

Numerische Verfahren zur Lösung des  
Transmissions-Problems in speziellen Billards  
und ihre Anwendung zur Konstruktion von  
Quantenkaskaden-Lasern

I N A U G U R A L – D I S S E R T A T I O N

zur

Erlangung des Doktorgrades  
der Mathematisch-Naturwissenschaftlichen Fakultät  
der Universität zu Köln

vorgelegt von

PASCAL HEIDER

aus Köln

Hundt Druck GmbH, Köln  
2005

Berichterstatter:

Prof. Dr. R. Seydel  
Prof. Dr. T. Küpper

Tag der mündlichen Prüfung:

9. November 2005

## Kurzzusammenfassung

Die physikalische Modellierung von Halbleiter-Lasern führt über die Maxwell-Gleichungen auf ein 2-dimensionales Transmissions-Problem, bei dem eine Lösung der Helmholtz-Gleichung im Inneren eines beschränkten Gebiets mit Wellenzahl  $nk$  und im Äußeren eine Lösung mit Wellenzahl  $k$  gesucht ist, so daß beide Lösungen über den Rand stetig differenzierbar sind. In dieser Arbeit wird ein Prädiktor-Korrektor-Algorithmus vorgestellt, der es im Unterschied zu früheren Verfahren gestattet, alle Eigenwerte in einem vorgegebenen Bereich  $a \leq |k| \leq b$  direkt, sukzessiv und effizient zu berechnen, so daß die Entwicklung von Suchstrategien wie bei früheren Algorithmen obsolet wird. Das Verfahren setzt auf einer Randintegraldiskretisierung der jeweiligen Helmholtz-Operatoren auf, deren Diskretisierungsfehler bei der Berechnung der Eigenfunktionen bzw. Eigenwerten exponentiell kleiner wird. Wenn das Ausgangsgebiet durch eine Billard-Kurve berandet wird, die durch eine Familie von  $p$ -periodischen Orbits beschrieben ist, können im quasi-klassischen Grenzübergang  $k \rightarrow \infty$  die Lösungen des Transmissions-Problems mittels Trajektorien dieses Billards approximiert werden. Die Existenz solcher Billards wurde in [BZ03] von Y. Baryshnikov und V. Zharnitsky bewiesen. In der vorliegenden Arbeit werden zwei numerische Konstruktionsverfahren für solche Billards angegeben. Insbesondere wird gezeigt, daß sie Lösungen des Transmissions-Problems besitzen, welche im Inneren des Billards lokalisiert sind. Die so entdeckten Billards definieren eine neue Klasse von Halbleiter-Lasern. Die mathematischen Vorhersagen über Direktionalität und Intensität solcher Laser konnten experimentell verifiziert werden.

## Abstract

The modelling of semi-conductor lasers via Maxwell's equations results in a 2-dimensional transmission-problem, for which a solution of Helmholtz's equation with wavenumber  $nk$  in the interior of a bounded domain and a solution with wavenumber  $k$  in the exterior has to be found, such that both solutions are continuously differentiable across the boundary. This thesis presents a predictor-corrector algorithm, which allows a direct, successive and efficient calculation of all eigenvalues in a prescribed range  $a \leq |k| \leq b$ , hence making the development of search strategies obsolete in contrast to former algorithms. The method is based on a boundary integral discretisation of the Helmholtz-operators, whose discretisation error of the eigenfunctions and eigenvalues decreases exponentially. In the case of a domain bounded by a billiard-curve, which is described by a family of  $p$ -periodic orbits, the solutions of the transmission-problem in the quasi-classical limit  $k \rightarrow \infty$  can be approximated by classical trajectories in the billiard. The existence of such billiards was shown by Y. Baryshnikov and V. Zharnitsky in [BZ03]. In this thesis two numerical methods for the explicit construction of such billiards are developed. Moreover, it will be shown that these billiards have solutions of the transmission-problem, which are localized in the interior of the billiard. These recently discovered billiards define a new class of semi-conductor lasers. The mathematical predictions about directionality and intensity of such lasers could be verified experimentally.



## Einleitung

Die Entwicklung des Quantenkaskaden-Lasers (QKL) bei den Bell-Labs Anfang der 90er Jahre ermöglichte eine Miniaturisierung in der Lasertechnik um mehrere Größenordnungen. Dieser große Fortschritt wurde ermöglicht durch die Verwendung spezieller Halbleitermaterialien und der Verwendung des Prinzips der Totalreflektion in den Resonatoren gegenüber der früheren Verwendung von Spiegeln. Der Brechungsindex  $n$  des Halbleitermaterials ist hierbei größer als der des umgebenden Mediums. In den ersten Quantenkaskaden-Lasern wurden kreisförmige Resonator-Querschnitte verwendet. Praktisch bedeutsame Leistungsdaten für die Direktionalität und Intensität wurden dagegen erst Mitte der 90er Jahre durch Querschnitte erreicht, bei denen quadrupolare Deformationen des Kreises als Resonator-Querschnittsfläche entdeckt wurden. Bei diesem Laser-Typ verlaufen die Lichtwege entlang eines stabilen 4-periodischen Orbits in Form einer Smoking-Fliege. Man spricht deshalb von *Bow-Tie Lasern*. Die Verbesserung der Output-Leistung eines Quantenkaskaden-Lasers kann erreicht werden durch eine Erhöhung der Anzahl der Reflektionen der Lichtstrahlen im Resonator. Diese Anzahl wird wesentlich bestimmt durch den Querschnitt des Resonators. Mathematisch kann ein derartiges Lichtsystem modelliert werden als ein ideales Billardsystem, bei dem die Billardkugel gemäß dem klassischen Gesetz „*Einfallswinkel=Ausfallswinkel*“ an der Bande reflektiert wird.

Y. Baryshnikov und V. Zharnitsky konnten 2003 die Existenz von Billardtischen zeigen, bei denen zu jedem Randpunkt ein Anstoßwinkel für eine Billardkugel existiert, so daß die Billardkugel nach genau  $p$  Reflektionen am Rand wieder in dieselbe Ausgangslage zurückkehrt. Solche sogenannten *p-Lazutkin-Billards* sind für die Herstellung von Resonatoren besonders gut geeignet. Triviale Beispiele für solche Billardtische sind Kreis und Ellipse, jedoch gelang die numerische Berechnung nicht-trivialer *p-Lazutkin-Billards* nicht.

Ein daran anschließendes generelles Problem besteht in der numerischen Berechnung der Leistungsmerkmale von Quantenkaskaden-Lasern mit beliebigen Randkurven. Die Lichtemission wird durch die *Maxwell-Gleichungen* bestimmt. Unter Berücksichtigung der Randbedingungen am Übergang zwischen Resonator und dem umgebendem Medium führen die zeitunabhängigen Maxwell-Gleichungen für die Wellenfunktion  $u$  im Resonator mit Querschnitt  $\Omega$  auf das 2-dimensionale Transmissions-Problem:

Finde  $k \in \mathbb{C}$ , so daß

$$\begin{aligned}\Delta u + k^2 n^2 u &= 0, \text{ für } x \in \Omega \\ \Delta u + k^2 u &= 0, \text{ für } x \notin \Omega\end{aligned}$$

erfüllt ist und  $u$  und die Normalenableitung  $\partial_n u$  stetig am Rand sind.

Im Laser-Problem sind nur solche Lösungen des Transmissions-Problems von Interesse, bei denen  $|k| \gg 1$  ist. Unter dieser Bedingung sind die bisher bekannten numerischen Verfahren zur Lösung

des Transmissions-Problems nicht effizient.

Jedes  $p$ -Lazutkin-Billard kann als dynamisches System aufgefaßt werden. Die besondere Eigenschaft dieser Billard-Klasse ist die Existenz einer invarianten Kurve im Phasenraum, die nur aus  $p$ -periodischen Orbits besteht. In dieser Arbeit werden zwei numerische Verfahren angegeben, mit denen erstmals nicht-triviale Beispiele von  $p$ -Lazutkin-Billard konstruiert werden konnten. Diese neuen Algorithmen erlauben insbesondere eine kontrollierte Einflußnahme auf die Gestalt der invarianten Kurve im Phasenraum, was sich für die technische Realisierung entsprechender Quantenkaskaden-Laser von großer Bedeutung erweist.

In einem  $p$ -Lazutkin-Billard bezeichnet man als dessen  $p$ -periodische Kaustik die Evolute aller  $p$ -periodischen Orbits, die den Billard-Rand beschreiben. Diese ist eine Kurve mit der Eigenschaft, daß ein Strahl, der tangential zu ihr ist, auch nach Reflektion am Rand tangential zu ihr bleibt. Für  $|k| \gg 1$  sind die Lösungen des Transmissions-Problems durch klassische Lichtwege im Billard charakterisiert. Dies führt insbesondere zur Existenz von Lösungen, die entlang der Kaustik lokalisiert sind. Diese neue Lösungsklasse des Transmissions-Problems nennt man *innere Whispering-Gallery Lösungen* wegen ihrer Analogie zur Schallwellenausbreitung in einer Flüster-Galerie.

Es stellte sich daher das Problem, solche Lösungen numerisch zu berechnen. Zur numerischen Behandlung des Transmissions-Problems mit großen Wellenzahlen werden in der physikalischen Literatur verschiedene Verfahren entwickelt, denen es jedoch an qualifizierten Fehlerabschätzungen mangelt oder bei denen die Laufzeit für gute Näherungen inakzeptabel anwächst.

Dem gegenüber wird in der vorliegenden Arbeit ein Laufzeit effizienter Algorithmus entwickelt, mit dem ein exponentielles Abfallen des Fehlers in der Berechnung von Eigenfunktionen und Eigenwerten erreicht wird. Hierzu wird das Transmissions-Problem auf die Lösung eines nicht-linearen Integralgleichungssystems zurückgeführt, bei dem sich die Integrale über den Rand  $\partial\Omega$  erstrecken. Es gelingt, die benötigten Integralkerne so in einen analytischen und singulären Teil aufzuspalten und für jeden Teil geeignet adaptierte Quadraturformeln zu entwickeln. Die Diskretisierung führt auf ein nicht-lineares Matrix-Eigenwertproblem. Es kann bewiesen werden, daß die approximativen Eigenfunktionen und Eigenwerten des diskreten Problems exponentiell gegen Lösungen des kontinuierlichen Problems konvergieren.

Üblicherweise wurden bisher zur Berechnung der Eigenwerte in einem vorgegebenen Bereich  $a \leq |k| \leq b$  sogenannte Sweep-Algorithmen entwickelt, die die sehr aufwendige Auswertungen von Testfunktionen erforderten. Überdies besteht bei diesen Algorithmen das Problem, daß nicht garantiert alle Lösungen gefunden werden. Dem gegenüber wird in der vorliegenden Arbeit ein neuer Sweep-Algorithmus entwickelt, der durch die Verwendung einer Prädiktor-Korrektor-Strategie die direkte, sukzessive und vollständige Berechnung aller Eigenwerte ermöglicht. Da in der Literatur keine Referenz-Beispiele bekannt sind, wurde die Leistungsfähigkeit des neuen Verfahrens anhand der Spezialisierung auf den Fall des klassischen Helmholtz-Problems gete-

stet. In den untersuchten Beispielen ist der neue Algorithmus circa 100 Mal schneller als die altbekannten Verfahren.

Mit den in dieser Arbeit entwickelten Algorithmen gelingt die effiziente Simulation von Quantenkaskaden-Lasern mit  $p$ -Lazutkin-Querschnitten einschließlich der numerischen Behandlung der Direktionalitäts- und Intensitätseigenschaften. Basierend auf solchen Simulationen wurde bei den Bell-Labs ein Quantenkaskaden-Laser mit einem 3-Lazutkin-Querschnitt gefertigt. Die numerischen Vorhersagen über Direktionalität und Intensität wurden experimentell bestätigt.

Die Arbeit ist in 5 Kapitel gegliedert. In Kapitel 1 dieser Arbeit werden die grundlegenden theoretischen Aussagen über mathematische Billards zusammengestellt. Insbesondere werden die konstruktiven Aspekte des Existenzbeweises für  $p$ -Lazutkin-Billards von Baryshnikov und Zharnitsky herausgearbeitet, die für die explizite numerische Konstruktion herangezogen worden sind.

Zwei unterschiedliche Konstruktionsverfahren, die einerseits auf einer Minimierungs-Strategie und andererseits auf einer Optimal-Control-Strategie beruhen, werden im zweiten Kapitel angegeben. Mit Hilfe dieser Verfahren wurde das oben genannte 3-Lazutkin-Billard für die technische Anwendung erzeugt. Zusätzlich werden weitere Beispiele berechnet, an denen die Konvergenz des direkten Sweep-Verfahrens zur Lösung des Transmissions-Problems getestet wurde.

Verwandt mit dem obigen Eigenwertproblem ist das bekannte Helmholtz-Eigenwertproblem  $\Delta u + k^2 u = 0$ , auf das viele Probleme der mathematischen Physik zurückgeführt werden können. In Kapitel 3 werden kurz Lösungsansätze zur Behandlung des „asymptotischen“ Helmholtz-Eigenwertproblems vorgestellt und auf  $p$ -Lazutkin-Billards angewendet.

Im vierten Kapitel werden die neuen Algorithmen zur effizienten numerischen Behandlung des Transmissions-Problems dargestellt und der zugehörige Konvergenzsatz bewiesen. Ferner werden die Ergebnisse der umfangreichen numerischen Tests dieser Algorithmen beschrieben.

Die Anwendung der neuen Algorithmen auf die numerische Simulation der physikalische Eigenschaften von Quantenkaskaden-Lasern erfolgt in Kapitel 5.

An dieser Stelle möchte ich mich ganz herzlich bei allen bedanken, die mich bei der Entstehung dieser Arbeit unterstützt haben. Yuliy Baryshnikov und Vadim Zharnitsky danke ich für die wunderbare Zusammenarbeit und Betreuung während meines Forschungsaufenthalts an den Bell-Labs, bei denen die ersten Ideen zu dieser Arbeit entstanden sind, sowie für zahlreiche Ratschläge, die den Fortgang der Arbeit vorangetrieben haben. Mein besonderer Dank gilt Herrn Prof. Dr. Seydel für seine Unterstützung und Förderung meiner Forschung.



## Inhaltsverzeichnis

<b>Einleitung</b>	<b>i</b>
<b>1 Mathematische Billards</b>	<b>1</b>
1.1 Eigenschaften von Billard-Systemen . . . . .	1
1.2 Billard-Systeme unter kleinen Deformationen von $\Omega$ . . . . .	6
1.3 $p$ -Lazutkin-Billards . . . . .	9
<b>2 Konstruktion von <math>p</math> - Lazutkin-Billards</b>	<b>19</b>
2.1 Minimierungs-Strategie . . . . .	19
2.2 Optimal-Control-Strategie . . . . .	23
2.3 Numerische Implementierung . . . . .	26
<b>3 Quantenmechanik und quasi-klassische Approximation</b>	<b>31</b>
3.1 Schrödinger-Gleichung und quasi-klassischer Übergang . . . . .	31
3.2 Quasi-klassische Approximationen . . . . .	36
3.3 Husimi-Projektion . . . . .	39
3.4 Galerie . . . . .	42
<b>4 Randintegralmethode</b>	<b>44</b>
4.1 Formulierung als Randintegralgleichung . . . . .	45
4.2 Diskretisierung des Operators $A(k)$ . . . . .	50
4.3 Der direkte Sweep-Algorithmus . . . . .	62
4.4 Konvergenz und Fehlerabschätzungen des Verfahrens . . . . .	68
4.5 Vergleich mit anderen numerischen Verfahren . . . . .	80
4.5.1 Verfahren zum quasi-klassischen Helmholtz-Problem . . . . .	80
4.5.2 Verfahren zum quasi-klassischen Transmissions-Problem . . . . .	85

4.6	Störung des Gebiets $\Omega$ . . . . .	89
4.7	Implementierung und CPU-Zeit . . . . .	91
<b>5</b>	<b>Eine neue Klasse von Microlasern</b>	<b>93</b>
5.1	Microlaser - Einführung und kurze historische Übersicht . . . . .	93
5.2	Physikalisches Modell von Halbleiter-Lasern . . . . .	94
5.3	Strahlenoptisches Modell und Konstruktion geeigneter asymmetrischer Resonanzkäfte . . . . .	97
5.4	Numerische Konstruktion von IWG-Lasern . . . . .	98
5.4.1	Berechnung des quasi-klassischen Spektrums . . . . .	100
5.4.2	Störung der Eingangsdaten . . . . .	101
5.5	Experimentelle Daten von IWG-Lasern . . . . .	104
<b>6</b>	<b>Anhang</b>	<b>108</b>
	<b>Literatur</b>	<b>111</b>
	<b>Erklärung</b>	<b>117</b>
	<b>Lebenslauf</b>	<b>119</b>

## 1 Mathematische Billards

Ein klassisches mechanisches Problem besteht in der Beschreibung der möglichen Wege, die eine reibungsfrei auf einem ideal ebenen Billardtisch laufende punktförmig gedachte Billardkugel bei vollelastischen Stößen mit der Bande nehmen kann. Letzteres bedeutet, daß die Kugel nach dem Gesetz „*Einfallswinkel=Ausfallswinkel*“ an der Bande reflektiert wird. In der mathematischen Abstraktion wird der Billardtisch beschrieben durch ein konvexes beschränktes Gebiet  $\Omega \subset \mathbb{R}^2$ , welches durch eine geschlossene glatte Kurve berandet ist.

Bekanntlich kann in diesem Fall die Kurve durch die Bogenlänge parametrisiert werden. Ist  $L$  die Länge der Randkurve, so gibt es also eine glatte Funktion  $c : [0, L] \rightarrow \mathbb{R}^2$  mit  $c(0) = c(L)$ , so daß

$$\partial\Omega = \{c(s) | s \in [0, L]\}$$

ist.

Die Funktion  $H(s, s') : [0, L] \times [0, L] \rightarrow \mathbb{R}$  mit

$$H(s, s') := \|c(s) - c(s')\|, \quad \|\cdot\| \text{ der Euklidische Abstand,}$$

heißt *erzeugende Funktion* des Billards. Eine Billardkugel, die den Rand im Punkt  $z = c(s)$  trifft und deren Stoßrichtung mit dem Tangentialvektor  $c'(s)$  den Einfallswinkel  $\theta$  bildet, wird so reflektiert, daß sie als nächstes den Rand im Punkt  $z' = c(s')$  unter dem Winkel  $\theta'$  mit dem Tangentialvektor  $c'(\theta')$  trifft.  $(z, \theta) \mapsto (z', \theta')$  definiert eine (von der gewählten Parametrisierung des Randes unabhängige) Abbildung

$$P : \partial\Omega \times (0, \pi] \rightarrow \partial\Omega \times (0, \pi].$$

Der Weg einer Billardkugel wird nun beschrieben durch eine diskrete Folge  $(z_i, \theta_i)_{i \geq 1}$  mit

$$(z_{i+1}, \theta_{i+1}) = P(z_i, \theta_i). \quad (1.1)$$

Die Abbildung  $P$  heißt *Billard-Abbildung* und beschreibt ein diskretes dynamisches System mit Phasenraum  $C := \partial\Omega \times (0, \pi]$ . Das dynamische System (1.1) zum Gebiet  $\Omega$  wird im folgenden oft auch als das zu  $\Omega$  gehörende *Billard-System*  $(\Omega, P)$  genannt. Wenn die Abhängigkeit von  $\Omega$  aus dem Zusammenhang erkennbar ist, wird nur die Bezeichnung *Billard-System* benutzt und  $\Omega$  als *Billard* bezeichnet.

### 1.1 Eigenschaften von Billard-Systemen

In diesem Kapitel werden zunächst einige bekannte Eigenschaften der Billard-Abbildung in Erinnerung gerufen (vgl. [KH95]). Ziel ist es, die Struktur des Phasenraumes  $C$  unter Anwendung der Billard-Abbildung  $P$  besser zu verstehen.

Unabhängig von der Geometrie von  $\Omega$  ist die Billard-Abbildung stets flächen- und orientierungserhaltend.

Nach der Definition ist

$$\cos \theta = \frac{\langle c'(s), z' - c(s) \rangle}{\|c'(s)\| \cdot \|z' - c(s)\|}.$$

Andererseits gilt

$$\begin{aligned} \frac{\partial H}{\partial s} &= \frac{1}{2} \frac{2[\langle c'(s), c(s) \rangle - \langle c'(s), z' \rangle]}{\|z' - c(s)\|} = -\frac{\langle c'(s), z' - c(s) \rangle}{\|c'(s)\| \cdot \|z' - c(s)\|} \\ &= -\cos \theta. \end{aligned} \tag{1.2}$$

Analog folgert man die Gleichung

$$\frac{\partial H}{\partial s'} = \cos \theta'. \tag{1.3}$$

Aus diesen beiden Gleichungen folgt, daß die Flächenform  $\sin \theta d\theta \wedge ds$   $P$ -invariant ist, denn aus

$$dH = \frac{\partial H}{\partial s} ds + \frac{\partial H}{\partial s'} ds' = -\cos \theta ds + \cos \theta' ds'$$

folgt

$$0 = d^2 H = -\sin \theta d\theta \wedge ds + \sin \theta' d\theta' \wedge ds'.$$

Für die graphische Darstellung des Phasenraumes werden oft die Koordinaten  $r = \cos \theta$  und  $s$  benutzt. In diesen Koordinaten folgen aus Gleichungen (1.2) und (1.3)

$$\begin{aligned} \frac{\partial}{\partial s} H(s, s') &= -r, \\ \frac{\partial}{\partial s'} H(s, s') &= r'. \end{aligned}$$

Als Korollar gewinnt man die folgenden Eigenschaften eines Billardsystems

### 1.1 Korollar (Eigenschaften eines Billard-Systems).

(B1) Die Billard-Abbildung besitzt keine asymptotisch stabilen Gleichgewichtspunkte.

**Beweis:** (B1) folgt aus der Flächenerhaltung.  $\square$

**1.2 Definition.**

- (i) Ein Billard-System  $\Omega$  heißt *konvex*, falls die Krümmung  $\kappa(z)$  des Randes  $\partial\Omega$  in jedem Punkt  $z$  die Ungleichung

$$\kappa(z) \geq 0$$

erfüllt.

- (ii) Ein Billard-System  $\Omega$  heißt *strikt konvex*, falls sogar

$$\kappa(z) > 0$$

gilt.

Um die Billard-Abbildung  $P$  gut zu verstehen, ist es notwendig,  $P$ -invariante Mengen, also Mengen  $M \subseteq C$  mit der Eigenschaft  $P(M) \subseteq M$ , zu erkennen und zu untersuchen. Beispiele für invariante Mengen sind die periodischen Orbits.

Das Tupel  $((z_1, \theta_1), \dots, (z_p, \theta_p)) \in C^p$  mit der Eigenschaft  $(z_{j+1}, \theta_{j+1}) = P(z_j, \theta_j)$  und  $(z_1, \theta_1) = P(z_p, \theta_p)$  beschreibt einen  $p$ -periodischer Orbit. Im folgenden werden nur spezielle  $p$ -periodische Orbits mit der zusätzlichen Eigenschaft betrachtet, daß die Punkte  $z_j$  den Rand entgegen dem Uhrzeigersinn durchlaufen (vergleiche Abbildung 1). Alle Sätze können aber leicht auf beliebige  $p$ -periodische Orbits verallgemeinert werden.

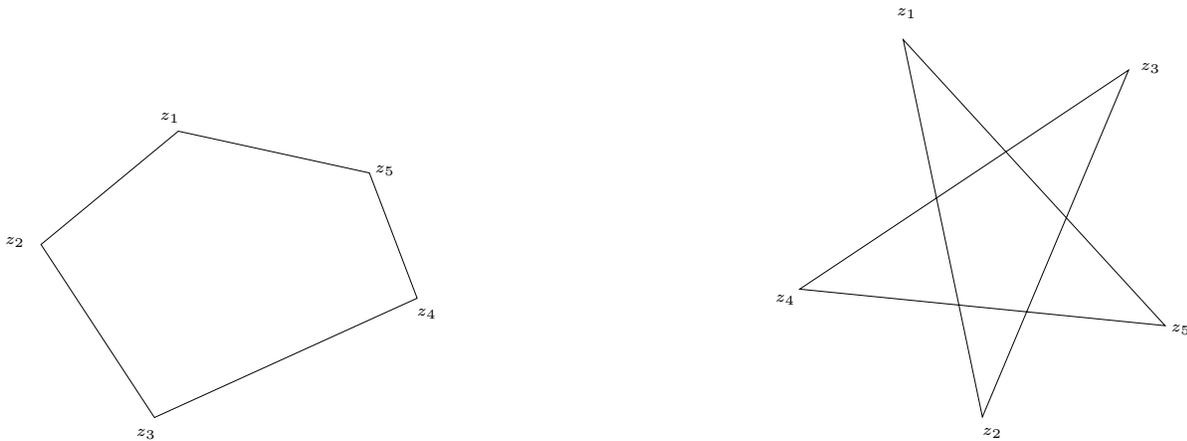


Abbildung 1: Links: Im folgenden werden nur noch periodische Orbits dieses Typs betrachtet. Die Punkte  $(z_1, \dots, z_5)$  durchlaufen den Billardrand entgegen dem Uhrzeigersinn. Rechts: Diese Typen von periodischen Orbits werden nicht betrachtet. Die Sätze können aber leicht auf diese Typen verallgemeinert werden.

Als Ausgangspunkt für das Auffinden von periodischen Orbits dient die folgende Beobachtung:

Liegen die Punkte  $z_{-1}, z_0, z_1$  auf einem Teil eines Orbits, dann bilden das Segment, das  $z_{-1}$  und  $z_0$  verbindet, und das Segment, das  $z_0$  mit  $z_1$  verbindet, den selben Winkel mit der Tangente im Punkt  $z_0$ . Sind  $s_i$  die zu den  $z_i$  gehörigen Längenparameter, so gilt wegen Gleichungen (1.2) und (1.3)

$$\frac{d}{ds}H(s_{-1}, s) + \frac{d}{ds}H(s, s_1) = 0, \quad \text{bei } s = s_0.$$

Also ist  $z_0$  ein kritischer Punkt des Funktionals  $s \mapsto H(s_{-1}, s) + H(s, s_1)$ .

Diese Beobachtung motiviert einen variationellen Ansatz zur Bestimmung von periodischen Orbits in Billard-Systemen. Die Grundlage für diese Methode ist die Untersuchung des *Längenfunctionals*

$$L(s_1, \dots, s_p) := \sum_{i=1}^p H(s_i, s_{i+1}),$$

wobei  $s_i$  die Bogenlängen-Parameter zu einer Folge von Punkte  $(z_i, \theta_i)$  unter der Billard-Abbildung sind. Ein  $p$ -periodischer Orbit entspricht einem kritischen Punkt dieses Funktionals.

Man kann folgenden Satz elementar beweisen, der auf Birkhoff zurück geht ([Sin76]).

**1.3 Satz (Birkhoff).** *Sei  $\Omega$  ein strikt konvexes Billard mit dem  $C^1$ -Rand  $\partial\Omega$ . Für jedes  $p \geq 2$ , existieren zwei geometrisch unterschiedliche  $p$ -periodische Orbits.*

Satz (1.3) liefert keine Aussage über das Stabilitätsverhalten der gefundenen periodischen Orbits. Für den Spezialfall  $p = 2$  kann jedoch das obige Ergebnis erweitert werden ([KH95]).

**1.4 Satz.** *Sei  $\Omega$  ein strikt konvexes Billard mit dem  $C^2$ -Rand  $\partial\Omega$ . Dann existieren mindestens zwei verschiedene 2-periodische Orbits: für den einen ist der Abstand zwischen den jeweiligen Randpunkten der Durchmesser  $\text{diam } \Omega$  von  $\Omega$ , für den anderen ist es die Weite  $\text{width } \Omega$  von  $\Omega$ . Der erste Orbit ist hyperbolisch, der zweite ist elliptisch.*

### 1.5 Definition.

- (i) Eine *invariante Kurve*  $\Gamma \subset C$  von  $P$  ist eine  $P$ -invariante Menge in  $C$ , die Graph einer stetigen Funktion  $\phi : \partial\Omega \rightarrow (0, \pi]$  ungleich der konstanten Funktionen 0 und  $\pi$  ist, also  $\Gamma = \{(z, \phi(z)) | z \in \partial\Omega\}$  und  $P(\Gamma) \subseteq \Gamma$ .
- (ii) Als *Kaustik* eines Billards  $\Omega$  bezeichnet man eine stückweise stetige, glatte Kurve  $\gamma$  im Inneren von  $\Omega$ , mit der Eigenschaft, daß eine Trajektorie, die tangential zu  $\gamma$  ist, auch nach Reflektion am Rand tangential zu  $\gamma$  bleibt.

- (iii) Eine Kaustik  $\gamma$  heißt *durch eine invariante Kurve definiert*, falls die Menge  $\Gamma$  aller Phasenraumpunkte  $(z, \theta)$ , wobei  $z$  Schnittpunkt einer Tangente von  $\gamma$  mit dem Rand  $\partial\Omega$  und  $\theta$  der zugehörige Schnittwinkel ist, eine invariante Kurve ist.

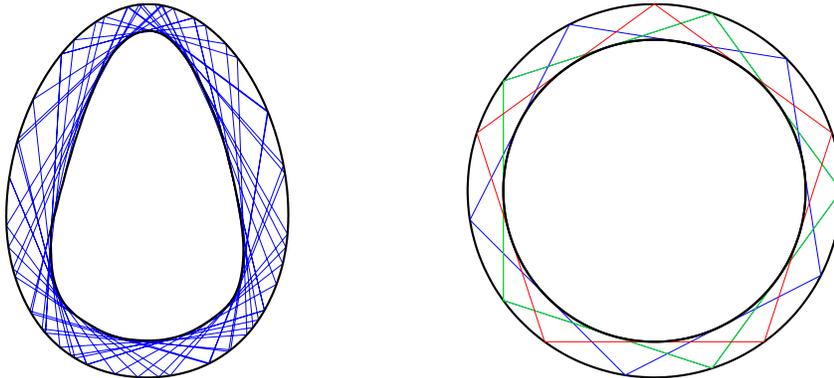


Abbildung 2: Links: Die Kurve im Inneren ist eine Kaustik. Ein Orbit bleibt auch nach Reflektion am Rand tangential zur Kurve im Inneren. Rechts: Eine durch eine invariante Kurve definierte Kaustik im Inneren des Kreises. Die invariante Kurve ist Graph der Funktion  $\phi(s) = \frac{2\pi}{5}$ .

Lazutkin ([Laz73b]) konnte für strikt konvexe Billard-Systeme die Existenz einer Familie von Kaustiken in der Nähe des Randes zeigen.

**1.6 Satz (Lazutkin).** *Sei  $\Omega$  ein strikt konvexes Billard-System mit glattem Rand. Dann existiert zu jedem hinreichend kleinen  $\varepsilon > 0$  eine Familie von invarianten Kurven  $\{\Gamma_\alpha\}$ , so daß stets*

$$\Gamma_\alpha \subset \partial\Omega \times (0, \varepsilon)$$

*ist.*

Zusätzlich zur Eigenschaft (B1) gelten für strikt konvexe Billards die:

**Eigenschaften eines strikt konvexen Billard-Systems:**

- (B2) Das System besitzt einen elliptischen und einen hyperbolischen 2-periodischen Orbit.
- (B3) Für jede natürliche Zahl  $p \geq 3$  besitzt das System mindestens 2 verschiedene  $p$ -periodische Orbits.
- (B4) Der Phasenraum  $C$  besitzt eine Familie  $\{\Gamma_\alpha\}$  von invarianten Kurven mit der Eigenschaft  $\Gamma_\alpha \subset \partial\Omega \times (0, \varepsilon)$  für hinreichend kleines  $\varepsilon > 0$ . Diesen Streifen nennt man *Lazutkin-Streifen*.

## 1.2 Billard-Systeme unter kleinen Deformationen von $\Omega$

Da bei der numerischen Konstruktion von Billards zwangsläufig kleine Fehler gemacht werden, ist es notwendig zu verstehen, wie sich die Billard-Abbildung  $P$  unter kleinen Deformationen von  $\Omega$  verhält.

Nach ([KH95], (9.2.1)) ist jede Billard-Abbildung eine sogenannte *Twist-Abbildung*. Insbesondere existiert eine Koordinatentransformation  $(z, \theta) \mapsto (J, \phi)$ , so daß sich das Billard-System (1.1) in den neuen Koordinaten als

$$(J_{i+1}, \phi_{i+1}) = (J_i, \phi_i + 2\pi\Lambda(J_i))$$

mit einer Funktion  $\Lambda(J)$  schreiben läßt, welche mit wachsendem  $J$  abnimmt ([GH96]). Die invarianten Kurven des Billard-Systems (1.1) gehen über in Geraden  $J = \text{const}$ .

Man betrachtet nun Störungen  $P_\varepsilon$  der Form

$$(J_{i+1}, \phi_{i+1}) = P_\varepsilon(J_i, \phi_i) = (J_i + \varepsilon f(J_i, \phi_i, \varepsilon), \phi_i + 2\pi\Lambda(J_i) + \varepsilon g(J_i, \phi_i, \varepsilon))$$

mit beschränkten und in der  $\phi$ -Komponente  $2\pi$ -periodischen,  $C^r$ -Funktionen  $f$  und  $g$ .

Für diese gilt der berühmte *Twist-Satz* von Moser.

**1.7 Satz (Twist-Satz, [GH96], [Mos73]).** *Sei  $\Lambda(J)$  eine  $C^r$ -Funktion mit  $r \geq 5$  und  $|\Lambda'(J)| \geq \nu > 0$  für  $a \leq J \leq b$  und  $\varepsilon > 0$ . Dann existiert ein  $\delta = \delta(\varepsilon, \Lambda(J))$ , so daß jede Abbildung  $P_\varepsilon$  mit*

$$\sup_{(J, \phi) \in R} \{ \|\varepsilon f\|_r + \|\varepsilon g\|_r \} < \nu\delta,$$

mit  $R = \{(J, \phi) | a \leq J \leq b\}$  eine invariante Kurve  $\Gamma_\varepsilon$  von der Form

$$J = J^0 + U(\psi), \quad \phi = \phi + V(\psi)$$

in  $R$  besitzt, wobei  $a < J^0 < b$  ist und  $U, V$  stetig differenzierbare  $2\pi$ -periodische Funktionen mit

$$\|U\|_1 + \|V\|_1 < \varepsilon$$

sind. Die induzierte Abbildung auf dieser Kurve ist gegeben durch

$$\phi \mapsto \phi + 2\pi\lambda$$

mit einer irrationalen Zahl  $\lambda$ , die unendlich viele Bedingungen der Form

$$\left| \lambda - \frac{n}{m} \right| \geq \beta m^{-\alpha} \tag{1.4}$$

mit  $\beta, \alpha > 0$  für alle  $n, m \in \mathbb{N}$  erfüllt. Zu jedem solchen  $\lambda$  im Bild von  $\Lambda$  gibt es eine invariante Kurve.

Dieser Satz besagt also, daß im Fall eines Billard-Systems eine invariante Kurve, die nur aus  $p$ -periodischen Orbits besteht (ein *resonanter Torus* gemäß [GH96, Seite 220]), nach einer kleinen Störung aufbricht und sich eine Kette von endlich vielen periodischen Orbits bildet. Diese Orbits sind abwechselnd elliptisch und hyperbolisch [GH96, Seite 221].

Im Phasendiagramm bildet sich demnach eine Kette von Inseln. Die zu den hyperbolischen Orbits gehörenden invarianten Mengen sind zwischen benachbarten „irrationalen“ Kurven eingesperrt, welche unter kleinen Störungen erhalten bleiben. Der Verlauf einer Trajektorie, die sich in einem solchen Gebiet befindet, ist sehr kompliziert. Sie bleibt in diesem Gebiet gefangen.

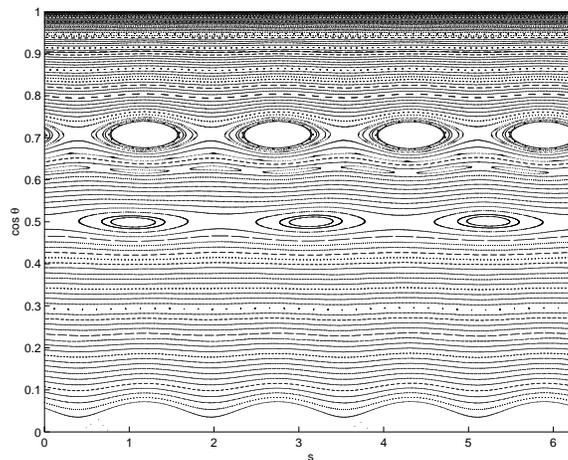


Abbildung 3: Phasenportrait eines leicht deformierten Kreises. Die Randkurve ist im Text definiert. Es wurden 250 Trajektorien mit jeweils 1000 Reflektionen berechnet. Als Koordinaten wurden  $(s, \cos \theta)$  gewählt.

In Abbildungen 3 - 5 ist ein numerisches Beispiel angegeben, das die Aussagen des Twist-Satzes veranschaulicht. Hierzu wurde das Phasendiagramm eines Billards berechnet, dessen Randkurve in Polarkoordinaten durch

$$r(\phi) = 1 + 0.001 \sin(4\phi) + 0.0002 \cos(3\phi)$$

gegeben ist. Diese kann als kleine Störung der Kreislinie  $r(\phi) = 1$  betrachtet werden.

Zur graphischen Darstellung des Phasenportraits wurden 250 Trajektorien in unterschiedlichen Bereichen des Phasenraumes mit jeweils 1000 Reflektionen am Rand numerisch simuliert (vgl. Abbildung 3). In Abbildung 4 ist eine Vergrößerung des Lazutkin-Streifens berechnet worden. Es sind deutlich einige invariante Kurven zu erkennen.

Die invarianten Kurven mit rationaler Rotation werden zerstört und es entstehen Ketten von Inseln wie oben beschrieben. Da diese Inseln im Phasenraum ein sehr kleines Lebesgue-Maß

haben, können sie im Lazutkin-Gebiet nur schwer numerisch erfaßt werden.

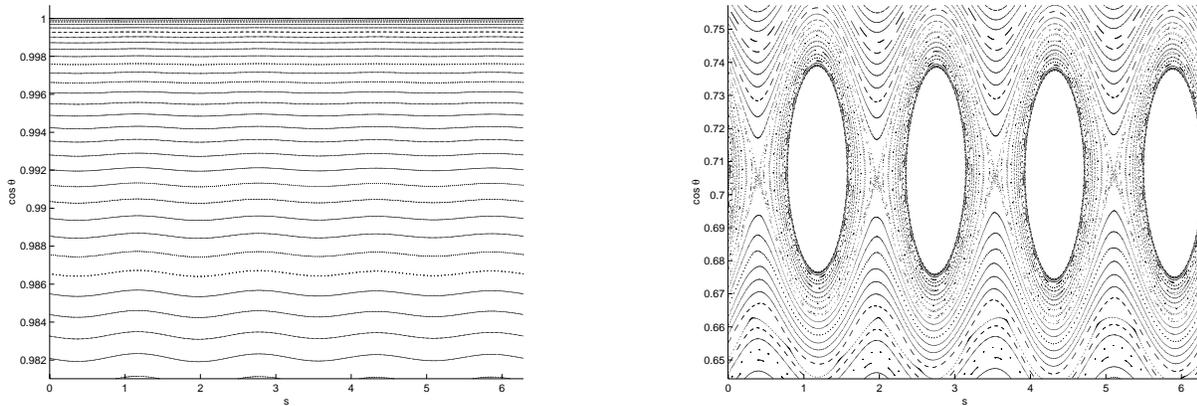


Abbildung 4: Links: Vergrößerung des Lazutkin-Streifens. Rechts: Vergrößerung des KAM-Streifens entlang der 4-periodischen Insel-Kette.

Jedoch kann das Aufbrechen in Inseln von invarianten Kurven, die aus 3- bzw. 4-periodischen Orbits bestehen, sehr deutlich beobachtet werden. Die Abbildung 4 enthält eine Vergrößerung der 4-periodischen Kette, bei der man die Abfolge von hyperbolischen und elliptischen Fixpunkten sehen kann.

Um die Phasenraumstruktur einer solchen Kette zu veranschaulichen, wurde eine weitere Vergrößerung der Inselkette berechnet. Man erkennt zwei hyperbolische Fixpunkte, die durch einen homoklinen Orbit miteinander verbunden sind. In der Mitte befindet sich ein elliptischer Fixpunkt, der von invarianten Kurven umgeben ist. An den Bildrändern werden invariante Kurven deutlich, die zu irrationalen Kurven des ungestörten Kreises gehören und die Inselkette von oben und unten einschließen. Je weiter man sich vom elliptischen Fixpunkt entfernt, desto komplizierter und chaotischer wird die Struktur des Phasenraumes.

### Numerische Anmerkung:

Die vorangehenden Bilder wurden wie folgt berechnet:

Da der Billard-Rand  $\partial\Omega$  analytisch in Polarkoordinaten angegeben ist,  $\partial\Omega = \{r(\phi)\}$ , läßt sich in jedem Randpunkt auch die Tangente analytisch berechnen. Im Unterschied zur oben benutzten Bogenlänge  $s$  zur Darstellung der Randpunkte wird nun der Polarwinkel  $\phi$  benutzt.

Die Polarkoordinatendarstellung des ausgehenden Strahls ist aus Steigung und Aufsprungspunkt unmittelbar zu berechnen. Dieser Strahl schneidet  $\partial\Omega$  in den Punkten  $\phi$  und  $P(\phi)$ . Der Schnittpunkt  $P(\phi)$  wird mit dem Newton-Verfahren bestimmt.

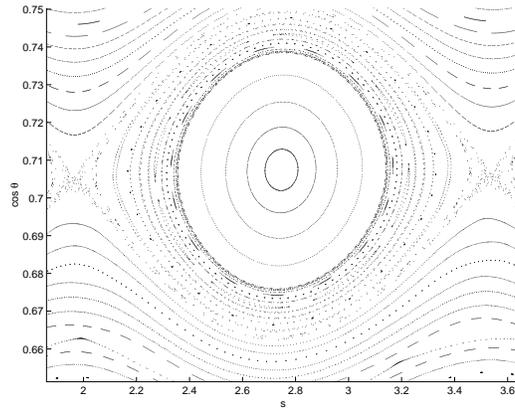


Abbildung 5: Vergrößerung einer 4-periodischen Insel.

### 1.3 $p$ -Lazutkin-Billards

Eine Kaustik, deren Strahlen  $p$ -periodische Orbits bilden, heißt  $p$ -periodische Kaustik. Hier wird wie oben angenommen, daß die Folge der Reflektionspunkte des periodischen Orbits am Rand entgegen dem Uhrzeigersinn verläuft. Man definiert ([BZ03])

**1.8 Definition.** Ein strikt konvexes Billard-System  $(\Omega, P)$  heißt  $p$ -Lazutkin-Billard-System, falls eine stetige, einparametrische, nicht-entartete, volle Familie von  $p$ -periodischen Orbits existiert, oder äquivalent hierzu, falls das System eine  $p$ -periodische Kaustik im Inneren von  $\Omega$  besitzt. Dabei heißt eine Familie von  $p$ -periodischen Orbits *nicht-entartet*, falls die Reflektionspunkte jedes Elements paarweise verschieden sind und nicht drei aufeinanderfolgende auf einer Geraden liegen. Eine solche Familie heißt *voll*, wenn jeder Punkt  $z \in \partial\Omega$  als Reflektionspunkt eines Elements der Familie vorkommt.

Ein  $p$ -Lazutkin-Billard besitzt somit zusätzlich zu den Eigenschaften (B1)-(B4) noch eine invariante Kurve  $\Gamma = \{(s, \phi(s)) | s \in \partial\Omega\}$  mit der Eigenschaft, daß jeder Punkt  $(z_1, \theta_1) \in \Gamma$  in einem  $p$ -periodischen Orbit  $((z_1, \theta_1), \dots, (z_p, \theta_p))$  mit  $(z_{i+1}, \theta_{i+1}) = P(z_i, \theta_i)$  liegt, wobei die Indizierung modulo  $p$  zu verstehen ist. Insbesondere dürfen die oben referierten Ergebnisse über kleine Deformationen des Randes  $\partial\Omega$  angewendet werden. Da bei der numerischen Konstruktion eines  $p$ -Lazutkin-Billards zwangsläufig durch die Diskretisierung kleine Fehler gemacht werden, wird diese invariante Kurve in einem numerisch bestimmten Phasenportrait nicht zu sehen sein. Jedoch läßt sich eine Inselkette – wie im obigen Abschnitt erläutert – beobachten. Der Phasenraum  $C$  zusammen mit der Billard-Abbildung  $P$  kann dann in diesem Bereich mit Hilfe der KAM-Theorie verstanden werden.

Zusammenfassend erhält man zu den Eigenschaften (B1)-(B4) zusätzlich die

### Eigenschaften eines $p$ -Lazutkin-Billards

- (B5) Der Phasenraum besitzt eine invariante Kurve aus  $p$ -periodischen Orbits. Diese entspricht einer  $p$ -periodischen Kaustik im Inneren von  $\Omega$ .
- (B6) Durch kleine Störungen (zum Beispiel bei der numerischen Konstruktion) entsteht eine Inselkette in einem KAM-Streifen entlang der invarianten Kurve.

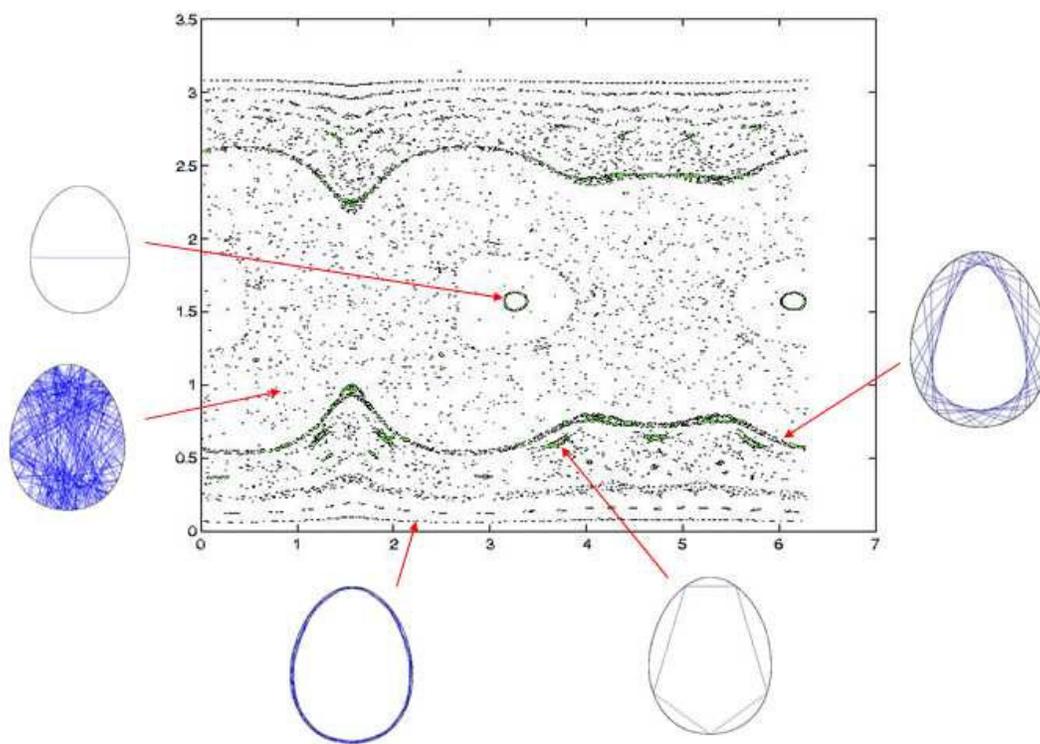


Abbildung 6: Illustration der Eigenschaften eines 4-Lazutkin-Billards anhand eines Phasenportraits. Der Billard wurde mit den Methoden aus Kapitel 2 konstruiert. Die Inselkette erkennt man in der Mitte des Bildes. Von Links nach rechts: der stabile 2-periodische Orbit, eine Trajektorie im chaotischen Regime, eine Trajektorie im Lazutkin-Streifen, ein (stabiler) 5-periodischer Orbit, eine Trajektorie im KAM-Streifen der invarianten Kurve zur 4-periodischen Kaustik.

