero. As during the averaging terms with opposite signs cancel we obtain a vanishing result in the end. Hence trajectories with reflections don't spoil the initial condition (3.10).

3. In the present approach the proof for the boundary condition becomes quite elegant. We need to consider $\boldsymbol{r}$ on the boundary but as a preparation let us consider $\boldsymbol{r}$ very close to the boundary. Trajectories ending at a $\boldsymbol{r}$ close to the boundary come in pairs, see Fig. 3.2. There are trajectories (denoted by γ_1) that stop just before hitting the boundary. Others (denoted by γ_2) hit the boundary and then go away from it by a small distance before reaching the final point $\boldsymbol{r}$. For each trajectory γ_1 one can obtain a trajectory γ_2 just by following the trajectory up to a bit after the next reflection, and then deforming the trajectory slightly to make the endpoint of γ_2 identical to the endpoint $\boldsymbol{r}$ of γ_1. The reflection numbers are related by

$$n_{\gamma_2} = n_{\gamma_1} + 1 \ .$$

Hence the signs in (3.8) are opposite due to

$$(-1)^{n_{\gamma_2}} = -(-1)^{n_{\gamma_1}} \ .$$

Now we will argue that in the limit where $\boldsymbol{r}$ is on the boundary the remaining factors in (3.8) are identical for the two trajectories, meaning that their contributions cancel exactly. Obviously in this limit the lengths are identical. For the stability matrices we have to take into account that

[3] Here we omit the special case of one single reflection with both $\boldsymbol{r}$ and $\boldsymbol{r}_0$ on the boundary.

γ_2 formally contains the reflection at the boundary whereas γ_1 does not. This means that M_{γ_2} is obtained from M_{γ_1} by multiplying with the stability matrix of a reflection. We thus get

$$M_{\gamma_2} = \begin{pmatrix} -1 & 0 \\ -\frac{2}{R\cos\beta} & -1 \end{pmatrix} \begin{pmatrix} M_{\gamma_1,11} & M_{\gamma_1,12} \\ M_{\gamma_1,21} & M_{\gamma_1,22} \end{pmatrix} = \begin{pmatrix} * & -M_{\gamma_1,12} \\ * & * \end{pmatrix} .$$

Hence we have indeed $|M_{\gamma_2,12}| = |M_{\gamma_1,12}|$ and as expected for $\boldsymbol{r} \in \partial B$ the contributions of all trajectories to the van Vleck propagator (3.8) cancel pairwise.

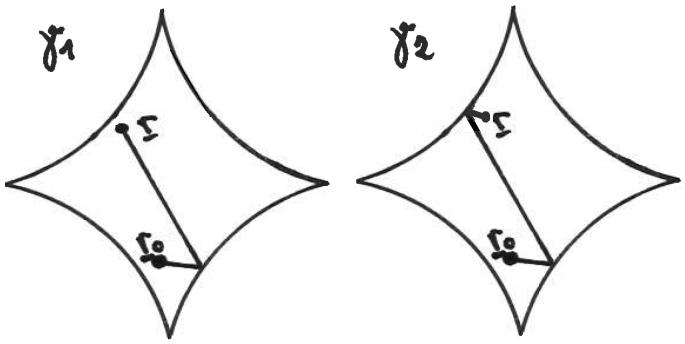


Figure 3.2: Trajectories giving compensating contributions to the van Vleck propagator in the limit where $\boldsymbol{r}$ is on the boundary.

1. We are now have to show that the van Vleck propagator $K_V(\boldsymbol{r}, \boldsymbol{r}_0, t)$ satisfies the Schrödinger equation in the two leading orders in $\frac{1}{\epsilon}$. In fact this is true even for the individual summands in $K_V(\boldsymbol{r}, \boldsymbol{r}_0, t)$, i.e. we will show that the summands in the van Vleck propagator (3.8)

$$K_{V,\gamma}(\boldsymbol{r}, \boldsymbol{r}_0, t) = \underbrace{\frac{(-1)^{n_\gamma}}{i\epsilon\pi|M_{\gamma,12}|^{1/2}}}_{=A} L_\gamma^{1/2} e^{\frac{i}{\epsilon}L_\gamma^2} \tag{3.12}$$

(with $\epsilon = \frac{2\hbar t}{m}$) satisfy

$$-\frac{\hbar^2}{2m}\nabla^2 K_{V,\gamma} = i\hbar\frac{\partial}{\partial t}K_{V,\gamma} + O(\epsilon^{-1}). \tag{3.13}$$

Similarly as in the proof of the free propagator (see Eq. (3.5)) we rewrite Eq. (3.13) as

$$\nabla^2 K_{V,\gamma} = -4i\frac{\partial}{\partial \epsilon}K_{V,\gamma} + O(\epsilon^{-1}). \tag{3.14}$$

To show that this holds we use the following lemma, to be proved later.

Lemma. Let us consider a trajectory from $\boldsymbol{r}_0$ to $\boldsymbol{r}$ that follows straight lines inside the billiard and obeys the reflection law. We write the length of the trajectory as a function of $\boldsymbol{r}$ and $\boldsymbol{r}_0$. Then we have

$$\frac{\partial L}{\partial \boldsymbol{r}} = \boldsymbol{d} \tag{3.15}$$

where $\boldsymbol{d}$ is the unit vector indicating the direction of the trajectory in the end.

The r.h.s. in Eq. (3.14) can be evaluated in the same way as the derivative of the free propagator w.r.t. ϵ (see Eq. (3.6)). We get

$$-4i\frac{\partial}{\partial \epsilon}K_{V,\gamma} = K_{V,\gamma}\left(\frac{4i}{\epsilon} + \left(\frac{2i}{\epsilon}L\right)^2\right) \tag{3.16}$$

where the first summand originates from the derivative acting on the prefactor and in the second summand the derivative acted on the exponent. We also left out the subscript γ of L. Compared to Eq. (3.5) we have specified the dimension n as 2 and generalised $|\boldsymbol{r}-\boldsymbol{r}_0|$ to the length of the trajectory L.

We now have to evaluate the l.h.s. of Eq. (3.14) using what we learnt about derivatives of the length. We will use nabla notation, i.e., write the gradient of a scalar function as $\operatorname{grad} f = \nabla f$ and the divergence of a vector valued function as $\operatorname{div} \boldsymbol{v} = \nabla \cdot \boldsymbol{v}$. The Laplacian is $\nabla^2 = \nabla \cdot \nabla$. Due to Eq. (3.12) we have to evaluate

$$\nabla^2 K_{V,\gamma} = \nabla^2 (A L^{1/2} e^{\frac{i}{\epsilon}L^2})$$

If instead of ∇ we had a derivative w.r.t. x we'd first apply this derivative to the product using the product rule, and then apply it again. The second derivative can act on the same factors as the first or on different ones. With ∇ instead of $\frac{d}{dx}$ we essentially have to do the same, we only have to pay a little bit of attention to see when nabla acts as a gradient and when it acts as a divergence. This determines where we have to put scalar products. I'll skip the algebra and just give the final result. This result will be ordered according to powers of $\frac{1}{\epsilon}$. We get no powers of $\frac{1}{\epsilon}$ if none of the derivatives acts on the exponential, a first power if only one derivative act on the exponential and two if both derivatives act on the exponential. The result reads

$$\begin{aligned}
& \nabla^2 K_{V,\gamma} \\
= \;& \nabla^2 \left(A L^{1/2} \right) e^{\frac{i}{\epsilon}L^2} \\
+ \;& \frac{2i}{\epsilon} \left(2 A L^{1/2} \underbrace{\nabla L \cdot \nabla L}_{\boldsymbol{d}\cdot\boldsymbol{d}=1} + \underbrace{2 L^{3/2} \nabla L \cdot \nabla A + A L^{3/2} \nabla^2 L}_{\equiv B} \right) e^{\frac{i}{\epsilon}L^2} \\
+ \;& \left(\frac{2i}{\epsilon} \right)^2 A L^{5/2} \underbrace{\nabla L \cdot \nabla L}_{=\boldsymbol{d}\cdot\boldsymbol{d}=1} e^{\frac{i}{\epsilon}L^2} \; . \qquad (3.17)
\end{aligned}$$

The first summand is of order ϵ^{-1} (as $A \propto \epsilon^{-1}$). It is responsible for the $O(\epsilon^{-1})$ in Eq. (3.13) and can be neglected compared to the second and third summand which are of order ϵ^{-2} and ϵ^{-3}. In the second summand we collected terms where only one derivative acted on the exponent returning a factor $\frac{2i}{\epsilon} L \nabla L$. The term with $\nabla^2 L$ results from the second derivative acting on ∇L, in all other cases we obtain scalar products with other gradients. We can simplify the term with $\nabla L \cdot \nabla L$ as ∇L gives the direction vector which is normalised. The remaining terms are grouped together in B which we will show to vanish. The term in the final line arises from the case that both derivatives act on the exponential; it can be simplified in a similar way. If we anticipate that B will be zero we can write our result as

$$\nabla^2 K_{V,\gamma}(\boldsymbol{r}, \boldsymbol{r}_0, t) = \underbrace{A L^{1/2} e^{\frac{i}{\epsilon}L^2}}_{=K_{V,\gamma}} \left(\frac{4i}{\epsilon} + \left(\frac{2i}{\epsilon} L \right)^2 \right) + O(\epsilon^{-1}) \; .$$

Up to the neglected terms this agrees with Eq. (3.16), implying the approximate Schrödinger equation (3.14).

However we still have to prove the lemma used above and we have to explain why B vanishes. This will be done below. The proof of the lemma involves some conceptually interesting ideas; the proof for $B = 0$ is presented for completeness in appendix A.4 and it is not examinable.

3.5 Derivatives of the length

We now prove the lemma above, and at the same time also a result about the derivative of the length with respect to the initial point.

Lemma. Let us consider a trajectory from $\boldsymbol{r}_0$ to $\boldsymbol{r}$ that follows straight lines inside the billiard and obeys the reflection law. We write the length of the trajectory as a function of $\boldsymbol{r}$ and $\boldsymbol{r}_0$. Then we have

$$\frac{\partial L}{\partial \boldsymbol{r}} = \boldsymbol{d}, \qquad \frac{\partial L}{\partial \boldsymbol{r}_0} = -\boldsymbol{d}_0 \tag{3.18}$$

where $\boldsymbol{d}$ is the unit vector indicating the direction of the trajectory in the end, and $\boldsymbol{d}_0$ is the unit vector indicating the direction in the beginning.

Proof. We first consider the case that **the trajectory is just a straight line**. In this case we have

$$L = |\boldsymbol{r} - \boldsymbol{r}_0|.$$

The direction is given by

$$\boldsymbol{d} = \boldsymbol{d}_0 = \frac{\boldsymbol{r} - \boldsymbol{r}_0}{|\boldsymbol{r} - \boldsymbol{r}_0|} \, .$$

Using $\frac{\partial |\boldsymbol{x}|}{\partial \boldsymbol{x}} = \frac{\boldsymbol{x}}{|\boldsymbol{x}|}$ we get

$$\frac{\partial L}{\partial \boldsymbol{r}} = \frac{\boldsymbol{r} - \boldsymbol{r}_0}{|\boldsymbol{r} - \boldsymbol{r}_0|} = \boldsymbol{d}$$

in line with the lemma, and similarly we have

$$\frac{\partial L}{\partial \boldsymbol{r}_0} = -\frac{\boldsymbol{r} - \boldsymbol{r}_0}{|\boldsymbol{r} - \boldsymbol{r}_0|} = -\boldsymbol{d}_0.$$

To cover the case of trajectories with **reflections** we proceed by induction. We assume that the statement holds for trajectories with up to ν reflections (where ν may be zero) and we show that this implies the statement for the case of $\nu + 1$ reflections. To do so we use that any trajectory with $\nu + 1$ reflections can be divided into a part A with ν reflections, the final reflection, and the part B without further reflections. The point of the final reflection will be denoted by $\tilde{\boldsymbol{r}}$. As $\tilde{\boldsymbol{r}}$ has to be on the boundary, the possible values of $\tilde{\boldsymbol{r}}$ can be parameterised by the arclength s. (The arclength gives the distance one has to travel along the boundary in a specified direction to get from a given reference point to $\tilde{\boldsymbol{r}}$.)

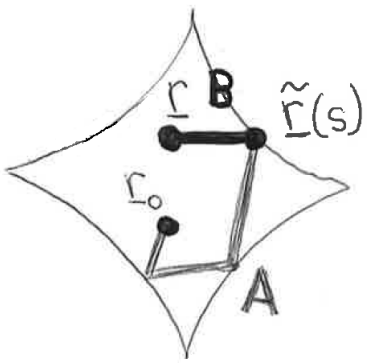


Figure 3.3: A trajectory with $\nu + 1 = 3$ reflections, decomposed into parts A and B as explained in the text.

The overall length of the trajectory can now be written as a sum of the length of part A from $\boldsymbol{r}_0$ to $\tilde{\boldsymbol{r}}(s)$ and the length of part B from $\tilde{\boldsymbol{r}}(s)$ to $\boldsymbol{r}$,

$$L(\boldsymbol{r}, \boldsymbol{r}_0) = L^B(\boldsymbol{r}, \tilde{\boldsymbol{r}}(s)) + L^A(\tilde{\boldsymbol{r}}(s), \boldsymbol{r}_0).$$

Here the reflection point has to be chosen such that the reflection law is satisfied. Hence s has to change if $\boldsymbol{r}$ or $\boldsymbol{r}_0$ is changed, to make sure that the trajectory with changed final or initial point still satisfies the reflection law. This dependence is not written out explicitly in the formula above.

