DIPLOMARBEIT

Exceptional Points and Inverse Scattering in Optical Interferometers

zur Erlangung des akademischen Grades

Diplom-Ingenieur

im Rahmen des Studiums

Technische Physik

eingereicht von

Michael Fraller
Matrikelnummer 00928090

ausgeführt am Institut für Theoretische Physik
der Fakultät für Physik der Technischen Universität Wien

Betreuung
Betreuer: Univ.Prof. Dipl.-Ing. Dr.techn. Stefan Rotter

Wien, 11.02.2020 (Unterschrift Verfasser) (Unterschrift Betreuer)

Contents

1 Introduction

The scattering of waves is a phenomenon that is literally all around us. From the beautiful interference patterns produced by sunlight in a rainbow to the propagation of acoustic waves in a concert hall, we encounter wave scattering on a daily basis. In spite of its already very rich history, there are still many open and interesting scattering problems to be solved from the scientific point of view. In this thesis we will address two specific problems of current interest, related to so-called Exceptional Points[1] and to the notoriously challenging inverse scattering problem. To get a handle on these phenomena, we will study them in the context of scattering in systems of reduced dimensionality. Specifically, we will work with optical interferometers in which light propagates and interferes along one-dimensional axes.

Generally speaking, the scattering problem can either be solved in forward or in the inverse direction. For the forward scattering problem the structure of the scatterer is known and one is interested in the connection between the outgoing and incoming parts. This can for example be used to predict the outcome of a scattering experiment where an incoming wave impinges on a target and one wants to know which outgoing wave pattern to expect. For the inverse scattering problem the structure of the scatterer is typically unknown. The goal here is to obtain information on the target scatterer based on the incoming and outgoing wave components. The inverse scattering problem has many different applications, ranging

from seismology[2] to bio-medical imaging[3].

In this thesis we are solving the forward and the inverse scattering problem for special cases. For the forward scattering problem we look at optical interferometers at a so-called Exceptional Point (EP)[1] and shall present strategies for obtaining these singularities in relatively simple setups that can easily be realized with current-day technology. EPs only occur in non-Hermitian problems where the eigenvalues are complex in general[4,5,6]. If two or more eigenvalues are degenerate the geometric multiplicity may not be equal to the algebraic one, unlike the Hermitian case leading to a range of unique properties[7]. EPs and their properties were already observed[8,9], for example in experiments with microcavities[10,11]. Compared to microcavities, optical interferometers have an easier experimental setup while maintaining the properties needed for the occurrence of an EP such as the existence of multiple modes and coupling between them. In this work we will discuss the occurrence of EPs in optical interferometers[9] and show their properties, specially the vanishing eigenvectors, the response to perturbations, the characteristic Riemann sheet structure around them and finally their impact on the scattering problem. The absence of eigenvectors was already observed in an experiment with an optical interferometer, which could be explained by linking it to an EP[12]. In particular, we will point out how to realize EPs in optical interferometers by just tuning a few parameters like the reflectivities of mirrors or beam-splitters and the distance between them. With these properties EPs could also be relevant for sensors like Sagnac interferometers, which could be enhanced by an EP[13,14,15].

In the second part of this thesis we will work on the inverse scattering problem in optical resonators containing a dielectric potential landscape. The task we want to address is to determine the shape of this potential just from the scattering information. Motivated by a different project in our group, where this problem was solved with a machine learning approach, we want to solve this problem with con-

ventional methods and compare the results. In the machine learning approach the function which maps the scattering matrix to the potential was obtained by training a neural network with precalculated samples obtained by solving the forward scattering problem on random potentials. The machine learning approach can be seen as high-dimensional data fitting, and therefore we will use similar conventional methods for comparison. Two approaches which address this problem in a similar way, are the well known gradient descent[16] and simulated annealing[17] method. In both cases we will transform the inverse scattering problem into a minimization problem. We will find the solutions with the named methods by tuning (fitting) the potential. We will discuss the quality of these approaches on several examples and shall estimate their performance, their advantages, disadvantages and limits.

2 Forward Scattering Problem

If one is interested in the scattering properties of a known system the forward scattering problem needs to be solved. It is a well known and popular problem and has applications in nearly every part of physics. The problem is defined by a scatterer and an incoming and outgoing wave field. These fields are linked together by the scattering matrix, which is uniquely defined by the properties of the scatterer and the frequency of the incoming wave field. We will look into the forward scattering problem to obtain the scattering matrix for optical interferometers at an Exceptional Point (EP)[9]. The scattering matrix allows us to predict the behavior of the optical interferometer under different circumstances, such as when small perturbations are applied to it. Since many sensors realized with optical interferometers rely on their behavior to perturbations it is particularly interesting to study them at an EP where the sensitivity is enhanced[10]. For this purpose we first study EPs and their properties in a quantum mechanical system. We will then show how these results apply to optical interferometers. We further need to obtain a method to find and identify EPs in optical interferometers, which will then allow us to study how an EP manifests in the scattering matrix.

2.1 Exceptional Points

Exceptional Points (EPs) typically occur in non-Hermitian systems featuring more than one mode[7]. They are characterized by two or more coalescing eigenvalues with one or more vanishing eigenvectors. The eigenvalues are complex in general but can also be real if, e.g., the system is PT-symmetric[18,5,8]. Compared to the already existing experiments with microcavities at an EP, we will study their occurrence in optical interferometers, since their experimental setup is easier compared to microcavities. EPs are linked to unique properties. At the EP one eigenstate vanishes and only one persists, which is directly linked to the chirality of EPs[19]. For every EP there exists a corresponding one with different chirality where the other eigenstate vanishes[20]. This chirality translates into a certain sense of direction in whispering-gallery mode resonators, where one mode is suppressed and only one survives. With a ring laser, which is similar to a whispering-gallery resonator, this chirality was already observed experimentally and can further also be linked to an EP[12]. A system at an EP reacts highly sensitively to perturbations[11]. The eigenvalue splitting follows a root behavior which has a high slope for small perturbation strength. Further the eigenvalues in the vicinity of an EP lie on complex surfaces also known as Riemann sheets[21], which are self-intersecting and have a pole point in its derivative which is directly located at the EP. This is linked to the root behavior of the eigenvalue splitting and the high sensitivity of a system in the vicinity of an EP. The characteristic topology of the Riemann sheets can further be useful to identify EPs.

2.1.1 Non-Hermitian Hamiltonians

To study Exceptional Points we consider a quantum mechanical system which is described by the Schrödinger equation:

$$i\frac{\partial}{\partial t}\psi = \hat{H}\psi \tag{2.1}$$

For simplicity we use natural units where we set $\hbar = 1$. The conservation of energy leads to a Hermitian Hamiltonian $\hat{H} = \hat{H}^\dagger$ with therefore real eigenvalues and an unitary time evolution operator $\hat{U}(t, t_0) = e^{-i\hat{H}(t-t_0)}$. However if one is only interested in a limited subspace of a system the Hamiltonian is not necessary Hermitian[4,5,1,18]. In this case the eigenvalues are complex in general and the time evolution operator is no longer unitary. This is directly linked to the gain or loss of this subsystem[9]. Moreover, if two or more eigenvalues coalesce in such a non-Hermitian subsystem, the geometric multiplicity may not equal the algebraic one, unlike in the Hermitian case[6]. These points are called Exceptional Points (EPs) and play a major role in non-Hermitian physics[22,7]. EPs offer unique properties such as the vanishing of eigenvectors, root behavior of the splitting of eigenvalues and a characteristic topology of the Riemann sheets around them[21,10].

If we now consider a two mode system we can choose a mode basis and the Schrödinger equation can be written as:[23]

$$\begin{aligned} i\frac{d}{dt}\begin{pmatrix} a_1(t) \\ a_2(t) \end{pmatrix} &= H \begin{pmatrix} a_1(t) \\ a_2(t) \end{pmatrix} \\ H &= \begin{pmatrix} s_1 & g_{12} \\ g_{21} & s_2 \end{pmatrix} \end{aligned} \tag{2.2}$$

where a_1 and a_2 are the modes and H is the Hamiltonian in this mode basis. The diagonal elements of H, $s_i = \omega_i + i\sigma_i$, are the eigenfrequencies and g_{ij} are the

coupling strengths of the modes which are chosen to be real. For non-vanishing coupling the modes are no longer independent of each other and are not the eigenstates of the system anymore. The eigenstates can be obtained by solving the eigenvalue problem of H.

$$\begin{aligned} det(H - I\lambda_i) = 0 \\ (H - I\lambda_i)v_i = 0 \end{aligned} \tag{2.3}$$

where λ_i are the eigenfrequencies and v_i are the eigenvectors (eigenstates). Demanding the coalescence of the eigenvalues leads to the equations:

$$\begin{aligned} \lambda_{\pm} = \frac{s_1 + s_2}{2} \pm \sqrt{\frac{(s_1 - s_2)^2}{4} + g_{12}g_{21}}, \\ (\frac{s_1 - s_2}{2})^2 + g_{12}g_{21} = 0. \end{aligned} \tag{2.4}$$

Solving the equations with g_{12} and g_{21} as parameters and introducing $s_+ = \frac{s_1+s_2}{2}$ and $s_- = \frac{s_1-s_2}{2}$ leads to the following solutions where we differentiate between vanishing and non-vanishing coupling strength. For vanishing coupling strength we obtain the Hamiltonian and eigenstates of a system at a so-called Diabolic Point (DP).

$$\begin{aligned} H_{\mathrm{DP}} &= \begin{pmatrix} s_+ & 0 \\ 0 & s_+ \end{pmatrix} \\ \lambda_{1,2} &= s_+ \\ v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad & v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{aligned} \tag{2.5}$$

At a DP the two modes are independent of each other and these points also occur in Hermitian physics.

For non-vanishing coupling strength we obtain the general form of the Hamiltonian

and the eigenstates of a system at an Exceptional Point (EP).

$$
\begin{aligned}
H_{\mathrm{EP}\pm} &= \begin{pmatrix} s_+ \mp i\sqrt{g_{12}g_{21}} & g_{12} \\ g_{21} & s_+ \pm i\sqrt{g_{12}g_{21}} \end{pmatrix} \\
\lambda_{1,2\pm} &= s_+ \\
v_{1\pm} &= \frac{1}{\sqrt{g_{12}+g_{21}}} \begin{pmatrix} \sqrt{g_{12}} \\ \pm i\sqrt{g_{21}} \end{pmatrix}
\end{aligned}
\tag{2.6}
$$

There only exists one eigenvector $v_{1\pm}$, which is a characteristic property of EPs, and the Hamiltonian is therefore not diagonalizable. Since every Hermitian matrix is similar to a diagonal matrix, this can only occur in the non-Hermitian case. Looking at the Hamiltonian of an EP the parameters g_{12} and g_{21} do not fully define the system, there are two possible solutions $H_{\mathrm{EP}+}$ and $H_{\mathrm{EP}-}$, which is the so-called chirality of EPs[19]. The diagonal elements of the Hamiltonian correspond to the eigenfrequencies of the coupled modes. if s_+ is real they have an imaginary part of same magnitude but opposite sign, which means one mode has a gain and the other one a loss of the same strength. Whether it is an EP+ or EP- depends on the placements of the gain and the loss. Looking at symmetric coupling, $g_{12} = g_{21} \neq 0$, an EP+ has the eigenvector $v_{1+} = \begin{pmatrix} 1, & i \end{pmatrix}^T$ and an EP- $v_{1-} = \begin{pmatrix} 1, & -i \end{pmatrix}^T$. In this case the eigenvector is a superposition of the modes and the chirality defines the phase between them. In the case of asymmetric coupling, where one coupling strength vanishes, the two modes have neither gain nor loss. Here the chirality is defined by which coupling is vanishing. The eigenvector of the Hamiltonian takes to form of $v_1 = \begin{pmatrix} 1, & 0 \end{pmatrix}^T$ for $g_{21} = 0$, and $v_1 = \begin{pmatrix} 0, & 1 \end{pmatrix}^T$ for $g_{12} = 0$. In this case not only one eigenvector but also one mode vanishes. For systems like whispering-gallery resonators or ring cavities the chirality translates into a certain sense of direction[24].

2.1.2 Eigenmode splitting

The eigenvalues of the Hamiltonian of a multiple mode system can be observed experimentally and manifest in the scattering properties. It is therefore crucial to study their behavior. We are now interested in how the system responds to small perturbations when it is located at an EP and shall compare this situation to the one encountered at a DP. We are particularly interested in the eigenvalue splitting as a result of small variations of the system parameters. For simplicity we will fix the parameter $s_+ = 0$ and assume a symmetric coupling mechanism, $g_{12} = g_{21} = 1$, for an EP and no coupling, $g_{12} = g_{21} = 0$, for a DP. Since we describe two resonators, we look at perturbations of their eigenfrequencies and coupling strengths. For the perturbation of the eigenfrequencies, P_ω, we only consider an asymmetric part since a symmetric one would only effect s_+ and will not result in any splitting of the eigenvalues. For the perturbation of the coupling strengths, P_g, we consider a symmetric one since we assumed a symmetric coupling mechanism.

$$P_\omega = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad P_g = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{2.7}$$

Adding one of the perturbation matrices with a perturbation strength of ϵ to the Hamiltonian, leads to the Hamiltonian of the perturbed system.

$$\begin{aligned} H &= H_{\{\mathrm{EP,DP}\}} + \epsilon P_{\{\omega,g\}} \\ H_{\mathrm{EP}} &= \begin{pmatrix} -i & 1 \\ 1 & i \end{pmatrix} \\ H_{\mathrm{DP}} &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \end{aligned} \tag{2.8}$$

Solving the eigenvalue problem of the perturbed Hamiltonian leads to the eigenvalues of the system. For P_ω we obtain the eigenvalue splitting of

$$\begin{aligned} |\Delta\lambda_{\mathrm{EP}}| &= 2\sqrt{\epsilon}\sqrt{-2i+\epsilon} \\ |\Delta\lambda_{\mathrm{DP}}| &= 2\epsilon \end{aligned} \tag{2.9}$$

and for P_g

$$\begin{aligned} |\Delta\lambda_{\mathrm{EP}}| &= 2\sqrt{\epsilon}\sqrt{2+\epsilon} \\ |\Delta\lambda_{\mathrm{DP}}| &= 2\epsilon \end{aligned} \tag{2.10}$$

Comparing the magnitude of the splitting, as shown in Figure 2.1, the eigenvalues at an EP split with a characteristic square root behavior while the splitting at a DP is linear. This results in a high sensitivity of systems at an EP. The high slope of the root function for small perturbation strengths could be exploited to create highly sensitive measuring systems[14,15,13], compared to ones at a DP. The direction of the splitting in the complex plane is shown in Figure 2.2. For a DP the perturbed Hamiltonian is still Hermitian, since the perturbations are Hermitian too, and therefore the eigenvalues remain real. For an EP the eigenvalues are splitting along the real axis with the perturbation P_g and into the complex plane with P_ω. The splitting along the real axis is explained by the PT-symmetry of P_g. The eigenvalues of a PT-symmetric Hamiltonian can be real even if the Hamiltonian is non-Hermitian[18,5,8]. The high sensitivity of EPs can again be observed by the distance between the eigenvalues.

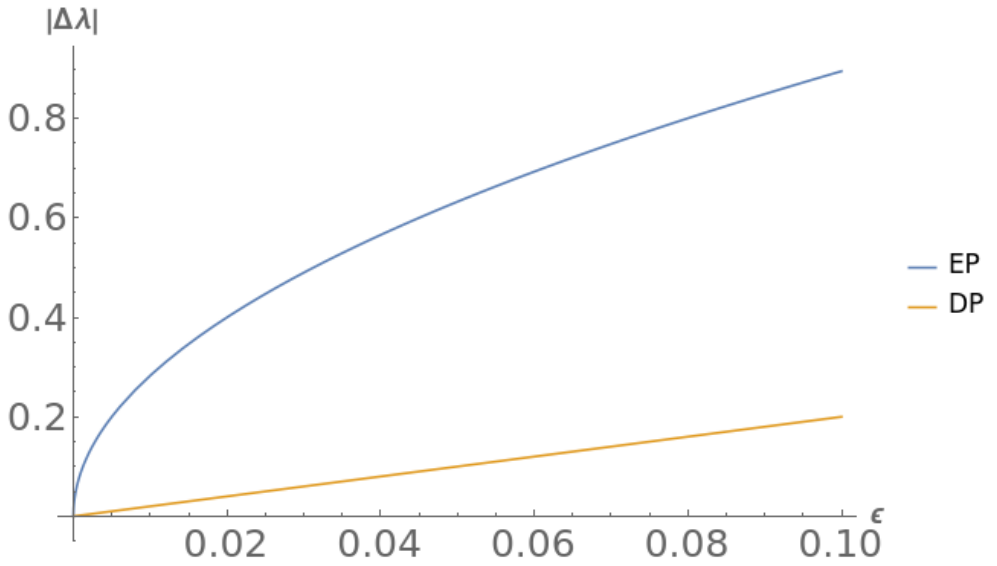
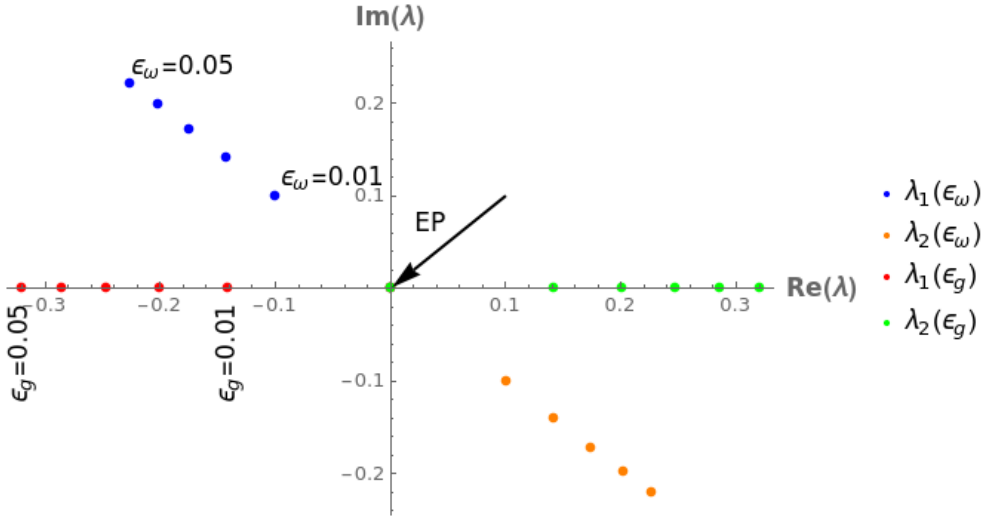


Figure 2.1: Magnitude of the eigenvalue splitting at an EP and at a DP. We chose the parameters $s_+ = 0$, $g_{\mathrm{EP}_{12}} = g_{\mathrm{EP}_{21}} = 1$, $g_{\mathrm{DP}_{12}} = g_{\mathrm{DP}_{21}} = 0$ with the perturbation P_ω with a varying strength of ϵ. For an EP the splitting follows a characteristic square root behavior which explains the high sensitivity of EPs. For a DP the splitting is linear.

