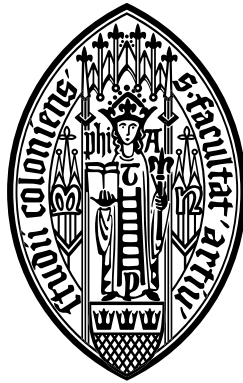


Nichtlineare Dynamik
von Oberflächeninseln
unter Elektromigrationseinfluss



Inauguraldissertation
zur Erlangung des Doktorgrades der
Mathematisch–Naturwissenschaftlichen Fakultät
der Universität zu Köln

vorgelegt von

Philipp Kuhn

aus Schweinfurt

Köln 2007

Berichterstatter: Prof. Dr. Joachim Krug
Prof. Dr. Axel Voigt

Tag der mündlichen Prüfung: 25.06.07

Kurzzusammenfassung

Die Arbeit befasst sich mit der Bewegung von Oberflächeninseln unter Elektromigrationseinfluss im Rahmen eines Kontinuumsmodells, das mittels numerischer und analytischer Methoden untersucht wird.

In dem kinetischen Regime, in dem der Massentransport ausschließlich entlang der Inselkante stattfindet, ergibt sich durch die Berücksichtigung der Anisotropie der Kristalloberfläche eine Vielfalt von Bewegungsformen, die neben schräg zum Feld laufenden Inselformen komplexes oszillatorisches Verhalten beinhaltet. Die dabei auftretenden Facetten-Orientierungen können mittels einer analytischen Theorie vorhergesagt werden. Es wird ferner ein Vergleich mit einem kinetischen Monte-Carlo-Modell durchgeführt, bei dem sich ähnliche Oszillationen zeigen, wie im Kontinuumsfall.

Findet der Massentransport über Materialaustausch mit den Atomen auf der Terrasse statt, so kann gezeigt werden, dass im Fall einer Leerstelleninsel mit unendlich hoher Stufenrandbarriere die kreisförmigen stationären Lösungen immer linear stabil sind. Das mit dem Stabilitätsproblem verbundene Phänomen nicht-orthogonaler Eigenmoden wird ausführlich diskutiert. Numerische Untersuchungen für diesen Fall zeigen eine nichtlineare Instabilität, bei der ein Teil der Insel einen Finger bildet, der sich von der Insel ablöst.

Abstract

This work addresses the motion of surface islands under electromigration within the framework of a continuum model, which is investigated by numerical and analytical methods.

In the kinetic regime, where mass transport is exclusively along the island edge, an inclusion of the crystal anisotropy of the surface results in a broad range of migration modes, including drift motion oblique to the field direction and complex oscillatory behaviour. The emerging facet orientations can be predicted from an analytical theory. Furthermore we compare the results of the continuum model with those of kinetic monte carlo simulations, where similar oscillations occur.

When mass transport is due to material exchange between terrace and island boundary, it can be shown that in the case of a vacancy island with infinitely strong step edge barrier, the stationary circular solution is always linearly stable. The phenomenon of non-orthogonal eigenmodes, which arises in this stability problem is discussed in detail. Numerical simulations for this case show a nonlinear instability, where a part of the island evolves into a finger, which detaches from the main part of the island.

Science is magic that *works*

(aus „Cat’s Cradle“ von Kurt Vonnegut, Jr.)

Inhaltsverzeichnis

1	Einführung	9
I	Das lokale Modell	17
2	Die dynamischen Gleichungen	19
3	Statische Stufenkonfigurationen	23
3.1	Stabilitätsanalyse	27
3.2	Simulationsergebnisse	29
4	Inseldynamik bei isotroper Beweglichkeit	33
4.1	Die Untersuchungen von Yang, Wang, Suo und Hao	33
4.2	Numerische Ergebnisse für das zeitabhängige Problem	41
5	Inseldynamik bei anisotroper Beweglichkeit	45
5.1	Stationäre Bewegungsformen	46
5.2	Oszillatorische Bewegung	50
5.3	Berechnung der ausgewählten Facettenorientierungen	56
6	Vergleich mit KMC-Simulationen	65
II	Das nichtlokale Modell	73
7	Die dynamischen Gleichungen	75
8	Lineare Stabilitätsanalyse	81
8.1	Die Eigenwertgleichung	81
8.2	Symmetriebetrachtungen	91
8.3	Das Spektrum	95
9	Pseudospektren und nichtnormale Matrizen	99
9.1	Störanfälligkeit des Eigenwertproblems	99
9.2	Theorie der Pseudospektren	105

9.3	Geometrische Form der Eigenmoden	113
9.4	Transientes Verhalten	120
9.4.1	Asymmetrischer Fall	127
9.4.2	Symmetrischer Fall	132
9.5	Kondition des Eigenwertproblems	133
10	Numerische Ergebnisse für die nichtlineare Zeitentwicklung	141
11	Zusammenfassung	149
A	Anhang	153
A.1	Wichtige geometrische Beziehungen	153
A.2	Einige nützliche geometrische Beziehungen in Polarkoordinaten	155
A.3	Einige Beziehungen für die modifizierten Besselfunktionen	156
A.4	Zusammenhang mit der Höendarstellung	159
A.5	Die Mathieufunktionen	161
A.6	Wechsel zur Fourierbasis	162
A.7	Elliptische Anfangsbedingung	165

Kapitel 1

Einführung

Die zentrale Größe der Thermodynamik ist die *Entropie* und die fundamentale Erkenntnis dieser Disziplin ist, dass diese Größe nur zunehmen kann und in einem abgeschlossenen System einem Maximum zustrebt. Dieser Zustand der maximalen Entropie ist der thermodynamische Gleichgewichtszustand, in dem jede Struktur aufgelöst und alle Aktivität zum Erliegen gekommen ist. Für diesen Zustand liefert die Thermodynamik im Prinzip ein abgeschlossenes Bild, das sich mit einigen wenigen universellen Gesetzen beschreiben lässt. Nun ist allerdings diese Idealisierung eines abgeschlossenen Systems in der Realität praktisch nie anzutreffen. Die meisten der in der Natur vorkommenden physikalischen Systeme sind offen, d.h. sie tauschen kontinuierlich Energie und Materie mit ihrer Umgebung aus. Und hier treten neue Phänomene auf, die unter den Bedingungen der ständig anwachsenden Entropie nicht möglich wären. Dadurch, dass das System Entropie an die Umgebung abgeben kann, wird es möglich, dass die Entropie lokal reduziert wird und sich insbesondere spontan Strukturen von teilweise erstaunlicher Komplexität bilden können. Es gab verschiedene Versuche, für diese *Musterbildung* in Nichtgleichgewichts-Systemen eine gemeinsame Beschreibung, analog zur klassischen Thermodynamik, zu finden. Es entstanden Schulen, wie z.B. die von Ilya Prigogine, der wohl einer der ersten war, der sich um eine allgemeine Theorie bemühte¹. Obwohl sich durchaus Gemeinsamkeiten und auch teilweise Prinzipien mit weitem Anwendungsbereich abzeichneten, ist der Versuch eine geschlossene, allgemeine Beschreibung zu finden, bislang gescheitert (oder zumindest noch nicht zum Ende gekommen)².

¹Prigogine prägte für diese Phänomene die Bezeichnung „dissipative Strukturen“ [41]. Weitere Benennungen und Erklärungsmodelle, die mit diesen Mustern verbunden sind, sind „Selbstorganisation“ [66], Synergetik [22] und im biologischen Kontext „Morphogenese“ [65].