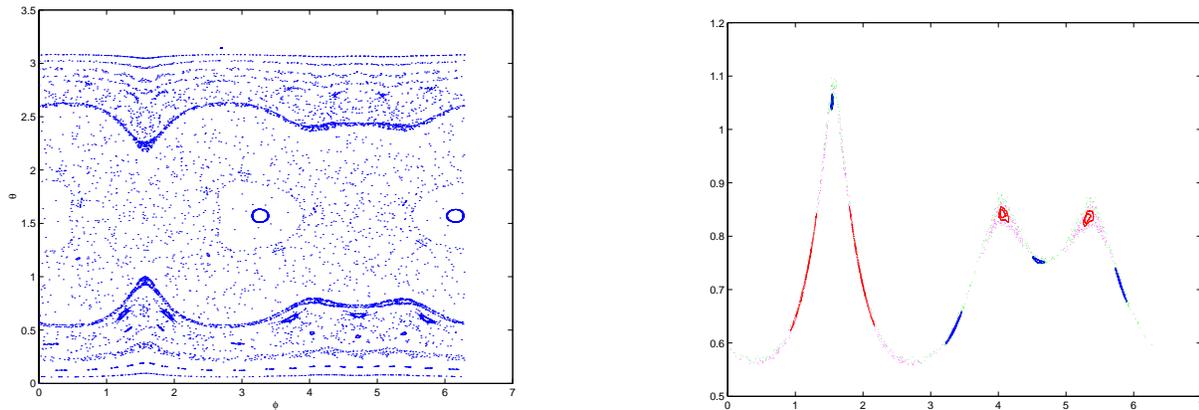


Abbildung 7: Links: Phasenportrait eines 4-Lazutkin-Billard. Rechts: Vergrößerung des KAM-Streifens.

Nicht-triviale Beispiele für  $p$ -Lazutkin-Billards sind Kurven mit konstanter Weite ([Inn88]), die eine Familie von 2-periodischen Orbits besitzen. Die Kaustik entspricht hier der Evolute des Randes  $\partial\Omega$ . Viel schwieriger ist es, Billards mit Familien von 3- oder mehr periodischen Orbits zu konstruieren. Eine sehr komplizierte Konstruktion für 3-periodische Orbits ist in [KG95] erschienen.

Allgemein ist die Frage nach der lokalen Beschaffenheit der Menge aller  $p$ -Lazutkin-Billards von großem Interesse.

Dieses Problem wurde kürzlich von Y. Baryshnikov und V. Zharnitsky gelöst ([BZ03]), die den folgenden allgemeinen Existenzsatz für  $p$ -Lazutkin-Billards beweisen konnten. Ein numerisches Konstruktionsverfahren wird zum erstenmal im nächsten Kapitel vorgestellt.

**1.9 Satz (Y. Baryshnikov, V. Zharnitsky).** *Sei  $\Omega$  ein  $p$ -Lazutkin-Billard mit glattem Rand  $\partial\Omega$ . Dann ist die Menge von Kurven, die bezüglich der Sobolev- $H^2$ -Topologie in einer hinreichend kleinen Umgebung von  $\partial\Omega$  liegen und ebenfalls eine stetige Familie von  $p$ -periodische Orbits besitzen, eine Mannigfaltigkeit in einem Hilbertraum, die diffeomorph zu einer Untermannigfaltigkeit der Kodimension  $p - 1$  im Sobolevraum aller geschlossenen Kurven im  $\mathbb{R}^p$  ist.*

Im folgenden wird eine Beweisskizze angegeben, die als Interpretation des in [BZ03] angegebenen Beweises zu verstehen ist, und insbesondere die konstruktiven Aspekte herausarbeitet, die für die späteren numerischen Konstruktionen wesentlich sind.

### Beweisskizze von Satz 1.9

Jeder  $p$ -periodische Orbit von  $\Omega$  wird durch die  $p$  Reflektionspunkte  $z_1, \dots, z_p$  auf dem Rand beschrieben. Diese Punkte können aufgefaßt werden als Eckpunkte eines  $p$ -Ecks, bei dem die

Indizierung der Ecken so gewählt ist, daß  $z_{j+1} = P(z_j)$  mit der Billardabbildung  $P$  ist, wobei die Indizes wieder modulo  $p$  zu verstehen sind.

Diese Anordnung der Punkte  $(z_1, \dots, z_p)$  heißt *Billard-Anordnung*. Sie ist zu unterscheiden von der *geometrischen Anordnung*, bei der die Reihenfolge beschrieben wird, mit der die Reflektionspunkte auf dem Rand auftreten, wenn dieser ausgehend von  $z_1$  im Gegenuhrzeigersinn durchlaufen wird.

Im folgenden kann man sich beschränken auf den Fall, daß Billard-Anordnung und geometrische Anordnung übereinstimmen, da man mittels einer eindeutig bestimmten Permutation aus  $\mathcal{S}^p$  die beiden Anordnungen ineinander überführen kann.

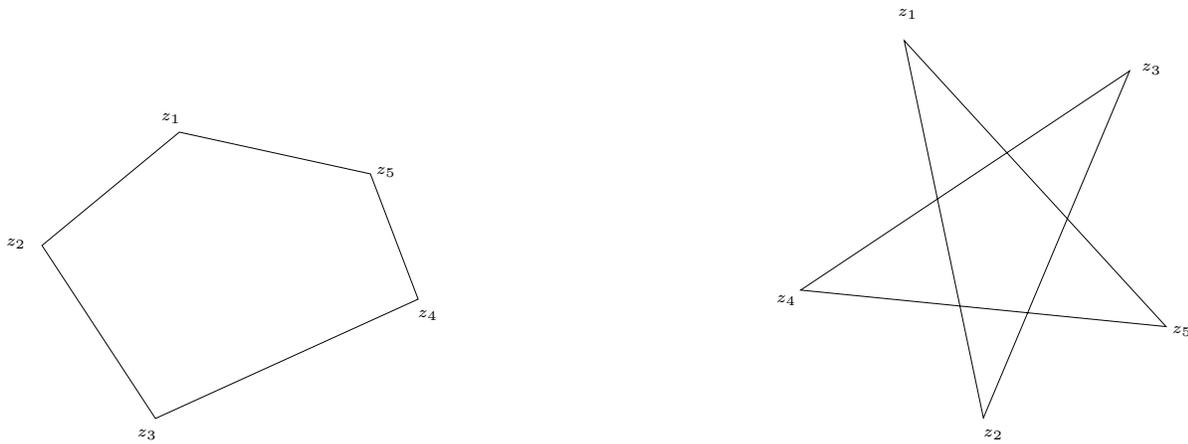


Abbildung 8: Links: Die Billardanordnung  $(z_1, \dots, z_5)$  stimmt mit der geometrischen Anordnung überein. Rechts: Hier ist die geometrische Anordnung  $(z_1, z_4, z_2, z_5, z_3)$ .

Sei  $V = \mathbb{C}^p$  und  $V' \subset V$  die Untermenge mit der Eigenschaft, daß  $z_j \neq z_{j+1}$  ist und die Strahlen  $z_{j-1}z_j$  und  $z_jz_{j+1}$  dreier aufeinanderfolgender Punkte nicht auf einer Geraden liegen. Nach der Voraussetzung gibt es eine stetige Familie

$$t \in I = [0, 1] \mapsto z(t) = (z_1(t), \dots, z_p(t)) \in V',$$

mit  $z_j(1) = z_{j+1}(0)$  für  $j = 1, \dots, p$ .

Sei  $\mathcal{L} : V \rightarrow \mathbb{R}$  definiert durch

$$\mathcal{L}(z_1, \dots, z_p) := \sum_{j=1}^p |z_j - z_{j+1}|.$$

Nach [Laz73b] ist  $\mathcal{L}(z_1(t), \dots, z_p(t)) = \text{const}$  für alle  $t \in I$ . Ohne Einschränkung kann durch Umskalieren im folgenden stets  $\text{const} = 1$  angenommen werden.

Man sieht durch Differenzieren des Funktionals  $\mathcal{L}$ , daß

$$M := \{z \in V' \mid \mathcal{L}(z) = 1\} \subset V$$

eine glatte,  $(2p - 1)$  - dimensionale reelle Untermannigfaltigkeit von  $V$  ist.

Die gegebene Familie von  $p$ -periodischen Orbits kann also als Kurve  $z : I \rightarrow M$  aufgefaßt werden.

Für alle  $t \in I$  gibt das Tupel  $(z_1(t), \dots, z_p(t))$  die Billard-Anordnung des entsprechenden  $p$ -periodischen Orbits an. Nach der Definition der Billard-Abbildung bedeutet dies, daß die Winkelhalbierende des Orbit-Segments  $z_{j-1}(t)z_j(t)z_{j+1}(t)$  orthogonal zur Tangente im Punkt  $z_j(t)$  ist. Dies bedeutet, daß für alle  $t \in I$  und  $j = 1, \dots, p$  die Gleichungen

$$\dot{z}_j(t) = u_j(t) \left( \frac{z_{j+1} - z_j}{|z_{j+1} - z_j|} + \frac{z_j - z_{j-1}}{|z_j - z_{j-1}|} \right) \quad (1.5)$$

mit reellen, glatten Funktionen  $u_j(t)$  erfüllt sind, die man als *Kontroll-Variablen* der Kurve  $z(t)$  bezeichnet.

### 1.10 Definition.

- (i) Eine *Distribution* ordnet auf stetige Weise jedem Element  $\mathfrak{x}$  aus einer glatten, zusammenhängenden Mannigfaltigkeit  $M$  einen linearen Unterraum  $\Delta_{\mathfrak{x}}$  des Tangentialraums  $T_{\mathfrak{x}}M$  von  $M$  in  $\mathfrak{x}$  zu.
- (ii) Eine Kurve  $\gamma : I \rightarrow M$  heißt *horizontal* zur Distribution  $\Delta$ , falls für alle  $t \in I$

$$\dot{\gamma}(t) \in \Delta_{\gamma(t)}$$

ist.

Für  $z = (z_1, \dots, z_p) \in M$  definieren

$$L_j := \frac{z_{j+1} - z_j}{|z_{j+1} - z_j|} + \frac{z_j - z_{j-1}}{|z_j - z_{j-1}|}, \quad j = 1, \dots, p \quad (1.6)$$

Tangentialvektoren von  $M$  im Punkt  $z$ . Die Zuordnung

$$z \mapsto \mathcal{B}_z := \langle L_1, \dots, L_p \rangle \subset T_z M \quad (1.7)$$

definiert dann eine Distribution auf  $M$ , die als *Birkhoff-Distribution* bezeichnet wird. Die Ausgangskurve  $z : I \rightarrow M$  ist wegen (1.5) dazu horizontal.

Die Birkhoff-Distribution besitzt die folgende bemerkenswerte Eigenschaft.

**1.11 Satz.** In jedem Punkt  $z \in M$  gilt

$$\langle L_1, \dots, L_p, [L_1, L_2], [L_2, L_3], \dots, [L_p, L_1] \rangle = T_z M,$$

wobei  $[\cdot, \cdot]$  die Lie-Klammer bezeichnet. In der Terminologie von [Mon02, Seite 23] bedeutet dies, daß die Birkhoff-Distribution vollständig nichtholonom ist.

**Beweis:** Zunächst werden einige Abkürzungen eingeführt, vergleiche Abbildung 9. Sei

$$\begin{aligned} r_j &:= z_{j+1} - z_j & l_j &:= |z_{j+1} - z_j| \\ e_j &:= \frac{z_{j+1} - z_j}{l_j} & v_j &:= e_j + e_{j-1} \\ \beta_j &:= \angle e_{j-1} e_j & \alpha_j &= \frac{\pi - \beta_j}{2}. \end{aligned}$$

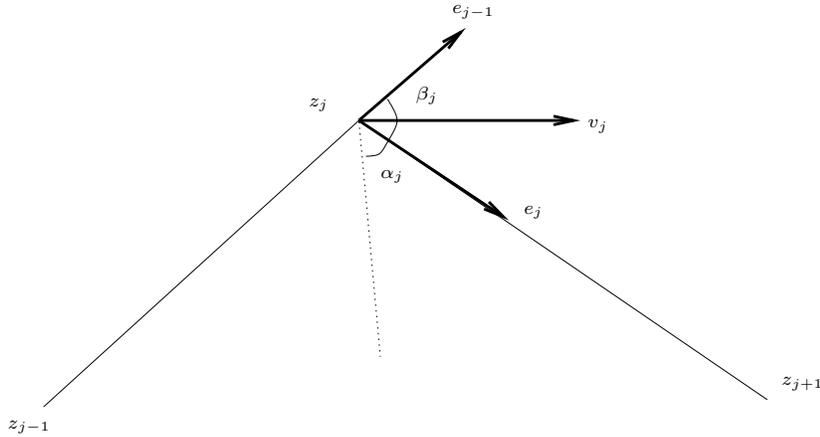


Abbildung 9: Geometrische Situation im Punkt  $z_j$ .

Mit den Koordinaten  $z_j = x_j + iy_j$  und den Standardkoordinaten-Differentialen als Basis schreiben sich die in (1.6) definierten Tangentialvektoren als

$$L_j = \left( \frac{x_{j+1} - x_j}{l_j} + \frac{x_j - x_{j-1}}{l_{j-1}} \right) \cdot \frac{\partial}{\partial x_j} + \left( \frac{y_{j+1} - y_j}{l_j} + \frac{y_j - y_{j-1}}{l_{j-1}} \right) \cdot \frac{\partial}{\partial y_j}. \quad (1.8)$$

$L_j$  wirkt also nur auf die  $j$ -te Komponente von  $z$ .

Im Punkt  $\tilde{z}_k = (0, \dots, z_k, \dots, 0)$  ergibt sich somit für  $j = 1, \dots, p$ ,  $k = 1, \dots, p$

$$L_j(\tilde{z}_k) = \begin{cases} 0 & j \neq k \\ \left( \frac{x_{j+1} - x_j}{l_j} + \frac{x_j - x_{j-1}}{l_{j-1}} \right) + i \cdot \left( \frac{y_{j+1} - y_j}{l_j} + \frac{y_j - y_{j-1}}{l_{j-1}} \right) = v_j & j = k \end{cases}.$$

Der Tangentialvektor  $L_j$  angewandt auf einen Punkt  $z \in M$  verrückt also die  $j$ -te Komponente von  $z$  orthogonal zur Winkelhalbierenden des Winkels  $\angle z_{j-1}z_jz_{j+1}$ . Für die Geschwindigkeit  $|L_j(\tilde{z}_k)| = |v_j| = |e_j + e_{j-1}|$  berechnet man mit dem Cosinus-Satz

$$\begin{aligned} |v_j|^2 &= |e_j|^2 + |e_{j-1}|^2 - 2|e_j||e_{j-1}|\cos(\pi - \beta_j) \\ &= 2(1 + \cos \beta_j) = 2(1 + \cos(\pi - 2\alpha_j)) = 2(1 - \cos^2 \alpha_j + \sin^2 \alpha_j) \\ &= 4\sin^2 \alpha_j. \end{aligned}$$

Faßt man die oben definierten  $l_j$  als Funktionen  $l_j : M \rightarrow \mathbb{R}, z \mapsto \sqrt{(x_{j+1} - x_j)^2 + (y_{j+1} - y_j)^2}$  auf, so können die Tangentialvektoren  $L_j$  als Differentialoperatoren auf  $l_j$  angewandt werden. Für  $j \neq k, k-1$  erhält man wegen (1.8)

$$L_j(l_k) = 0.$$

Für  $j = k$  ergibt sich wegen  $\frac{\partial}{\partial x_j} l_j = -\frac{x_{j+1} - x_j}{l_j}$  bzw.  $\frac{\partial}{\partial y_j} l_j = -\frac{y_{j+1} - y_j}{l_j}$  in der komplexen Schreibweise

$$\begin{aligned} L_j(l_j) &= -\left\{ \frac{x_{j+1} - x_j}{l_j} \left( \frac{x_{j+1} - x_j}{l_j} + \frac{x_j - x_{j-1}}{l_{j-1}} \right) + \frac{y_{j+1} - y_j}{l_j} \left( \frac{y_{j+1} - y_j}{l_j} + \frac{y_j - y_{j-1}}{l_{j-1}} \right) \right\} \\ &= -\operatorname{Re}(e_j \cdot \bar{v}_j) = -\operatorname{Re}(e_j \cdot \overline{e_j + e_{j-1}}) \\ &= -(1 + \cos \beta_j) = -2\sin^2 \alpha_j. \end{aligned}$$

Analog erhält

$$L_j(l_{j-1}) = 2\sin^2 \alpha_j.$$

Mit der Quotientenregel kann nun die Wirkung von  $L_j$  auf die  $e_k$  berechnet werden. Zunächst erhält man wegen (1.8) für  $j \neq k, k+1$

$$L_j e_k = 0.$$

Man findet weiter mit den obigen Ergebnissen

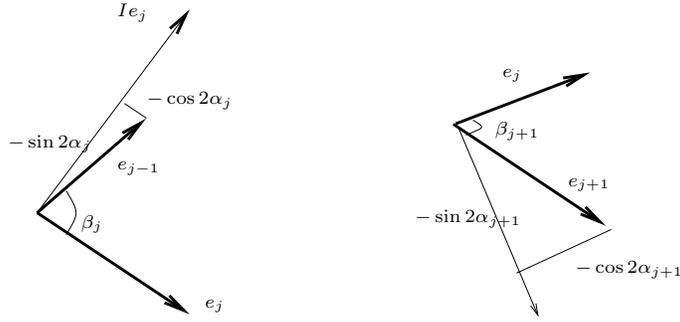
$$\begin{aligned} L_j e_j &= \frac{l_j L_j(z_{j+1} - z_j) + L_j(l_j)(z_{j+1} - z_j)}{l_j^2} = \frac{-l_j v_j + 2\sin^2 \alpha_j (z_{j+1} - z_j)}{l_j^2} \\ &= \frac{-(e_j + e_{j-1}) + 2\sin^2 \alpha_j \cdot e_j}{l_j} = -\frac{\cos 2\alpha_j \cdot e_j + e_{j-1}}{l_j} \end{aligned}$$

und analog

$$L_j e_{j+1} = \frac{\cos 2\alpha_{j+1} \cdot e_j + e_{j+1}}{l_j}.$$

Sei mit  $I$  die Drehung um  $\frac{\pi}{2}$  in  $\mathbb{C}$  abgekürzt. Anhand von Abbildung 10 sieht man die Gleichungen

$$\begin{aligned} \cos 2\alpha_j \cdot e_j + e_{j-1} &= -\sin 2\alpha_j \cdot I e_j \\ \cos 2\alpha_{j+1} \cdot e_j + e_{j+1} &= -\sin 2\alpha_{j+1} \cdot I e_j. \end{aligned}$$

Abbildung 10: Skizze der geometrischen Situation im Punkt  $z_j$ .

Zur Abkürzung sei  $C_j := [L_j, L_{j+1}]$ . Dann gilt offenbar  $C_j \tilde{z}_k = 0$  für  $k \neq j, j+1$ . Weiter findet man

$$\begin{aligned} C_j \tilde{z}_j &= L_j(L_{j+1}z_j) - L_{j+1}(L_jz_j) = -L_{j+1}e_j = -\frac{\sin 2\alpha_{j+1}}{l_j} Ie_j \\ C_j \tilde{z}_{j+1} &= L_j(L_{j+1}z_{j+1}) - L_{j+1}(L_jz_{j+1}) = L_je_j = \frac{\sin 2\alpha_j}{l_j} Ie_j. \end{aligned}$$

$C_j$  verrückt also die  $j$ -te Komponente von  $z$  orthogonal zum Segment  $z_{j+1}z_j$  und die  $j+1$ -te Komponente genau entgegengesetzt.

Da  $\mathcal{B}$  das lineare Erzeugnis der  $L_j$  ist, ist die Dimension des Raumes der von  $C_j$  aufgespannten Klassen in  $T_zV/\mathcal{B}$  offenbar mindestens  $p-1$  und das lineare Erzeugnis  $\langle L_1, \dots, L_p, C_1, \dots, C_p \rangle$  hat die Dimension  $2p-1$ .  $\square$

Damit erfüllt die Birkhoff-Distribution die Voraussetzungen des Satzes von Chow-Rashevskii [Mon02, Theorem 2.2], so daß also je zwei Punkte auf  $M$  durch eine zu der Birkhoff-Distribution horizontale Kurve verbunden werden können. Definiert man den Abstand zweier Punkte  $\mathfrak{a}$  und  $\mathfrak{b}$  auf  $M$  als das Infimum der Bogenlängen von zur Birkhoff-Distribution horizontalen Kurven, die  $\mathfrak{a}$  und  $\mathfrak{b}$  verbinden, so stimmt die dadurch auf  $M$  induzierte Topologie mit der gewöhnlichen Topologie auf  $M$  überein ([Mon02, Theorem 2.3]).

Auf der Menge aller Kurven auf  $M$  kann auf verschiedene Weise eine Mannigfaltigkeitsstruktur im Sinne von [Lan66] eingeführt werden. Sei  $\Omega^{(1)} := H^1(I, M)$  der Raum aller Kurven  $z : I \rightarrow M$  derart, daß für jede lokale Karte  $(U, \phi)$  von  $M$  mit einer offenen Teilmenge  $U \subset M$  und einer Abbildung  $\phi : U \rightarrow \mathbb{R}^{2p-1}$  und für jedes abgeschlossen Intervall  $J \subset z^{-1}(U)$  die Abbildung  $\phi \circ (z|_J)$  im Sobolev-Raum  $H^1(J, \mathbb{R}^{2p-1})$  liegt. Dann ist nach [PT01]  $\Omega^{(1)}$  eine glatte, unendlich dimensionale  $H^1(I, \mathbb{R}^{2p-1})$ -Mannigfaltigkeit. Betrachtet man statt der Sobolev  $H^1$ -Norm die Sobolev  $H^2$ -Norm, so gilt nach dem Einbettungssatz von Sobolev ([Eva98])  $H^2(I, \mathbb{R}^{2p-1}) \subset C^1(I, \mathbb{R}^{2p-1})$ . Genau aus diesem Grund verwenden Baryshnikov und Zharnitsky die feinere  $H^2$ -Norm statt der  $H^1$ -Norm zur Einführung einer Mannigfaltigkeitsstruktur auf dem Raum aller

Kurven auf  $M$ . Definiert man  $\Omega := H^2(I, M)$  als den Raum aller Kurven  $z : I \rightarrow M$  derart, daß für jede lokale Karte  $(U, \phi)$  von  $M$  mit einer offenen Teilmenge  $U \subset M$  und einer Abbildung  $\phi : U \rightarrow \mathbb{R}^{2p-1}$  und für jedes abgeschlossen Intervall  $J \subset z^{-1}(U)$  die Abbildung  $\phi \circ (z|_J)$  im Sobolev-Raum  $H^2(J, \mathbb{R}^{2p-1})$  liegt, so wird dies eine glatte, unendlich-dimensionale  $H^2(I, \mathbb{R}^{2p-1})$ -Mannigfaltigkeit.

Man betrachtet folgenden Teilmengen von Kurven in  $\Omega$ :

$$\begin{aligned}\Omega_{z_0}(\mathcal{B}) &:= \{z \in \Omega \mid z \text{ horizontal zu } \mathcal{B}, z(0) = z_0\} \\ \Omega_{z_0, v_0}(\mathcal{B}) &:= \{z \in \Omega \mid z \text{ horizontal zu } \mathcal{B}, z(0) = z_0, \dot{z}(0) = v_0, z_j(1) = z_{j+1}(0)\}\end{aligned}$$

Nach [Mon02, Theorem D.1] ist  $\Omega_{z_0}(\mathcal{B})$  eine Hilbert-Mannigfaltigkeit mit den durch

$$z \mapsto (u_1(t), \dots, u_p(t)),$$

definierten lokalen Karten, wobei  $u_1(t), \dots, u_p(t)$  die oben definierten Kontroll-Variablen sind.

Als nächstes betrachtet man die *Endpunktabbildung*  $\text{end} : \Omega_{z_0}(\mathcal{B}) \rightarrow M \times \mathcal{B}$ , die jeder zur Birkhoff-Distribution horizontalen Kurve  $z$  mit  $z(0) = z_0$  den Endpunkt  $\text{end}(z) = (z(1), \dot{z}(1))$  zuordnet. Die Endpunktabbildung  $\text{end}$  ist glatt bezüglich der Mannigfaltigkeitsstruktur von  $\Omega$  ([Mon02]).

Baryshnikov und Zharnitsky zeigen nun, daß  $\text{end}|_{\Omega_{z_0, v_0}(\mathcal{B})}$  eine Submersion ist, d.h. daß die zu  $\text{end}$  gehörende Tangentialabbildung für jede Kurve  $z \in \Omega_{z_0, v_0}(\mathcal{B})$  surjektiv ist.

Als Krümmung der Distribution  $\mathcal{B}$  bezeichnet man den durch

$$F(X, Y) := -[X, Y] \quad \text{mod } \mathcal{B}, \quad X, Y \in \mathcal{B}$$

definierten Vektorbündel-Morphismus  $F : \wedge^2 \mathcal{B} \rightarrow TM/\mathcal{B}$  (vgl. [Mon02, Chap. 4.1]). Für  $X \in \mathcal{B}$  induziert  $F(X, \cdot)$  einen Vektorbündel-Morphismus  $\mathcal{B} \rightarrow TM/\mathcal{B}$ . Nach Satz (1.11) ist dies lokal eine lineare Abbildung  $\Psi : \mathbb{R}^p \rightarrow \mathbb{R}^{p-1}$ . Sei  $X = \sum_{j=1}^p u_j L_j \in \mathcal{B}$  so ist  $\Psi(v_1, \dots, v_p) = -[\sum_{j=1}^p u_j L_j, \sum_{j=1}^p v_j L_j] \quad \text{mod } \mathcal{B}$ . Durch explizite Rechnung zeigen Baryshnikov und Zharnitsky in [BZ03], daß der Kern von  $\Psi$  eindimensional ist im Falle  $X = \dot{z}$ . Damit ist insbesondere  $\Psi$  surjektiv. Ferner Zeigen sie, daß sich der Beweis der Surjektivität der Tangentialabbildung von  $\text{end}$  lokal auf den Beweis der Surjektivität von  $\Psi$  reduziert.

### Konsequenzen für die numerische Konstruierbarkeit von $p$ -Lazutkin-Billards

Zur Konstruktion von  $p$ -Lazutkin-Billards können zwei Wege beschritten werden. Zum einen, kann versucht werden, die Elemente des Raumes  $\Omega_{z_0, v_0}(\mathcal{B})$  geeignet zu diskretisieren. Weil die  $p$ -Lazutkin-Billards in einer hinreichend kleinen Umgebung eine glatte Mannigfaltigkeit bilden, können dann mit einer Homotopie ausgehend von einem Kreis, der ein  $p$ -Lazutkin-Billard für

jedes  $p$  ist, weitere nicht-triviale  $p$ -Lazutkin-Billardards in dieser Mannigfaltigkeit konstruiert werden. Zum anderen kann direkt versuchen, die lokale Koordinatendarstellung  $(u_1(t), \dots, u_p(t))$  zu berechnen. Diese beiden Zugänge werden im nächsten Kapitel detailliert beschrieben.

## 2 Konstruktion von $p$ - Lazutkin-Billards

Nach dem Satz (1.9) kann man ein  $p$  - Lazutkin Billard stets so deformieren, daß die  $p$ -Lazutkin-Eigenschaft erhalten bleibt, wenn nur die Deformation „klein“ genug ist. Da der Kreis für jedes  $p$  ein  $p$ -Lazutkin-Billard ist, gibt es (nicht-triviale) Deformationen des Kreises, die  $p$ -Lazutkin-Billards sind. Es liegt nahe, den Deformationsbegriff anschaulich zu nehmen, und ausgehend von einer Diskretisierung des Kreises unter Verwendung von Homotopie-Methoden Billards  $\Omega$  zu konstruieren, deren Randkurve  $\partial\Omega$  und Billard-Abbildung  $P$  die folgende Bedingung erfüllen:

$$(BC) \quad \text{für jedes } z_1 \in \partial\Omega \text{ existieren } z_2, \dots, z_p \in \partial\Omega, \text{ so daß } z_{j+1} = Pz_j \text{ und } Pz_p = z_1.$$

In diesem Kapitel werden zwei konstruktive Methoden angegeben, die zur erstmaligen Herstellung von Billards mit  $p$ -periodischen Kaustiken verwendet wurden.

Der erste Algorithmus bestimmt  $N$  Mitglieder der  $p$ -periodischen Familie. Der Rand wird durch die Menge aller Reflektionspunkte diskretisiert. Dieser Ansatz führt auf ein Minimierungsproblem mit nichtlinearen Nebenbedingungen.

Der zweite Algorithmus basiert auf dem Konstruktionsansatz des Beweises von Satz (1.9), und versucht, eine horizontale Kurve zu der dort beschriebenen Birkhoff-Distribution zu finden, indem die Koordinatendarstellung gemäß Gleichung (1.5) berechnet wird. Dazu startet man mit einem  $p$ -Eck in der Ebene und bewegt die Ecken mit positiver Geschwindigkeit orthogonal zu den jeweiligen Winkelhalbierenden. Über diesen Ansatz gelangt man zu einem Zwei-Punkt-Randwertproblem.

### 2.1 Minimierungs-Strategie

Der Rand  $\partial\Omega$  wird durch  $p \cdot N$  viele Punkte

$$z_i^j = (x_i^j, y_i^j), \quad i = 1, \dots, p, \quad j = 1, \dots, N$$

diskretisiert. Für jedes feste  $j$  beschreibt das Tupel  $(z_1^j, \dots, z_p^j)$  ein Mitglied der Familie periodischer Orbits. Die Anordnung der  $z_i^j$  wird so gewählt, daß für jedes feste  $j$  die Punkte  $z_1^j, \dots, z_p^j$  der Billard-Bedingung (BC) genügen. Der Punkt  $z_i^{j+1}$  soll der Nachbar des Punktes  $z_i^j$  entlang des Randes gegen Uhrzeigerichtung sein. Um die Indexnumerierung korrekt zu halten, definiert man formal zusätzlich  $z_i^{N+1} = z_{i+1}^1$  für  $i = 1, \dots, p-1$  und  $z_p^{N+1} = z_1^1$ .

Die  $2 \cdot p \cdot N$  Unbekannten  $x_i^j$  und  $y_i^j$  müssen bestimmt werden.

Die Index Numerierung muß in einem kreisförmigen Sinne verstanden werden, d. h. die tiefer gestellten Indices sind immer modulo  $p$  zu verstehen.

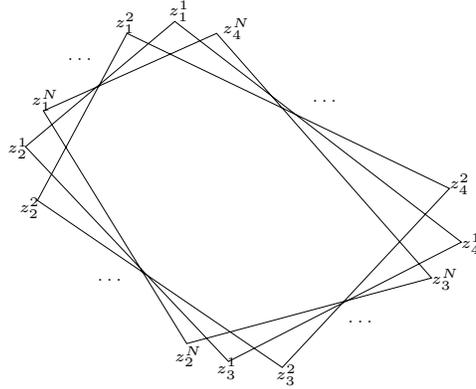


Abbildung 11: Diskretisierung des Billard Randes  $\partial\Omega$  für die Minimierungs-Strategie

Der Vektor

$$h_i^j := z_i^j - z_{i-1}^j, \quad i = 1, \dots, p \text{ und } j = 1, \dots, N$$

gibt die Richtung des einfallenden Strahls im Punkt  $z_i^j$  an, der zum  $j$ -ten periodischen Orbit gehört. Der Vektor  $h_{i+1}^j$  ist dann die Richtung des reflektierten Strahls im Punkt  $z_i^j$ . Da die Punkte  $z_1^j, \dots, z_p^j$  einen periodischen Orbit unter der Billard-Abbildung  $P$  bilden, ist die Richtung der ins Innere zeigenden Normalen von  $\partial\Omega$  im Punkt  $z_i^j$  gegeben durch

$$n_i^j = \frac{h_{i+1}^j}{\|h_{i+1}^j\|} - \frac{h_i^j}{\|h_i^j\|}. \quad (2.1)$$

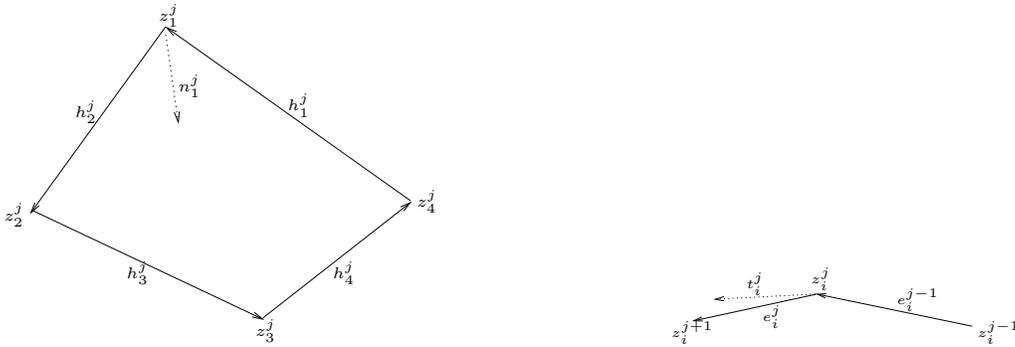


Abbildung 12: Berechnung von Normale und Tangente im Punkt  $z_i^j$ .

Der Vektor

$$e_i^j := z_i^{j+1} - z_i^j, \quad i = 1, \dots, p \text{ und } j = 1, \dots, N$$

beschreibt die Richtung vom Punkt  $z_i^j$  zu seinem nächsten Nachbarn entlang des Randes. Die

Tangente an  $\partial\Omega$  im Punkt  $z_i^j$  ist somit gegeben durch

$$t_i^j = \frac{e_i^j}{\|e_i^j\|} + \frac{e_i^{j-1}}{\|e_i^{j-1}\|}. \quad (2.2)$$

Da die Normalen und Tangenten orthogonal sind, hat man

$$(n_i^j, t_i^j) = 0 \quad (2.3)$$

für die Unbekannten  $z_i^j$ .

Natürlich ist das Gleichungssystem (2.3) nicht eindeutig lösbar, weil ein Kreis (diskretisiert in obiger Art) die Gleichung erfüllen wird. Um eine bestimmte Lösung auszuwählen, wählt man ein Tupel  $(Z_1, \dots, Z_p)$  und minimiert den Euklidischen Abstand zwischen diesem Punkt und einem Mitglied  $(z_1^1, \dots, z_p^1)$  der Familie periodischer Orbits. Durch eine geschickte Wahl des Tupels  $(Z_1, \dots, Z_p)$  vermeidet man die triviale Kreislösung. (Vorsicht: Jedoch liegen für  $p = 3$  stets 3 Punkte auf einem Kreis, so daß die Methode für diesen Fall nicht anwendbar ist.)

Die Konstruktion läßt sich also als folgendes Minimierungsproblem mit nichtlinearen Nebenbedingungen formulieren:

**Problem A:**

Minimiere  $\|(Z_1, \dots, Z_p) - (z_1^1, \dots, z_p^1)\|$  so daß

$$\forall i = 1, \dots, p, j = 1, \dots, N : (n_i^j, t_i^j) = 0.$$

Mehrere Algorithmen zur numerische Lösung von Problem A sind bekannt. Die meisten von diesen sind in fertigen Programmpaketen wie zum Beispiel AMPL frei erhältlich. Jeder dieser Algorithmen benötigt eine Startnäherung für die gesuchte Lösung. Entscheidend für die Effizienz und die Konvergenz dieser Verfahren ist die geschickte Wahl einer Startlösung. Im folgenden wird eine Strategie beschrieben, wie gute Startnäherungen für Algorithmen zur Lösung von Problem A gewonnen werden können.

### 2.1 Algorithmus (Minimierungs-Strategie).

- 1 Wähle als Anfangsnäherung Punkte  $z_i^j$ ,  $i = 1, \dots, p$ ,  $j = 1, \dots, N$  die entlang eines Kreises gleichmäßig verteilt sind.
- 2 Wähle das Tupel  $(Z_1, \dots, Z_p)$  so, daß die Punkte  $Z_j$  in der Nähe der Kreislinie liegen.
- 3 Löse Problem A mit diesen Daten.
- 4 Verwende diese Lösung als neue Anfangsnäherung und wähle ein neues Tupel  $(Z_1, \dots, Z_p)$  so, daß die Punkte  $Z_i$  nicht auf einem Kreis, aber in der Nähe des vorher gewählten Tupels liegen. Gehe dann zu Schritt 3. In dieser Weise wird eine Folge von  $p$ -Lazutkin-Billardards erzeugt. Stoppe, wenn ein gewünschtes  $p$ -Lazutkin-Billard erreicht ist.

Schritt 4 ist der wichtigste Schritt der Strategie. Es gibt keine explizite Regel, wie das Tupel gewählt werden muß. In zahlreichen numerischen Experimenten hat sich erwiesen, daß man als Faustregel beachten sollte, daß das Tupel einem der periodischen Orbits ähnlich sein soll. Man kann auf diese Weise die Konstruktion beeinflussen und das  $p$ -Lazutkin-Billard nach seinen Vorstellungen formen. Weitere Details zur Implementierung und Lösungsstrategie findet man in Abschnitt 2.3.

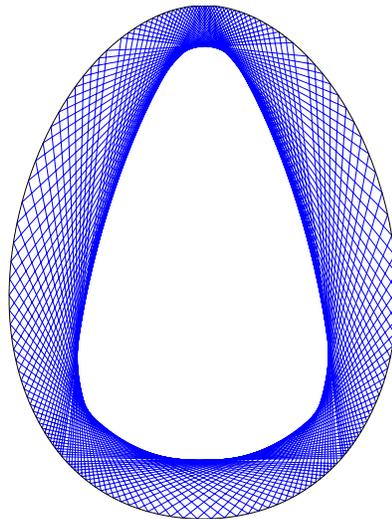


Abbildung 13: Ein 4 - Lazutkin-Billard konstruiert mit der Minimierungs-Strategie. Der Rand wurde durch 40 Familienmitglieder diskretisiert. Mehr Details zur Konstruktion findet man in Abschnitt 2.3.

## 2.2 Optimal-Control-Strategie

Der Rand  $\partial\Omega$  wird bei diesem Ansatz durch eine Vereinigung von  $p$  Bögen  $z_j$ ,  $j = 1, \dots, p$  dargestellt. Jeder dieser Bögen wird durch die Bewegung der Ecken eines anfangs gewählten  $p$ -Ecks definiert. Für die Bögen wird die Parametrisierung  $z_j(t) : [0, 1] \rightarrow \mathbb{C}$  gewählt. Um eine geschlossene Umrandung zu erhalten, müssen die folgenden Randbedingungen für  $j = 1, \dots, p$  erfüllt werden:

$$z_j(1) = z_{j+1}(0), \quad (2.4)$$

falls man die Indexnumerierung des vorherigen Abschnittes beibehält.

Die Bögen  $z_j(t)$  müssen so bestimmt werden, daß für jedes  $t \in [0, 1]$  die Punkte  $z_1(t), \dots, z_p(t)$  der Billardbedingung (BC) genügen.

Die Richtung der Tangente im Punkt  $z_j(t)$  ist gegeben durch

$$\frac{z_{j+1}(t) - z_j(t)}{\|z_{j+1}(t) - z_j(t)\|} + \frac{z_j(t) - z_{j-1}(t)}{\|z_j(t) - z_{j-1}(t)\|}.$$

Daher schreibt sich

$$\dot{z}_j(t) = \left( \frac{z_{j+1}(t) - z_j(t)}{\|z_{j+1}(t) - z_j(t)\|} + \frac{z_j(t) - z_{j-1}(t)}{\|z_j(t) - z_{j-1}(t)\|} \right) \cdot u_j(t) \quad (2.5)$$

mit positivem reellen  $u_j(t)$ . Sei nun das Ausgangs  $p$ -Eck beschrieben durch das Tupel  $(z_1^0, \dots, z_p^0)$ . Dieses liefert die Anfangsbedingungen

$$z_j(0) = z_j^0 \quad (2.6)$$

für die obigen Differentialgleichungen. Ferner muß  $z_j(t)$  die obigen Geschlossenheitsbedingung (2.4) erfüllen. Die skalaren Funktionen  $u_j(t)$  werden als Kontroll-Variablen behandelt, die die Geschwindigkeit der entstehenden Bögen steuern. Man fordert an die Lösungstrajektorie und die Kontroll-Variablen die Minimierungsbedingung:

$$\min - \int_0^1 \sum_{j=1}^p \log u_j(t) dt \quad (2.7)$$

Diese Wahl verhindert zum einen, daß die Kontrollvariable zu Null wird, und vereinfacht andererseits die nachfolgenden Berechnungen. Die Lösungen des Systems (2.5) mit Anfangswerten (2.6) sind Elemente aus dem Raum  $\Omega_{z_0, v_0}(\mathcal{B})$  von Kapitel 1. Die Kontroll-Variablen entsprechen den jeweiligen Koordinatenfunktionen. Durch die zusätzliche Bedingung (2.7) wird also eine Kurve aus der Mannigfaltigkeit  $\Omega_{z_0, v_0}(\mathcal{B})$  ausgewählt.

Bis jetzt wurde die Konstruktion als Optimal-Control-Problem mit steuernder Differentialgleichung (2.5), Rand- und Anfangsbedingungen (2.6) und (2.4) und Zielfunktional (2.7) formuliert.

Im folgenden wird dieses Problem in ein Randwertproblem transformiert, das viel einfacher zu lösen ist.

Zunächst wird der komplexe Bogen  $z_j(t)$  in Real und Imaginärteil aufgespalten,

$$z_j(t) = x_j(t) + i \cdot y_j(t)$$

mit reell-wertigen Kurven  $x_j(t)$  und  $y_j(t)$ . Nachdem auch System (2.5) in Real- und Imaginärteil aufgespalten wurde, erhält man ein System der Gestalt

$$\begin{aligned}\dot{x}_j &= f_j(x_1, \dots, x_p, y_1, \dots, y_p, u_1, \dots, u_p) \\ \dot{y}_j &= g_j(x_1, \dots, x_p, y_1, \dots, y_p, u_1, \dots, u_p).\end{aligned}$$

Die Randbedingungen übersetzen sich in dieser Notation in offensichtlicher Weise. Die expliziten Ausdrücke für die Funktionen  $f$  und  $g$  sind in Abschnitt 2.3 zu finden.

Um die Notation weiter zu vereinfachen, setzt man  $x = (x_1, \dots, x_p)^T$ ,  $y = (y_1, \dots, y_p)^T$  und  $u = (u_1, \dots, u_p)^T$ ,  $\xi = (x, y)$  und definiert

$$\begin{aligned}F &:= \begin{pmatrix} f \\ g \end{pmatrix} : \mathbb{R}^{2p} \times \mathbb{R}^p \rightarrow \mathbb{R}^{2p}, \\ (\xi, u) &\mapsto \begin{pmatrix} f(x, y, u) \\ g(x, y, u) \end{pmatrix}.\end{aligned}$$

Somit kann Gleichung (2.5) zusammen mit (2.7) als  $2p$ -dimensionales, reelles, nichtlineares, Optimal-Control-Problem formuliert werden:

Minimiere das Funktional

$$-\int_0^1 \sum_{j=1}^p \log u_j(t) dt,$$

so daß  $\xi(t)$  und  $u(t)$  das autonome Differentialgleichungssystem

$$\dot{\xi} = F(\xi, u) \tag{2.8}$$

lösen und den obigen Randbedingungen genügen.

Sei

$$H(\xi, u, \dot{\xi}, \dot{u}) := -\sum_{j=1}^p \log u_j + \lambda^T \cdot (F(\xi, u) - \dot{\xi})$$

das zugehörige Hamiltonsche Funktional mit Lagrange-Multiplikatoren

$$\lambda(t) = (\lambda_x(t), \lambda_y(t))^T \in \mathbb{R}^p \times \mathbb{R}^p.$$

Dann kann das Variationsproblem mit Nebenbedingungen als folgendes Variationsproblem ohne Nebenbedingungen behandelt werden

$$\min \int_0^1 H(\xi, u, \dot{\xi}, \dot{u}) dt.$$

Die Euler-Lagrange-Gleichung hierzu liefert dann

$$D_{(\xi, u)} H - \frac{d}{dt} D_{(\dot{\xi}, \dot{u})} H = 0,$$

also

$$\begin{aligned} D_{\xi} H - \frac{d}{dt} D_{\dot{\xi}} H &= 0 \\ D_u H - \frac{d}{dt} D_{\dot{u}} H &= 0. \end{aligned} \tag{2.9}$$

Aus der ersten Gleichung von (2.9) folgt

$$-\dot{\lambda}^T = \lambda^T D_{\xi} F.$$

Teilt man in  $x, y$ -Koordinaten auf, so gelangt man zu

$$\dot{\lambda}_x = -(\lambda_x D_x f + \lambda_y D_x g) \tag{2.10}$$

$$\dot{\lambda}_y = -(\lambda_x D_y f + \lambda_y D_y g). \tag{2.11}$$

Beide Gleichungen sind immer noch von  $u$  abhängig. Die zweite Gleichung von (2.9) liefert

$$-\frac{1}{u_j} = -\lambda^T D_{u_j} F,$$

was in  $x, y$ -Koordinaten auf

$$\frac{1}{u_j} = \sum_{k=1}^p (\lambda_x)_k \frac{\partial f_k}{\partial u_j} + \sum_{k=1}^p (\lambda_y)_k \frac{\partial g_k}{\partial u_j} \tag{2.12}$$

führt.

Wie man leicht feststellen kann<sup>1</sup>, sind die partiellen Ableitungen  $\frac{\partial f_k}{\partial u_j}$  und  $\frac{\partial g_k}{\partial u_j}$  entweder gleich 0 oder hängen nur von den  $x_j$ 's und  $y_j$ 's ab. Somit lassen sich die  $u_j$  in Termen ausdrücken, die ausschließlich von  $x_j, y_j$  und  $\lambda_j$  abhängen. Daher können sie aus den Gleichungen (2.8), (2.10) und (2.11) eliminiert werden. An diesem Punkt spielt die spezielle Wahl des Funktionals (2.7) eine fundamentale Rolle, da eine andere Wahl verursacht, daß die  $u_j$ 's nicht mehr eliminiert

---

<sup>1</sup>Details im Abschnitt 2.3

werden können. Die Konstruktion ist somit auf folgendes  $4p$ -dimensionale nichtlineare Randwertproblem reduziert:

**Problem B:**

$$\begin{aligned}\dot{\xi} &= F(\xi, u) \\ \dot{\lambda}^T &= -\lambda^T D_{\xi} F\end{aligned}$$

mit Randbedingungen

$$\xi(0) = \xi^0,$$

und

$$\xi_1(1) = \xi_2^0, \dots, \xi_p(1) = \xi_1^0, \quad \xi_{p+1}(1) = \xi_{p+2}^0, \dots, \xi_{2p}(1) = \xi_{p+1}^0.$$

Für die numerische Lösung von Problem B existieren wiederum mehrere Algorithmen, wie zum Beispiel Schieß- oder Kollokationsverfahren. Die Nichtlinearität und die hohe Dimension des Problem erfordern eine geeignete Kontinuationsstrategie (Homotopie), um sich an die richtige Lösung heranzutasten.