(a) Eigenvalue splitting in the complex plane at an EP.



(b) Eigenvalue splitting in the complex plane at a DP.

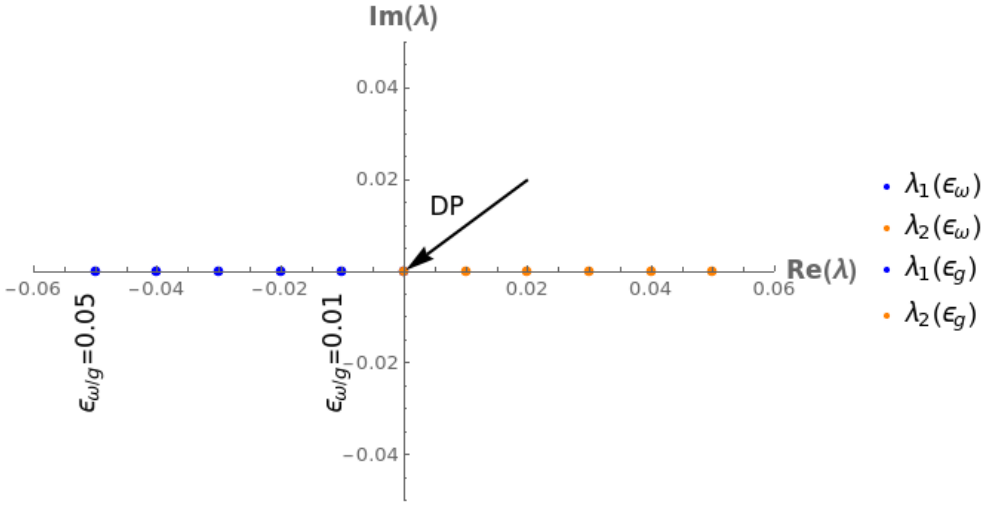


Figure 2.2: Comparison of the eigenvalue splitting at an EP and at a DP in the complex plane with the perturbations P_ω and P_g. We again chose the parameters $s_+ = 0$, $g_{\mathrm{EP}_{12}} = g_{\mathrm{EP}_{21}} = 1$, $g_{\mathrm{DP}_{12}} = g_{\mathrm{DP}_{21}} = 0$. The perturbation strengths ϵ_ω and ϵ_g were varied from 0 to 0.05 with a step size of 0.01. At an EP the eigenvalues are splitting in opposite direction, along the real axes for P_g and into the complex plane for P_ω. At a DP the eigenvalues split in opposite direction along the real axis for both perturbations with a linear behavior.

2.1.3 Riemann sheets

For a better understanding of how a system reacts to perturbations we now look at the two perturbations of system parameters combined rather than independently. This allows us to obtain the behavior of the system for more complex perturbations where both the eigenfrequencies and coupling strengths are perturbed simultaneously. By plotting the eigenvalues in the two dimensional perturbation space $(\epsilon_\omega, \epsilon_g)$, the eigenvalues lie on complex planes, also known as Riemann sheets[21]. These sheets reveal the behavior of the eigenvalues and have a different topology in the vicinity of an EP or a DP. The Riemann sheets of a system are interesting to study since with their characteristic topology we can decide whether the system is at an EP or a DP.

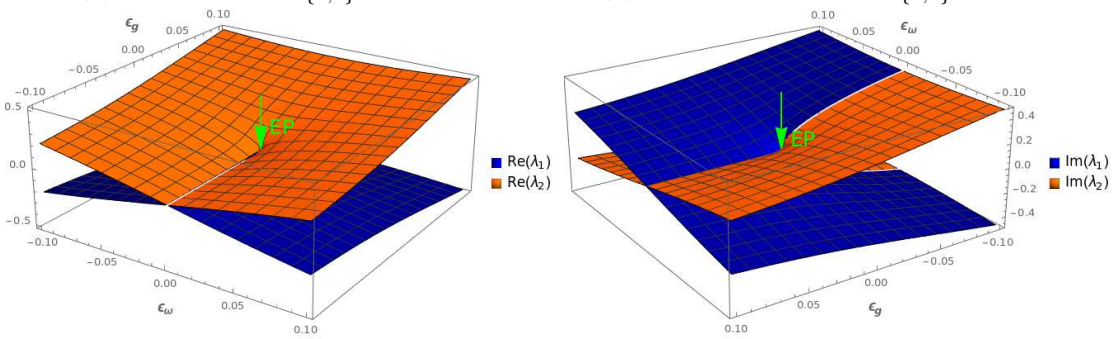
The Riemann sheets are obtained by solving the eigenvalue problem of the perturbed Hamiltonian with two perturbation strengths, ϵ_ω and ϵ_g , as parameters.

$$\begin{aligned} H(\epsilon_\omega, \epsilon_g) &= H_{\{\mathrm{EP,DP}\}} + \epsilon_\omega P_\omega + \epsilon_g P_g \\ det(H(\epsilon_\omega, \epsilon_g) &- I\lambda_i(\epsilon_\omega, \epsilon_g)) = 0 \end{aligned} \tag{2.11}$$

This leads to the eigenvalue functions $\lambda_i(\epsilon_\omega, \epsilon_g)$ which are plotted in Figure 2.3. At a DP the Riemann sheets have a conical form, hence the name Diabolic Point, and the eigenvalue splitting is linear in every direction. The sheets have the topology of the absolute function, hence the linear behavior. For an EP the Riemann sheets are self-intersecting, with the topology of the complex square root function. The EP is located at the intersection of the sheets. The derivative of the sheets has a pole point located directly at the EP, which is again linked to the high sensitivity.

(a) Real part of $\lambda_{\{1,2\}}$ at an EP

(b) Imaginary part of $\lambda_{\{1,2\}}$ at an EP



(c) Real part of $\lambda_{\{1,2\}}$ at a DP

(d) Real part of $\lambda_{\{1,2\}}$ at a DP, different perspective

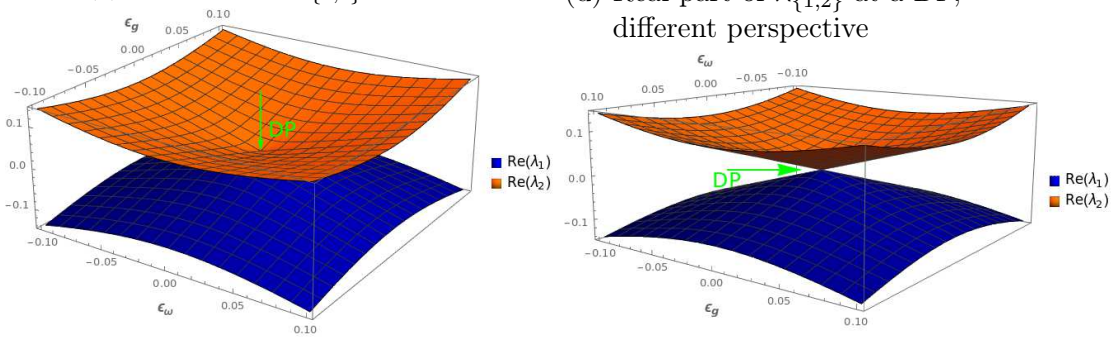


Figure 2.3: Riemann Sheets of a system in the vicinity of an EP and a DP. At an EP the Riemann sheets are self-intersecting. The high sensitivity at an EP manifests as a pole point in the derivative of the sheets located directly at the EP. At a DP the Riemann sheets have a conical form, and the eigenvalues have a linear splitting behavior. Since the perturbed Hamiltonian of a DP is Hermitian, the eigenvalues remain real and have no imaginary part.

2.2 Optical Interferometers

For years, people have been looking at EPs in complicated resonator arrangements where tuning the system to the EP was very difficult[10]. Here we propose an interesting and easily accessible alternative to study EPs that just involves the easy setup of an optical interferometer. Optical interferometers have a wide range of applications and are often used as sensors, e.g. for distance measuring or determining refractive indices, and are used in spectroscopy. EPs in optical interferometers are particularly interesting to study since the properties of an EP, specially the high sensitivity, can be exploited to create highly sensitive measuring systems[14,15,13]. Further EPs itself and their properties can be studied with a simple setup of mirrors on an optical table, compared to a more complex experiment with microcavities. Optical interferometers are systems of mirrors and optical parts with a given geometry. Each part is fully described by its number of ports and its scattering matrix. The ports of each part are connected to other optical parts with free space between them and consist of an incoming and outgoing wave part. The optical wave at each port is given by the superposition of the incoming and outgoing part which are linked together by the scattering matrix of the corresponding optical part. A port is called closed when it is internally connected to another optical part, and opened otherwise. Open ports can be seen as external interfaces and can be monitored in experiments. We will restrict our systems to a set of fundamental optical parts given in Table 2.1. The path of the optical waves are defined by the geometry and placements of the mirrors (M1, M2, STM) and beam splitters (BS) which are connected to each other by free space. To study EPs in interferometers we show the analogy to quantum mechanical systems and introduce a mathematical model to obtain the scattering properties and eigenstates. We then introduce an optical interferometer which is similar to the well-known coupled microcavities system with equally strong gain and loss, where an EP was

already observed experimentally[10]. The equivalent interferometer consists of two optical cavities coupled with a mirror. To achieve an equally strong gain and loss we additionally introduce a gain/loss medium (G) which is placed inside the cavities. Except this gain/loss medium the scattering matrices of all parts are unitary, which is directly linked to the conservation of energy in Hermitian systems. But even without a gain/loss a system may still be non-Hermitian due to the interaction with the environment at the open ports, which couple out radiation. We further introduce an optical ring cavity, similar to a whispering-gallery resonator, which can be tuned to an EP without the need of an active gain medium. This interferometer was already studied experimentally with interesting results[12], and we will show that they can be linked to EPs. At last we will look into interferometers with an EP of higher order[25], where more than one eigenvector vanishes and the sensitivity is enhanced even more.

Symbol	Description	S-Matrix	Parameters
G	gain/loss medium	$\begin{pmatrix} 0 & e^h \\ e^h & 0 \end{pmatrix}$	h
M1	fully reflecting mirror	(-1)	-
M2	fully reflecting mirror at an angle	$\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$	-
STM	semi-transparent mirror	$\begin{pmatrix} -r_{\mathrm{STM}} & i\sqrt{1-r_{\mathrm{STM}}^2} \\ i\sqrt{1-r_{\mathrm{STM}}^2} & -r_{\mathrm{STM}} \end{pmatrix}$	r_{STM}
BS	beam splitter	$\begin{pmatrix} 0 & -r_{\mathrm{BS}} & i\sqrt{1-r_{\mathrm{BS}}^2} & 0 \\ -r_{\mathrm{BS}} & 0 & 0 & i\sqrt{1-r_{\mathrm{BS}}^2} \\ i\sqrt{1-r_{\mathrm{BS}}^2} & 0 & 0 & -r_{\mathrm{BS}} \\ 0 & i\sqrt{1-r_{\mathrm{BS}}^2} & -r_{\mathrm{BS}} & 0 \end{pmatrix}$	r_{BS}

Table 2.1: Fundamental parts of our optical interferometers. The mirrors M1 and M2 are fully reflecting. They are not tunable and only define the path of the optical waves. The semi-transparent mirror STM and the beam-splitter BS have a tunable reflectivity (r_{STM}, r_{BS}). Their transmission is defined by the reflectivity and the fact that these parts have a unitary scattering matrix. We additionally introduce a gain/loss medium which is the only non-Hermitian one, with the gain h as parameter. The parameters are used to bring the interferometers to an EP and further to introduce perturbations.

2.2.1 Mathematical description

To bring an optical interferometer to an EP and obtain its properties we need a mathematical description of interferometers in general. For this purpose we assume light propagation and interference only along one-dimensional axes, which allows us to employ the so-called paraxial approximation. This approximation has many applications in optics and is specially useful to e.g. describe the propagation of light emitted by a laser. A paraxial wave can be written as

$$u(x,z) = E(x,z)e^{ikz} \tag{2.12}$$

where $E(k,z)$ is the envelope of the electric field and k the wavenumber. The envelope $E(k,z)$ has to fulfill the paraxial optical wave equation[8,26,27].

$$i\frac{\partial}{\partial z}E(x,z) + (\frac{1}{2k}\frac{\partial^2}{\partial x^2} + k_0 n(x))E(x,z) = 0 \tag{2.13}$$

where $n(x) = n_r(x) + in_i(x)$ is the complex refractive index, $k = k_0 n_0$ the wavenumber and k_0 the corresponding wavenumber in free space. Comparing this equation to the Schrödinger equation in natural units for a quantum particle:

$$i\frac{\partial}{\partial t}\psi(x,t) + (\frac{1}{2m}\frac{\partial^2}{\partial x^2} - V(x))\psi(x,t) = 0 \tag{2.14}$$

we see a canonical isomorphism between the two equations obtained by the replacements $z \to t$, $k \to m$, $k_0 n(x) \to -V(x)$ and $E(x,z) \to \psi(x,t)$[8]. This isomorphism introduces a correspondence between an optical interferometer and a quantum mechanical system and defines an (effective) Hamiltonian. Further the optical interferometer is at an EP if the corresponding quantum mechanical system is at one.

To obtain the eigenstates of an optical interferometer we start by solving the scat-

tering problem. All ports of the interferometer are arbitrarily enumerated and we refer to them by their port number i and to their outgoing and incoming wave part as a_i and b_i. As boundary condition we choose incoming plane waves with wavenumber k at each open port i with a complex amplitude f_i.

$$a_i = f_i \tag{2.15}$$

To obtain the wave inside the interferometer we split it into its free space components. We address each part of free space by their adjacent port numbers m and n, and make the ansatz of forward an backward propagating plane waves with wavenumber k and complex amplitudes $c_{m,n}$, $d_{m,n}$. The wave in each free space part can be written as:

$$u_{m,n}(z) = c_{m,n}e^{ik(z-\frac{l}{2})} + d_{m,n}e^{-ik(z-\frac{l}{2})} \tag{2.16}$$

Since we are looking at each part of free space independently we choose the zero point of z in the middle between the two adjacent optical parts. With this ansatz we can obtain the amplitudes at the ports m and n, which are located at $z_m = -\frac{l}{2}$ and $z_n = \frac{l}{2}$.

$$\begin{aligned} a_n &= c_{m,n}e^{ik\frac{l}{2}} \\ b_n &= d_{m,n}e^{-ik\frac{l}{2}} \\ a_m &= d_{m,n}e^{ik\frac{l}{2}} \\ b_m &= c_{m,n}e^{-ik\frac{l}{2}} \end{aligned} \tag{2.17}$$

Eliminating $c_{m,n}$ and $d_{m,n}$ leads to the relation of the incoming and outgoing waves at port m and n.

$$\begin{pmatrix} a_m \\ a_n \end{pmatrix} = \begin{pmatrix} 0 & e^{ikl} \\ e^{ikl} & 0 \end{pmatrix} \begin{pmatrix} b_m \\ b_n \end{pmatrix} \tag{2.18}$$

We will refer to this equation as compatibility condition at free space.
Each optical part, which links together ports $p_1 \dots p_{n_p}$, where n_p denotes the parts number of ports, is described by a scattering matrix $S_{p_1,\dots,p_{n_p}}$. The scattering matrix of each part is defined by the parts type (M1, M2, STM, BS or G) and its parameters, see Table 2.1. This scattering matrix leads to the compatibility conditions at the optical part.

$$\begin{pmatrix} b_{p_1} \\ \vdots \\ b_{p_n} \end{pmatrix} = S_{p_1,\dots,p_{n_p}} \begin{pmatrix} a_{p_1} \\ \vdots \\ a_{p_n} \end{pmatrix} \tag{2.19}$$

Combining all equations of boundary conditions, compatibility conditions at free space and compatibility conditions at optical parts leads to a system of linear equations, written in matrix form:

$$\mathbf{M}(k)\vec{v} = \vec{f} \tag{2.20}$$

The matrix $\mathbf{M}(k)$ is the coefficient matrix of the equations, $\vec{v} = \begin{pmatrix} a_1, & b_1, & \dots, & a_n, & b_n \end{pmatrix}^T$ contains the incoming and outgoing wave parts of each port and $\vec{f} = \begin{pmatrix} f_1, & \dots, & f_n \end{pmatrix}^T$ is a constant vector which holds the amplitudes of the incoming waves given by the boundary conditions. Solving this equation with a specific k and $\vec{f}$ leads to the scattering states $\vec{v}_{scat}(k,\vec{f}) = \begin{pmatrix} a_1(k,\vec{f}), & b_1(k,\vec{f}), & \dots, & a_n(k,\vec{f}), & b_n(k,\vec{f}) \end{pmatrix}^T$ of the interferometer. Since only the open ports can be monitored in an experiment we introduce the scattering matrix $\mathbf{S_{tot}}(k)$ of the whole interferometer which links them together.

$$\begin{pmatrix} b_{op_1} \\ \vdots \\ b_{op_n} \end{pmatrix} = \mathbf{S_{tot}}(k) \begin{pmatrix} a_{op_1} \\ \vdots \\ a_{op_n} \end{pmatrix} \tag{2.21}$$

Here $op_1 \dots op_n$ denotes the port numbers of the open ports. The elements of the total scattering matrix $(\mathbf{S_{tot}}(k))_{i,j}$ define the connection between the amplitude of the incoming wave at port op_j and the outgoing wave at port op_i. They can be expressed in terms of the scattering solutions.

$$(\mathbf{S_{tot}}(k))_{i,j} = b_{op_i}(k, \vec{f}_{op_j}) \tag{2.22}$$

where $b_{op_i}(k, \vec{f}_{op_j})$ denotes the amplitude of the outgoing wave at port op_i, and $\vec{f}_{op_j}$ represents the boundary conditions with only one incoming wave at port op_j with an amplitude of 1.