Now to evaluate $\frac{\partial L}{\partial \boldsymbol{r}}$ we have to take derivatives w.r.t. all dependencies on $\boldsymbol{r}$ in the formula above, including the dependence of s on $\boldsymbol{r}$. Using the chain rule we obtain

$$\frac{\partial L}{\partial \boldsymbol{r}} = \frac{\partial L^B}{\partial \boldsymbol{r}} + \left(\frac{\partial L^B}{\partial \tilde{\boldsymbol{r}}} + \frac{\partial L^A}{\partial \tilde{\boldsymbol{r}}} \right) \cdot \frac{d\tilde{\boldsymbol{r}}}{ds} \frac{\partial s}{\partial \boldsymbol{r}}.$$

As by assumption the statement holds for A and B we know that $\frac{\partial L^B}{\partial \boldsymbol{r}}$ is the direction in the end of part B and hence in the end of the whole trajectory, denoted by $\boldsymbol{d}$. As $\tilde{\boldsymbol{r}}$ is the end of A and the beginning of B, $\frac{\partial L^A}{\partial \tilde{\boldsymbol{r}}}$ is the final direction of A denoted by $\boldsymbol{d}^A$ and $\frac{\partial L^B}{\partial \tilde{\boldsymbol{r}}}$ is the negative of the initial direction of B denoted by $\boldsymbol{d}_0^B$. We thus have

$$\frac{\partial L}{\partial \boldsymbol{r}} = \boldsymbol{d} + (-\boldsymbol{d}_0^B + \boldsymbol{d}^A) \cdot \frac{d\tilde{\boldsymbol{r}}}{ds} \frac{\partial s}{\partial \boldsymbol{r}}.$$

We would like everything after the first term to cancel, and this indeed follows from the reflection law. $\frac{\partial \tilde{\boldsymbol{r}}}{\partial s}$ points in the direction tangential to the trajectory, as $\tilde{\boldsymbol{r}}$ changes in that direction if is increased. If s is chosen as an arclength as above, $\frac{\partial \tilde{\boldsymbol{r}}}{\partial s}$ is the unit vector in tangential direction. $\boldsymbol{d}_0^B$ and $\boldsymbol{d}^A$ are also unit vectors, and the scalar product between two unit vectors is simply the angle they enclose. From simple geometric considerations we see that the angle enclosed between $\boldsymbol{d}^A$ and the tangential vector is equal to $\frac{\pi}{2} - \beta_{\text{in}}$ where β_{in} is the angle of incidence. Hence we have

$$\boldsymbol{d}^A \cdot \frac{d\tilde{\boldsymbol{r}}}{ds} = \cos\left(\frac{\pi}{2} - \beta_{\text{in}}\right) = \sin(\beta_{\text{in}})$$

and analogously

$$\boldsymbol{d}_0^B \cdot \frac{d\tilde{\boldsymbol{r}}}{ds} = \cos\left(\frac{\pi}{2} - \beta_{\text{out}}\right) = \sin(\beta_{\text{out}})$$

where β_{out} is the angle of reflection. The reflection law $\beta_{\text{in}} = \beta_{\text{out}}$ thus implies that

$$(-\boldsymbol{d}_0^B + \boldsymbol{d}^A) \cdot \frac{d\tilde{\boldsymbol{r}}}{ds} = 0$$

and hence

$$\frac{\partial L}{\partial \boldsymbol{r}} = \boldsymbol{d}.$$

To obtain $\frac{\partial L}{\partial \boldsymbol{r}_0} = -\boldsymbol{d}_0$ we could carry through this argument in a similar manner. However the result

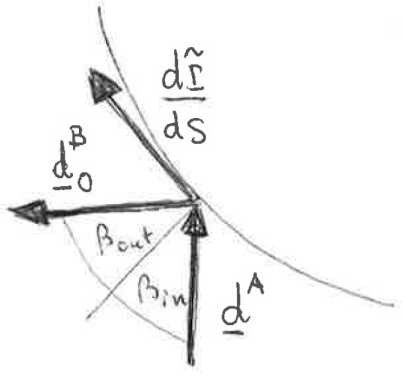


Figure 3.4: Relation of $\boldsymbol{d}^A, \boldsymbol{d}_0^B, \frac{d\tilde{\boldsymbol{r}}}{ds}$ to the angle of incidence and angle of reflection appearing the reflection law.

follows more if we use $\frac{\partial L}{\partial \boldsymbol{r}} = \boldsymbol{d}$ and see what happens if the sense of motion in a trajectory is changed: This change interchanges beginning and end and flips the signs of the directions, leading to $\frac{\partial L}{\partial \boldsymbol{r}_0} = -\boldsymbol{d}_0$ as desired.

Chapter 4

Stationary phase approximation

Stationary-phase approximations are an important technique that can be used approximate quantum mechanical quantities in terms of classical quantities. They allow to evaluate integrals in the limit where a parameter entering the integral becomes large, which can be helpful for example in the semiclassical limit. We will apply a stationary-approximation to sketch an alternative derivation of the van Vleck propagator, and later to perform a Laplace transform of the van Vleck propagator that will help us to extract information about the energy levels.

As a preparation we need some integrals.

Gauss integral. We have $\int_{-\infty}^{\infty} e^{-bx^2} dx = \sqrt{\frac{\pi}{b}}$ for $b > 0$.

Fresnel integral. We have

$$\int_{-\infty}^{\infty} e^{iax^2} = \sqrt{\frac{\pi}{|a|}} e^{i\frac{\pi}{4}\operatorname{sgn} a} \tag{4.1}$$

"Proof". If we formally set $b = -ia$ in the Gauss integral we get

$$\int_{-\infty}^{\infty} e^{iax^2} = \sqrt{\frac{\pi}{-ia}} = \sqrt{\frac{\pi}{-i|a|\operatorname{sgn} a}} = \sqrt{\frac{\pi}{|a|}} \sqrt{i \operatorname{sgn} a}.$$

If we now take, e.g., $a > 0$ and $\sqrt{i} = \sqrt{e^{i\frac{\pi}{2}}} = e^{i\frac{\pi}{4}}$ we obtain the desired result. However it is not clear whether we should be using the positive or the negative square root, and this approach ignores the condition $b > 0$ for the Gauss integral. A more careful derivation clarifying which square root should be considered is possible using integration over contours in the complex plane.

Aim. In the following our aim will be to evaluate integrals of the type

$$I = \int_a^b A(x) e^{iN\phi(x)} dx$$

in the limit $N \to \infty$. Here $\phi(x)$ is called the phase, and $A(x)$ is called the amplitude.

Idea. We will see that these integrals are dominated by the vicinity of stationary points of the phase, i.e., points with $\phi'(x) = 0$.

Why? I will just give an intuitive argument (referring to the Asymptotics course for details). First I will explain why most values of x do *not* give important contributions to the integral. If we change x slightly, this usually also changes $\phi(x)$ and in view of the large parameter $N \to \infty$ it leads to a large change of $N\phi(x)$. A large change of $N\phi(x)$ can change the sign of $e^{iN\phi(x)}$. Hence in the limit $N \to \infty$ the contributions of nearby x have a tendency to compensate each other due to opposite signs.

However this compensation is avoided if we are close to a stationary point. In the case $\phi'(x) = 0$ a change of x will to leading order leave $\phi(x)$ and thus $e^{iN\phi(x)}$ unchanged. Hence contributions from x close to a stationary point do not have the same tendency to cancel. The integral is thus dominated by the vicinity of the stationary points.

Special case. We will first consider the case that the phase has only one stationary point x_* inside the integration interval, and that this point occurs neither at the integration limits nor at a singularity of $A(x)$, nor at a point with $\phi''(x) = 0$.

The integral is dominated by the vicinity of x_*. For the relevant x we can thus write

$$x = x_* + \delta x,$$

where δx is small. We now perform a Taylor expansion up to quadratic order, leading to

$$\phi(x) = \phi(x_* + \delta x) \approx \phi(x_*) + \underbrace{\phi'(x_*)}_{=0} \delta x + \frac{1}{2}\phi''(x_*)\,(\delta x)^2 \ .$$

Here the linear term vanishes because we are expanding around a stationary point. We also Taylor expand $A(x)$. However, $A(x)$ does not involve the large parameter N, so we don't have to be as careful as for the phase factor. We thus take only the leading term in the Taylor expansion, replacing $A(x)$ by its value at the stationary point

$$A(x) \approx A(x_*) \ .$$

With these approximations the integral over the vicinity of the stationary point x_* becomes

$$I \approx \int_{\text{small } \delta x} A(x_*) e^{iN\phi(x_*)} e^{\frac{1}{2}iN\phi''(x_*)(\delta x)^2} d\,(\delta x) \ . \tag{4.2}$$

We can now enlarge the integration interval to the whole real axis as the contributions of δx further away from the stationary point δx will mostly cancel. We thus obtain

$$I \approx \int_{-\infty}^{\infty} A(x_*) e^{iN\phi(x_*)} e^{\frac{1}{2}iN\phi''(x_*)(\delta x)^2} d\,(\delta x) \ . \tag{4.3}$$

The first two factors in the integrand are constant and the integration over the third factor boils down to a Fresnel integral (4.1) with $a = \frac{1}{2}N\phi''(x_*)$. We thus obtain

$$I \approx A(x_*) e^{iN\phi(x_*)} \sqrt{\frac{\pi}{|\frac{1}{2}N\phi''(x_*)|}} e^{i\frac{\pi}{4}\operatorname{sgn}\phi''(x_*)}$$

or after rearrangement

$$I \approx \sqrt{\frac{2\pi}{N|\phi''(x_*)|}} A(x_*) e^{iN\phi(x_*) + i\frac{\pi}{4}\operatorname{sgn}\phi''(x_*)}. \tag{4.4}$$

Note. Here we excluded stationary points at the integration limits as only half of their vicinity is in the integration interval, i.e., their contribution would be halved. For stationary points at singularities (such as jumps and divergences) of $A(x)$ the approximation of $A(x)$ by $A(x_*)$ would not work. For stationary points with $\phi''(x_*) = 0$ the above approximation would diverge, and one would have to go beyond the quadratic approximation to obtain a finite result.

General case. We now consider $\phi(x)$ with several stationary points x_j. We again exclude the case of stationary points at the integration limits or singularities of $\phi(x)$. We now have to integrate over the vicinity of each of these points, each time obtaining a result as in Eq. (4.4). Summation over the stationary points gives

$$I \approx \sum_j \sqrt{\frac{2\pi}{N|\phi''(x_j)|}} A(x_j) e^{iN\phi(x_j) + i\frac{\pi}{4}\operatorname{sgn}\phi''(x_j)}. \tag{4.5}$$

Note. One might wonder whether in the presence of several stationary points it is still permissible to extend the integration limits from the vicinity of one stationary point to the real axis as done in the step from (4.2) to (4.3). However this extension happens after the quadratic expansion of the exponent; hence at the point where the extension is made $\delta x = 0$ is the only stationary point of the exponent. Therefore neither (4.2) nor (4.3) contain the contributions from other stationary points, and the extension is permissible.

Example. We consider the integral

$$I = \int_{-\infty}^{\infty} e^{iN\cosh x} dx \tag{4.6}$$

where $N \to \infty$. The phase is $\phi(x) = \cosh x$ and the amplitude is $A(x) = 1$. The stationary point x_* of the phase has to satisfy

$$0 = \phi'(x_*) = \sinh x_* \;\Rightarrow\; x_* = 0\,.$$

With

$$\begin{aligned} \phi(x_*) &= \cosh x_* = 1 \\ \phi''(x_*) &= \cosh x_* = 1 \end{aligned}$$

Eq. (4.4) thus boils down to

$$I \approx \sqrt{\frac{2\pi}{N}} e^{iN + i\frac{\pi}{4}}\,. \tag{4.7}$$

Higher dimensions. In $\mathbb{R}^n$ the Fresnel integral takes the form

$$\int_{\mathbb{R}^n} e^{i\boldsymbol{x}\cdot\mathcal{A}\boldsymbol{x}} d^n x = \frac{\pi^{n/2}}{\sqrt{|\det\mathcal{A}|}} e^{i\nu\frac{\pi}{4}}$$

where $\mathcal{A}$ is a symmetric matrix with nonzero eigenvalues, and ν is the difference between the number of positive eigenvalues and the number of negative eigenvalues. If we generalise the stationary-phase approximation to $\mathbb{R}^n$ we obtain

$$\int A(\boldsymbol{x}) e^{iN\phi(\boldsymbol{x})} d^n x \approx \sum_j \left(\frac{2\pi}{N}\right)^{n/2} \frac{1}{\sqrt{|\det H(\boldsymbol{x}^{(j)})|}} A(\boldsymbol{x}^{(j)}) e^{iN\phi(\boldsymbol{x}^{(j)}) + i\nu^{(j)}\frac{\pi}{4}}$$

where the sum is taken over the stationary points of $\phi(\boldsymbol{x})$ in the integration range, H is the Hessian matrix with entries $H_{kl} = \frac{\partial^2\phi}{\partial x_k \partial x_l}$ and $\nu^{(j)}$ is the difference between the numbers of positive and negative eigenvalues of the Hessian matrix at $\boldsymbol{x}^{(j)}$. The conditions are analogous to the one-dimensional case, and they involve $\det H(\boldsymbol{x}^{(j)}) \neq 0$.[1]

[1] We note that this means that the stationary points have to be isolated. For example the result does not hold if the stationary points form a line in a higher-dimensional space. One can show that the direction of such a line would have be an eigenvector of $H(\boldsymbol{x}^{(j)})$ with eigenvalue 0 and then the determinant would vanish.

Example: Derivation of the van Vleck propagator from Feynman's path integral. For general Lagrangian systems Feynman showed that $K(\boldsymbol{r}, \boldsymbol{r}_0, t)$ is given by an integral over all trajectories going in some way $\boldsymbol{r}_0$ to $\boldsymbol{r}$, i.e., an integral over all functions $\boldsymbol{r}(t')$ with $\boldsymbol{r}(0) = \boldsymbol{r}_0$ and $\boldsymbol{r}(t) = \boldsymbol{r}$. (Note that here the functions are used as integration variables!) These trajectories need not satisfy the classical laws of motion. The integrand involves a factor $e^{iS/\hbar}$ where S is the classical action of the trajectory, which in Lagrangian mechanics is defined as the integral over the Lagrangian $S = \int_0^t L(\boldsymbol{r}(t'), \dot{\boldsymbol{r}}(t'))dt'$. The propagator is thus given by an expression of the type

$$K(\boldsymbol{r}, \boldsymbol{r}_0, t) = \int \ldots e^{iS/\hbar}$$

In the limit $\hbar \to 0$ one can now perform a stationary phase approximation, and this approximation leads to a sum over trajectories for which the action becomes stationary. As seen in Mechanics 2/23 trajectories with stationary action are just the physical trajectories, satisfying Lagrange's equation of motion. If one works out the second derivatives of the action one gets an expression of the type

$$K(\boldsymbol{r}, \boldsymbol{r}_0, t) \approx \sum_\gamma \frac{1}{\sqrt{|M_{\gamma,12}|}} e^{iS_\gamma/\hbar - i\mu_\gamma \frac{\pi}{2}} \ldots$$

where the sum goes over all physical trajectories from $\boldsymbol{r}_0$ to $\boldsymbol{r}$. This formula generalises the van Vleck propagator derived for dispersing billiards: In particular the stability matrix element $M_{\gamma,12}$ contributes in the same way as above; the term $iS_\gamma/\hbar$ is analogous to $\frac{iL_\gamma^2}{\epsilon}$; and the phase factor $e^{-i\mu_\gamma \frac{\pi}{2}}$ generalises the sign $(-1)^{n_\gamma}$.

I am omitting a lot of details (and some further factors) here. The main mathematical challenge is to properly define an integral where the integration variable itself is a function, and then to generalise the stationary phase approximation to such integrals. However it has hopefully become clear that path integrals and stationary phase approximations provide a conceptually elegant way to perform semiclassical approximations for general chaotic systems.

Chapter 5

The trace formula

In this chapter we will discuss the so-called trace formula, which relates the energy levels of a chaotic systems (e.g. dispersing billiards) to classical trajectories.

5.1 Level density

In the trace formula, the energy levels are represented through the level density $d(E)$. This is a function of the energy E with delta functions at values of E that coincide with energy levels.