Die folgende Strategie liefert sehr gute Ergebnisse, wenn als Lösungsalgorithmus für Problem B ein Einzelschießverfahren benutzt wird:

## 2.2 Algorithmus (Optimal-Control-Strategie).

- 1 Wähle  $p$  entlang des Einheitskreises äquidistant verteilte Punkte, um  $\xi^0$  zu erhalten und wähle eine Anfangsnäherung<sup>2</sup> für  $\lambda(0)$ .
- 2 Löse Problem B mit einem Einzelschießverfahren.
- 3 Ändere die  $\xi^0$  ein wenig und benutze die berechneten  $\lambda(0)$  als neue Anfangsnäherung.
- 4 Gehe zu Schritt 2 und halte, falls das gewünschte  $p$ -Lazutkin-Billard erreicht ist.

## 2.3 Numerische Implementierung

Dieser Abschnitt enthält einige Hinweise zur praktischen Implementierung der oben beschriebenen Ansätze. Die Hinweise beziehen sich zwar auf ein 3 - Lazutkin-Billard und ein 4 - Lazutkin-Billard, sind aber leicht verallgemeinerbar.

Die beiden speziell konstruierten Billards werden im folgenden noch öfters als Referenzbeispiele verwendet. Das 3-Lazutkin-Billard wird auf Grund seiner Form im Folgenden oft als *Kartoffel-Billard* bezeichnet, das 4-Lazutkin-Billard als *Ei-Billard*.

---

<sup>2</sup>Details zur Wahl von  $\lambda(0)$  findet man im nächsten Abschnitt

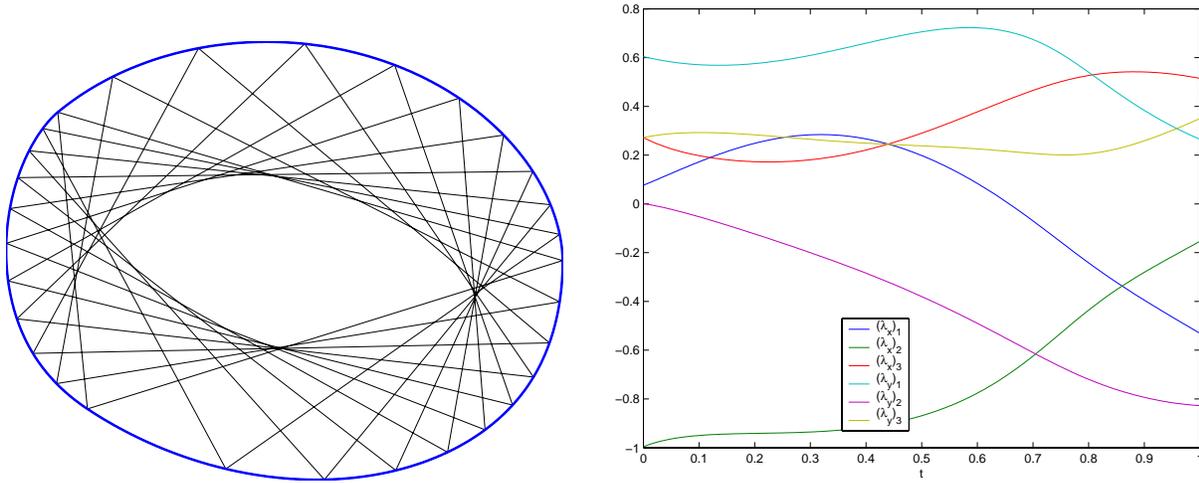


Abbildung 14: Links: Billard mit einer Familie von 3-periodischen Orbits. Einige Familienmitglieder sind geplottet. Die 3-periodische Kaustik ist im Inneren des Billards sichtbar als Evolute der periodischen Orbits. Rechts: Der Verlauf der Lagrange-Multiplikatoren zu dem Billard links.

### Minimierungs-Strategie

Das Minimierungsproblem wurde mit der Software AMPL von Bell-Labs zusammen mit dem Löser SNOPT gelöst. Die im Abschnitt 2.1 vorgestellte Strategie beruht auf vielen numerischen Experimenten. Die von mir benutzte Implementierung benötigt 5 Punkte zum Darstellen eines periodischen Orbits  $(z_1, \dots, z_4)$ . Der Punkt  $z_5$  stimmt mit dem Punkt  $z_1$  überein, um Fallunterscheidungen durch die periodische Indizierung zu vermeiden. Für einen periodischen Orbit werden somit 10 Unbekannte benötigt. Der Rand wird durch 30 periodische Orbits diskretisiert, es werden also insgesamt 300 Unbekannte benötigt.

### Optimal-Control-Strategie

Exemplarisch wird zunächst die Herleitung für die Funktionen  $f_j$  und  $g_j$  für eine Familie von 4-periodischen Orbits vorgeführt, die oben weggelassen worden ist. Mit der periodischen Indizierung  $z_5(t) = z_1(t)$  und  $z_0(t) = z_4(t)$ , gilt für  $j = 1, \dots, 4$

$$\dot{z}_j(t) = \left( \frac{z_{j+1}(t) - z_j(t)}{\|z_{j+1}(t) - z_j(t)\|} + \frac{z_j(t) - z_{j-1}(t)}{\|z_j(t) - z_{j-1}(t)\|} \right) \cdot u_j(t)$$

Aufspalten in Real- und Imaginärteil gibt dann

$$\dot{z}_j = \left( \frac{x_{j+1} - x_j}{l_j} + \frac{x_j - x_{j-1}}{l_{j-1}} \right) u_j + i \cdot \left( \frac{y_{j+1} - y_j}{l_j} + \frac{y_j - y_{j-1}}{l_{j-1}} \right) u_j$$

mit

$$l_j := \sqrt{(x_{j+1} - x_j)^2 + (y_{j+1} - y_j)^2}.$$

Also, folgt für die Funktionen  $f_j$  und  $g_j$  aus Abschnitt 2.2

$$\begin{aligned} f_j(x, y, u) &= \left( \frac{x_{j+1} - x_j}{l_j} + \frac{x_j - x_{j-1}}{l_{j-1}} \right) u_j \\ g_j(x, y, u) &= \left( \frac{y_{j+1} - y_j}{l_j} + \frac{y_j - y_{j-1}}{l_{j-1}} \right) u_j \end{aligned}$$

Für die benötigten partiellen Ableitungen von  $f_j$  findet man

$$\begin{aligned} \frac{\partial f_j}{\partial x_j} &= \left( -\frac{1}{l_j} + \frac{1}{l_{j-1}} + \frac{(x_{j+1} - x_j)^2}{l_j^3} - \frac{(x_j - x_{j-1})^2}{l_{j-1}^3} \right) \cdot u_j \\ \frac{\partial f_j}{\partial x_{j+1}} &= \left( \frac{1}{l_j} - \frac{(x_{j+1} - x_j)^2}{l_j^3} \right) \cdot u_j \\ \frac{\partial f_j}{\partial x_{j-1}} &= \left( \frac{1}{l_{j-1}} + \frac{(x_j - x_{j-1})^2}{l_{j-1}^3} \right) \cdot u_j \\ \frac{\partial f_j}{\partial y_j} &= \left( \frac{(x_{j+1} - x_j)(y_{j+1} - y_j)}{l_j^3} - \frac{(x_j - x_{j-1})(y_j - y_{j-1})}{l_{j-1}^3} \right) \cdot u_j \\ \frac{\partial f_j}{\partial y_{j+1}} &= \left( -\frac{(x_{j+1} - x_j)(y_{j+1} - y_j)}{l_j^3} \right) \cdot u_j \\ \frac{\partial f_j}{\partial y_{j-1}} &= \left( \frac{(x_j - x_{j-1})(y_j - y_{j-1})}{l_{j-1}^3} \right) \cdot u_j \end{aligned}$$

und

$$\frac{\partial f_j}{\partial u_k} = \begin{cases} \frac{x_{j+1} - x_j}{l_j} + \frac{x_j - x_{j-1}}{l_{j-1}} & : j = k \\ 0 & : j \neq k \end{cases}.$$

Berücksichtigt man  $g_j(x, y, u) = f_j(y, x, u)$  so findet man mit den obigen Gleichungen

$$\begin{aligned}\frac{\partial g_j}{\partial x_j} &= \left( \frac{(x_{j+1} - x_j)(y_{j+1} - y_j)}{l_j^3} - \frac{(x_j - x_{j-1})(y_j - y_{j-1})}{l_{j-1}^3} \right) \cdot u_j \\ \frac{\partial g_j}{\partial x_{j+1}} &= \left( -\frac{(x_{j+1} - x_j)(y_{j+1} - y_j)}{l_j^3} \right) \cdot u_j \\ \frac{\partial g_j}{\partial x_{j-1}} &= \left( \frac{(x_j - x_{j-1})(y_j - y_{j-1})}{l_{j-1}^3} \right) \cdot u_j \\ \frac{\partial g_j}{\partial y_j} &= \left( -\frac{1}{l_j} + \frac{1}{l_{j-1}} + \frac{(y_{j+1} - y_j)^2}{l_j^3} - \frac{(y_j - y_{j-1})^2}{l_{j-1}^3} \right) \cdot u_j \\ \frac{\partial g_j}{\partial y_{j+1}} &= \left( \frac{1}{l_j} - \frac{(y_{j+1} - y_j)^2}{l_j^3} \right) \cdot u_j \\ \frac{\partial g_j}{\partial y_{j-1}} &= \left( \frac{1}{l_{j-1}} + \frac{(y_j - y_{j-1})^2}{l_{j-1}^3} \right) \cdot u_j\end{aligned}$$

und

$$\frac{\partial g_j}{\partial u_k} = \begin{cases} \frac{y_{j+1} - y_j}{l_j} + \frac{y_j - y_{j-1}}{l_{j-1}} & : j = k \\ 0 & : j \neq k \end{cases}.$$

Wie im Abschnitt 2.2 beschrieben lassen sich die  $u_j$ 's in Termen von  $x_j$ 's und  $y_j$ 's ausdrücken. Nun hat man alle Formeln zusammen, um das in Abschnitt 2.2 beschriebene Randwertproblem (B) zu behandeln.

Wählt man die 4 Punkte  $(1, 0), (-1, 0), (0, 1), (0, -1)$ , also

$$\xi_0 = (1, 0, -1, 0, 0, 1, 0, -1)^T$$

zusammen mit

$$\lambda_0 = (0.1075, -0.6960, -0.1075, 0.6960, 0.7345, -0.0685, -0.7345, 0.0685)^T$$

und integriert das System (2.8) von 0 bis 1, so erhält man eine gute Näherungslösung für den Einheitskreis, von dem man dann ausgehend eine Kontinuation gemäß Strategie 2.2 starten kann.

Für eine 3-periodische Billardkurve können die Punkte  $(1, 0), (0, 0), (1, -1)$ , also

$$\xi_0 = (1, 0, 0, 0, 1, -1)^T$$

zusammen mit

$$\lambda_0 = (-0.4419, -0.5134, 1.0597, 0.2810, -0.3681, -1.4285)^T.$$

gewählt werden.

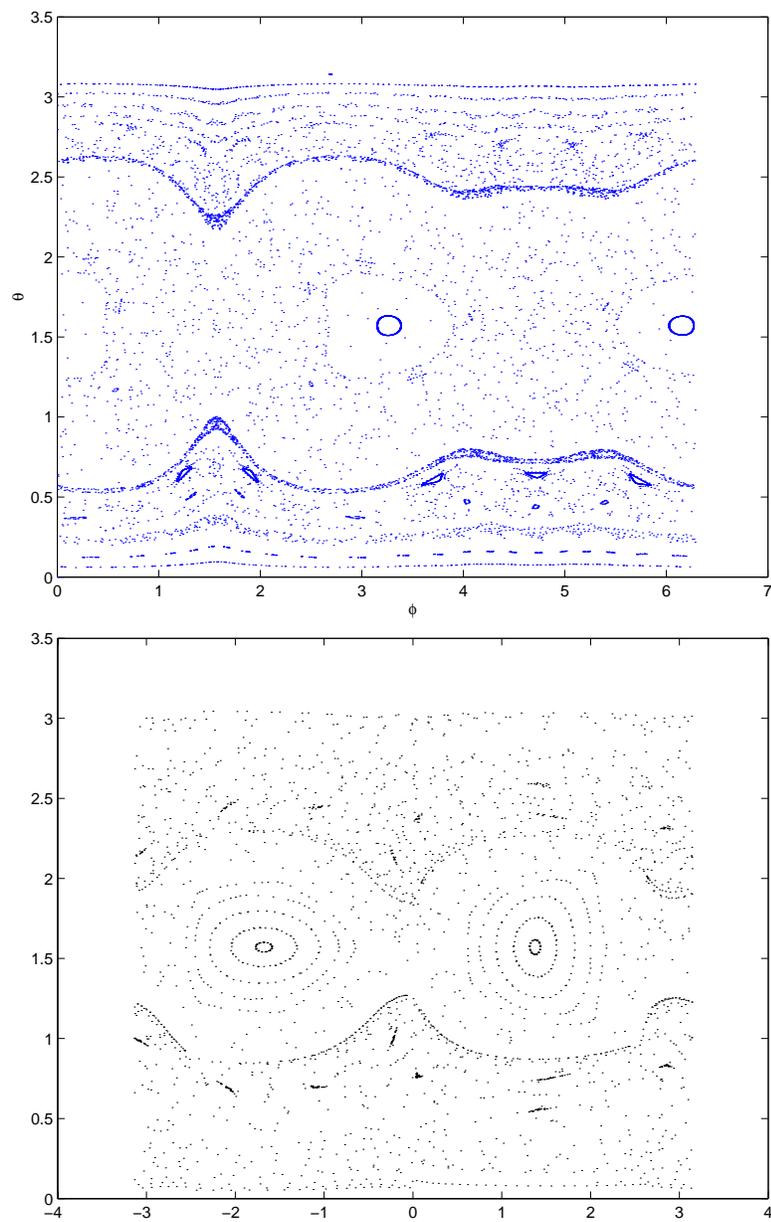


Abbildung 15: Oben: Phasenportrait des Ei-Billard. Links: Phasenportrait des Kartoffel-Billard. In beiden Abbildungen ist der KAM-Streifen deutlich zu erkennen.

### 3 Quantenmechanik und quasi-klassische Approximation

Über die quasi-klassische Approximation wird in der Quantenmechanik ein Zusammenhang zwischen den Lösungen der Helmholtz-Gleichung

$$\Delta u + k^2 u = 0$$

mit  $k^2 \gg 1$  in einem Billard  $\Omega \subset \mathbb{R}^2$  mit der Randbedingung

$$u = 0 \quad \text{auf } \partial\Omega$$

und den Orbits des zu  $\Omega$  gehörenden Billard-Systems hergestellt. Einige Ansätze hierzu werden im folgenden Kapitel kurz referiert.

Im Falle eines  $p$ -Lazutkin-Billard  $\Omega$  ergeben sich Typen von Lösungen der Helmholtz-Gleichung, die zum ersten Mal 2004 in [BHZP04] beschrieben worden sind.

Lösungen der Helmholtz-Gleichung existieren nur für ein diskretes Spektrum  $\{k_n^2\}$ . Über das Spektrum des Laplace-Operators für große Eigenwerte  $k$  ist im allgemeinen wenig bekannt. Die Weyl-Formel Satz (3.1) liefert eine asymptotische Aussage für die Anzahl von Eigenwerten unterhalb einem vorgegebenem  $k$ . Da analytische Darstellungen der Lösungen der Helmholtz-Gleichung mit großen  $k$  im allgemeinen nicht möglich sind, muß man sich auf approximative Beschreibungen bzw. auf numerische Ergebnisse beschränken.

In der Physik wird die sogenannte Husimi-Projektion dazu verwendet, Lösungen der Helmholtz-Gleichung im Phasenraum des zugehörigen Billard-Systems darzustellen. Die in Kapitel 4 entwickelten neuen numerischen Verfahren zur Lösung der Helmholtz-Gleichung gestatten nun eine effiziente Berechnung der Husimi-Projektion.

#### 3.1 Schrödinger-Gleichung und quasi-klassischer Übergang

Die folgenden Überlegungen sind wohl bekannt, wurden aber zur besseren Orientierung in der physikalischen Begriffswelt aufgenommen. Die Bewegung eines quantenmechanischen Teilchens der Masse  $m$  in einem Potential  $U(x, t)$  ( $x \in \mathbb{R}^3$ ) wird in der Quantenmechanik durch eine Wellenfunktion  $\psi(x, t)$  beschrieben, die die *Schrödinger-Gleichung*,

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \Delta \psi + U(x, t) \tag{3.1}$$

erfüllt. Falls das Potential  $U$  nicht von der Zeit abhängt, führt der Ansatz  $\psi(x, t) = u(x)f(t)$  auf die Gleichung

$$\frac{i\hbar}{f} \frac{df}{dt} = \frac{1}{u} \left( -\frac{\hbar^2}{2m} \Delta u + U(x)u \right),$$

deren linke Seite nur noch von  $t$  und deren rechte Seite nur noch von  $x$  abhängt, so daß beide Seiten gleich einer Separationskonstanten  $E$  sein müssen. Damit gilt

$$\psi(x, t) = u(x)e^{-iEt/\hbar},$$

wobei  $u$  der zeitunabhängige Schrödinger-Gleichung

$$\frac{\hbar^2}{2m}\Delta u + (E - U(x))u = 0 \quad (3.2)$$

genügt.

Ein einfacher 2-dimensionaler Spezialfall ist die Bewegung eines Teilchens in einem „Potentialtopf“,

$$U(x) = \begin{cases} 0 & : x \in \Omega \\ \infty & : x \notin \Omega \end{cases} \quad (3.3)$$

In diesem Fall muß  $u(x) = 0$  für  $x \in \partial\Omega$  sein ([Sch68]), so daß die Bewegung des Teilchens durch die *Helmholtz-Gleichung*

$$\begin{aligned} \Delta u + \frac{2m}{\hbar^2}Eu &= 0 & x \in \Omega \\ u &= 0 & x \in \partial\Omega \end{aligned}$$

beschrieben wird. In der weiteren Behandlung werden die Maß-Einheiten so gewählt, daß der Faktor  $\frac{2m}{\hbar^2}$  zu 1 wird. Aus der Theorie der partiellen Differentialgleichungen ist bekannt, daß die Helmholtz-Gleichung für die diskrete Folge  $(E_n)$  nicht-triviale Lösungen besitzt, die gerade die Eigenwerte des Laplace-Operators auf dem Gebiet  $\Omega$  mit Dirichlet Randbedingungen sind ([Eva98]). Nur für spezielle Gebiete  $\Omega$  existiert eine explizite analytische Darstellung der Eigenfunktionen bzw. eine explizite Formel zur Berechnung der Eigenwerte ([Eva98]), so daß man im allgemeinen Fall auf numerische Näherungen zurückgreifen muß.

Physikalisch interessant sind insbesondere Eigenfunktionen zu großen Eigenwerten, also  $E_n \gg \text{vol}(\Omega)$ , vergleiche hierzu ([Gut90]).

Obwohl es nicht möglich ist, für beliebige Gebiete  $\Omega$  explizite Formeln für die Eigenwerte  $E_n$  und Darstellungen für die Eigenfunktionen anzugeben, können Aussagen über das asymptotische Verhalten der Eigenwerte  $E_n$  gemacht werden. Es gilt

**3.1 Satz (Weyl-Formel).** *Sei*

$$N(E) := \#\{n \mid 0 \leq E_n \leq E\}$$

und  $L$  die Länge des Randes  $\partial\Omega$ . Dann gilt

$$N(E) \sim \frac{\text{vol}(\Omega)}{4\pi}E - \frac{L}{4\pi}\sqrt{E}, \quad \text{für } E \rightarrow \infty.$$

Einen sehr eleganten Beweis der Weyl-Formel findet man in [Kac66]. Die Weyl-Formel kann als Ausgangspunkt einer ganzen Reihe von sogenannten *Spurformeln* verstanden werden ([Hej76]).

Wegen  $E = k^2$  erhält man durch Differenzieren der Weyl-Formel nach  $k$  für den mittleren Eigenwertabstand bei  $k$

$$\Delta_k = \left( \frac{dN}{dE} \frac{dE}{dk} \right)^{-1} = \left( \frac{\text{vol}(\Omega)k}{2\pi} - \frac{L}{4\pi} \right)^{-1} = \frac{4\pi}{2\text{vol}(\Omega)k - L}. \quad (3.4)$$

Eine physikalische Interpretation der Lösungen der Helmholtz-Gleichung mit großem  $E_n$  erhält man wie folgt. Mit dem Ansatz

$$\psi(x, t) = A(x, t)e^{\frac{i}{\hbar}S(x, t)},$$

worin  $A(x, t)$  und  $S(x, t)$  reell-wertige Funktionen sind, liefert die Schrödinger-Gleichung die Beziehung

$$i\hbar \left( \frac{\partial A}{\partial t} + \frac{i}{\hbar} A \frac{\partial S}{\partial t} \right) = -\frac{\hbar^2}{2m} \left( \Delta A + \frac{2i}{\hbar} \nabla A \nabla S + \frac{i}{\hbar} A \Delta S - \frac{1}{\hbar^2} A \|\nabla S\|^2 \right) + A \cdot U. \quad (3.5)$$

Die Konstante  $\hbar$  hat die Dimension einer Wirkung und hängt natürlich von den gewählten Maßeinheiten ab, im SI-System gilt zum Beispiel

$$\hbar = 1.0545887 \cdot 10^{-34} J_s.$$

Die relative Größe dieser Konstanten bezüglich der anderer Größen eines physikalischen Systems bestimmt den Grad der quantenmechanischer Natur. Ein System, dessen charakteristische Wirkung in der Größenordnung von  $\hbar$  liegt, muß mit den Methoden der Quantenphysik beschrieben werden, umgekehrt kann ein System mit charakteristischer Wirkung  $\gg \hbar$  durch die klassische Physik beschrieben werden. Die klassische Mechanik der Massepunkte kann also in dem Sinne  $\hbar \rightarrow 0$  als Grenzwert der Quantenmechanik aufgefaßt werden, die sich in der Beschreibung (3.1) niederschlägt. Dieses Prinzip heißt *Korrespondenzprinzip* der Quantenmechanik (vgl. [Sch68]).

Faßt man nun in Gleichung (3.5)  $\hbar$  als kleine Größe auf und sammelt dann alle Terme ohne explizite  $\hbar$ -Abhängigkeit, so erhält man die Gleichung

$$\frac{\partial S}{\partial t} + \frac{1}{2m} \|\nabla S\|^2 + U(x) = 0. \quad (3.6)$$

Dies ist die aus der klassischen Mechanik bekannte *Hamilton-Jacobi-Gleichung*, für die Hamiltonsche Wirkungsfunktion  $S(x, t)$ . Man erhält somit eine Interpretation der Phase der Wellenfunktion  $\psi$ :  $\hbar$  mal der Phase entspricht näherungsweise der Hamiltonschen Hauptfunktion, sofern  $\hbar$  als kleine Größe angesehen werden kann.

Die Hamilton-Jacobi-Gleichung mit dem Potential (3.3) beschreibt die freie Bewegung eines Masseteilchens im Gebiet  $\Omega$ , das am Rand  $\partial\Omega$  reflektiert wird. Die klassische Bewegung ist demnach gerade ein Billard-System.

Der quasi-klassische Übergang  $\hbar \rightarrow 0$  in Gleichung (3.2) entspricht dem Lösen der Helmholtz-Gleichung mit  $E \gg 1$ . Nach dem vorangehenden Gedankengang liegt es nahe, einen Zusammenhang zwischen Lösungen mit  $E \gg 1$  und den klassischen Trajektorien des Billard-Systems herzustellen.

Die Grundannahme zur asymptotischen Lösung der Helmholtz-Gleichung besteht darin, daß Lösungen für großes  $k$  die Form

$$u(x) = \sum_{j=1}^N e^{ikS_j(x)} \left( A_j(x) + O\left(\frac{1}{k}\right) \right) \quad (3.7)$$

haben. Die Funktionen  $S_j$  und  $A_j$  sind vom Ort abhängig,  $N$  ist eine natürliche Zahl und jeder Term in der Summe stellt eine Welle dar, wobei  $A_j$  die Amplitude und  $S_j$  die Phase ist. Setzt man (3.7) in die Helmholtz-Gleichung ein, so erhält man durch Nullsetzen der Terme zu  $k$  und  $k^2$  für jedes  $j = 1, \dots, N$  die beiden Gleichungen

$$\begin{aligned} \|\nabla S_j\|^2 &= 1 \\ 2\nabla S_j \cdot \nabla A_j + A_j \Delta S_j &= 0. \end{aligned}$$

Die erste Gleichung ist die *Eikonal*-Gleichung der geometrischen Optik. Die zweite Gleichung ist als Transportgleichung bekannt. Die Lösungen der Eikonal-Gleichung sind durch gerade Linien gegeben. Diese sind die orthogonalen Trajektorien der *Wellenfronten*  $S_j = \text{const}$ .

Sei  $t$  die Bogenlänge entlang einer solchen Linie. Die Eikonal-Gleichung impliziert dann, daß

$$S_j(t) = S_j^0 \pm t$$

gilt, wobei  $S_j^0$  den Wert von  $S_j$  bestimmt, von dem aus die Bogenlänge  $t$  gemessen wird. Wird die Bogenlänge positiv in die Richtung in der  $S$  wächst gemessen, so folgt  $S_j(t) = S_j^0 + t$ .

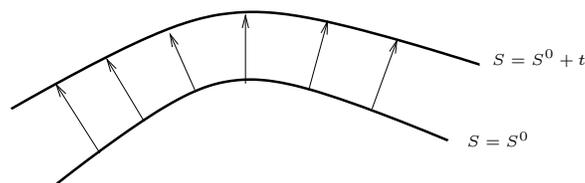


Abbildung 16: Entwicklung einer Wellenfront. Die Strahlen sind alle orthogonal zur Wellenfront  $S = S^0$  und haben die Länge  $t$ .

Die in der Transport-Gleichung auftretende Ableitung von  $A_j$  ist die Richtungsableitung entlang der Linie. Kürzt man diese mit  $\frac{dA}{dt}$  ab, so läßt sich die Transport-Gleichung als

$$2\frac{dA}{dt} + A\Delta S = 0$$

formulieren. Die Lösung dieser Gleichung ist gegeben durch ([KR60])

$$A_j(t) = A_j^0 \left( \frac{\rho_1 \rho_2}{(\rho_1 + t)(\rho_2 + t)} \right)^{\frac{1}{2}}, \quad (3.8)$$

wobei  $A_j^0$  der Wert von  $A_j$  bei  $t = 0$  und  $\rho_1, \rho_2$  die Haupt-Krümmungsradien der Wellenfront  $S_j = S_j^0$  sind.

Setzt man (3.7) in die Randbedingung  $u = 0$  für  $x \in \partial\Omega$  ein und setzt die Terme 0-ter Ordnung gleich Null, so erhält man

$$\sum_{j=1}^N e^{ikS_j} A_j = 0, \quad \text{auf } \partial\Omega.$$

Man nimmt weiter an, daß die Terme jeweils in Paaren verschwinden, d.h. daß für jedes  $j$  ein  $j' \neq j$  mit

$$e^{ikS_j} A_j + e^{ikS_{j'}} A_{j'} = 0 \quad (3.9)$$

existiert. Diese Annahme entspricht der physikalischen Erfahrung, daß jede Welle bzw. Strahl, der den Rand trifft, eine reflektierte Welle bzw. Strahl verursacht. Es wird nun gezeigt, daß diese Welle dem Reflexionsgesetz (*Einfallswinkel = Reflektionswinkel*) gehorcht.

Da Gleichung (3.9) für ein ganzes Intervall von Werten für  $k$  gültig ist, folgt

$$S_j = S_{j'} \quad \text{auf } \partial\Omega,$$

also auch

$$A_j = -A_{j'} \quad \text{auf } \partial\Omega.$$

Die Eikonal-Gleichung liefert dann die Beziehung

$$\frac{\partial S_j}{\partial \nu} = \pm \frac{\partial S_{j'}}{\partial \nu} \quad \text{auf } \partial\Omega.$$

Angenommen, das +-Zeichen gilt. Dann folgt sofort, daß  $S_j = S_{j'}$  auch in ganz  $\Omega$  gilt und somit  $A_j = -A_{j'}$  auf dem Rand  $\partial\Omega$ . Wegen der Darstellungs-Formel (3.8) von  $A_j$  folgt dann auch  $A_j = -A_{j'}$  in  $\Omega$ , so daß der Term  $e^{ikS_j} A_j + e^{ikS_{j'}} A_{j'}$  identisch Null in  $\Omega$  ist. Man kann somit die Terme für  $j$  und  $j'$  in der Darstellung (3.7) weglassen.

Somit darf O.B.d.A angenommen werden, daß

$$\begin{aligned} \frac{\partial S_j}{\partial \nu} &= -\frac{\partial S_{j'}}{\partial \nu} \quad \text{auf } \partial\Omega \\ S_j &= S_{j'} \end{aligned}$$

gilt. Dies ist das Reflexionsgesetz.

### 3.2 Quasi-klassische Approximationen

Für große  $k$  besteht also ein Zusammenhang zwischen den Lösungen der Helmholtz-Gleichung und dem zugehörigen Billard-System, so daß approximative Lösungen der Helmholtz-Gleichung konstruiert werden können, die bestimmte Merkmale des Billard-Systems ausnutzen. Dieser Zugang zur Helmholtz-Gleichung wird im folgenden an einigen Beispielen erläutert.

Gegenüber früheren Arbeiten kann aufgrund der Konstruktionsmethoden für  $p$ -Lazutkin-Billard aus Kapitel 2 ein erheblich umfangreicher Beispiel-Zoo für Lösungstypen von quasi-klassischen Approximationen angegeben werden. Verschiedene Lösungstypen werden im folgenden durch jeweils ein typisches Beispiel beschrieben. Die Berechnung der Beispiele erfolgte auf der Basis der neu entwickelten Algorithmen von Kapitel 4. Neu aufgenommen in den Zoo von quasi-klassischen Approximationen werden Lösungen, die entlang der  $p$ -periodischen Kaustik von  $p$ -Lazutkin-Billard verlaufen.

#### a. Lazutkin-Gebiet

Jedes strikt konvexe Billard besitzt einen Lazutkin-Streifen wie in Kapitel 1 erläutert wurde. Der Phasenraum des Billard-Systems besitzt eine Familie von invarianten Kurven in der Nähe des Randes  $\partial\Omega \times (0, \varepsilon)$  (vgl. Eigenschaft (B4)). Lazutkin konnte zeigen ([Laz73a]), daß es quasi-klassische Lösungen der Helmholtz-Gleichung gibt, die entlang des Randes lokalisiert sind. Solche Lösungen werden als *Whispering-Gallery-Moden* bezeichnet. Die ersten Untersuchungen zur Existenz solcher Moden wurden von Lord Rayleigh ([Ray64]) durchgeführt.

Es gibt mehrere Ansätze zur quasi-klassischen Konstruktion von Lösungen. Am häufigsten benutzt wird der von J.B. Keller und S. I. Rubinow ([KR60]), bei dem diese als Lösungen der oben beschriebenen Eikonal-Gleichung konstruiert werden.

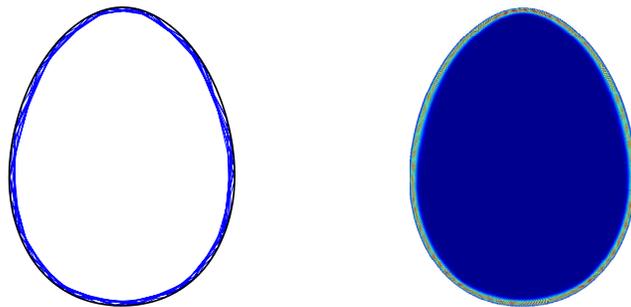


Abbildung 17: Links: Eine Trajektorie des Billard-Systems im Lazutkin-Streifen. Rechts: Eine Whispering-Gallery Lösung der Helmholtz-Gleichung. Der Betrag der Lösung  $u(x)$  ist abgebildet, die Farbskalierung ist so gewählt, daß blau dem Betrag Null entspricht und rot dem größten Betrag.

### b. 2-periodische Orbits

Nach Eigenschaft (B2) besitzt jedes Billard einen stabilen 2-periodischen Orbit. Nach dem Konstruktionsprinzip von Keller und Rubinow erhält man dann sogenannte *Bouncing-Ball-Moden*, die entlang dieses stabilen 2-Orbits lokalisiert sind.

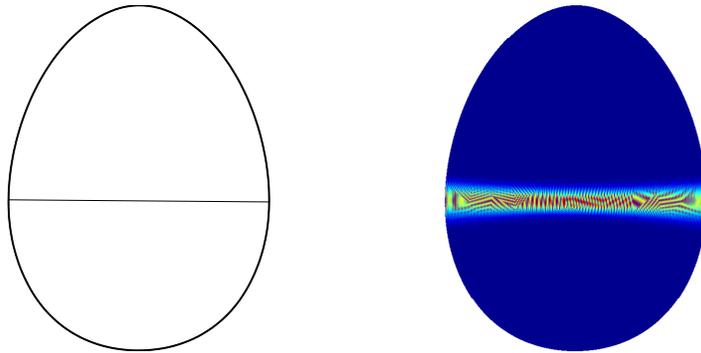


Abbildung 18: Links: Der stabile 2-periodische Orbit entsprechend der Weite des Billards (vgl. Eigenschaft (B2)). Rechts: Eine Bouncing-Ball Lösung der Helmholtz-Gleichung.

### c. Stabile periodische Orbits

Allgemeiner können stabilen periodischen Orbits quasi-klassische Lösungen zugeordnet werden. Ein stabiler Orbit verursacht eine Insel-Bildung im Phasenportrait, mit deren Hilfe zum Beispiel wiederum mit dem Keller-Rubinow Prinzip eine approximative Lösung konstruiert werden kann. Eine etwas andere Konstruktionsmethode wird in ([TSS02]) beschrieben. Hier wird die Helmholtz-Gleichung durch spezielle Annahmen der Lösungsstruktur in eine parabolische Gleichung transformiert, die anschließend (approximativ) gelöst wird.

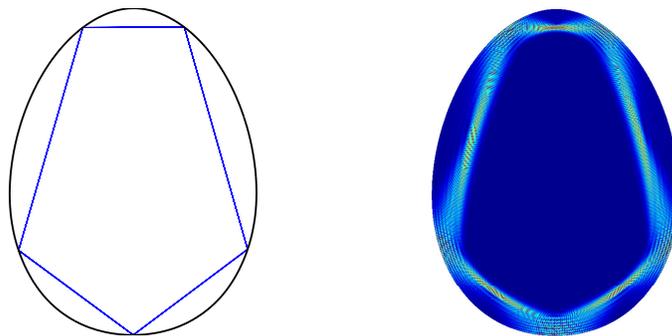


Abbildung 19: Links: Ein stabiler 5-periodischer Orbit des Billards. Rechts: Eine Lösung der Helmholtz-Gleichung, die entlang dieses stabilen 5-periodischen Orbits lokalisiert ist.

#### d. Instabile Orbits

In einer Reihe von Arbeiten ([Hel86], [Hel87]) konnte E. Heller zeigen, daß sogar hyperbolische periodische Orbits eine Konstruktion von bestimmten Näherungslösungen erlauben. Ein instabiler Orbit hinterläßt im Wellenbild einen *narbenartigen* Abdruck. Lösungen mit dieser Eigenschaft werden deshalb auch *Scars* genannt. Zur Konstruktion benutzt Heller mehrere Gaußsche Wellen-Pakete, die entlang des instabilen Orbits verteilt sind, und betrachtet die Entwicklung der Pakete unter dem Einfluß der Schrödinger-Gleichung. Die Narbenbildung im Wellenbild läßt sich anschaulich durch numerische Berechnungen im Billard-System erläutern: Startet eine Trajektorie auf einem hyperbolischen periodischen Orbit, so wird die Trajektorie schon nach wenigen Reflektionen am Rand durch numerische Fehler vom (invarianten) Orbit abgelenkt. Der Verlauf der Trajektorie wird mehr und mehr chaotisch. Die Gaußschen Wellen-Pakete können als numerische Störungen bei der Berechnung der Trajektorie aufgefaßt werden.

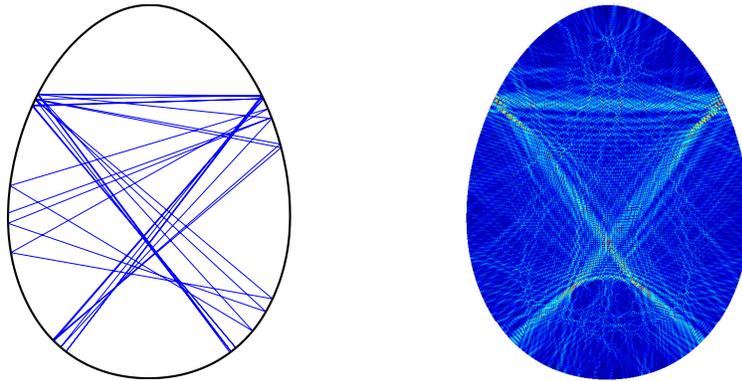


Abbildung 20: Links: Eine Trajektorie, die durch numerische Fehler schon nach wenigen Reflektionen vom zu grundlegenden hyperbolischen Orbit abgelenkt wird. Rechts: Ein Scar zu diesem 6-periodischen hyperbolischen Orbit. Der periodische Orbit ist im chaotisch wirkenden Hintergrund noch deutlich zu erkennen.

Die aufgeführten Beispiele zeigen, daß die von der Billard-Abbildung  $P$  induzierten Strukturen des Phasenraums zur Konstruktion approximativer Lösungen benutzt werden können, und umgekehrt aus der Lösungs-Struktur (zum Beispiel numerisch berechneter Lösungen) auf die zugrunde liegende klassische Dynamik im Billard zurück geschlossen werden kann.

#### e. $p$ -periodische Kaustik

Die in Kapitel 1 definierten  $p$ -Lazutkin-Billardards führen durch kleine Störungen (zum Beispiel bei der numerischen Konstruktion) zu Billards mit einem KAM-Streifen (Eigenschaft (B6)). Dieses Merkmal im Phasenraum kann ebenfalls zur Konstruktion quasi-klassischer Approximationen benutzt werden wie in [BHZP04] dargelegt. Die resultierenden Moden werden im Folgenden als

*innere Whispering-Gallery-Moden* (IWG) bezeichnet, weil sie im Inneren des Billards lokalisiert sind. In einer Kirche mit Chorraum in der Form eines  $p$ -Lazutkin-Billards, gibt es also zu jedem Wandpunkt einen Winkel, so daß ein Flüstern in Richtung dieses Winkels entlang der gesamten Kaustik im Inneren des Chorraums vernommen werden kann.

Eine direkte Anwendung der Keller-Rubinow-Methode ist in diesem Fall solange nicht möglich, wie eine approximative Parametrisierung des KAM-Streifens fehlt.

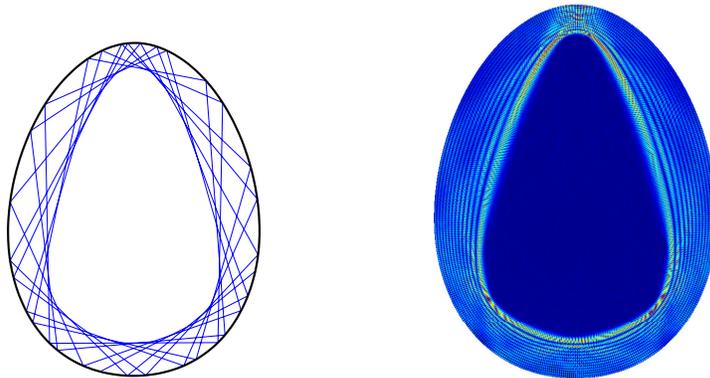


Abbildung 21: Links: Eine Trajektorie im KAM-Streifen nach Eigenschaft (B6) von  $p$ -Lazutkin-Billards. Rechts: Eine IWG-Lösung der Helmholtz-Gleichung in einem 4-Lazutkin-Billard.

### 3.3 Husimi-Projektion

Als Anwendung der effizienten numerischen Methoden von Kapitel 4 können nunmehr auch die für die Quantenphysiker wichtigen Husimi-Plots effizient berechnet werden. Der folgende Abschnitt geht kurz auf die Definitionen und numerischen Verfahren ein und zeigt einige Beispielrechnungen.

Es wurden Beispiele für die Konstruktion von approximativen Lösungen der Helmholtz-Gleichung aus klassischen Lösungen angegeben. In der Physik ist es üblich, die sogenannten Wigner- oder Husimifunktionen zur Beschreibung der Phasenraumstruktur von Quanten-Systemen zu verwenden ([Wig32], [Hus40], [BFS04]).

Im Phasenraumpunkt  $(p, q)$ , in dem der jeweilige Zustand des Quantensystems lokalisiert ist, ist die *Husimi-Funktion*  $H_n^B(p, q)$  definiert als Projektion der Eigenfunktion  $u_n$  zum Eigenwert  $k_n$  auf den sogenannten kohärenten Zustand ([Per95])

$$u_{(p,q),k}^B(x) := \left(\frac{k}{\pi}\right)^{\frac{1}{2}} (\det \operatorname{Im} B)^{\frac{1}{4}} e^{ik[\langle p, x-q \rangle + \frac{1}{2}\langle x-q, B(x-q) \rangle]}$$

mit einer symmetrischen, komplexen  $2 \times 2$  Matrix  $B$ , die die Gestalt des Zustandes bestimmt, d.h es ist

$$H_n^B(p, q) = \left( \frac{k_n}{2\pi} \right)^2 |\langle u_{(p,q),k_n}^B, u_n \rangle_\Omega|^2.$$

Zur Berechnung von  $H_n^B(p, q)$  ist mit dieser Definition ein 4-fach-Integral auszuführen ([Gut90]). In [BFS04] wird daher eine verwandte Funktion mit ähnlicher Aussagekraft aber leichter Berechenbarkeit eingeführt.

Dazu wählt man im Phasenraum  $C$  des Billards  $\Omega$  die Bogenlänge  $s$  und  $r = \cos \theta$  als Koordinaten (vgl. Kapitel 1). Setzt man  $\psi_n(\sigma) := \langle n(\sigma), \nabla u_n(x(\sigma)) \rangle$  für die Normalenableitung von  $u_n(x)$  im Punkt  $x(\sigma)$ , so definiert diese mittels der Greenschen Darstellungsformel schon vollständig die Funktion  $u_n(x)$  ([Eva98])

Falls die Funktion  $u_n(x)$  normiert ist, gilt die Beziehung

$$\frac{1}{2} \int_{\partial\Omega} |\psi_n(\sigma)|^2 \langle n(\sigma), x(\sigma) \rangle d\sigma = k_n^2.$$

In Abhängigkeit von einem Parameter  $b \in \mathbb{C}$  mit  $\text{Im } b > 0$  definiert man

$$c_{(s,r),k}^b(\sigma) := \left( \frac{k}{\pi} \right)^{\frac{1}{4}} (\text{Im } b)^{\frac{1}{4}} \sum_{m \in \mathbb{Z}} e^{ik[r(\sigma-s+mL) + \frac{b}{2}(\sigma-s+mL)^2]},$$

wobei  $L$  die Länge von  $\partial\Omega$  ist.

Als *Poincaré-Husimi-Funktion* definieren [BFS04] nun

$$h_n^b(s, r) := \frac{1}{2\pi k_n} \left| \int_{\partial\Omega} c_{(s,r),k_n}^b(\sigma) \psi_n(\sigma) d\sigma \right|^2.$$

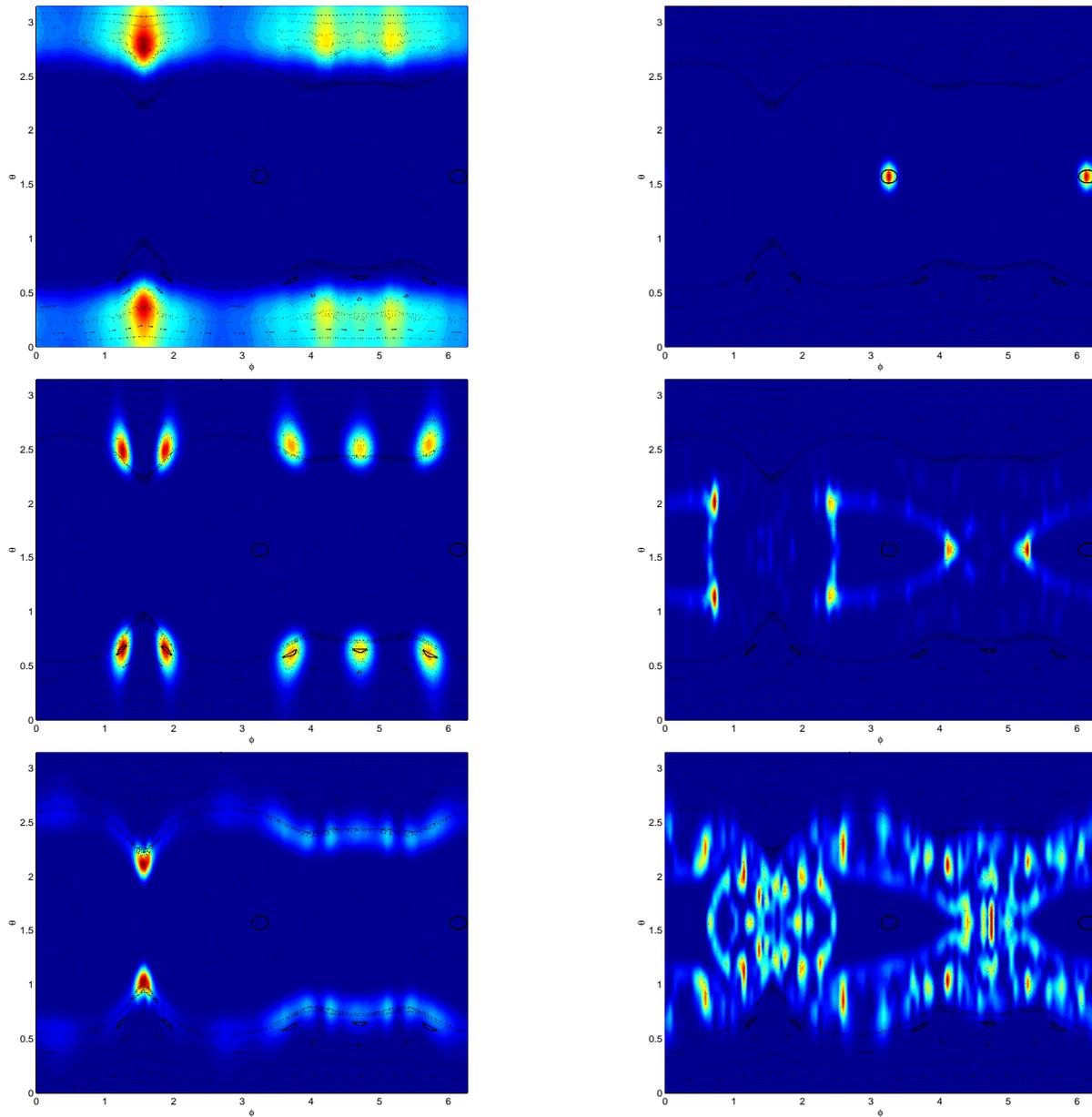


Abbildung 22: Die Poincaré-Husimi Funktion für einige Lösungen des Ei-Billards. Die Koordinaten  $(s, r)$  wurden transformiert auf  $(\phi, \theta)$ , wobei  $\phi$  der Polarwinkel des Randpunktes ist. Von oben links nach unten rechts: eine Whispering-Gallery-Mode, die Bouncing-Ball-Mode, eine Lösung entlang eines stabilen 5-periodischen Orbits, eine Scar-Lösung, eine innere Whispering-Gallery-Mode, eine chaotische Lösung. In der Husimi-Projektion der Lösungen auf den Phasenraum  $C$  der Billard-Abbildung kann man deutlich den oben beschriebenen quasi-klassischen Zusammenhang dieser Lösungstypen erkennen.