Even if the matrix $\mathbf{M}(k)$ is regular in general, for some interferometers it can be singular for a specific k, which leads to non-trivial solutions with no external input $\vec{f} = 0$. These solutions $\vec{v}_{qb}(k_i) = \vec{v}_{scat}(k_i, 0)$ are called quasi-bound states and need to fulfill the equations

$$\begin{aligned} det(M(k_i)) &= 0 \\ \vec{v}_{qb}(k_i) &\in ker(\mathbf{M}(k_i)) \end{aligned} \tag{2.23}$$

By considering the isomorphism between the paraxial wave equation and the Schrödinger equation, the plane waves with constant amplitude in z-direction correspond to quantum mechanical wave functions with time-independent amplitude. Since these stationary states are the eigenstates of the Hamiltonian we can identify the quasi-bound states as eigenstates of the interferometer. Similar to the eigenvalues of the Hamiltonian we can assign an algebraic and geometric multiplicity to the quasi-bound states.

$$\begin{aligned} a_{k_i} &= \min_{l \in \mathbb{N}}(\lim_{k \to k_i} \frac{det(\mathbf{M}(k))}{(k - k_i)^l} \neq 0) \\ g_{k_i} &= dim(ker(\mathbf{M}(k_i))) \end{aligned} \tag{2.24}$$

With the introduced method we can describe optical interferometers, solve the scattering problem and obtain the eigenstates. With the definitions of the algebraic and geometric multiplicity we can determine if an interferometer is at an EP. This further allows us to study the behavior of the eigenvalues and also how they manifest in the scattering matrix. In the following sections we will look into specific interferometers, tune them to EPs and study their properties.

2.2.2 Double cavity

The occurrence of EPs in optics was already studied experimentally, but not in optical interferometers. They were mainly observed in experiments with microcavities with a more difficult experimental setup. A well-known example for an EP in optics is an experiment with two coupling microcavities[28]. The two equally shaped microcavities have perfectly balanced gain and loss values and are coupled to each other in a symmetric way. They are described by an operator that is very similar to the Hamiltonian in Equation 2.6. To introduce EPs into optical interferometers we choose a setup similar to the coupled microcavities. We need two cavities with gain and loss, and a symmetric coupling. The cavities are formed by a system of two mirrors where gain or loss is inserted between them. To introduce a symmetric coupling we place them next to each other where the share one mirror. Choosing a semi-transparent mirror the two cavities couple to each other with a coupling strength defined by the reflectivity of the common mirror. The schematic of this interferometer is shown in Figure 2.4. The mirrors STM1, STMC form the first cavity and SMT2, STMC the second one. Similar to the microcavities the two cavities need to have the same shape and we therefore choose the same distance between their mirrors. We further choose a high reflectivity of the out-coupling mirrors, STM1 and STM2, since they introduce an additional loss to the system. With these restrictions only two parameters, the magnitude of the gain/loss and

the reflectivity of the coupling mirror, remain undefined and will be used to tune the system to an EP. Similar to subsection 2.1.1 the interferometer can also be tuned to a DP. DPs were already studied in optical interferometers, but they are nevertheless interesting for the comparison to the system at an EP.

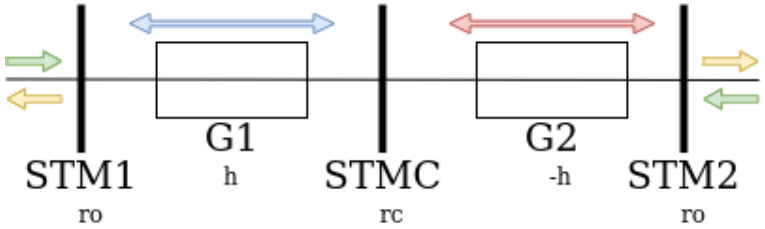


Figure 2.4: Schematic of a coupled cavity system, consisting of three semi-transparent mirrors and two gain/loss media. The first cavity is formed by STM1 and STMC, and the second by STM2 and STMC. Gain is placed inside the first cavity and loss with equal strength in the second one. The coupling mirror, STMC, is responsible for the coupling between the modes. The blue and red arrows represent the modes in the first and second cavity. The green and yellow arrows are the external in and out-coupling of the interferometer at the open ports, which can be monitored in an experiment.

Vanishing eigenvector

The coalescence of two or more eigenvalues of an interferometer is necessary but not sufficient for the occurrence of an EP. Additionally at least one eigenvector has to vanish which is an essential property that occurs at every EP. We therefore rely on this property to find and identify EPs in optical interferometers. In the previous section we introduced a method to obtain the eigenvalues and eigenvectors. To tune the interferometer to an EP we obtain the eigenstates as function of the two undefined parameters r_c and h. Demanding the coalescence of both the eigenvalues and eigenvectors will then lead to the equations which have to be fulfilled to reach an EP.

We start with the eigenvalues that are obtained by solving the eigenvalue equation

of optical interferometers, Equation 2.23, with respect to k,

$$1 - r_c r_o e^{2l(ik+h)} - r_c r_o e^{2l(ik-h)} + r_o^2 e^{4ikl} = 0. \tag{2.25}$$

We did not assign a numerical value to r_o and l yet, but treat them as constants and will not tune them to reach an EP. Demanding the coalescence of the eigenvalues leads to the connection of r_c and h,

$$h = \frac{1}{4l} ln(\frac{1+t_c}{1-t_c}). \tag{2.26}$$

where $t_c = \sqrt{1-r_c^2}$ is the transmission of the coupling mirror. The eigenvalues are now degenerate, which means the system is either at an EP or a DP,

$$k_{1,2} = \frac{i}{2l} ln(r_o). \tag{2.27}$$

Comparing the eigenvalues with the ones in subsection 2.1.1, $\lambda_{1,2} = 0$, only their real part is vanishing but not the imaginary one. This is linked to the additional loss introduced by the out-coupling mirrors and does not occur with fully reflecting mirrors and no out-coupling, $r_o = 1$. It is notable that the eigenvalues no longer depend on r_c itself, only on other parameters of the system.

We again differentiate between vanishing and non-vanishing coupling and compare the geometric multiplicity of the eigenvalues for these cases. Based on the geometric multiplicity we can decide whether the interferometer is at an EP or a DP.

$$\begin{aligned} r_c = 1: &\qquad g_{k_{1,2}} = 2 \\ r_c < 1: &\qquad g_{k_{1,2}} = 1 \end{aligned} \tag{2.28}$$

For vanishing coupling the interferometer is at a DP with one eigenvalue and two eigenvectors. The eigenvectors are the two modes in the cavities and are indepen-

dent of each other. With non-vanishing coupling there is only one eigenvector left and the system is at an EP. The remaining eigenvector is a superposition of the two modes in the cavities and they are no longer independent. We can conclude that EPs can indeed occur in optical interferometers and we will further study their properties.

Eigenvalue splitting

As we will see in the following sections the eigenvalues of an optical interferometer manifest as pole point in the scattering matrix. If the eigenvalues are real, which is only possible with a gain medium, the system can be seen as a laser where the lasing frequencies corresponding to the eigenvalues. Many sensors realized with optical interferometers rely on the behavior of the eigenvalues at perturbations of the system. It is therefore essential to study their behavior. To reinforce our claim of the existence of EPs in optical interferometers we will use perturbations similar to the ones in subsection 2.1.2, an asymmetric perturbation of the eigenfrequencies of the cavities and a symmetric perturbation of the coupling strengths. The real part of the eigenfrequencies of optical cavities depends on the distance between the mirrors. By moving the coupling mirror, STMC, we create a perturbation similar to P_ω. Since the coupling strength in the cavity system only depends on the transmission of STMC we choose $t_c = t_{c0} + dt$ as second perturbation which corresponds to P_g. To obtain the eigenvalues of the perturbed interferometer we use numerical root finding methods to solve the eigenvalue equation, since it is no longer solvable analytically. Comparing the splitting of the eigenvalues, by varying the position of STMC, at an EP and at a DP, shown in Figure 2.5, we obtain the square root behavior at an EP and the linear behavior at a DP. Not only the splitting of the magnitude but also the splitting in the complex plane, shown in Figure 2.6, have the same behavior as in the calculations of quantum

mechanical systems in subsection 2.1.2. This observation is again explained by the correspondence between optical interferometers and quantum mechanical systems.

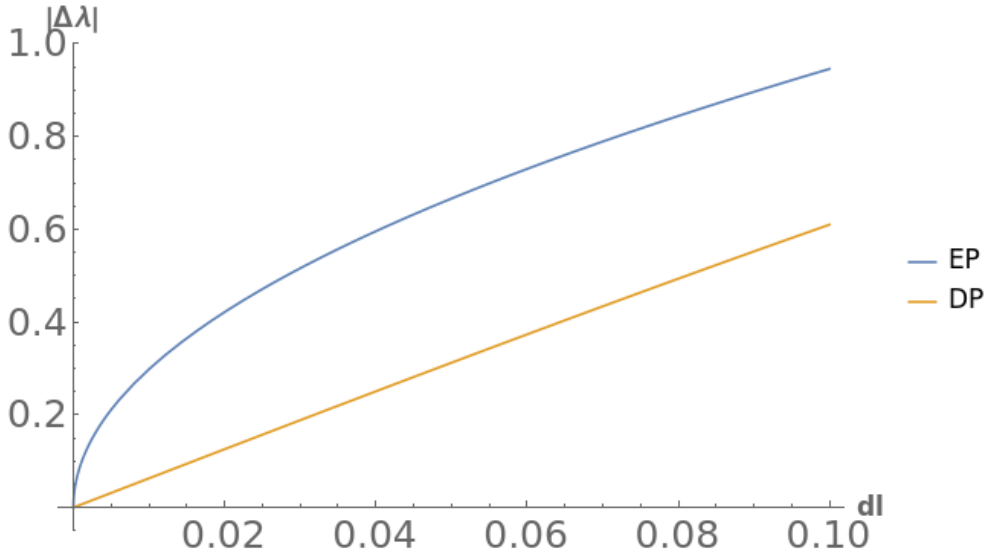
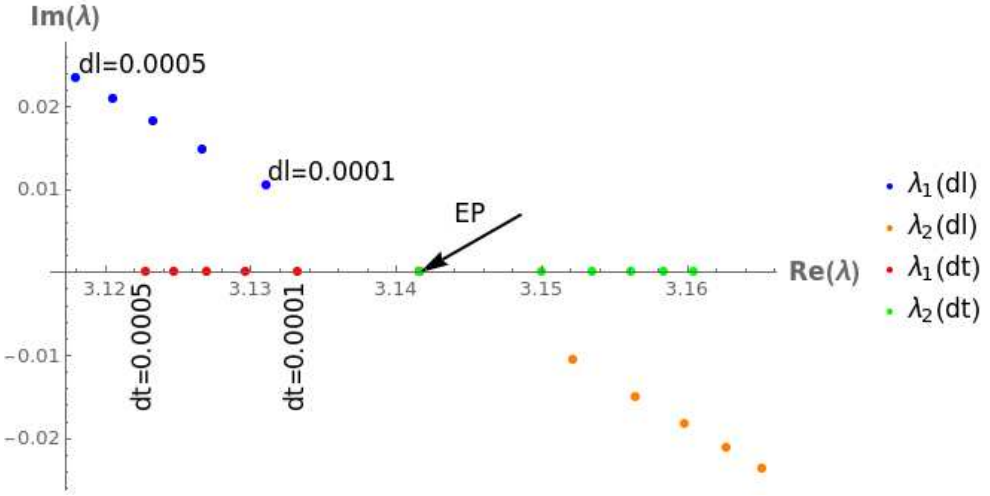


Figure 2.5: Magnitude of the eigenvalue splitting of a double cavity interferometer at an EP and DP, with dl as perturbation. At an EP the eigenvalues have a square root behavior with a high slope for small perturbations. At a DP the splitting is linear.

(a) Eigenvalue splitting in the complex plane at an EP.



(b) Eigenvalue splitting in the complex plane at a DP.

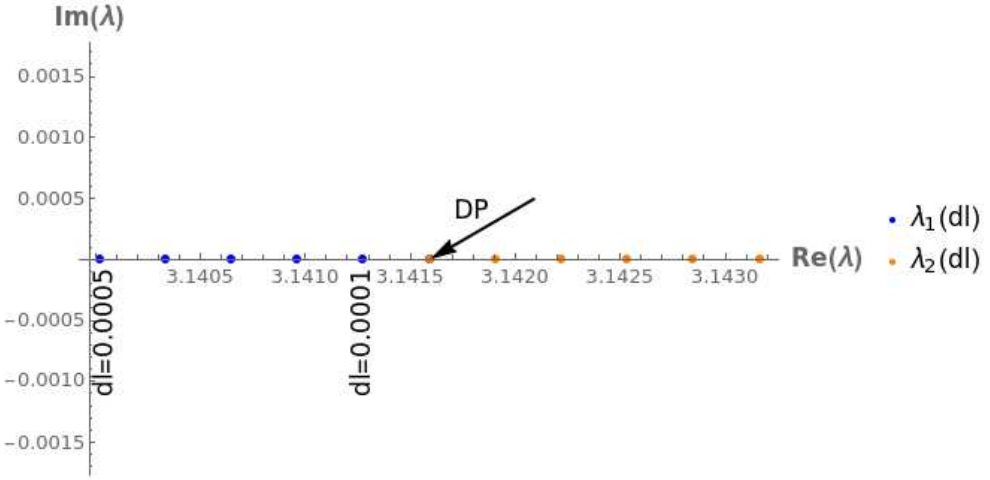


Figure 2.6: Comparison of the eigenvalue splitting of an double cavity interferometer at an EP and a DP in the complex plane. We chose the parameters $r_o = 1$, $l = 1$ and $t_c = 0$ for a DP, $t_c = \frac{1}{\sqrt{2}}$ for an EP. The perturbations dl and dt were varied form 0 to 0.0005 with a step size of 0.0001. The splitting has the same behavior as in Figure 2.2. At a DP the two perturbations have again the same splitting behavior and we therefore only plotted one.

Riemann Sheets

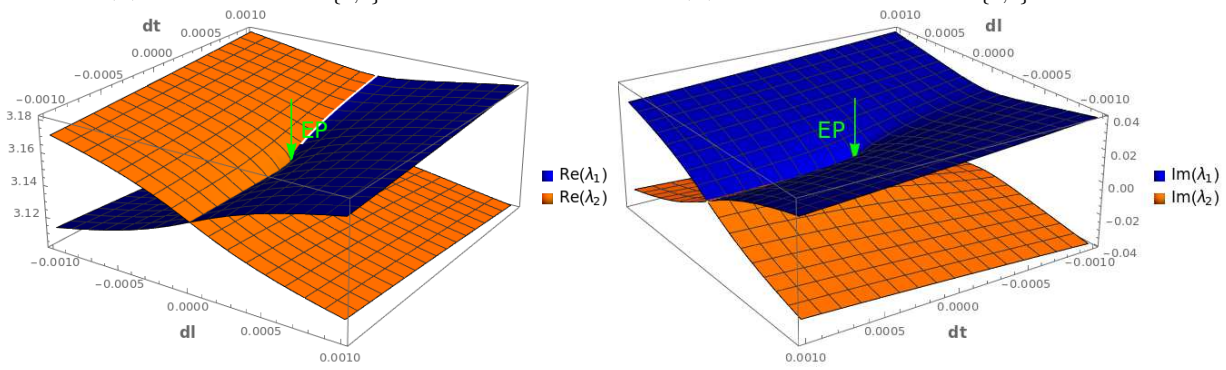
If one is interested to study more complex perturbations it is not sufficient to obtain the behavior with only one perturbation parameter, but one needs to obtain the Riemann sheets of a double cavity system in the two dimensional perturbation space (dl, dt). We are particularly interested in the topology of the Riemann sheets in the vicinity of an EP and a DP, which will be an additional indication of EPs in optical interferometers. The Riemann sheets are again obtained by solving the eigenvalue equation with dl and dt as parameters numerically, and the eigenvalues are plotted in the perturbations space shown in Figure 2.7. As in the predictions in subsection 2.1.3 the Riemann sheets have a square root topology in the vicinity of an EP and are self-intersecting. In the vicinity of a DP the sheets have a conical form and the topology of the absolute function. The high sensitivity of an EP manifests as pole point in the derivative of the sheets located directly at the EP, while the derivative of the sheets in the vicinity of a DP is finite.

Scattering properties

We obtained the behavior of the eigenvalues of a double cavity system under different perturbations. However the eigenvalues themselves can not be directly observed, but manifest in the scattering properties. To look into the impact of the eigenvalues into scattering we solve and discuss the scattering problem of the interferometer. Since the eigenvalues at an EP react highly sensitively to perturbations, it is particularly interesting how they can be used in optical sensing. With the introduced mathematical model we are able to obtain the scattering matrix of the whole system. Since we want to know what impact the eigenvalues have on the scattering problem, we compare the eigenvalue equation to the scattering matrix, which leads to the conclusion that the eigenvalues of the interferometers correspond to pole points in the scattering matrix. Considering that the eigen-

(a) Real part of $\lambda_{\{1,2\}}$ at an EP.

(b) Imaginary part of $\lambda_{\{1,2\}}$ at an EP.



(c) Real part of $\lambda_{\{1,2\}}$ at a DP.

(d) Real part of $\lambda_{\{1,2\}}$ at a DP, Different perspective.