Definition. The **level density** of a quantum system with energy levels E_j is given by

$$d(E) = \sum_j \delta(E - E_j) \, .$$

As we have already expressed the propagator in terms of classical trajectories we will try to obtain a similar result for the level density by relating it to the propagator. To do so is convenient to first express the propagator in terms of energy levels and energy eigenfunctions.

Theorem. For a system with discrete energy levels E_j and energy eigenfunctions $\psi_j(\boldsymbol{r})$ the propagator can be written as

$$K(\boldsymbol{r}, \boldsymbol{r}_0, t) = \sum_j \psi_j(\boldsymbol{r}) e^{-iE_j t/\hbar} \psi_j^*(\boldsymbol{r}_0). \tag{5.1}$$

Proof. (Also considered in a problem.) The claimed formula satisfies all the properties required for the propagator.

- It satisfies the Schrödinger equation as

$$\begin{aligned} \hat{H} K(\boldsymbol{r}, \boldsymbol{r}_0, t) &= \sum_j E_j \psi_j(\boldsymbol{r}) e^{-iE_j t/\hbar} \psi_j^*(\boldsymbol{r}_0) \\ &= \left(-\frac{i}{\hbar}\right)^{-1} \frac{\partial}{\partial t} \sum_j \psi_j(\boldsymbol{r}) e^{-iE_j t/\hbar} \psi_j^*(\boldsymbol{r}_0) \\ &= i\hbar \frac{\partial}{\partial t} K(\boldsymbol{r}, \boldsymbol{r}_0, t) \end{aligned}$$

- The initial condition is satisfied due to[1]

$$K(\boldsymbol{r}, \boldsymbol{r}_0, 0) = \sum_j \psi_j(\boldsymbol{r})\psi_j^*(\boldsymbol{r}_0) = \delta(\boldsymbol{r} - \boldsymbol{r}_0) \ .$$

- For the boundary condition we just consider the case of a system with Dirichlet boundary conditions for all $\boldsymbol{r} \in \partial B$. In such a system all energy eigenfunctions must satisfy $\psi_j(\boldsymbol{r}) = 0$ for $\boldsymbol{r} \in \partial B$. Then Eq. (5.1) implies $K(\boldsymbol{r}, \boldsymbol{r}_0, t) = 0$ for $\boldsymbol{r} \in \partial B$.

Note. The propagator is also defined for systems whose energy levels are not discrete, such as the system without potential or boundaries and the system with Dirichlet boundary conditions at $x = 0$. Then Eq. (5.1) doesn't hold in the form given here.

Green's function. To access the level density from the propagator we first make a Laplace transformation to switch from the time domain to the energy domain. The Laplace transform of the propagator is called the Green function[2]. It is given by

$$G(\boldsymbol{r}, \boldsymbol{r}_0, \underbrace{E + i\eta}_{=E^+}) = -\frac{i}{\hbar}\int_0^\infty e^{i(E+i\eta)t/\hbar} K(\boldsymbol{r}, \boldsymbol{r}_0, t)dt \ . \tag{5.2}$$

Here the energy argument is taken with an imaginary part $\eta > 0$ to make the integral convergent. We will later take η to zero.

If we evaluate the Laplace transform we get

$$G(\boldsymbol{r}, \boldsymbol{r}_0, E + i\eta) = \sum_j \psi_j(\boldsymbol{r})\psi_j^*(\boldsymbol{r}_0) \left(-\frac{i}{\hbar}\int_0^\infty e^{i(E-E_j+i\eta)t/\hbar} dt \right)$$

where

$$-\frac{i}{\hbar}\int_0^\infty e^{i(E-E_j+i\eta)t/\hbar} dt = -\frac{i}{\hbar}\left(\frac{i(E - E_j + i\eta)}{\hbar} \right)^{-1} e^{i(E-E_j+i\eta)t/\hbar}\Big|_0^\infty = (E - E_j + i\eta)^{-1} \ .$$

Hence the Green function can be written in terms of energy levels and energy eigenfunctions as

$$G(\boldsymbol{r}, \boldsymbol{r}_0, E + i\eta) = \sum_j \frac{\psi_j(\boldsymbol{r})\psi_j^*(\boldsymbol{r}_0)}{E - E_j + i\eta} \ .$$

The Green function allows to access the level density as follows:

Theorem.

$$d(E) = -\frac{1}{\pi} \lim_{\eta \to 0} \mathrm{Im} \int G(\boldsymbol{r}, \boldsymbol{r}, E + i\eta) d^n r \ . \tag{5.3}$$

[1] The identity between the second and third term will be derived in an exercise.

[2] In a different convention, the propagator is called the time dependent Green function and the Green function as called the energy dependent Green function.

Proof. We have

$$\begin{aligned}
& -\frac{1}{\pi}\mathrm{Im}\int G(\boldsymbol{r},\boldsymbol{r},E+i\eta)d^n r \\
= & -\frac{1}{\pi}\mathrm{Im}\sum_j \frac{\int \psi_j(\boldsymbol{r})\psi_j^*(\boldsymbol{r})d^n r}{E+i\eta-E_j} \\
= & -\frac{1}{\pi}\mathrm{Im}\sum_j \frac{1}{E+i\eta-E_j} \\
= & -\frac{1}{\pi}\mathrm{Im}\sum_j \frac{E-i\eta-E_j}{(E-E_j)^2+\eta^2} \\
= & \frac{1}{\pi}\sum_j \frac{\eta}{(E-E_j)^2+\eta^2} \\
\to & \sum_j \delta(E-E_j) = d(E) \quad \text{for} \quad \eta\to 0
\end{aligned}$$

Here we used that the energy eigenfunctions are normalised, we multiplied with the complex conjugate of the denominator $E+i\eta-E_n$, and we used that the Lorentz function $\frac{1}{\pi}\frac{\epsilon}{x^2+\eta^2}$ converges to $\delta(x)$ in the limit $\eta\to 0$. (See Eq. (3.3).)

Note. Inserting identical arguments for $\boldsymbol{r}$ and $\boldsymbol{r}_0$ and then integrating over them is the continuous analogue of taking a trace of a matrix. (As in a trace we use identical indices and sum over them.) This is the reason why the formula we will get in the end is called the trace formula.

5.2 Application to billiards

We now apply the procedure outlined above to dispersing billiards. We start with the van Vleck propagator

$$K_V(\boldsymbol{r},\boldsymbol{r}_0,t) = \sum_\gamma \frac{1}{i\epsilon\pi}\sqrt{\frac{L_\gamma}{|M_{\gamma,12}|}}(-1)^{n_\gamma}e^{\frac{i}{\epsilon}L_\gamma^2}\,.$$

where $\epsilon=\frac{2\hbar t}{m}$. Then the Green function can be approximated as a sum over contributions of trajectories γ going from $\boldsymbol{r}_0$ to $\boldsymbol{r}$. The summands are obtained Laplace transformation of the summands K_γ in the van Vleck propagator. We get

$$\begin{aligned}
G_\gamma(\boldsymbol{r},\boldsymbol{r}_0,E^+) &= -\frac{i}{\hbar}\int_0^\infty e^{iE^+t/\hbar}K_\gamma(\boldsymbol{r},\boldsymbol{r}_0,t) \qquad (5.4)\\
&= \underbrace{-\frac{i}{\hbar}\frac{1}{i\pi}\frac{m}{2\hbar}}_{=-\frac{m}{2\pi\hbar^2}}\sqrt{\frac{L_\gamma}{|M_{\gamma,12}|}}(-1)^{n_\gamma}\underbrace{\int_0^\infty \frac{1}{t}\exp\left(\frac{i}{\hbar}\left(E^+t+\frac{\frac{m}{2}L_\gamma^2}{t}\right)\right)dt}_{=I}
\end{aligned}$$

where $E^+=E+i\eta$ denotes the energy with an imaginary part η that we will take to zero. We now want to write the integral in a more symmetric form of the type studied in Eq. (4.6),

$$\frac{i}{\hbar}\left(Et+\frac{\frac{m}{2}L_\gamma^2}{t}\right) = iN\cosh x = \frac{iN}{2}(e^x+e^{-x})\,.$$

Here the term proportional to t is supposed to give rise to $\frac{N}{2}e^x$ and the term proportional to $\frac{1}{t}$ should give rise to $\frac{N}{2}e^{-x}$. To do this we choose $\frac{N}{2}$ as the square root of the product of the coefficients in front of t and $\frac{1}{t}$,

$$\frac{N}{2} = \frac{1}{\hbar}\sqrt{E\frac{m}{2}L_\gamma^2}\,.$$

Similarly e^x has to be the square root of the ratio of the two coefficients,

$$e^x = t\sqrt{\frac{E}{\frac{m}{2}L_\gamma^2}} \; .$$

This implies that

$$\frac{dt}{t} = d\ln t = dx \; .$$

The coefficient N assumes a nicer form if we write the energy as

$$E = \frac{p^2}{2m} = \frac{\hbar^2 k^2}{2m}$$

where p is the momentum and k is the wavenumber (related to the wavelength by $k = \frac{2\pi}{\lambda}$). We then obtain

$$\frac{N}{2} = \frac{1}{\hbar}\sqrt{\frac{\hbar^2 k^2}{2m}\frac{m}{2}L_\gamma^2} = \frac{kL_\gamma}{2}$$

and the integral becomes

$$I = \int_{-\infty}^{\infty} \exp(ikL_\gamma \cosh x)dx \; .$$

We now have to evaluate the integral and then determine the resulting contributions to the level density via Eq. (5.3). When doing so we have to distinguish between trajectories that go directly from $\boldsymbol{r}_0$ to $\boldsymbol{r}$ and those that involve reflections.

5.3 Trajectories without reflections

Green's function. For trajectories going directly from $\boldsymbol{r}_0$ to $\boldsymbol{r}$ Eq. (5.4) strongly simplifies. We have $M_{\gamma,12} = L_\gamma$ and hence $\sqrt{\frac{L_\gamma}{|M_{\gamma,12}|}} = 1$. We also have $n_\gamma = 0$ and thus $(-1)^{n_\gamma} = 1$. To access the level density we need only the limiting case where $\boldsymbol{r}_0$ coincides with $\boldsymbol{r}$ and η is taken to zero. If we denote the contribution from trajectories without reflection by a subscript 0 we obtain

$$\begin{aligned} \mathrm{Im}\, G_0(\boldsymbol{r},\boldsymbol{r},E) &= -\frac{m}{2\pi\hbar^2} \underbrace{\lim_{L\to 0} \mathrm{Im} \int_{-\infty}^{\infty} \exp(ikL\cosh x)dx}_{=\pi} \\ &= -\frac{m}{2\hbar^2} \; . \end{aligned}$$

Here I am just quoting the $L \to 0$ limit which is nontrivial to show.[3]

Level density. Due to Eq. (5.3) the contribution of trajectories without reflections to the level density is

$$\begin{aligned} d_0(E) &= -\frac{1}{\pi}\mathrm{Im} \int G_0(\boldsymbol{r},\boldsymbol{r},E)d^2r \\ &= \frac{m}{2\pi\hbar^2}\int d^2r \; . \end{aligned}$$

The integral $\int d^2r$ simply gives the area $A = |B|$ of the billiard. We thus get

[3] The integrand is related to a Hankel function $H_0^{(1)}(y) = \frac{1}{\pi i}\int_{-\infty}^{\infty} \exp(iy\cosh x)dx$. In the limit $y \to 0$ the real part of the Hankel function turns to 1, meaning that the imaginary part of $\int_{-\infty}^{\infty} \exp(iy\cosh x)dx$ turns to π.

$$d_0(E) = \frac{mA}{2\pi\hbar^2}.$$

This term is called the **Weyl term**. We will later see that all other contributions to the level density oscillate around zero as the energy is changed. Hence the Weyl term can be interpreted as the **average level density**.

5.4 Trajectories with reflections

We can access more information about the level density than just its average. This additional information comes from trajectories that contain reflections.

Green's function. The contribution of trajectories to the Green's function is given by (5.4) where we have later rewritten the integral I as

$$I = \int_{-\infty}^{\infty} \exp(ikL_\gamma \cosh x) dx.$$

This integral looks just like the example from Eq. (4.6) with the replacement $N = kL_\gamma$. Hence we might be tempted to evaluate it in the limit where k and thus kL_γ goes to infinity. Physically, taking this limit makes a lot of sense as for $k \to \infty$ the wavelengths $\lambda = \frac{2\pi}{k}$ become much smaller than the size of the system so in the light of our previous discussion we would expect that quantum mechanical properties can be approximated well in terms of classical ones. The limit $k \to \infty$ also means that the energy $E = \frac{\hbar^2 k^2}{2m}$ becomes large. A stationary phase approximation in this limit wouldn't have worked in the case of trajectories without reflections as k is multiplied with the length of the trajectory which was zero in the previous case. For the trajectories with reflections we can simply use the stationary phase approximation of I already given in (4.7) with the identification $N = kL_\gamma$. This leads to

$$I \approx \sqrt{\frac{2\pi}{kL_\gamma}} e^{ikL_\gamma + i\frac{\pi}{4}}$$

and the Green function (5.4) turns into

$$G_\gamma(\boldsymbol{r}, \boldsymbol{r}_0, E) = \underbrace{-\frac{m}{\hbar^2} \frac{1}{\sqrt{2\pi k |M_{\gamma,12}|}} (-1)^{n_\gamma} e^{i\frac{\pi}{4}}}_{=B_\gamma} e^{ikL_\gamma} \,. \tag{5.5}$$

Level density. The level density is obtained from the Green function as in Eq. (5.3). We have to equate the arguments $\boldsymbol{r}$ and $\boldsymbol{r}_0$ and integrate over them, let the η go to zero, and apply $-\frac{1}{\pi}\mathrm{Im}$. If we do this for the sum over all G_γ where γ involves reflections we obtain the term

$$d_{\mathrm{osc}}(E) = -\frac{1}{\pi}\mathrm{Im} \int \sum_{\gamma \text{ w. refl.}} G_\gamma(\boldsymbol{r}, \boldsymbol{r}, E) d^2 r$$

which has to be added to $d_0(E)$ to give an approximation of the level density. The index "osc" was chosen as the result will turn out to give oscillations of the level density around its average. The limit $\eta \to 0$ was taken by dropping η from the formula. The sum goes over all trajectories that start at $\boldsymbol{r}$, undergo some reflections, and then return to $\boldsymbol{r}$ (see Fig. 5.1). Hence the trajectories are **closed**. However the initial directions $\boldsymbol{d}_0$ and the final directions $\boldsymbol{d}$ may still differ, i.e, the trajectories may still have an edge at $\boldsymbol{r}$. If we insert Eq. (5.5) for the Green function we get

$$d_{\mathrm{osc}}(E) = \mathrm{Im} \int \sum_{\gamma \text{ w. refl.}} \underbrace{\left(-\frac{B_\gamma}{\pi}\right)}_{=\tilde{B}_\gamma} e^{ikL_\gamma} d^2 r. \tag{5.6}$$

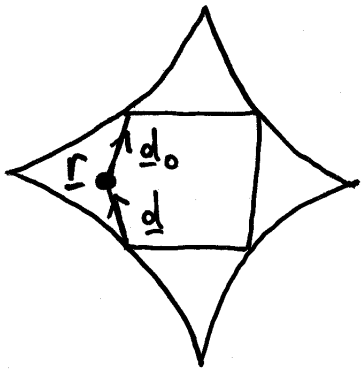


Figure 5.1: A trajectory starting and ending at position $\boldsymbol{r}$.