²Vielleicht trifft hier das Novalis-Zitat zu:

„Mannigfache Wege gehen die Menschen. Wer sie verfolgt und vergleicht, wird wunderliche Figuren entstehen sehn; Figuren, die zu jener großen Chifferschrift zu gehören scheinen, die man überall, auf Flügeln, Eierschalen, in Wolken, im Schnee, in Kristallen und in Steinbildungen, auf gefrierenden Wassern, im Innern und Äußern der Gebirge, der Pflanzen, der Tiere, der Menschen, in den Lichtern des Himmels, auf berührten und gestrichenen Scheiben von Pech und Glas, in den Feilspänen um den Magnet her, und sonderbaren Konjunkturen des

So bilden die Untersuchungen von Musterbildung ein weites und vielfältiges Gebiet³. Es umfasst so verschiedene Bereiche wie die klassischen Beobachtungen von Strömungsmustern in Flüssigkeiten und die, mit dem Namen von Alan Turing verbundenen Reaktions-Diffusions-Systeme. Wir wollen an dieser Stelle lediglich die Theorie des Kristallwachstums aus der Gasphase erwähnen⁴, die gewissermaßen benachbart zu unserem Thema liegt. Eine für dieses Gebiet zentrale Arbeit war das 1951 von Burton, Cabrera und Frank vorgestellte Modell [11]. Die Erkenntnis, die diesem Modell zugrunde liegt, ist die, dass Atome, die aus der Gasphase auf die Oberfläche des wachsenden Kristalls abgeschieden werden⁵, vorrangig an Inhomogenitäten eingebaut werden. Selbst auf einer von Verunreinigungen freien Oberfläche mit perfekter Kristallstruktur existieren solche Inhomogenitäten in Form von Stufenkanten⁶. Da das Wachstum des Kristalls vornehmlich an diesen Kanten vor sich geht⁷, lässt sich ein Modell formulieren, das neben der Konzentration an frei diffundierenden Atomen, die geometrische Form der (als kontinuierlich angenommenen) Stufenkante als dynamische Größe enthält. Diese Art von Beschreibung wird auch unserer Arbeit zugrunde liegen. Jedoch ist hier die Ursache für das Nichtgleichgewicht nicht das Wachstum, sondern das Phänomen der sogenannten *Elektromigration*. Da dieser Effekt für unsere Arbeit zentral ist, werden wir einen kurzen Streifzug durch die historische Entwicklung dieses Begriffs machen.

Die wahrscheinlich frühesten Beobachtungen gehen auf den französischen Wissenschaftler Gerardin zurück [17]. Er bemerkte, dass nach der Elektrolyse von geschmolzenen Blei-Zinn- und Natrium-Quecksilber-Legierungen nach dem Erstarren die beiden Enden des Materials unterschiedliche physikalische, bzw. chemische Eigenschaften zeigten⁸. Offenbar bewirkte also die angelegte Spannung einen Materialtransport, der die Schmelze entmischt. Dieser Transport findet im festen Zustand natürlich mit einer wesentlich geringeren Geschwindigkeit statt und wurde erst in den dreißiger Jahren nachgewiesen. Ein besonders illustratives Beispiel sind hier die Untersuchungen von R.P. Johnson von 1938, bei denen Wolframdrähte aus durchgebrannten Glühbirnen verschiedene Ober-

Zufalls, erblickt. In ihnen ahndet man den Schlüssel dieser Wunderschrift, die Sprachlehre derselben, allein die Ahndung will sich selbst in keine feste Formen fügen, und scheint kein höherer Schlüssel werden zu wollen.“

(Aus „Die Lehrlinge zu Sais“).

³Für eine umfassende Übersicht, siehe [12].

⁴Das Nichtgleichgewicht besteht hier darin, dass aus einer übersättigten Gasphase beständig Material auf der Oberfläche deponiert wird. Für eine detaillierte Darstellung, siehe [39, 45].

⁵D.h. es handelt sich hier also um adsorbierte Atome, was häufig zu der Bezeichnung *Adatom* verkürzt wird.

⁶Das mikroskopische Bild einer solchen Oberfläche ist eine aus Lagen aufgebaute Landschaft. Die Ränder der nicht vollständig geschlossenen Lagen bilden hier Stufen, die von Terrassen voneinander getrennt sind (vgl. Abbildung 1.3).

⁷Natürlich besteht daneben noch der wichtige Prozess der Keimung auf der Terrasse.

⁸Im Falle der Blei-Zinn-Legierung war das eine Ende weich und das andere brüchig, während bei der Natrium-Quecksilber-Verbindung nur ein Ende Reaktivität bei Kontakt mit Wasser zeigte.

flächen-Beschaffenheit zeigten, je nachdem ob sie mit Gleich- oder Wechselstrom betrieben worden waren [29]. Abbildung 1.1 zeigt einen derartigen Vergleich. Während der von Gleichstrom durchflossene Draht glatt ist, weist die Oberfläche des mit Wechselstrom betriebenen Drahtes Rippen auf. Die ersten systematischen Untersuchungen wurden jedoch erst in den fünfziger Jahren von Seith und Wever durchgeführt [54]. Durch die Beobachtung der Bewegung kleiner, mechanisch erzeugter Eindrücke in einem stromdurchflossenen Draht konnten sie nachweisen, dass die Richtung des Massentransports damit zusammenhängt, welche Ladungsträger (Elektronen oder Löcher) in dem entsprechenden Material vorherrschen. Dies stellte insofern einen Wendepunkt dar, als sich damit zeigte, dass die treibende Kraft nicht alleine elektrostatischer Natur sein kann („direkte Kraft“), sondern dass eine Wechselwirkung mit den Ladungsträgern vorliegen muss. Das Bild, das wir heute von diesem Prozess haben, ist in der Tat ein Impulsübertrag der Ladungsträger auf die Gitter- oder Fremdatome („Windkraft“⁹).

Die Hauptanwendung der Elektromigration bestand bis in die sechziger Jahre hinein hauptsächlich darin, Metalle von Verunreinigungen (besonders in Gasform) zu befreien. Die Forschung nahm eine neue Richtung und gewann erheblich an Bedeutung, als Elektromigrations-Effekte als einer der Hauptschädigungsmechanismen von Aluminium-Leiterbahnen in integrierten Schaltkreisen identifiziert wurden¹⁰ [7, 6]. Durch den Stromfluss wird hier Material abgetragen, bzw. angehäuft, so dass es entweder zur Unterbrechung oder zum Kurzschluss in den leitenden Elementen kommt. In Abbildung 1.2 ist ein solcher Elektromigrationsschaden zu sehen, der unter kontrollierten experimentellen Bedingungen erzeugt wurde. Diese Leiterbahnen bestehen aus polykristallinem Material und der Massentransport verläuft vorrangig entlang der Korngrenzen, die Zonen erhöhter Beweglichkeit darstellen [5].

Mit dem einsetzenden intensiven Interesse an dem Entwicklungsmechanismus dieser Schäden begann auch die theoretische Auseinandersetzung mit der Bewegung kleiner Hohlräume (Poren) unter Elektromigrationseinfluss. Hier wurden Modelle entwickelt, die dem in dieser Arbeit behandelten Thema verwandt sind. Die technologische Bedeutung der Elektromigration ist auch heute noch gegeben¹¹, wir werden uns jedoch nun einem Thema zuwenden, das näher an dem unseren liegt. Spaltet man einen Kristall in einem kleinen Winkel gegenüber sei-

⁹Diese Vorstellung geht ursprünglich auf F. Skaupy zurück [55]. Die relative Stärke von direkter Kraft und Windkraft ist von Material zu Material verschieden, wobei bei Metallen im Allgemeinen die Windkraft dominiert.

¹⁰Hier herrschen aufgrund der geringen Abmessungen entsprechend hohe Stromdichten. In der Praxis unterliegen die Elemente über Jahre diesem Einfluss, so dass sich eine erhöhte Anfälligkeit gegenüber derartigen Schäden ergibt.

¹¹Anfang der siebziger Jahre wurde entdeckt, dass eine Beimischung von Kupfer die Beweglichkeit von Aluminium entlang der Korngrenzen herabsetzt. Zusammen mit der besseren Leitfähigkeit von Kupfer führte dies dazu, dass Kupfer-Aluminium-Legierungen (oder seit einiger Zeit auch Kupfer alleine) bei der Herstellung von integrierten Schaltkreisen zunehmend an Bedeutung gewonnen haben.

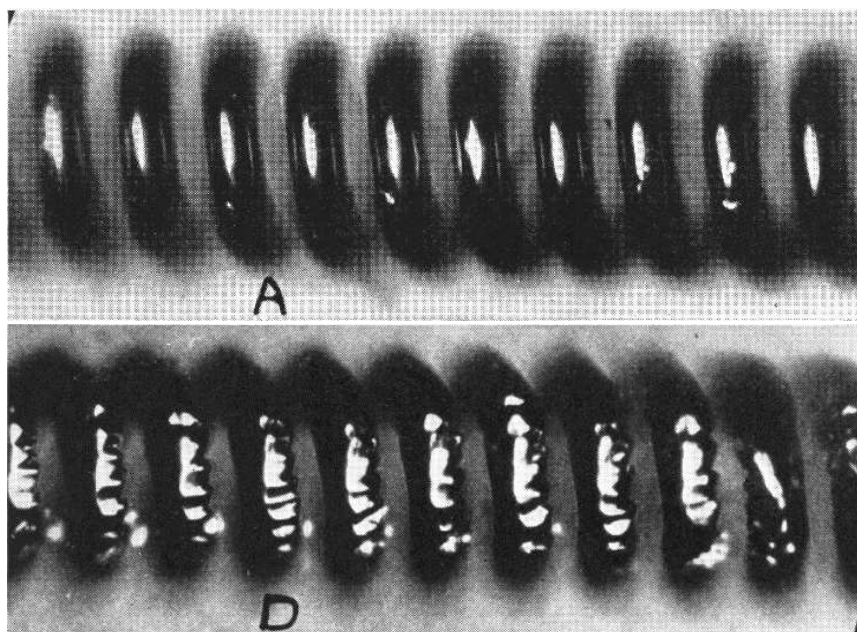


Abbildung 1.1: Wolframdrähte aus durchgebrannten Glühlampen. Für den Draht im oberen Bild wurde Wechselstrom verwendet, für den im unteren Bild Gleichstrom (aus [29]).