### Numerische Berechnung der Husimi-Funktion

Der Algorithmus zur Berechnung von Eigenwerten und Funktionen für große  $k$  aus Kapitel 4 liefert ein skalares Vielfaches  $\tilde{\psi}_n(\sigma)$  der Funktion  $\psi_n(\sigma)$ . Die Normalisierung  $\psi_n(\sigma)$  erhält man durch ([Bäc02])

$$\psi_n(\sigma) := \frac{\sqrt{2}k}{\sqrt{\int_{\partial\Omega} n(\sigma)x(\sigma)|\tilde{\psi}_n(\sigma)|^2 d\sigma}} \tilde{\psi}_n(\sigma).$$

Zu ihrer Berechnung verwendet man die Trapezregel.

Die unendliche Summe zur Berechnung von  $c_{(s,r),k}^b(\sigma)$  wird durch eine endliche Summe approximiert. Das Integral zur Berechnung von  $h_n^b(s,r)$  schließlich wird wiederum mit einer Trapezsumme berechnet.

Der Phasenraum wird dann durch ein Gitter  $(s_i, r_j)$  diskretisiert und die Funktion  $h_n^b(s,r)$  wird an den Gitterpunkten numerisch ausgewertet.

### 3.4 Galerie

Zur Veranschaulichung der beschriebenen Resultate und als Referenz für die Beurteilung der im folgenden zu beschreibenden Algorithmen wurden für 6 verschiedene Klassen von Lösungen der Helmholtz-Gleichung über dem Ei-Billard repräsentative Fälle berechnet.

In Abbildung 22 sind die jeweiligen Husimi-Projektionen abgebildet. Simulationen der zugrunde liegenden klassischen Trajektorien des Billard-Systems findet sich im obigen Text.

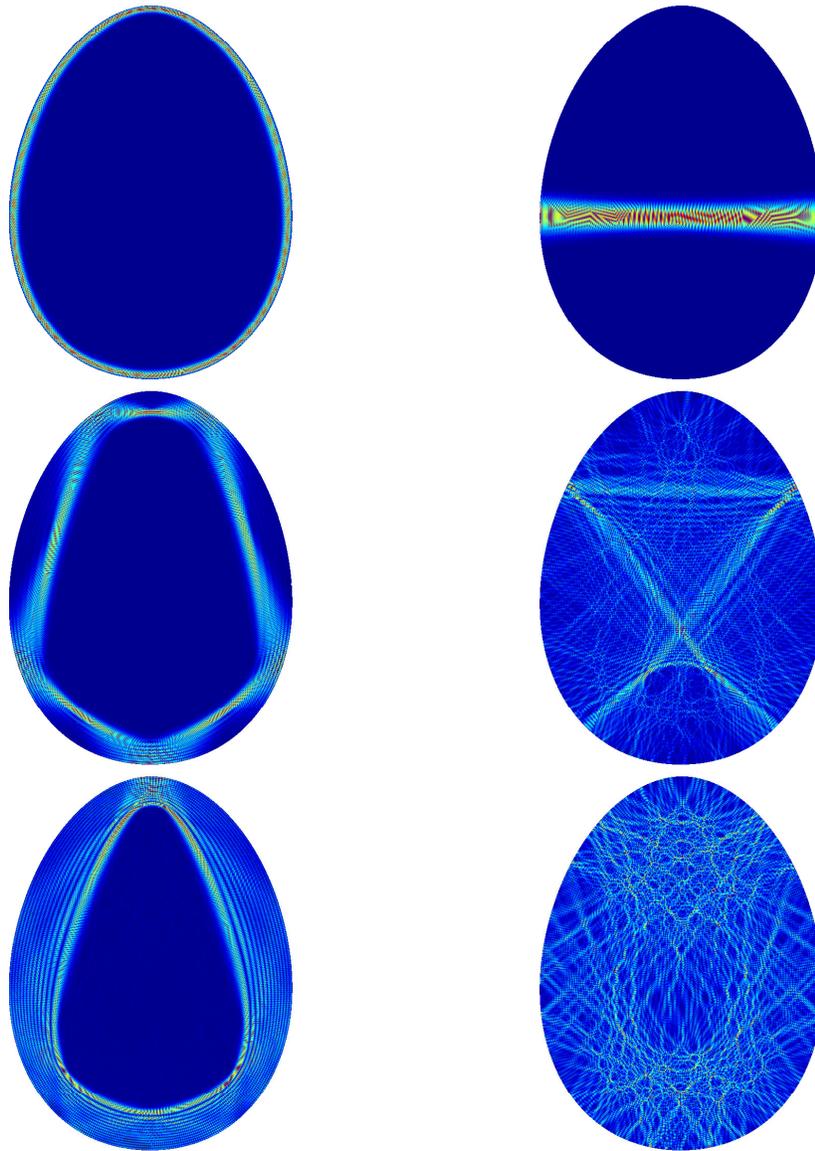


Abbildung 23: 6 Lösungstypen, die im folgenden als Referenz benutzt werden. In Abbildung 22 sind die Husimi-Projektionen dieser Lösungen abgebildet (in der gleichen Reihenfolge).

## 4 Randintegralmethode

Wie in Kapitel 3 gesehen, ist die effiziente numerische Behandlung des *quasi-klassischen Helmholtz-Eigenwertproblems* (HP) gesucht:

Bestimme alle  $k \in \mathbb{R}$  mit  $k \gg 1$  und  $a \leq k \leq b$ , so daß die Helmholtz-Gleichung

$$\Delta u + k^2 u = 0 \quad \text{in } \Omega \quad (4.1)$$

eine nicht-triviale Lösung mit

$$u|_{\partial\Omega} = 0$$

besitzt. Hierbei ist  $\Omega \subset \mathbb{R}^2$  ein beschränktes Gebiet mit geschlossenem  $\mathcal{C}^2$ -Rand. Ist  $\Omega$  insbesondere ein  $p$ -Lazutkin-Billard, so wurde in Kapitel 3 gezeigt, daß es Lösungen geben muß, die entlang der zugehörigen  $p$ -periodischen Kaustik lokalisiert sind.

Ein mit (HP) verwandtes Problem der mathematischen Physik ist das *quasi-klassische Transmissions-Problem* (TP):

Bestimme alle  $k \in \mathbb{C}$  mit  $|k| \gg 1$  und  $a \leq |k| \leq b$ , so daß die Gleichungen

$$\Delta u_0 + (kn_0)^2 u_0 = 0 \quad \text{in } \Omega \quad (4.2)$$

$$\Delta u + k^2 u = 0 \quad \text{in } \mathbb{R}^2 \setminus \bar{\Omega} \quad (4.3)$$

mit den Randbedingungen

$$u_0 - u = 0 \quad \text{auf } \partial\Omega \quad (4.4)$$

$$\frac{\partial u_0}{\partial \nu} - \frac{\partial u}{\partial \nu} = 0 \quad \text{auf } \partial\Omega \quad (4.5)$$

nicht-triviale Lösungen  $u_0 \in \mathcal{C}^2(\Omega) \cap \mathcal{C}(\bar{\Omega})$  und  $u \in \mathcal{C}^2(\mathbb{R}^2 \setminus \bar{\Omega}) \cap \mathcal{C}(\mathbb{R}^2 \setminus \Omega)$  besitzen. Hierbei ist  $\Omega \subset \mathbb{R}^2$  ein beschränktes Gebiet mit geschlossenem  $\mathcal{C}^2$ -Rand und  $n_0 \in \mathbb{R}$  eine Konstante. Zusätzlich soll für  $|x| \rightarrow \infty$  die *Sommerfeld'sche*-Ausstrahlungsbedingung

$$u(x) = O\left(\frac{1}{r}\right), \quad \frac{\partial u}{\partial r} - iku = o\left(\frac{1}{r}\right), \quad \text{für } r = |x| \rightarrow \infty \quad (4.6)$$

erfüllt sein.

In Kapitel 5 wird eine Anwendung aus der Halbleiter-Technik behandelt, die auf dieses Problem reduziert werden kann.

Trotz der eher einfachen Struktur der Gleichungen (4.1), (4.2) und (4.3) bereiten die Probleme (HP) und (TP) enorme numerische Schwierigkeiten, die durch die zusätzliche quasi-klassische Anforderung  $k \gg 1$  bzw.  $|k| \gg 1$  verursacht werden. Die mathematische, besonders aber die physikalische Literatur, schlägt einige Verfahren zur Lösung von (HP) vor, eine Übersicht und einen

Vergleich mit dem weiter unten vorgestellten Algorithmus befinden sich in 4.5. Für das (TP) hingegen existieren nur wenige Ansätze, für die zudem keine Fehlerabschätzungen existieren.

Diese Lücken werden in diesem Kapitel geschlossen. Es wird ein Algorithmus vorgestellt, der sowohl (HP) als auch (TP) effizient löst. Es wird dann gezeigt, daß der Diskretisierungsfehler exponentiell schnell fällt.

Ausgangspunkt des Algorithmus ist eine äquivalente Formulierung von (HP) und (TP) als Integralgleichung. Die auftretenden Integraloperatoren werden auf einer von Kussmaul ([Kus68]) und Kress ([Kre95]) aufbauenden Darstellung numerisch effektiv diskretisiert. Diese beiden Arbeiten beschäftigen sich jedoch nicht mit den oben beschriebenen Eigenwert-Problemen.

Für die eigentliche Eigenwertsuche im Intervall  $a \leq |k| \leq b$  wird dann eine Prädiktor-Korrektor-Strategie entwickelt, die gegenüber früheren Verfahren die folgenden Vorteile bietet:

- Der Algorithmus ist sehr robust, es werden keine Eigenwerte in dem Intervall  $[a, b]$  übersehen.
- Der Algorithmus findet die Eigenwerte im Intervall  $[a, b]$  *direkt*, er benötigt keine Suche mit Hilfe von Testfunktionen.

Zum Schluß wird ein Algorithmus für die Lösung von nicht-linearen Eigenwertproblemen der Form

$$A(k(\varepsilon), \varepsilon)x = 0,$$

wobei  $A(k, \varepsilon)$  eine voll-besetzte, nicht-linear von  $k$  und  $\varepsilon$  abhängige Matrix ist. Probleme dieses Types tauchen bei der numerischen Behandlung von (HP) und (TP) unter kleinen Deformationen des Randes  $\partial\Omega$  auf.

#### 4.1 Formulierung als Randintegralgleichung

In diesem Abschnitt werden die analytischen Hilfsmittel zur Formulierung von (HP) und (TP) als Integralgleichungen referiert. Die Darstellung folgt im wesentlichen dem Vorgehen in [Kre99].

Sei

$$\Phi_k(x, y) := \frac{i}{4} H_0^{(+)}(k|x - y|), \quad x \neq y$$

die 2-dimensionale *Fundamentallösung* der Helmholtz-Gleichung, wobei  $H_0^{(+)}(z)$  die Hankelfunktion erster Art ist. Für festes  $y \in \mathbb{R}^2$  erfüllt  $\Phi_k(x, y)$  die Helmholtz-Gleichung auf  $\mathbb{R}^2 \setminus \{y\}$ .

Man erhält zunächst Lösungen von (4.1) ohne Berücksichtigung der Randbedingungen aufgrund des folgenden Satzes

**4.1 Satz.** Sei  $\phi$  eine beliebige Funktion aus  $C(\partial\Omega)$ . Die Funktionen

$$u(x) = \int_{\partial\Omega} \Phi_k(x, y) \phi(y) ds_y, \quad x \in \mathbb{R}^2 \setminus \partial\Omega \quad (4.7)$$

bzw.

$$v(x) = \int_{\partial\Omega} \frac{\partial\Phi_k}{\partial\nu_y}(x, y) \phi(y) ds_y, \quad x \in \mathbb{R}^2 \setminus \partial\Omega \quad (4.8)$$

erfüllen die Helmholtz-Gleichung (4.1) in  $\mathbb{R}^2 \setminus \partial\Omega$  und sind dort unendlich oft differenzierbar. Man nennt  $u(x)$  bzw.  $v(x)$  das Einfachschicht-Potential bzw. das Doppelschicht-Potential. Die Funktion  $\phi$  heißt Belegung.

**Beweis:** Für festes  $x \notin \partial\Omega$  sind die zweiten Ableitungen von  $\Phi_k(x, y)$  bzw.  $\frac{\partial\Phi_k}{\partial\nu_y}(x, y)$  nach  $x$  gleichmäßig stetig, so daß die Differentiation unter das Integral gezogen werden kann und die Differentiationsreihenfolge vertauscht werden darf. Der Satz folgt dann aus den Lösungseigenschaften der Fundamentallösung.  $\square$

Am Rand  $\partial\Omega$  werden die obigen Integranden singulär. Das Verhalten am Rand wird durch die beiden folgenden Sätze beschrieben, die in der Literatur als *Sprungbedingung* bekannt sind.

**4.2 Satz (Sprungbedingung für das Einfachschicht-Potential, ([Kre99, Thm. 6.18])).**

Für das Einfachschicht-Potential  $u$  mit  $C^2$ -stetigem Rand  $\partial\Omega$  und stetiger Belegung  $\phi$  gilt

$$\frac{\partial u_{\pm}}{\partial\nu}(x) = \int_{\partial\Omega} \frac{\partial\Phi_k}{\partial\nu_x}(x, y) \phi(y) ds_y \mp \frac{1}{2} \phi(x), \quad x \in \partial\Omega. \quad (4.9)$$

Hierbei wird

$$\frac{\partial u_{\pm}}{\partial\nu}(x) := \lim_{h \rightarrow +0} \nu(x) \cdot \nabla u(x \pm h\nu(x))$$

im Sinne gleichmäßiger Konvergenz auf dem Rand  $\partial\Omega$  verstanden. Das Integral existiert als uneigentliches Integral.

**4.3 Satz (Sprungbedingung für das Doppelschicht-Potential, ([Kre99, Thm. 6.17])).**

Das Doppelschicht-Potential  $v$  mit  $C^2$ -stetigem Rand  $\partial\Omega$  und stetiger Belegung  $\phi$  läßt sich stetig von  $\Omega$  nach  $\bar{\Omega}$  und  $\mathbb{R}^2 \setminus \bar{\Omega}$  nach  $\mathbb{R}^2 \setminus \Omega$  fortsetzen. Für die Grenzwerte gilt

$$v_{\pm}(x) = \int_{\partial\Omega} \frac{\partial\Phi_k}{\partial\nu_y} \phi(y) ds_y \pm \frac{1}{2} \phi(x), \quad x \in \partial\Omega. \quad (4.10)$$

Hierbei ist

$$v_{\pm}(x) := \lim_{h \rightarrow +0} v(x \pm h\nu(x))$$

und das Integral existiert als uneigentliches Integral.

Zusätzlich gilt

$$\lim_{h \rightarrow +0} \nu(x) \cdot \{\nabla v(x + h\nu(x)) - \nabla v(x - h\nu(x))\} = 0.$$

Zur Lösung des Helmholtz-Problems kann also ein Ansatz als Einfach- bzw. Doppelschicht-Potential oder als Kombination beider gemacht werden. Zur Lösung des Transmissions-Problems wird eine Kombination eines Einfach- bzw. Doppelschicht-Potential gemacht.

In beiden Fällen muß die Belegung  $\phi$  so gewählt werden, daß die vorgegebenen Randbedingungen erfüllt werden. Dies führt im Falle von (HP) zu

**4.4 Satz.** *Das Doppelschicht-Potential*

$$u(x) = \int_{\partial\Omega} \frac{\partial\Phi_k}{\partial\nu_y}(x, y)\phi(y)ds_y$$

mit stetiger Belegung  $\phi$  ist eine nicht-triviale Lösung von (HP), falls  $\phi$  eine nicht-triviale Lösung der Integralgleichung

$$\phi(x) - 2 \int_{\partial\Omega} \frac{\partial\Phi_k}{\partial\nu_y}(x, y)\phi(y)ds_y = 0, \quad x \in \partial\Omega$$

ist.

**Beweis:** Anwendung von Satz (4.3).  $\square$

**Bemerkung:** In dieser Arbeit werden ausschließlich Eigenwert-Probleme behandelt. Nur im Fall, daß  $k$  ein Eigenwert von (HP) bzw. (TP) ist, existieren nicht-triviale Lösungen ([Eva98]). Ist  $k$  hingegen kein Eigenwert, so ist die Null-Lösung die einzige Lösung von (HP) bzw. (TP).

Der Integraloperator  $K : \mathcal{C}(\partial\Omega) \rightarrow \mathcal{C}(\partial\Omega)$  mit

$$(K\psi)(x) := 2 \int_{\partial\Omega} \frac{\partial\Phi_k}{\partial\nu_y}\psi(y)ds_y, \quad x \in \partial\Omega \tag{4.11}$$

ist kompakt auf  $\mathcal{C}(\partial\Omega)$  und  $\mathcal{C}^{0,\alpha}(\partial\Omega)$  für  $0 < \alpha < 1$  und bildet  $\mathcal{C}^{0,\alpha}(\partial\Omega)$  nach  $\mathcal{C}^{1,\alpha}(\partial\Omega)$  ab ([Kre99, Seite 82]).

Damit kann man Satz (4.4) umformulieren:

$$u(x) = \int_{\partial\Omega} \frac{\partial\Phi_k}{\partial\nu_y}(x, y)\phi(y)ds_y$$

ist eine Lösung von (HP), falls  $\phi$  eine nicht-triviale Lösung der Integralgleichung

$$(I - K)\phi = 0 \tag{4.12}$$

ist, wobei  $I$  der Identitätsoperator auf  $\mathcal{C}(\partial\Omega)$  ist.

Für die Behandlung des Transmissions-Problems werden zusätzlich Integraloperatoren definiert.

Nach [Kre99] ist

$$(K^*\phi)(x) := 2 \int_{\partial\Omega} \frac{\partial\Phi_k}{\partial\nu_x} \phi(y) ds_y, \quad x \in \partial\Omega \quad (4.13)$$

der zu  $K$  bezüglich

$$\langle \phi, \psi \rangle := \int_{\partial\Omega} \phi\psi ds, \quad \phi, \psi \in \mathcal{C}(\partial\Omega)$$

adjungierte Integraloperator, d.h. es gilt  $\langle K\phi, \psi \rangle = \langle \phi, K^*\psi \rangle$ .

Der durch

$$(S\phi)(x) := 2 \int_{\partial\Omega} \Phi_k(x, y) \phi(y) ds_y, \quad x \in \partial\Omega \quad (4.14)$$

definierte Integraloperator  $S : \mathcal{C}(\partial\Omega) \rightarrow \mathcal{C}(\partial\Omega)$  ist selbstadjungiert und besitzt die gleichen Kompaktheits- und Abbildungseigenschaften wie  $K$ , ([Kre99]).

Sei nun  $\mathcal{A}(\partial\Omega)$  der Raum aller stetigen Funktionen  $\psi$  mit der Eigenschaft, daß das Doppelschicht-Potential  $v$  mit der Belegung  $\psi$  stetige Normalenableitungen auf beiden Seiten von  $\partial\Omega$  hat. Da beide Normalenableitungen übereinstimmen und  $\mathcal{A}(\partial\Omega)$  nicht leer ist, wird durch

$$(T\psi)(x) := 2 \frac{\partial}{\partial\nu_x} \int_{\partial\Omega} \frac{\partial\Phi_k}{\partial\nu_y} \psi(y) ds_y, \quad x \in \partial\Omega \quad (4.15)$$

ein Integraloperator  $T : \mathcal{A}(\partial\Omega) \rightarrow \mathcal{C}(\partial\Omega)$  definiert.

Sei nun  $k_0 := k \cdot n_0$  und

$$\Phi_0(x, y) := \frac{i}{4} H_0^{(+)}(k_0|x - y|)$$

die Fundamentallösung zu Gleichung (4.2). Analog zu den Integraloperatoren  $K, K^*, S$  und  $T$  werden nun durch Ersetzung der Funktion  $\Phi(x, y)$  durch  $\Phi_0(x, y)$  in den Definitionen (4.11), (4.13), (4.14) und (4.15) die Integraloperatoren  $K_0, K_0^*, S_0$  und  $T_0$  definiert.

Für die Lösung des Transmissions-Problems wird eine Kombination aus Einfach- und Doppelschichtpotential angesetzt, die auf eine Arbeit von H. Brakhage und P. Werner ([BW65]) zurückgeht. Man macht den Ansatz

$$u(x) = \int_{\partial\Omega} \left\{ \frac{\partial\Phi(x, y)}{\partial\nu_y} \psi(y) + \Phi(x, y) \phi(y) \right\} ds_y = \frac{1}{2}(K\psi + S\phi), \quad x \in \mathbb{R}^2 \setminus \bar{\Omega} \quad (4.16)$$

und

$$u_0(x) = \int_{\partial\Omega} \left\{ \frac{\partial\Phi_0(x, y)}{\partial\nu_y} \psi(y) + \Phi_0(x, y) \phi(y) \right\} ds_y = \frac{1}{2}(K_0\psi + S_0\phi), \quad x \in \Omega \quad (4.17)$$

mit Funktionen  $\phi(x), \psi(x) \in \mathcal{C}(\partial\Omega)$ . Nach der Konstruktion erfüllt  $u_0$  die Gleichung (4.2) und  $u$  die Gleichung (4.3) sowie die Sommerfeld-Bedingung. Es müssen also nun die Funktionen  $\phi(x)$  und  $\psi(x)$  so bestimmt werden, daß zusätzlich auch noch die Randbedingungen erfüllt werden.

Die Sprungbedingungen (4.9) und (4.10) liefern die Gleichungen

$$\begin{aligned} u_+ &= \frac{1}{2}K\psi + \frac{1}{2}\psi + \frac{1}{2}S\phi, & (u_0)_- &= \frac{1}{2}K_0\psi - \frac{1}{2}\psi + \frac{1}{2}S_0\phi \\ \frac{\partial u_+}{\partial\nu_x} &= \frac{1}{2}T\psi + \frac{1}{2}K^*\phi - \frac{1}{2}\phi, & \frac{\partial (u_0)_-}{\partial\nu_x} &= \frac{1}{2}T_0\psi + \frac{1}{2}K_0^*\phi + \frac{1}{2}\phi. \end{aligned}$$

Aus den Stetigkeitsforderungen  $u - u_0 = 0$  und  $\frac{\partial u}{\partial\nu} - \frac{\partial u_0}{\partial\nu} = 0$  am Rand erhält man ein Integralgleichungssystem für die unbekanntenen Belegungen  $\phi$  und  $\psi$ ,

$$2\psi + (K - K_0)\psi + (S - S_0)\phi = 0 \quad (4.18)$$

$$2\phi - (T - T_0)\psi - (K^* - K_0^*)\phi = 0, \quad (4.19)$$

welches sich in der Form

$$\left( 2 \cdot \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} - \begin{pmatrix} -(K - K_0) & -(S - S_0) \\ T - T_0 & K^* - K_0^* \end{pmatrix} \right) \begin{pmatrix} \psi \\ \phi \end{pmatrix} = 0 \quad (4.20)$$

schreiben läßt.

Die Darstellungen (4.12) und (4.20) werden zusammengefaßt im

**4.5 Satz.** *Sei im Falle des Helmholtz-Problems durch  $A : \mathcal{C}(\partial\Omega) \rightarrow \mathcal{C}(\partial\Omega)$  und*

$$A(k) := (I - K) \quad (4.21)$$

*bzw. im Falle des Transmissions-Problems durch  $A(k) : \mathcal{C}(\partial\Omega) \times \mathcal{C}(\partial\Omega) \rightarrow \mathcal{C}(\partial\Omega) \times \mathcal{C}(\partial\Omega)$  und*

$$A(k) := 2 \cdot \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} - \begin{pmatrix} -(K - K_0) & -(S - S_0) \\ T - T_0 & K^* - K_0^* \end{pmatrix} \quad (4.22)$$

*ein Integraloperator definiert. Seien  $a, b$  wie in (HP) bzw. (TP) vorgegeben. Dann gilt:*

- *Das Paar  $k$  und  $u$  ist Lösung von (HP), falls  $a \leq k \leq b$  ist und die Integralgleichung*

$$A(k)\phi = 0$$

*mit dem Integraloperator aus (4.21) eine nicht-triviale Lösung  $\phi$  besitzt. In diesem Fall ist dann*

$$u(x) = \int_{\partial\Omega} \frac{\partial\Phi_k}{\partial\nu_y}(x, y)\phi(y)ds_y.$$

- Das Tripel  $k, u_0$  und  $u$  ist Lösung von (TP), falls  $a \leq |k| \leq b$  ist und die Integralgleichung

$$A(k) \begin{pmatrix} \psi \\ \phi \end{pmatrix} = 0$$

mit dem Integraloperator aus (4.22) eine nicht-triviale Lösung  $\begin{pmatrix} \psi \\ \phi \end{pmatrix}$  besitzt. In diesem Fall ist dann

$$u_0(x) = \int_{\partial\Omega} \left\{ \frac{\partial\Phi_0(x, y)}{\partial\nu_y} \psi(y) + \Phi_0(x, y)\phi(y) \right\} ds_y \quad x \in \Omega$$

und

$$u(x) = \int_{\partial\Omega} \left\{ \frac{\partial\Phi(x, y)}{\partial\nu_y} \psi(y) + \Phi(x, y)\phi(y) \right\} ds_y \quad x \in \mathbb{R}^2 \setminus \bar{\Omega}.$$

**4.6 Definition.** Sind  $k$  und  $u$  Lösungen von (HP) bzw.  $k, u_0$  und  $u$  Lösungen von (TP), so nennt man in beiden Fällen  $k$  die *Eigenresonanz* und  $u$  bzw.  $u_0$  und  $u$  die *Eigenfunktion*.

## 4.2 Diskretisierung des Operators $A(k)$

Der Operator  $A(k)$  muß für die numerische Behandlung diskretisiert werden. Hierzu wird eine geeignete Darstellung der jeweiligen Integralkerne entwickelt, die die Ergebnisse von Kussmaul und Kress auf die Probleme (HP) bzw. (TP) verallgemeinern. Die jeweiligen Integralkerne werden in einen analytischen Teil und in einen Teil mit einer Singularität der Form  $\log(4 \sin^2 \frac{t-\tau}{2})$  aufgespalten. Für beide Teile werden dann gesonderte Quadraturverfahren entwickelt. In der Aufspaltung der Kerne in einen analytischen bzw. singulären Teil liegt die Effizienz dieser Diskretisierungsform, da der singuläre Teil durch ein geeignetes Quadraturverfahren integriert werden kann.

Der Rand  $\partial\Omega$  sei parametrisiert in der Form

$$\partial\Omega = \{x(t) | x(t) = (x_1(t), x_2(t)), 0 \leq t \leq 2\pi\},$$

so daß  $|x'(t)| > 0$  ist und  $\partial\Omega$  gegen den Uhrzeigersinn orientiert ist. Es bezeichne  $\nu(x(t))$  die äußere Einheitsnormale und  $n(t)$  den Normalenvektor  $n(t) := |x'(t)| \cdot \nu(x(t)) = (x'_2(t), -x'_1(t))$ .

Zur Vereinfachung der Notation wird die Abhängigkeit der Integraloperatoren bzw. der jeweiligen Integralkerne von  $k$  unterdrückt.

Zunächst werden die einzelnen Aufspaltungen der Integralkerne vorgestellt.

**a. Aufspaltung des Operators  $K$** 

Unter Berücksichtigung von  $H_0^{(+)'}(z) = -H_1^{(+)}(z)$ , folgt zunächst

$$\nabla_y \Phi_k(x, y) = \frac{ik}{4} \frac{x-y}{|x-y|} H_1^{(+)}(k|x-y|).$$

Für den Operator  $K$  erhält man

$$\begin{aligned} (K\phi)(x(t)) &= \int_0^{2\pi} \frac{ik}{2} \frac{n(\tau) \cdot (x(t) - x(\tau))}{|x'(\tau)| \cdot |x(t) - x(\tau)|} H_1^{(+)}(k|x(t) - x(\tau)|) \phi(x(\tau)) |x'(\tau)| d\tau \\ &= \int_0^{2\pi} H(t, \tau) \phi(x(\tau)) d\tau, \end{aligned}$$

mit dem Integrkern

$$H(t, \tau) := \frac{ik}{2} \frac{n(\tau) \cdot (x(t) - x(\tau))}{|x(t) - x(\tau)|} H_1^{(+)}(k|x(t) - x(\tau)|). \quad (4.23)$$

Für  $t \rightarrow \tau$  besitzt der Kern  $H(t, \tau)$  eine hebbare Singularität. Um numerische Schwierigkeiten bei der Auswertung des Kerns  $H(t, \tau)$  zu vermeiden, verwendet man die von R. Kussmaul ([Kus68]) angegebene Darstellung

$$H(t, \tau) = H_1(t, \tau) \log \left( 4 \sin^2 \frac{t - \tau}{2} \right) + H_2(t, \tau)$$

mit den analytischen Funktionen

$$\begin{aligned} H_1(t, \tau) &:= -\frac{k}{2\pi} \frac{n(\tau) \cdot (x(t) - x(\tau))}{|x(t) - x(\tau)|} J_1(k|x(t) - x(\tau)|) \\ H_2(t, \tau) &:= H(t, \tau) - H_1(t, \tau) \log \left( 4 \sin^2 \frac{t - \tau}{2} \right), \end{aligned}$$

wobei  $J_1(z)$  die erste Besselfunktion ist. Für die Diagonalterme gilt

$$H_1(t, t) = 0, \quad H_2(t, t) = \frac{1}{2\pi} \frac{n(t) \cdot x''(t)}{|x'(t)|^2}. \quad (4.24)$$

Die Details der Herleitung dieser Darstellung sind im Anhang angegeben.

**b. Aufspaltung des Operators  $K - K_0$** 

Ähnlich wie oben folgt

$$\begin{aligned} (K - K_0)(\psi)(x(t)) &= \int_0^{2\pi} \left\{ \frac{i}{2} \frac{n(\tau)}{|x'(\tau)|} \frac{x(t) - x(\tau)}{|x(t) - x(\tau)|} \left[ kH_1^{(+)}(k|x(t) - x(\tau)|) \right. \right. \\ &\quad \left. \left. - k_0H_1^{(+)}(k_0|x(t) - x(\tau)|) \right] \cdot |x'(\tau)| \right\} \psi(x(\tau)) \, d\tau \\ &= \int_0^{2\pi} \tilde{H}(t, \tau) \psi(x(\tau)) \, d\tau, \end{aligned} \quad (4.25)$$

mit dem Integralkern

$$\tilde{H}(t, \tau) := \frac{i}{2} n(\tau) \frac{x(t) - x(\tau)}{|x(t) - x(\tau)|} \left[ kH_1^{(+)}(k|x(t) - x(\tau)|) - k_0H_1^{(+)}(k_0|x(t) - x(\tau)|) \right].$$

Analog zu oben wird der singuläre Anteil dieses Kerns abgespalten. Der Kern  $\tilde{H}(t, \tau)$  läßt sich dann schreiben als

$$\tilde{H}(t, \tau) = \tilde{H}_1(t, \tau) \log \left( 4 \sin^2 \frac{t - \tau}{2} \right) + \tilde{H}_2(t, \tau),$$

wobei  $\tilde{H}_1(t, \tau)$  durch

$$\tilde{H}_1(t, \tau) := -\frac{1}{2\pi} n(\tau) \frac{x(t) - x(\tau)}{|x(t) - x(\tau)|} \left[ kJ_1(k|x(t) - x(\tau)|) - k_0J_1(k_0|x(t) - x(\tau)|) \right]$$

gegeben ist und der Kern  $\tilde{H}_2(t, \tau)$  durch

$$\tilde{H}_2(t, \tau) = \tilde{H}(t, \tau) - \tilde{H}_1(t, \tau) \log \left( 4 \sin^2 \frac{t - \tau}{2} \right).$$

Für die Diagonalterme erhält man gemäß (4.24)

$$\tilde{H}_1(t, t) = 0 \quad \text{und} \quad \tilde{H}_2(t, t) = 0.$$

**c. Aufspaltung des Operators  $K^* - K_0^*$** 

Für den Operator  $(K^* - K_0^*)$  findet man

$$\begin{aligned} (K^* - K_0^*)(\phi)(x(t)) &= \int_0^{2\pi} \left\{ -\frac{i}{2} \frac{n(t)}{|x'(t)|} \frac{x(t) - x(\tau)}{|x(t) - x(\tau)|} \left[ kH_1^{(+)}(k|x(t) - x(\tau)|) \right. \right. \\ &\quad \left. \left. - k_0H_1^{(+)}(k_0|x(t) - x(\tau)|) \right] \cdot |x'(\tau)| \right\} \phi(x(\tau)) \, d\tau \\ &= \frac{1}{|x'(t)|} \int_0^{2\pi} \tilde{H}^*(t, \tau) \phi(x(\tau)) \, d\tau, \end{aligned} \quad (4.26)$$

mit dem Integralkern

$$\tilde{H}^*(t, \tau) = -\frac{i}{2}n(t) \frac{x(t) - x(\tau)}{|x(t) - x(\tau)|} \left[ kH_1^{(+)}(k|x(t) - x(\tau)|) - k_0H_1^{(+)}(k_0|x(t) - x(\tau)|) \right] \cdot |x'(\tau)|.$$

Wiederum ergibt Abspalten des singulären Anteils die Gestalt

$$\tilde{H}^*(t, \tau) = \tilde{H}_1^*(t, \tau) \log \left( 4 \sin^2 \frac{t - \tau}{2} \right) + \tilde{H}_2^*(t, \tau),$$

wobei diesmal

$$\tilde{H}_1^*(t, \tau) := \frac{1}{2\pi}n(t) \frac{x(t) - x(\tau)}{|x(t) - x(\tau)|} \left[ kJ_1(k|x(t) - x(\tau)|) - k_0J_1(k_0|x(t) - x(\tau)|) \right] \cdot |x'(\tau)|$$

und

$$\tilde{H}_2^*(t, \tau) = \tilde{H}^*(t, \tau) - \tilde{H}_1^*(t, \tau) \log \left( 4 \sin^2 \frac{t - \tau}{2} \right)$$

ist. Für die Diagonalterme erhält man wie oben

$$\tilde{H}_1^*(t, t) = 0 \quad \text{und} \quad \tilde{H}_2^*(t, t) = 0.$$

#### d. Aufspaltung des Operators $S - S_0$

Zunächst wird die Aufspaltung des Operators  $S$  diskutiert. Mit der obigen Parametrisierung des Randes findet man

$$\begin{aligned} (S\phi)(x(t)) &= \int_0^{2\pi} \frac{i}{2} H_0^{(+)}(k|x(t) - x(\tau)|) |x'(\tau)| \phi(x(\tau)) \, d\tau \\ &= \int_0^{2\pi} M(t, \tau) |x'(\tau)| \phi(x(\tau)) \, d\tau \end{aligned}$$

mit dem Integralkern

$$M(t, \tau) = \frac{i}{2} H_0^{(+)}(k|x(t) - x(\tau)|).$$

Auch bei diesem Kern kann der singuläre Anteil analog zu den obigen Kernen abgespalten werden. Es gilt

$$M(t, \tau) = M_1(t, \tau) \log \left( 4 \sin^2 \frac{t - \tau}{2} \right) + M_2(t, \tau)$$

mit

$$M_1(t, \tau) = -\frac{1}{2\pi} J_0(k|x(t) - x(\tau)|)$$

und

$$M_2(t, \tau) = M(t, \tau) - M_1(t, \tau) \log \left( 4 \sin^2 \frac{t - \tau}{2} \right).$$

Die Diagonalterme ergeben sich zu

$$M_1(t, t) = -\frac{1}{2\pi} \quad \text{und} \quad M_2(t, t) = \frac{i}{2} - \frac{C}{\pi} - \frac{1}{\pi} \log \frac{k|x'(t)|}{2}.$$

Hieraus läßt sich nun leicht die gesuchte Zerlegung des Operators  $S - S_0$  bestimmen. Es gilt

$$\begin{aligned} (S - S_0)(\phi)(x(t)) &= \int_0^{2\pi} \frac{i}{2} \left( H_0^{(+)}(k|x(t) - x(\tau)|) - H_0^{(+)}(k_0|x(t) - x(\tau)|) \right) |x'(\tau)| \phi(x(\tau)) \, d\tau \\ &= \int_0^{2\pi} \tilde{M}(t, \tau) |x'(\tau)| \phi(x(\tau)) \, d\tau \end{aligned}$$

mit

$$\tilde{M}(t, \tau) = \frac{i}{2} \left( H_0^{(+)}(k|x(t) - x(\tau)|) - H_0^{(+)}(k_0|x(t) - x(\tau)|) \right).$$

Man definiert

$$\tilde{M}_1(t, \tau) = -\frac{1}{2\pi} \left( J_0(k|x(t) - x(\tau)|) - J_0(k_0|x(t) - x(\tau)|) \right)$$

und

$$\tilde{M}_2(t, \tau) = \tilde{M}(t, \tau) - \tilde{M}_1(t, \tau) \log \left( 4 \sin^2 \frac{t - \tau}{2} \right).$$

Da die Resonanz  $k$  im Diagonalterm  $M_2(t, t)$  auftaucht, hebt sich dieser nicht weg, man findet

$$\tilde{M}_1(t, t) = 0 \quad \text{und} \quad \tilde{M}_2(t, t) = \frac{1}{\pi} \log n.$$

### e. Aufspaltung des Operators $T - T_0$

Wie in [Kre95] gezeigt wird, besitzt der Operator  $T$  die Darstellung

$$(T\phi)(x(t)) = \left( \frac{d}{ds} S \frac{d\phi}{ds} \right) (x(t)) + k^2 \nu \cdot S(\nu\phi)(x(t)). \quad (4.27)$$

Mittels partieller Integration und wegen  $\phi(x(0)) = \phi(x(2\pi))$  findet man

$$\begin{aligned} \left( \frac{d}{ds} S \frac{d\phi}{ds} \right) (x(t)) &= \frac{i}{2} \frac{1}{|x'(t)|} \int_0^{2\pi} \frac{\partial}{\partial t} H_0^{(+)}(k|x(t) - x(\tau)|) \frac{d\phi(x(\tau))}{d\tau} \, d\tau \\ &= -\frac{i}{2} \frac{1}{|x'(t)|} \int_0^{2\pi} \frac{\partial^2}{\partial t \partial \tau} H_0^{(+)}(k|x(t) - x(\tau)|) \phi(x(\tau)) \, d\tau. \end{aligned}$$

Sei nun

$$N(t, \tau) := \frac{\partial^2}{\partial t \partial \tau} \left\{ \frac{i}{2} H_0^{(+)}(k|x(t) - x(\tau)|) + \frac{1}{2\pi} \log \left( 4 \sin^2 \frac{t - \tau}{2} \right) \right\}, \quad (4.28)$$

dann folgt mit erneuter partieller Integration

$$\begin{aligned} \left( \frac{d}{ds} S \frac{d\phi}{ds} \right) (x(t)) &= -\frac{1}{|x'(t)|} \int_0^{2\pi} \left\{ N(t, \tau) - \frac{\partial^2}{\partial t \partial \tau} \left( \frac{1}{2\pi} \log \left( 4 \sin^2 \frac{t-\tau}{2} \right) \right) \right\} \phi(x(\tau)) d\tau \\ &= -\frac{1}{|x'(t)|} \int_0^{2\pi} \left\{ -\frac{1}{2\pi} \cot \frac{t-\tau}{2} \cdot \frac{d\phi(x(\tau))}{d\tau} + N(t, \tau) \phi(x(\tau)) \right\} d\tau. \end{aligned}$$

Im Anhang wird gezeigt, daß auch der Kern  $N(t, \tau)$  eine ähnliche Aufspaltung besitzt wie die vorher besprochenen Kerne. Es gilt nämlich

$$N(t, \tau) = N_1(t, \tau) \log \left( 4 \sin^2 \frac{t-\tau}{2} \right) + N_2(t, \tau),$$

mit den analytischen Funktionen

$$N_1(t, \tau) = -\frac{1}{2\pi} \frac{\partial^2}{\partial t \partial \tau} J_0(k|x(t) - x(\tau)|)$$

und

$$N_2(t, \tau) = N(t, \tau) - N_1(t, \tau) \log \left( 4 \sin^2 \frac{t-\tau}{2} \right).$$

Die Diagonalterme ergeben sich zu

$$N_1(t, t) = -\frac{k^2|x'(t)|^2}{4\pi}$$

und

$$\begin{aligned} N_2(t, t) &= \left( \pi i - 1 - 2C - 2 \log \frac{k|x'(t)|^2}{2} \right) \frac{k^2|x'(t)|^2}{4\pi} \\ &\quad + \frac{1}{12\pi} + \frac{(x'(t)x''(t))^2}{2\pi|x'(t)|^4} - \frac{|x''(t)|^2}{4\pi|x'(t)|^2} - \frac{x'(t)x'''(t)}{6\pi|x'(t)|^2}. \end{aligned}$$

Bei  $N_2(t, t)$  ist zu beachten, daß die hinteren Terme nicht von  $k$  abhängen, so daß sie sich später bei der Berechnung des modifizierten Kerns  $\tilde{N}_2$  wegheben werden.

Berücksichtigt man die Formeln (vgl. [AS70])

$$\begin{aligned} J_{n+1}(z) &= \left( \frac{2n}{z} \right) J_n(z) - J_{n-1}(z) & H_{n+1}^{(+)}(z) &= \left( \frac{2n}{z} \right) H_n^{(+)}(z) - H_{n-1}^{(+)}(z) \\ J_n'(z) &= -J_{n+1}(z) + \frac{n \cdot J_n(z)}{z} & H_n^{(+)'}(z) &= -H_{n+1}^{(+)}(z) + \frac{n \cdot H_n^{(+)}(z)}{z}, \end{aligned}$$

für  $n = 0, 1, \dots$ , so folgt mit einer kleinen Rechnung

$$\begin{aligned}
N_1(t, \tau) &= -\frac{1}{2\pi} \frac{\partial^2}{\partial t \partial \tau} J_0(k|x(t) - x(\tau)|) \\
&= -\frac{1}{2\pi} \frac{\partial}{\partial t} \left\{ k \frac{(x(t) - x(\tau))x'(\tau)}{|x(t) - x(\tau)|} J_1(k|x(t) - x(\tau)|) \right\} \\
&= -\frac{1}{2\pi} \left\{ k \frac{(x(t) - x(\tau))x'(\tau)}{|x(t) - x(\tau)|} k \frac{(x(t) - x(\tau))x'(t)}{|x(t) - x(\tau)|} J_1'(k|x(t) - x(\tau)|) \right. \\
&\quad \left. + k J_1(k|x(t) - x(\tau)|) \left\{ -\frac{(x(t) - x(\tau))x'(t)}{|x(t) - x(\tau)|^3} \cdot (x(t) - x(\tau))x'(\tau) + \frac{x'(t)x'(\tau)}{|x(t) - x(\tau)|} \right\} \right\} \\
&= -\frac{1}{2\pi} \bar{N}(t, \tau) \left\{ k^2 J_1'(k|x(t) - x(\tau)|) - \frac{k J_1(k|x(t) - x(\tau)|)}{|x(t) - x(\tau)|} \right\} \\
&\quad - \frac{k x'(t)x'(\tau)}{2\pi |x(t) - x(\tau)|} J_1(k|x(t) - x(\tau)|) \\
&= \frac{1}{2\pi} \bar{N}(t, \tau) k^2 J_2(k|x(t) - x(\tau)|) - \frac{k x'(t)x'(\tau)}{2\pi |x(t) - x(\tau)|} J_1(k|x(t) - x(\tau)|) \\
&= -\frac{1}{2\pi} \bar{N}(t, \tau) \left\{ k^2 J_0(k|x(t) - x(\tau)|) - \frac{2k J_1(k|x(t) - x(\tau)|)}{|x(t) - x(\tau)|} \right\} \\
&\quad - \frac{k x'(t)x'(\tau)}{2\pi |x(t) - x(\tau)|} J_1(k|x(t) - x(\tau)|),
\end{aligned}$$

wobei zur Abkürzung

$$\bar{N}(t, \tau) := \frac{(x(t) - x(\tau))x'(t)(x(t) - x(\tau))x'(\tau)}{|x(t) - x(\tau)|^2}$$

gesetzt wurde. Vertauscht man nun die Besselfunktionen durch die Hankelfunktionen und beachtet

$$\frac{\partial^2}{\partial t \partial \tau} \log \left( 4 \sin^2 \frac{t - \tau}{2} \right) = \frac{1}{2} \frac{1}{\sin^2 \frac{t - \tau}{2}},$$

so erhält man in analoger Weise

$$\begin{aligned}
N(t, \tau) &= \frac{i}{2} \bar{N}(t, \tau) \left\{ k^2 H_0^{(+)}(k|x(t) - x(\tau)|) - \frac{2k H_1^{(+)}(k|x(t) - x(\tau)|)}{|x(t) - x(\tau)|} \right\} \\
&\quad + \frac{i}{2} \frac{k x'(t)x'(\tau)}{|x(t) - x(\tau)|} H_1^{(+)}(k|x(t) - x(\tau)|) + \frac{1}{4\pi} \frac{1}{\sin^2 \frac{t - \tau}{2}}
\end{aligned}$$

Durch Ersetzen von  $k$  durch  $k_0$  in den obigen Gleichungen gelangt man schließlich zu einer Darstellung des Operators  $\left( \frac{d}{ds} S_0 \frac{d\phi}{ds} \right) (x(t))$ .

Durch Subtraktion wird dann eine Darstellung des Operators  $\left( \frac{d}{ds} (S - S_0) \frac{d\phi}{ds} \right) (x(t))$  gefunden.

Sei  $\tilde{N}(t, \tau)$  der entsprechende Integralkern und

$$\tilde{N}(t, \tau) = \tilde{N}_1(t, \tau) \log \left( 4 \sin^2 \frac{t - \tau}{2} \right) + \tilde{N}_2(t, \tau)$$

die Aufspaltung mit Diagonaltermen

$$\tilde{N}_1(t, t) = \frac{|x'(t)|^2(k_0^2 - k^2)}{4\pi}$$

und

$$\tilde{N}_2(t, t) = \frac{|x'(t)|^2}{4\pi} \left\{ (k^2 - k_0^2)(\pi i - 1 - 2C) - 2k^2 \log \frac{k|x'(t)|^2}{2} + 2k_0^2 \log \frac{k_0|x'(t)|^2}{2} \right\}.$$

Auf eine explizite Darstellung der Kerne wird an dieser Stelle verzichtet, es muß lediglich die Differenz der  $k$ -Terme und der  $k_0$ -Terme gebildet werden.