Stationary-phase approximation for the r-integral. Let us evaluate the integral over $\boldsymbol{r}$. Again we use a stationary-phase approximation for large k. We thus have to find out for which $\boldsymbol{r}$ the length L becomes stationary meaning that $\frac{\partial L}{\partial \boldsymbol{r}} = 0$. Recall that if we write the length as a function of the initial point $\boldsymbol{r}_0$ and the final point $\boldsymbol{r}$ we have (see section 3.5)

$$\frac{\partial}{\partial \boldsymbol{r}} L(\boldsymbol{r}, \boldsymbol{r}_0) = \boldsymbol{d}$$

where $\boldsymbol{d}$ is the final direction of the trajectory. We have also seen that the derivative w.r.t. $\boldsymbol{r}_0$ is

$$\frac{\partial}{\partial \boldsymbol{r}_0} L(\boldsymbol{r}, \boldsymbol{r}_0) = -\boldsymbol{d}_0$$

where $\boldsymbol{d}_0$ is the direction of the trajectory in the beginning. However now the initial and the final point are the same and we have to take the derivative of $L(\boldsymbol{r}, \boldsymbol{r})$. According to the chain rule this means that we have to take the derivative w.r.t. both occurrences of $\boldsymbol{r}$ and sum over them, leading to

$$\frac{\partial}{\partial \boldsymbol{r}} L(\boldsymbol{r}, \boldsymbol{r}) = \boldsymbol{d} - \boldsymbol{d}_0 \ .$$

The length thus becomes stationary for trajectories with $\boldsymbol{d}_0 = \boldsymbol{d}$. In other words the trajectory does not only come back to the same point in space, but also with the same direction and thus no edge. Such trajectories are called **periodic orbits**.

Hence the integral in (5.6) goes over closed trajectories but in the sense of a stationary phase approximation it is dominated by closed trajectories that are similar to periodic orbits. These are closed trajectories that still have slight edges. We will parameterise these trajectories as follows:

- By the periodic orbit p they are close to.
- By the position of the edge along the periodic orbit. We indicate the position of the edge along the orbit by a variable $r_\parallel$. Normally the variable $r_\parallel$ should run from 0 to the length of the orbit L_p. However, some periodic orbits (see Fig. 5.2) just involve multiple traversals of a shorter orbit. In this cases the possible positions for edges can already by reached if we let $r_\parallel$ run between 0 and the length at the shorter orbit, which can be obtained by dividing the length of p by the number of repetitions. Thus in general $r_\parallel$ runs from zero up to the primitive length L_p^{prim} defined by

$$L_p^{\text{prim}} = \begin{cases} L_p & \text{if } p \text{ does not follow multiple reptitions of a shorter orbit} \\ \frac{L_p}{r} & \text{if } p \text{ follows } r \text{ repetitions of a shorter orbit} \end{cases} \tag{5.7}$$

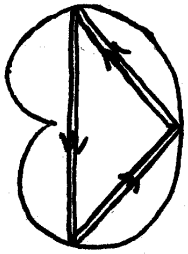


Figure 5.2: A periodic orbit following two repetitions of a shorter periodic orbit.

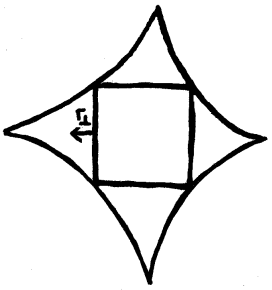


Figure 5.3: Perpendicular deviation from a periodic orbit.

- Finally we need a coordinate $r_\perp$ which indicates how far the edge is moved away from the orbit. See Fig. 5.3.

If we use these parameters Eq. (5.6) turns into

$$d_{\rm osc}(E) = \mathrm{Im} \sum_p \int_0^{L_p^{\rm prim}} dr_\parallel \int dr_\perp \tilde{B}_\gamma e^{ikL_\gamma} \tag{5.8}$$

where the properties in the integrand refer to the closed trajectory γ and thus depend on p, $r_\parallel$, and $r_\perp$. They are equal to the corresponding values for the periodic orbit if $r_\perp = 0$.

Integration over $r_\perp$. We first do the integral over $r_\perp$. We use a stationary phase approximation

$$\int_a^b A(x) e^{iN\phi(x)} dx \approx \sqrt{\frac{2\pi}{N|\phi''(x_*)|}} A(x_*) e^{iN\phi(x_*) + i\frac{\pi}{4}\,\mathrm{sgn}\,\phi''(x_*)}. \tag{5.9}$$

In (5.8) the role of the amplitude $A(x)$ is played by $\tilde{B}_\gamma$, the role of the phase $\phi(x)$ is played by the length L_γ seen as a function of $r_\perp$, and the role of the large parameter N is played by k. At the stationary point $r_\perp = 0$ the trajectory γ is identical to the periodic orbit p so replacing the indices by p is equivalent to inserting x_* in (5.9). (The only nontrivial consequence of this is that the stability matrix entering $\tilde{B}_\gamma$ is evaluated on the periodic orbit. The other terms in $\tilde{B}_\gamma$ are constants, and the number of reflections which also does not change when varying $r_\perp$.)

The stationary phase approximation thus gives

$$\int dr_\perp \tilde{B}_\gamma e^{ikL_\gamma} \approx \sqrt{\frac{2\pi}{k|\frac{\partial^2 L}{\partial r_\perp^2}|}} \tilde{B}_p e^{ikL_p + i\frac{\pi}{4}\,\mathrm{sgn}\,\frac{\partial^2 L}{\partial r_\perp^2}}. \tag{5.10}$$

where the derivative $\frac{\partial^2 L}{\partial r_\perp^2}$ is evaluated for $r_\perp = 0$ which is the value of $r_\perp$ corresponding to the periodic orbit.

Final result. Now we use the following additional ingredients:

- We need the second derivative $\frac{\partial^2 L}{\partial r_\perp^2}$. As before the length L can formally be written as a function $L(\boldsymbol{r}, \boldsymbol{r}_0)$ of the initial point $\boldsymbol{r}_0$ and the final point $\boldsymbol{r}$. As we are interested in trajectories whose initial and final point coincide we only have to consider $L(\boldsymbol{r}, \boldsymbol{r})$. When computing the first derivative we have to sum the derivatives w.r.t. the two occurrences of $\boldsymbol{r}$, leading to

$$\frac{\partial}{\partial r_\perp} L(\boldsymbol{r}, \boldsymbol{r}) = \left[\frac{\partial}{\partial r_\perp} L(\boldsymbol{r}, \boldsymbol{r}_0) + \frac{\partial}{\partial r_{\perp 0}} L(\boldsymbol{r}, \boldsymbol{r}_0) \right]_{\boldsymbol{r}_0 = \boldsymbol{r}}.$$

  For the second derivative we again take into account both arguments of L. We then obtain one term where both derivatives act on the first argument, two identical mixed terms where one derivative acts on the first argument and the other one on the second argument, and one term where both derivatives act on the second argument. We thus have

$$\frac{\partial^2}{\partial r_\perp^2} L(\boldsymbol{r}, \boldsymbol{r}) = \left[\frac{\partial^2}{\partial r_\perp^2} L(\boldsymbol{r}, \boldsymbol{r}_0) + 2 \frac{\partial^2}{\partial r_\perp \partial r_{\perp 0}} L(\boldsymbol{r}, \boldsymbol{r}_0) + \frac{\partial^2}{\partial r_{\perp 0}^2} L(\boldsymbol{r}, \boldsymbol{r}_0) \right]_{\boldsymbol{r}_0 = \boldsymbol{r}}. \tag{5.11}$$

  Now one can show that for any trajectory the following holds (see appendix **??**)

$$\frac{\partial^2 L}{\partial r_\perp^2}(\boldsymbol{r}, \boldsymbol{r}_0) = \frac{M_{22}}{M_{12}}, \qquad \frac{\partial^2 L}{\partial r_\perp \partial r_{\perp 0}}(\boldsymbol{r}, \boldsymbol{r}_0) = -\frac{1}{M_{12}}, \qquad \frac{\partial^2 L}{\partial r_{\perp 0}^2}(\boldsymbol{r}, \boldsymbol{r}_0) = \frac{M_{11}}{M_{12}}.$$

  Here M is the stability matrix of the orbit under consideration. This implies

$$\frac{\partial^2}{\partial r_\perp^2} L(\boldsymbol{r}, \boldsymbol{r}) = \frac{\mathrm{tr} M_p - 2}{M_{p,12}} \tag{5.12}$$

  After inserting this result the $M_{p,12}$ in (5.10) will be cancelled. This result also implies that the second derivative is positive: We had seen in chapter 1 that all elements of the stability matrix in a dispersing billiard have the same sign. Moreover we have $|\mathrm{tr}\, M| > 2$ and thus the sign of the numerator $\mathrm{tr}\, M - 2$ is determined by the sign of $\mathrm{tr}\, M$. Hence the signs of the numerator and the denominator coincide and their quotient is positive.

- Moreover (as will be investigated in exercises) the trace of the stability matrix of a periodic orbit does not depend on which of the points along point along this orbit is taken as the the start and end point, i.e. on how we choose $r_\parallel$. This means that the $r_\parallel$ integral just goes over a constant and leads to multiplication with a factor L_p^{prim}.

Altogether we thus get

$$\begin{aligned} d_{\mathrm{osc}}(E) &\approx \sum_p \sqrt{\frac{2\pi}{k \left| \frac{\mathrm{tr} M_p - 2}{M_{p,12}} \right|}} \tilde{B}_p e^{ikL_p + i\frac{\pi}{4}} \\ &= \mathrm{Im} \frac{m}{\pi \hbar^2 k} \sum_p \frac{L_p^{\mathrm{prim}} (-1)^{n_p}}{\sqrt{|\mathrm{tr}\, M_p - 2|}} e^{ikL_p + i\frac{\pi}{2}} \end{aligned}$$

where the prefactors from (5.10) were combined with those from $\tilde{B}_p$. Due to $\mathrm{Im} e^{ikL_p + i\frac{\pi}{2}} = \mathrm{Im}\, i e^{ikL_p} = \cos(kL_p)$ this result simplifies to

$$d_{\mathrm{osc}}(E) \approx \frac{m}{\pi \hbar^2 k} \sum_p \frac{L_p^{\mathrm{prim}} (-1)^{n_p}}{\sqrt{|\mathrm{tr}\, M_p - 2|}} \cos(kL_p). \tag{5.13}$$

Eq. (5.13) is called the **trace formula**. It was derived by by Balian and Bloch for chaotic billiards, and by Gutzwiller in its form for general chaotic systems.

The overall level density is approximated (for large k and thus large E) by the sum of contributions from trajectories with and without reflections,

$$d(E) \approx d_0(E) + d_{\text{osc}}(E).$$

Here the first term gives the average level density, and the second one gives fluctuations about that average. When evaluating $d_{\text{osc}}(E)$ numerically one can't include all the (infinitely many) periodic orbits of the system. But the more orbits one includes the closer one gets from the average level density to the proper level density displaying peaks at the energy levels.

5.5 Alternative notation

In many references the trace formula is given in a different notation. This notation has the advantage of being generalisable to chaotic systems that are not billiards. It involves the following quantities:

- The (reduced) **action** S. For billiards this is just the length multiplied with the absolute value of the momentum
$$S = pL = \hbar k L = \sqrt{2mE}L \; . \tag{5.14}$$
Here we used that the absolute value of the momentum is equal to $\hbar$ times the wavenumber. For billiards we also have $E = \frac{p^2}{2m}$ (as there is only the kinetic energy), implying that $p = \sqrt{2mE}$. With the action the kL_p in the cosine of (5.13) can be written as $\frac{S_p}{\hbar}$.[4]
- The **period** T. The period of a periodic orbit is the time it takes for the particle to come back to the initial point. It is related to the length L via the velocity v. We have
$$v = \frac{L}{T} \;\; \Rightarrow \;\; T = \frac{L}{v} = \frac{mL}{p} = \frac{mL}{\hbar k} = \frac{mL}{\sqrt{2mE}} \; ,$$
where we used $p = mv$ as well as the relations above. Together with a $\frac{m}{\hbar k}$ from the prefactor in (5.13) the primitive length L_p^{prim} can be grouped into the primitive period T_p^{prim}.
- Instead of the number of reflections n one can use the **Maslov index** $\mu = 2n$. As the cosine flips its sign when the argument is changed by π we can then write $(-1)^n \cos\frac{S}{\hbar} = \cos\left(\frac{S}{\hbar} - n\pi\right) = \cos\left(\frac{S}{\hbar} - \mu\frac{\pi}{2}\right)$. For general chaotic systems also odd values of μ are possible.

It is also helpful to note that the derivative of the reduced action w.r.t. the energy is the period:

$$\frac{dS}{dE} = \frac{d}{dE}\left(\sqrt{2mE}L\right) = \frac{mL}{\sqrt{2mE}} = T \tag{5.15}$$

With these definitions (5.13) can be rewritten as

$$d(E) \approx \underbrace{\frac{mA}{2\pi\hbar^2}}_{=d_0=\bar{d}} + \underbrace{\frac{1}{\pi\hbar}\sum_p \underbrace{\frac{T_p^{\text{prim}}}{\sqrt{|\text{tr}\, M_p - 2|}}}_{=A_p(E)} \cos\left(\frac{S_p}{\hbar} - \mu_p\frac{\pi}{2}\right)}_{=d_{\text{osc}}(E)} \; . \tag{5.16}$$

[4] For systems that are not billiards the absolute value of the momentum can change along the orbit, and the reduced action is defined as its integral over the orbit. The reduced action is closely related to but different from the action introduced in Mechanics 2.

The factor $A_\gamma(E)$ involving the primitive period and the stability matrix is called the stability amplitude.

Chapter 6

Spectral statistics

If the classical behaviour of a system is chaotic, this has important consequences for statistics of its quantum mechanical energy levels. In particular these energy levels don't 'like' to come close each other, they tend to 'repel'. To illustrate this, Fig. 6.1 displays the energy spectra of three chaotic systems and one integrable system. In the picture arrows indicate that two or more energy levels are so close to each other that they can't be resolved in the plot. We see that this happens much less frequently for the chaotic systems where the levels rarely come close.