ner Hochsymmetrieachse, so besteht die dabei entstehende *vizinale* Oberfläche aus geraden Stufenkanten die durch (je nach Winkel) breitere oder schmalere Terrassen voneinander getrennt sind (Abbildung 1.3). Unter Elektromigrationseinfluss können nun Instabilitäten eines solchen gleichmäßigen Stufenzugs entstehen. Bei dem in Abbildung 1.3 (b) gezeigten Profil haben sich sogenannte *Stufenbündel* gebildet, bei der sich in Teilen der Oberfläche die Stufen auf einen relativ engen Raum konzentrieren und zu einem Bündel, manchmal auch Makro-Stufe genannt zusammen gefunden haben. Auf einer etwas größeren Skala zeigt sich dies als ein Streifenmuster (siehe Abbildung 1.4 links). Bei dem Prozess in Abbildung 1.3 (c) ergibt sich eine Instabilität in der Form der einzelnen Stufen. Die entstehenden Mäander organisieren sich in Phase zueinander, so dass die Oberfläche auf größerer Skala eine typische Rippenform zeigt (Abbildung 1.4 rechts).

Beide Musterbildungsprozesse waren schon vorher aus Wachstumssituationen bekannt (siehe z.B. [31, 39]). Dass sich die Stufenbewegung auch mittels Stromfluss manipulieren lässt, wurde Ende der achziger Jahre von Latyshev *et al.* entdeckt¹² [35]. Dies ist insofern bedeutsam, als sich dadurch eine neue Möglichkeit

¹²Dabei handelte es sich nicht um Experimente zur Elektromigration. Eine Oberfläche einem Gleichstromfluss auszusetzen, ist eine gängige Methode, die Oberfläche zu heizen, um den Diffusionsprozess zu beschleunigen.

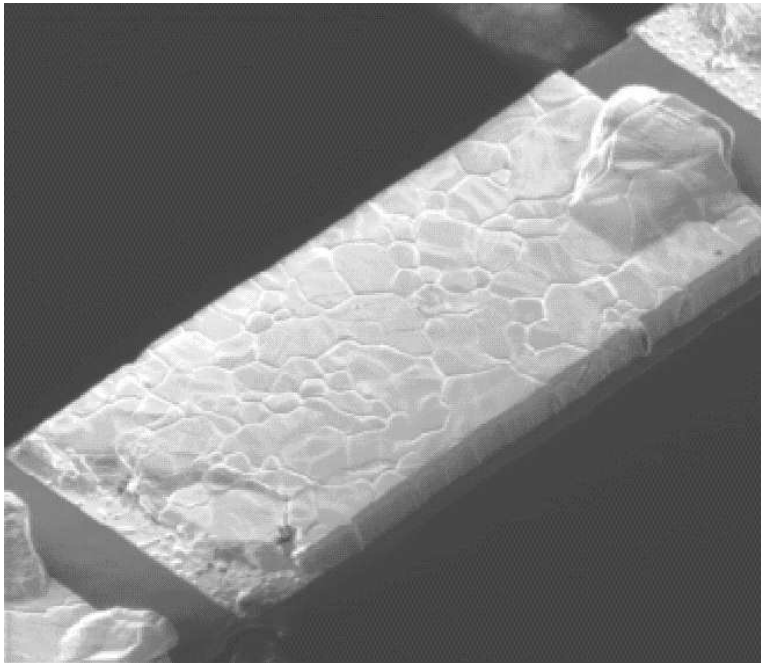


Abbildung 1.2: Aluminium-Segmente (Breite 800nm, Dicke $10\mu m$) auf TiN, belastet mit einer Stromdichte von $1\frac{MA}{cm^2}$ bei 500 K. Typischer Elektromigrati-onsschaden durch Abtragen von Material am kathodischen Ende und Anhäufung an der Anode. An der Oberfläche der Leiterbahn sind deutlich die Korngrenzen des Materials zu erkennen, die die bevorzugten Transport-Pfade für den Massenstrom bilden. (aus [58]).

ergibt, einen kontrollierten Nichtgleichgewichtszustand zu schaffen, der morphologische Instabilitäten hervorbringt.

Vor dem Hintergrund unserer Arbeit sind diese Systeme in konzeptioneller Hinsicht interessant, weil sich hier, ganz im Geiste der Burton-Cabrera-Frank-Theorie die Entwicklung der Oberfläche durch die Bewegung der einzelnen Stufenkanten verstehen lässt. Tatsächlich werden theoretische Ansätze, die denen im Kristallwachstum ähneln, hier mit Erfolg angewendet [57, 14, 46]. Bei vizinalen Oberflächen handelt es sich um Multilag-Strukturen, d.h. um Gebilde, bei denen mehrere Kristalllagen (und damit Stufenkanten) eine Rolle spielen. Sie sind als Modellsysteme geeignet, weil die Stufen sich einander in Form und Verhalten gleichen. Im Allgemeinen enthalten Oberflächen natürlich vielfältige Stufenkonfigurationen, die auf komplizierte Weise aufeinander einwirken. Bevor eine theoretische Beschreibung für das Verhalten solcher Morphologien unter Elektromigration angegangen werden kann, bilden einfachere Systeme die Möglichkeit, ein Verständnis der grundlegenden Dynamik zu entwickeln. Und hier kommen wir schließlich auf das eigentliche Thema unserer Arbeit. Statt

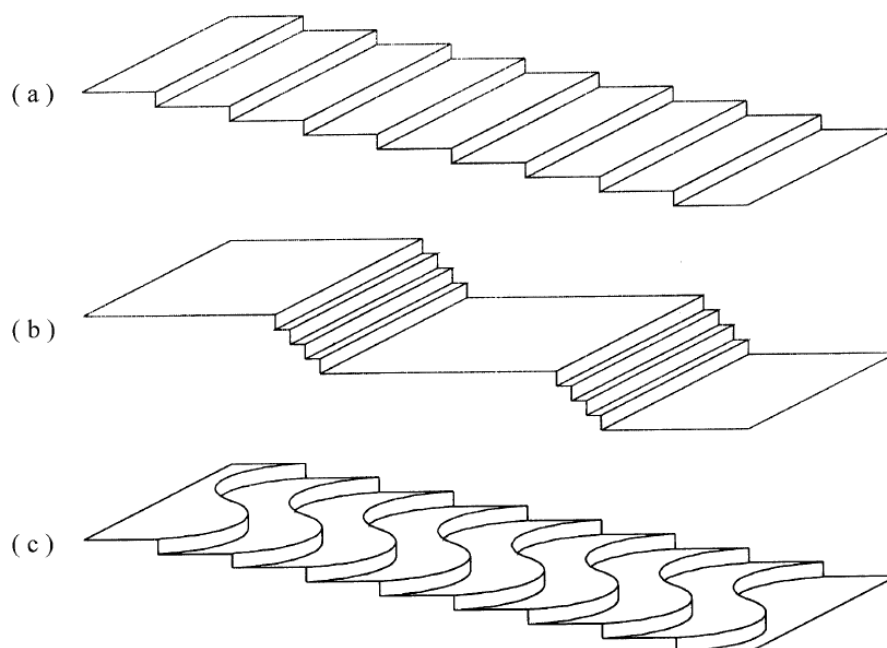


Abbildung 1.3: Konfigurationen einer vyzinalen Oberfläche: Das obere Bild (a) zeigt eine gleichmäßige Anordnung der Terrassenbreiten, die mittlere Abbildung (b) zeigt *Stufenbündel*, die durch breite Terrassen voneinander getrennt sind und die untere Skizze (c) zeigt das Phänomen des *Stufenmäanderns* (aus [69]).

unendlich ausgedehnte Stufen zu betrachten, bietet es sich an, als eine sehr einfache Konfiguration eine geschlossene Stufenkante, d.h. eine Oberflächeninsel zu untersuchen. Der durch die Elektromigration verursachte Massentransport bewirkt hier, dass die Insel als Ganzes zu driften beginnt¹³. Wir bauen für die theoretische Beschreibung auf den Arbeiten von Olivier Pierre-Louis und Theodore L. Einstein aus dem Jahr 2000 auf [44]. Die Inselbewegung wird dabei von drei Arten von Prozessen bestimmt, die in Abbildung 1.5 dargestellt sind. In der mit PD („periphery diffusion“) bezeichneten Bewegung wandern Atome an der Inselkante entlang, im Prozess TD („terrace diffusion“) diffundieren sie über das Substrat, bevor sie wieder in den Inselrand eingebaut werden und EC („evaporation-condensation“) bezeichnet den Ablöse- bzw. Anlagerungsvorgang des Atoms an die Insel. Wird die Inselbewegung von einem dieser Prozesse dominiert, so erhält man Grenzfälle, bzw. kinetische Regimes, von denen wir die beiden Fälle PD und TD untersuchen werden. Im ersten Teil der Arbeit konzentrieren wir uns auf den Grenzfall der reinen Kantendiffusion. Dies bringt

¹³Experimentell wurde eine solche Driftbewegung in [38, 50] beobachtet.