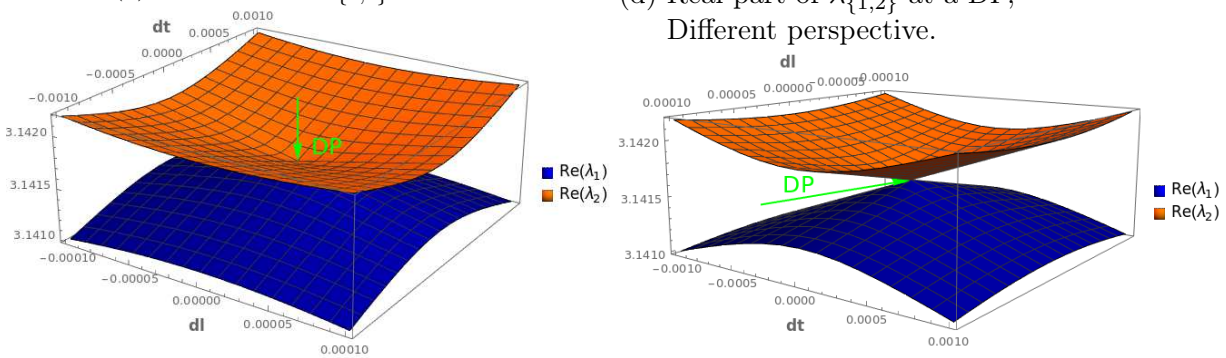
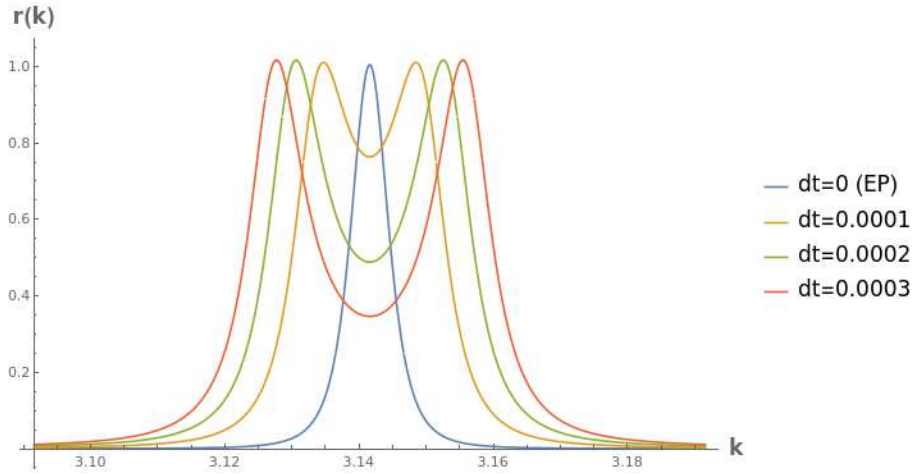


Figure 2.7: Riemann sheets of the double cavity system at an EP and DP, by varying dl and dt. We chose the parameters $r_o = 1$, $l = 1$ and $t_c = 0$ for a DP, $t_c = \frac{1}{\sqrt{2}}$ for an EP.

states correspond to the quasi-bound states, which have outgoing waves without incoming ones, also explains that the eigenvalues correspond to pole points in the scattering matrix. The eigenvalues can therefore be seen as peaks in the transmission or reflectivity of the interferometer. The peaks are located at the real part of the eigenvalues and their width depends on their imaginary part. This can be used to study EPs in interferometers experimentally and confirm their sensitive behavior for perturbations. At an EP the peaks of the reflectivity, shown in Figure 2.8, can be distinguished at a lower perturbation compared to a DP. Since they correspond to the eigenvalues the position of the two peaks also have the characteristic root behavior and the linear behavior at a DP.

(a) Normalized reflectivity at an EP.



(b) Reflectivity at a DP.

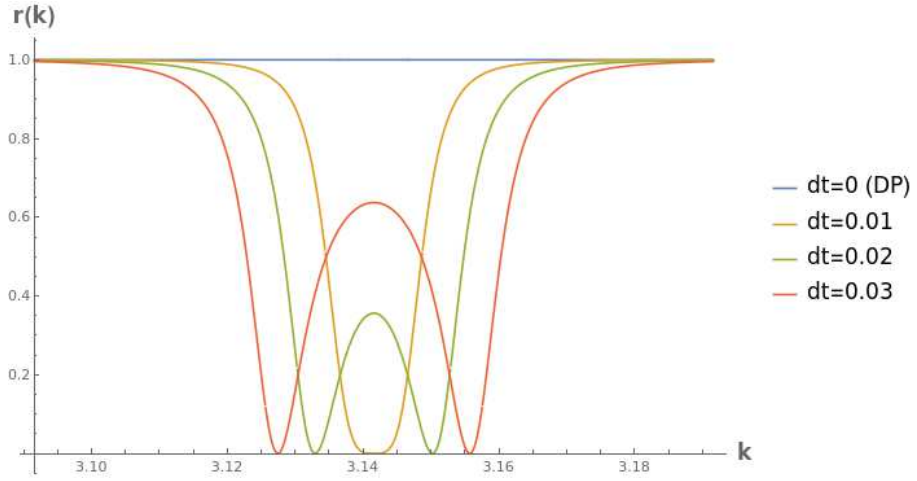


Figure 2.8: Comparison of the reflectivity of a double cavity interferometer at an EP and DP. We chose the parameters $r_o = 0.99$, $l = 1$ and $t_c = 0$ for a DP, $t_c = \frac{1}{\sqrt{2}}$ for an EP. The real part of the eigenvalues manifest as peaks in the reflectivity, and can be observed experimentally. Compared to a DP at an EP a lower perturbation is sufficient to observe the two different peaks. The reflectivity at an EP is normalized since due to the gain medium it can exceed 1.

2.2.3 Ring cavity

We introduced EPs into optical interferometers by the example of the double cavity system, but we still needed an active gain medium to bring the system to an EP. To make the experimental setup even easier we need to introduce an interferometer which can be brought to an EP without the need of a gain medium. For this purpose we look at the Hamiltonian at an EP, given by Equation 2.6:

$$\begin{aligned} H_{\mathrm{EP}} &= \begin{pmatrix} s_+ - i\sqrt{g_{12}g_{21}} & g_{12} \\ g_{21} & s_+ + i\sqrt{g_{12}g_{21}} \end{pmatrix} \\ \lambda_{1,2} &= s_+ \\ v_1 &= \frac{1}{\sqrt{g_{12}+g_{21}}} \begin{pmatrix} \sqrt{g_{12}} \\ i\sqrt{g_{21}} \end{pmatrix} \end{aligned} \tag{2.29}$$

The reason why we needed equal gain and loss values in each cavity is explained by the diagonal elements of H_{EP}. For non-vanishing coupling strengths the imaginary parts have the same magnitude with different sign. On the other side both coupling strengths vanishing will just lead to a DP. The interesting case is an asymmetric coupling where only one coupling strength vanishes, $g_{21} = 0, g_{12} \neq 0$, and the other one persists. In this case the diagonal elements have the same value but the system is still at an EP. Further the eigenvector v_1 reveals that in this case the eigenstate of the EP is not a superposition of both modes, like in the case of symmetric coupling. Only one mode, which is the eigenstate of the system, persists and the other one vanishes. In other words we are looking for a system with two modes, with equal eigenfrequencies, and an asymmetric coupling mechanism. Such systems were already realized and studied in experiments with microcavities[29,30,24,31]. A microcavity can be seen as two mode system with a clockwise (CW) and counter-clockwise (CCW) mode. Without any coupling between them the microcavity is

at a DP. Placing two nano scatterers near the microcavity, with a specific angle between them, introduces an asymmetric coupling between the modes, and the microcavity is at an EP. Only the mode which is not backscattered remains as an eigenstate and the other one vanishes. This sense of direction is a consequence of the chirality of EPs[29]. The experimental setup of microcavities is fairly difficult, since once manufactured they are not tunable anymore. We therefore ask the question if there exists a simpler setup where EPs and their chiral behavior can be observed, and are looking into optical interferometers. To create an interferometer similar to a microcavity, we choose a ring cavity as shown in Figure 2.9. The ring cavity has a clockwise (CW) and counter-clockwise (CCW) mode. Without any coupling between them, the ring cavity is at a DP similar to the microcavity. To reach an EP we need to introduce an asymmetric coupling mechanism, which is achieved by a back-reflecting mirror (BRM). The back-reflecting mirror only couples the CCW mode into the CW one. Due to the similarity of the two systems, we expect to be able to tune the ring cavity to an EP and observe the chiral and directional behavior.

The ring cavity we are looking into, shown in Figure 2.10, consist of two mirrors and a beam-splitter which form the ring cavity itself. Similar to a whispering-gallery resonator the cavity has a clockwise (CW) and counter clockwise (CCW) mode. Since the both modes in this system propagate along the same paths but in opposite direction, they have the same eigenfrequencies. The two back-reflecting mirrors, BRM1 and BRM2, introduce an asymmetric coupling between the modes. At the mirror BRM1 the CW mode couples into the CCW mode and vice versa at the mirror BRM2. Since we demand one vanishing coupling strength, only one of the back-reflecting mirrors is present at an EP, which we choose to be BRM1. The other one (BRM2) will act as perturbation of the system to again study the behavior of the eigenvalues. As it turns out, this ring cavity was already studied

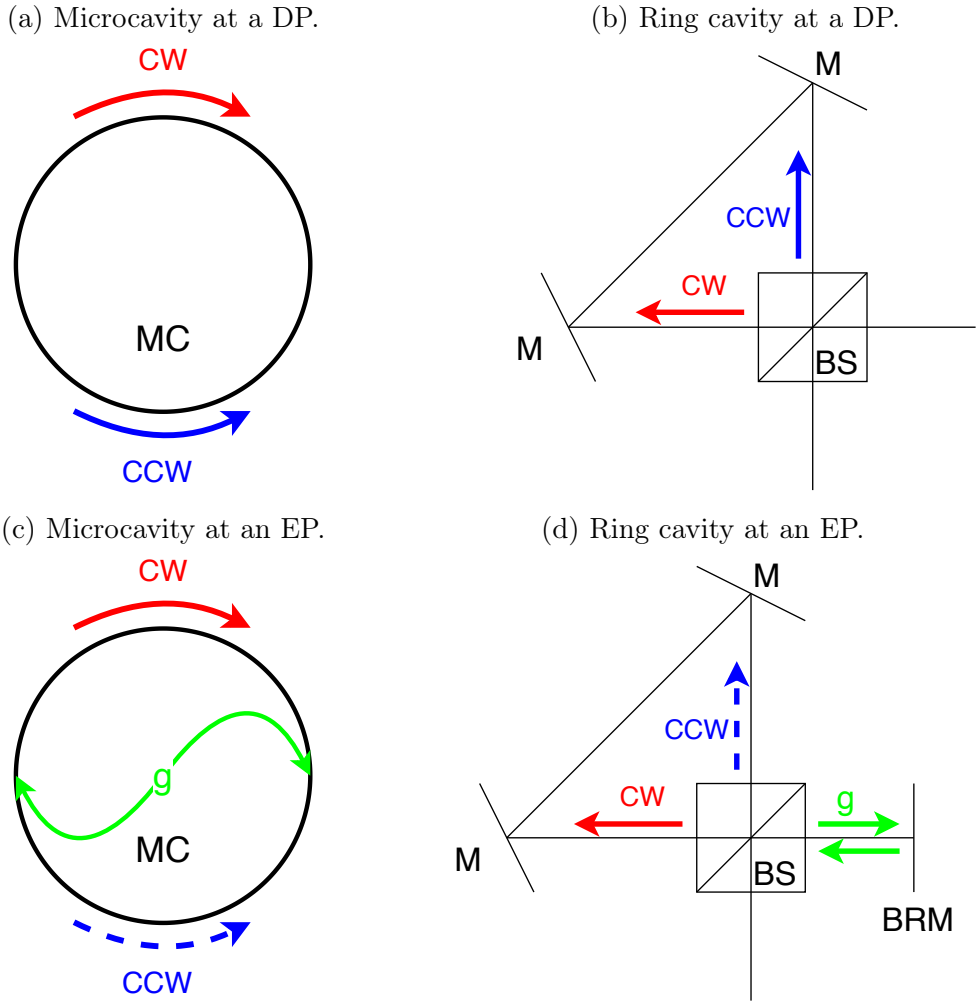


Figure 2.9: Schematic of a microcavity (MC) and a ring cavity at a DP and an EP. Both cavities have two modes, a clockwise (CW) and a counter-clockwise (CCW) mode, represented by the red and blue arrows. With no coupling mechanism (g) the modes are independent of each other and are the eigenstates of the systems. The systems are at a DP. By introducing an asymmetric coupling from the CCW mode into the CW mode, represented by the green arrows, the cavities are at an EP where only the CW mode remains as eigenstate.

experimentally, under the aspect of phase locking of the laser modes, where the vanishing of the coupling mode was observed[12]. We will study this interferometer with our mathematical model and link this observation to EPs, and further obtain again the behavior of the eigenvalues.

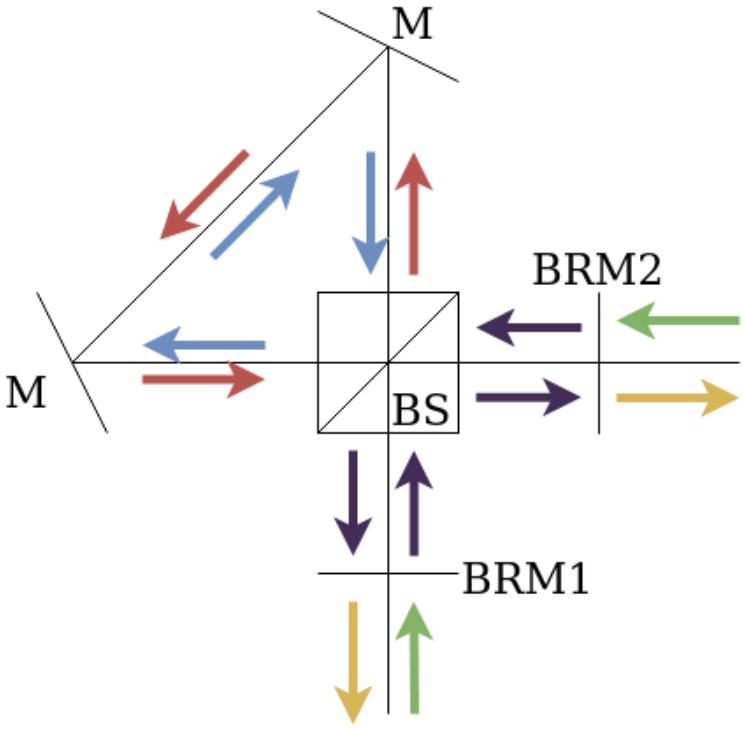


Figure 2.10: Schematic of a ring cavity. The cavity is formed by the two mirrors and the beam-splitter, which is also responsible for the in- and out-coupling. The blue and red arrows represent the clockwise (CW) and counter clockwise (CCW) modes. The green and yellow arrows at the open ports are the incoming and outgoing waves. The two back-reflecting mirrors, BRM1 and BRM2, introduce an asymmetric coupling between the modes, illustrated by the purple arrows.

Vanishing eigenvector

To study the occurrence of an EP in the ring cavity we again use the property of the vanishing eigenvector. Without BRM1 and BRM2, $r_1 = 0$ and $r_2 = 0$, the

CW and CCW mode are independent of each other since there is no coupling. Both propagate along the same path with opposite direction and have therefore the same eigenfrequencies. The ring cavity is at a DP, with both the CW and CCW mode as eigenstates. Bringing in one back-reflecting mirror, BRM1, with an arbitrary reflectivity $r_1 \neq 0$, leads to a coupling from the CW mode into the CCW one. The CCW mode is not effected by this mirror at all. Looking at the Hamiltonian and eigenvector of the corresponding quantum mechanical system, H_{EP} and v_1 with $g_{21} = 0, g_{12} \neq 0$, we would expect that the interferometer is at an EP with the CW mode as eigenvector. To confirm this explanation we obtain the eigenstates with our mathematical model of interferometers. The eigenvalue equation of the ring cavity has a very simple form:

$$(1 + e^{ikl} r_{bs})^2 = 0 \tag{2.30}$$

It is particular interesting that the eigenvalues are not affected by the value of r_1, but it affects the geometric multiplicity.

$$\begin{aligned} r_1 = 0 : &\qquad g_{k_{1,2}} = 2 \\ r_1 > 0 : &\qquad g_{k_{1,2}} = 1 \end{aligned} \tag{2.31}$$

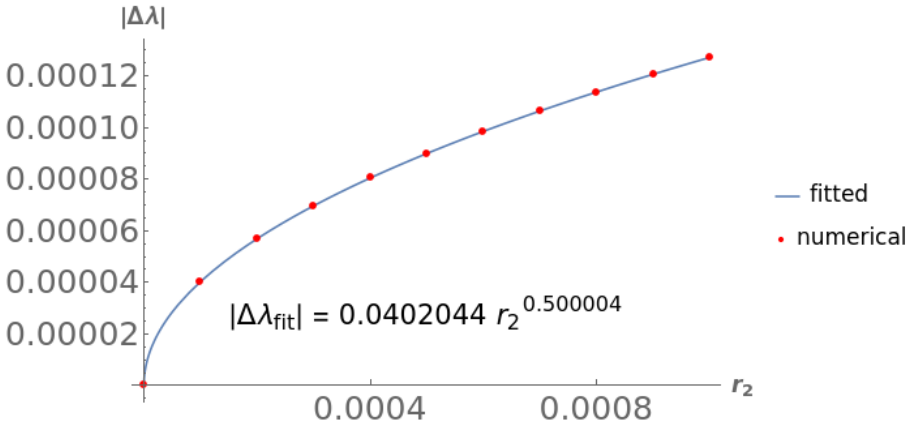
With an arbitrary non-vanishing reflectivity r_1 the system is at an EP. Solving the eigenvector equation, Equation 2.23, reveals that the remaining eigenvector is indeed the CW mode, which is in agreement with the explanation of the corresponding quantum mechanical system. Compared to the double cavity, the ring cavity is easily brought to an EP and does not need a gain medium or tunable mirrors. This leads to a simple experimental setup. This ring cavity setup was already studied in an experiment, where a high suppression of the donor (in our case CCW) mode was observed[12]. The suppression was 4 to 5 orders of magni-

tude higher than previously reported. This observation is now explained with the vanishing mode at an EP, and is linked to the chiral behavior of the system.