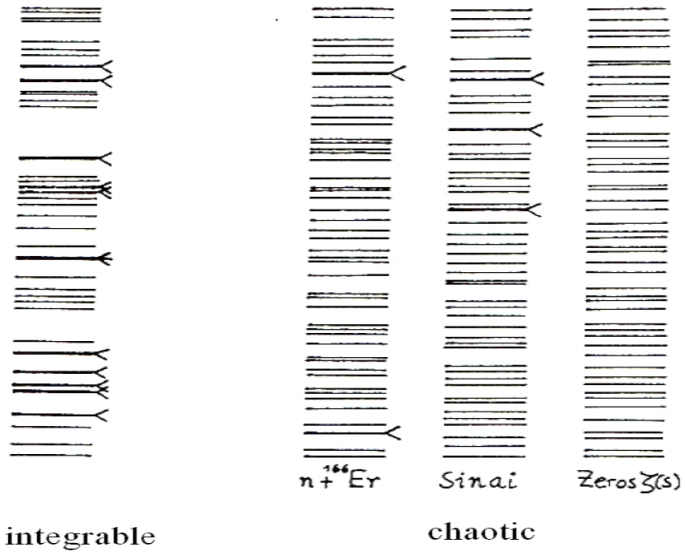


Figure 6.1: Spectra of one integrable system and three chaotic systems. Arrows indicate energy levels that can't be distinguished in the plot.

In this section we will introduce a quantity that characterises the statistics of energy levels, the spectral form factor. We will then evaluate the spectral form factor using the trace formula within the so called diagonal approximation. Afterwards I will give a brief overview over connections of these results to random matrix theory and some recent research.

6.1 Spectral form factor

To characterise the statistics of the energy levels we could use the following integral over a product of level densities with slightly different arguments.

$$\int_{-\infty}^{\infty} d\left(E+\frac{y}{2}\right) d\left(E-\frac{y}{2}\right) dE\,. \tag{6.1}$$

To understand what this integral expresses we just insert the definition of the level density in terms of delta functions and simplify. We get

$$\begin{aligned}\int_{-\infty}^{\infty} d\left(E+\frac{y}{2}\right)\left(E-\frac{y}{2}\right) dE &= \int_{-\infty}^{\infty} \sum_j \delta\left(E+\frac{y}{2}-E_j\right) \sum_k \delta\left(E-\frac{y}{2}-E_k\right) dE \\ &= \sum_{j,k} \delta\left(E_k+\frac{y}{2}+\frac{y}{2}-E_j\right) \\ &= \sum_{j,k} \delta\left(y-(E_j-E_k)\right)\,.\end{aligned}$$

Here we have used that the integral over a product involving a delta function gives back the rest of the integrand evaluated at the position where the argument of the delta function is zero. We applied this rule and inserted the zero of the second delta function, $E = E_k + \frac{y}{2}$, into the rest of the integrand which happens to involve a delta function as well. We see that the final result is a sum over delta peaks for all possible differences between energy levels $E_j - E_k$. Hence our integral measures the likelihood of energy differences in the system.

In the following we will not use (6.1) directly but improve the definition in several aspects. Most of these are matters of taste but the fourth point (averaging) is crucial.

- We want to measure the energy differences by dimensionless numbers. To find a suitable energy unit to achieve this we note that an energy interval of width ΔE contains $\int_E^{E+\Delta E} d(E')dE'$ levels. This number can be be approximated by $\bar{d}\Delta E$; here $\bar{d} = d_0$ is the mean level density. Hence an interval of size $\frac{1}{\bar{d}}$ contains on average one level. So this seems to be a useful energy unit, and we will thus write the energy difference as $y = \frac{\epsilon}{\bar{d}}$ where ϵ is dimensionless. (Note that this is different from the $\epsilon \propto \hbar t$ used earlier.)

- To obtain a dimensionless quantity we divide (6.1) by $\bar{d}^2$.

- The remaining calculation becomes easier if we consider take a Fourier transform $\int_{-\infty}^{\infty} d\epsilon\, e^{2\pi i\epsilon\tau} \ldots$. Our quantity then becomes a function of the dimensionless time variable τ.

- Importantly (6.1) is still a sum over delta functions. Hence it is very singular and system specific. To get something more useful we should average. As we have done a Fourier transformation we are taking this additional average over a small interval of length $\Delta\tau \ll 1$ for the variable τ. By convention we also replace the energy integral from (6.1) by an average over the energy range from 0 to $E_{\max}$ where $E_{\max}$ is taken to infinity. To abbreviate these averages we define

$$\langle \ldots \rangle \equiv \lim_{E_{\max}\to\infty} \frac{1}{E_{\max}} \int_0^{E_{\max}} dE \frac{1}{\Delta\tau} \int_{\tau-\frac{\Delta\tau}{2}}^{\tau+\frac{\Delta\tau}{2}} d\tau \ldots \tag{6.2}$$

  To keep the notation simple we here use the same symbol τ both for the integration variable and for the centre of the averaging interval.

- Finally, we only consider the oscillatory part of the level density $d_{\rm osc}(E)$. To see that this does

not change much we insert $d(E) = \bar{d} + d_{\text{osc}}(E)$ into the averaged product of level densities:

$$\begin{aligned} & \left\langle d\left(E+\frac{y}{2}\right) d\left(E-\frac{y}{2}\right)\right\rangle \\ \approx\ & \langle \bar{d}^2\rangle + \underbrace{\left\langle \bar{d}\, d_{\text{osc}}\left(E-\frac{y}{2}\right)\right\rangle}_{\approx 0} + \underbrace{\left\langle d_{\text{osc}}\left(E+\frac{y}{2}\right)\bar{d}\right\rangle}_{\approx 0} + \left\langle d_{\text{osc}}\left(E+\frac{y}{2}\right) d_{\text{osc}}\left(E-\frac{y}{2}\right)\right\rangle . \end{aligned}$$

Here we have used that $d_{\text{osc}}(E)$ oscillates around zero and hence becomes neglible after averaging. Hence using the product of the oscillatory parts instead of the product of the full level densities only changes the result by a constant.

Definition. We thus define the **spectral form factor** as

$$K(\tau) = \frac{1}{\bar{d}^2}\left\langle \int_{-\infty}^{\infty} d_{\text{osc}}\left(E+\frac{\epsilon}{2\bar{d}}\right) d_{\text{osc}}\left(E-\frac{\epsilon}{2\bar{d}}\right) e^{2\pi i\epsilon\tau} d\epsilon \right\rangle . \tag{6.3}$$

In the following we will consider only $\tau > 1$, but the spectral form factor for negative τ can be determined in an analogous way.

Note. If in our definition we had just dropped the Fourier transform we would have obtained the (two point) **correlation function** of the level density.

6.2 Diagonal approximation

We will now use the trace formula to evaluate $K(\tau)$. For the two oscillatory level densities in (6.3) we write

$$d_{\text{osc}}(E \pm \tfrac{\epsilon}{2\bar{d}}) = \frac{1}{\pi\hbar}\sum_p A_p(E \pm \tfrac{\epsilon}{2\bar{d}}) \cos\left(\frac{S_p(E \pm \frac{\epsilon}{2\bar{d}})}{\hbar} - \mu_p\frac{\pi}{2}\right) . \tag{6.4}$$

For small energy differences we approximate $A_p(E \pm \frac{\epsilon}{2\bar{d}}) \approx A_p(E)$. For the action we can't make the same approximation. As the action appears in the cosine any change of the action of the order $\hbar$ will affect the sign of the summands. Instead we perform a Taylor expansion to linear order in ϵ. As the derivative of the action is the period we get

$$S_p(E \pm \tfrac{\epsilon}{2\bar{d}}) \approx S_p(E) \pm T_p(E)\tfrac{\epsilon}{2\bar{d}} .$$

Eq. (6.4) is thus approximated by

$$d_{\text{osc}}(E \pm \tfrac{\epsilon}{2\bar{d}}) \approx \frac{1}{\pi\hbar}\sum_p A_p(E) \cos\left(\frac{S_p(E)}{\hbar} \pm \frac{T_p(E)\epsilon}{2\hbar\bar{d}} - \mu_p\frac{\pi}{2}\right) . \tag{6.5}$$

We insert this result for the two level densities in $K(\tau)$, renaming the summation variable for the second orbit sum into p'. This gives

$$\begin{aligned} K(\tau) \approx\ & \frac{1}{(\pi\hbar\bar{d})^2}\left\langle \int_{-\infty}^{\infty} \sum_{p,p'} A_p A_{p'} \cos\left(\frac{S_p}{\hbar} + \frac{T_p\epsilon}{2\hbar\bar{d}} - \mu_p\frac{\pi}{2}\right) \right. \\ & \left. \times \cos\left(\frac{S_{p'}}{\hbar} - \frac{T_{p'}\epsilon}{2\hbar\bar{d}} - \mu_{p'}\frac{\pi}{2}\right) e^{2\pi i\epsilon\tau} d\epsilon \right\rangle \\ =\ & \frac{1}{2(\pi\hbar\bar{d})^2}\left\langle \int_{-\infty}^{\infty} \sum_{p,p'} A_p A_{p'} \left(\cos\left(\frac{S_p + S_{p'}}{\hbar} + \frac{(T_p - T_{p'})\epsilon}{2\hbar\bar{d}} - (\mu_p + \mu_{p'})\frac{\pi}{2}\right) \right. \right. \\ & \left. \left. + \cos\left(\frac{S_p - S_{p'}}{\hbar} + \frac{(T_p + T_{p'})\epsilon}{2\hbar\bar{d}} - (\mu_p - \mu_{p'})\frac{\pi}{2}\right)\right) e^{2\pi i\epsilon\tau} d\epsilon \right\rangle . \end{aligned} \tag{6.6}$$

Here we dropped the arguments E, and in the second line we used a trigonometric identity for the product of two cosines. The expression still looks rather complicated but we can strongly simplify it as we are averaging over the energy. As the actions depend on the energy (for billiards they are equal to $pL = \sqrt{2mE}L$) the cosines will usually oscillate as the energy is varied and contributions of energies leading to different signs will (nearly) cancel each other out. However in the second summand, involving a difference of actions, we can avoid such cancellations if we make the actions (and thus for billiards the lengths) identical. In this case the action difference drops out of the argument of the cosine. Hence for present calculation we just consider orbits with identical actions. (We will later also consider pairs of orbits with very small but nonzero action differences. These pairs give systematic contributions as well.)

The actions are identical if we just take identical orbits $p = p'$. In addition for a billiard it also possible to follow the same periodic orbit with opposite (time reversed) direction of motion. See Fig. 6.2 for an example of two periodic orbits that just differ in their sense of motion. Such orbits will also have identical action and can be taken as p and p'. If we take only identical and time reversed orbits we obtain the **diagonal approximation** for the spectral form factor. This approximation goes back to work by Michael Berry, and by John Hannay and Alfredo Ozorio de Almeida in the 1980s.

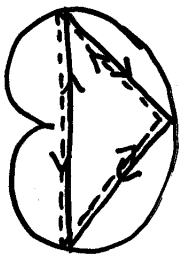


Figure 6.2: Two mutually time-reversed periodic orbits.

To implement this approximation we drop the first summand in (6.6). When writing down the second summand we just keep the sum over p and consider p' equal to p or its time reversed. For $p = p'$ all subscripts p' in (6.6) can be replaced by p. The same applies if p' is time reversed w.r.t. p, as time reversed orbits have identical stability amplitudes, lengths, actions as well as numbers of reflection and hence Maslov indices. The summation over p' thus only leads to a factor 2 accounting for the two possibilities of chosing the second orbit; this cancels the divisor 2 in the prefactor of (6.6). We thus obtain the approximation

$$K_{\mathrm{diag}}(\tau) = \frac{1}{(\pi\hbar\bar{d})^2} \left\langle \int_{-\infty}^{\infty} \sum_p A_p^2 \cos\left(\frac{T_p\epsilon}{\hbar\bar{d}}\right) e^{2\pi i\epsilon\tau} d\epsilon \right\rangle .$$

As $\hbar$ and $\bar{d}$ appear in this formula only as a product, we introduce the parameter $T_H = 2\pi\hbar\bar{d}$ known as the Heisenberg time. With this parameter we get

$$K_{\mathrm{diag}}(\tau) = \frac{4}{T_H^2} \left\langle \int_{-\infty}^{\infty} \sum_p A_p^2 \cos\left(\frac{2\pi T_p\epsilon}{T_H}\right) e^{2\pi i\epsilon\tau} d\epsilon \right\rangle .$$

If we use the expression for the cosine in terms of exponentials we get

$$K_\tau(\epsilon) = \frac{2}{T_H^2} \left\langle \int_{-\infty}^{\infty} \sum_p A_p^2 \left[\exp\left(2\pi i\left(\tau + \frac{T_p}{T_H}\right)\epsilon\right) + \left(2\pi i\left(\tau - \frac{T_p}{T_H}\right)\epsilon\right)\right] d\epsilon \right\rangle .$$

With $\int_{-\infty}^{\infty} \exp(2\pi i x\epsilon) d\epsilon = \delta(x)$ this gives

$$K_{\mathrm{diag}}(\tau) = \frac{2}{T_H^2} \left\langle \sum_p A_p^2 \left[\delta\left(\tau + \frac{T_p}{T_H}\right) + \delta\left(\tau - \frac{T_p}{T_H}\right)\right] \right\rangle .$$

Here the first delta function is zero as we are only considering $\tau > 0$ and T_p is also positive so the argument of this delta function is always positive. We are thus left with

$$K_{\text{diag}}(\tau) = \frac{2}{T_H^2} \left\langle \sum_p A_p^2 \delta\left(\tau - \frac{T_p}{T_H}\right) \right\rangle . \tag{6.7}$$

Now the problem has been reduced to evaluating a sum over orbits weighed with their squared stability amplitudes and with a delta function selecting that the period of the orbits should be $T_p = \tau T_H$. We also have the average $\langle \ldots \rangle$ which we would expect to indicate that the period doesn't have to be precisely τT_H but is allowed to vary in an interval around this value. Evaluating this sum is a purely classical problem and requires use of the classically chaotic dynamics, in particular ergodicity. The sum has been evaluated by Hannay and Ozorio de Almeida. Here we will just state their result, first in a more straightforward notation (without averages over delta functions) and then translated to the notation of (6.7).