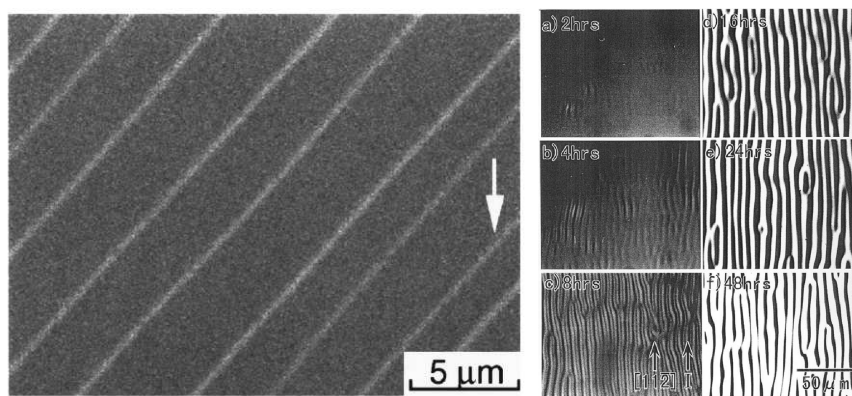


Abbildung 1.4: Elektromigrationsinduzierte Stufeninstabilitäten auf vizinalen Oberflächen. Das linke Bild zeigt die Elektronenmikroskop-Aufnahme einer Si(111)-Oberfläche, auf der sich Stufenbündel gebildet haben (im Bild als helle Streifen sichtbar). Der Pfeil zeigt die Stromrichtung an, die hier schräg zu der Richtung der Stufen verläuft (aus [15]). Das rechte Bild zeigt die Entstehung eines Rippenmusters aus Stufenmäandern auf Si(111), die zueinander in Phase liegen. Der Pfeil zeigt an, dass der Strom hier in Richtung der Stufenkante fließt (aus [40]).

als Vereinfachung mit sich, dass die Bewegung der Atome auf eine räumliche Dimension eingeschränkt wird, so dass die entsprechenden Gleichungen jeweils eine räumliche und eine zeitliche Koordinate enthalten. Die Erweiterung gegenüber den Arbeiten in [44] wird darin bestehen, dass wir den Einfluss der Kristallanisotropie auf der Oberfläche mit berücksichtigen werden. Im zweiten Teil kommen wir auf den Fall der reinen Terrassen-Diffusion zu sprechen, wo sich die Bewegung der Atome nun in zwei Dimensionen vollzieht. Die entsprechenden Gleichungen sind wesentlich komplizierter und die numerische Behandlung anspruchsvoller.

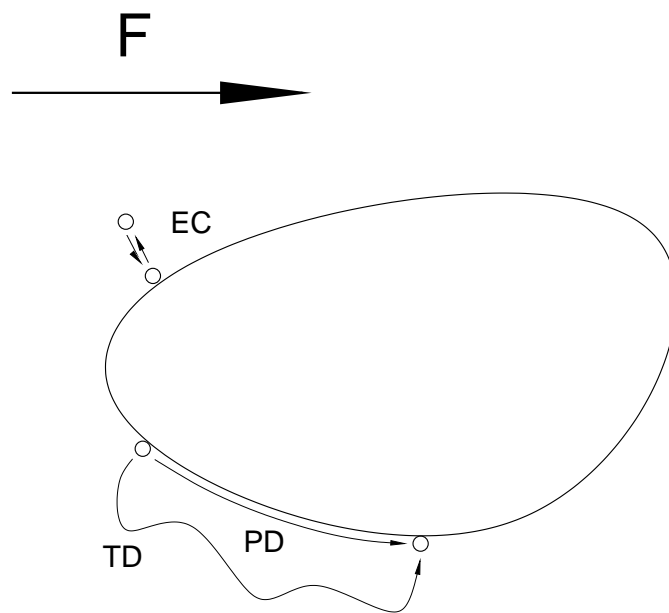


Abbildung 1.5: Die drei Hauptmechanismen des Massentransports: Diffusion entlang der Inselkante (PD), Diffusion auf der Terrasse (TD) und Anlagerung/Ablösung vom Inselrand (EC). Der Pfeil mit dem Buchstaben F kennzeichnet die Richtung des elektrischen Feldes.

Teil I

Das lokale Modell

Kapitel 2

Die dynamischen Gleichungen

Die grundlegende Annahme für das lokale Modell besteht darin, nur Massentransport entlang des Inselrandes zuzulassen. Letzteren stellen wir mathematisch durch eine geschlossene Kurve in der Ebene dar, die wir durch die Bogenlänge s parametrisieren. Die Richtung, d.h. die lokale Orientierung dieser Kurve beschreiben wir durch den Winkel θ zwischen der Normalen und der (positiven) y -Richtung. Die Richtung der Normalen soll dabei von der Insel weg in den Außenraum zeigen und wir legen den Winkel *im Uhrzeigersinn* als positiv fest¹. Die Krümmung κ ergibt sich dann in der üblichen Weise als Ableitung des Orientierungswinkels θ nach der Bogenlänge²

$$\kappa = \partial_s \theta. \quad (2.1)$$

Entlang dieser Kurve fließt nun ein Massenstrom³ $j(s)$, der die Bewegung des Inselrandes bewirkt. Er wird von zwei Einfüssen bestimmt. Zum einen besteht die Tendenz der Stufenkante, zur thermodynamischen Gleichgewichtsform zu relaxieren. Diese ist dadurch gekennzeichnet, dass das *chemische Potential* μ entlang des Randes einen konstanten Wert annimmt [39]. Das chemische Potential berechnet sich aus der lokalen Geometrie als das Produkt aus Steifigkeit⁴ $\tilde{\gamma}$ und Krümmung κ (siehe z.B. [39])

¹D.h. entgegengesetzt zum mathematisch positiven Durchlaufsin.

²Wir legen die Krümmung einer *konvexen* Insel als positiv fest. Daraus folgt mit (2.1), dass s mit dem Winkel θ anwächst, d.h. die Bogenlänge wächst beim Umlauf im Uhrzeigersinn an.

³Da die Fläche eines Inselstücks der darin enthaltenen Masse entspricht, hat die Größe $j(s)$ die Dimension Fläche pro Zeit.

⁴Die Steifigkeit hängt mit der freien Energie pro Länge (Oberflächenspannung) γ der Stufenkante über

$$\tilde{\gamma}(\theta) = \gamma(\theta) + \gamma''(\theta)$$

zusammen. Sie hängt im Allgemeinen von der Orientierung ab, wodurch sich anisotrope Gleichgewichtsformen ergeben.

$$\mu = \tilde{\gamma}\kappa.$$

Den (Nichtgleichgewichts-) Strom j_d , der die Relaxation bewirkt, werden wir als proportional zum Gradienten des chemischen Potentials annehmen [26]

$$j_d = -\sigma\partial_s\mu.$$

Die Proportionalitätskonstante σ ist die *Adatombeweglichkeit*, die angibt, wie stark die Atome bei vorgegebener Kraft $\partial_s\mu$ in Bewegung gesetzt werden. Sie hängt entlang der Stufenkante davon ab, wie schwer es für ein Teilchen ist, den Gitterplatz zu wechseln und ist damit im Allgemeinen von der Orientierung der Kante abhängig⁵.