Nun kann endlich die gesuchte Darstellung des Operators  $T - T_0$  angegeben werden. Zusammensetzen der bisherigen Ergebnisse liefert

$$\begin{aligned} (T - T_0)(\phi)(x(t)) &= \left( \frac{d}{ds}(S - S_0) \frac{d\phi}{ds} \right) (x(t)) + \nu (k^2 S(\nu\phi) - k_0^2 S_0(\nu\phi)) (x(t)) \\ &= \nu (k^2 S(\nu\phi) - k_0^2 S_0(\nu\phi)) (x(t)) - \frac{1}{|x'(t)|} \int_0^{2\pi} \tilde{N}(t, \tau) \phi(x(\tau)) d\tau. \end{aligned}$$

Mit den obigen Ergebnissen für den Operator  $S$  läßt sich nun eine Zerlegung des Ausdrucks  $(k^2 S(\nu\phi) - k_0^2 S_0(\nu\phi))$  angeben. Sei hierzu der Integralkern

$$\tilde{L}(t, \tau) = \frac{i}{2} \left( k^2 H_0^{(+)}(k|x(t) - x(\tau)|) - k_0^2 H_0^{(+)}(k_0|x(t) - x(\tau)|) \right)$$

definiert. Analog zu oben gilt

$$\tilde{L}_1(t, \tau) = -\frac{1}{2\pi} \left( k^2 J_0(k|x(t) - x(\tau)|) - k_0^2 J_0(k_0|x(t) - x(\tau)|) \right)$$

bzw.

$$\tilde{L}_2(t, \tau) = \tilde{L}(t, \tau) - \tilde{L}_1(t, \tau) \log \left( 4 \sin^2 \frac{t - \tau}{2} \right).$$

Für die Diagonalterme ergibt sich schließlich

$$\tilde{L}_1(t, t) = \frac{k_0^2 - k^2}{2\pi}$$

und

$$\tilde{L}_2(t, t) = \left( \frac{i}{2} - \frac{C}{\pi} - \frac{1}{\pi} \log \frac{|x'(t)|}{2} \right) (k^2 - k_0^2) - \frac{1}{\pi} (k^2 \log k - k_0^2 \log k_0).$$

Nun kann endlich die Diskretisierung vorgenommen werden. Hierzu wird ein geeignetes Quadraturverfahren zur Berechnung der singulären Integralteile entwickelt. Durch die obige Parametrisierung können die gesuchten Belegungen  $\phi, \psi \in \mathcal{C}(\partial\Omega)$  mittels

$$\tilde{\phi}(t) := \phi(x(t)) \quad \text{und} \quad \tilde{\psi}(t) := \psi(x(t))$$

als  $2\pi$ -periodische Funktion aufgefaßt werden. Es liegt somit nahe, die Quadratur auf trigonometrische Interpolation zu begründen. Hierzu wird das Intervall  $[0, 2\pi]$  äquidistant in  $2N$  Teilintervalle aufgeteilt. Seien

$$t_j := \frac{j\pi}{N}, \quad j = 0, \dots, 2N - 1$$

die dazu gehörigen Knotenpunkte. Bezeichne  $T_N$  den  $2N$ -dimensionalen Raum der trigonometrischen Polynome  $N$ -ter Ordnung, also Funktionen der Form

$$v(t) = \sum_{m=0}^N a_m \cos mt + \sum_{m=1}^{N-1} b_m \sin mt.$$

Der Operator  $P_N : \mathcal{C}[0, 2\pi] \rightarrow T_N$  sei der dazugehörige Interpolationsoperator. Man verwendet

$$\int_0^{2\pi} \log \left( 4 \sin^2 \frac{t-\tau}{2} \right) f(\tau) d\tau \simeq \int_0^{2\pi} \log \left( 4 \sin^2 \frac{t-\tau}{2} \right) (P_N f)(\tau) d\tau$$

als Quadraturregel für die singulären Integralteile der jeweiligen Operatoren.

Die Lagrange-Basis  $L_0(t), \dots, L_{2N-1}(t)$  mit  $L_k(t_j) = \delta_{jk}$  für die trigonometrische Interpolation ist gegeben durch

$$L_j(t) := \frac{1}{2N} \left\{ 1 + 2 \sum_{m=1}^{N-1} \cos m(t - t_j) + \cos N(t - t_j) \right\}.$$

Somit folgt zunächst

$$(P_N f)(\tau) = \sum_{j=0}^{2N-1} f(t_j) L_j(\tau)$$

und weiter

$$\int_0^{2\pi} \log \left( 4 \sin^2 \frac{t-\tau}{2} \right) f(\tau) d\tau \simeq \sum_{j=0}^{2N-1} R_j(t) f(t_j), \quad (4.29)$$

mit

$$R_j(t) = \int_0^{2\pi} \log \left( 4 \sin^2 \frac{t-\tau}{2} \right) L_j(\tau) d\tau.$$

**4.7 Lemma** ([Kre99, Seite 208]). *Es gilt*

(i)

$$R_j(t) = -\frac{2\pi}{N} \sum_{m=1}^{N-1} \frac{1}{m} \cos m(t - t_j) - \frac{\pi}{N^2} \cos N(t - t_j).$$

(ii)

$$R_j(t_k) = R_k(t_j) =: R_{|k-j|}, \quad j, k = 0, \dots, 2N-1$$

mit

$$R_j := -\frac{2\pi}{N} \sum_{m=1}^{N-1} \frac{1}{m} \cos \frac{mj\pi}{N} - \frac{(-1)^j \pi}{N^2}, \quad j = 0, \dots, 2N-1.$$

Für den analytischen Teil des Integralkerns wird die Trapezsumme

$$\int_0^{2\pi} f(t) dt \simeq \int_0^{2\pi} (P_N f)(t) dt = \frac{\pi}{N} \sum_{j=0}^{2N-1} f(t_j) \quad (4.30)$$

verwendet.

**Diskretisierung von  $A(k)$  im Falle (HP)**

Die Gleichungen (4.29) und (4.30) liefern die Näherung

$$\begin{aligned} (K\phi)(x(t)) &= \int_0^{2\pi} H(t, \tau) \tilde{\phi}(\tau) d\tau \\ &\simeq \sum_{l=0}^{2N-1} \left\{ R_l(t) H_1(t, t_l) + \frac{\pi}{N} H_2(t, t_l) \right\} \tilde{\phi}(t_l). \end{aligned} \quad (4.31)$$

Die Approximation der Integralgleichung  $A(k)(\phi)(x(t)) = 0$  schreibt sich somit als

$$\tilde{\phi}(t) - \sum_{l=0}^{2N-1} \left\{ R_l(t) H_1(t, t_l) + \frac{\pi}{N} H_2(t, t_l) \right\} \tilde{\phi}(t_l) = 0. \quad (4.32)$$

Setzt man zur Abkürzung  $\tilde{\phi}_j := \tilde{\phi}(t_j)$  für die unbekanntenen Funktionswerte der Belegung  $\tilde{\phi}$  an den Knotenstellen  $t_j$  und verwendet Lemma (4.7, (ii)), so sieht man, daß Gleichung (4.32) auf das lineare Gleichungssystem

$$\tilde{\phi}_j - \sum_{l=0}^{2N-1} \left\{ R_{|j-l|} H_1(t_j, t_l) + \frac{\pi}{N} H_2(t_j, t_l) \right\} \tilde{\phi}_l = 0, \quad j = 0, \dots, 2N-1$$

führt.

Definiert man die Matrix  $A = (A_{jl}) \in \mathbb{R}^{2N \times 2N}$  durch

$$A_{jl} := R_{|j-l|} H_1(t_j, t_l) + \frac{\pi}{N} H_2(t_j, t_l),$$

so läßt sich die Integralgleichung  $A(k)(\phi)(x(t)) = 0$  durch die lineare, endlich-dimensionale Gleichung

$$(I - A) \begin{pmatrix} \tilde{\phi}_0 \\ \vdots \\ \tilde{\phi}_{2N-1} \end{pmatrix} = 0$$

diskretisieren.

Die Matrix

$$\tilde{A}^N(k) := I - A$$

ist somit die Diskretisierung des Operators  $A(k)$  für den Fall (HP).

Zur Vereinfachung der Notation wird sowohl im diskreten Fall  $\tilde{\phi} = (\tilde{\phi}_0, \dots, \tilde{\phi}_{2N-1})$  als auch im kontinuierlichen Fall  $\tilde{\phi}$  als Belegung verwendet. Die Matrix  $\tilde{A}^N(k)$  hängt selbstverständlich von dem Parameter  $k$  ab, der bisher zur Vereinfachung der Notation unterdrückt wurde. Man schreibt einfach  $\tilde{A}(k)$ , falls der Diskretisierungsparameter  $N$  aus dem Zusammenhang ersichtlich ist.

#### Diskretisierung von $A(k)$ im Falle (TP)

Die gesuchten Funktionswerte der Belegungen  $\tilde{\phi}$  und  $\tilde{\psi}$  an den Stellen  $x(t_j)$  werden mit  $\tilde{\phi}_j := \phi(x(t_j))$  und  $\tilde{\psi}_j := \psi(x(t_j))$  bezeichnet. Zusätzlich wird noch  $\gamma_j := |x'(t_j)|$  eingeführt.

Mit den oben eingeführten Aufspaltungen wird die erste Gleichung von (4.18) zu

$$\begin{aligned} 2\psi(x(t)) &+ \int_0^{2\pi} \left\{ \log \left( 4 \sin^2 \frac{t-\tau}{2} \right) \tilde{H}_1(t, \tau) + \tilde{H}_2(t, \tau) \right\} \psi(x(\tau)) d\tau \\ &+ \int_0^{2\pi} \left\{ \log \left( 4 \sin^2 \frac{t-\tau}{2} \right) \tilde{M}_1(t, \tau) + \tilde{M}_2(t, \tau) \right\} |x'(\tau)| \phi(x(\tau)) d\tau = 0. \end{aligned}$$

Analog zum Fall (HP) gelangt man mittels Gleichungen (4.29) und (4.30) für den Fall (TP) zu einer endlich-dimensionalen Gleichung für die unbekanntenen  $\tilde{\phi}_j$  und  $\tilde{\psi}_j$

$$\begin{aligned} 2\tilde{\psi}_j &+ \sum_{l=0}^{2N-1} \left\{ R_{|j-l|} \tilde{H}_1(t_j, t_l) + \frac{\pi}{N} \tilde{H}_2(t_j, t_l) \right\} \tilde{\psi}_l \\ &+ \sum_{l=0}^{2N-1} \left\{ R_{|j-l|} \tilde{M}_1(t_j, t_l) + \frac{\pi}{N} \tilde{M}_2(t_j, t_l) \right\} \gamma_l \tilde{\phi}_l = 0, \quad j = 0, \dots, 2N-1. \end{aligned} \quad (4.33)$$

Die zweite Gleichung von (4.18) führt zunächst zu

$$\begin{aligned}
2\phi(x(t)) &- \frac{x'(t)^\perp}{|x'(t)|} \int_0^{2\pi} \left\{ \log \left( 4 \sin^2 \frac{t-\tau}{2} \right) \tilde{L}_1(t, \tau) + \tilde{L}_2(t, \tau) \right\} |x'(\tau)| \frac{x'(\tau)^\perp}{|x'(\tau)|} \psi(x(\tau)) d\tau \\
&+ \frac{1}{|x'(t)|} \int_0^{2\pi} \left\{ \log \left( 4 \sin^2 \frac{t-\tau}{2} \right) \tilde{N}_1(t, \tau) + \tilde{N}_2(t, \tau) \right\} \psi(x(\tau)) d\tau \\
&- \frac{1}{|x'(t)|} \int_0^{2\pi} \left\{ \log \left( 4 \sin^2 \frac{t-\tau}{2} \right) \tilde{H}_1^*(t, \tau) + \tilde{H}_2^*(t, \tau) \right\} \phi(x(\tau)) d\tau = 0.
\end{aligned}$$

Mit den obigen Quadraturformeln erhält man schließlich das endlich-dimensionale System

$$\begin{aligned}
2\tilde{\phi}_j \gamma_j &- \sum_{l=0}^{2N-1} \left\{ R_{|j-l|} \left( \tilde{L}_1(t_j, t_l) x'_j x'_l - \tilde{N}_1(t_j, t_l) \right) + \frac{\pi}{N} \left( \tilde{L}_2(t_j, t_l) x'_j x'_l - \tilde{N}_2(t_j, t_l) \right) \right\} \tilde{\psi}_l \\
&- \sum_{l=0}^{2N-1} \left\{ R_{|j-l|} \tilde{H}_1^*(t_j, t_l) + \frac{\pi}{N} \tilde{H}_2^*(t_j, t_l) \right\} \tilde{\phi}_l = 0, \quad j = 0, \dots, 2N-1. \quad (4.34)
\end{aligned}$$

Mit den durch

$$\begin{aligned}
(A_{11})_{jl} &:= - \left\{ R_{|j-l|} \tilde{H}_1(t_j, t_l) + \frac{\pi}{N} \tilde{H}_2(t_j, t_l) \right\} \\
(A_{12})_{jl} &:= - \left\{ R_{|j-l|} \tilde{M}_1(t_j, t_l) \gamma_l + \frac{\pi}{N} \tilde{M}_2(t_j, t_l) \gamma_l \right\} \\
(A_{21})_{jl} &:= \left\{ R_{|j-l|} \left( \tilde{L}_1(t_j, t_l) x'_j x'_l - \tilde{N}_1(t_j, t_l) \right) + \frac{\pi}{N} \left( \tilde{L}_2(t_j, t_l) x'_j x'_l - \tilde{N}_2(t_j, t_l) \right) \right\} \\
(A_{22})_{jl} &:= \left\{ R_{|j-l|} \tilde{H}_1^*(t_j, t_l) + \frac{\pi}{N} \tilde{H}_2^*(t_j, t_l) \right\},
\end{aligned}$$

definierten Matrizen  $A_{11}, A_{12}, A_{21}, A_{22} \in \mathbb{C}^{2N \times 2N}$  wird das lineare Gleichungssystem (4.33) und (4.34) kompakt in Matrixschreibweise umformuliert zu

$$\left( 2 \cdot \begin{pmatrix} I_{2N \times 2N} & & 0 \\ & \gamma_1 & \\ 0 & & \ddots \\ & & & \gamma_{2N-1} \end{pmatrix} - \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \right) \cdot \begin{pmatrix} \tilde{\psi}_0 \\ \vdots \\ \tilde{\psi}_{2N-1} \\ \tilde{\phi}_0 \\ \vdots \\ \tilde{\phi}_{2N-1} \end{pmatrix} = 0. \quad (4.35)$$

Die Matrix auf der linken Seite ist schließlich die gesuchte Diskretisierung des Operators  $A(k)$ . Sie wird im folgenden stets mit  $\tilde{A}^N(k)$  bezeichnet.

### 4.3 Der direkte Sweep-Algorithmus

Die Eigenresonanzen können nun approximiert werden, indem man  $k$  so bestimmt, daß die Matrix  $\tilde{A}(k)$  singulär wird.

Das Verschwinden der Determinante von  $\tilde{A}(k)$  ist äquivalent zur Singularität von  $\tilde{A}(k)$ . In Folge der Diskretisierung wird die Determinante zwar klein, aber niemals Null sein für Werte von  $k$ , die hinreichend nahe an Eigenwerten des Laplace-Operators liegen. Überdies ist die numerische Berechnung von Determinanten instabil, weil unterschiedliche Skalierungen der Matrix unterschiedliche Skalierungen im Werte der Determinante bewirken ([Sey94]). Statt der Determinante wird besser die Singulärwertzerlegung der Matrix  $\tilde{A}(k)$  berechnet. Allgemein ist diese für eine Matrix  $A$  gegeben durch unitäre Matrizen  $U$  und  $V$  und eine Diagonalmatrix  $S$ , so daß

$$A = USV^*$$

gilt ([Gv96]). Die Matrix  $S$  enthält die Singulärwerte der Matrix  $A$ .

Allgemeiner kann die Singularität einer Matrix mit Hilfe sogenannter Testfunktionen festgestellt werden. Dazu definiert man

**4.8 Definition.** Eine Abbildung  $\tau : \mathbb{C}^{n \times n} \rightarrow \mathbb{C}$  mit der Eigenschaft

$$\tau(A) = 0 \quad \Leftrightarrow \quad A \text{ ist singulär}$$

heißt *Testfunktion*.

Eine Familie von Testfunktionen erhält man zum Beispiel durch eine Idee von Seydel ([Sey94]). Im Falle  $A = \tilde{A}(k)$  dient der kleinste Singulärwert als Maß für die Singularität der Matrix  $\tilde{A}(k)$ .

Bei den in der Literatur bisher beschriebenen Verfahren erfolgt die Eigenwertbestimmung (für den (HP)-Fall) in einem vorgegebenen Intervall mittels eines sogenannten *Sweep-Algorithmus*, der das Intervall nach Minima von geeigneten Testfunktionen untersucht ([Bar00], [Bäc02]). Hierzu wird die Testfunktion an bestimmten Stellen (z.B. unterteilt man das Intervall äquidistant) ausgewertet und man erhält durch Interpolation Näherungen für die Minima. Dieser Ansatz hat jedoch gravierende Nachteile:

- (i) Er bedingt einen hohen Rechenaufwand durch die aufwendige Auswertung der Testfunktion.
- (ii) Durch die äquidistante Unterteilung des Intervalls werden nahe bei einanderliegende Eigenwerte übersehen. Tests ergeben, daß ca. 10 Prozent der Eigenwerte übersehen werden ([BRH98]). Eine feinere Auflösung des Intervalls erhöht im Gegenzug wiederum die Rechenzeit.

- (iii) Er ist unbrauchbar für das Transmissions-Problem (TP), wo die gesuchten Eigenresonanzen nicht reell sind. Für (TP) wurden bisher noch keine effizienten Sweep-Algorithmen beschrieben.
- (iv) Es sind keine Fehlerabschätzungen für die berechneten Eigenwerte möglich.

Im folgenden wird ein neuer *direkter* Sweep-Algorithmus vorgeschlagen, der diese Nachteile größtenteils vermeidet und gänzlich ohne die Auswertung von Testfunktionen arbeitet.

Der neue Algorithmus beruht auf einer *Prädiktor - Korrektor*-Strategie.

Für die folgende Darstellung wird eine Anzahl von  $2N$  diskreten Punkte auf dem Rand fest gewählt. Gesucht sind also nicht-triviale Lösungen der Gleichung

$$\tilde{A}(k)x = 0 \quad (4.36)$$

mit  $x \neq 0$ . Im Fall (HP) ist  $x = \tilde{\phi}$  im Fall (TP) ist  $x = \begin{pmatrix} \tilde{\psi} \\ \tilde{\phi} \end{pmatrix}$ . Die Behandlung dieses nichtlinearen Eigenwertproblems geht von der Analyse eines von A. Neumaier ([Neu85]) vorgeschlagenen Lösungsverfahrens aus.

Sei  $k^*$  ein Eigenwert und  $x^*$  der dazugehörige Eigenvektor. Der Prädiktor liefert eine Näherung  $k_0$  für  $k^*$  und eine Näherung  $x_0$  für  $x^*$ . Der Einfachheit halber wird zunächst angenommen, daß der Nullraum der Matrix  $\tilde{A}(k^*)$  eindimensional ist.

Der Korrektor erzeugt eine Folge von Paaren  $k_j$  und  $x_j$ , die gegen  $k^*$  und  $x^*$  konvergieren. Die Anzahl der benötigten Korrektorschritte zum Erreichen einer geforderten Genauigkeit dient als Maß für die Güte der Approximation des Prädiktors.

Der Korrektor soll eine nicht-triviale Lösung der Gleichung

$$\tilde{A}(k^*)x^* = 0$$

bestimmen und zusätzlich eine geeignete Normierung des Lösungsvektors  $x^*$  liefern. Bei Verwendung der Maximumsnorm wird dies durch die zusätzliche Bedingung

$$e^T x^* - 1 = 0$$

erreicht, wobei  $e$  ein geeigneter Einheitsvektor ist. Mit der durch

$$F(k, x) := \begin{pmatrix} \tilde{A}(k)x \\ e^T x - 1 \end{pmatrix}$$

definierten Funktion muß die Gleichung

$$F(k^*, x^*) = 0 \quad (4.37)$$

gelöst werden. Wenn der Prädiktor eine hinreichend gute Startnäherung liefert, scheint sich nun das Newton-Verfahren zur Lösung der Gleichung (4.37) anzubieten. Dazu müßte in jedem Newton-Schritt eine  $LU$ -Zerlegung der vollbesetzten Matrix  $\tilde{A}(k_j)$  durchgeführt werden. In der obigen Notation ergäbe sich folgender Newton-Algorithmus

$$\begin{aligned}
 (NA) \quad y_{j+1} &:= \tilde{A}(k_j)^{-1} \tilde{A}'(k_j) x_j, \\
 x_{j+1} &:= \frac{y_{j+1}}{e^T y_{j+1}} \\
 k_{j+1} &:= k_j - \frac{1}{e^T y_{j+1}}
 \end{aligned}$$

Die Idee, die für jeden Schritt notwendige  $LU$ -Zerlegung dadurch zu umgehen, daß man die Matrizen  $\tilde{A}(k_j)$  durch einen festen Shift  $\tilde{A}(\sigma)$  ersetzt, führt nicht zum Ziel, weil nämlich dann das Newton-Verfahren gegen die Lösung des falschen Problems  $\tilde{A}(k)x = (k - \sigma)\tilde{A}'(k)x$  konvergiert.

Daher schlägt A. Neumaier eine Abwandlung des Newton-Verfahrens vor, bei der die  $LU$ -Zerlegung wie folgt umgangen wird. Es bezeichne  $dx_j$  das Residuum im  $j - 1$ -ten Schritt, also

$$x_{j+1} = x_j - dx_j.$$

Dann gilt unter der Voraussetzung, daß  $\tilde{A}(k)$  zweimal stetig differenzierbar ist

$$\begin{aligned}
 dx_j = x_j - x_{j+1} &= x_j + (k_{j+1} - k_j)y_{j+1} \\
 &= x_j + (k_{j+1} - k_j)\tilde{A}(k_j)^{-1}\tilde{A}'(k_j)x_j \\
 &= \tilde{A}^{-1}(k_j) \left( \tilde{A}(k_j) + (k_{j+1} - k_j)\tilde{A}'(k_j) \right) x_j \\
 &= \tilde{A}^{-1}(k_j)\tilde{A}(k_{j+1})x_j + O(|k_{j+1} - k_j|^2).
 \end{aligned}$$

Vernachlässigt man nun Terme zweiter Ordnung, so erhält man

$$x_{j+1} = x_j - \tilde{A}(k_j)^{-1}\tilde{A}(k_{j+1})x_j \quad (4.38)$$

als Iterationsvorschrift, die die neue Approximation  $k_{j+1}$  enthält. Diese Näherung muß zuerst bestimmt werden. Es läßt sich zeigen ([Neu85, Theorem 1]), daß man in Gleichung (4.38) die Matrix  $\tilde{A}(k_j)$  durch eine feste Matrix  $\tilde{A}(\sigma)$  ersetzen darf, für die das so modifizierte Newton-Verfahren gegen eine Lösung des richtigen Problems konvergiert.

Damit ergibt sich der folgende Algorithmus von A. Neumaier

**4.9 Algorithmus (Residual Inverse Iteration).**

- 1 Setze  $l = 0$ .
- 2  $e$  sei der Einsvektor mit der Eins in der Komponente, an der  $\|x_0\|_\infty$  das Maximum annimmt.
- 3 Bestimme eine Approximation für  $k_{l+1}$  mit Hilfe eines Newton-Schrittes angewandt auf die Gleichung

$$e^T \tilde{A}(k_0)^{-1} \tilde{A}(k_{l+1}) x_l = 0.$$

- 4 Berechne das Residuum

$$r_l := \tilde{A}(k_{l+1}) x_l.$$

- 5 Berechne eine verbesserte Approximation  $x_{l+1}$  von  $x$  durch Lösen der Gleichung

$$\tilde{A}(k_0) dx_l = r_l$$

und

$$\bar{x}_{l+1} = x_l - dx_l, \quad x_{l+1} = \frac{\bar{x}_{l+1}}{e^T \bar{x}_{l+1}}.$$

- 6 Erhöhe  $l$  um Eins und gehe zu 3.

Speziell im linearen Fall  $\tilde{A}(k) = A - kI$  folgt aus der Update-Gleichung in Schritt 5 und der Residuungsgleichung aus Schritt 4

$$\begin{aligned} (A - k_0 I) \bar{x}_{l+1} &= (A - k_0 I) x_l - (A - k_0 I) dx_l \\ &= (A - k_0 I) x_l - (A - k_{l+1} I) x_l \\ &= (k_{l+1} - k_0) x_l. \end{aligned}$$

Der Vektor  $\bar{x}_{l+1}$  ist somit parallel zu  $(A - k_0 I)^{-1} x_l$ , und der Algorithmus erweist sich in diesem Fall als äquivalent zur bekannten inversen Iteration mit Shift  $k_0$ . Hinsichtlich der Konvergenzeigenschaften des Neumaier-Algorithmus gilt der

**4.10 Satz ([Neu85, Theorem 1]).** *Sei  $k^*$  eine einfache Nullstelle von  $\det \tilde{A}(k) = 0$  und  $x^*$  ein zugehöriger Eigenvektor mit Normierung  $e^T x^* = 1$ . Dann konvergiert der Algorithmus 4.9 für alle  $k_0$ , die hinreichend nahe bei  $k^*$  liegen, und es gilt*

$$\frac{\|x_{l+1} - x^*\|}{\|x_l - x^*\|} = O(|k_0 - k^*|), \quad |k_{l+1} - k^*| = O(\|x_l - x^*\|).$$

Die von Neumaier vorgeschlagene Version des Algorithmus könnte schon als Korrektor benutzt werden. Jedoch impliziert Satz (4.10) nur lineare Konvergenz, was im allgemeinen zu viele Iterationen erfordert.

Man könnte nun versuchen, in jedem Update-Schritt den Shift  $k_0$  durch den Shift  $k_l$  zu ersetzen und eine  $LU$ -Zerlegung in Kauf zu nehmen, um quadratische Konvergenz wie im Newton-Algorithmus (NA) zu erzwingen. Da dies den selben Aufwand wie das Newton-Verfahren hätte, bietet sich dem gegenüber an, einen Kompromiß zwischen speziell gewählten Update- und Iterationsschritten zu suchen. Ferner verbietet es sich, zu viele Update-Schritte durchzuführen, weil sich die Konditionierung der jeweiligen  $LU$ -Zerlegungen wegen der Singularität von  $\tilde{A}(k^*)$  verschlechtert (vgl. 4.4).

Ein praktisches Maß für die Konvergenzgeschwindigkeit des Korrektors wird von der Folge  $\frac{|k_{l+1}-k^*|}{|k_l-k^*|}$  geliefert. Es gibt keine verwendbaren a priori-Schranken, so daß die Werte a posteriori bestimmt werden müssen. Im folgenden wird die mittlere Konvergenzrate von  $\frac{|k_{l+1}-k^*|}{|k_l-k^*|}$  mit  $\beta$  bezeichnet.

Die Konvergenzrate wird durch den folgenden Algorithmus erheblich beschleunigt. Es wird sich zeigen, daß  $\beta$  durch einen einzigen Update-Schritt um einen Faktor bis zu  $10^{-3}$  verkleinert werden kann.

Der in Kapitel 3, Gleichung (3.4) eingeführte mittlere Abstand  $\Delta_k$  zweier aufeinanderfolgender Eigenwerte im Bereich um  $k$  dient als Skalierungsmaß. Gilt

$$|k_l - k_{l+1}| < \frac{1}{10} \Delta_k,$$

so wird  $k_0$  durch  $k_{l+1}$  ersetzt und der Algorithmus arbeitet mit diesem neuen Shift weiter.

#### 4.11 Algorithmus (Direkter Sweep-Korrektor).

- 1 Führe solange Residual Inverse Iteration mit Eingabe  $k_0$  und  $x_0$  aus, bis

$$|k_l - k_{l+1}| < \frac{1}{10} \Delta_k$$

erfüllt oder eine maximale Anzahl von Schritten übertroffen worden ist.

- 2 Starte mit dieser Approximation von  $k$  und  $x$  die Residual Inverse Iteration erneut, bis

$$|k_l - k_{l+1}| < \varepsilon,$$

gilt, wobei  $\varepsilon$  eine vorgegebene Fehlerschranke ist, oder eine maximal Anzahl von Schritten übertroffen worden ist.

Die Abbruchbedingung in Schritt 2 kann natürlich durch andere Bedingungen ersetzt werden. Es ist zum Beispiel eine relative Fehlerschranke denkbar oder eine, die den mittleren Abstand berücksichtigt, wie zum Beispiel  $|k_l - k_{l+1}| < \varepsilon \cdot \Delta_k$ .

Der Prädiktor wird aus einem einfachen Taylorreihen-Ansatz für  $\tilde{A}(k)$  hergeleitet. Es gilt

$$\tilde{A}(k + \delta k) = \tilde{A}(k) + \delta k \cdot \tilde{A}'(k) + O((\delta k)^2).$$

Sei nun  $\delta k$  so bestimmt, daß  $k + \delta k$  ein Eigenwert ist und sei  $\bar{x}$  der zu  $k + \delta k$  gehörige Eigenvektor. Dann gilt

$$\tilde{A}(k)\bar{x} + \delta k \cdot \tilde{A}'(k)\bar{x} + O((\delta k)^2\bar{x}) = 0.$$

Vernachlässigt man Terme höhere Ordnung, so erhält man das verallgemeinerte Eigenwertproblem

$$-\tilde{A}(k)\bar{x} = \delta k \cdot \tilde{A}'(k)\bar{x}.$$

Diese Idee motiviert den

**4.12 Satz (Direkter Sweep-Prädiktor).** *Es gelte  $\tilde{A}(k_0)x_0 = 0$  und  $\delta k$  und  $\bar{x}$  seien Lösungen des verallgemeinerten Eigenwertproblems*

$$-\tilde{A}(k_0)\bar{x} = \delta k \cdot \tilde{A}'(k_0)\bar{x}$$

mit  $|\delta k| \ll 1$ . Dann existiert ein Eigenwert  $k^*$  mit

$$|k^* - (k_0 + \delta k)| = O(|\delta k|^2).$$

**Beweis:** Sei  $\tilde{k} := k_0 + \delta k$ . Die Taylorentwicklung liefert

$$\tilde{A}(\tilde{k}) = \tilde{A}(k_0) + \delta k \tilde{A}'(k_0) + \frac{(\delta k)^2}{2} R(k_0, \delta k),$$

mit einem vom  $k_0$  und  $\delta k$  abhängigen Restglied  $R$ . Setzt man  $\mathcal{C}^2$ -Stetigkeit von  $\tilde{A}$  voraus, so bleibt dieser Ausdruck beschränkt. Sei  $y^*$  ein Linkseigenvektor von  $\tilde{A}(k^*)$ . Definiert man nun die Funktion

$$f(k) := y^* \tilde{A}(k) \bar{x},$$

so gilt  $f(k^*) = 0$ . Der Mittelwertsatz liefert dann die Existenz eines  $\kappa \in [k^*, \tilde{k}]$  mit

$$0 = f(\tilde{k}) + (k^* - \tilde{k})f'(\kappa).$$

Wegen

$$\begin{aligned} f(\tilde{k}) &= y^* \tilde{A}(\tilde{k}) \bar{x} = y^* \left\{ \tilde{A}(k_0) + \delta k \tilde{A}'(k_0) + \frac{(\delta k)^2}{2} R(k_0, \delta k) \right\} \bar{x} \\ &= \frac{(\delta k)^2}{2} y^* R(k_0, \delta k) \bar{x} \end{aligned}$$

folgt somit

$$k^* - \tilde{k} = -\frac{(\delta k)^2}{2f'(\kappa)} y^* R(k_0, \delta k) \bar{x} \quad (4.39)$$

und daraus die Behauptung.  $\square$

Die Konstante  $\alpha$ , die sich in dem  $O(|\delta k|^2)$ -Term versteckt, ist für die Praxis relevant. Sie darf nicht zu groß werden. Für die Konvergenzanalyse im folgenden Abschnitt definiert man

**4.13 Definition.** Die Zahl

$$\alpha := \frac{|k^* - \tilde{k}|}{|k_0 - \tilde{k}|^2}$$

heißt *Approximationsgüte des Prädiktors*.

Aus Satz (4.12) und Algorithmus (4.11) läßt sich ein Algorithmus zur Berechnung der Nachbargenwerte und Eigenvektoren zu einem gegebenem Paar  $k$  und  $x$  herleiten.

**4.14 Algorithmus (Direkter Sweep-Algorithmus).**

- 1 Berechne die Lösung des verallgemeinerten Eigenwertproblems

$$-\tilde{A}(k)\bar{x} = \delta k \cdot \tilde{A}'(k)\bar{x}.$$

Sei  $0 \leq |\delta k_1| \leq |\delta k_2| \leq \dots$ , wobei  $\delta k_j$  die Eigenwerte sind und  $x_1, x_2, \dots$  die zugehörigen Eigenvektoren.

- 2 Setze  $j = 1$ .
- 3 Das Paar  $k + \delta k_j$  und  $x_j$  ist eine Näherung für den  $j$ -ten Nachbarn von  $k$  und  $x$ . Korrigiere diese Näherung mit Algorithmus (4.11).
- 4 Falls der Korrektor weniger als 6 Schritte benötigt hat, erhöhe  $j$  um Eins und gehe zu 3.

Numerische Beispiele für (HP) wurden schon in Kapitel 3 angegeben. Für (TP) werden einige Beispiele in Kapitel 5 vorgestellt.

## 4.4 Konvergenz und Fehlerabschätzungen des Verfahrens

Die Aussagen dieses Abschnittes beziehen sich auf das Helmholtz-Problem, um die Notation übersichtlich zu halten, da bei der Diskretisierung dieses Problems nur ein Integraloperator verwendet werden muß. Sie können aber leicht auf das Transmission-Problem verallgemeinert werden, da die Diskretisierungsmatrix des Problem (TP) dieselben Eigenschaften besitzt wie die

Diskretisierungsmatrix des Problems (HP), die für die Konvergenzaussagen benötigt werden. Um dennoch einen Einblick von der Effizienz des Verfahrens zu geben, werden am Ende des Abschnittes einige numerische Resultate angegeben, die die Konvergenzaussagen auch für das Transmissions-Problem belegen.

Im vorherigen Abschnitt wurde gezeigt, wie mit dem direkten Sweep-Algorithmus zu einer vorgegebenen Diskretisierung  $N$  nicht-triviale Näherungen  $k^N$  und  $\tilde{\phi}^N$  für die Lösungen  $k^*$  und  $\tilde{\phi}^*$  der Operatorgleichung  $A(k)(\phi)(x(t)) = 0$  gewonnen werden.

In diesem Abschnitt werden folgende Konvergenz-Aspekte untersucht:

- Konvergenz der Folge  $(\tilde{\phi}^N, k^N)$  für  $N \rightarrow \infty$ ,
- eng damit verbunden ist die Frage, wie groß  $N$  zu wählen ist, um eine vorgegebene Genauigkeit von  $k$  zu gewährleisten,
- Konditionierung der Diskretisierungsmatrix in der Nähe des gesuchten Eigenwertes,
- Güte  $\alpha$  des Prädiktors,
- Konvergenzrate  $\beta$  des Korrektors.

### Konvergenz der Folge $(\tilde{\phi}^N, k^N)$ für $N \rightarrow \infty$

Für die Konvergenzanalyse wird das oben beschriebene numerische Schema zur Approximation der Operatorgleichung  $(I - K)(\phi)(x(t)) = 0$  genauer untersucht. Die auftretenden Operatoren sind alle von  $k$  abhängig, jedoch wurde diese Abhängigkeit in der Notation unterdrückt, um die Darstellung übersichtlicher zu gestalten. In diesem Abschnitt wird die Abhängigkeit von  $k$  jeweils explizit in die Notation aufgenommen.

Mit den beiden Quadraturoperatoren

$$(K_{1,N}\tilde{\phi})(t) := \int_0^{2\pi} \log\left(4 \sin^2 \frac{t-\tau}{2}\right) (P_N H_1(t, \tau)\tilde{\phi})(\tau) d\tau = \sum_{l=0}^{2N-1} R_l(t) H_1(t, t_l)\tilde{\phi}(t_l),$$

$$(K_{2,N}\tilde{\phi})(t) := \int_0^{2\pi} (P_N H_2(t, \tau)\tilde{\phi})(\tau) d\tau = \frac{\pi}{N} \sum_{l=0}^{2N-1} H_2(t, t_l)\tilde{\phi}(t_l)$$

schreibt sich die Approximation von  $(I - K)(\phi)(x(t)) = 0$  als

$$I\tilde{\phi} - K_{1,N}\tilde{\phi} - K_{2,N}\tilde{\phi} = 0. \quad (4.40)$$

Um eine endlich-dimensionale Diskretisierungsgleichung zu erhalten, wurde aus Gleichung (4.40) durch Kollokation mit dem Operator  $P_N$  das Diskretisierungsschema

$$I\tilde{\phi} - P_N K_{1,N}\tilde{\phi} - P_N K_{2,N}\tilde{\phi} = 0 \quad (4.41)$$

für  $\tilde{\phi} \in \mathcal{C}^{1,\alpha}[0, 2\pi]$  abgeleitet. Jede Lösung  $\tilde{\phi}^N$  dieser Gleichung gehört automatisch zu dem Raum  $T_N$  der trigonometrischen Polynome.

Sei  $\tilde{\phi}^*, k^*$  Lösung von (HP) und seien  $\tilde{\phi}^N, k^N$  Approximationen gemäß obigem Schema zur Diskretisierung  $N$ . Dann gelten die folgenden Gleichungen

$$(I - K(k^*))\tilde{\phi}^* = 0 \quad (4.42)$$

$$(I - (K_{1,N}(k^N) + K_{2,N}(k^N)))\tilde{\phi}^N = 0. \quad (4.43)$$

Sei nun  $X = \mathcal{C}[0, 2\pi] \times \mathbb{C}$  versehen mit der Norm  $\|\cdot\| = \|\tilde{\phi}\|_\infty + |k|$  für ein Element  $(\tilde{\phi}, k) \in X$ . Dann werden durch

$$\mathcal{K}(\tilde{\phi}, k) = \begin{pmatrix} K(k)\tilde{\phi} \\ \|\tilde{\phi}\|^2 \end{pmatrix} \quad (4.44)$$

$$\mathcal{K}_N(\tilde{\phi}, k) = \begin{pmatrix} (K_{1,N}(k) + K_{2,N}(k))\tilde{\phi} \\ \|\tilde{\phi}\|^2 \end{pmatrix} \quad (4.45)$$

(nichtlineare) Operatoren auf  $X$  definiert. Im folgenden wird gezeigt, daß eine konvergente Folge von Näherungslösungen  $(\tilde{\phi}^N, k^N)$  von (4.45) eine exakte Lösung  $(\tilde{\phi}^*, k^*)$  garantiert. Die Fréchet-Ableitung eines Operators  $K$  wird im folgenden mit  $DK$  bezeichnet und  $L(X)$  sei der Raum der linearen Operatoren auf  $X$ . Es gilt der Konvergenzsatz

#### 4.15 Satz. Gelte

$$\begin{aligned} (I - \mathcal{K})(\tilde{\phi}^*, k^*) &= 0, \\ (I - D\mathcal{K})(\tilde{\phi}^*, k^*) &\in L(X) \\ \left\| \left( (I - D\mathcal{K})(\tilde{\phi}^*, k^*) \right)^{-1} \right\| &\leq \beta \end{aligned}$$

mit einem  $\beta > 0$ . Dann existiert für alle hinreichend großen  $N$  ein Paar  $(\tilde{\phi}^N, k^N) \in X$ , so daß

$$(I - \mathcal{K}_N)(\tilde{\phi}^N, k^N) = 0, \quad \|(\tilde{\phi}^N, k^N) - (\tilde{\phi}^*, k^*)\| \rightarrow 0$$

gilt. Darüberhinaus existieren positive Konstanten  $C, s$ , so daß

$$\|(\tilde{\phi}^N, k^N) - (\tilde{\phi}^*, k^*)\| \leq Ce^{-sN}$$

gilt.

Der Beweis des Satzes beruht auf der Approximationstheorie für nichtlineare Operatoren von R. Moore ([Moo68, Theorem 2], vgl. auch [Ans71, Chap. 6.6]).

Es werden zunächst die folgenden beiden Fakten über die Qualität der Approximation durch trigonometrische Interpolation benötigt.

**4.16 Lemma ([Kre99, Theorem 11.7]).** Sei  $g : \mathbb{R} \rightarrow \mathbb{R}$  eine analytische  $2\pi$ -periodische Funktion. Dann existiert ein Streifen  $U = \mathbb{R} \times [-\sigma, \sigma] \subset \mathbb{C}$  mit  $\sigma > 0$ , so daß  $g$  zu einer holomorphen,  $2\pi$ -periodischen und beschränkten Funktion  $g : U \rightarrow \mathbb{C}$  fortgesetzt werden kann. Der Fehler bei trigonometrischer Interpolation kann durch

$$\|P_N g - g\|_\infty \leq C e^{-\sigma N}$$

abgeschätzt werden, wobei die Konstante  $C$  nur von der Schranke der Funktion  $g$  auf  $U$  abhängt.

Speziell für die Abschätzung des Quadraturoperators  $K_{1,N}$  benötigt man das Lemma

**4.17 Lemma ([Kre95, Lemma 4.1]).** Sei  $f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{C}$  eine stetige Funktion, die  $2\pi$ -periodisch bezüglich beider Variablen und stetig differenzierbar bezüglich der ersten Variablen ist. Dann gilt für

$$u(t) := \int_0^{2\pi} \log \left( 4 \sin^2 \frac{t-\tau}{2} \right) f(t, \tau) d\tau, \quad 0 \leq t \leq 2\pi$$

die Abschätzung

$$\|u\|_{0,\alpha} \leq \gamma \left( \|f\|_\infty + \left\| \frac{\partial f}{\partial t} \right\|_\infty \right),$$

für alle  $0 < \alpha < 1$  und eine Konstante  $\gamma$ , die nur von  $\alpha$  abhängt.

Als erstes wird nun gezeigt, daß punktweise  $\mathcal{K}_N \rightarrow \mathcal{K}$  konvergiert, und daß dies auch für die entsprechenden Fréchet-Ableitungen gilt.

**4.18 Lemma.** Für alle  $(\tilde{\phi}, k) \in X$  und  $(h, \kappa) \in X$  existieren positive Konstanten  $C, s$  derart, daß

$$\|\mathcal{K}_N(\tilde{\phi}, k) - \mathcal{K}(\tilde{\phi}, k)\| \leq C e^{-sN}$$

und

$$\|D\mathcal{K}_N(\tilde{\phi}, k)(h, \kappa) - D\mathcal{K}(\tilde{\phi}, k)(h, \kappa)\| \leq C e^{-sN}$$

gilt.

**Beweis:** Sei  $(\tilde{\phi}, k) \in X$ . Definiert man kontinuierliche Operatoren  $K_1$  und  $K_2$  zu den Quadraturoperatoren  $K_{1,N}$  und  $K_{2,N}$  durch

$$\begin{aligned} (K_1 \tilde{\phi})(t) &:= \int_0^{2\pi} \log \left( 4 \sin^2 \frac{t-\tau}{2} \right) H_1(t, \tau) \tilde{\phi}(\tau) d\tau, \\ (K_2 \tilde{\phi})(t) &:= \int_0^{2\pi} H_2(t, \tau) \tilde{\phi}(\tau) d\tau, \end{aligned}$$

so ist zunächst  $(K\tilde{\phi})(t) = (K_1\tilde{\phi} + K_2\tilde{\phi})(t)$ . Die Anwendung von Lemma (4.17) auf  $f(t, \tau) = (P_N H_1(t, \tau)\tilde{\phi})(\tau) - H_1(t, \tau)\tilde{\phi}(\tau)$  ergibt für jedes  $0 < \alpha < 1$  die Abschätzung

$$\begin{aligned} \|K_{1,N}\tilde{\phi} - K_1\tilde{\phi}\|_{0,\alpha} &\leq \gamma \|(P_N H_1(t, \tau)\tilde{\phi})(\tau) - H_1(t, \tau)\tilde{\phi}(\tau)\|_\infty \\ &\quad + \gamma \left\| \left( P_N \frac{\partial H_1(t, \tau)}{\partial t} \tilde{\phi} \right) (\tau) - \frac{\partial H_1(t, \tau)}{\partial t} \tilde{\phi}(\tau) \right\|_\infty, \end{aligned}$$

mit einer von  $\alpha$  abhängigen Konstante  $\gamma$ . Wegen  $\|\cdot\|_\infty \leq \|\cdot\|_{0,\alpha}$  folgt mit Lemma (4.16) somit

$$\|K_{1,N}\tilde{\phi} - K_1\tilde{\phi}\|_\infty \leq \gamma_1 e^{-\sigma_1 N} \quad (4.46)$$

mit einer nur von  $\alpha, k$  und  $\tilde{\phi}$  abhängigen Konstanten  $\gamma_1$  und einer von  $k, \tilde{\phi}$  abhängigen Konstante  $\sigma_1$ . Im folgenden sei  $\alpha = \frac{1}{2}$  fest gewählt

Es ergibt sich mit Lemma (4.16)

$$\begin{aligned} \|K_{2,N}\tilde{\phi} - K_2\tilde{\phi}\|_\infty &= \sup_{t \in [0, 2\pi]} \left| \int_0^{2\pi} [(P_N H_2(t, \tau)\tilde{\phi})(\tau) - H_2(t, \tau)\tilde{\phi}(\tau)] d\tau \right| \\ &\leq 2\pi \|(P_N H_2(t, \tau)\tilde{\phi})(\tau) - H_2(t, \tau)\tilde{\phi}(\tau)\|_\infty \\ &\leq \gamma_2 e^{-\sigma_2 N} \end{aligned}$$

mit nur von  $k, \tilde{\phi}$  abhängigen Konstanten  $\gamma_2$  und  $\sigma_2$ .

Setzt man diese beiden Ungleichungen zusammen, so folgt

$$\|(K_{1,N}\tilde{\phi} + K_{2,N}\tilde{\phi}) - K\tilde{\phi}\|_\infty \leq \gamma e^{-\sigma N}$$

mit nur von  $k, \tilde{\phi}$  abhängigen Konstanten  $\gamma$  und  $\sigma$ . Hieraus folgt die erste Behauptung.