Theorem. In the limit of large T, the sum over periodic orbits weighed with A_p^2 with periods in the interval $(T - \Delta T/2, T + \Delta/2)$ gives

$$\sum_{p \text{ with } T-\Delta T/2<T_p<T+\Delta T/2} A_p^2 \quad \approx T\Delta T . \tag{6.8}$$

Corollary. To translate to the notation of (6.7) we apply the sum rule to the case $T = \tau T_H$, $\Delta T = \Delta\tau T_H$. Here we have $T_H = 2\pi\hbar\bar{d} = 2\pi\hbar\frac{mA}{2\pi\hbar^2}$. If we formulate our limit physically, we have fixed $\hbar$ but increase the size of the system compared to the wavelength. This gives $A \to \infty$. Or we formally take $\hbar \to 0$. In either case the Heisenberg time and thus T and the period of the relevant orbits increases, as demanded for the application of the sum rule. For our case the sum rule thus boils down to

$$\sum_{p \text{ with } (\tau-\Delta\tau/2)T_H<T_p<(\tau+\Delta\tau/2)T_H} A_p^2 \quad \approx \Delta\tau\, \tau\, T_H^2 . \tag{6.9}$$

This can be rewritten as

$$\sum_p A_p^2 \int_{\tau-\Delta\tau/2}^{\tau+\Delta\tau/2} \delta\left(\tau - \frac{T_p}{T_H}\right) d\tau \approx \Delta\tau\, \tau\, T_H^2 . \tag{6.10}$$

Indeed if we do the integral above we pick up a contribution 1 from the delta function at $\tau = \frac{T_p}{T_H}$, but only if $\frac{T_p}{T_H}$ is in the integration interval. This selects orbits in the interval for T_p in (6.9). If we divide out $\frac{1}{\Delta\tau}$ the l.h.s. of (6.10) can be written in the form of the τ average in (6.2). We are also free to do an E average as in (6.2) which won't change anything. We can thus rewrite the sum rule as

$$\left\langle \sum_p A_p^2 \delta\left(\tau - \frac{T_p}{T_H}\right) \right\rangle = \tau T_H^2 .$$

If we insert this into (6.7) we obtain the diagonal approximation of the spectral form factor

$$K_{\text{diag}}(\tau) = 2\tau.$$

Systems without time reversal invariance. There are also systems where in contrast to Fig. 6.2 one can't just follow a trajectory with opposite sense of motion and expect a valid physical trajectory. This applies in particular to many systems with magnetic fields (if the magnetic field leads to trajectories different from, say, the billiard trajectories studied here). Some of these systems are still chaotic. In this case we can also apply the diagonal approximation but we only keep pairs of identical trajectories. This means a factor 2 will be omitted from the calculation an we get $K_{\rm diag}(\tau) = \tau$.

On the other hand in systems with geometric symmetries there are additional reasons why orbits can have identical lengths. For example in a system with reflection symmetry reflection of a periodic orbit at a symmetry line gives another periodic orbit. In this case the number of orbit pairs contributing to $K_{\rm diag}(\tau)$ increases accordingly.

6.3 Random matrix theory (not examinable)

This brief section is intended to give an idea of how the problem of spectral statistics in treated in Random Matrix Theory. Details are given in the Random Matrix Theory course.

Random Matrix Theory provides a way to get predictions for, e.g., the spectral form factor $K(\tau)$. Here matrices come in because they represent Hamiltonians. If one uses a basis of wavefunctions $\psi_n(\boldsymbol{r})$ (denoted by $|n\rangle$ in bra-ket notation) one can represent the Hamiltonian $\hat{H}$ by a matrix H with elements

$$H_{mn} = \langle m|\hat{H}|n\rangle = \int \psi_m^*(\boldsymbol{r})\hat{H}\psi_n(\boldsymbol{r})d^n r \ .$$

For a complicated system it is difficult to work with the precise $\hat{H}$ or H. It is much easier to instead consider different possible matrices, and average over all of them.

Which matrices do we have to consider? As the Hamiltonian $\hat{H}$ is Hermitian the matrices should be **Hermitian** as well. However for billiards (and many other systems) there is a further restriction: The billiard Hamiltonian $\hat{H} = -\frac{\hbar^2}{2m}\nabla^2$ is real. One can show that the same applies for all systems that are time reversal invariant, i.e., all system for which one can go through the same trajectory with opposite sense of motion and still obtain a valid trajectory. Hence for time-reversal invariant systems we should average over **real** matrices. If the matrices are both real and Hermitian that also implies that they are **symmetric**. The elements of these matrices satisfy

$$H_{\alpha\beta} = H_{\beta\alpha} \in \mathbb{R}.$$

In order to average over such matrices we need a weight. It is convenient to use the **'Gaussian' weight**

$$P(H) = \frac{1}{\mathcal{N}} e^{-a\,\mathrm{tr}\,H^2}$$

where a is a constant and $\mathcal{N}$ is chosen such that an integral over $P(H)$ alone is normalised, i.e.

$$\int dH_{11}dH_{12}\ldots dH_{NN}P(H) = 1.$$

Then a prediction for the spectral form factor is obtained by taking the same quantity as in (6.3), i.e.,

$$\frac{1}{\bar{d}^2}\int_{-\infty}^{\infty} d_{\rm osc}\left(E+\tfrac{\epsilon}{2\bar{d}}\right) d_{\rm osc}\left(E-\tfrac{\epsilon}{2\bar{d}}\right) e^{2\pi i\epsilon\tau} d\epsilon$$

but instead of the averages considered earlier we average over the real symmetric matrices with the weight $P(H)$. The prediction for the spectral form factor thus reads

$$\begin{aligned} K_{\rm RMT}(\tau) \;&=\; \int dH_{11}dH_{12}\ldots dH_{NN}\ P(H) \\ &\qquad \frac{1}{\bar{d}^2}\int_{-\infty}^{\infty} d_{\rm osc}\left(E+\tfrac{\epsilon}{2\bar{d}}\right) d_{\rm osc}\left(E-\tfrac{\epsilon}{2\bar{d}}\right) e^{2\pi i\epsilon\tau} d\epsilon\,. \end{aligned} \tag{6.11}$$

Here the level densities are sums over delta functions located at the eigenvalues of the matrices. One can choose different values for the matrix size N but usually calculations are done in the limit $N \to \infty$, which surprisingly makes things easier.

Ensembles.

- The real symmetric matrices with Gaussian weight as used above are referred to as the **Gaussian Orthogonal Ensemble (GOE)**. The word "orthogonal" instead of "real symmetric" is a little confusing. The ensemble is called orthogonal because multiplication of a real symmetric matrix with an orthogonal matrix and its inverse from both sides gives another real symmetric matrix; also the weight $e^{-a \operatorname{tr} H^2}$ remains the same if H is multiplied with orthogonal matrices in this way.
- Some systems are not time reversal invariant. Examples are systems that involve a magnetic field. In these cases there is no reason to restrict the average to real matrices and we instead consider arbitrary Hermitian matrices. The Hermitian matrices with Gaussian weight are referred to as the **Gaussian Unitary Ensemble (GUE)**. Similarly as for the GOE this name was chosen because multiplying a Hermitian matrix from both sides with a unitary matrix gives another Hermitian matrix.

It turns out that the spectral form factor $K(\tau)$ for chaotic systems agrees with the random matrix prediction.

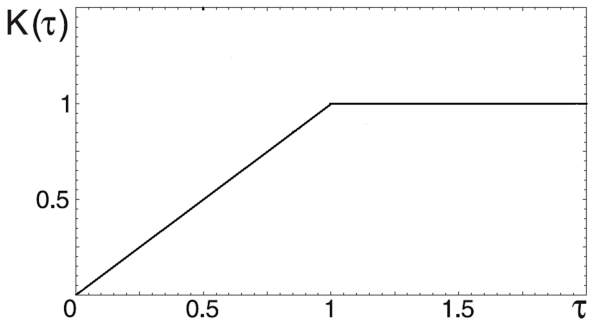
Bohigas-Giannoni-Schmit conjecture. The spectral statistics of time reversal invariant chaotic systems, e.g. their spectral form factor, agree with the predictions from the Gaussian Orthogonal Ensemble, and for chaotic systems without time reversal invariance we obtain agreement with the Gaussian Unitary Ensemble.

(Here I am omitting special cases such as spin systems and systems with geometric symmetries.)

Before discussing in more detail the reasons for this agreement I want to list the random matrix predictions for $K(\tau)$.

Predictions. If we follow this approach we obtain the following results for the spectral form factor:

- For systems without time reversal invariance we just get τ for $0 < \tau < 1$, in line with the diagonal approximation. However the result remains equal to 1 for $\tau > 1$.



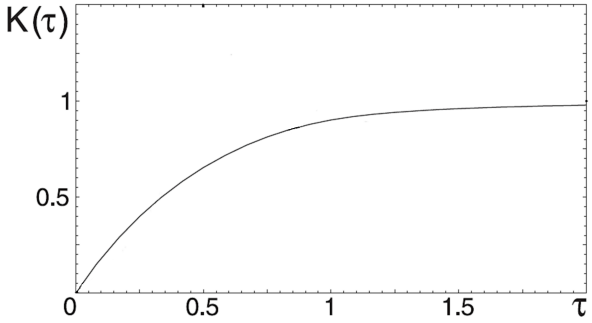
- For systems with time reversal invariance and $0 < \tau < 1$ we obtain

$$K(\tau) = 2\tau - \tau \ln(1 + 2\tau)$$

which can be expanded into the following power series

$$K(\tau) = 2\tau - 2\tau^2 + 2\tau^3 - \frac{8}{3}\tau^4 + ...$$

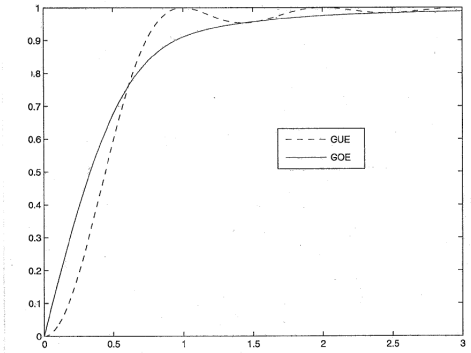
The leading term agrees with the diagonal approximation but the remaining ones must arise from a different mechanism. For $\tau > 1$ the behaviour is different, and I am omitting the corresponding formula.



Level repulsion. We expect that energy levels will 'repel' each other, i.e., small energy differences ϵ should be unlikely. To see this behaviour more prominently in our predictions it's better to omit the Fourier transform and work with the full level density instead of including only the oscillary part. The resulting correlation function

$$\frac{1}{\bar{d}^2}\left\langle d\left(E+\frac{\epsilon}{2\bar{d}}\right) d\left(E-\frac{\epsilon}{2\bar{d}}\right)\right\rangle$$

is similar to (6.1) and it gives the likelihood of energy differences $\frac{\epsilon}{d}$. If we plot it we get the following:



Hence indeed small energy differences are much less likely than larger ones. As expected the random matrix results display level repulsion.

6.4 Beyond the diagonal approximation (not examinable)

At first sight, it is not at all clear why chaotic systems should conform with random matrix predictions. A way to show this is to build a theory that is valid for chaotic quantum systems, using the trace formula, and then show that we get the same results as in random matrix theory. The diagonal approximation was the first step in this direction. As seen above the diagonal approximation agrees with the terms linear in τ in the random matrix result for $0 < \tau < 1$.

The remaining terms in $K(\tau)$ can be understood in terms of periodic orbits as well. Recall that we expressed $K(\tau)$ as a double sum over periodic orbits p and p'. For the diagonal approximation we argued that important contributions to that double sum arise from pairs of orbits whose actions are identical because the orbits are either identical or mutually time reversed. However also pairs of orbits with very similar actions give contributions. Such pairs of orbits were identified by Martin Sieber and Klaus Richter. A sketch is shown in Fig. 6.3. Here one orbit (the red/dashed one) crosses itself with a small angle. Then one can show that there is a partner orbit narrowly avoiding this crossing (the black

one with a full line). The two orbits differ only near the crossing; otherwise they are very close. They have very similar lengths and thus very similar actions.

Fig. 6.3 is a sketch only. A realistic picture of such an orbit pair in a billiard is depicted in Fig. 6.4. On the l.h.s. we see a periodic orbit inside the billiard, composed of straight lines and obeying the reflection law when hitting the boundary. In the picture a part of the billiard is cut off, in reality in continues further to the right. A circle hightlights a crossing between two of the straight lines. (There is also a third line that happens to go through the circle, but this is a mere coincidence.) The picture on the r.h.s. shows a partner orbit that is very similar but avoids the crossing; the two lines in the circle now just come close but don't cross.

Orbit pairs like this exist only in time reversal invariant systems as a part of the dashed orbit follows a part of the orbit depicted by a full line with opposite sense of motion. Sieber and Richter showed that orbit pairs like this are responsible for the next-to-leading term in the τ expansion, equal to $-2\tau^2$. This term also agrees with the predictions of random matrix theory for the Gaussian Orthogonal Ensemble.

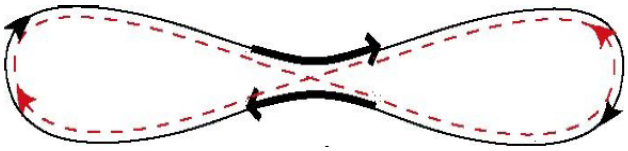


Figure 6.3: Pair of orbits with a crossing and an avoided crossing (sketch).

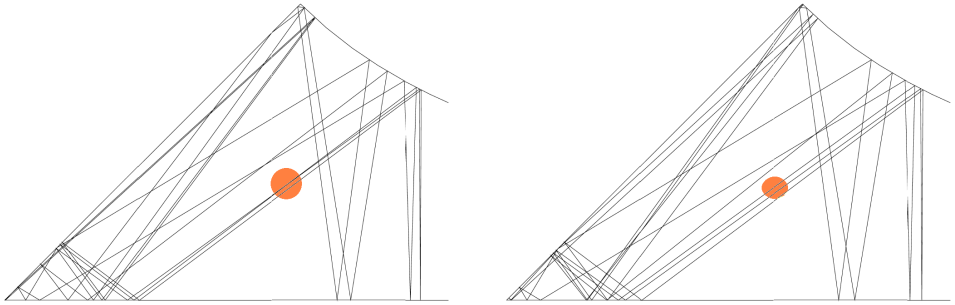


Figure 6.4: Pair of orbits with a crossing and an avoided crossing (realistic picture).

For even higher orders in τ one has to consider more and more complicated orbit pairs based on the same mechanism. For example one can take pairs where one orbit has two crossings and its partner avoids both, or one orbit has a crossing at one place and an avoided crossing at a different place, and in the partner orbit it is the other way around. Examples are shown on the l.h.s. of Fig. 6.5. Or we can have generalisations of a crossing where not two but three or more parts of an orbit come close. Then in the partner orbit connections in this part of the orbit are changed, and it is very close to the original orbit everywhere else. Examples are shown on the r.h.s. of Fig. 6.5. The types of orbit pairs sketched in Fig. 6.5 together give the τ^3 contribution to $K(\tau)$ in agreement with the GOE.

One can get arbitrarily high orders in τ if one includes orbit pairs with arbitrarily many crossings involving arbitrarily many parts of the orbit. For example Fig. 6.6 shows orbits with one crossing, and two generalised crossings where three parts of an orbit come close. The crossing can be replaced by an avoided crossing. For the generalised crossing we have six different ways of drawing connections, depicted in the inset. This allows to generate many different orbits with similar actions. If we insert any two of these orbits in (6.6) we get systematic contributions to the spectral form factor. By evaluating all such contributions to (6.6) one can access the full expansion of $K(\tau)$ in powers of τ.