Der zweite Einfluss, der der Relaxation entgegenwirkt und das System aus dem Gleichgewicht her austreibt, ist die Elektromigrationskraft F_{el} . Von ihr werden wir annehmen, dass sie nicht von den lokalen Gegebenheiten der Stufenkante abhängt, sondern an jedem Punkt mit konstanter Stärke und Richtung wirkt. Wir legen die x -Achse in Feldrichtung und berechnen die Kraft entlang der Stufenkante durch Projektion auf die Tangentialrichtung des Inselrandes

$$F_{el} = F \cos \theta, \quad (2.2)$$

woraus sich der Elektromigrationsstrom j_{el} als

$$j_d = \sigma F \cos \theta$$

ergibt. Die Kombination dieser beiden Ströme ergibt für den gesamten Massenstrom die für uns zentrale Gleichung

$$j = \sigma \left(-\frac{\partial}{\partial s}(\tilde{\gamma}\kappa) + F \cos \theta \right).$$

Die Bewegung des Inselrandes folgt daraus mittels der Massenerhaltung. Führt der Strom j an einer Stelle zu einer Anhäufung oder zu Abbau von Material, ist also die (eindimensionale) Divergenz $\partial_s j$ positiv oder negativ, so verschiebt sich an dieser Stelle der Inselrand in Normalenrichtung nach außen oder innen. Die Normalengeschwindigkeit des Randes berechnet sich dann nach

$$v_n + \frac{\partial}{\partial s} j = 0. \quad (2.3)$$

⁵Zum Beispiel ist in Richtung der Kristallachse die Stufenkante (abgesehen von der thermischen Rauigkeit) atomar glatt. Hier wird ein Adatom im Allgemeinen eine hohe Beweglichkeit besitzen und σ dementsprechend in dieser Richtung ein Maximum aufweisen.

Aus den beiden Gleichungen (2.4) und (2.3) kann nun die Bewegung der Insel bestimmt werden. Die Geometrie der Insel (ausgedrückt durch $\theta(s)$ und $\kappa(s)$, bzw. deren Ableitungen nach der Bogenlänge) bestimmt den Strom (2.4), der wiederum die geometrische Form der Insel nach (2.3) verändert.

Der Ansatz, diese Gleichungen als Modell für die Elektromigration von Oberflächeninseln zu verwenden, geht auf die Arbeit [44] von O. Pierre-Louis und T.L. Einstein zurück. Allerdings sind sie äquivalent zu einem Grenzfall der Gleichungen für die Elektromigration von (zweidimensionalen) Poren im Innern eines leitenden Materials, der schon früher, unter anderem in den Arbeiten [27],[60], [59], [68] und [70] behandelt wurde.

Die wesentliche Neuerung, die wir hier einführen wollen, besteht darin, die Kristallanisotropie der Oberfläche zu berücksichtigen. Wir lassen dazu für die Beweglichkeit σ (und teilweise auch für die Steifigkeit $\tilde{\gamma}$) eine Orientierungsabhängigkeit zu. Der Strom schreibt sich dann als

$$j = \sigma(\theta) \left(-\frac{\partial}{\partial s} (\tilde{\gamma}(\theta)\kappa) + F \cos \theta \right). \quad (2.4)$$

Als Form für die Beweglichkeit $\sigma(\theta)$ nehmen wir einen einfachen Ausdruck an, der eine Kristallanisotropie mit n -zähliger Symmetrie nachahmt

$$\sigma(\theta) = \sigma_0 \frac{1 + S \cos^2 \left(\frac{n}{2} (\theta + \alpha) \right)}{1 + S}, \quad (2.5)$$

dabei ist S der Parameter, der die Stärke der Anisotropie bestimmt und σ_0 die maximale Beweglichkeit. Der Winkel α gibt wieder, inwieweit die Symmetrieachse gegenüber dem Feld gedreht ist. Abbildung 2.1 zeigt $\sigma(\theta)$ für den drei- und sechszähligen Fall. Die meisten unserer Untersuchungen wurden für $n = 6$ und $\alpha = 0$ durchgeführt.

Wir führen an dieser Stelle noch die charakteristische Längen- und Zeitskala ein, die sich aus den Größen $\tilde{\gamma}$, σ_0 und F bilden lässt

$$\begin{aligned} l_E &= \sqrt{\frac{\tilde{\gamma}}{F}} \\ t_E &= \frac{l_E^4}{\sigma_0 \tilde{\gamma}}. \end{aligned} \quad (2.6)$$

Die Länge l_E beschreibt das Verhältnis von Steifigkeit zu Feldstärke und ist damit ein Maß für die relative Stärke der Elektromigration. Auf Skalen, die klein gegen l_E sind, wird also die Relaxation zur Gleichgewichtskonfiguration dominieren, während auf großen Skalen die Elektromigrationseffekte wichtig werden.

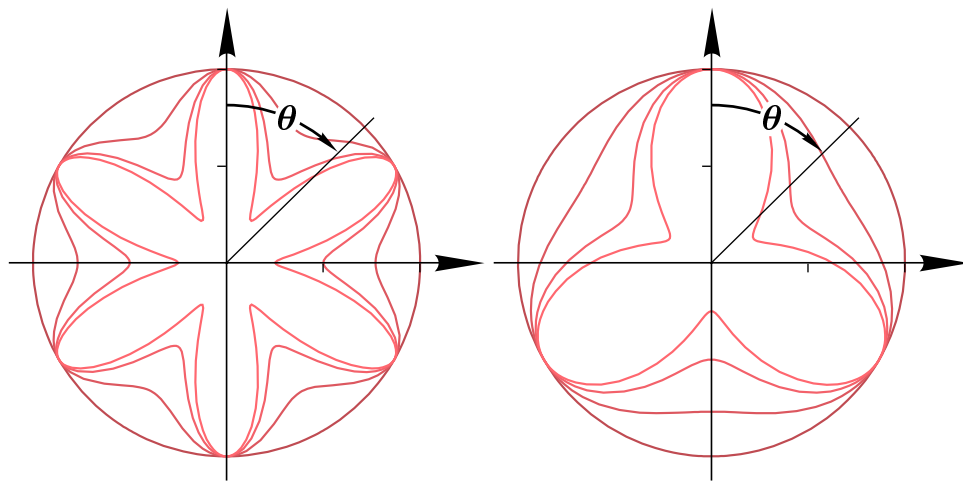


Abbildung 2.1: Orientierungsabhängigkeit der Beweglichkeit $\sigma(\theta)$ mit $S = 0, 0.3, 1, 3$ für $n = 6$ (links) und $n = 3$ (rechts). Der Winkel α ist Null. Die Orientierung $\theta = 0$ mit maximaler Beweglichkeit entspricht einer Kante, die in Feldrichtung ausgerichtet ist. Das Maximum liegt jeweils bei σ_0 .

Kapitel 3

Statische Stufenkonfigurationen

Bevor wir uns unserem eigentlichen Thema, nämlich der Inselbewegung zuwenden, untersuchen wir den etwas einfacheren Fall einer unendlich ausgedehnten Stufenkante. Im Folgenden entwickeln wir dazu die Untersuchung aus [32] für ein eindimensionales Oberflächenprofil unter Elektromigrationseinfluss und wenden sie auf das äquivalente Problem der Stufenkante an.

Betrachten wir also eine unendlich lange Stufenkante der mittleren Orientierung $\langle \theta \rangle$ und Feld in horizontaler Richtung. Statische Konfigurationen sind Lösungen der Bewegungsgleichung (2.3) mit $v_n \equiv 0$. Nach (2.3) impliziert das $j = \text{const.}$, d.h. unter Verwendung von (2.4).

$$\sigma(\theta) \left(-\frac{\partial}{\partial s} (\tilde{\gamma}(\theta) \kappa) + F \cos \theta \right) = \text{const.} = j^*. \quad (3.1)$$

Wir werden von nun an voraussetzen, dass die Steifigkeit keine Orientierungsabhängigkeit aufweist, d.h. $\tilde{\gamma} = \text{const.}$. Benutzen wir nun für die Krümmung $\kappa = \partial_s \theta$, und formen die Gleichung etwas um erhalten wir

$$\frac{\partial}{\partial s} \left(\tilde{\gamma} \frac{\partial \theta}{\partial s} \right) = F \cos \theta - \frac{j^*}{\sigma(\theta)} \quad (3.2)$$

was formal einer Newtonschen Bewegungsgleichung für ein Teilchen in einem Potential entspricht. Es gelten dabei folgende Entsprechungen:

$$\begin{aligned} \text{Teilchenkoordinate} &\leftrightarrow \theta \\ \text{Zeit} &\leftrightarrow s \\ \text{Teilchengeschwindigkeit} &\leftrightarrow \frac{\partial \theta}{\partial s} = \kappa \\ \text{Teilchenmasse} &\leftrightarrow \tilde{\gamma}. \end{aligned}$$

Ferner entsprechen den Wendepunkten der Bewegung Stellen, an denen die Krümmung ihr Vorzeichen wechselt und die Fixpunkte der Dynamik korrespondieren mit Facetten, d.h. Konfigurationen mit $\theta(s) \equiv \text{const.} = \theta^*$.