Die Fréchet-Ableitung des Operators  $\mathcal{K}$  im Punkt  $(\tilde{\phi}, k)$  angewandt auf  $(h, \kappa)$  berechnet sich zu

$$\begin{aligned} DK(\tilde{\phi}, k)(h, \kappa) &= \begin{pmatrix} (D_{\tilde{\phi}} K)(\tilde{\phi}, k) & (D_k K)(\tilde{\phi}, k) \\ 2\langle \tilde{\phi}, \cdot \rangle & 0 \end{pmatrix} \begin{pmatrix} h \\ \kappa \end{pmatrix} \\ &= \begin{pmatrix} (D_{\tilde{\phi}} K)(\tilde{\phi}, k)h + \kappa(D_k K)(\tilde{\phi}, k) \\ 2\langle \tilde{\phi}, h \rangle \end{pmatrix} \\ &=: \begin{pmatrix} DK(\tilde{\phi}, k) \begin{pmatrix} h \\ \kappa \end{pmatrix} \\ 2\langle \tilde{\phi}, h \rangle \end{pmatrix} \end{aligned}$$

mit einem linearen Operator  $DK(\tilde{\phi}, k) : X \rightarrow \mathcal{C}[0, 2\pi]$  definiert durch

$$DK(\tilde{\phi}, k) \begin{pmatrix} h \\ \kappa \end{pmatrix} = \int_0^{2\pi} H(t, \tau)h(\tau)d\tau + \kappa \int_0^{2\pi} \frac{\partial H(t, \tau)}{\partial k} \tilde{\phi}(\tau)d\tau. \quad (4.47)$$

Der Kern

$$H(t, \tau) = \frac{ik n(\tau) \cdot (x(t) - x(\tau))}{2 |x(t) - x(\tau)|} H_1^{(+)}(k|x(t) - x(\tau)|)$$

wurde in Gleichung (4.23) definiert. Die Ableitung nach  $k$  berechnet sich zu

$$\begin{aligned} \frac{\partial H(t, \tau)}{\partial k} &= \frac{i n(\tau)(x(t) - x(\tau))}{2 |x(t) - x(\tau)|} \left( H_1^{(+)}(k|x(t) - x(\tau)|) + k|x(t) - x(\tau)| H_1^{(+)\prime}(k|x(t) - x(\tau)|) \right) \\ &= \frac{i}{2} n(\tau)(x(t) - x(\tau)) \left\{ \frac{H_1^{(+)}(k|x(t) - x(\tau)|)}{|x(t) - x(\tau)|} \right. \\ &\quad \left. + k \left( -H_2^{(+)}(k|x(t) - x(\tau)|) + \frac{H_1^{(+)}(k|x(t) - x(\tau)|)}{k|x(t) - x(\tau)|} \right) \right\} \\ &= \frac{ik}{2} n(\tau)(x(t) - x(\tau)) \left( -H_2^{(+)}(k|x(t) - x(\tau)|) + 2 \frac{H_1^{(+)}(k|x(t) - x(\tau)|)}{k|x(t) - x(\tau)|} \right) \\ &= \frac{ik}{2} n(\tau)(x(t) - x(\tau)) H_0^{(+)}(k|x(t) - x(\tau)|). \end{aligned}$$

Nach Kapitel 4.2 Abschnitt a. und d. existiert somit auch eine Aufspaltung des Integrals in Gleichung (4.47) in einen analytischen und singulären Teil.

Die im ersten Teil des Beweises benutzte Technik zum Beweis der punktweisen Konvergenz von  $K_N \rightarrow K$  läßt sich somit auch zum Beweis der punktweisen Konvergenz von  $DK_N(\tilde{\phi}, k) \rightarrow DK(\tilde{\phi}, k)$  benutzen. Analog folgt also

$$\|DK_N(\tilde{\phi}, k)(h, \kappa) - DK(\tilde{\phi}, k)(h, \kappa)\| \leq \gamma_D e^{-\sigma_D N}$$

mit von  $N$  unabhängigen Konstanten  $\gamma_D$  und  $\sigma_D$ . Hieraus folgt die Behauptung.  $\square$

Im allgemeinen sind die Folgen  $(\mathcal{K}_N)$  und  $(DK_N)$  nur punktweise konvergent gegen  $\mathcal{K}$  bzw.  $DK_N$  aber nicht Norm konvergent ([Kre99, Seite 203]). Um diesen Nachteil auszugleichen benötigt man zusätzliche Kompaktheitseigenschaften der Folgen  $(\mathcal{K}_N)$  und  $D(\mathcal{K}_N)$ .

**4.19 Definition.** Eine Menge  $\mathcal{A}$  von nicht-linearen Operatoren auf  $X$  heißt *kollektiv kompakt*, falls für jede beschränkte Menge  $S \subset X$  die Menge  $\mathcal{A}S$  relativ kompakt ist. Eine Folge nicht-linearer Operatoren ist kollektiv kompakt, falls die zugehörige Menge kollektiv kompakt ist.

Allgemein zeigt Moore, daß Folgen von Quadratoroperatoren und deren Fréchet-Ableitungen, die bestimmte Integraloperatoren approximieren, stets kollektiv kompakt sind. Mit diesen Vorarbeiten kann nun der Satz (4.15) bewiesen werden.

**Beweis von Satz (4.15):**

Nimmt man die Voraussetzungen von Satz (4.15) und die in Lemma (4.18) bewiesenen Eigenschaften zusammen, so sind die Voraussetzungen des Satzes von Moore ([Ans71, Theorem 6.8])

erfüllt. Es existiert daher eine Konstante  $C$ , so daß mit

$$d_N := \sup_{\|(\tilde{\phi}, k)\|=1} \|\mathcal{K}_N(\tilde{\phi}, k) - \mathcal{K}(\tilde{\phi}, k)\|$$

$$d_{N0}^* := \sup_{\substack{\|(\tilde{\phi}, k)\|=1 \\ (h, \kappa) \in X}} \|DK_N(\tilde{\phi}, k)(h, \kappa) - DK(\tilde{\phi}, k)(h, \kappa)\|.$$

die Abschätzung

$$\|(\tilde{\phi}^N, k^N) - (\tilde{\phi}^*, k^*)\| \leq C \cdot (d_N + d_{N0}^*)$$

erfüllt ist, was zu zeigen war.  $\square$

Diese theoretische Fehlerabschätzung und Konvergenzrate wird auch in der Praxis bestätigt. Die Abbildung 24 zeigt Konvergenzmessungen für Eigenresonanzen des Einheitskreis. Die Eigenresonanz in der linken Seite ist die erste Nullstelle der 5-ten Besselfunktion  $k^* = 8.7714838$ , die Eigenresonanz in rechten Seite ist die 9-te Nullstelle der 9-ten Besselfunktion  $k^* = 40.6285537$ . Auf der  $y$ -Achse ist jeweils  $\ln(|k^* - k^N|)$  aufgetragen, auf der  $x$ -Achse die Diskretisierungszahl  $N$ . In beiden Fällen sieht man deutlich, das der Logarithmus des Fehlers linear in  $N$  abnimmt, wie Satz (4.15) es fordert.

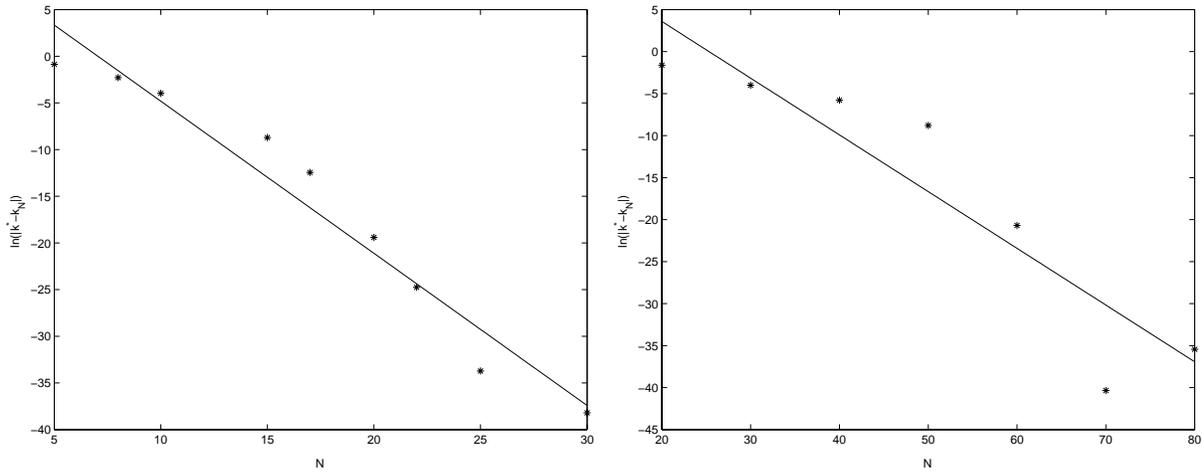


Abbildung 24: Die Konvergenz ist exponentiell. Die in Satz 4.15 auftretenden Konstanten  $C$  und  $s$  nehmen in der Praxis vernünftige Werte an.

Auch für asymmetrische  $C^\infty$ -Gebiete  $\Omega$  ist Satz (4.15) gültig. Auch bei solchen Gebieten bleiben die Konstanten  $C$  und  $s$  vernünftig, wie die Abbildung 25 für die verschiedenen Referenz-Beispiele aus Kapitel 3 zeigt.

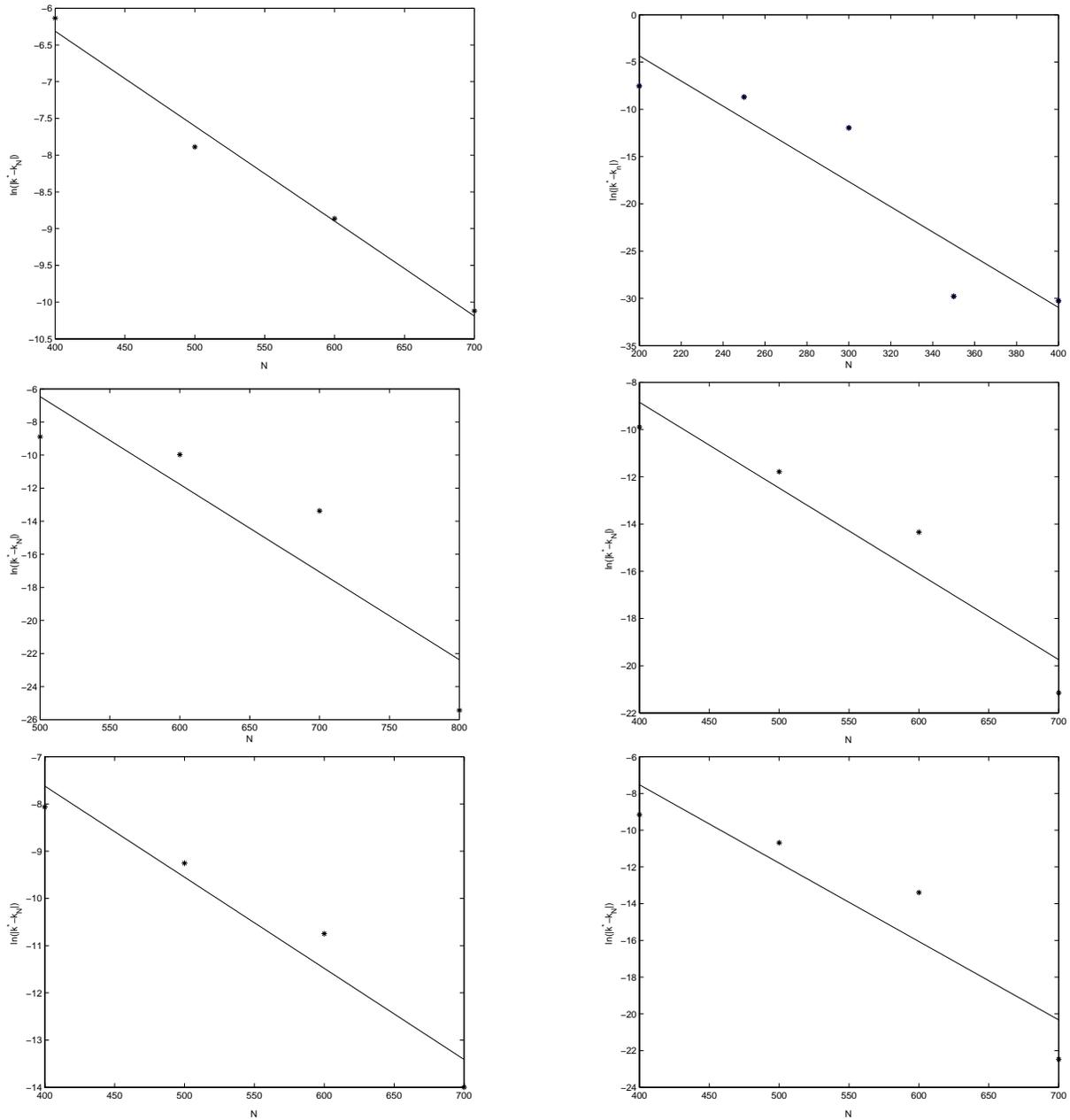


Abbildung 25: Konvergenz in Abhängigkeit von  $N$  der Referenz-Beispiele aus Kapitel 3. Die Reihenfolge ist die selbe wie in Abbildung 23. Die Konvergenz ist auch bei asymmetrischen Gebiet  $\Omega$  und irregulären Eigenmoden exponentiell.

### Die Wahl von $N$

Eine geschickte Wahl von  $N$  ist von großer Bedeutung. Wählt man  $N$  zu klein, so ist der Fehler  $|k^* - k^N|$  viel zu groß. Wird  $N$  jedoch zu groß gewählt, ist der numerische Aufwand unnötig hoch. Die Wahl von  $N$  hängt nicht von der Lösungsgestalt der Eigenmode ab, sondern nur von der Wellenzahl  $k$  und der Geometrie des Lösungsgebiets. Berechnet man alle Eigenwerte in einem kleinen Intervall, so muß  $N$  nicht geändert werden, wenn man vor der Berechnung einmal  $N$  so gewählt hat, daß eine vorgegebene Genauigkeit erreicht wird.

Im eindimensionalen Fall reduziert sich das Helmholtz-Problem zu

$$u_{xx} + k^2 u = 0, \quad u(0) = u(L) = 0,$$

wobei  $L$  die Länge des Gebietes ist. Bekanntermaßen ist für jedes  $m \in \mathbb{Z}$  und  $k := \frac{2\pi}{L}m$  durch  $u(x) = e^{ikx}$  eine Lösung gegeben. Die Wellenlänge beträgt  $\lambda = \frac{2\pi}{k}$ , und somit passen  $m = \frac{kL}{2\pi}$  viele Wellenlängen in das Gebiet. Diese Überlegungen lassen sich auf den zweidimensionalen Fall übertragen. Ist  $L$  die Länge des Randes  $\partial\Omega$ , so passen höchstens  $\frac{kL}{2\pi}$  viele Wellenlängen auf den Rand. Setzt man

$$N \approx \frac{kL}{2\pi} \quad \text{bzw.} \quad N \approx \frac{nkL}{2\pi}$$

so läßt sich die Wellenfunktion entlang des Randes gut durch eine Funktion aus dem Raum  $T_N$  der trigonometrischen Polynome approximieren. Mit einem *Sicherheitsfaktor*  $\rho = 1.5$ , also  $N \approx 1.5 \frac{kL}{2\pi}$  bzw.  $N \approx 1.5 \frac{nkL}{2\pi}$  erreicht man hervorragende Fehlerabschätzungen für  $|k^* - k^N|$  wie man in Abbildungen 24 und 25 sehen kann.

Um eine vorgegebene Genauigkeit des zu bestimmenden Eigenwertes zu garantieren, bestimmt man zunächst  $k^N$  mit  $N \approx \frac{kL}{2\pi}$  und erhöht dann  $N$  schrittweise, bis die gewünschte Genauigkeit erreicht ist.

$N$	$k^N$
20	$40.54268295835170 - 1.787438685151035 \cdot 10^1 i$
30	$40.62871418447914 + 1.832740070399995 \cdot 10^{-2} i$
40	$40.62859364720026 + 3.073176431383128 \cdot 10^{-3} i$
50	$40.62855571087057 + 1.514321392844854 \cdot 10^{-4} i$
60	$40.62855371897813 + 1.024774541453136 \cdot 10^{-9} i$
70	$40.62855371896453 + 3.030503112013308 \cdot 10^{-18} i$

Tabelle 1: Konvergenz eines Eigenwertes des Einheitskreis. Der korrekte Wert beträgt  $k^* = 40.628553718964527622$  (alle Stellen korrekt). Für  $N = 40 \approx \frac{kL}{2\pi}$  sind schon 3 Stellen hinter dem Komma korrekt. Mit einem Sicherheitsfaktor  $\rho = 1.5$  stimmen schon 9 Stellen hinter dem Komma.

### Kondition der Diskretisierungsmatrix in der Nähe des gesuchten Eigenwerts

Zu einer Diskretisierung  $N$  bestimmt der Korrektor ein  $k^N$ , so daß die Matrix  $\tilde{A}^N(k^N)$  singular ist. Dazu wird eine Folge  $k_0, k_1, \dots$  gebildet, die gegen  $k^N$  konvergiert. Der Korrektor führt 2  $LU$ -Zerlegungen aus. Die erste wird für  $\tilde{A}(k_0)$  durchgeführt, die zweite für  $\tilde{A}(k_l)$  mit  $|k_l - k^N| \approx \Delta_k/10$ . Da die Matrix  $\tilde{A}(k^N)$  singular ist, gilt  $\text{cond}(\tilde{A}(k^N)) = \infty$ . Um die Stabilität des Korrektors gewährleisten zu können, müssen die Matrizen  $\tilde{A}(k_0)$  und  $\tilde{A}(k_l)$  vernünftig konditioniert sein. Analytische Abschätzungen für die Kondition existieren nicht, so daß man sich auf numerisch bestimmte Ergebnisse stützen muß. Dazu wurde  $N = 500$  gewählt und  $\text{cond}(\tilde{A}(k^N - \frac{\Delta_k}{5}))$ ,  $\text{cond}(\tilde{A}(k^N + \frac{\Delta_k}{10}))$  und  $\text{cond}(\tilde{A}(k^N - \frac{\Delta_k}{10}))$  numerisch bestimmt. Für  $k^*$  wurden die 6 Referenz-Beispiele aus Kapitel 3 genommen. Die Ergebnisse finden sich in Tabelle 2.

$N = 500$	$\text{cond}(\tilde{A}(k^N - \frac{\Delta_k}{5}))$	$\text{cond}(\tilde{A}(k^N + \frac{\Delta_k}{10}))$	$\text{cond}(\tilde{A}(k^N - \frac{\Delta_k}{10}))$
Beispiel 1	1390.1756	2331.5236	2178.7876
Beispiel 2	768.9324	2050.8742	2223.9629
Beispiel 3	743.9041	974.4677	1487.0091
Beispiel 4	539.72861	1082.5865	1079.0200
Beispiel 5	819.1820	1297.2396	1248.6215
Beispiel 6	528.68401	1055.8187	1055.6206

Tabelle 2: Die Kondition der jeweiligen Matrizen ist in allen Fällen vernünftig. Durch die jeweiligen  $LU$ -Zerlegungen werden also keine verfälschenden Rundungsfehler erwartet. Die Typen sind in der Reihenfolge der Referenzbeispiele aus Kapitel 3, vgl. Abbildung 23 .

### Güte des Prädiktors

Für die Qualität des Prädiktors ist wesentlich, daß die in dem  $O(|k^* - \tilde{k}|^2)$ -Term versteckte Konstante  $\alpha$  nicht zu groß ist. Leider liefert die Abschätzung (4.39) keine vernünftigen analytischen Resultate, da die in der Praxis erwartete Größenordnung von  $\alpha$  erheblich kleiner ist. Um eine für die Praxis relevante Größenordnung von  $\alpha$  zu bestimmen, wurden für verschiedene Billards und Start-Eigenwerte  $k_0$  numerische Messungen für  $\alpha$  vorgenommen.

Im ersten Beispiel wurde für das Ei-Billard  $N = 750$  gewählt und als Starteigenwert  $k_0 = 100.045026119225$ . Der direkte Sweep-Algorithmus 4.14 liefert 71 Nachbarn. Für alle diese Nachbarn wurde dann die Approximationsgüte  $\alpha$  bestimmt. Der Durchschnittswert beträgt 2.70778, der maximale Wert beträgt 3.0472301.

Im zweiten Beispiel wurde für das Kartoffel-Billard mit  $N = 600$  und Starteigenwert  $k_0 = 99.99891622$  der direkte Sweep-Algorithmus gestartet. Er liefert 20 Nachbarn mit einer mittleren Approximationsgüte 1.40027292 und einer maximalen Güte von 1.6572521.

In beiden Beispielen ist die Approximationsgüte des Prädiktors akzeptabel groß. Für weitere Beispiele am Ei- bzw. Kartoffel-Billard ergaben sich ähnliche Gütewerte, d.h. ein  $\alpha \approx 3$  für das Ei-Billard und  $\alpha \approx 1.5$  für das Kartoffel-Billard.

### Konvergenzrate $\beta$ des Korrektors

Die Konvergenzrate des Korrektors wurde ebenfalls für Eigenwerten bestimmt, die in praktischen Anwendungen relevant sind. Es wurden wiederum die Eigenwerte aus Kapitel 3 gewählt,  $N = 900$  gewählt und der Korrektor wurde jeweils bei  $k^N - \frac{\Delta k}{5}$  gestartet. In allen Fällen zeigte sich, daß die Konvergenzrate nach dem Update deutlich besser wird, in Abhängigkeit des Eigenwertes um einen Faktor von 3 bis 500.

### Konvergenzaussagen für das Transmissions-Problem

Der theoretisch vorhergesagte exponentielle Abfall des Fehlers konnte auch in praktischen Rechnungen bestätigt werden. Die auftretenden Konstanten  $C$  und  $\sigma$  nehmen in praktischen Beispielen vernünftige Werte an. Hierzu wurde zunächst die Konvergenz eines Kreis-Billards mit  $n_0 = 1.5$  numerisch untersucht. Die untersuchte Eigenresonanz war  $k^* = 9.489027663741322 - 0.3981092205268944i$ .

Im Falle des für die Praxis relevanten (vgl. Kapitel 5) Ei-Billards wurde eine Resonanz  $k^* = 29.27656365159416 - 0.03164897713314i$  mit  $n_0 = 1.5$ , im Falle des Kartoffel-Billards eine Resonanz  $k^* = 30.08695202122208 - 0.01159580267233i$  mit  $n_0 = 3.14$  berechnet. Die Konvergenz in Abhängigkeit von  $N$  ist in Abbildung 26 gezeigt.

Da die Methode sogar für nicht-konvexe  $\Omega$  anwendbar ist, wurde auch ein nicht-konvexes Testbeispiel gewählt. Hier war  $n_0 = 1.5$  und  $k^* = 10.06175555972444 - 0.30897032641919i$ . Die zu dieser Eigenresonanz gehörende Eigenfunktion ist in Abbildung 28, Seite 88 dargestellt. Die Konvergenz in Abhängigkeit von  $N$  ist in Abbildung 26 zusammengefaßt.

Die praktischen Ergebnisse bestätigen die theoretischen Vorhersagen der außerordentlich schnellen Konvergenz der in dieser Arbeit vorgestellten Diskretisierungsmethode.

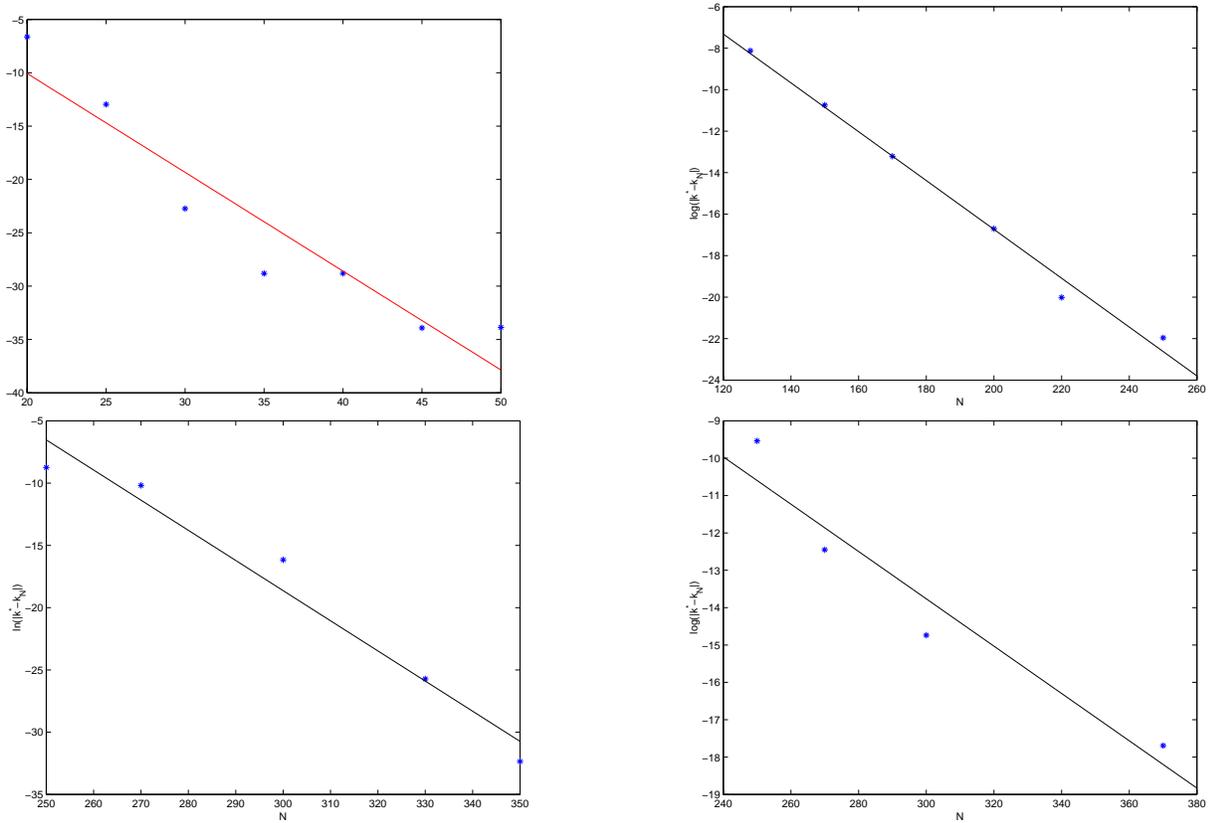


Abbildung 26: Oben Links: Kreisförmiges Billard mit  $n_0 = 1.5$ , Oben Rechts: nicht-konvexes Billard mit  $n_0 = 1.5$ , Unten Links: Ei-Billard mit  $n_0 = 1.5$ , Unten Rechts: Kartoffel-Billard mit  $n_0 = 3.14$ . Die Konvergenz der Randintegral-Diskretisierung mit Splitten des Integral-Kerns ist exponentiell.

## 4.5 Vergleich mit anderen numerischen Verfahren

Um die Effizienz des vorgestellten Verfahrens zu verdeutlichen, werden die Eigenschaften des neuen Algorithmus mit bisher verwendeten Algorithmen verglichen. Der Vergleich wird für (HP) und (TP) getrennt durchgeführt, da einige der bisher benutzten Verfahren nicht für beide Fälle geeignet waren.

### 4.5.1 Verfahren zum quasi-klassischen Helmholtz-Problem

Das Helmholtz-Problem gehört zu den Standard-Problemen der numerischen Mathematik und der mathematischen Physik. Wegen der besonderen Relevanz sind verschiedene Methoden und Ansätze entwickelt worden. Ein kurzer Überblick und eine Einteilung der verschiedenen Algorithmen in zwei Klassen findet sich in [Bar00]. In diesem Abschnitt wird diese Einteilung übernommen und der oben beschriebene direkte Sweep-Algorithmus zusammen mit der Randintegral-Diskretisierung in dieses Schema eingeordnet und verglichen.

Sei  $u(x)$  eine Lösung des Helmholtz-Problems  $\Delta u + k^2 u = 0$  im Gebiet  $\Omega \subset \mathbb{R}^2$  mit Dirichlet-Bedingung.  $u(x)$  kann als Linearkombination von *Basis-Funktionen* aus einem geeigneten Funktionenraum dargestellt werden,

$$u(x) = \sum_{j=1}^N x_j \phi_j(x). \quad (4.48)$$

Dies ist numerisch notwendig, da der unendlich-dimensionale Funktionenraum, in dem sich  $u$  befindet, durch einen endlich-dimensionalen Approximationsraum ersetzt werden muß. Ist die Basis  $\{\phi_j\}$  vollständig, so wird  $u$  im Limit  $N \rightarrow \infty$  exakt durch Gleichung (4.48) beschrieben. Für numerische Zwecke müssen  $N$  und die Basis-Funktionen geschickt gewählt werden, um den Rechenaufwand so klein wie möglich zu halten. Für die Basiswahl ergeben sich prinzipiell zwei Möglichkeiten:

- Die Basis-Funktionen hängen nicht von  $k$  ab.
- Die Basis-Funktionen sind von  $k$  abhängig.

Es ergeben sich zwei verschiedene numerische Ansätze, die die Algorithmen in zwei Klassen teilen.

	Klasse (A) – <i>Basis-Diagonalisierung</i>	Klasse (B) – <i>Matching der Fundamentallösung</i>
<i>Basis- funktionen</i>	<ul style="list-style-type: none"> <li>• unabhängig von <math>k</math></li> <li>• erfüllen Randbedingungen</li> <li>• Erfüllen Helmholtz-Gleichung nicht</li> </ul>	<ul style="list-style-type: none"> <li>• abhängig von <math>k</math></li> <li>• erfüllen Randbedingungen nicht</li> <li>• Sind Lösungen der Helmholtz Gleichung</li> </ul>
<i>Basisgröße</i>	skaliert wie $\text{vol}(\Omega)$	skaliert wie $\text{vol}(\partial\Omega)$
<i>resultierende diskrete Gleichung</i>	$Hx = Ex$  Diagonalisierung der Matrix $H$ , findet viele Paare $(k_j, x_j)$ gleichzeitig	$A(k)x = 0$  Suche im $k$ -Intervall nach $\det A(k) = 0$ mit Sweep- Algorithmus
<i>Beispiele</i>	<ul style="list-style-type: none"> <li>• Finite-Elemente [SF73]</li> <li>• Finite-Differenzen [TOS01]</li> <li>• Conformal Mapping [Rob84]</li> </ul>	<ul style="list-style-type: none"> <li>• Randintegral-Methoden [KS97]</li> <li>• Plane-Wave-Decomposition [Hel91]</li> <li>• Bogomolny - Methode [Bog92]</li> </ul>

### Algorithmen der Klasse A

Die Unabhängigkeit der Basisfunktionen von  $k$  bewahrt die Linearität des Eigenwertproblems und resultiert in einem Matrix-Eigenwertproblem. Auf diese Weise werden viele Eigenwerte und Vektoren durch eine Diagonalisierung gewonnen. Prinzipiell können so alle Eigenwerte vom 1. bis zu einem maximalen (abhängig von der gewünschten Genauigkeit) bestimmt werden. Die Anzahl von nutzbaren Eigenwerten ist abhängig von der Anzahl  $N$  der Basisfunktionen in der Darstellung (4.48). Für ein Gebiet  $\Omega \subset \mathbb{R}^2$  skaliert sich  $N$  somit wie  $k^2$ , falls ein Zustand mit Wellenzahl  $k$  gesucht ist.

Als Basisfunktionen können zum Beispiel analytisch bekannte Eigenfunktionen einer einfacher zu lösenden Gleichung gewählt werden. Die Basis ist dann orthonormal und vollständig. Alternativ kann die Basis so gewählt werden, daß die resultierende Matrix  $H$  dünn-besetzt ist und die notwendige Diagonalisierung schnell durchführbar ist. Es wird eine breite Klasse von Basisfunktionen zugelassen, die zum Beispiel Gitterfunktionen, Finite-Elemente Ansatzfunktion, B-Splines, ecetera, enthält ([BCO81], [Gru87], [SF91], [Dav93], [SF73], [Str86]). Je glatter die Basisfunktionen sind, desto schneller ist die Konvergenz. Andererseits führen glatte Basisfunktionen meistens zu einer voll-besetzten Matrix  $H$ . Gittermethoden (zum Beispiel Finite-Differenzen) führen zu dünn-besetzten Matrizen, jedoch sind die Konvergenzeigenschaften nicht sehr gut.

Für spezielle Gebiete lassen sich unter speziellen Voraussetzung sehr spezielle Basisfunktionen

wählen, die dann zu dünn-besetzten Matrizen führen und gute Konvergenzeigenschaften haben ([Rob84]).

Für die Diagonalisierung der Matrix  $H$  kann anschließend ein Standardverfahren der numerischen linearen Algebra benutzt werden, das die jeweiligen Strukturen der Matrix optimal ausnutzt.

### Algorithmen der Klasse B

Oftmals ist die Fundamentallösung der zu lösenden Gleichung für spezielle Gebiete bekannt. In diesem Fall ist es dann möglich, die Funktion  $u$  im Inneren von  $\Omega$  nur aus Kenntnis der Funktionswerte auf dem Rand  $\partial\Omega$  zu konstruieren. Die Dimension des Problems wird somit auf die Dimension des Randes reduziert. Das Problem ist auf diese Weise zu einem Rand-Matching Problem geworden, das äquivalent zu einer *Null-Determinanten-Bedingung* ist. Durch einen derartigen Ansatz gibt man zwangsläufig die Linearität in der Diskretisierungsmatrix auf. Für gewöhnlich ist es dann nicht mehr vorhersagbar, für welche  $k$  die Determinante verschwinden wird, so daß eine Suche (Sweep) im  $k$ -Raum notwendig wird. Die numerische Aufgabe besteht also nicht nur in der Wahl einer Basis, sondern auch in der effizienten Durchführung des notwendigen Sweeps.

Die oben vorgestellte Formulierung der Helmholtz-Gleichung als Integralgleichung führt auf einen Algorithmus der Klasse B. Die Diskretisierung hängt von der Wahl des numerischen Quadraturverfahrens ab. In der ingenieurwissenschaftlichen- und physikalischen Literatur werden meistens auf der Trapez- oder Simpsonsregel basierende Quadraturformeln benutzt ([Bäc02], [Boa94], [Bre81], [KS97]). Diese Ansätze stellen keine stark einschränkende Bedingungen an die Beschaffenheit des Randes  $\partial\Omega$  da, lediglich eine hinreichende Glätte wird vorausgesetzt. Ein spezieller quasi-klassischer Integraloperator wird zur Diskretisierung nach Bogomolny ([Bog92]) verwendet. Eine weitere viel benutzte Methode ist die sogenannte *Plane-Wave-Decomposition* von E. Heller ([Hel91], [Hel86]).

Für alle Methoden ist nach erfolgreicher Diskretisierung ein Sweep zum Aufspüren der Eigenwerte notwendig. Die in der Literatur bisher benutzten Sweep-Verfahren benutzen eine Testfunktion als Maß für die Regularität der Diskretisierungsmatrix. Die Determinante ist wegen der numerischen Instabilität bei der Auswertung unbrauchbar, so daß im allgemeinen eine Singulärwertzerlegung benutzt wird ([Bar00], [Bäc02]). Dann wird das zu durchsuchende Intervall nach Minima der Testfunktion abgesucht.

### Andere Algorithmen

Vergini und Saraceno ([VS95]) entwickelten eine Art Hybrid-Algorithmus zur Bestimmung der Eigenwerte. Ihr Algorithmus gehört zur Klasse B, liefert aber mit einer Diagonalisierung mehrere Eigenwerte – vergleichbar mit Algorithmen der Klasse A. Für spezielle Billards (zum Beispiel

dem Stadium-Billard) konnten Vergini und Saraceno besonders gut geeignete Basisfunktionen angeben, für diese Billards ist dieses Verfahren die effizienteste Methode zur Bestimmung von Eigenwerten im quasi-klassischen Bereich. Jedoch ist das Problem ungelöst, eine derartige Basis für beliebige Billards zu bestimmen ([Bar00]).

### Vergleich

Die Algorithmen der Klasse A eignen sich besonders gut zur Bestimmung der kleinsten Eigenwerte, da alle Eigenwerte bis zu einem bestimmten, durch die vorgegebene Genauigkeit festgelegten, berechnet werden. Sollen jedoch nur einige Eigenwerte im quasi-klassischen Bereich berechnet werden, ist der numerische Aufwand zu groß. Prinzipiell sind auch Algorithmen der Klasse B geeignet, Eigenwerte für moderat großes  $k$  zu bestimmen. Liegen jedoch Eigenwerte dichter bei einander als die Auflösung des benutzten Sweep-Algorithmus, so werden diese bei einer Suche übersehen. Numerische Experimente in [BRH98] ergaben bei Verwendung der Plane-Wave-Decomposition zusammen mit einem SVD-Sweep, daß ca. 10 Prozent der Eigenwerte übersehen werden. Experimente mit der oben vorgeschlagenen Randintegral-Diskretisierung ebenfalls zusammen mit einem SVD-Sweep konnten diese Resultate auch auf andere Diskretisierung-Methoden übertragen. Das Problem liegt in der mangelnden Fähigkeit des SVD-Sweeps, eng benachbarte Eigenwerte zu trennen. *Diese Probleme hat der in dieser Arbeit vorgestellte Direkte-Sweep nicht.* Abbildung 27 zeigt die Anzahl  $N(E)$  der berechneten Eigenwerte bis zu einer Energie  $E = k^2$ . Die Gerade entspricht der durch die Weyl-Formel (3.1) bestimmten theoretischen Asymptote. Die Kurve längs der Asymptote entspricht der mit dem Direkten-Sweep berechneten Anzahl, die untere Kurve der mit einem SVD-Sweep berechneten Anzahl. Die Auflösung des SVD-Sweeps wurde so gewählt, daß beide Algorithmen die gleiche Laufzeit hatten.

Ein großer Vorteil der Klasse B - Algorithmen ist die hohe erreichbare Genauigkeit. Wie oben bewiesen wurde, besitzt die in dieser Arbeit vorgestellten Randintegral-Diskretisierung einen exponentiell abfallenden Fehler. Somit kann jeder Eigenwert mit beliebiger (Maschinengenauigkeit) bestimmt werden.

Die Aufspaltung des singulären Integralkerns in einen analytischen Teil und einem numerisch effizient handbaren singulären Teil ermöglicht eine sehr effiziente Randintegral-Diskretisierung. Bei den bisher verwendeten Verfahren konnte man nur ein polynomielles Abfallen des Fehlers (meistens auch nur quadratisch) erwarten, bei dem in dieser Arbeit verwendeten jedoch ein exponentielles. Ein wesentlicher Vorteile der Randintegral-Diskretisierung gegenüber anderen Verfahren besteht darin, daß sie leicht für eine große Vielfalt zulässiger Billard-Gebieten nutzbar ist, d.h. die Flexibilität in der Wahl der zugrunde liegenden Randbedingungen. Die anderen oben besprochenen Methoden lassen sich – wenn überhaupt – nur sehr schwer verallgemeinern. Der SVD-Sweep ist im übrigen nur für reelle  $k$  geeignet, da bei einer zwei-dimensionalen Minima-Suche der rechnerische Aufwand zu groß würde.

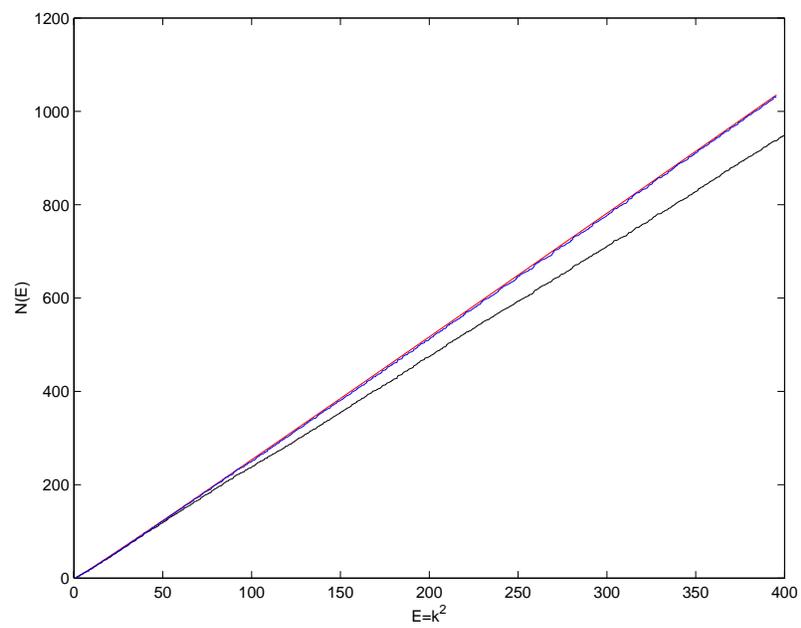


Abbildung 27: Vergleich des SVD-Sweeps (untere Kurve) mit Direkten-Sweep für die Berechnung der ersten 1000 Eigenwerte des Ei-Billards. Der SVD-Sweep *übersieht* circa 10 Prozent der Eigenzustände. Die Gerade entspricht der Weyl-Asymptote.

Mit speziellen Basisfunktionen für bestimmte Billards sind die PWD und die Methode von Vergini der Randintegralmethode mit direktem Sweep überlegen. Insbesondere die Vergini-Methode arbeitet äußerst effizient bei dem sogenannten Stadium-Billard, ein Rechteck mit aufgesetzten Halbkreisen am rechten bzw. linken Rand. Für dieses Billard konnten sehr effizient arbeitende Basisfunktionen angegeben werden. Jedoch ist es nicht so leicht, für beliebige Billards effizient funktionierende Basisfunktionen zu finden ([Bar00]).

#### 4.5.2 Verfahren zum quasi-klassischen Transmissions-Problem

Zur Berechnung von Eigenmoden und Resonanzen mit den hier vorgegebenen Randbedingungen werden in der Literatur hauptsächlich zwei verschiedene Ansätze verfolgt. Der erste Ansatz zur Diskretisierung des Differentialoperators ist der hier vorgestellte Weg über die Randintegralmethode. Der zweite Ansatz benutzt eine geeignete Reihenentwicklung der Lösungsfunktion.

##### Randintegral-Diskretisierungen

In [Bäc02], [Wie03] wird die gleiche äquidistante Diskretisierung des Randes zur numerischen Quadratur benutzt. Jedoch wird die Mittelpunktsregel zur Approximation der Integrale verwendet. An den Singularitäten wird der Grenzwert der Kernfunktion verwendet. Dieses Vorgehen limitiert die Effizienz des Verfahrens drastisch. Zum einen kann höchstens eine quadratische Konvergenz aufgrund der Mittelpunktsregel erwartet werden, zum anderen wirkt sich die Singularität auf die Berechnung der Kernfunktion in der Nähe der Singularität aus. Die numerischen Ergebnisse sind stark mit Rundungs- und Auslöschungsfehlern behaftet. Die oben vorgestellte Diskretisierung umgeht dieses Problem, indem die Singularität abgespalten und eine speziell konstruierte Quadraturregel benutzt wird. Dieser Vorteil wird zwar durch eine etwas schwierigere Implementierung erkauft, der Rechenaufwand bleibt aber gleich.

Zur Berechnung der Resonanzen wird in [Wie03] ein Newton-Verfahren zur Berechnung von  $\det \tilde{A}^N(k) = 0$  benutzt. In jedem Newton-Schritt muß dann die Inverse der Matrix  $\tilde{A}^N(k_l)$  berechnet werden. Dies ist einerseits numerisch sehr aufwendig, andererseits verschlechtert sich die Kondition  $\text{cond}(\tilde{A}^N(k_l))$ , je näher  $k_l$  an die gesuchte Resonanz  $k^*$  rückt. Wie oben gezeigt, treten diese Probleme bei dem Korrektor des direkten Sweep-Algorithmus nicht auf. Der im direkten Sweep-Algorithmus benutzte Prädiktor wird in [Wie03] nicht beschrieben.

##### Matching-Methode

Die Grundidee der Matching-Methode ist aus der Quantisierung für Helmholtz-Probleme mit Dirichlet-Randbedingungen entstanden. Es wird angenommen, daß der Rand  $\partial\Omega$  eindeutig in Polarkoordinaten darstellbar ist, d.h. es existiert eine Funktion  $R(\phi)$  mit  $\partial\Omega = R([0, 2\pi])$  und zu jedem Punkt  $x \in \partial\Omega$  existiert genau ein  $\phi_x \in [0, 2\pi)$  mit  $x = R(\phi_x)$ . Für die Innen- bzw.

Außenlösung wird ein Ansatz in Zylinderfunktionen mit konstantem  $k$  gemacht, also

$$\begin{aligned}\psi_1(r, \phi) &= \sum_{m=-\infty}^{\infty} \left( \alpha_m H_m^{(+)}(nkr) + \beta_m H_m^{(-)}(nkr) \right) e^{im\phi} & r < R(\phi), \\ \psi_2(r, \phi) &= \sum_{m=-\infty}^{\infty} \left( \gamma_m H_m^{(+)}(kr) + \delta_m H_m^{(-)}(kr) \right) e^{im\phi} & r > R(\phi).\end{aligned}$$

Jeder der Summanden erfüllt die entsprechende Helmholtz-Gleichung (für die Innen- bzw. Außenlösung), jedoch werden die Randbedingungen nicht erfüllt. Da die Funktionen  $\{H_m^{(\pm)}(nkr)\}$  und  $\{H_m^{(\pm)}(kr)\}$  eine Normalbasis bilden, ist die Entwicklung für  $r < R_{\min}$  und  $r > R_{\max}$  exakt, solange die Summe über unendlich viele Terme erstreckt wird. Die Grenzen  $R_{\min}$  und  $R_{\max}$  ergeben sich aus der unteren bzw. oberen Schranke von  $R(\phi)$ . Die Annahme, daß sich die Entwicklung im Gebiet  $R_{\min} < r < R_{\max}$  analytisch fortsetzen läßt, ist als *Rayleigh-Hypothese* bekannt. Es läßt sich zeigen, daß diese Vermutung für eine Familie von Deformationen, die durch  $\epsilon$  parametrisiert wird, erst für ein kritisches  $\epsilon_c$  zusammenbricht. Dies geschieht aber erst, nachdem der entsprechende Rand durch die Deformation  $\epsilon_c$  konkav geworden ist. Jedoch führt dieses Problem schon für  $\epsilon < \epsilon_c$  zu numerischen Schwierigkeiten, so daß die Matching-Methode nur für Gebiete benutzt werden kann, die aus einer kleinen Störung eines integrierbaren Gebiets (wie zum Beispiel einer Ellipse) hervorgehen.

In der Entwicklung darf  $\delta_m = 0$  angenommen werden, da keine eingehenden Wellen betrachtet werden. Die oben hergeleiteten Randbedingungen

$$\begin{aligned}\psi_1(R(\phi), \phi) &= \psi_2(R(\phi), \phi) \\ \left. \frac{\partial \psi_1}{\partial r} \right|_{(R(\phi), \phi)} &= \left. \frac{\partial \psi_2}{\partial r} \right|_{(R(\phi), \phi)}\end{aligned}$$

liefern Bedingungen an die restlichen Koeffizienten  $\alpha_m, \beta_m$  und  $\gamma_m$ . Ausgeschrieben gelangt man zu den Gleichungen

$$\begin{aligned}\sum_{m=-\infty}^{\infty} \left( \alpha_m H_m^{(+)}(nkR(\phi)) + \beta_m H_m^{(-)}(nkR(\phi)) \right) e^{im\phi} &= \sum_{m=-\infty}^{\infty} \gamma_m H_m^{(+)}(kR(\phi)) e^{im\phi} \\ n \sum_{m=-\infty}^{\infty} \left( \alpha_m H_m^{(+)'}(nkR(\phi)) + \beta_m H_m^{(-)'}(nkR(\phi)) \right) e^{im\phi} &= \sum_{m=-\infty}^{\infty} \gamma_m H_m^{(+)'}(kR(\phi)) e^{im\phi}.\end{aligned}$$

Multipliziert man beide Seiten mit  $e^{-in\phi}$  und integriert, so erhält man eine Matrix Gleichung für die Vektoren  $\alpha, \beta$  und  $\gamma$ ,

$$\begin{aligned}\mathcal{H}_1^+ \alpha + \mathcal{H}_1^- \beta &= \mathcal{H}_2^+ \gamma \\ \mathcal{D}\mathcal{H}_1^+ \alpha + \mathcal{D}\mathcal{H}_1^- \beta &= \frac{1}{n} \mathcal{D}\mathcal{H}_2^+ \gamma.\end{aligned}$$

Eliminiert man  $\gamma$ , so gelangt man zu

$$S(k)\alpha = \beta$$

mit einer Matrix  $S(k)$ . Da die Lösung im Ursprung regulär sein muß, folgt  $\alpha = \beta$  und man erhält schließlich an  $k$  die Quantisierungsbedingung,

$$S(k)\alpha = \alpha.$$

Analog zum Randintegral-Ansatz sind für die Matrix  $T(k) := S(k) - I$  die Werte  $k$  zu bestimmen, für die diese Matrix singulär wird.