Eigenvalue splitting

Due to the simple setup of the ring cavity, which makes it suitable for experiments, we want to look into the behavior of their eigenvalues. We already concluded that the reflectivity of BRM1 does not lead to a splitting of the eigenvalues. We bring in a second back-reflecting mirror, BRM2, with its reflectivity, r_2 as perturbation parameter. The additional back-reflection can for example be interpreted as reflection of a particle and the interferometer is therefore suitable for optical sensing. Since both of the modes are coupling to each other, while none has gain or loss, the interferometer can no longer be at an EP. To obtain the behavior of the interferometer we solve the eigenvalue equation with r_2 as parameter. Since we need to solve it numerically we need a fitting function to meaningfully describe its behavior. We choose the function $f(x) = ax^b$ with a and b as parameters, since it can describe both, the behavior at an EP and at a DP. The eigenvalue splitting, shown in Figure 2.11, has a square root behavior as expected for EPs. In the complex plane the eigenvalues split along the imaginary axis. This can be obstructive since the corresponding peaks in the scattering matrix are at the same position with a different width, which could be difficult to distinguish.

(a) Magnitude of the eigenvalue splitting of a ring cavity at an EP.



(b) Splitting of the eigenvalues of a ring cavity in the complex plane at an EP.

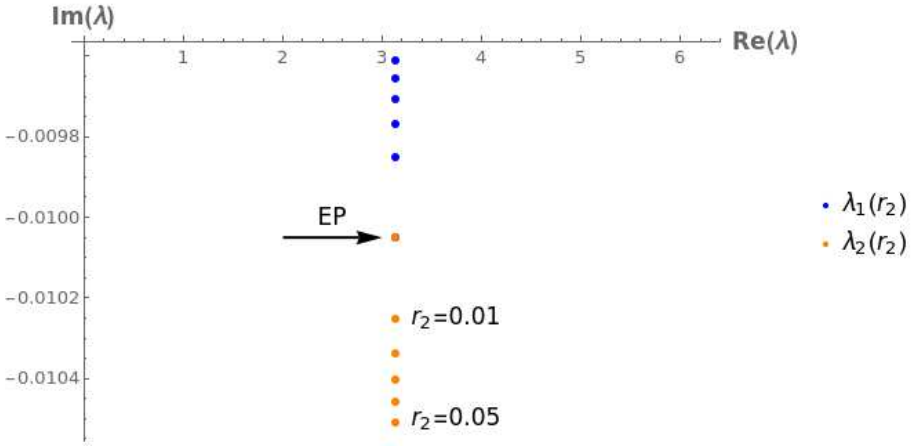
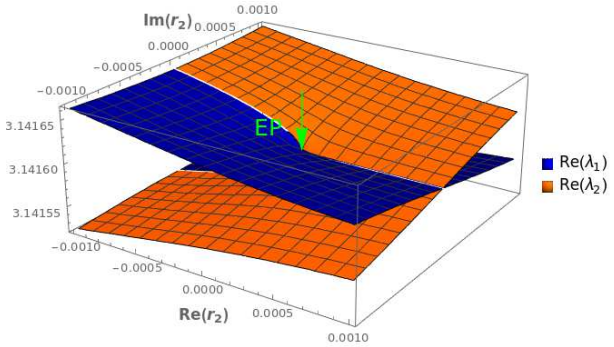


Figure 2.11: Eigenvalue splitting of a ring cavity at an EP, with the parameters $r_{bs} = 0.99$, $r_1 = 0.01$ and $l = 1$. The fitting function reveals the square root behavior of the splitting. The eigenvalues in the complex plain are only splitting along the imaginary axis, which could be difficult to observe in the scattering properties.

Riemann Sheets

To reinforce the identification of the EP in the ring cavity, we again look into the topology of the Riemann sheets. For this purpose we extend the parameter r_2 into the complex plane which leads to a perturbation of the magnitude and phase of the coupling strength. Since the phase of the coupling is also depending on the distance between the back-reflecting mirror and the beam-splitter, it can also be interpreted as perturbation of the position of the mirror BRM2. The Riemann sheets, shown in Figure 2.12, have the topology of the complex square root function which is characteristic for an EP.

(a) Real part of $\lambda_{1,2}$ at an EP.



(b) Imaginary part of $\lambda_{1,2}$ at an EP.

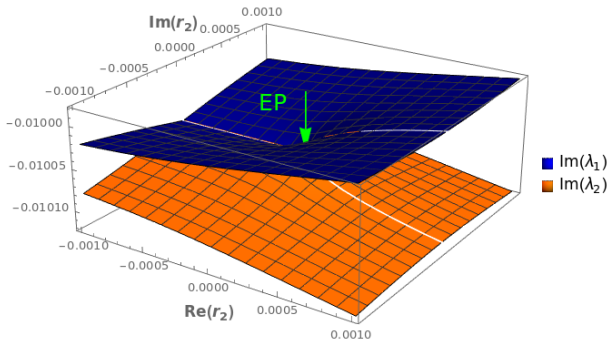


Figure 2.12: Riemann Sheets of the ring cavity interferometer, with the parameters $r_{bs} = 0.99$, $r_1 = 0.01$ and $l = 1$, with the characteristic square root topology in the vicinity of an EP.

2.2.4 Higher order Exceptional Points

Higher order exceptional points can occur when not only two but more eigenvalues are degenerate[32,33,34]. In general an EP has the order m when m degenerated eigenvalues have one corresponding eigenvector. Higher order Exceptional Points where already studied in experiments with microcavities, where an enhanced sensitivity was observed[25,35]. Generally speaking the eigenvalue splitting at an EP of order m has an m-th root behavior. They can be used to create sensors with an extraordinarily high sensitivity. Since the ring cavity was easily brought to an EP we now consider a system of two coupled ring cavities, show in Figure 2.13. This system was also already considered by the same group as the single ring cavity.[12] The two cavities are symmetric and have the same properties. The interferometer has four eigenmodes, a clockwise (CW) and counter clockwise (CCW) mode in each cavity. We will refer to them as CW1, CCW1, CW2 and CCW2. Due to the geometric placements of the ring cavities some of the modes have an intrinsic coupling, from CCW1 into CCW2 and from CW2 into CW1. Comparing this coupling to the single ring cavity we expect that the interferometer is already at an EP without any back-reflection. The two tunable back-reflecting mirrors, BRM1 and BRM2, introduce an additional coupling and the interferometer can be brought to an EP of order four.

Vanishing eigenvector

To study the occurrence of an EP we will look into the corresponding quantum mechanical system. This system has four coupling modes, which correspond to CW1, CW2, CCW2 and CCW1, and is described by a 4x4 Hamiltonian. As in the previous section the eigenfrequencies of the two modes of one ring cavity are the same. Further assuming identical properties of the two ring cavities leads to the same eigenfrequencies of all four modes denoted by s_0. Looking at the schematic

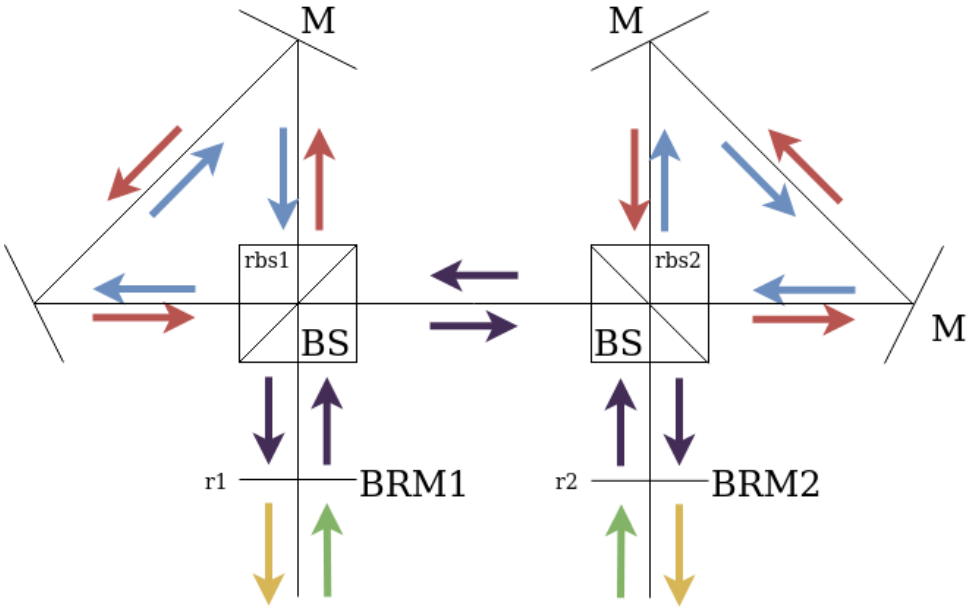


Figure 2.13: Schematic of the coupled ring cavity interferometer. It consists of two separate ring cavities as in Figure 2.10. The blue and red arrows indicate the clockwise (CW) and counter clockwise (CCW) modes in each cavity. The green and yellow arrows represent the in- and out-coupling of the interferometer at the open ports. The purple arrows indicate coupling between modes. Beside the coupling introduced by the back-reflecting mirrors this setup has also an intrinsic coupling from CCW1 into CCW2 and from CW2 into CW1, which can not be eliminated.

of the coupling ring cavities, Figure 2.13, we see an intrinsic coupling from CCW1 into CCW2 and also from CW2 into CW1, which we assign a coupling strength of g_i. The two back-reflecting mirrors, BRM1 and BRM2, introduce an additional coupling which we neglect for now. With these considerations we can obtain the Hamiltonian,

$$H = \begin{pmatrix} s_0 & g_i & 0 & 0 \\ 0 & s_0 & 0 & 0 \\ 0 & 0 & s_0 & g_i \\ 0 & 0 & 0 & s_0 \end{pmatrix}. \tag{2.32}$$

Since the Hamiltonian is an upper triangle matrix the eigenvalues are the diagonal elements, and therefore all four are still equal. It is also similar to a Jordan normal form which allows us to obtain the two eigenvectors, CW1 and CCW2. The two remaining eigenmodes do not couple into any other mode. This Hamiltonian describes a system with two EPs of order two. This observation was already predicted by the Davidson group as they stated that the coupling ring cavities are the "unfolded" versions of the single ring cavity with a back-reflecting mirror[12]. Since the single ring cavity with a back-reflecting mirror is at an EP of order two and the coupling ring cavities are at two EPs of order two this is in perfect agreement. To study EPs of higher order we need to tune the ring cavities and introduce an additional coupling mechanism, BRM2. By again considering all coupling mechanisms we can obtain the Hamiltonian,

$$H = \begin{pmatrix} s_0 & g_{12} & g_{13} & g_{14} \\ 0 & s_0 & g_{23} & g_{24} \\ 0 & 0 & s_0 & g_{34} \\ 0 & 0 & 0 & s_0 \end{pmatrix}. \tag{2.33}$$

The form of an upper triangle matrix of the Hamiltonian lets us again conclude that all four eigenvalues are still degenerate, and due to the similarity to a Jordan normal form only one eigenvector, CW1 is remaining. The system is now at an EP of order four.

We again want to compare the results described by the quantum mechanical system to our introduced model. The eigenvalue equation of the coupled ring cavities, with only one back-reflecting mirror BRM2, has a very similar form compared to the single ring cavity.

$$\begin{aligned} (1 + e^{ikl} r_{bs})^4 &= 0 \\ k_{1,2,3,4} &= \frac{\pi}{l} + i\frac{ln(r_{bs})}{l} \end{aligned} \tag{2.34}$$

All four eigenvalues are degenerated, and we need the geometric multiplicity to clarify whether it is an EP or a DP. Further solving the eigenvector equation leads to the remaining modes.

$$\begin{aligned} r_2 = 0: &\qquad g_{k_{1,2,3,4}} = 2 \\ r_2 > 0: &\qquad g_{k_{1,2,3,4}} = 1 \end{aligned} \tag{2.35}$$

The coupled ring cavities have only one eigenvalue. Without a back-reflecting mirror the modes CW1 and CCW2 remain as eigenvectors and the interferometer has two EPs of order two. Bringing in one back-reflecting mirror, BRM2, CCW2 also vanishes and the interferometer is at an EP of order four. The only remaining eigenvector is the CW1 mode. This is in agreement with the explanation of the corresponding quantum mechanical system.

Eigenvalue splitting

Since the interferometer is at an EP of order four we expect a fourth root behavior of the eigenvalue splitting compared to the square root behavior at an EP of order two. Directly at an EP the slope of the splitting is infinite in both cases, but in its vicinity it is higher for an EP of fourth order. This enhanced sensitivity was

experimentally observed in microcavities[25] and we will introduce it into optical interferometers.

Similar to the case of the single ring cavity, we look at the coupled ring cavities with one back-reflecting mirror, BRM2. The interferometer is at an EP of order four and we introduce a second back-reflecting mirror, BRM1, as perturbation. The eigenvalue splitting follows indeed a fourth root behavior, shown in Figure 2.14. Since in both interferometers, single ring cavity and coupled ring cavities, we use the reflectivity of a mirror as perturbation parameter, we can compare the magnitude of the splitting. For this purpose we choose the same reflectivity of the beam-splitter and the mirrors in both interferometers.

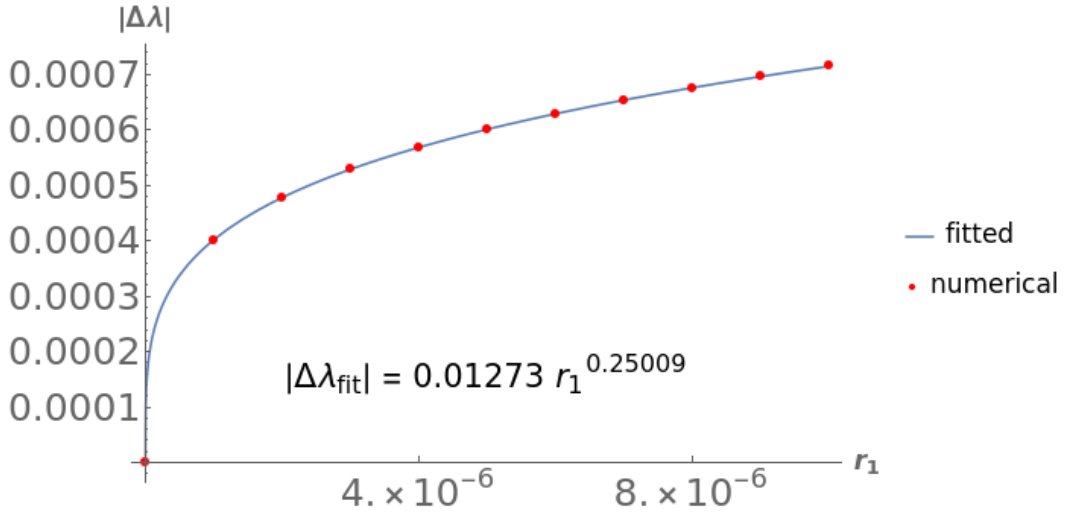
$$\begin{aligned} r_{bs} = 0.99 \quad r_1 = 0.01 \quad r_2 = 0.001 \\ \text{single ring cavity:} \quad |\Delta\lambda| = 0.00012713 \\ \text{coupled ring cavities:} \quad |\Delta\lambda| = 0.0022679 \end{aligned} \tag{2.36}$$

The splitting of a coupled ring cavities interferometer is of larger magnitude and the sensitivity is enhanced.

Riemann Sheets

For a better understanding of an EP of order four it is again crucial to study the Riemann sheets. Similar to the single ring cavity we extend the perturbation parameter r_1 into the complex plane, and the interpretation of the imaginary part as perturbation of the position of mirror BRM1 remains. Since four eigenvalues are splitting up, the Riemann sheets consist of four surfaces rather than two. The sheets, shown in Figure 2.15, have the topology of the complex fourth root function and reveal that the interferometer is indeed at an EP of order four. They are self-intersecting and have a pole-point in their derivative located at the EP.

(a) Magnitude of the eigenvalue splitting of coupled ring cavities.



(b) Eigenvalues splitting of coupled ring cavities in the complex plane.

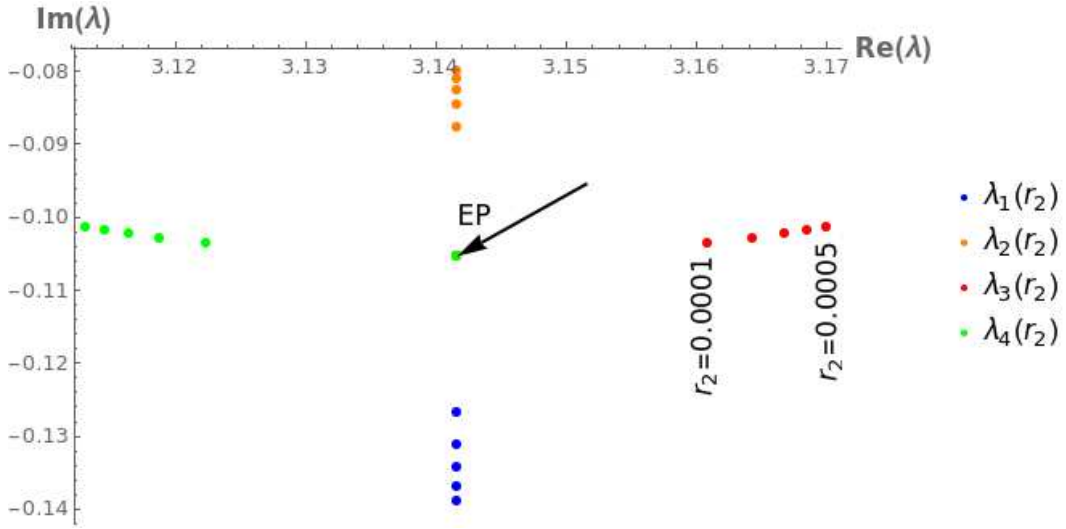
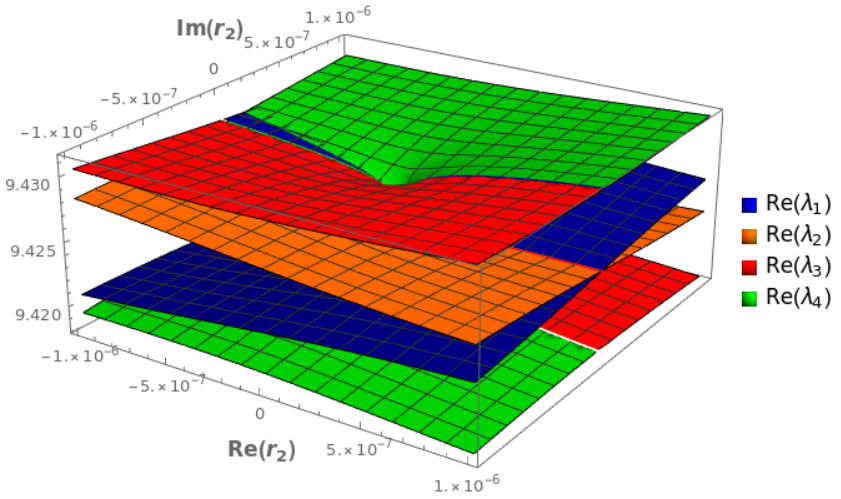


Figure 2.14: Eigenvalue splitting of coupled ring cavities at an EP of order four, with r_1 as perturbation parameter. We chose the parameters $r_{bs} = 0.99$, $r_2 = 0.01$. The fitting function of the magnitude of the splitting reveals the fourth root behavior.