In dieser Analogie ist die rechte Seite von (3.2) ein ortsabhängiges Kraftfeld, das durch die Ableitung des Potentials

$$V(\theta) = V(0) - \int_0^\theta \left(F \cos t - \frac{j^*}{\sigma(t)} \right) dt \quad (3.3)$$

gegeben ist. Im Folgenden werden wir jeweils zwischen dem geometrischen Bild der Stufenkonfiguration und dem mechanischen Bild der Teilchenbewegung hin- und herwechseln. Um die Form der Stufenkante zu gewinnen, sind nun zwei Integrationen nötig. Die erste erhalten wir direkt durch Ausnutzung der Energieerhaltung

$$\begin{aligned} \frac{\tilde{\gamma}}{2} \kappa^2 + V(\theta) &= \text{const.} = E \\ \Leftrightarrow \kappa &= \sqrt{\frac{2}{\tilde{\gamma}} (E - V(\theta))} \end{aligned}$$

und die zweite durch Integration der beiden geometrischen Relationen

$$\begin{aligned} \frac{d}{d\theta} x(\theta) &= \frac{\cos \theta}{\kappa} \\ \frac{d}{d\theta} y(\theta) &= \frac{\sin \theta}{\kappa} \end{aligned}$$

Aus (3.3) ist Folgendes über die Form des Potentials zu sehen: Da σ immer positiv ist, besteht $V(\theta)$ aus einer Überlagerung eines streng monotonen Anteils (steigend oder fallend, je nach Vorzeichen von j^*) und einer Sinusfunktion (siehe Abbildung 3.1).

Es sind zwei Lösungstypen möglich: Zum einen die ungebundenen Zustände, bei denen das Teilchen gewissermaßen einen endlosen Abhang hinunterrollt und seine Geschwindigkeit κ immer größer wird. Die Form der Stufenkante wäre demnach eine Spirale, an die sich nach einem Wendepunkt symmetrisch dazu eine weitere Spirale anschließt¹. Der für uns interessante Lösungstyp besteht in den gebundenen Zuständen, bei denen sich das Teilchen in einer Mulde zwischen zwei Maxima hin- und herbewegt. Das geometrische Analogon ist eine periodische Zickzacklinie, deren mittlere Steigung dem zeitlich mittleren Aufenthaltsort des Teilchens entspricht. Im nächsten Abschnitt werden wir Kantenprofile endlicher Länge mit periodischen Randbedingungen untersuchen. Dadurch wird dem

¹In diesem Zusammenhang sei am Rande erwähnt, dass der feldfreie Fall eine Clothoide (Cornu-Spirale) liefert, d.h. eine solche Spirale ist bezüglich der Wirkung der Kapillarität zeitlich invariant.

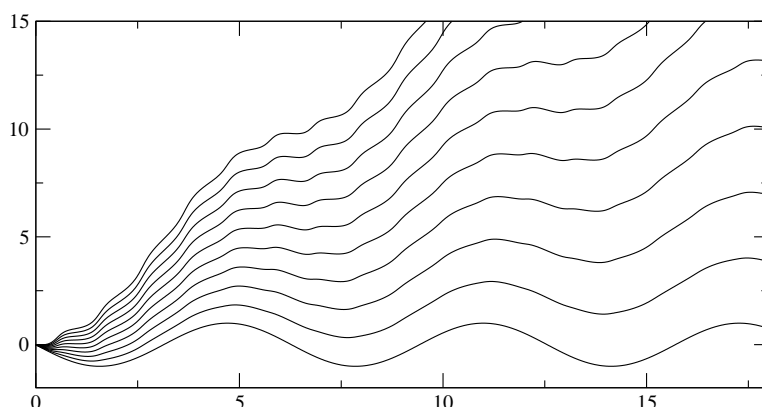


Abbildung 3.1: Teilchenpotential $V(\theta)$ für $n = 6$, $S = 2.0$. Die Werte für j^* sind (von unten nach oben) $0.0, 0.1, 0.2 \dots 0.9$

System die mittlere Steigung $\langle \theta \rangle = 0$ aufgezwungen, das heißt, der Strom j^* und die Wellenlänge der statischen Konfiguration sind nicht mehr unabhängig voneinander, sondern müssen dieser Bedingung gemäß aufeinander abgestimmt sein.

Unter diesen Kantenkonfigurationen ist nun im Bezug auf die Inseloszillationen ein bestimmter Typ von besonderer Bedeutung: Die Bewegung zwischen zwei entarteten Maxima². Insbesondere gibt es hier Oszillationen mit beliebig langer Periodendauer, bei denen sich das Teilchen die meiste Zeit in der Nähe der Maxima aufhält, ohne diese allerdings zu erreichen³. Das heißt, die zugehörige Kantenkonfiguration enthält lange facettenähnliche Stücke mit niedriger Krümmung. Als Grenzfall ergibt sich ein unendlich großer Keil (oder auch „Kink“), dessen Flanken sich asymptotisch an die Facettenorientierungen annähern. Entartete Maxima dieser Art gibt es nur bei bestimmten Werten von j^* , was die möglichen Facettenorientierungen auf eine endliche Auswahl einschränkt, die durch die Forderung nach Entartung des zugehörigen Potentials darüberhinaus explizit bestimmt werden kann⁴. Es wird sich im Folgenden herausstellen, dass die Stufenkante diesem Kink-Lösungstyp entgegenstrebt⁵, so dass sich in diesem Fall die Orientierung der Facetten voraussagen lässt.

²Der „Kink“-Lösungstyp aus [32].

³Allgemein gilt die Aussage, dass ein Teilchen unendlich lange braucht, um ein Potentialmaximum so zu erreichen, dass es dort zum Liegen kommt. Am einfachsten sieht man dies folgendermaßen: Wegen der Zeitumkehrinvarianz von (3.2) ist die Zeit, um das Maximum zu erreichen, genauso lang wie die Zeit, die ein am Maximum ruhendes Teilchen brauchen würde um herabzurollen. Da jedoch das Teilchen am Maximum einfach liegen bleibt, muss diese Zeit unendlich lang sein.

⁴Dazu wird bei vorgegebener Beweglichkeit $\sigma(\theta)$ der Wert von j^* so lange variiert, bis sich aus (3.3) ein Potential $V(\theta)$ ergibt, das zwei benachbarte Maxima derselben Höhe enthält. Die Positionen θ_1 und θ_2 dieser Maxima entsprechen dann den Facettenorientierungen.

⁵Genauer gesagt entwickelt sich in einem endlichen System ein Kink-Antikink-Paar, da die Randbedingungen eine bestimmte mittlere Steigung erzwingen (siehe weiter unten).

Im Prinzip lassen sich diese Ergebnisse auf den Fall mit anisotroper Steifigkeit verallgemeinern. Die Gleichung (3.2) ändert sich nur insoweit, als dass die „Masse“ $\tilde{\gamma}$ ortsabhängig wird. Die Lage der Fixpunkte der Teilchenbewegung wird dadurch allerdings nicht verändert⁶, so dass sich für diesen Fall die *gleichen* Facettenorientierungen ergeben wie für isotrope Steifigkeit.

Symmetrien

Fällt eine der Spiegelachsen der Beweglichkeit mit einer Spiegelachse des Kosinus-Terms in (3.3), d.h. $\theta = 0, \pi$ zusammen, dann ist das resultierende Potential an dieser Stelle θ_s punktsymmetrisch⁷, das heißt Maxima und Minima des Potentials liegen symmetrisch zueinander. Wenn nun beide durch eine weitere Symmetrie miteinander verknüpft sind, zeigt sich eine Symmetrie in den möglichen Facetten-Orientierungen. Die Beweglichkeit weist in unserem Fall eine diskrete Drehsymmetrie bezüglich des Winkels $\frac{2\pi}{n}$ auf, die sich für unser Problem allerdings nur dann auswerten lässt, wenn n eine gerade Zahl ist. Dann enthält die Symmetriegruppe insbesondere die Drehung um π , d.h.

$$\sigma(\theta + \pi) = \sigma(\theta). \quad (3.4)$$

In diesem Fall wechselt die rechte Seite von (3.3) und damit das Potential $V(\theta)$ das Vorzeichen, wenn eine Drehung um π ausgeführt und gleichzeitig j^* durch $-j^*$ ersetzt wird. Diese Operation macht aber gewissermaßen aus einer aufsteigenden eine absteigende Stufenkante bzw. vertauscht im Falle einer Insel Ober- und Unterkante⁸. Der Vorzeichenwechsel verwandelt Maxima in Minima und umgekehrt, so dass, falls eine Spiegelsymmetrie bezüglich θ_s vorliegt, die ausgewählten Orientierungen von Ober- und Unterkante symmetrisch zu θ_s liegen. Neben dieser gibt es noch eine weitere Symmetrie: Die rechte Seite von (3.3) wechselt ihr Vorzeichen, wenn F und j^* gleichzeitig ihr Vorzeichen wechseln. In Verbindung mit einer Spiegelsymmetrie ergibt sich daraus, dass das Kantenprofil invariant bleibt, wenn Ober- und Unterkante vertauscht und die Richtung des Feldes umgedreht wird.