Ein Vorteil dieses Diskretisierungs-Ansatzes besteht in der physikalische Interpretierbarkeit der Reihendarstellung als Streuungsmatrix. Jedoch verursacht dieser Ansatz schon bei kleinen Störungen des Randes starke numerische Schwierigkeiten. Bei numerischen Experimenten mit diesem Ansatz für das Kartoffel-Billard zeigte sich dieses Problem besonders deutlich. Der Autor konnte nur mit erheblichen Aufwand einige wenige Bouncing-Ball-Moden berechnen. Der Matching-Ansatz ist somit unbrauchbar für größere Deformationen eines integrierbaren Gebiets. Diese Unzulänglichkeit schränkt das Verfahren natürlich auch für die Konstruktion geeigneter Laser-Resonatoren und deren numerischer Simulation ein.

Darüberhinaus ist es sehr schwierig, Fehlerabschätzungen für berechnete Lösungen zu gewinnen. Diese sind aufgrund der Anfälligkeit des Verfahrens für numerische Instabilität von großer Bedeutung, da man in solchen Fällen unbedingt ein Maß für Brauchbarkeit des numerischen Resultates braucht.

### Spezielle Gebiete

Wie bereits erwähnt, läßt sich die Matching-Methode nur für sehr kleine Deformation benutzen, wodurch die Nutzbarkeit sehr eingeschränkt wird. Die in dieser Arbeit vorgestellte Diskretisierung unterliegt solch starken Einschränkungen nicht. Für die Gültigkeit der verwendeten Integralgleichungen muß nur hinreichende Glätte des Randes vorausgesetzt werden. Zur Anwendbarkeit der Methode wurden numerische Simulationen an einem nicht-konvexen Gebiet und einem Gebiet mit einer Ecke durchgeführt. Der Rand des nicht-konvexen Gebietes wird durch

$$\partial\Omega = \{(\cos t + 0.65 \cdot \cos 2t - 0.65, 1.5 \sin t) | 0 \leq t \leq 2\pi\}$$

beschrieben. Als Ecken-Gebiet wurde ein 3/4-Kreis mit aufgesetzter Spitze genommen. Die entstehende Ecke wurde durch einen Kreisbogen ersetzt (vgl. [Wie03]), um die notwendige Glätte von  $\partial\Omega$  zu erreichen.

Für beide Gebiete liefert der Algorithmus hervorragende Ergebnisse, die obigen Fehler- und Konvergenzaussagen ließen sich ebenfalls numerisch bestätigen (vgl. Abbildung 28).

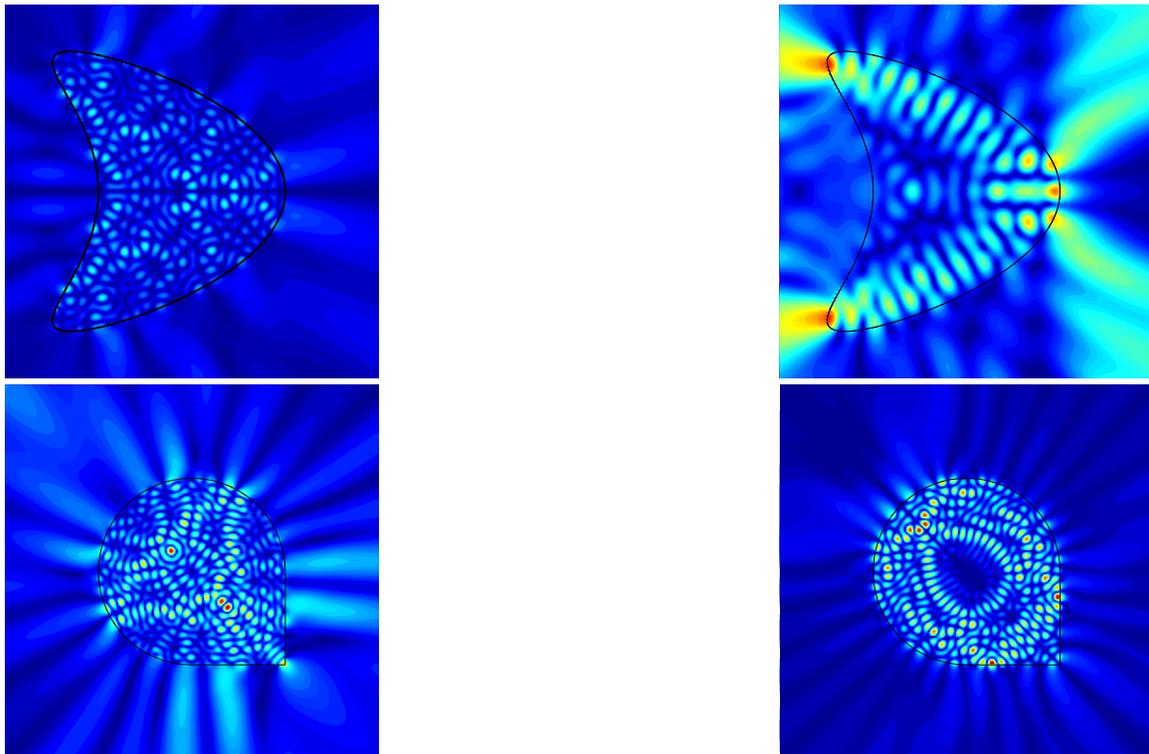


Abbildung 28: Oben: Zwei Eigenmoden eines nicht konvexen Gebietes. Der Brechungsindex beträgt  $n_0 = 3.2$  (links) und  $n_0 = 1.5$  (rechts). Unten: Zwei Eigenmoden eines Gebietes mit Ecke, der Brechungsindex beträgt hier  $n_0 = 3.2$ .

## 4.6 Störung des Gebiets $\Omega$

Von physikalischer Relevanz ist die Entwicklung der Eigenresonanzen unter Störung des Randes.

Die Bestimmung von Eigenresonanzen und Eigenmoden läßt sich wie in den vorherigen Kapiteln beschrieben als nichtlineares Eigenwertproblem  $\tilde{A}(k)x = 0$  formulieren, wobei die Matrix  $\tilde{A}(k)$  von der Diskretisierung und des zu behandelnden Problems abhängt.

Zusätzlich wird nun die Matrix  $\tilde{A}(k)$  abhängig von einem Parameter  $\varepsilon$ , der die Störung des Randes beschreibt. Es muß also das Problem  $\tilde{A}_\varepsilon(k_\varepsilon)x_\varepsilon = 0$  in Abhängigkeit des Parameters  $\varepsilon$  untersucht werden. Ist  $k_0$  ein einfacher Eigenwert zu  $\varepsilon = 0$ , so folgt aus dem Satz über implizite Funktionen, daß es Funktionen  $k(\varepsilon)$  und  $x(\varepsilon)$  gibt, so daß für hinreichend kleine  $\varepsilon$  die Gleichung

$$\tilde{A}_\varepsilon(k(\varepsilon))x(\varepsilon) = 0$$

erfüllt ist.

Dieser Abschnitt beschäftigt sich mit der numerischen Berechnung solcher Eigenwert-Lösungsäste. Der weiter unten vorgeschlagene *Tangenten-Prädiktor* - Algorithmus zusammen mit dem Korrektor 4.11 ermöglicht eine zuverlässige Fortsetzung von Eigenwertästen. In der vorliegenden Arbeit wird ausschließlich der Korrektor 4.11 benutzt, da die auftretenden Matrizen  $A_\varepsilon$  voll besetzt sind. Bei dünn besetzten Matrizen bietet es sich an, andere Korrektor-Verfahren zu benutzen, wie zum Beispiel das Arnoldi- oder Jacobi-Davidson-Verfahren ([Vos04]).

Sei  $\Omega_\varepsilon$  eine Familie von Gebieten und  $\partial\Omega_\varepsilon$  die dazugehörige Familie von Randkurven. Die Abbildung  $\varepsilon \mapsto \partial\Omega_\varepsilon$  sei stetig differenzierbar. Zu jedem  $\varepsilon$  ist das Helmholtz-Problem bzw. das Transmissions-Problem im Gebiet  $\Omega_\varepsilon$  zu lösen. Eine Lösung  $k_\varepsilon$  kann als Näherung für eine Lösung  $k_{\tilde{\varepsilon}}$  mit  $\tilde{\varepsilon} \approx \varepsilon$  dienen. Aus dieser Beobachtung läßt sich sofort ein einfaches Sekanten-Prädiktor Verfahren zur Berechnung der  $k_\varepsilon$  und  $u_\varepsilon$  konstruieren.

### 4.20 Algorithmus (Sekanten-Prädiktor).

- 1 Sei  $k_0$  ein Eigenwert des Helmholtz-Problems zu  $\varepsilon = 0$ .
- 2 Wähle ein  $\Delta\varepsilon$ .
- 3 Bestimme eine Lösung  $k_{\Delta\varepsilon}$  mit Hilfe des Korrektors 4.11 und Startnäherung  $k_0$ .
- 4 Setze  $\varepsilon \leftarrow \Delta\varepsilon$ .
- 5 Bestimme eine Lösung  $k_{\varepsilon+\Delta\varepsilon}$  mit Hilfe des Korrektors 4.11 und Startnäherung  $2k_\varepsilon - k_{\varepsilon-\Delta\varepsilon}$ .
- 6 Setze  $\varepsilon \leftarrow \varepsilon + \Delta\varepsilon$ .

7 Gehe zu Schritt 5, bis ein vorgegebenes  $\varepsilon$  erreicht ist.

Dieser Algorithmus kann verbessert werden, indem die Schrittweite  $\Delta\varepsilon$  dynamisch angepaßt wird. So kann zum Beispiel die Schrittlänge vergrößert werden, wenn der Korrektor eine bestimmte Anzahl von Korrekturschritten nicht überschreitet, und sie kann reduziert werden, falls eine bestimmte Anzahl von Korrekturschritten überschritten wird.

Im Algorithmus 4.20 wird der Differenzenquotient  $\frac{k_\varepsilon - k_{\varepsilon - \Delta\varepsilon}}{\Delta\varepsilon}$  zur Approximation benutzt. Eine Verbesserung der Approximation wird erzielt, indem der Differenzenquotient durch den Differentialquotienten  $\frac{dk_\varepsilon}{d\varepsilon}$  ersetzt wird. Als Näherung wird dann  $k_\varepsilon + \frac{dk_\varepsilon}{d\varepsilon} \cdot \Delta\varepsilon$  verwendet. Zu einer vorgegebenen Diskretisierung  $N$  sei  $\tilde{A}_\varepsilon(k)$  die Diskretisierungsmatrix und das Paar  $k_\varepsilon, x_\varepsilon$  erfülle die Gleichung

$$\tilde{A}_\varepsilon(k_\varepsilon)x_\varepsilon = 0. \quad (4.49)$$

Sei weiter  $\tau_\varepsilon$  ein Linkseigenvektor von  $\tilde{A}_\varepsilon(k_\varepsilon)$ . Differenziert man Gleichung (4.49) nach  $\varepsilon$  und multipliziert von links mit  $\tau_\varepsilon^*$ , so folgt

$$\tau_\varepsilon^* \left\{ \left( \frac{\partial \tilde{A}_\varepsilon(k_\varepsilon)}{\partial \varepsilon} + \frac{\partial \tilde{A}_\varepsilon(k_\varepsilon)}{\partial k_\varepsilon} \frac{dk_\varepsilon}{d\varepsilon} \right) x_\varepsilon + \tilde{A}_\varepsilon(k_\varepsilon)x'_\varepsilon \right\} = 0.$$

Zur Abkürzung seien

$$\begin{aligned} \frac{\partial \tilde{A}_\varepsilon(k_\varepsilon)}{\partial \varepsilon} &= \mathcal{A} + i\mathcal{B} \\ \frac{\partial \tilde{A}_\varepsilon(k_\varepsilon)}{\partial k_\varepsilon} &= \mathcal{C} + i\mathcal{D}. \end{aligned}$$

Diese Matrizen können jeweils durch einen Differenzenquotienten numerisch bestimmt werden. Dann folgt nach Aufspalten in Real- und Imaginärteil die Gleichung

$$\begin{pmatrix} \tau_\varepsilon^* \mathcal{C} \tilde{\phi}_\varepsilon & -\tau_\varepsilon^* \mathcal{D} \tilde{\phi}_\varepsilon \\ \tau_\varepsilon^* \mathcal{D} \tilde{\phi}_\varepsilon & \tau_\varepsilon^* \mathcal{C} \tilde{\phi}_\varepsilon \end{pmatrix} \begin{pmatrix} \operatorname{Re} \frac{dk_\varepsilon}{d\varepsilon} \\ \operatorname{Im} \frac{dk_\varepsilon}{d\varepsilon} \end{pmatrix} = - \begin{pmatrix} \tau_\varepsilon^* \mathcal{A} \tilde{\phi}_\varepsilon \\ \tau_\varepsilon^* \mathcal{B} \tilde{\phi}_\varepsilon \end{pmatrix}. \quad (4.50)$$

Aus diesen Ergebnissen läßt sich ein *Tangenten-Prädiktor-Verfahren* herleiten.

#### 4.21 Algorithmus (Tangenten-Prädiktor-Algorithmus).

- 1 Sei  $k_0$  ein Eigenwert des Helmholtz-Problems zu  $\varepsilon = 0$ .
- 2 Wähle ein  $\Delta\varepsilon$ .
- 3 Setze  $\varepsilon \leftarrow 0$ .
- 4 Bestimme  $\left. \frac{dk_\varepsilon}{d\varepsilon} \right|_\varepsilon$  mit Gleichung (4.50) .

- 5 Bestimme eine Lösung  $k_{\varepsilon+\Delta\varepsilon}$  mit Hilfe des Korrektors 4.11 und Startnäherung  $k_\varepsilon + \Delta\varepsilon \cdot \left. \frac{dk_\varepsilon}{d\varepsilon} \right|_\varepsilon$ .
- 6 Setze  $\varepsilon \leftarrow \varepsilon + \Delta\varepsilon$ .
- 7 Gehe zu Schritt 4, bis ein vorgegebenes  $\varepsilon$  erreicht ist.

Für diesen Algorithmus ist ebenfalls eine Schrittweitensteuerung möglich und naheliegend.

Für die Schrittweitensteuerung kann man eine kleine Erweiterung vornehmen. Der Algorithmus berechnet in Schritt 5 mit Hilfe des Korrektors sowohl  $k_{\varepsilon+\Delta\varepsilon}$  als auch  $x_{\varepsilon+\Delta\varepsilon}$ . O.B.d.A. dürfen die Vektoren  $x_\varepsilon$  und  $x_{\varepsilon+\Delta\varepsilon}$  als Einheitsvektoren bezüglich des Standardskalarprodukts angenommen werden. In diesem Fall muß dann

$$|\langle x_\varepsilon, x_{\varepsilon+\Delta\varepsilon} \rangle| \approx 1$$

gelten. Weicht dieser Wert stark von 1 ab, so hat der Korrektor in Schritt 5 einen falschen Wert geliefert. In diesem Fall muß Schritt 5 mit einer kleineren Schrittweite  $\Delta\varepsilon$  wiederholt werden. Umgekehrt, ist der Wert sehr nahe bei 1 darf die Schrittweite erhöht werden. Mit dieser Strategie wurden in Kapitel 5.4.2 sehr gute Ergebnisse erzielt.

#### 4.7 Implementierung und CPU-Zeit

Die oben vorgestellten Algorithmen wurden in MATLAB implementiert. Gravierende Vorteile gegenüber einer C- oder Fortran-Implementierung bestehen in die Verfügbarkeit von optimierten LAPACK-Routinen und der automatische Speicherplatzverwaltung von MATLAB. Dieser Vorteil wird durch eine (in CPU-Zeit gemessene) langsamere Berechnung der Routinen erkaufte. Die unten angegebenen CPU-Zeiten wurden auf einem Intel Pentium(R) 4 mit 3.06 GHz, 512 MB-RAM und MATLAB 6.1.0.45 Release 12.1 auf einem Windows 2000 System gemessen.

Die berechneten Eigenfunktionen können mit Hilfe der in MATLAB vorhandenen Graphikfunktionen dargestellt werden. Dadurch erspart man sich eine aufwendige Transformation von Daten etwa in das von Gnuplot verwendete Format. Für die Entwicklung neuer Algorithmen erweist sich MATLAB deshalb am besten geeignet.

Im Fall (HP) erzeugt der Algorithmus (4.14) vollbesetzte Matrizen aus  $\mathbb{C}^{2N \times 2N}$  und im Fall (TP) vollbesetzte Matrizen aus  $\mathbb{C}^{4N \times 4N}$ , was einen enormen Speicherplatzbedarf zur Folge hat. Hier liegt eine Beschränkung des Algorithmus, denn sobald der Speicherbedarf den vorhandenen Hauptspeicher übersteigt, müssen Daten auf der Festplatte zwischen gespeichert werden und die Laufzeit sinkt rapide ab.

Zur Berechnung einer Eigenresonanz mit Hilfe des Korrektors (4.11) werden nur 2 *LU*-Zerlegungen benötigt, für die die LAPACK-Routine ZGETRF verwendet wird. Die meiste CPU-Zeit wird

zur Assemblierung der jeweiligen Diskretisierungsmatrix  $\tilde{A}(k)$  benutzt, da  $O(N^2)$  viele (teure) Besselfunktionen-Auswertungen berechnet werden müssen. Im Fall (HP) benötigte MATLAB mit  $N = 300$  für eine Assemblierung 1.859 Sekunden, mit  $N = 600$  für die Assemblierung 7.343 Sekunden, also wie erwartet 4 mal so lange. Der Speicherplatz-Effekt wird deutlich bei einer Diskretisierung mit  $N = 900$ , hier benötigt MATLAB zur Assemblierung 52.515 Sekunden, also etwa 29 mal so lange wie für den Fall  $N = 300$ . Zum Vergleich: für eine  $LU$ -Zerlegung einer assemblierten Diskretisierungsmatrix benötigt MATLAB mit  $N = 300$  nur 0.531 Sekunden.

Mit der oben beschriebenen Update-Strategie des Korrektors werden im Durchschnitt 2 Assemblierungen vor dem Update und 4 Assemblierungen nach dem Update benötigt, um eine Genauigkeit auf 6 Nachkomma-Stellen der Referenzbeispiele aus Kapitel 3 zu erreichen. Zusammen mit den 2  $LU$ -Zerlegungen entspricht dies bei  $N = 600$  etwa einer Rechenzeit von 48 Sekunden. Um diese Genauigkeit zu erreichen, benötigt der herkömmliche  $SVD$ -Sweep etwa 100 Testfunktionsauswertungen, d.h. 100 Assemblierungen und zusätzlich 100  $SVD$ -Auswertungen, für die jeweils 22 Sekunden benötigt werden. Insgesamt benötigt der  $SVD$ -Sweep somit ca. 49 Minuten.

Zur Berechnung von Lösungen des Transmissions-Problems werden ebenfalls im Durchschnitt nur 6 Assemblierungen und 2  $LU$ -Zerlegungen benötigt. Im Gegensatz hierzu wächst der Aufwand des  $SVD$ -Sweeps jedoch quadratisch, da nun ein 2-dimensionales Gebiet durchsucht werden muß. Es würden dann etwa 40 Stunden benötigt, um eine vergleichbare Genauigkeit zu erhalten.

Zusammen mit der Prädiktor-Strategie entfaltet der Algorithmus seine volle Effizienz. Im Fall (HP) wurden bei einer Beispielrechnung mit  $N = 600$  zur Berechnung des Prädiktors 395 Sekunden benötigt. Anschließend konnten dann 71 Eigenwerte mit jeweils 2  $LU$ -Zerlegungen und weniger als 6 Assemblierungen berechnet werden. Jeder der 71 Eigenwerte war auf 6 Nachkommastellen genau. Für diese Berechnung benötigt der direkte Sweep ca. 63 Minuten. Ein  $SVD$ -Sweep mit vergleichbarer Genauigkeit würde 57 Stunden benötigen.

Für den Fall (TP) werden mit  $N = 300$  für die Berechnung des Prädiktors ebenfalls 395 Sekunden benötigt. Bei einer Beispielrechnung mit dem Ei-Billard konnten dann 8 Eigenwerte mit jeweils 2  $LU$ -Zerlegungen und 6 Assemblierungen auf 6 Nachkommastellen berechnet werden. Dies entspricht einer Rechenzeit von ca. 13 Minuten. Mit einem herkömmlichen  $SVD$ -Sweep ist solch eine Berechnung (in dieser Genauigkeit) gar nicht möglich.

## 5 Eine neue Klasse von Microlasern

Die Existenz von  $p$ -Lazutkin-Billardards und deren numerische Behandlung bilden die Grundlage für die Konstruktion einer neuen Klasse von Microlasern vom sogenannten *Quanten-Kaskaden-Lasertyp*. Die in dieser Arbeit entwickelten algorithmischen Ideen können zu einer vollständigen Simulation der physikalischen Verhältnisse und der Eigenschaften solcher Laser verwendet werden.

Zum besseren Verständnis wird der mathematischen Behandlung eine kurze Synopse der physikalischen Eigenschaften und Beschreibung von Microlasern vorangestellt.

### 5.1 Microlaser - Einführung und kurze historische Übersicht

Laser bestehen im wesentlichen aus einem *aktiven Material*, in dem Licht einer bestimmten Wellenlänge durch eine externe Energiequelle erzeugt wird, und dem sogenannten *Resonator*, der das aktive Material enthält und für stimulierte Lichtemission sorgt.

Der Resonator bestimmt die Eigenschaften des emittierten Lichts, wie Leistung, Direktionalität, Intensität und spektrale Eigenschaften. Außerdem bestimmt er die Größe und Form des Lasers. Die heute in Wissenschaft und Technik meist gebrauchten Laser sind Halbleiterlaser, bei denen der Resonator die Form eines sogenannten *Fabry-Perot-Käfigs* hat, in dem zwei Halbleiterplatten als parallele Spiegel dienen und das Licht durch das aktive Material hindurch hin und her reflektieren. Die Optimierung der Resonatoreigenschaften, d.h. die Konstruktion von Resonatoren mit bestimmten geforderten Eigenschaften, ist seit der Entwicklung der ersten Laser ein aktives Gebiet der Forschung.

Ein Meilenstein der neueren Forschung ist die Entwicklung der Microdisk-Halbleiterlaser ([YS93], [MLS<sup>+</sup>92]), deren Prinzip auf der totalen inneren Reflektion von Licht beruht. Diese Microdisk-, -Zylinder und -Droplet-Laser besitzen einen kreisförmigen symmetrischen Resonator, dessen Emissions-Moden vom Typ der Whispering-Gallery-Moden des elektromagnetischen Feldes sind. In solchen Moden läuft das Licht entlang des Randes und wird an den Rändern des Resonators in einem Winkel reflektiert, der unterhalb des kritischen Winkels für die totale innere Reflektion liegt. Das Licht bleibt auf diese Weise in dem Resonator gefangen. Kleine Verluste werden nur durch Unebenheiten der Resonatoroberfläche und quantenmechanische Tunneleffekte hervorgerufen. Dieses Reflektionsprinzip ermöglichte die Herstellung der bis heute kleinsten Laser ([MLS<sup>+</sup>92]). Diese winzigen Laser bilden fundamentale Komponenten für die Entwicklung von optischen Computern und Netzwerken.

Große Nachteile der Microdisk-Laser waren ursprünglich die schwache Output-Leistung und die fehlende Direktionalität des emittierten Lichtes. Diese Nachteile konnten durch eine quadrupo-

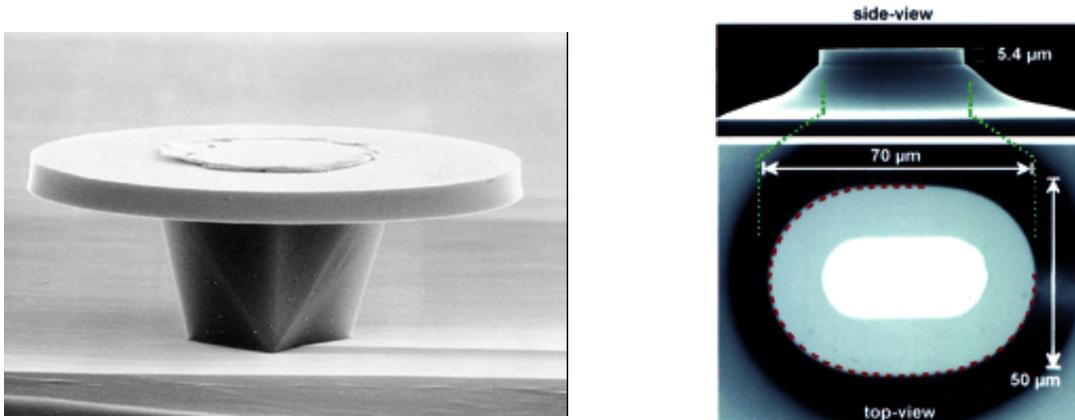


Abbildung 29: Links: Schrägaufsicht eines Microlasers unter dem Elektronen Mikroskop. Der Resonator ist die obere Scheibe. Rechts: Seitenaufsicht eines Bow-Tie-Lasers. Der Resonator befindet sich an der Spitze des Kegels. (Quelle: Bell-Labs)

lare Deformation des kreisförmigen Resonators umgangen werden ([GCN<sup>+</sup>98]), bei der die dem emittierten Licht zugrunde liegende Mode zu einem 4-periodischen Orbit gehört, der die Gestalt einer Smoking-Fliege hat. Der Laser-Typ wird deshalb auch *Bow-tie-Laser* genannt.

Ziel dieser Arbeit ist die Simulation der physikalischen Eigenschaften von Lasern, deren Resonatoren *p*-Lazutkin-Billardards sind. Am Ende des Kapitels werden die Ergebnisse der numerischen Simulationen mit experimentellen Daten verglichen, die von Wolfgang Parz an entsprechend gefertigten Lasern gewonnen wurden ([BHZP04]).

## 5.2 Physikalisches Modell von Halbleiter-Lasern

Die verwendeten Halbleiterlaser sind sogenannte *Quantenkaskaden-Laser*. Dabei ist der Resonator ein Zylinder, dessen Querschnitt die physikalischen Eigenschaften bestimmt. So wird zum Beispiel für den Whispering-Gallery-Laser ein kreisförmiger Querschnitt verwendet, für den Bow-tie-Laser gerade eine quadrupolare Deformation des Kreises. Die Höhe des Zylinders beträgt etwa  $6\mu m$  und der mittlere Durchmesser liegt zwischen  $75\mu m$  und  $150\mu m$ . Der Zylinder selbst besteht aus wechselnden Schichten von Halbleitermaterial InGaAs und InAlAs.

Elektronen fallen von Schicht zu Schicht und geben dabei jeweils eine gewisse Energie-Menge ab. Die Schichten sind so bestimmt, daß immer ein Vielfaches einer Frequenz abgestrahlt wird. Die Wellenlänge des Lasers wird durch die unterschiedlichen Dicken der Materialschichten bestimmt, so kann ein Wellenlängenbereich von  $3.4\mu m$  bis  $17\mu m$  erreicht werden.

Der Querschnitt des Zylinders wird im folgenden mit  $\Omega$  bezeichnet. Der effektive Brechungsindex des Halbleitermaterials sei  $n_0 > 1$ . Die den Laser umgebende Luft hat den Brechungsindex  $n = 1$ . Im Falle von InGaAs und InAlAs Schichten ist  $n_0 = 3.3$ .

Der Querschnitt  $\Omega \subset \mathbb{R}^2$  sei als konvex und sternförmig angenommen. O.B.d.A. darf der Sternpunkt als Koordinatenursprung angenommen werden.  $R = \Omega \times h$  beschreibt den Resonator im 3-dimensionalen Raum. Die Funktion

$$n(x) = \begin{cases} n_0 & : x \in R \\ 1 & : x \notin R \end{cases}$$

gibt den Brechungsindex im Raumpunkt  $x = (x_1, x_2, x_3)$  an.

Die von dem Laser erzeugten elektromagnetischen Felder werden durch die Maxwell'schen Gleichungen beschrieben. Man unterscheidet zwischen einem gestreuten äußeren Feld und einem Feld innerhalb des Resonators. Es wird später gezeigt, daß beide Felder über die diskontinuierliche Grenze des Dielektrikums hinweg stetig differenzierbar sein müssen. Man darf annehmen, daß alle Felder die Zeitabhängigkeit  $e^{-i\omega t}$  haben. Sei also

$$\mathbf{E}(t, x) = e^{-i\omega t} \cdot \begin{pmatrix} E^1(x) \\ E^2(x) \\ E^3(x) \end{pmatrix}$$

das elektrische Feld und

$$\mathbf{H}(t, x) = e^{-i\omega t} \cdot \begin{pmatrix} H^1(x) \\ H^2(x) \\ H^3(x) \end{pmatrix}$$

das zugehörige magnetische Feld.

Die Maxwell'schen Gleichungen liefern dann ([LL71])

$$\begin{aligned} \nabla \times \mathbf{E} &= ik\mathbf{H} \\ \nabla \times \mathbf{H} &= -ikn(x)^2\mathbf{E}, \end{aligned}$$

mit Wellenzahl

$$k = \frac{\omega}{c}. \tag{5.1}$$

In Gebieten mit konstantem Brechungsindex gilt

$$\nabla \cdot \mathbf{E} = 0,$$

zusammen mit der Formel

$$\nabla^2 \mathbf{F} = \nabla(\nabla \cdot \mathbf{F}) - \nabla \times \nabla \times \mathbf{F},$$

folgt aus den Maxwellschen Gleichungen

$$\nabla^2 \mathbf{E} + k^2 n(x)^2 \mathbf{E} = 0.$$

Man kann annehmen, daß das elektrische Feld stets exakt parallel zur  $x_3$ -Achse ist ([GCN<sup>+</sup>98]), so daß die Komponentenfunktionen  $E^i$  nicht von  $x_3$  abhängen. Die Ausbreitungsrichtung ist also orthogonal zur Zylinderachse. Man spricht auch von *TM-Polarisation*, da es sich um ein zur Zylinderachse transversales Magnetfeld handelt.

Betrachtet man nur die  $x_3$ -Komponente und schreibt  $E(x, y) = E^3(x_1, x_2)$ , so ist die Gleichung

$$\Delta E + k^2 n(x, y)^2 E = 0 \tag{5.2}$$

in  $\mathbb{R}^2$  zu lösen, wobei  $n(x, y)$  durch

$$n(x, y) = \begin{cases} n_0 & : (x, y) \in \Omega \\ 1 & : (x, y) \in \mathbb{R}^2 \setminus \Omega \end{cases}$$

gegeben ist.

Wegen Gleichung (5.1) hat das Feld  $E(x, y)$  die Zeitabhängigkeit  $e^{-ickt}$ . Setzt man  $k = \operatorname{Re} k + i \cdot \operatorname{Im} k$  ein, so folgt, daß sich die Amplitude des Feldes  $E$  wie  $e^{c \operatorname{Im} kt}$  mit der Zeit verhält. Da stets  $\operatorname{Im} k < 0$  ist, ist  $|\operatorname{Im} k|$  ein Maß für die Lebenszeit des Feldes.

Als nächstes werden die passenden Randbedingungen hergeleitet.

In Zylinderkoordinaten  $(r, \phi, x_3)$  ist  $\nabla \times \mathbf{E}$  gegeben durch

$$\nabla \times \mathbf{E} = \begin{vmatrix} \vec{r} & \vec{\phi} & \vec{x}_3 \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial \phi} & \frac{\partial}{\partial x_3} \\ E^{(r)} & rE^{(\phi)} & E \end{vmatrix}.$$

Somit folgt insbesondere für die  $\phi$ -Komponente

$$(\nabla \times \mathbf{E})^{(\phi)} = \frac{\partial E^{(r)}}{\partial x_3} - \frac{\partial E}{\partial r} = -\frac{\partial E}{\partial r}.$$

Aus den ersten Maxwell-Gleichungen folgt dann

$$\mathbf{H}^{(\phi)} = -\frac{i}{k} (\nabla \times \mathbf{E})^{(\phi)} = \frac{i}{k} \frac{\partial E}{\partial r}.$$

Da  $\mathbf{H}$  stetig sein muß, muß auch die Ableitung  $\frac{\partial E}{\partial r}$  stetig sein. Diese Bedingung bedeutet, daß die Normalen-Ableitung von  $E(x, y)$  am Rand  $\partial\Omega$  stetig sein muß.

Mathematisch ist somit das in Kapitel 4 behandelte quasi-klassische Transmissions-Problem zu lösen:

Gesucht sind Funktionen  $u_0 \in \mathcal{C}^2(\Omega) \cap \mathcal{C}(\bar{\Omega})$  und  $u \in \mathcal{C}^2(\mathbb{R}^2 \setminus \bar{\Omega}) \cap \mathcal{C}(\mathbb{R}^2 \setminus \Omega)$  und  $k \in \mathbb{C}$  mit  $|k| \gg 1$ , so daß  $u_0$  und  $u$  die Helmholtz-Gleichungen

$$\Delta u_0 + (kn_0)^2 u_0 = 0 \quad \text{in } \Omega \quad (5.3)$$

$$\Delta u + k^2 u = 0 \quad \text{in } \mathbb{R}^2 \setminus \bar{\Omega} \quad (5.4)$$

mit den Randbedingungen

$$u_0 - u = 0 \quad (5.5)$$

$$\frac{\partial u_0}{\partial \nu} - \frac{\partial u}{\partial \nu} = 0 \quad \text{auf } \partial\Omega \quad (5.6)$$

erfüllt. Zusätzlich soll für  $|x| \rightarrow \infty$  die *Sommerfeld'sche*-Ausstrahlungsbedingung

$$u(x) = O\left(\frac{1}{r}\right), \quad \frac{\partial u}{\partial r} - iku = o\left(\frac{1}{r}\right), \quad \text{für } r = |x| \rightarrow \infty \quad (5.7)$$

erfüllt sein. Die Bedingung (5.7) beschreibt auslaufende Wellen ([SS04]).

### 5.3 Strahlenoptisches Modell und Konstruktion geeigneter asymmetrischer Resonanzkäfte

Wie bereits in Kapitel 4 hervorgehoben ist dieses Transmissions-Problem nur für sehr spezielle Gebiete  $\Omega$  analytisch lösbar, so daß man sich auf numerische Lösungen beschränken muß.

Der Gleichungstyp (5.3) beschreibt die Ausbreitung einer Lichtwelle in einem Medium, das in  $\Omega$  den Brechungsindex  $n_0$  und sonst den Index 1 hat. Bei konstantem Brechungsindex breiten sich die Wellenfronten längs gerader Linien aus. Das Ausbreitungsverhalten dieser Lichtstrahlen ist Gegenstand der geometrischen oder Strahlen-Optik. Nimmt man für kleine Wellenlängen, also große  $|k|$  die Lösung der Gleichung (5.3) in der Form

$$u(x) = A(x)e^{ikn(x)S(x)}$$

an, so gilt für  $S$  die *Eikonal-Gleichung*

$$\|\nabla S\|^2 = n^2$$

der geometrischen Optik.

Am Übergang zweier Medien mit konstantem Brechungsindex wird der Strahl entweder reflektiert oder er dringt in das andere Medium ein. Das Verhalten des Strahls ist abhängig vom Inzidenz-Winkel  $\chi$  zwischen Strahl und der Normale am Übergang. Der Winkel

$$\chi_0 = \arcsin \frac{1}{n}$$

ist der Winkel für totale Reflexion; ist der Inzidenz-Winkel  $\chi$  kleiner als  $\chi_0$ , so verläßt der Strahl das Medium, ist der Winkel größer, so wird der Strahl reflektiert. Dies gilt exakt nur für ebene Wellen und für einen unendlich langen und glatten Übergang, jedoch kann diese Beschreibung als Näherung verwendet werden. Man vernachlässigt durch diese Näherung die sogenannten *Tunnel-Effekte*.

Diese Dynamik läßt sich sehr einfach im Poincaré-Schnitt des Billards veranschaulichen. Es muß nur die Linie, die dem Winkel  $\chi_0$  entspricht, eingezeichnet werden. Ein Strahl bleibt solange im Gebiet  $\Omega$  gefangen, bis die  $\theta$ -Komponente im Poincaré-Schnitt der zugehörigen Trajektorie oberhalb dieser Linie verläuft. Da im klassischen Limit (hier  $\text{Re } k \rightarrow \infty$ ) das Verhalten der Lösungen der Maxwell-Gleichung durch das Verhalten der klassischen Trajektorien beschrieben werden, kann man ähnlich wie im Dirichlet-Fall aus den Eigenschaften des Poincaré-Schnitts Rückschlüsse auf die Lösungsstruktur bestimmter Lösungen der Maxwell- bzw. Helmholtz-Gleichung mit obigen Randbedingungen schließen. Eine rigorose mathematische Beschreibung ähnlich dem Dirichlet-Fall fehlt jedoch zur Zeit und ist Gegenstand aktueller Forschung.

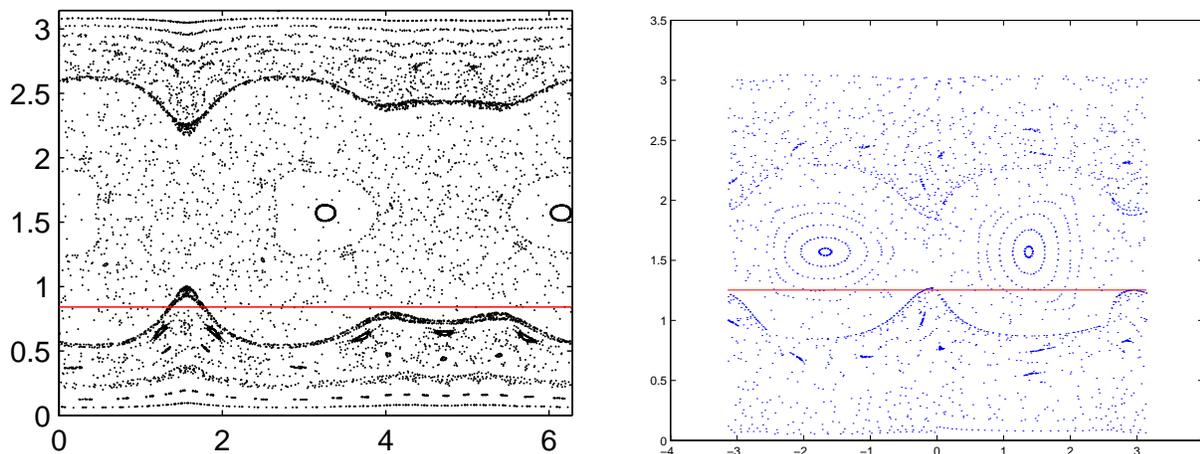


Abbildung 30: Links: Brechungsindexlinie mit  $n_0 = 1.5$  im Phasenportrait des Ei-Billards. Rechts: Brechungsindexlinie mit  $n_0 = 3.3$  im Phasenportrait des Kartoffel-Billards.

#### 5.4 Numerische Konstruktion von IWG-Lasern

In Kapitel 1 wurde gezeigt, daß ein numerisch konstruierter  $p$ -Lazutkin-Billard einen KAM-Streifen besitzen, vgl. Eigenschaft (B7). Sei  $\Omega$  nun ein  $p$ -Lazutkin-Billard und  $R = \Omega \times [0, h]$  der zugehörige Resonator. Sei  $\Gamma = \{(s, \theta(s)) | s \in [0, L]\} \subset C$  die zur  $p$ -periodischen Kaustik gehörende invariante Kurve im Phasenraum, entlang der sich der KAM-Streifen geformt hat. Im strahlenoptischen Modell entspricht dies der Eigenschaft, daß ein Lichtstrahl, der den Resonator

im Randpunkt  $s$  unter dem Winkel  $\theta \approx \theta(s)$  trifft, bei folgenden Reflektionen im Randpunkt  $s'$  unter dem Winkel  $\theta \approx \theta(s')$  reflektiert wird.

Mit Hilfe der in Kapitel 2 vorgestellten Konstruktionsmethoden wurde ein  $p$ -Lazutkin-Billard konstruiert, für das ein kleiner Teil des resultierenden KAM-Streifens oberhalb der Brechungslinie zu einem vorgegebenen Index lag.

Zwei Beispiele wurden im Verlauf dieser Arbeit öfters vorgestellt: Das Ei-Billard ist für einen Brechungsindex  $n_0 = 1.5$  konstruiert, das Kartoffel-Billard für einen Index  $n_0 = 3.3$ , vgl. Abbildung 30.

Das strahlenoptische Modell fordert nun, daß ein Lichtstrahl, dessen Reflektionswinkel am Rand oberhalb des kritischen Winkels liegt, an dieser Stelle den Resonator verläßt. Nach Konstruktion geschieht das für die konstruierten  $p$ -Lazutkin-Billards an den Stellen, an denen die jeweiligen KAM-Streifen oberhalb der Brechungslinie liegen.

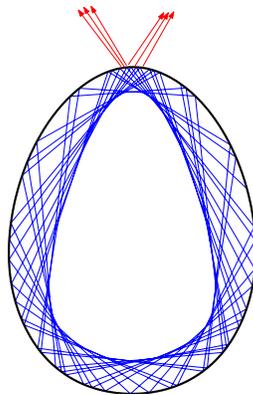


Abbildung 31: Gemäß dem strahlenoptischen Modell verläßt ein Lichtstrahl den Resonator, falls der Reflektionswinkel über dem kritischen Winkel liegt.

Mit den numerischen Hilfsmitteln aus Kapitel 4 ist es nun möglich, diese approximative Vorhersage konkret zu verifizieren. Ähnlich wie in Kapitel 3 müssen die Lösungen von (TP), die der oben beschriebenen strahlenoptischen Beschreibung entsprechen, entlang der  $p$ -periodischen Kaustik im Inneren von  $\Omega$  lokalisiert sein, d.h. die innere Lösung  $u_0(x)$  ist eine innere Whispering-Gallery-Mode. Ferner muß das äußere Feld der Lösung, also die Funktion  $u(x)$  an den Stellen von  $\partial\Omega$ , an denen das KAM-Gebiet oberhalb der Brechungslinie liegt, gerichtet von  $\partial\Omega$  abstrahlen. Der Winkel mit der Tangente ist um so größer, je weiter der KAM-Streifen oberhalb der Brechungslinie liegt.

Für das Ei-Billard erwartet man dann, daß das Licht bei ungefähr  $\phi = \frac{\pi}{2}$  den Resonator verläßt, bei dem Kartoffel-Laser bei  $\phi = 0$  und  $\phi = \pi$ .

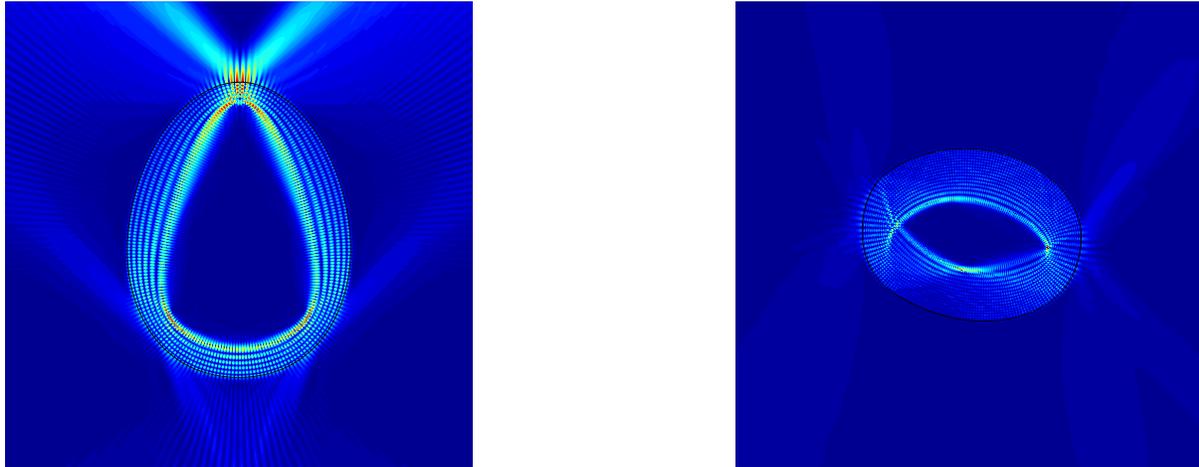


Abbildung 32: Links: Simulation des Ei-Billards mit  $n_0 = 1.5$ . Im Inneren erkennt man deutlich die IWG, an der Spitze tritt das Licht gerichtet aus dem Resonator. Rechts: Simulation des Kartoffel-Billards mit  $n = 3.14$ . Im Inneren erkennt man wiederum die IWG. Das Licht tritt rechts und links am Resonator aus.

#### 5.4.1 Berechnung des quasi-klassischen Spektrums

Mit Hilfe des direkten Sweep-Algorithmus und der Diskretisierung für (TP) wurden Teile des quasi-klassischen Spektrums des Ei- bzw. Kartoffel-Billards berechnet. Für das Ei-Billard wurde der Brechungsindex  $n_0 = 1.5$  benutzt, für das Kartoffel-Billard  $n_0 = 3.14$ . In beiden Fällen wurde mit  $N = 400$  diskretisiert.

In beiden Fällen wurde  $\text{Re } k$  so gewählt, daß die berechneten Resonanzen auf 6 Stellen hinter dem Komma korrekt sind. Auffallend ist im Ei-Spektrum die Regelmäßigkeit, die oberen 4 Eigenwerte entsprechen Whispering-Gallery-Moden am Rand, die Moden mit  $\text{Im } k \approx -0.03$  entsprechen den inneren Whispering-Gallery-Moden. Die Moden mit  $\text{Im } k \approx -0.05$  entsprechen Bouncing-Ball-Moden.

Das Spektrum des Kartoffel-Billards weist keine so deutlichen Strukturen auf. Die Eigenmoden mit  $\text{Im } k \approx 0$  entsprechen ebenfalls Whispering-Gallery-Moden am Rand. Der untere Bereich entspricht Bouncing-Ball-Moden. Die Resonanzen zwischen diesen beiden Gebieten kann man den inneren Whispering-Gallery-Moden zuordnen.