(a) Real part of $\lambda_{1,2,3,4}$ at an EP.



(b) Imaginary part of $\lambda_{1,2,3,4}$ at an EP.

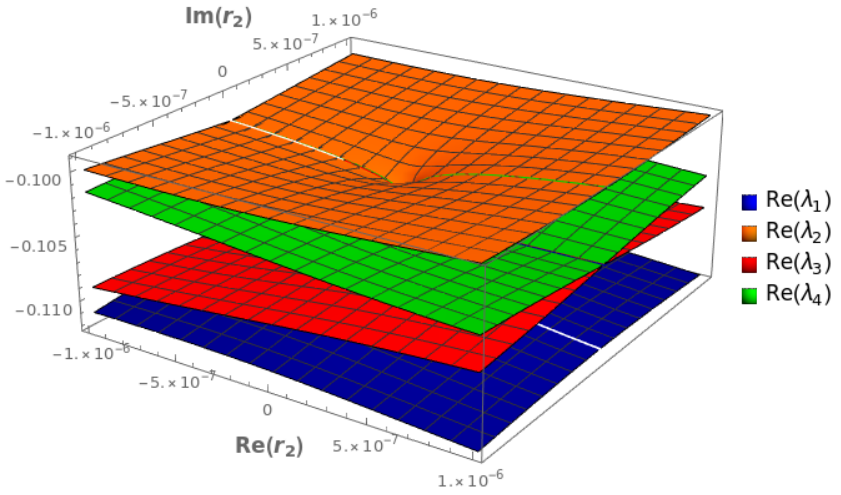


Figure 2.15: Riemann Sheets of the coupled ring cavities at an EP of order four, plotted in the perturbation space $(Re(r_1), Im(r_1))$. We chose the parameters $r_{bs} = 0.99$, $r_2 = 0.01$. All four sheets have the topology of the complex fourth root.

2.3 Conclusion

While EPs are already well-studied, their occurrence in optical interferometers was not. They are particularly interesting systems to study EPs due to their straightforward setup. We introduced a mathematical model of interferometers in general which is able to find and identify EPs. This method relies on the isomorphism of the paraxial wave equation and the Schrödinger equation. We introduced an interferometer, a double cavity, with a similar setup as the known microcavity system, where EPs were already observed experimentally. We could show that one is able to tune the system to an EP and obtain all properties linked to it, namely the vanishing eigenvector, the characteristic splitting of the eigenvalues and the Riemann sheets. By further looking into an interferometer system proposed by the Davidson group[12], an optical ring cavity, we could show that their obtained results can directly be linked to EPs. This interferometer could be tuned to an EP without any gain medium or tunable mirrors which results in an easy experimental setup, which could be built on an optical table. Sensors which are realized with optical interferometers could benefit of the high sensitivity of EPs and their characteristic eigenvalue splitting. A particular interesting sensor is the so-called Sagnac interferometer, which has the same geometry as the introduced ring cavity[14,15,13]. In a rotating frame the optical wave of the clockwise and counter clockwise direction experience different lengths of the cavity. An EP can be exploited to enhance the response of the system to these perturbations, which leads to a better resolution of the sensor. We further introduced higher order EPs into optical interferometers on the example of the coupled ring cavities, where the sensitivity is even enhanced compared to the single ring cavity. This setup is still simple and suitable for experiments, with the only remaining challenge being the construction of two ring cavities of the same geometry. In theory an arbitrary amount of ring cavities can be coupled, which allows to create an optical interferometer at an EP of arbitrary

high order while maintaining the property of a simple setup.

3 Inverse Scattering Problem

The inverse scattering problem[36] is about obtaining the internal structure and parameters of a system by only knowing its scattering properties. The inverse scattering problem has a wide range of applications not only in physics but also in medicine and many other fields[37]. It is a particularly hard problem to solve since most models for forward scattering are not invertible and the solutions are not unique in many cases[38,39]. We will solve the inverse scattering problem of optical resonators containing a dielectric potential landscape. With the canonical isomorphism between an optical and a quantum mechanical system, see subsection 2.2.1, the dielectric potential landscape can be transformed into a quantum mechanical potential and it is sufficient to solve the problem for the latter one. We therefore look into the inverse scattering problem of a one-dimensional quantum mechanical system. This problem was also addressed in a different project of our group where a machine learning approach was chosen. Machine learning is a rather new and unconventional approach and we therefore want to compare it to conventional approaches. The machine learning approach solves the inverse scattering problem by training a neural network to generated samples. It can be seen as data fitting in a high dimensional space. For comparison we choose conventional approaches which work in a similar way. We therefore transform the inverse scattering problem into a fitting problem by introducing an error function, and minimize the error with numerical methods. The two chosen methods for minimization are gradient

descent and simulated annealing, which are both well-known in physics. Gradient decent[16] is a deterministic and easy-to-implement method but can suffer from poor convergence when local minima exist. Simulated annealing[40] on the other hand is a stochastic method with a more complex implementation, but resolves the problem of local minima. In this section we will give an introduction to the inverse scattering problem, choose a numerical model, transform it into a minimization problem and finally solve it with the named methods.

3.1 Mathematical model

We first look into the scattering problem of quantum mechanical systems in general. For simplicity, we restrict ourselves to one-dimensional smooth potentials with a compact carrier. To describe the scattering properties of such system we introduce the scattering matrix. The scattering matrix describes the transmission and reflectivity of an incoming wave depending on its momentum. It is obtained by solving the stationary Schrödinger equation in natural units:

$$(\frac{\partial^2}{\partial x^2} - 2mV(x) + k^2)\psi_k(x) = 0 \tag{3.1}$$

with the potential $V(x)$ and the momentum of the wave k. The momentum is directly linked to the energy with the relation $E = \frac{k^2}{2m}$, where m denotes the mass of the particle. We choose the boundary conditions of asymptotic incoming and outgoing plane waves:

$$\begin{aligned} x \to -\infty : \quad & \psi_k(x) = A_{1k}e^{ikx} + B_{1k}e^{-ikx} \\ x \to \infty : \quad & \psi_k(x) = B_{2k}e^{ikx} + A_{2k}e^{-ikx} \end{aligned} \tag{3.2}$$

Here A_{1k}, B_{1k} denotes the amplitudes of the incoming and outgoing waves on the left side and A_{2k}, B_{2k} on the right one. The solution of the Schrödinger equation links the amplitudes together and leads to the scattering matrix:

$$\begin{pmatrix} B_{1k} \\ B_{2k} \end{pmatrix} = S(k) \begin{pmatrix} A_{1k} \\ A_{2k} \end{pmatrix} \tag{3.3}$$

The elements of the scattering matrix are the reflectivity of each side and the transmission from one side to the other:

$$S(k) = \begin{pmatrix} r_1(k) & t_{21}(k) \\ t_{12}(k) & r_2(k) \end{pmatrix} \tag{3.4}$$

The forward scattering problem can be expressed as function which links a given potential to its scattering matrix.

$$S = f_{FSP}(V) \tag{3.5}$$

As the name suggests the inverse scattering problem can be expressed as the inverse function of the forward scattering problem and links a given scattering matrix to its potential.

$$V \in f_{ISP}(S) = \{V' : f_{FSP}(V') = S\} \tag{3.6}$$

In general the forward scattering problem is not injective, e.g. if the potential allows boundary states, and therefore the solution of the inverse scattering problem may not be unique. With only the information of the scattering matrix, we can not decide which one of the solutions is the true potential and we therefore are satisfied with any potential where the corresponding scattering matrix is equal to the given one.

To solve the inverse scattering problem numerically we introduce a finite element model. We transform the continuous spaces of position and momentum into discrete finite subspaces. We denote variables in the finite element model with a tilde symbol ($\tilde{}$). For the position space we choose a finite interval $[\tilde{x}_{min}, \tilde{x}_{max}]$ with an equidistant grid of $\tilde{n}_x$ points. This translates to a position space of $\tilde{x} = (\tilde{x}_1, \ldots, \tilde{x}_{\tilde{n}_x})$, where $\tilde{x}_1 = \tilde{x}_{min}$, $\tilde{x}_{\tilde{n}_x} = \tilde{x}_{max}$ and the distance of two points is given by $\Delta\tilde{x} = \tilde{x}_{i+1} - \tilde{x}_i = \frac{\tilde{x}_{max} - \tilde{x}_{min}}{\tilde{n}_x - 1}$. The size of the interval depends on the

compact carrier of the true potential and should be big enough to include it. The number of grid points heavily depends on the slope of the potential and should be chosen high enough not to lose information. The potential in this model is represented by the potential on the finite position space $\tilde{V} = (\tilde{V}_1, \ldots, \tilde{V}_{\tilde{n}_x})$ with $\tilde{V}_i = V(\tilde{x}_i)$. For the momentum space we choose a similar approach where we choose an interval $[0, \tilde{k}_{max}]$ with an equidistant grid of $\tilde{n}_k$ points, with $\tilde{k}_1 = 0$, $\tilde{k}_{\tilde{n}_k} = \tilde{k}_{max}$ and the distance $\Delta\tilde{k} = \frac{\tilde{k}_{max}}{\tilde{n}_k - 1}$ between two points. For the inverse scattering problem the scattering matrix in the momentum space is given and therefore the parameters of the momentum space can be chosen accordingly. The scattering matrix in this model is again the scattering matrix on the finite momentum space $\tilde{S} = (\tilde{S}_1, \ldots, \tilde{S}_{\tilde{n}_k})$ with $\tilde{S}_i = S(\tilde{k}_i)$. Similar to the continuous model we can express the forward and inverse scattering problem as functions.

$$\begin{aligned} \tilde{S} &= \tilde{f}_{FSP}(\tilde{V}) \\ \tilde{V} &\in \tilde{f}_{ISP}(\tilde{S}) = \{\tilde{V}' : \tilde{f}_{FSP}(\tilde{V}') = \tilde{S}\} \end{aligned} \tag{3.7}$$

While the definition of the forward scattering problem is still valid, a problem arises for the inverse scattering problem. Due to the finite and discrete position and momentum spaces, there may not exist a potential with a scattering matrix equal to the given true one. We will resolve this problem in the following section, where we define the inverse scattering problem as a minimization problem and its solution is the potential whose scattering matrix is "nearest" to the true one.

3.2 Minimization problem

To solve the inverse scattering problem numerically we transform it into a minimization problem. For this purpose we introduce an error function which describes the disagreement between the scattering matrix of a potential, V, and the

true scattering matrix, S_{true}. The error function should be positive-semidefinite, $\epsilon(V, S_{\text{true}}) \geq 0$, and only zero if the potential is a solution to the inverse scattering problem, $\epsilon(V, S_{\text{true}}) = 0 \Leftrightarrow f_{FSP}(V) = S_{\text{true}}$. In the continuous model we define the error function as:

$$\epsilon(V, S_{\text{true}}) = \int_{k=0}^{\infty} ||(f_{FSP}(V))(k) - S_{\text{true}}(k)||_F dk \tag{3.8}$$

Here $||\cdot||_F$ denotes the Frobenius norm, which is the sum of the absolute squares of all elements of the matrix. In the finite element model the integral transforms into a sum and the error function takes the form:

$$\tilde{\epsilon}(\tilde{V}, \tilde{S}_{\text{true}}) = \sum_{i_k=1}^{\tilde{n}_k} ||(\tilde{f}_{FSP}(\tilde{V}))_{i_k} - (\tilde{S}_{\text{true}})_{i_k}||_F \Delta\tilde{k} \tag{3.9}$$

With this error function we can write the inverse scattering problem, with a given scattering matrix $\tilde{S}_{\text{true}}$, as minimization problem:

$$\tilde{\epsilon}(\tilde{V}, \tilde{S}_{\text{true}}) = \min_{\forall \tilde{V}'}(\tilde{\epsilon}(\tilde{V}', \tilde{S}_{\text{true}})) \tag{3.10}$$

with the solution $\tilde{V}$ where the error is minimal. The inverse scattering problem is now again well defined. To solve it we need two different algorithms, one for the forward scattering problem and one for the minimization.

3.3 Forward scattering problem

Due to the nature of the finite element model we follow the divide and conquer paradigm to solve the forward scattering problem. We will divide the potential into small subspaces, which was already done in our finite element model, and solve the forward scattering problem for each part (divide step). We then put the solutions

of all parts together to solve the forward scattering problem of the whole potential (conquer step). We start by obtaining the scattering properties of a subspace $[\tilde{x}_i, \tilde{x}_{i+1}]$. In our finite element model we have no information of the potential between two points and we therefore need a suitable approximation, which we choose to be the Heaviside function. For this approximation the scattering properties are easily calculated analytically and it still offers a good approximation if the step size $\Delta\tilde{x}$ is chosen small enough.

$$\tilde{x}_i \leq x \leq \tilde{x}_{i+1}: \quad V_{\text{approx}}(x) = \tilde{V}(\tilde{x}_i) + (\tilde{V}(\tilde{x}_{i+1}) - \tilde{V}(\tilde{x}_i))H(x - \frac{\tilde{x}_{i+1} + \tilde{x}_i}{2}) \quad (3.11)$$

Here the function H denotes the Heaviside function and not the Hamiltonian. We now introduce the transfer matrix. While the scattering matrix links the amplitudes of the incoming waves on both sides to the outgoing ones, the transfer matrix links the amplitudes of the incoming and outgoing waves of one side to the ones on the other side,

$$\begin{pmatrix} b_2 \\ a_2 \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}. \quad (3.12)$$

Both the scattering matrix and the transfer matrix lead to the same equations and can be transformed into each other. For a potential of the form of a Heaviside function the transfer matrix can be calculated analytically.

$$(\tilde{T}_{i,i+1})_j = T_{i,i+1}(\tilde{k}_j) = \frac{1}{2} \begin{pmatrix} (1 + \frac{\tilde{k'}_i}{\tilde{k'}_{i+1}}) e^{i(\tilde{k'}_i + \tilde{k'}_{i+1})\frac{\Delta\tilde{x}}{2}} & (1 - \frac{\tilde{k'}_i}{\tilde{k'}_{i+1}}) e^{i(\tilde{k'}_i - \tilde{k'}_{i+1})\frac{\Delta\tilde{x}}{2}} \\ (1 - \frac{\tilde{k'}_i}{\tilde{k'}_{i+1}}) e^{-i(\tilde{k'}_i - \tilde{k'}_{i+1})\frac{\Delta\tilde{x}}{2}} & (1 + \frac{\tilde{k'}_i}{\tilde{k'}_{i+1}}) e^{-i(\tilde{k'}_i + \tilde{k'}_{i+1})\frac{\Delta\tilde{x}}{2}} \end{pmatrix}$$
$$\tilde{k'}_i = \sqrt{\tilde{k}_j^2 - 2m\tilde{V}_i} \quad (3.13)$$

$(\tilde{T}_{i,i+1})_j$ denotes the transfer matrix between the points $\tilde{x}_i$, $\tilde{x}_{i+1}$, with an wave of momentum $\tilde{k}_j$. The transfer matrix of the whole potential can be easily obtained

and is the product of the transfer matrices of each subspace.

$$\tilde{T}_i = (\tilde{T}_{1,n})_i = \prod_{j=1}^{\tilde{n}_x - 1} (\tilde{T}_{j,j+1})_i \tag{3.14}$$

Finally we can transform the transfer matrix into the scattering matrix.

$$\begin{pmatrix} b_n \\ a_n \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}$$

$$\begin{pmatrix} b_1 \\ b_n \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} a_1 \\ a_n \end{pmatrix} \tag{3.15}$$

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} = \begin{pmatrix} -\frac{T_{21}}{T_{22}} & \frac{1}{T_{22}} \\ \frac{1}{T_{22}} & \frac{T_{12}}{T_{22}} \end{pmatrix}$$

$$\tilde{S}_i = S(\tilde{k}_i) = \begin{pmatrix} (\tilde{S}_{11})_i & (\tilde{S}_{12})_i \\ (\tilde{S}_{21})_i & (\tilde{S}_{22})_i \end{pmatrix} = \begin{pmatrix} -\frac{(\tilde{T}_{21})_i}{(\tilde{T}_{22})_i} & \frac{1}{(\tilde{T}_{22})_i} \\ \frac{1}{(\tilde{T}_{22})_i} & \frac{(\tilde{T}_{12})_i}{(\tilde{T}_{22})_i} \end{pmatrix}$$

This is only valid for transfer matrices with $det(T) = 1$ and is not fulfilled by each individual transfer matrix, but since we demand a potential with compact support, and therefore $\tilde{V}_1 = \tilde{V}_{\tilde{n}_x} = 0$, this condition is fulfilled for $\tilde{T}$ if $\tilde{x}_{min}$ and $\tilde{x}_{max}$ are chosen accordingly. With the introduced transfer matrix method we are now able to solve the forward scattering problem of our potentials. To solve the inverse scattering problem we additionally need an algorithm for the minimization. We chose two popular algorithms, gradient descent and simulated annealing, which will be discussed in the following sections.