Für das Beispiel $n = 6$ mit Spiegelsymmetrie bezüglich $\theta_s = 0$ funktioniert das Symmetrieargument folgendermaßen: Ein Maximum an der Stelle θ^* ist durch

⁶Die Steifigkeit $\tilde{\gamma}(\theta)$ wird weder Null noch singulär. In der mechanischen Analogie bleibt damit die Masse nach oben und unten beschränkt und die Ruhepunkte der Bewegung sind weiterhin die Extrema des Potentials.

⁷Spiegelsymmetrie der Ableitung impliziert Punktsymmetrie der Funktion

⁸Am einfachsten lässt sich dieser Zusammenhang an einer kreisrunden Insel verdeutlichen: Startet man an der Oberkante am Punkt mit $\theta = 0$ bewirkt eine Drehung des Winkels um π dass der Punkt um die Insel herumwandert, bis er auf der gegenüberliegenden Seite an der Unterkante landet. Die positive Richtung des Massenstroms j ist immer im Uhrzeigersinn (das folgt letztlich aus der Festlegung, dass die Krümmung einer kreisrunden Insel positiv ist), so dass ein in x -Richtung weisender Strom an der Oberkante positiv, an der Unterkante dagegen negativ ist.

die Spiegelsymmetrie mit einem Minimum an der Stelle $-\theta^*$ verknüpft, was wiederum einem Maximum an der Stelle $\pi - \theta^*$ entspricht. Anders ausgedrückt, zu jeder Facettenorientierung θ^* gibt es eine symmetrisch auf der Gegenseite liegende Orientierung $\pi - \theta^*$.

3.1 Stabilitätsanalyse

Nachdem wir die geometrische Form der statischen Kantenprofile untersucht haben, wenden wir uns der Frage nach der Stabilität dieser Lösungen zu. Dazu versehen wir das statische Profil $\theta_0(s)$ mit einer infinitesimalen Störung

$$\theta(s) = \theta_0(s) + \varepsilon\theta_1(s)$$

und entwickeln den Strom (2.4) in ε (wir beschränken uns hier auf den Fall $\tilde{\gamma} \equiv \text{const.}$)

$$\begin{aligned} j &= (\sigma(\theta_0) + \varepsilon\sigma'(\theta_0)\theta_1) \left(-\tilde{\gamma}\partial_s^2(\theta_0 + \varepsilon\theta_1) + F \cos \theta_0 - \varepsilon F \sin \theta_0\theta_1 \right) \\ &= \sigma(\theta_0) \left(-\tilde{\gamma}\partial_s^2\theta_0 + F \cos \theta_0 \right) \\ &\quad + \varepsilon \left(\sigma(\theta_0) \left(-\tilde{\gamma}\partial_s^2\theta_1 - F \sin \theta_0\theta_1 \right) + \sigma'(\theta_0)\theta_1 \underbrace{\left(-\tilde{\gamma}\partial_s^2\theta_0 + F \cos \theta_0 \right)}_{\frac{1}{\sigma(\theta_0)}j^*} \right). \end{aligned}$$

Mit (A-9) lässt sich die Zeitentwicklung der Störung θ_1 nun folgendermaßen schreiben⁹

$$\begin{aligned} \partial_t\theta_1 &= \partial_s v_n = -\partial_s^2 j_1 \\ &= -\partial_s^2 \left(\sigma(\theta_0) \left(-\tilde{\gamma}\partial_s^2\theta_1 - F \sin \theta_0\theta_1 \right) + j^* \frac{\sigma'(\theta_0)}{\sigma(\theta_0)}\theta_1 \right) \\ &= \partial_s^2 \left(\sigma(\theta_0) \left(\tilde{\gamma}\partial_s^2\theta_1 + V''(\theta_0)\theta_1 \right) \right). \end{aligned} \quad (3.5)$$

Die Frage der Stabilität ist demnach die Frage nach der Existenz von positiven Eigenwerten des Differentialoperators auf der rechten Seite. Wir werden dieses Problem hier nicht in voller Allgemeinheit angehen, sondern begnügen uns

⁹Im letzten Schritt benutzen wir

$$\begin{aligned} -V'(\theta_0) &= F \cos \theta_0 - \frac{j^*}{\sigma(\theta_0)} \\ \Rightarrow -V''(\theta_0) &= -F \sin \theta_0 + j^* \frac{\sigma'(\theta_0)}{\sigma(\theta_0)^2} \\ \Leftrightarrow j^* \frac{\sigma'(\theta_0)}{\sigma(\theta_0)} &= \sigma(\theta_0) (F \sin \theta_0 - V''(\theta_0)) \end{aligned}$$

mit folgender Beobachtung: Da die ursprüngliche Dynamik (2.3) translationsinvariant ist, enthält das Spektrum immer den Eigenwert Null. Die zugehörige Eigenmode ist $\partial_s \theta_0$. Zum Beweis¹⁰ differenzieren wir (3.2) und erhalten

$$\begin{aligned}\tilde{\gamma} \partial_s^3 \theta_0 &= -V''(\theta_0) \partial_s \theta_0 \\ \Leftrightarrow \tilde{\gamma} \partial_s^2 (\partial_s \theta_0) + V''(\theta_0) \partial_s \theta_0 &= 0 \\ \Rightarrow \partial_s^2 \left(\sigma(\theta_0) (\tilde{\gamma} \partial_s^2 (\partial_s \theta_0) + V''(\theta_0) \partial_s \theta_0) \right) &= 0.\end{aligned}$$

Wählen wir als Profil $\theta_0 = \text{const.} = \theta^*$ vereinfacht sich Gleichung 3.5 zu

$$\partial_t \theta_1 = -\tilde{\gamma} \sigma(\theta^*) \partial_s^4 \theta_1 - V''(\theta^*) \sigma(\theta^*) \partial_s^2 \theta_1 \quad (3.6)$$

woraus durch Fouriertransformation direkt die Dispersionsrelation

$$\omega(k) = -\tilde{\gamma} \sigma(\theta^*) k^4 + V''(\theta^*) \sigma(\theta^*) k^2 \quad (3.7)$$

folgt. Aus (3.3) und (3.1) folgt nun außerdem für $\theta = \theta^* \equiv \text{const.}$

$$\begin{aligned}V''(\theta^*) &= F \sin \theta^* - j^* \frac{\sigma'(\theta^*)}{\sigma^2(\theta^*)} \\ j^* &= F \sigma(\theta^*) \cos \theta^*\end{aligned}$$

und damit

$$\begin{aligned}\omega(k) &= F (\sigma(\theta^*) \sin \theta^* - \sigma'(\theta^*) \cos \theta) k^2 - \tilde{\gamma} \sigma(\theta^*) k^4 \\ &= A(\theta^*) k^2 - B(\theta^*) k^4\end{aligned} \quad (3.8)$$

Der negative k^4 -Term wirkt stets stabilisierend während der quadratische Term für $V''(\theta^*) > 0$ bzw.

$$\sigma(\theta^*) \sin \theta^* > \sigma'(\theta^*) \cos \theta^*$$

zu einer Instabilität von langwelligen Störungen (kleiner Wellenzahlvektor k) führt. Wie wir im vorangegangenen Abschnitt gesehen haben, ist die notwendige und hinreichende Bedingung für die Existenz einer Facette mit der Orientierung θ^* das Verschwinden der Ableitung $V'(\theta^*) = 0$, so dass sich das eben Gesagte folgendermaßen zusammenfassen lässt: *Mit der Orientierung θ^* lässt sich genau dann eine stabile Facette bilden, wenn*

$$V'(\theta^*) = 0 \quad (3.9)$$

$$V''(\theta^*) < 0 \quad (3.10)$$

d.h. wenn θ^ der Position eines Maximums des Potentials $V(\theta)$ entspricht.*

¹⁰Dieses Argument stammt aus [62].