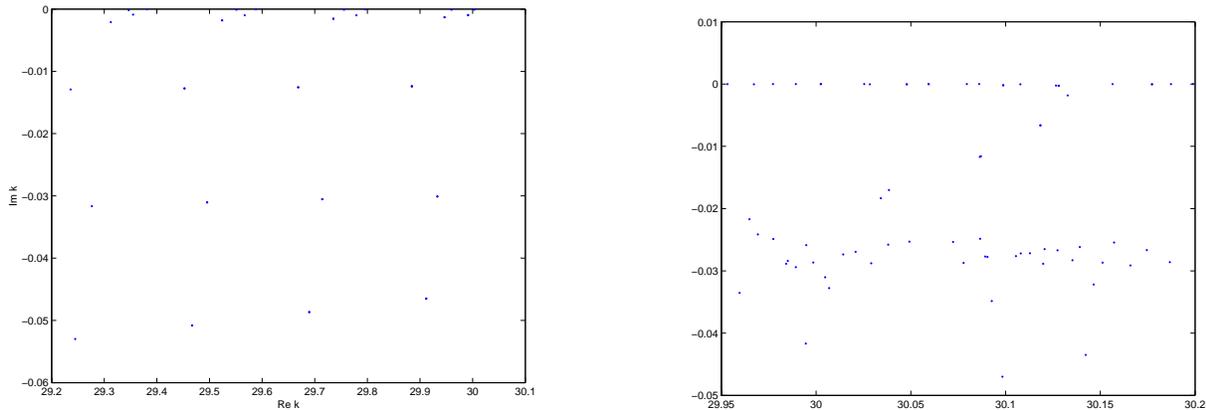


Abbildung 33: Links: Ausschnitt aus dem Spektrum des Ei-Billard. Rechts: Ausschnitt aus dem Spektrum des Kartoffel-Billard.

#### 5.4.2 Störung der Eingangsdaten

Von physikalischem und mathematischem Interesse ist unter anderem das Verhalten der Eigenresonanzen unter kleinen (stetigen) Änderungen im zugrunde liegenden Gebiet  $\Omega$  und bei Veränderungen im Brechungsindex des Resonators  $n_0$ . Der erste Fall ist interessant, da bei der technischen Fertigung der Halbleiter-Laser stets Ungenauigkeiten gemacht werden. Die Abweichung von Endprodukt und Vorlage liegt zwischen 2 und 5 Prozent. Der Brechungsindex hängt von dem verwendeten Halbleitermaterial ab, so daß man durch eine Analyse der Abhängigkeit von  $n_0$  das Verhalten mit anderen Materialien untersuchen kann.

In Kapitel 4 wurde bereits ein Algorithmus zur Eigenwert-Kontinuation unter Störung des Randes vorgestellt. Der dort vorgestellte Algorithmus kann ohne Veränderung übernommen werden. Der Algorithmus entfaltet für das hier untersuchte Problem erst richtig seine Effizienz, da nun erst die in Kapitel 4 weit gefaßten Voraussetzungen an die Matrix  $\tilde{A}$  benötigt werden. Algorithmus 4.21 ermöglicht eine Eigenwert-Kontinuation des nichtlinearen Transmissions-Problems.

Numerische Ergebnisse am Ei- bzw. Kartoffel-Billard deuten darauf hin, daß die inneren Whispering-Gallery Moden sehr robust unter kleinen Störungen sind und diese spezielle Modenstruktur durch Störungen in der Fertigung nicht zerstört werden. Dieses Ergebnis läßt sich auch mit der KAM-Theorie und dem strahlenoptischen Modell erklären, da das für die IWG verantwortliche Gebiet im Poincaré-Schnitt durch kleine Störungen entsteht und für kleine Störungen erhalten bleibt. Werden die Störungen jedoch größer, wird die IWG zerstört, das Gebiet im Poincaré-Schnitt wird ebenfalls zerstört. Die Output-Leistung des Lasers kann durch

$$I(\phi) = |u(R, \phi)|^2$$

simuliert werden. Hierbei wird das Quadrat des Absolutwert des Außenfeldes in einem festen Abstand  $R$  und Polarwinkel  $\phi$  außerhalb des Resonators bestimmt. Schon durch kleine Störungen des Randes fällt dieser Wert ab. Obwohl die Moden-Struktur im Inneren des Resonators trotz kleiner Störungen erhalten bleibt, ändert sich die Intensität des äußeren (meßbaren) Feldes. Insofern ist die Gestalt des Billard mit  $p$ -periodischer Kaustik optimal, da kleine Störungen im Rand zu einer Verringerung der Output-Leistung führen.

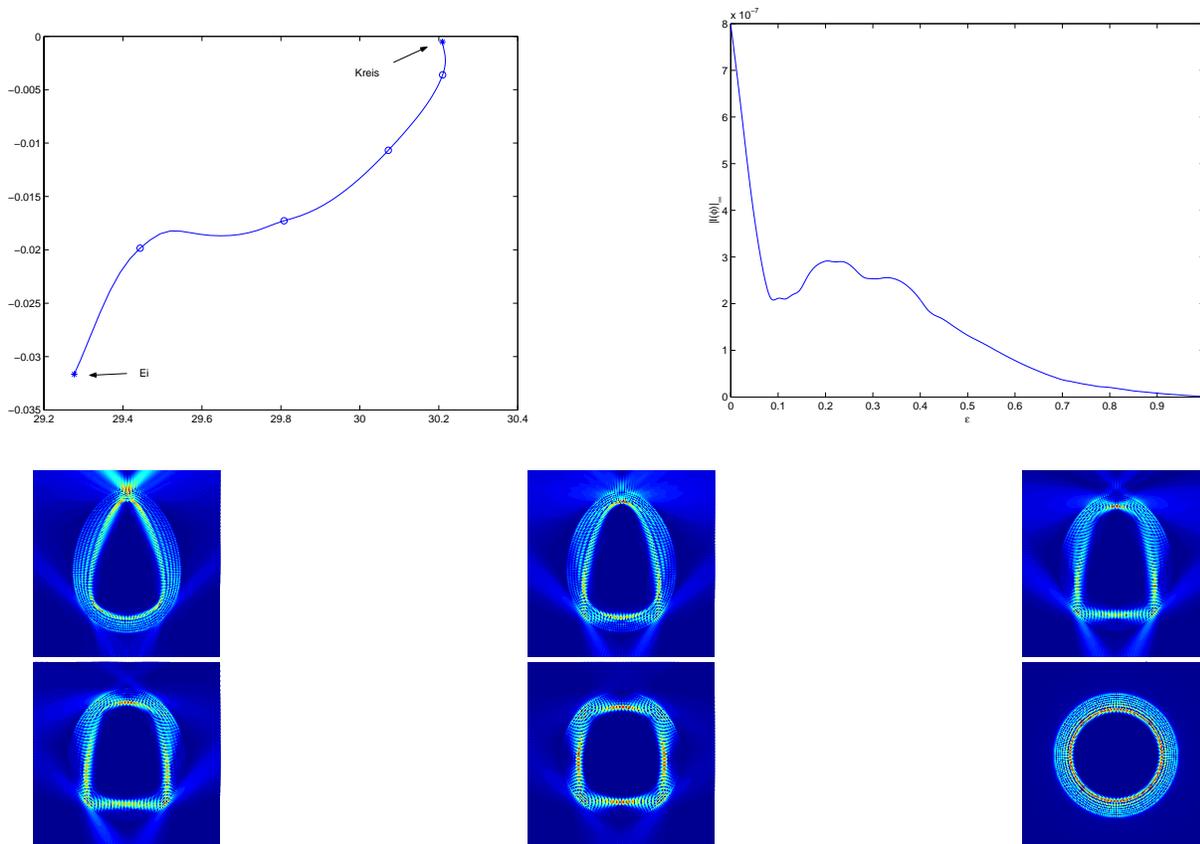


Abbildung 34: Oben: Eigenwert-Pfad unter Störung des Randes. Der Rand wurde stetig vom Ei  $\varepsilon = 0$  zum Kreis  $\varepsilon = 1$  deformiert. Die Markierungen entsprechen  $\varepsilon = 0.1, 0.325, 0.55, 0.805$ . Die jeweiligen Eigenmoden sind von links nach rechts abgebildet. Auffällig ist, daß der starke gerichtete Output bei  $\varepsilon = 0$  sehr schnell durch die Deformation zerstört wird. Das KAM-Gebiet übersteht wie im Text beschrieben nur kleine Störungen. Oben rechts: Simulierter gemessener maximale Output des Lasers in einem festen Abstand. Für das Ei-Billard ist der Output am stärksten und fällt dann schnell ab.

Der Einfluß einer Störung des Randes auf die Eigenwerte ist viel stärker als der Einfluß einer kleinen Veränderung im Brechungsindex. Deshalb ist die oben beschriebene Kontinuation nicht unbedingt für diesen Fall notwendig. Numerische Simulationen am Ei- bzw. Kartoffel-Billard haben gezeigt, daß ein einfaches Sekanten-Prädiktor-Verfahren ausreichend ist. Diese Verfahren sind sehr einfach und schnell zu implementieren.

Die numerischen Simulationen (siehe Abbildung 35) bestätigen die physikalischen Erwartungen. Für wachsenden Brechungsindex  $n_0$  verschiebt sich die Brechungslinie und das invariante Gebiet liegt vollständig unterhalb dieser Linie, so daß die Lebenszeit einer Mode größer wird, also  $|\text{Im } k|$  kleiner. Außerdem fällt die Output-Leistung des Lasers ab, da bei größerem  $n_0$  auch weniger Energie den Resonator verlassen kann. Die Moden bleiben im Resonator gefangen.

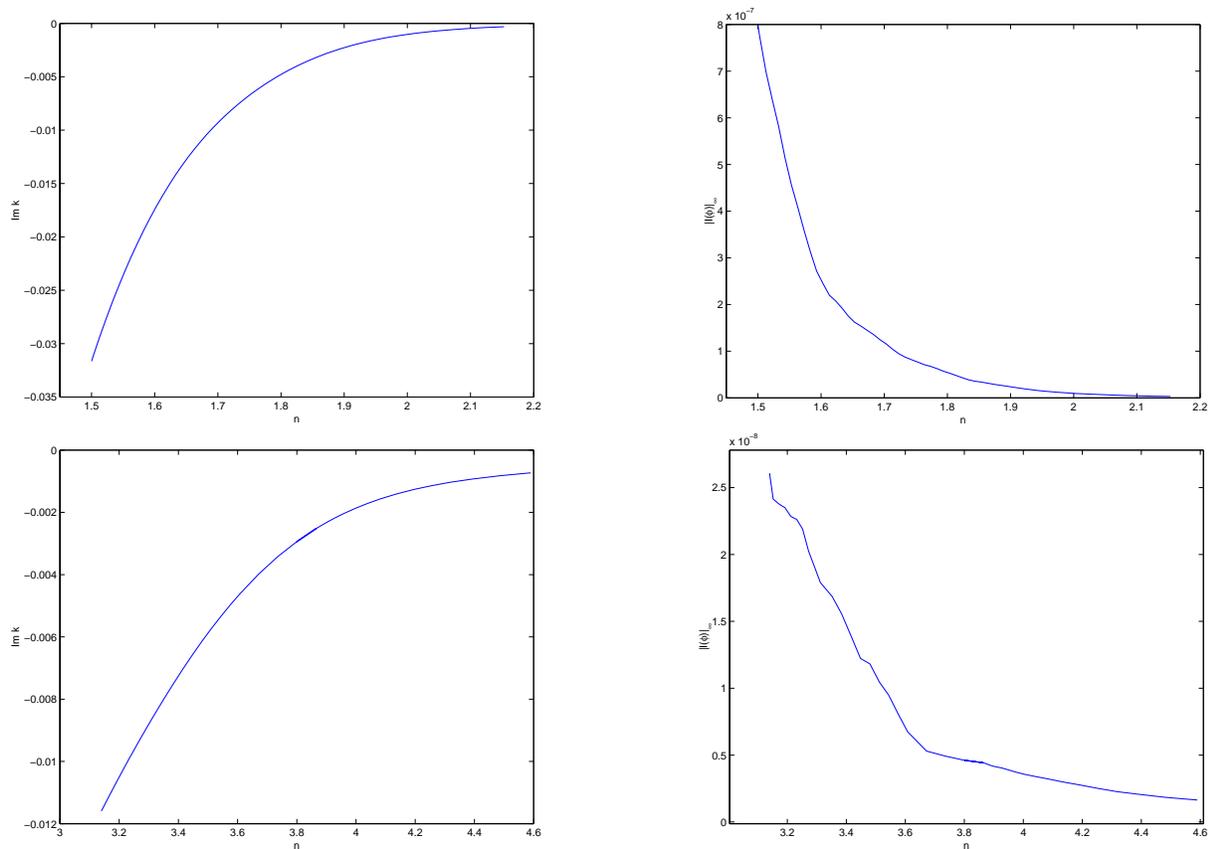


Abbildung 35: Oben: Änderung des Brechungsindex für das Ei-Billard. Unten: Änderung des Index für das Kartoffel-Billard. Jeweils rechts daneben ist die Output-Leistung des jeweiligen Lasers abgebildet.

## 5.5 Experimentelle Daten von IWG-Lasern

Für die erfolgreiche technische Anwendung von Lasern mit inneren Whispering-Gallery-Moden müssen einige Vorgaben erfüllt werden. Das in einem Quantenkaskaden-Laser verwendete Halbleitermaterial besitzt einen effektiven Brechungsindex von  $n_0 = 3.14$ , so daß ein  $p$ -Lazutkin-Billard mit  $p$ -periodischer Kaustik, die teilweise oberhalb der Brechungslinie liegt, konstruiert werden muß.

Hierzu wurde eine Familie von 3-periodischen Orbits gewählt, so daß jedes Familienmitglied ein sehr spitzwinkliges Dreieck im Billard-Phasenraum bildet. Wie in Kapitel 2 gezeigt, kann dann zur Konstruktion nur der Optimal-Control-Ansatz verwendet werden<sup>3</sup>. Als Startdreieck wurde dann ein spitzwinkliges Dreieck gewählt, so daß der spitze Winkel kleiner als der kritische Winkel zur Brechungslinie  $n_0 = 3.14$  war. Auf diese Weise kann man erzwingen, daß ein Teil der invarianten Kurve oberhalb der Brechungslinie liegen wird. Das resultierende Billard ist das schon öfters in der Arbeit aufgetretene Kartoffel-Billard.

Eine zweite Forderung ist, daß die Resonanzen im Spektrum separiert liegen. Auf diese Weise wird verhindert, daß zwei benachbarte Moden gleichzeitig *lasen*. Wie oben gesehen erfüllt das Kartoffel-Billard diese Forderung, der Abstand zwischen inneren WG-Moden und Bouncing-Ball-Moden bzw. Rand-WG-Moden ist groß genug.

Mit dem Kartoffel-Billard zur Grundlage wurde bei den Bell-Labs ein Quantenkaskaden-Laser hergestellt, um die numerischen Simulationen im tatsächlichen Experiment zu bestätigen. Der Durchmesser dieses Laser beträgt  $\approx 135\mu m$  mit einer Oberflächenrauigkeit von nur  $20nm$  am Rand. Die Betriebstemperatur des Lasers beträgt  $10K$ . Zwei aufeinanderfolgende Eigenmoden wurden mit einer Wellenlänge von  $\lambda = 7.5496\mu m$  bzw.  $\lambda = 7.5900\mu m$  gemessen. In Abbildung 36 ist das gemessen bzw. simulierte Farfield-Emissions-Muster zusehen.

Zur Bestimmung des Farfield-Emissions-Musters wird das elektrische Feld in einem Abstand von ca.  $1.5mm$  um den Mittelpunkt des Lasers gemessen. Man kann 4 Maxima im Farfield erkennen. Diese Maxima resultieren aus der gerichteten Abstrahlung des Kartoffel-Lasers an den Spitzen der Längs-Achse, wie in Abbildung 32 deutlich wurde. Die Winkelangaben in Abbildung 36 geben den Polarwinkel des Meßgerätes an, man kann aus diesen Angaben nicht bestimmen, an welchen Stellen das Licht aus dem Resonator tritt. Beim Vergleich des simulierten Feldes mit dem gemessenen Feld muß man zusätzlich beachten, daß bei dem einen die Winkel entgegen des Uhrzeigersinnes gemessen werden, bei dem anderen mit dem Uhrzeigersinn.

---

<sup>3</sup>der aus diesem Grund entwickelt worden ist

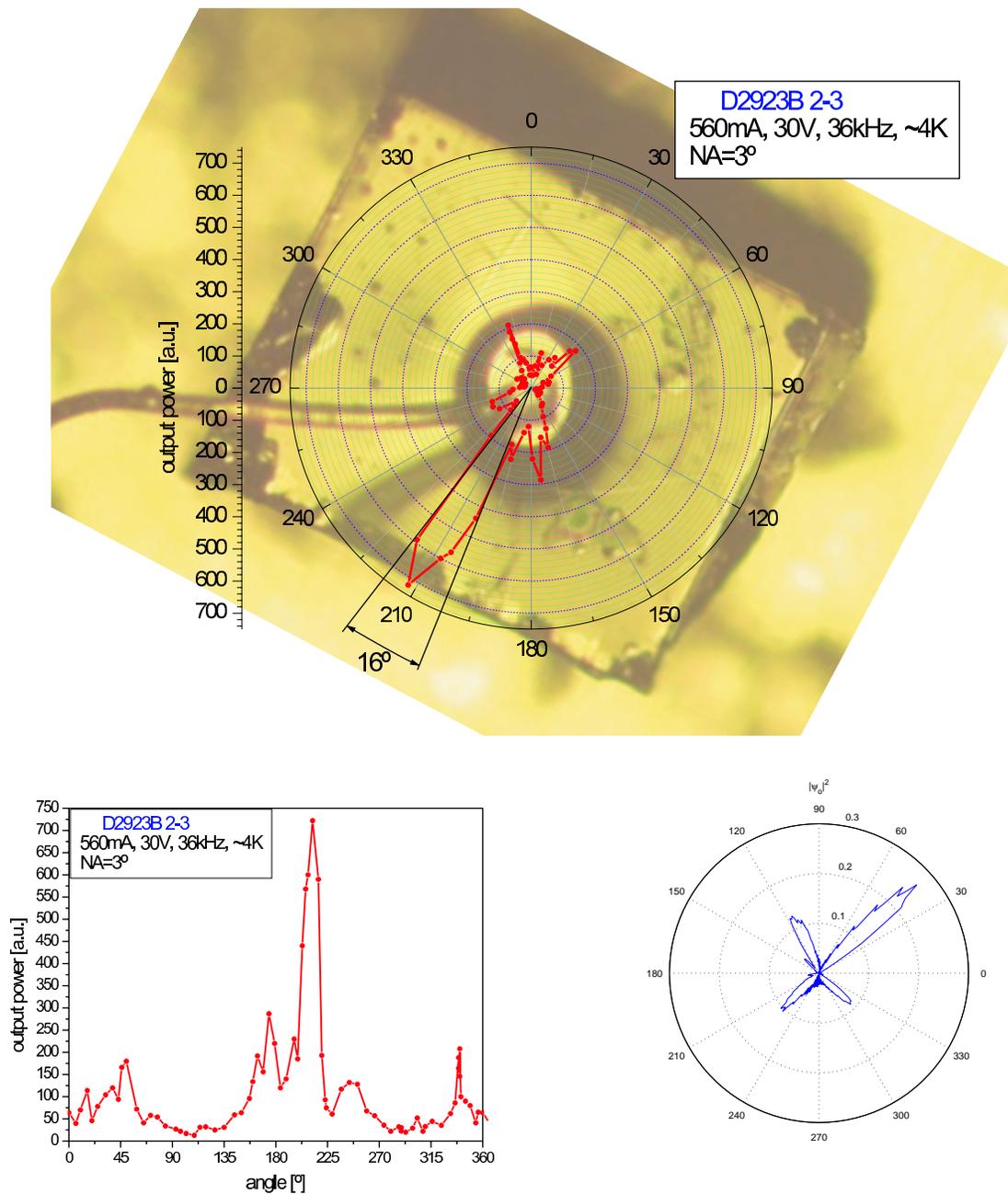


Abbildung 36: Oben: Gemessenes Farfield-Muster des Kartoffel-Lasers zusammen mit Resonatoraufsicht. Unten links: Gemessenes Farfield-Muster ohne Überblendung. Man erkennt 4 Maxima. Unten rechts: Simuliertes Farfield-Muster. Man kann auch hier 4 Maxima erkennen.

Weitere Messungen wurden an einem leicht deformierten Kreis durchgeführt. Es wurden experimentell Eigenmoden gefunden, die eine hohe Direktionalität im Farfield-Emissions-Muster aufwiesen. Für die numerische Simulation wurde die Randlinie des Resonators durch

$$r(\phi) = 1 + 0.005 \sin 4\phi + 0.003 \cos 3\phi$$

angenommen (vgl. rechten Teil der Abbildung 37).

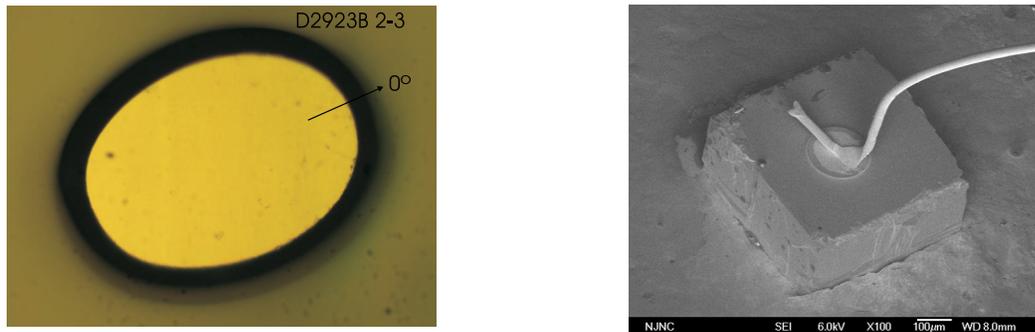


Abbildung 37: Links: Aufsicht auf den Kartoffel-Laser. Rechts: Aufsicht auf den deformierten kreisförmigen Laser.

In Abbildung 38 ist ein kleiner Ausschnitt des quasi-klassischen Spektrums abgebildet. Die Eigenmoden mit  $\text{Im } k \approx -0.02$  entsprechen inneren Whispering-Gallery-Moden, sie liegen gut getrennt von den übrigen Eigenmoden.

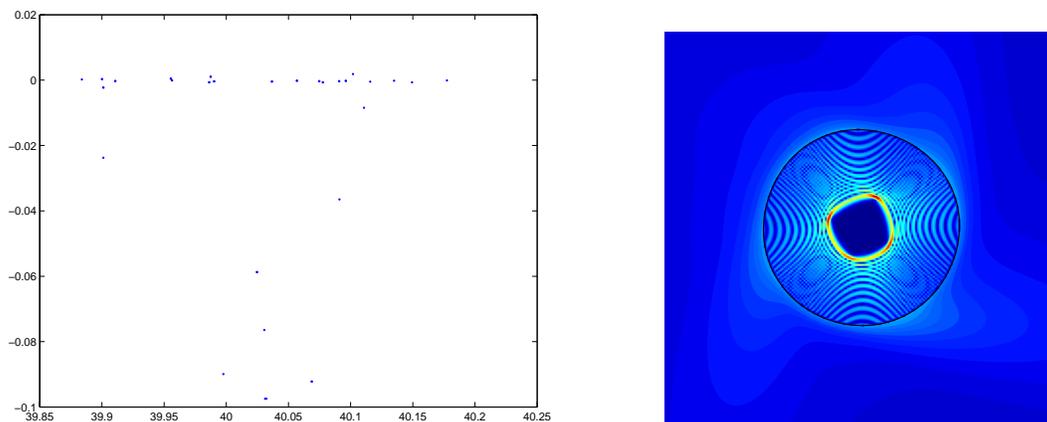


Abbildung 38: Links: Ausschnitt des quasi-klassisches Spektrums eines leicht deformierten kreisförmigen Billard. Die Moden mit  $\text{Im } k \approx -0.02$  entsprechen IWG-Moden. Rechts: Eine IWG-Mode.

Das Farfield-Emissions-Muster der IWG-Moden ist in Abbildung 39 zusehen. Verglichen mit den gemessenen Daten (rechte Bildhälfte) ist dies eine sehr gute Übereinstimmung zwischen numerischer Simulation und Experiment.

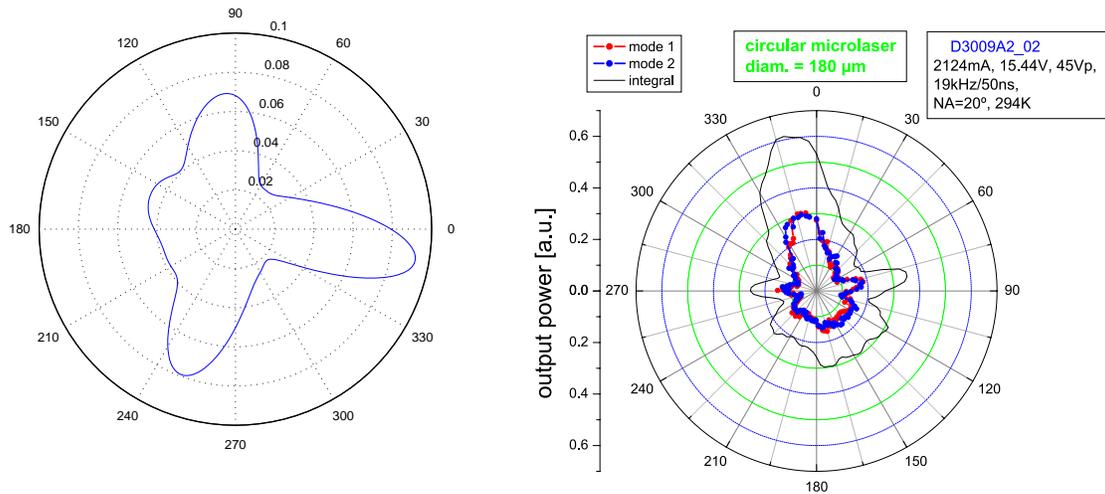


Abbildung 39: Links: Simuliertes Farfield-Muster des deformierten Kreises. Rechts: Gemessenes Farfield-Muster des deformierten Kreises.

## 6 Anhang

In diesem Abschnitt werden die in Kapitel 4 benötigten Aufspaltungen der Integralkerne hergeleitet. Es wird gezeigt, daß sich alle Kerne in der Form

$$A_1(t, \tau) \log \left( 4 \sin^2 \frac{t - \tau}{2} \right) + A_2(t, \tau)$$

mit analytischen Funktionen  $A_1$  und  $A_2$  schreiben lassen.

Es ist bekannt, daß die Hankelfunktion  $H_0^{(+)}(z)$  bei  $z = 0$  eine logarithmische Singularität besitzt. Um diese Singularität besser zu verstehen, betrachtet man die Reihenentwicklungen der Besselfunktion  $J_0(z)$  und der Neumannfunktion  $N_0(z)$ , denn es gilt  $H_0^{(+)}(z) = J_0(z) + iN_0(z)$ . Es gilt (vgl. [AS70])

$$J_0(z) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(n!)^2} \left( \frac{z}{2} \right)^{2n} \quad (6.1)$$

$$N_0(z) = \frac{2}{\pi} \left( \log \frac{z}{2} + C \right) \cdot J_0(z) - \frac{2}{\pi} \sum_{n=1}^{\infty} \left\{ \sum_{m=1}^n \frac{1}{m} \right\} \frac{(-1)^n}{(n!)^2} \left( \frac{z}{2} \right)^{2n}, \quad (6.2)$$

wobei  $C$  die Euler Konstante  $C = 0.57721 \dots$  ist. Aus diesen Reihendarstellungen kann man sofort ablesen, daß  $J_0(z)$  analytisch ist,  $J_0(0) = 1$  und die logarithmische Singularität der Hankelfunktion von der Neumannfunktion herkommt.

Der Integralkern  $M(t, \tau)$  war gegeben durch  $M(t, \tau) = \frac{i}{2} H_0^{(+)}(k|x(t) - x(\tau)|)$ . Zur Abkürzung sei im folgenden  $\alpha := k|x(t) - x(\tau)|$ , die Abhängigkeit von  $k$ ,  $t$  und  $\tau$  wird unterdrückt. Eine kleine Rechnung liefert nun,

$$\begin{aligned} M(t, \tau) &= -\frac{1}{2\pi} J_0(\alpha) \left( 2 \log \frac{\alpha}{2} \right) + \frac{i}{2} J_0(\alpha) - \frac{C}{\pi} + O(|x(t) - x(\tau)|^2) \\ &= -\frac{1}{2\pi} J_0(\alpha) \log |x(t) - x(\tau)|^2 + \frac{i}{2} J_0(\alpha) - \frac{C}{\pi} - \frac{1}{\pi} \log \frac{k}{2} + O(|x(t) - x(\tau)|^2) \\ &= -\frac{1}{2\pi} J_0(\alpha) \log \left( 4 \sin^2 \frac{t - \tau}{2} \right) \\ &\quad - \frac{1}{2\pi} J_0(\alpha) \log \frac{|x(t) - x(\tau)|^2}{4 \sin^2(t - \tau)/2} + \frac{i}{2} J_0(\alpha) - \frac{C}{\pi} - \frac{1}{\pi} \log \frac{k}{2} + O(|x(t) - x(\tau)|^2) \\ &=: M_1(t, \tau) \log \left( 4 \sin^2 \frac{t - \tau}{2} \right) + M_2(t, \tau), \end{aligned}$$

mit analytischen Funktionen  $M_1(t, \tau)$  und  $M_2(t, \tau)$ .

Um die Diagonalterme auszurechnen, werden noch einige geometrische Abschätzungen benötigt, die auch bei der Behandlung der anderen Kerne Anwendung finden. Für  $t \approx \tau$  gilt

$$x(t) = x(\tau) + x'(\tau)(t - \tau) + \frac{x''(\tau)}{2}(t - \tau)^2 + O((t - \tau)^3).$$

Hieraus folgert man

$$|x(t) - x(\tau)| = |t - \tau| \cdot |x'(\tau)| + O(|t - \tau|^2), \quad (6.3)$$

mit anderen Worten, für  $t \rightarrow \tau$  verhält sich  $|x(t) - x(\tau)|$  wie  $O(t - \tau)$ .

Wegen  $n(t) = (x_2'(t), -x_1'(t))$  gilt außerdem

$$\begin{aligned} n(t) \cdot (x(t) - x(\tau)) &= x_2'(t)(x_1(t) - x_1(\tau)) - x_1'(t)(x_2(t) - x_2(\tau)) \\ &= x_2'(t) \left( x_1'(\tau)(t - \tau) + \frac{x_1''(\tau)}{2}(t - \tau)^2 \right) \\ &\quad - x_1'(t) \left( x_2'(\tau)(t - \tau) + \frac{x_2''(\tau)}{2}(t - \tau)^2 \right) + O((t - \tau)^3), \end{aligned}$$

also

$$n(t) \cdot (x(t) - x(\tau)) = \frac{n(t)x''(\tau)}{2}(t - \tau)^2 + O((t - \tau)^3). \quad (6.4)$$

Mit anderen Worten, für  $t \rightarrow \tau$  verhält sich  $n(t) \cdot (x(t) - x(\tau))$  wie  $O((t - \tau)^2)$ .

Wegen  $J_0(0) = 1$  folgt sofort für den ersten Diagonalterm

$$M_1(t, t) = -\frac{1}{2\pi}.$$

Mit (6.3) findet man zunächst

$$\lim_{t \rightarrow \tau} \frac{|x(t) - x(\tau)|^2}{4 \sin^2(t - \tau)/2} = |x'(\tau)|^2,$$

und somit also

$$M_2(t, t) = \frac{i}{2} - \frac{C}{\pi} - \frac{1}{\pi} \log \frac{k|x'(t)|}{2}.$$

Durch termweise Differentiation von (6.2) erhält man die Reihendarstellung von  $N_1(z)$ ,

$$N_1(z) = -N_0'(z) = -\frac{2}{\pi z} J_0(z) + \frac{2}{\pi} \left( \log \frac{z}{2} + C \right) J_1(z) + O\left(\frac{z}{2}\right).$$

Wegen  $H_1(z) = J_1(z) + iN_1(z)$  gewinnt man leicht daraus eine Darstellung für die Hankelfunktion  $H_1(z)$ ,

$$H_1(z) = \frac{2i}{\pi} \left( -\frac{J_0(z)}{z} + J_1(z) \log \frac{z}{2} \right) + J_1(z) + \frac{2iC}{\pi} J_1(z) + O\left(\frac{z}{2}\right),$$

wobei die ersten beiden Terme eine lineare bzw. eine logarithmische Singularität besitzen, die hinteren Terme jedoch analytisch sind.

Der Kern  $H(t, \tau)$  war gegeben durch

$$H(t, \tau) = \frac{ik}{4} \frac{n(\tau) \cdot (x(t) - x(\tau))}{|x(t) - x(\tau)|} H_1(k|x(t) - x(\tau)|).$$

Eine kleine Rechnung liefert nun

$$\begin{aligned}
H(t, \tau) &= \frac{ik n(\tau) \cdot (x(t) - x(\tau))}{2 |x(t) - x(\tau)|} \left\{ \frac{2i}{\pi} \left( -\frac{J_0(\alpha)}{\alpha} + J_1(\alpha) \log \frac{\alpha}{2} \right) \right. \\
&\quad \left. + J_1(\alpha) + \frac{2iC}{\pi} J_1(\alpha) \right\} + O(|x(t) - x(\tau)|) \\
&= \frac{ik n(\tau) \cdot (x(t) - x(\tau))}{2 |x(t) - x(\tau)|} \left\{ \frac{2i}{\pi} \left( -\frac{J_0(\alpha)}{\alpha} + J_1(\alpha) \log \frac{\alpha}{2} \right) \right\} + \underbrace{O(|x(t) - x(\tau)|)}_{\text{wegen (6.3) und (6.4)}} \\
&= \frac{n(\tau) \cdot (x(t) - x(\tau))}{|x(t) - x(\tau)|} \left\{ -\frac{k}{2\pi} J_1(\alpha) \log |x(t) - x(\tau)|^2 + \frac{k}{\pi} \frac{J_0(\alpha)}{\alpha} \right\} + O(|x(t) - x(\tau)|) \\
&= -\frac{k}{2\pi} \frac{n(\tau) \cdot (x(t) - x(\tau))}{|x(t) - x(\tau)|} J_1(\alpha) \log \left( 4 \sin^2 \frac{t - \tau}{2} \right) \\
&\quad + \frac{n(\tau) \cdot (x(t) - x(\tau))}{|x(t) - x(\tau)|} \left\{ -\frac{k}{2\pi} J_1(\alpha) \log \frac{|x(t) - x(\tau)|^2}{4 \sin^2(t - \tau)/2} + \frac{k}{\pi} \frac{J_0(\alpha)}{\alpha} \right\} + O(|x(t) - x(\tau)|) \\
&=: H_1(t, \tau) \log \left( 4 \sin^2 \frac{t - \tau}{2} \right) + H_2(t, \tau),
\end{aligned}$$

wobei  $H_1(t, \tau)$  und  $H_2(t, \tau)$  wegen (6.3) und (6.4) analytisch sind. Ebenfalls wegen (6.3), (6.4) und  $J_1(0) = 0$  folgt sofort, daß  $H_1(t, t) = 0$  gelten muß. Aus dem gleichen Grund verschwindet auch der erste Term von  $H_2(t, \tau)$  für  $t \rightarrow \tau$ , es bleibt

$$\begin{aligned}
H_2(t, t) &= \lim_{t \rightarrow \tau} \frac{1}{\pi} \frac{n(\tau) \cdot (x(t) - x(\tau))}{|x(t) - x(\tau)|^2} \\
&= \frac{1}{2\pi} \frac{n(t)x''(t)}{|x'(t)|^2}
\end{aligned}$$

übrig. Die Aufspaltung des Kerns  $H^*(t, \tau)$  verläuft analog, es muß zusätzlich der Term  $-|x'(\tau)|$  beachtet werden.

## Literatur

- [Ans71] Philip M. Anselone. *Collectively Compact Operator Approximation Theory And Applications to Integral Equations*. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1971.
- [AS70] M. Abramowitz and I. Stegun. *Handbook of Mathematical Functions*. Dover Publications, Inc., New York, 1970.
- [Bar00] A. H. Barnett. *Dissipation in Deforming Chaotic Billiards*. PhD thesis, Havard University, 2000.
- [Bäc02] A. Bäcker. Numerical aspects of eigenvalue and eigenfunction computations for chaotic quantum systems. *preprint, arXiv:nlin.CD/0204061*, 2002.
- [BCO81] E. B. Becker, G. F. Carey, and J. T. Oden. *Finite Elements*. Prentice-Hall, Englewood Cliffs, New Jersey, 1981.
- [BFS04] A. Bäcker, S. Fürstenberg, and R. Schubert. Poincaré Husimi representation of eigenstates in quantum billiards. *Physical Review E*, 70(036204):10, 2004.
- [BHZP04] Y. Baryshnikov, P. Heider, V. Zharnitsky, and W. Parz. Whispering gallery modes inside the asymmetric resonant cavities. *Phys. Rev. Lett.*, 93:4839–4843, 2004.
- [Boa94] P. A. Boasman. Semiclassical accuracy for billiards. *Nonlinearity*, 7:485–537, 1994.
- [Bog92] E. B. Bogomolny. Semiclassical quantization of multidimensional systems. *Nonlinearity*, 5:805–866, 1992.
- [Bre81] C. A. Brebbia, editor. *Progress in Boundary Element Methods*, volume 1. Halstead Press, Broadway, Australien, 1981.
- [BRH98] L. Baowen, M. Robnik, and B. Hu. Relevance of chaos in numerical solutions of quantum billiards. *Physical Review E*, 57(4):4095–4105, 1998.
- [BW65] H. Brakhage and P. Werner. Über das Dirichletsche Außenraumproblem für die Helmholtzsche Schwingungsgleichung. *Arch. Math.*, 16:325–329, 1965.
- [BZ03] Y. Baryshnikov and V. Zharnitsky. *Billiards and Non-holonomic systems*. Zapiski Nauchnyh Seminarov POMI, 2003.
- [Dav93] J. B. Davies. Finite element analysis of waveguides and cavities - a review. *IEEE Transactions on Magnetism*, 29(2):1578, 1993.
- [Eva98] L. C. Evans. *Partial Differential Equations*. American Mathematical Society, Providence, Rhode Island, 1998.

- [GCN<sup>+</sup>98] C. Gmachl, F. Capasso, E. E. Narimanov, J. U. Nöckel, A. Douglas Stone, J. Faist, D. L. Sivco, and A. Y. Cho. High-power directional emission from microlasers with chaotic resonators. *Science*, 280:1556–1564, 1998.
- [GH96] J. Guckenheimer and P. Holmes. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. Springer-Verlag, New York, 1996.
- [Gru87] H. Gruber, editor. *Finite Elements in Physics*, 1987.
- [Gut90] M. C. Gutzwiller. *Chaos in Classical and Quantum Mechanics*. Springer-Verlag, New York, 1990.
- [Gv96] G. H. Golub and C. F. van Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore, Maryland, 1996.
- [Hej76] D. A. Hejhal. The Selberg trace formula and the Riemann zeta function. *Duke Math. Journal*, 43(3):441–482, 1976.
- [Hel86] E. J. Heller. Qualitative properties of eigenfunctions of classically chaotic hamiltonian systems. In T.H. Seligmann and H. Nishioka, editors, *Quantum Chaos and Statistical Nuclear Physics*, Lecture Notes in Physics, pages 162–181. Springer-Verlag, Berlin, 1986.
- [Hel87] E. J. Heller. Quantum localization and the rate of exploration of phase space. *Physical Review A*, 35(3):1360–1370, 1987.
- [Hel91] E. J. Heller. Semiclassical wave packet dynamics and chaos in quantum mechanics. In M. J. Giannoni, A. Voros, and J. Zinn-Justin, editors, *Proceedings of the 1989 Les Houches Summer School on „Chaos and Quantum Physics“*, pages 547–663. Elsevier Science Publishers B.V., 1991.
- [Hus40] K. Husimi. Some formal properties of the density matrix. *Proc. Phys. Math. Soc. Jpn.*, 22:264–314, 1940.
- [Inn88] N. Innami. Convex curves whose points are vertices of billiard triangles. *Kodai Math. J.*, 11:17–24, 1988.
- [Kac66] M. Kac. Can one hear the shape of a drum? *Amer. Math. Monthly*, 73(4):1–23, 1966.
- [KG95] A. Katok and E. Gutkin. Caustics for inner and outer billiards. *Comm. in Math. Phys.*, 173:101–133, 1995.
- [KH95] A. Katok and B. Hasselblatt. *Introduction to the Modern Theory of Dynamical Systems*. Cambridge University Press, Cambridge, 1995.

- [KR60] J. B. Keller and S. I. Rubinow. Asymptotic solution of eigenvalue problems. *Annals of Physics*, 9:24–75, 1960.
- [Kre95] R. Kress. On the numerical solution of a hypersingular integral equation in scattering theory. *Journal of Computational and Applied Mathematics*, 61:345–360, 1995.
- [Kre99] R. Kress. *Linear Integral Equations*. Springer-Verlag, New York, 1999.
- [KS97] I. Kosztin and K. Schulten. Boundary integral method for stationary states of two-dimensional quantum systems. *Int. J. Mod. Phys. C*, 8:293–325, 1997.
- [Kus68] R. Kussmaul. Ein numerisches Verfahren zur Lösung des Neumannschen Außenraumproblems für die Helmholtzsche Schwingungsgleichung. *Computing*, 4:246–273, 1968.
- [Lan66] S. Lang. *Introduction to Differentiable Manifolds*. Intersciences Publishers, New York, 1966.
- [Laz73a] V. F. Lazutkin. Asymptotics of the eigenvalues of the Laplacian and quasimodes. A series of quasimodes corresponding to a system of caustics close to the boundary of the domain. *Math. USSR Izvestija*, 7(2):439–466, 1973.
- [Laz73b] V. F. Lazutkin. The existence of caustics for billiards in convex domains. *Math. USSR Izvestija*, 7(2):185–214, 1973.
- [LL71] L. D. Landau and E. M. Lifschitz. *Lehrbuch der theoretischen Physik*, volume II. Akademie-Verlag, Berlin, 1971.
- [MLS<sup>+</sup>92] S. L. McCall, A. F. J. Levi, R. E. Slusher, S. J. Pearton, and R. A. Logan. Whispering-gallery mode microdisk lasers. *Appl. Phys. Lett.*, 60(3):289–291, 1992.
- [Mon02] R. Montgomery. *A Tour of Subriemannian Geometries, Their Geodesics and Applications*. American Mathematical Society, Providence, Rhode Island, 2002.
- [Moo68] R. H. Moore. Approximations to nonlinear operator equations and newton’s method. *Numerische Mathematik*, 12:23–34, 1968.
- [Mos73] J. Moser. *Stable and Random Motions in Dynamical Systems*. Princeton University Press, Princeton, New Jersey, 1973.
- [Neu85] A. Neumaier. Residual inverse iteration for the nonlinear eigenvalue problem. *SIAM J. Numer. Anal.*, 22(5):914–923, 1985.
- [Per95] A. Peres. *Quantum Theory: Concepts and Methods*. Kluwer Academic Publishers, 1995.

- [PT01] P. Piccione and D. V. Tausk. Variational aspects of the sub-Riemannian geodesic problem. *J. Diff. Geom.*, (30):183–206, 2001.
- [Ray64] J. W. Rayleigh. The problem of Whispering Gallery. In *Scientific Papers*, volume 5. Dover Publications, Inc., New York, 1964.
- [Rob84] M. Robnik. Quantising a generic family of billiards with analytic boundaries. *J. Phys. A*, 17(5):1049–1074, 19984.
- [Sch68] L. I. Schiff. *Quantum Mechanics*. McGraw-Hill, Inc., New York, 1968.
- [Sey94] R. Seydel. *Practical Bifurcation and Stability Analysis: From Equilibrium to Chaos*. Springer-Verlag, New York, 1994.
- [SF73] G. Strang and G. J. Fix. *An Analysis of the Finite Element Method*. Prentice-Hall, Englewood Cliffs, New Jersey, 1973.
- [SF91] P. P. Silvester and R. L. Ferrari. *Finite Elements for Electrical Engineers*. Cambridge University Press, Cambridge, 1991.
- [Sin76] Ya. G. Sinai. *Topics in Ergodic Theory*. Princeton University Press, Princeton, New Jersey, 1976.
- [SS04] S. Sauter and C. Schwab. *Randelementmethoden*. Teubner Verlag, Stuttgart, 2004.
- [Str86] G. Strang. *Introduction to Applied Mathematics*. Wellesley-Cambridge Press, Wellesley, M.A., 1986.
- [TOS01] U. Trottenberg, C. Oosterlee, and A. Schüller. *Multigrid*. Academic Press, London, 2001.
- [TSS02] H.E. Tureci, H. G. Schwefel, and A. D. Stone. A Gaussian-optical approach to stable periodic orbits resonances of partially chaotic dielectric micro-cavities. *The Intern. Electr. Journ. of Optics*, 10(6):752–776, 2002.
- [Vos04] H. Voss. Numerical methods for sparse nonlinear eigenvalue problems. In *Proc. XVth Summer School on Software and Algorithms of Numerical Mathematics*, 2004.
- [VS95] E. Vergini and M. Saraceno. Calculation by scaling of highly excited states of billiards. *Phys. Rev. E*, 52:2204–2207, 1995.
- [Wie03] J. Wiersig. Boundary element method for resonances in dielectric microcavities. *J. Opt. A: Pure Appl. Opt.*, 5:53–60, 2003.
- [Wig32] E. P. Wigner. On the quantum correction for thermodynamic equilibrium. *Phys. Rev.*, 40:749–750, 1932.

- [YS93] Y. Yamamoto and R. E. Slusher. Optical processes in microcavities. *Phys. Today*, 46:66–73, 1993.



# Erklärung

Ich versichere, daß ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; daß diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; daß sie – abgesehen von unten angegebenen Teilpublikationen – noch nicht veröffentlicht worden ist sowie, daß ich eine solche Veröffentlichung vor Abschluß des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen dieser Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. Seydel betreut worden.

Köln, im Juli 2005

## Teilpublikation

[BHZP04] Y. Baryshnikov, P. Heider, V. Zharnitsky and W. Parz. Whispering gallery modes inside the asymmetric resonant cavities. *Phys. Rev. Lett.*, 93:4839-4834, 2004.



# Lebenslauf

## Persönliche Angaben:

Name	Pascal Heider
Geburtsdatum	16.07.1978
Geburtsort	Köln
Staatsangehörigkeit	deutsch
Familienstand	ledig
Eltern	Dr. Franz-Peter Heider und Monika Heider, geb. Rosentreter

## Schulbildung:

1985-1989	Katholische Grundschule, Hürth-Efferen
1989-1998	Hildegard-von-Bingen Gymnasium, Köln, Abschluß Abitur (Note: 1,2)

## Stipendien und Preise

1996, 1997, 1998	Bundessieger im Bundeswettbewerb Mathematik
1996 - 2002	Stipendium der Studienstiftung des Deutschen Volkes

## Wehrdienst:

Jul. 1998 - Apr. 1999	Stabsdienst im Führungsstab der Luftwaffe, Bonn
-----------------------	-------------------------------------------------

## Studium:

ab April 1999	Studium der Mathematik mit Nebenfach Informatik an der Universität zu Köln
Oktober 2000	Vordiplom (Note: <i>sehr gut</i> )
September 2002	Diplom (Note: <i>mit Auszeichnung</i> )

## Auslandsaufenthalt:

Okt. 2002 - Jul. 2003	Forschungsaufenthalt bei den Bell-Labs, New Jersey
-----------------------	----------------------------------------------------

## Berufliche Tätigkeit:

seit Aug. 2003	Wissenschaftlicher Mitarbeiter am Mathematischen Institut der Universität zu Köln, Lehrstuhl Prof. Dr. Seydel
----------------	------------------------------------------------------------------------------------------------------------------