3.4 Gradient descent

The first algorithm for minimization is gradient descent[16]. Gradient descent is a widely used heuristic approach for general optimization problems. It is therefore also used for data fitting, which makes it a good candidate to compare it to the machine learning approach. It is deterministic and works best for high dimensional spaces[41]. It follows the greedy paradigm, which means that it chooses the locally best option in each step, which may not lead to the global optimum. This problem can be partially eliminated by choosing an parameter space with a high dimension, where the occurrence of local optima is unlikely. However this will in return then require a higher computational power. Further the starting point has a high impact on the quality of this method. A good starting point will lead to faster and better results, while a bad one can lead to a high number of iterations or a solution of low quality.

The goal of this algorithm is to find a potential $\tilde{V}$ for a given scattering matrix, S_{true}, where the error function $\tilde{\epsilon}(\tilde{V}, \tilde{S}_{\text{true}})$ is at least at a local minimum. In every step of this algorithm we have a current potential, denoted by $\tilde{V}_{act}$, which represents the best solution so far. At the beginning this potential is initialized with a starting point, $\tilde{V}_{init}$. In every step we want to modify this potential to lower the error. For this purpose we look at how a small modification of the potential, $\tilde{V}_{act} \to \tilde{V}_{act} + \Delta\tilde{V}$, impacts the error,

$$\tilde{\epsilon}(\tilde{V}_{act} + \Delta\tilde{V}, \tilde{S}_{\text{true}}) - \tilde{\epsilon}(\tilde{V}_{act}, \tilde{S}_{\text{true}}) = \nabla\tilde{\epsilon}(\tilde{V}_{act}, \tilde{S}_{\text{true}})\Delta\tilde{V} + \mathcal{O}((\Delta\tilde{V})^2). \tag{3.16}$$

Here $\nabla\tilde{\epsilon}(\tilde{V}_{act}, \tilde{S}_{\text{true}}) = (\frac{\partial\tilde{\epsilon}}{\partial\tilde{V}_i})(\tilde{V}_{act}, \tilde{S}_{\text{true}})$ is the gradient of the error function evaluated at $\tilde{V}_{act}$. In order to obtain a smaller error the inner product of the gradient of the error function and the modification of the potential should be negative. We

therefore choose a modification of:

$$\Delta\tilde{V} = -\alpha\nabla\tilde{\epsilon}(\tilde{V}_{act}, \tilde{S}_{\text{true}}). \tag{3.17}$$

where α is a step size which should be chosen small to not leave the linear range of the approximation. This step is repeated until we reach a local minimum and the gradient of the error function is theoretically vanishing. Since it will numerically never be exactly zero, we introduce an additional parameter $D\tilde{\epsilon}_{min}$, which is the lower bound of the gradient of the error function. The algorithm terminates when

$$||\nabla\tilde{\epsilon}(\tilde{V}_{act}, \tilde{S}_{\text{true}})|| < D\tilde{\epsilon}_{min} \tag{3.18}$$

In pseudo code the gradient descent algorithm is written as:

function GRADIENT_DESCENT($\tilde{S}_{\text{true}}$, $\tilde{V}_{init}$, α, $D\tilde{\epsilon}_{min}$)
  $\tilde{V}_{act} \leftarrow \tilde{V}_{init}$
  while $||\nabla\tilde{\epsilon}(\tilde{V}_{act}, \tilde{S}_{\text{true}})|| \geq D\tilde{\epsilon}_{min}$ **do**
    $\tilde{V}_{act} \leftarrow \tilde{V}_{act} - \alpha\nabla\tilde{\epsilon}(\tilde{V}_{act}, \tilde{S}_{\text{true}})$
  end while
  return $\tilde{V}_{act}$
end function

For a better understanding of the method we illustrated its convergence, shown in Figure 3.1, on the example of a Gaussian peak potential. The reconstructed potential is plotted after a various number of steps, and we started with a constant potential of $V_{init} = 0$. In each step the gradient descent method chooses a slightly better potential until the gradient of the error function falls below a threshold. The final obtained potential is well reconstructed with only a minimal error.

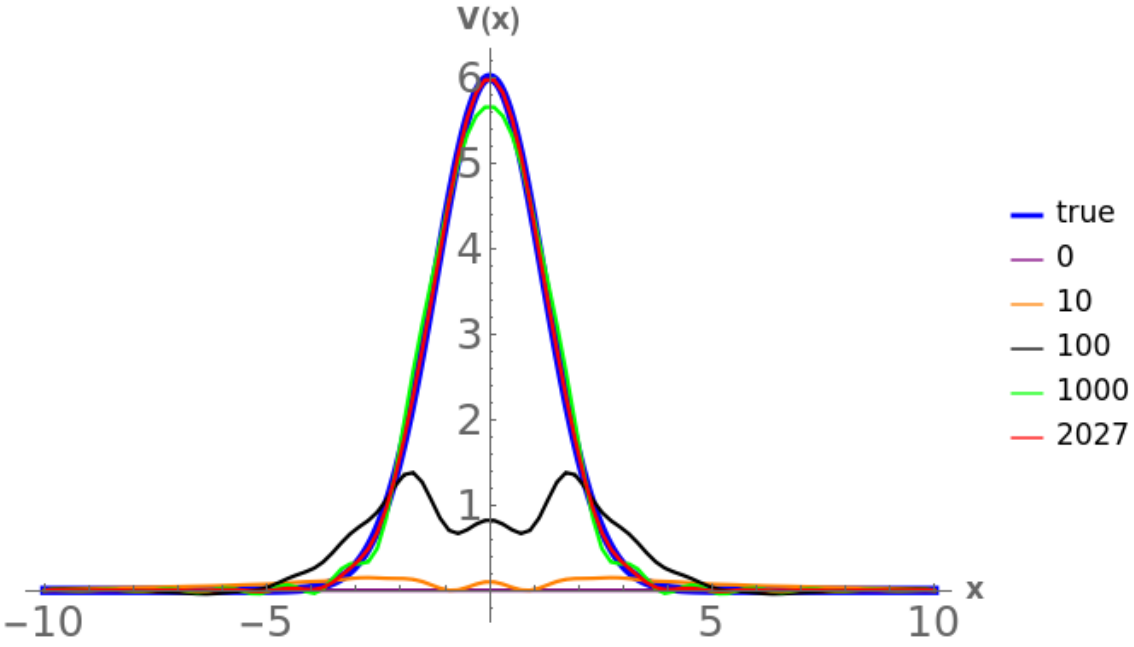


Figure 3.1: A Gaussian peak potential reconstructed with the gradient descent method. We used an initial potential of $V_{init} = 0$ and a step size of $\alpha = 0.1$. The algorithm terminates when the gradient of the error function falls below $D\tilde{\epsilon}_{min} = 10^{-8}$, which is achieved after 2027 steps. The plot shows the true potential and the reconstructed ones after 0, 10, 100, 1000 and 2027 steps.

3.5 Simulated Annealing

Another algorithm we look into is simulated annealing[17]. Similar to gradient descent it is widely used for minimization problems and therefore also suitable to compare it to machine learning. It is specially used in physics and is able to find the global minimum of a function. As the name suggests it simulates the annealing process of solids, where the energy states are distributed according to Boltzmann statistics. At the thermal equilibrium for a given temperature T the probability that the system has a microscopic configuration with an energy greater or equal than E_i is given by:

$$p(E_i) \propto e^{-\frac{E_i - E_0}{k_B T}} \tag{3.19}$$

where E_0 is the ground state (minimum energy) and therefore $E_i - E_0 \geq 0$. By cooling the solid and decreasing the temperature the probability to encounter the ground state increases and is 1 at $T = 0$. We simulate this process where the error function takes the role of the energy and the potential the role of the microscopic configuration.

We start with a temperature of $T = T_{init}$ and an initial potential $\tilde{V}_{act} = \tilde{V}_{init}$ which is generated randomly. In every step we bring the system to thermal equilibrium. For this purpose we use the Metropolis algorithm[42] where we perturb the current potential $\tilde{V}_{act}$ slightly and generate a new potential $\tilde{V}_{new} = \tilde{V}_{act} + \Delta\tilde{V}$, where $\Delta\tilde{V}$ is a small perturbation generated randomly. We then look into the change of the error function $\Delta\tilde{\epsilon} = \tilde{\epsilon}(\tilde{V}_{new}, \tilde{S}_{\text{true}}) - \tilde{\epsilon}(\tilde{V}_{act}, \tilde{S}_{\text{true}})$ and accept the new potential with a probability of $p = min(1, e^{-\frac{\Delta\tilde{\epsilon}}{T}})$. Better potentials, with a smaller error, are always accepted while potentials with higher error can be rejected with a probability depending on the temperature. Repeating this step n_{sweeps} times will lead to the thermal equilibrium. We further decrease the temperature by a factor of λ and repeat the whole step until the temperature reaches a lower bound of

T_{min}. The potential $\tilde{V}_{act}$ is then our approximation of the global minimum. We again illustrated this algorithm by its pseudo code:

function SIMULATED_ANNEALING($\tilde{S}_{\text{true}}$, T_{init}, T_{min}, λ, n_{sweeps})
  $\tilde{V}_{act} \leftarrow$ random_potential()
  $T \leftarrow T_{init}$
  while $T \geq T_{min}$ **do**
    for $1 \dots n_{sweeps}$ **do**
      $\Delta\tilde{V} \leftarrow$ random_perturbation()
      $\tilde{V}_{new} \leftarrow \tilde{V}_{act} + \Delta\tilde{V}$
      $\Delta\tilde{\epsilon} \leftarrow \tilde{\epsilon}(\tilde{V}_{new}, \tilde{S}_{\text{true}}) - \tilde{\epsilon}(\tilde{V}_{act}, \tilde{S}_{\text{true}})$
      if random_number$(0, 1) \leq e^{-\frac{\Delta\tilde{\epsilon}}{T}}$ **then**
        $\tilde{V}_{act} \leftarrow \tilde{V}_{new}$
      end if
    end for
    $T \leftarrow \frac{T}{\lambda}$
  end while
  return $\tilde{V}_{act}$
end function

3.6 Comparison and Conclusion

We now compare the three methods and look into their abilities to solve the inverse scattering problem. For this purpose we solved the problem for three different randomly generated sample potentials. The potentials were taken from the machine learning project. It is worth to mention that the sample potentials are subject to the same statistics as the potentials which were used to train the neural network. The true and reconstructed potentials by the methods are shown in Figure 3.2, Figure 3.3 and Figure 3.4. We first want to compare the quality of the methods. We therefore introduce an error function of the potential and the

scattering matrix:

$$\begin{aligned} err_V &= \int_{x=-\infty}^{\infty} |(V_m(x) - V_{true}(x))|^2 dx \\ err_S &= \int_{k=0}^{\infty} ||S_m(k) - S_{true}(k)||_F dk \end{aligned} \tag{3.20}$$

where V_m, V_{true} are the reconstructed and true potential, and S_m, S_{true} their scattering matrices. For all methods and potentials the errors, shown in Table 3.1, are fairly low and of the same scale. The machine learning method has the smallest error of the potential (err_V) and therefore best solutions in all three samples. For the second potential the error was even one order of magnitude smaller compared to gradient descent and simulated annealing. It is likely that the neural network was trained to a very similar potential and therefore was able to reconstruct it very well. Except for the second potential the error of the scattering matrix (err_S) was lower for gradient descent and simulated annealing compared to the machine learning method. The machine learning method was trained to potentials of the same statistic as the sample potentials and was therefore biased, while gradient descent and simulated annealing had no information of the potentials at all and only had the information of the scattering matrix.

Not only the quality is interesting to compare, but also the properties of the methods itself. While the machine learning approach needs a higher computational power, all work is done by training the neural network. After the training it is able to solve an instance of the inverse scattering problem with minimal computational power and in a short time. It is therefore suitable for real-time applications. For both, the gradient descent and simulated annealing method, no preparation is needed, and every instance of the problem needs nearly the same computational power. While the machine learning methods needs to be trained to similar potentials, gradient descent and simulated annealing are generic solutions and work for

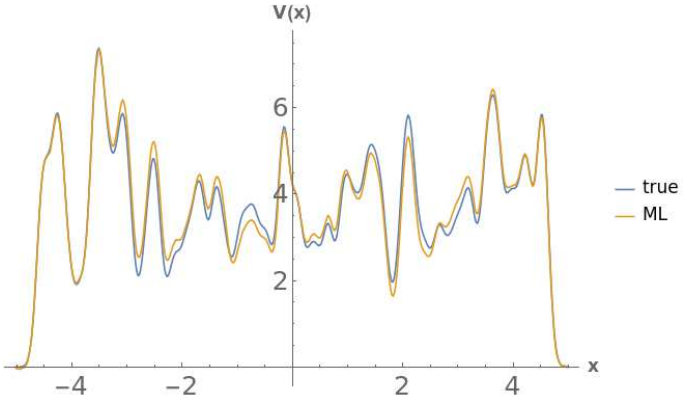
all potentials equally well.

We can conclude that the machine learning method compares very well and can even outperform conventional methods, and is specially useful if the form of the potential is known or if a single instance needs to be solved fast. It would be further interesting how these methods compare in a higher dimensional space or at other physical problems.

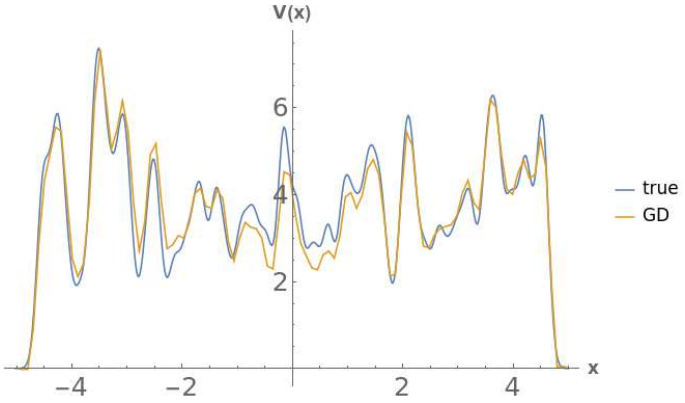
Method	Potential 1		Potential 2		Potential 3	
	err_V	err_S	err_V	err_S	err_V	err_S
Machine Learning	0.549	0.0539	0.042	0.0014	1.544	0.132
Gradient Descent	1.242	0.042	0.314	0.044	0.627	0.054
Simulated Annealing	1.637	0.0413	0.809	0.044	1.555	0.053

Table 3.1: Comparison of machine learning, gradient descent and simulated annealing on three sample potentials shown in Figure 3.2, Figure 3.3 and Figure 3.4. err_V, err_S are the errors of the potential and scattering matrix.

(a) Potential 1, reconstructed with machine learning.



(b) Potential 1, reconstructed with gradient descent.



(c) Potential 1, reconstructed with simulated annealing.

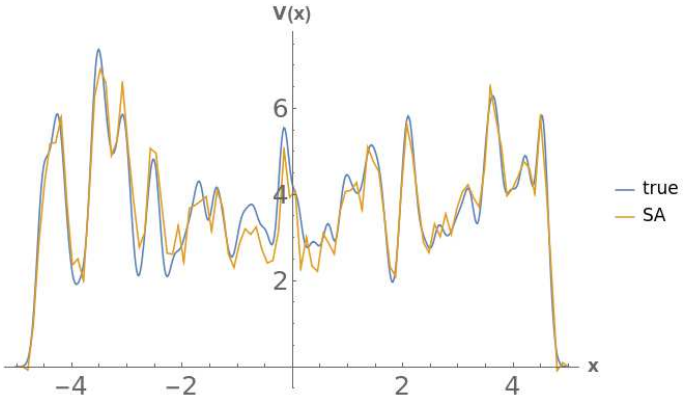
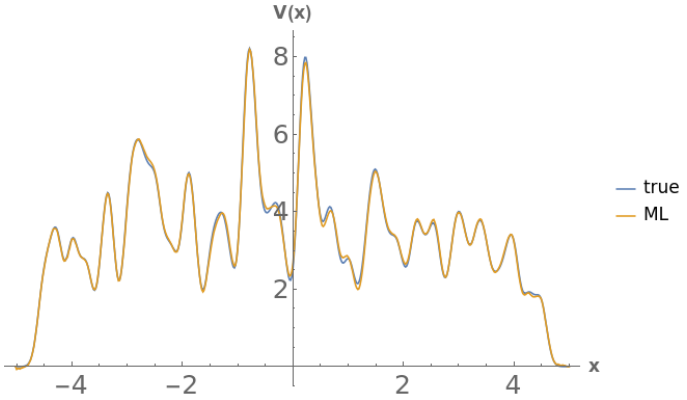
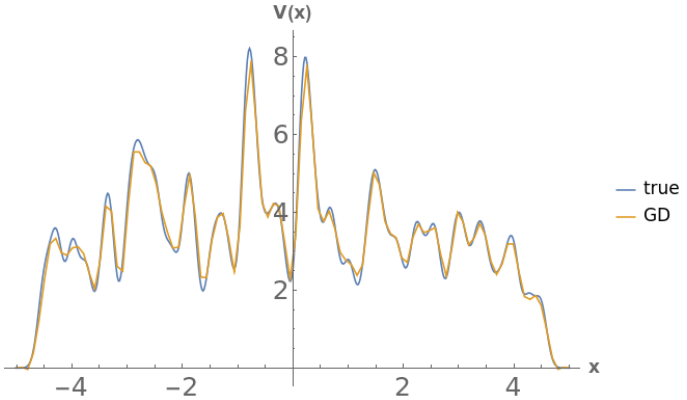


Figure 3.2: Potential 1 reconstructed with the different approaches. For gradient descent and simulated annealing we chose $\tilde{x}_{min} = -5$, $\tilde{x}_{max} = 5$, $\tilde{n}_x = 100$ and $\tilde{k}_{max} = 10$, $\tilde{n}_k = 200$. For gradient descent we chose the parameters $\alpha = 0.5$, $D\tilde{\epsilon}_{min} = 10^{-10}$, $\tilde{V}_{init} = 1$. For simulated annealing we chose $T_{init} = 10^{-4}$, $T_{min} = 10^{-6}$, $\lambda = 1.2$, $n_{sweeps} = 5000$.