For systems that are **not time reversal invariant** only some types of orbit pairs are possible. In the diagonal approximation, only identical orbits have to be considered. As it is not possible to go through

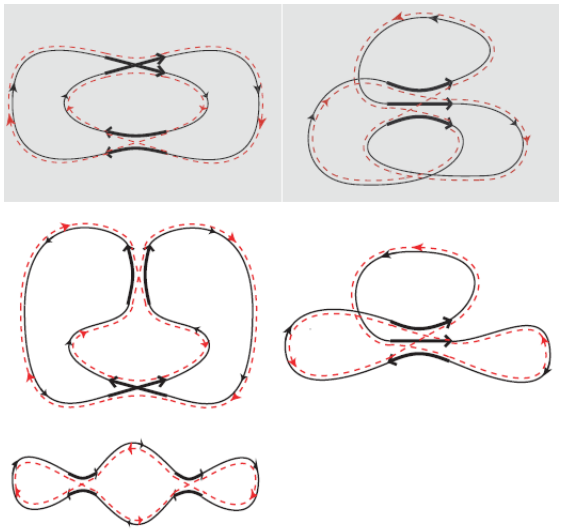


Figure 6.5: Orbit pairs responsible for the τ^3 contribution to $K(\tau)$.

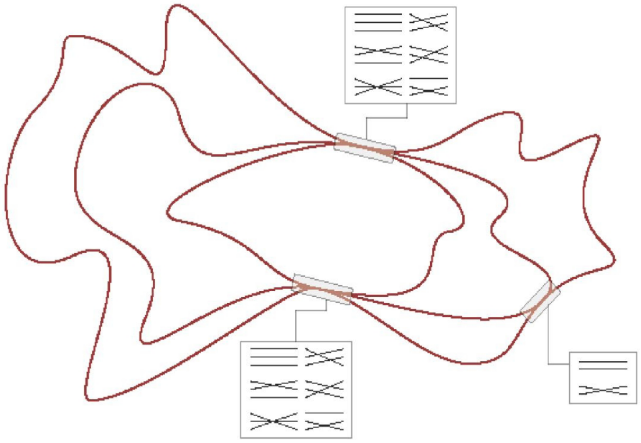


Figure 6.6: Orbits differing in one crossing and two generalised crossings where three parts of an orbit come close.

the same orbit with opposite direction of motion p and p' cannot be mutually time reversed. We have already seen that this halves result of the diagonal approximation we just get τ. As argued before the pairs from Fig. 6.3 don't exist in this case either. Many of the other types of orbit pairs differing in crossings are also impossible without time reversal invariance. The others are still possible but one can show that their contributions cancel, leaving (for $0 < \tau < 1$) only the diagonal approximation.

To understand the behaviour for $\tau > 1$ as well, one has to be more careful when implementing the semiclassical approximation, and use a variant of the semiclassical approximation that preserves more quantum mechanical information than the trace formula.

A Appendix (not examinable)

A.1 Stability matrix for a reflection

We now prove that the deviations $r_\perp$ and α after the reflection are related to the deviations $r_{\perp 0}$ and α_0 before the reflection by

$$\begin{pmatrix} r_\perp \\ \alpha \end{pmatrix} \approx \begin{pmatrix} -1 & 0 \\ -\frac{2}{R\cos\beta} & -1 \end{pmatrix} \begin{pmatrix} r_{\perp 0} \\ \alpha_0 \end{pmatrix} . \tag{A.1}$$

To do so we compare the original trajectory to a slightly changed trajectory with angle of incidence and angle of reflection $\tilde{\beta}$. The angle difference between the two trajectories before the reflections is denoted by α_0, and the angle difference after the reflections is denoted by α. Similarly the perpendicular separation before the reflections is denoted by $r_{\perp 0}$, and the corresponding separation after the reflections is $r_\perp{}'$. Normally we would have to say carefully whether $r_{\perp 0}$ and $r_\perp$ are taken perpendicular to the original or the changed trajectory and at which point close to the reflections the distance will be measured. However this choice is not important for our results, as as we are only interested in an approximation for small deviations $r_{\perp 0}$ and α_0. We just follow the most convenient convention, shown in Fig. A.1.

As the reflection points are close, the part of the circle between these points can be approximated by a straight line. Then the two shaded triangles in Fig. A.1 are similar: They agree in the length of one side, and they both contain one right angle and one angle β. Hence the remaining side lengths and angles coincide as well. This means that within our approximation $|r_\perp| = |r_{\perp 0}|$. However the directions of the perpendicular deviations are opposite. In the figure, the changed trajectory is to the left of the original trajectory before the reflections, and to the right after the reflections. Therefore we have

$$r_\perp = -r_{\perp 0} . \tag{A.2}$$

We now want to determine α in terms of $r_{\perp 0}$ and α_0. We first observe that the line between the two shaded triangles has the length $\frac{r_{\perp 0}}{\cos\beta}$. If the radius of the circle is R this means that the angle γ in the figure can be approximated by

$$\gamma \approx \frac{r_{\perp 0}}{R\cos\beta} . \tag{A.3}$$

To relate $\alpha_0, \alpha, \beta, \tilde{\beta}$, and γ it is helpful to consider Fig. A.2 where the three points marked by dots in Fig. A.1 have been identified. We can immediately identify two angles β, two angles $\tilde{\beta}$, and one angle γ. The angle inclosed between the two ingoing lines must be α_0. The sign is positive as the ingoing line of the changed trajectory is further to the left compared to the ingoing line of the original trajectory. The angle between the two outgoing lines is related to α. However we need to include a minus sign as the direction of the outgoing line of the changed trajectory is further to the right compared to the outgoing line of the original trajectory. Now we see from Fig. A.2 that the angle $-\alpha$ can be obtained

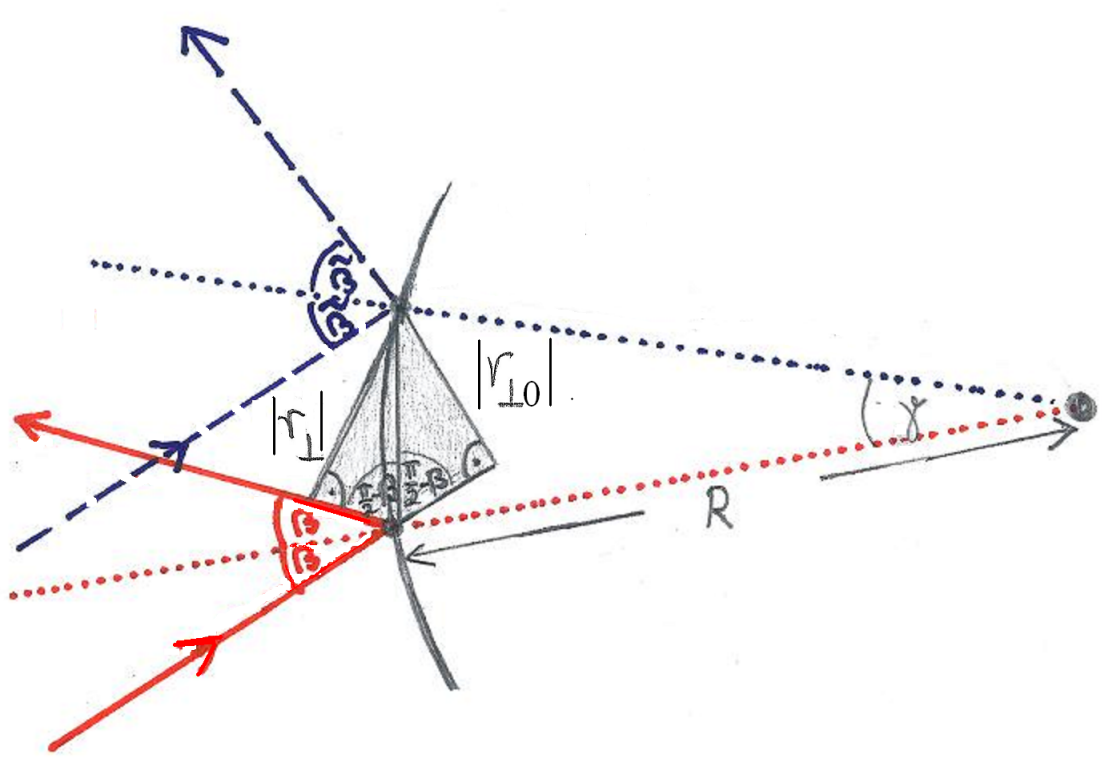


Figure A.1: Initial deviations $r_{\perp 0}, \alpha_0$ and final deviations $r_\perp, \alpha$ for a reflection at a circle of radius R.

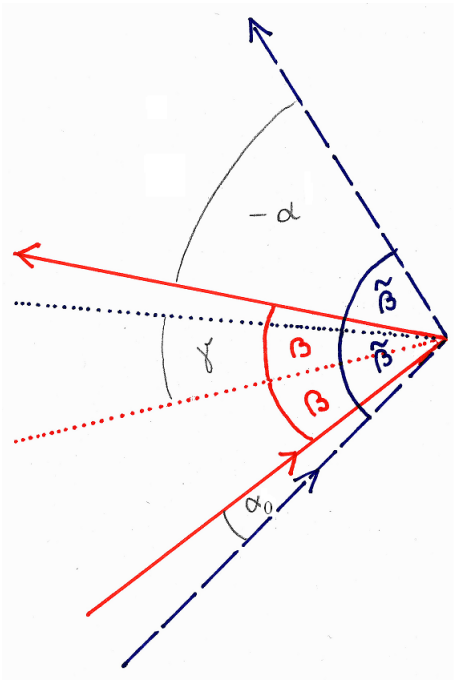


Figure A.2: The angles involved in a reflection.

as follows:

$$\begin{aligned} -\alpha &= \tilde{\beta} - (\beta - \gamma) \\ &= (\alpha_0 + \beta + \gamma) - (\beta - \gamma) \\ &= \alpha_0 + 2\gamma \\ &\approx \alpha_0 + \frac{2}{R\cos\beta} r_{\perp 0} \, . \end{aligned} \tag{A.4}$$

In the first line we used that $-\alpha$ can be obtained from the upper angle $\tilde{\beta}$ by subtracting an angle $\beta - \gamma$. In the second line we used that the lower angle $\tilde{\beta}$ is a sum of α_0, β, and γ. In the fourth line we cancelled the β's and in the fourth line we inserted (A.3). Writing (A.2) and (A.4) in matrix notation we obtain the desired result

$$\begin{pmatrix} r_\perp \\ \alpha \end{pmatrix} \approx \begin{pmatrix} -1 & 0 \\ -\frac{2}{R\cos\beta} & -1 \end{pmatrix} \begin{pmatrix} r_{\perp 0} \\ \alpha_0 \end{pmatrix} \, . \tag{A.5}$$

A.2 More about elliptic/stable dependence on initial conditions

We now discuss in more detail the case $|\mathrm{tr}\, M| < 2$. As mentioned before this leads to **eigenvalues** $m_{1,2} = \frac{1}{2}\left(\mathrm{tr}\, M \pm \sqrt{(\mathrm{tr}\, M)^2 - 4}\right)$ that are complex and complex conjugate with respect to each other, i.e., $m_2 = m_1^*$. As we also have $m_2 = \frac{1}{m_1}$ this means that $|m_1|^2 = 1$. Hence the eigenvalues are complex numbers with absolute value 1 and can be written in the form

$$m_1 = e^{i\phi}, \qquad m_2 = e^{-i\phi} \, .$$

The **eigenvectors** are mutually complex conjugate as well. To show this, we start from the eigenvalue equation for $m_1 = e^{i\phi}$,

$$M\boldsymbol{u}_1 = e^{i\phi}\boldsymbol{u}_1 \, . \tag{A.6}$$

We now complex conjugate the whole equation, using that M is real. This leads to

$$M\boldsymbol{u}_1^* = e^{-i\phi}\boldsymbol{u}_1^* \, .$$

Hence the eigenvector corresponding to $e^{-i\phi}$ is $\boldsymbol{u}_1^*$.

We now investigate how the deviation between two nearby trajectories evolves in time. To do so we split the **initial deviation** $(r_{\perp 0}, \alpha_0)$ into components pointing in direction $\mathrm{Re}\, \boldsymbol{u}_1$ and $-\mathrm{Im}\, \boldsymbol{u}_1$. (Why? We should split into components connected to the eigenvectors. As the eigenvectors are complex and the problem is real it makes sense to consider basis vectors $\mathrm{Re}\, \boldsymbol{u}_1, -\mathrm{Im}\, \boldsymbol{u}_1 \in \mathbb{R}^2$. The minus sign in front of the imaginary part is just for convenience.) If we denote the coefficients by a_0 and b_0 and also introduce $z_0 = a_0 + ib_0$ we thus write the initial deviation as

$$a_0\, \mathrm{Re}\, \boldsymbol{u}_1 + b_0(-\mathrm{Im}\, \boldsymbol{u}_1) = \mathrm{Re}\, z_0\, \mathrm{Re}\, \boldsymbol{u}_1 - \mathrm{Im}\, z_0\, \mathrm{Im}\, \boldsymbol{u}_1 = \mathrm{Re}\, (z_0 \boldsymbol{u}_1) \, .$$

The **deviation in the end of the trajectory** is similarly decomposed as

$$a\, \mathrm{Re}\, \boldsymbol{u}_1 + b(-\mathrm{Im}\, \boldsymbol{u}_1) = \mathrm{Re}\, z\, \mathrm{Re}\, \boldsymbol{u}_1 - \mathrm{Im}\, z\, \mathrm{Im}\, \boldsymbol{u}_1 = \mathrm{Re}\, (z\boldsymbol{u}_1)$$

where $z = a + ib$. The final deviation can be obtained by applying the stability matrix to the initial deviation. We thus obtain

$$\mathrm{Re}\, (z\boldsymbol{u}_1) = M\, \mathrm{Re}\, (z_0\boldsymbol{u}_1) = \mathrm{Re}\, (z_0 M\boldsymbol{u}_1) = \mathrm{Re}\, (z_0 e^{i\phi}\boldsymbol{u}_1) \, .$$

In the third line we used that M is real and can thus be written inside the real part. In the fourth line we used the eigenvalue equation (A.6). Now comparison between the first and fourth line gives

$$z = e^{i\phi} z_0 \, .$$

In the complex plane multiplication with $e^{i\phi}$ increases the phase of z by ϕ. This means that deviation z is just rotated in the complex plane by an angle ϕ. Thus the deviation is rotated in a plane spanned by $\mathrm{Re}\, \boldsymbol{u}_1$ and $-\mathrm{Im}\, \boldsymbol{u}_1$, but neither increased nor decreased.