(a) Potential 2, reconstructed with machine learning.



(b) Potential 2, reconstructed with gradient descent.



(c) Potential 2, reconstructed with simulated annealing.

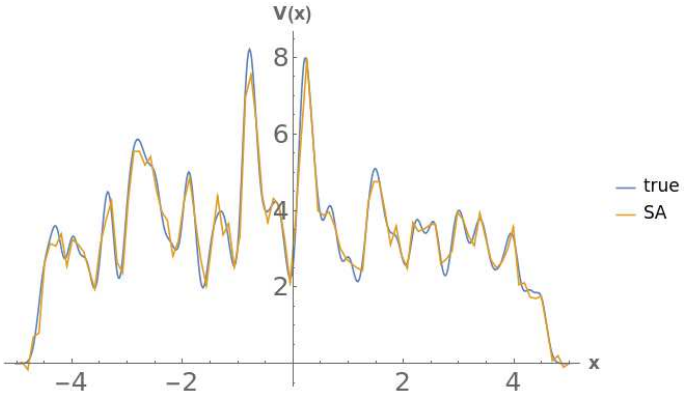
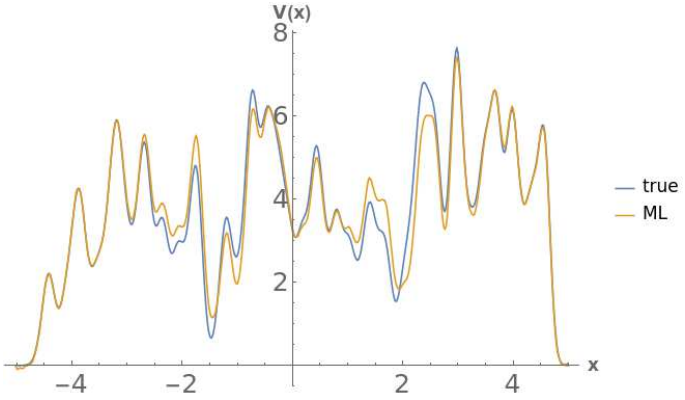
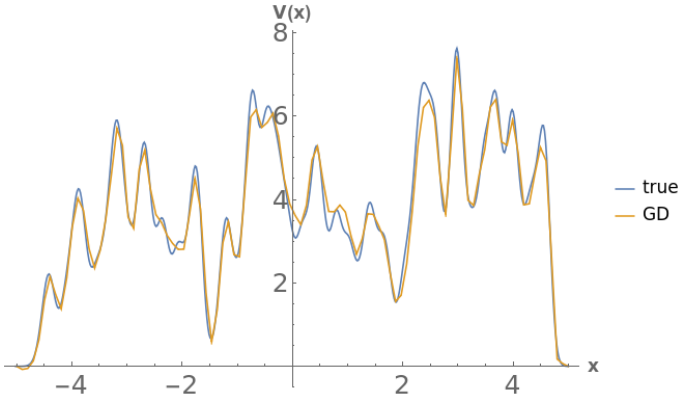


Figure 3.3: Potential 2 reconstructed with the different approaches. For gradient descent and simulated annealing we chose $\tilde{x}_{min} = -5$, $\tilde{x}_{max} = 5$, $\tilde{n}_x = 100$ and $\tilde{k}_{max} = 10$, $\tilde{n}_k = 200$. For gradient descent we chose the parameters $\alpha = 0.5$, $D\tilde{\epsilon}_{min} = 10^{-10}$, $\tilde{V}_{init} = 1$. For simulated annealing we chose $T_{init} = 10^{-4}$, $T_{min} = 10^{-6}$, $\lambda = 1.2$, $n_{sweeps} = 5000$.

(a) Potential 3, reconstructed with machine learning.



(b) Potential 3, reconstructed with gradient descent.



(c) Potential 3, reconstructed with simulated annealing.

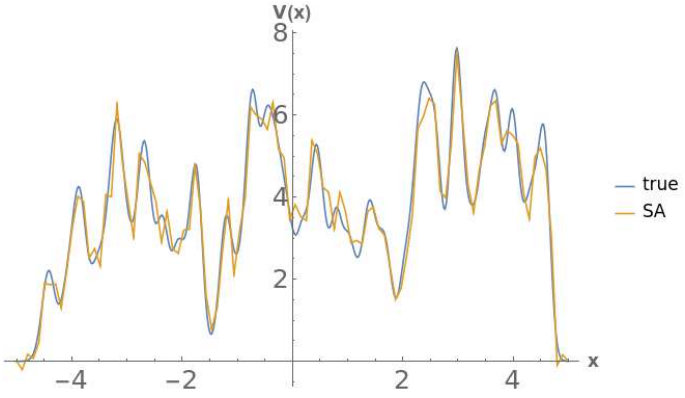


Figure 3.4: Potential 3 reconstructed with the different approaches. For gradient descent and simulated annealing we chose $\tilde{x}_{min} = -5$, $\tilde{x}_{max} = 5$, $\tilde{n}_x = 100$ and $\tilde{k}_{max} = 10$, $\tilde{n}_k = 200$. For gradient descent we chose the parameters $\alpha = 0.5$, $D\tilde{\epsilon}_{min} = 10^{-10}$, $\tilde{V}_{init} = 1$. For simulated annealing we chose $T_{init} = 10^{-4}$, $T_{min} = 10^{-6}$, $\lambda = 1.2$, $n_{sweeps} = 5000$.

4 Summary and Outlook

We introduced a method to find and identify EPs in interferometers, and showed their properties, such as the vanishing eigenvector, the characteristic eigenvalue splitting, the characteristic topology of the Riemann sheets and their impact into the scattering problem. We gave an example of an interferometer, consisting of two coupled cavities, which is similar to the coupled microcavity setup and could show that all properties of EPs occur. We further looked into interferometers at an EP with a very simple setup, only consisting of a ring cavity, without any gain or loss media. Due to their simple setup these system are very suitable to study EPs. For the ring cavity interferometer, where a high suppression of one mode was already experimentally observed, we could link this observation to an EP. We further looked into higher order EPs and could show that they can occur in interferometers and also enhance the sensitivity compared to EPs of second order. A particular interesting interferometer which could use the properties of an EP is the Sagnac interferometer, which measures the angular velocity. The root behavior of the eigenvalue splitting can be used to lower the detection limit and enhance the resolution.

We solved the inverse scattering problem with two fitting methods, gradient descent and simulated annealing, and compared it to a machine learning approach, where a neural network was trained to generated samples. The machine learning approach compares very well to our methods and can even outperform them. It

has the advantage that it can be biased by choosing specific training data, which is useful when the general form of the potential is known. Further it only needs a high computational power to train the network, while solving a single instance of the inverse scattering problem is done in a short time. For completely unknown potentials the gradient descent and simulated annealing method are more suitable since they do not rely on a specific training set. They do not need any preparation work, but need higher computational power to solve a single instance. It would be interesting how machine learning compares to gradient descent and simulated annealing in higher dimensions. Due to the generic nature of those methods they can additionally be compared with other physical problems.

Acknowledgement

I want to thank all the people who have supported me and made the completion of my diploma thesis possible. I mostly thank my supervisor, Prof. Stefan Rotter, for the opportunity to work in an interesting field and for the constant support and help during my thesis. I also want to thank my family and friends, without whom this would not have been possible.

Bibliography

[1] W D Heiss. "Exceptional points of non-Hermitian operators". In: *Journal of Physics A: Mathematical and General* 37.6 (Jan. 2004), pp. 2455–2464. DOI: 10.1088/0305-4470/37/6/034. URL: https://doi.org/10.1088%2F0305-4470%2F37%2F6%2F034.

[2] MAARTEN V de Hoop. "Microlocal analysis of seismic inverse scattering". In: *Inside out: inverse problems and applications* 47 (2003), pp. 219–296.

[3] Zhong Qing Zhang and Qing Huo Liu. "Three-dimensional nonlinear image reconstruction for microwave biomedical imaging". In: *IEEE Transactions on Biomedical Engineering* 51.3 (2004), pp. 544–548.

[4] Carl M. Bender. "Making sense of non-Hermitian Hamiltonians". In: *Reports on Progress in Physics* 70.6 (June 2007), pp. 947–1018. DOI: 10.1088/0034-4885/70/6/R03. arXiv: hep-th/0703096 [hep-th].

[5] Ramy El-Ganainy et al. "Non-Hermitian physics and PT symmetry". In: *Nature Physics* 14 (Jan. 2018). Review Article, pp. 11–19. URL: http://dx.doi.org/10.1038/nphys4323.

[6] M.V. Berry. "Physics of Nonhermitian Degeneracies". In: *Czechoslovak Journal of Physics* 54.10 (Oct. 2004), pp. 1039–1047. ISSN: 1572-9486. DOI: 10.1023/B:CJOP.0000044002.05657.04. URL: https://doi.org/10.1023/B:CJOP.0000044002.05657.04.

[7] W D Heiss. "The physics of exceptional points". In: *Journal of Physics A: Mathematical and Theoretical* 45.44 (2012), p. 444016. URL: http://stacks.iop.org/1751-8121/45/i=44/a=444016.

[8] S. K. Özdemir et al. "Parity-time symmetry and exceptional points in photonics". In: *Nature Materials* 18.8 (2019), pp. 783–798. ISSN: 1476-4660. DOI: 10.1038/s41563-019-0304-9. URL: https://doi.org/10.1038/s41563-019-0304-9.

[9] Mohammad-Ali Miri and Andrea Alù. "Exceptional points in optics and photonics". In: *Science* 363.6422 (2019). ISSN: 0036-8075. DOI: 10.1126/science.aar7709. eprint: http://science.sciencemag.org/content/363/6422/eaar7709.full.pdf. URL: http://science.sciencemag.org/content/363/6422/eaar7709.

[10] Weijian Chen et al. "Exceptional points enhance sensing in an optical microcavity". In: *Nature* 548 (Aug. 2017), pp. 192–196. URL: http://dx.doi.org/10.1038/nature23281.

[11] Jan Wiersig. "Enhancing the Sensitivity of Frequency and Energy Splitting Detection by Using Exceptional Points: Application to Microcavity Sensors for Single-Particle Detection". In: *Phys. Rev. Lett.* 112 (20 May 2014), p. 203901. DOI: 10.1103/PhysRevLett.112.203901. URL: https://link.aps.org/doi/10.1103/PhysRevLett.112.203901.

[12] E. Ronen et al. "Phase locking of lasers with self-stabilized minimal coupling". In: *Opt. Express* 20.27 (Dec. 2012), pp. 28163–28170. DOI: 10.1364/OE.20.028163. URL: http://www.opticsexpress.org/abstract.cfm?URI=oe-20-27-28163.

[13] Satoshi Sunada. "Large Sagnac frequency splitting in a ring resonator operating at an exceptional point". In: *Phys. Rev. A* 96 (3 Sept. 2017), p. 033842.

DOI: 10.1103/PhysRevA.96.033842. URL: https://link.aps.org/doi/10.1103/PhysRevA.96.033842.

[14] J. Ren et al. "Ultrasensitive micro-scale parity-time-symmetric ring laser gyroscope". In: *Opt. Lett.* 42.8 (Apr. 2017), pp. 1556–1559. DOI: 10.1364/OL.42.001556. URL: http://ol.osa.org/abstract.cfm?URI=ol-42-8-1556.

[15] Yu-Hung Lai et al. "Enhanced sensitivity operation of an optical gyroscope near an exceptional point". In: *arXiv e-prints*, arXiv:1901.08217 (Jan. 2019), arXiv:1901.08217. arXiv: 1901.08217 [physics.optics].

[16] Sebastian Ruder. "An overview of gradient descent optimization algorithms". In: *arXiv preprint arXiv:1609.04747* (2016).

[17] Emile Aarts and Jan Korst. "Simulated annealing and Boltzmann machines". In: (1988).

[18] Liang Feng, Ramy El-Ganainy, and Li Ge. "Non-Hermitian photonics based on parity-time symmetry". In: *Nature Photonics* 11.12 (2017), pp. 752–762. ISSN: 1749-4893. DOI: 10.1038/s41566-017-0031-1. URL: https://doi.org/10.1038/s41566-017-0031-1.

[19] W.D. Heiss and H.L. Harney. "The chirality of exceptional points". In: *The European Physical Journal D - Atomic, Molecular, Optical and Plasma Physics* 17.2 (Nov. 2001), pp. 149–151. ISSN: 1434-6079. DOI: 10.1007/s100530170017. URL: https://doi.org/10.1007/s100530170017.

[20] Bo Peng et al. "Chiral modes and directional lasing at exceptional points". In: *Proceedings of the National Academy of Sciences* 113.25 (2016), pp. 6845–6850.

[21] Jörg Doppler et al. "Dynamically encircling an exceptional point for asymmetric mode switching". In: *Nature* 537.7618 (2016), p. 76.

[22] W.D. Heiss. "Exceptional Points – Their Universal Occurrence and Their Physical Significance". In: *Czechoslovak Journal of Physics* 54.10 (Oct. 2004), pp. 1091–1099. ISSN: 1572-9486. DOI: 10.1023/B:CJOP.0000044009.17264.dc. URL: https://doi.org/10.1023/B:CJOP.0000044009.17264.dc.

[23] B. Peng et al. "Loss-induced suppression and revival of lasing". In: *Science* 346.6207 (2014), pp. 328–332. ISSN: 0036-8075. DOI: 10.1126/science.1258004. eprint: http://science.sciencemag.org/content/346/6207/328.full.pdf. URL: http://science.sciencemag.org/content/346/6207/328.

[24] Jan Wiersig, Sang Wook Kim, and Martina Hentschel. "Asymmetric scattering and nonorthogonal mode patterns in optical microspirals". In: *Phys. Rev. A* 78 (5 Nov. 2008), p. 053809. DOI: 10.1103/PhysRevA.78.053809. URL: https://link.aps.org/doi/10.1103/PhysRevA.78.053809.

[25] Hossein Hodaei et al. "Enhanced sensitivity at higher-order exceptional points". In: *Nature* 548 (Aug. 2017), pp. 187–191. URL: http://dx.doi.org/10.1038/nature23280.

[26] R El-Ganainy et al. "Theory of coupled optical PT-symmetric structures". In: *Optics letters* 32 (Oct. 2007), pp. 2632–4. DOI: 10.1364/OL.32.002632.

[27] K. G. Makris et al. "Beam Dynamics in $\mathcal{PT}$ Symmetric Optical Lattices". In: *Phys. Rev. Lett.* 100 (10 Mar. 2008), p. 103904. DOI: 10.1103/PhysRevLett.100.103904. URL: https://link.aps.org/doi/10.1103/PhysRevLett.100.103904.

[28] Long Chang et al. "Parity-time symmetry and variable optical isolation in active-passive-coupled microresonators". In: *Nature Photonics* 8.7 (2014), pp. 524–529. ISSN: 1749-4893. DOI: 10.1038/nphoton.2014.133. URL: https://doi.org/10.1038/nphoton.2014.133.

[29] C. Dembowski et al. "Observation of a Chiral State in a Microwave Cavity". In: *Phys. Rev. Lett.* 90 (3 Jan. 2003), p. 034101. DOI: 10.1103/PhysRevLett.90.034101. URL: https://link.aps.org/doi/10.1103/PhysRevLett.90.034101.

[30] Jan Wiersig et al. "Nonorthogonal pairs of copropagating optical modes in deformed microdisk cavities". In: *Phys. Rev. A* 84 (2 Aug. 2011), p. 023845. DOI: 10.1103/PhysRevA.84.023845. URL: https://link.aps.org/doi/10.1103/PhysRevA.84.023845.

[31] Jan Wiersig. "Structure of whispering-gallery modes in optical microdisks perturbed by nanoparticles". In: *Phys. Rev. A* 84 (6 Dec. 2011), p. 063828. DOI: 10.1103/PhysRevA.84.063828. URL: https://link.aps.org/doi/10.1103/PhysRevA.84.063828.

[32] Hossein Hodaei et al. "Higher-order exceptional points in photonic systems (Conference Presentation)". In: *Proc.SPIE* 10345 (2017). DOI: 10.1117/12.2274649. URL: https://doi.org/10.1117/12.2274649.

[33] Kun Ding et al. "Emergence, Coalescence, and Topological Properties of Multiple Exceptional Points and Their Experimental Realization". In: *Phys. Rev. X* 6 (2 Apr. 2016), p. 021007. DOI: 10.1103/PhysRevX.6.021007. URL: https://link.aps.org/doi/10.1103/PhysRevX.6.021007.

[34] Gilles Demange and Eva-Maria Graefe. "Signatures of three coalescing eigenfunctions". In: *Journal of Physics A: Mathematical and Theoretical* 45.2 (Dec. 2011), p. 025303. DOI: 10.1088/1751-8113/45/2/025303. URL: https://doi.org/10.1088%2F1751-8113%2F45%2F2%2F025303.

[35] Julius Kullig et al. "Exceptional points of third-order in a layered optical microdisk cavity". In: *New Journal of Physics* 20.8 (Aug. 2018), p. 083016.

DOI: 10.1088/1367-2630/aad594. URL: https://doi.org/10.1088%2F1367-2630%2Faad594.

[36] Heinrich P Baltes, Mario Bertero, and Res Jost. *Inverse scattering problems in optics*. Vol. 20. Springer, 1980.

[37] Tyler S Ralston et al. "Inverse scattering for optical coherence tomography". In: *JOSA A* 23.5 (2006), pp. 1027–1037.

[38] Mario Bertero. "Linear inverse and III-posed problems". In: *Advances in electronics and electron physics*. Vol. 75. Elsevier, 1989, pp. 1–120.

[39] Anthony J Devaney. "Nonuniqueness in the inverse scattering problem". In: *Journal of Mathematical Physics* 19.7 (1978), pp. 1526–1531.

[40] Léon Bottou. "Stochastic gradient descent tricks". In: *Neural networks: Tricks of the trade*. Springer, 2012, pp. 421–436.

[41] Jason D Lee et al. "Gradient descent converges to minimizers". In: *arXiv preprint arXiv:1602.04915* (2016).

[42] Isabel Beichl and Francis Sullivan. "The metropolis algorithm". In: *Computing in Science & Engineering* 2.1 (2000), p. 65.