A.3 Stability islands

With the results from the previous appendix we can give a (slightly) more thorough discussion of the stability islands arising in Poincaé plots of mixed systems, see Fig. 1.8. These islands are related to stable periodic orbits of the system. One stable periodic orbit of the corresponding to Fig. 1.8 is depicted in Fig. 1.9. This orbit just involves two reflections. Hence it corresponds to only two points in the plot. (In the picture one is at $s = 0, p = \sin\beta = 0$; the other one is shown twice at $s = 1, p = \sin\beta = 0$ and $s = -1, p = \sin\beta = 0$ as the points $s = 1$ and $s = -1$ are identified.) Now what happens if we start a trajectory close to this periodic orbit? We first look at this situation in terms of $(r_\perp, \alpha)$. If we work in these coordinates the resulting deviation after one traversal of the orbit would be rotated compared to the initial one by an angle ϕ in the plane spanned by $\mathrm{Re}\,\boldsymbol{u}_1$ and $-\mathrm{Im}\,\boldsymbol{u}_1$. The deviation after two revolutions is rotated by 2ϕ, and so on. Altogether these points form a circle. Now one can show that the same circles appear if instead we consider a Poincaré plot and work with s and $\sin\beta$. They explain the circles around the two points corresponding to the periodic orbit. The other stability islands in the plot are related to more complicated stable periodic orbits. The structures closer to the edge of the stability islands are also quite interesting, but this is even beyond the scope of this appendix.

A.4 Proof for $B = 0$

We now show that, as required for the derivation of the van Vleck propagator, B as defined in Eq (3.17) vanishes. This is not needed for the rest of the course and will not be included in the lectures or the examinable material. As a preparation we compute the Laplacian of the length.

Lemma. If L is the length of a physical billiard trajectory the Laplacian w.r.t. the final point $\boldsymbol{r}$ can be expressed through elements of the stability matrix M as

$$\nabla^2 L = \frac{M_{22}}{M_{12}} \,. \tag{A.7}$$

Proof. Due to $\mathrm{grad}\, L = \boldsymbol{d}$ we have

$$\nabla^2 L = \mathrm{div}\,\mathrm{grad}\, L = \mathrm{div}\,\boldsymbol{d} = \frac{\partial d_1}{\partial r_1} + \frac{\partial d_2}{\partial r_2} \,. \tag{A.8}$$

This equation holds in any system of coordinates. For our purposes it is convenient to use a system of coordinates $r_\parallel$, $r_\perp$ with coordinate axes parallel and perpendicular to the direction of the trajectory in the end. In these coordinates Eq. (A.8) boils down to

$$\nabla^2 L = \frac{\partial d_\parallel}{\partial r_\parallel} + \frac{\partial d_\perp}{\partial r_\perp} \,.$$

To evaluate the first summand we have to investigate how $\boldsymbol{d}$ changes if the final point $\boldsymbol{r}$ is just moved further into the direction of the trajectory. In this case the final direction does not change at all, hence the derivative is zero. This leaves to evaluate $\frac{\partial d_\perp}{\partial r_\perp}$, i.e., we have to determine the change of the direction vector perpendicular to the old direction if the final point is moved in perpendicular direction. It is convenient to relate this change to the change of the direction angle. If the final direction changes by an angle α, this leads to a change $d_\perp = \sin\alpha$ in perpendicular direction, see Fig A.3. As in Chapter 1 we consider only small α and thus approximate $d_\perp = \sin\alpha$ by α. This implies

$$\nabla^2 L = \frac{\partial \alpha}{\partial r_\perp} \,.$$

Reassuringly the variables in this expression are the same as those considered in Chapter 1 when studying the dependence of the motion in a billiard on small changes of the initial conditions. $r_\perp$ and α describe the deviations in the end of a trajectory resulting from a change of the initial position and angle. Let us denote the changes of the initial position and angle by $r_{\perp 0}$ and α_0. Then the final deviations are obtained from the initial deviations by multiplying with the stability matrix M,

$$\begin{pmatrix} r_\perp \\ \alpha \end{pmatrix} = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} r_{\perp 0} \\ \alpha_0 \end{pmatrix} . \tag{A.9}$$

We might hope that the derivative $\frac{\partial \alpha}{\partial r_\perp}$ is related to elements of the stability matrix. However it is important that the derivative in $\frac{\partial \alpha}{\partial r_\perp}$ is taken w.r.t. a final change in position assuming that the initial position is fixed. Hence to compute the derivative we first have to express α in terms of $r_\perp$ and $r_{\perp 0}$ and then take the derivative w.r.t. $r_\perp$ for $r_{\perp 0} = 0$. We thus take the second line of Eq. (A.12),

$$\alpha = M_{21} r_{\perp 0} + M_{22} \alpha_0 \, .$$

We then insert α_0 as obtained from the first line,

$$\begin{aligned} r_\perp &= M_{11} r_{\perp 0} + M_{12} \alpha_0 \\ \Rightarrow \alpha_0 &= \frac{1}{M_{12}} r_\perp - \frac{M_{11}}{M_{12}} r_{\perp 0} \, . \end{aligned}$$

This gives

$$\alpha = M_{21} r_{\perp 0} + \frac{M_{22}}{M_{12}} r_\perp - \frac{M_{11} M_{22}}{M_{12}} r_{\perp 0}$$

Now we can take the derivative w.r.t. $r_\perp$ and obtain the desired result

$$\nabla^2 L = \frac{\partial \alpha}{\partial r_\perp} = \frac{M_{22}}{M_{12}} \, .$$

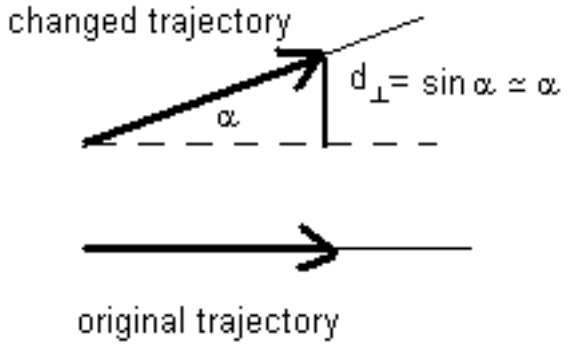


Figure A.3: Relation between $d_\perp$ and α.

We can now show why the B defined in (3.17) vanishes.

$$B = 2L^{3/2} \nabla L \cdot \nabla A + A L^{3/2} \nabla^2 L = 0$$

Proof. We use that B is proportional to

$$\frac{2 \nabla L \cdot \nabla A}{A} + \nabla^2 L \tag{A.10}$$

where $AL^{3/2}$ was divided out. Now we simplify the term with the scalar product as

$$\begin{aligned}\frac{2\nabla L\cdot\nabla A}{A} &= \frac{2\boldsymbol{d}\cdot\nabla A}{A}\\ &= \frac{2\frac{\partial A}{\partial r_\parallel}}{A}\\ &= 2\frac{\partial}{\partial r_\parallel}\ln|A|\\ &= 2\frac{\partial}{\partial r_\parallel}\left(\ln\left|\frac{(-1)^n}{i\epsilon\pi}\right|-\frac{1}{2}\ln|M_{12}|\right)\\ &= -\frac{\frac{\partial}{\partial r_\parallel}M_{12}}{M_{12}}\,.\end{aligned}$$

Here we first used that $\nabla L=\boldsymbol{d}$. Then we used that the scalar product of a gradient with a unit vector gives the partial derivative w.r.t. the coordinate increasing in the direction of the unit vector. E.g. the scalar product of a gradient with the vector $(1,0)$ gives the partial derivative in x-direction. Here we face a scalar product with the direction of the trajectory and hence we have to take the partial derivative w.r.t. $r_\parallel$. Then we used the formula for the derivative of the logarithm. Afterwards we used the definition of A (omitting the subscripts γ), and wrote the logarithm of a product as the sum of the logarithms of the two factors. Finally we dropped the derivative of the constant term and took the derivative of $\ln|M_{12}|$.

The problem has thus been reduced to computing $\frac{\partial M_{12}}{\partial r_\parallel}$, i.e, studying how M_{12} changes as $r_\parallel$ is increased. Increasing $r_\parallel$ simply means that we move further along the direction of the trajectory, increasing the final straight-line segment. It is thus convenient to decompose the stability matrix M into the stability matrix for a straight-line segment of length $r_\parallel$ and the stability matrix M^0 for the case $r_\parallel=0$. The product is

$$M=\begin{pmatrix}1 & r_\parallel\\ 0 & 1\end{pmatrix}\begin{pmatrix}M^0_{11} & M^0_{12}\\ M^0_{21} & M^0_{22}\end{pmatrix}=\begin{pmatrix}M^0_{11}+r_\parallel M^0_{21} & M^0_{12}+r_\parallel M^0_{22}\\ M^0_{21} & M^0_{22}\end{pmatrix}$$

which implies $M_{12}=M^0_{12}+r_\parallel M^0_{22}$ and $M_{22}=M^0_{22}$. Hence we have

$$\frac{\partial M_{12}}{\partial r_\parallel}=M^0_{22}=M_{22}\,.$$

If we insert this result as well as $\nabla^2 L=\frac{M_{22}}{M_{12}}$ in (A.10) we obtain

$$B\propto\frac{2\nabla L\cdot\nabla A}{A}+\nabla^2 L=-\frac{\frac{\partial}{\partial r_\parallel}M_{12}}{M_{12}}+\nabla^2 L=-\frac{M_{22}}{M_{12}}+\frac{M_{22}}{M_{12}}=0$$

as desired. This completes the missing bit of the proof.

Remark. Importantly if $M_{\gamma,12}$ is zero the prefactor in $K_{V,\gamma}$ diverges. Also the steps in our proof where we divided by $M_{\gamma,12}$ don't work any more. Actually this is the reason why the proof is restricted to dispersing billiards. For these billiards we showed in Chapter 1 that for trajectories with at least one reflection the elements of the stability matrix are nonzero (and even that they all have the same sign). For trajectories without reflections we don't have to worry either as these trajectories have $M_\gamma=\begin{pmatrix}1 & L_\gamma\\ 0 & 1\end{pmatrix}$ and thus $M_{\gamma,12}\neq 0$ as well. If one wants to amend the problem of diverging $M_{\gamma,12}$ one obtains the additional phase factor mentioned earlier (third remark in section 3.4).

A.5 Second derivatives of the length

We proceed to derive the formulas for second derivatives of the length w.r.t. $r_\perp$ and $r_{\perp 0}$ used in the derivation of the trace formula earlier. We use that $\frac{\partial L}{\partial \boldsymbol{r}} = \boldsymbol{d}$ implies that $\frac{\partial L}{\partial r_\perp} = d_\perp$ where $d_\perp$ is the component of the final direction vector that is perpendicular to the orbit with $r_\perp = 0$. Hence it is the sine of the angle enclosed between the trajectory under consideration and the orbit with $r_\perp = 0$. This angle is just the α we introduced earlier. Hence for small α we just have $\frac{\partial L}{\partial r_\perp} = d_\perp = \sin\alpha \approx \alpha$.

Similarly we have $\frac{\partial L}{\partial \boldsymbol{r}_0} = -\boldsymbol{d}_0$ where $\boldsymbol{d}_0$ is the initial direction of the trajectory. By the same argument as above this leads to $\frac{\partial L}{\partial r_{\perp 0}} \approx -\alpha_0$.

These results imply

$$\frac{\partial^2 L}{\partial r_\perp^2}(\boldsymbol{r},\boldsymbol{r}_0) \approx \frac{\partial \alpha}{\partial r_\perp}, \quad \frac{\partial^2 L}{\partial r_\perp \partial r_{\perp 0}}(\boldsymbol{r},\boldsymbol{r}_0) \approx \frac{\partial \alpha}{\partial r_{\perp 0}} \approx -\frac{\partial \alpha_0}{\partial r_\perp}, \quad \frac{\partial^2 L}{\partial r_{\perp 0}^2}(\boldsymbol{r},\boldsymbol{r}_0) = -\frac{\partial \alpha_0}{\partial r_{\perp 0}}\,. \tag{A.11}$$

At $r_\perp = r_{\perp 0} = 0$ the approximate sign can actually be replaced by an equality as the approximation was correct to leading order in the deviations $r_\perp, r_{\perp 0}$ so first derivatives at $r_\perp = r_{\perp 0} = 0$ are evaluated without error.

The occurrence of derivatives involving $r_\perp, r_{\perp 0}, \alpha, \alpha_0$ is reassuring as these quantities are related to the stability matrix that we expect to arise in the claimed result. As discussed in chapter 1 multiplication with the stability matrix M turns the initial deviations into final deviations

$$\begin{pmatrix} r_\perp \\ \alpha \end{pmatrix} \approx \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} r_{\perp 0} \\ \alpha_0 \end{pmatrix}\,. \tag{A.12}$$

To compute the derivatives above we need the angles as functions of $r_\perp$ and $r_{\perp 0}$. The first line of (A.12) implies

$$\begin{aligned} r_\perp &\approx M_{11} r_{\perp 0} + M_{12}\alpha_0 \\ \Rightarrow \alpha_0 &\approx \frac{1}{M_{12}} r_\perp - \frac{M_{11}}{M_{12}} r_{\perp 0}\,. \end{aligned} \tag{A.13}$$

The second line then gives

$$\begin{aligned} \alpha &\approx M_{21} r_{\perp 0} + M_{22}\alpha_0 \\ &\approx M_{21} r_{\perp 0} + \frac{M_{22}}{M_{12}} r_\perp - \frac{M_{11}M_{22}}{M_{12}} r_{\perp 0} \\ &= \frac{M_{12}M_{21} - M_{11}M_{22}}{M_{12}} r_{\perp 0} + \frac{M_{22}}{M_{12}} r_\perp \\ &= -\frac{1}{M_{12}} r_{\perp 0} + \frac{M_{22}}{M_{12}} r_\perp \end{aligned} \tag{A.14}$$

where in the last line we used that $\det M = M_{11}M_{22} - M_{12}M_{21} = 1$. The required derivatives are thus

$$\frac{\partial \alpha}{\partial r_\perp} \approx \frac{M_{22}}{M_{12}}, \qquad \frac{\partial \alpha}{\partial r_{0\perp}} \approx -\frac{1}{M_{12}}, \qquad \frac{\partial \alpha_0}{\partial r_\perp} \approx \frac{1}{M_{12}}, \qquad \frac{\partial \alpha_0}{\partial r_{0\perp}} \approx -\frac{M_{11}}{M_{12}} \tag{A.15}$$

As (A.13) and (A.14) are approximations to leading order in $r_\perp$ and $r_{\perp 0}$, the results for the derivatives are again exact if the derivatives are taken at $r_\perp = 0$ and $r_{\perp 0} = 0$. If we combine (A.11) and (A.15) and use that both approximations are exact at $r_\perp = r_{\perp 0} = 0$ we thus obtain the desired result

$$\frac{\partial^2 L}{\partial r_\perp^2}(\boldsymbol{r},\boldsymbol{r}_0) = \frac{M_{22}}{M_{12}}, \qquad \frac{\partial^2 L}{\partial r_\perp \partial r_{\perp 0}}(\boldsymbol{r},\boldsymbol{r}_0) = -\frac{1}{M_{12}}, \qquad \frac{\partial^2 L}{\partial r_{\perp 0}^2}(\boldsymbol{r},\boldsymbol{r}_0) = \frac{M_{11}}{M_{12}}\,.$$

Note that the middle result is the same regardless of which of the two corresponding expressions from (A.11) we use.