3.2 Simulationsergebnisse

In diesem Abschnitt vergleichen wir die Vorhersage für die Facettenorientierungen mit den numerischen Ergebnissen. Die Simulationen wurden von Frank Haußer durchgeführt. Das zugrundeliegende Finite-Elemente-Verfahren wird in Abschnitt 5 kurz skizziert.

Wir betrachten eine Stufenkante der horizontalen Ausdehnung L mit periodischen Randbedingungen. Wie zuvor erwähnt, stehen dem System damit nicht mehr alle statischen Konfigurationen zur Verfügung, sondern lediglich diejenigen, deren mittlere Steigung Null ist. Zunächst wählen wir die Parameter so, dass $\theta^* = 0$ einer im Sinne der vorangegangenen Stabilitätsanalyse stabilen Facette entspricht und wählen als Anfangskonfiguration eine Kante mit kleiner, zufälliger Störung. In Abbildung (3.2) ist das Profil zu verschiedenen Zeitpunkten wiedergegeben. Wie zu erwarten gehen die anfänglichen Störungen zurück, die angestrebte Konfiguration ist also tatsächlich die sich über die gesamte Systemlänge erstreckende Facette.

In der ebenfalls zufällig gestörten Kante in Abbildung 3.3 sind die Parameter dagegen so gewählt, dass die horizontale Orientierung instabil ist. Die anfängliche Störung entwickelt sich zu der facettierten Berg- und Talandschaft, die im Abschnitt über die statischen Lösungen beschrieben wurde. Um Facetten zu bilden und gleichzeitig die mittlere Steigung beibehalten zu können, entwickelt die Oberfläche eine Aufeinanderfolge von facettierten Bereichen mit abwechselnd positiver und negativer Steigung. In der Folge vergrößert das Profil dadurch, das sich jeweils ein Kink und Antikink-Paar gegenseitig auslöschen, wodurch die Facetten wachsen und die Wellenlänge der Konfiguration sich vergrößert. Dieser Vorgang wird sich solange fortsetzen, bis die maximale mit den Randbedingungen vereinbare Wellenlänge erreicht ist (Abbildung 3.4).

Wie in Abschnitt 5.3 beschrieben wird, lassen sich die Orientierungen der Endkonfiguration explizit berechnen. Für den in Abbildung 3.4 gezeigten Fall existieren nur für zwei Werte von j^* entartete Maxima. Die beiden zugehörigen Paare von Facettenorientierungen (ausgedrückt als Steigung) sind

$$\begin{aligned} m_1 &= 0.9561 \quad \text{und} \quad -0.3491 \\ m_2 &= -0.41 \quad \text{und} \quad -2.79 \end{aligned}$$

Das zweite Paar ist für unseren Fall ausgeschlossen, da sich mit zwei negativen Steigungen die Randbedingungen nicht erfüllen lassen. Wie Abbildung (3.5) zeigt, wird jedoch das erste Wertepaar mit sehr guter Übereinstimmung realisiert. Interessanterweise treten die m_2 -Orientierungen ebenfalls vorübergehend auf, bevor sie durch Vergrößerung verschwinden (Abbildung 3.6 und 3.7).

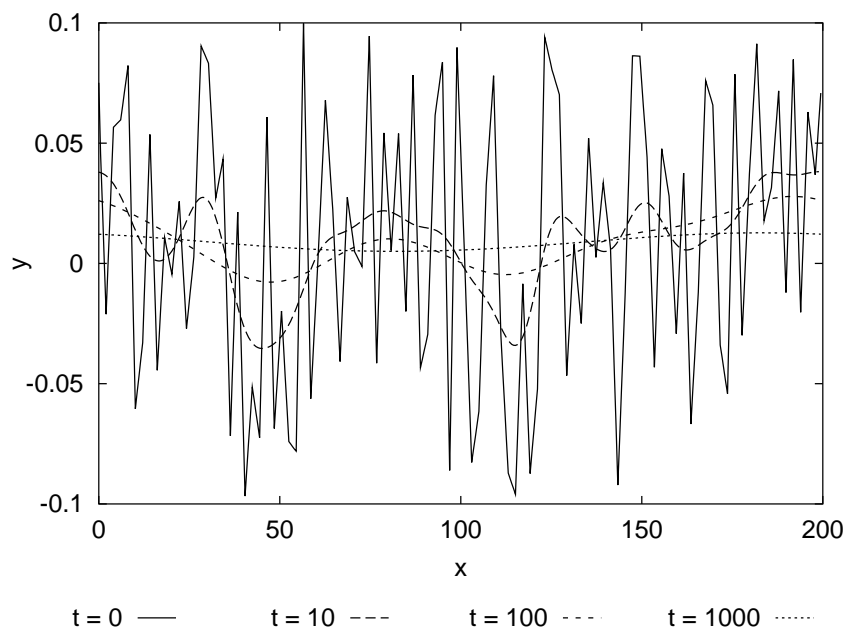


Abbildung 3.2: Zeitentwicklung einer (zufällig) gestörten Stufenkante mit Steigung 0. Die Parameter $n = 6$, $S = 1.0$, $\alpha = \pi/36$ wurden so gewählt, dass eine Facette dieser Orientierung linear stabil ist

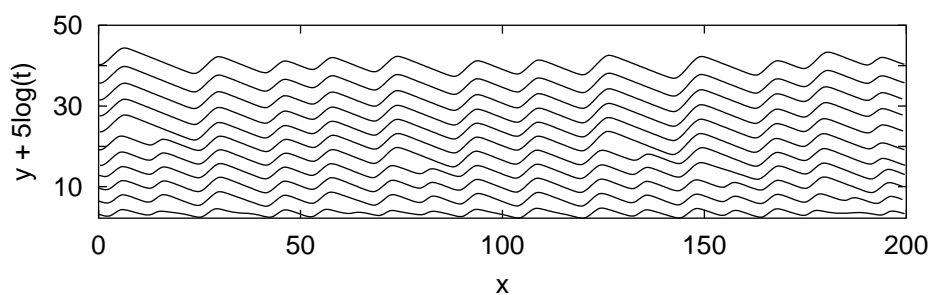


Abbildung 3.3: Facettenbildung und Vergrößerung einer (zufällig) gestörten Stufenkante mit Steigung 0 (zur Zeit $t = 2^n \times 10$, $n = 1, 2, \dots, 10$). Die Parameter $n = 6$, $S = 1.0$, $\alpha = -\pi/36$ wurden so gewählt, dass eine Facette dieser Orientierung linear instabil ist.

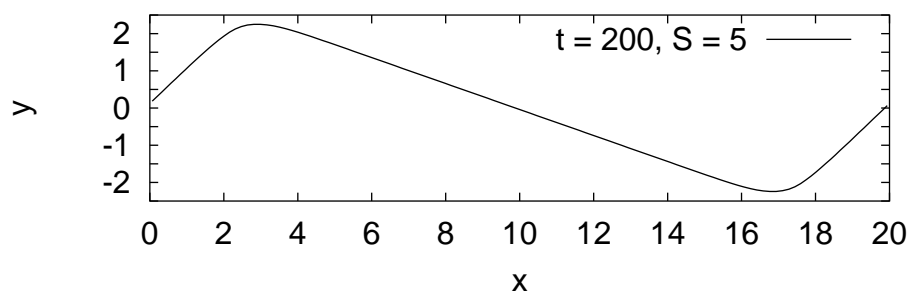


Abbildung 3.4: Endkonfiguration nach 200 Zeitschritten. Anfangskonfiguration war ein Kink-Antikink-Paar mit Steigung ± 1 ($n=6, S=5$)

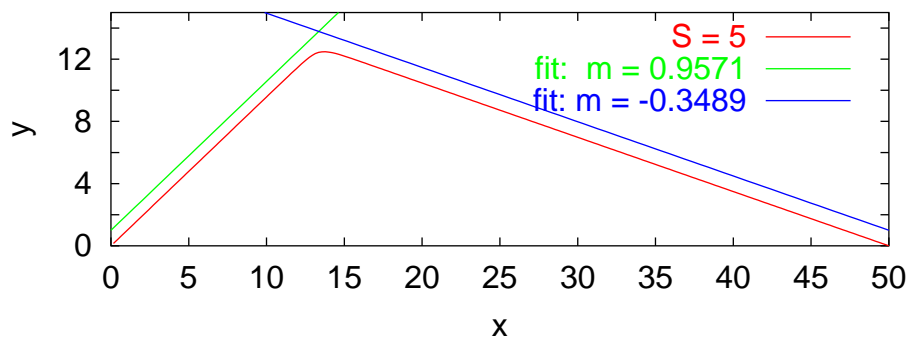


Abbildung 3.5: Linearer Fit an die Facetten. Die vorausgesagten Steigungen sind $m_1 = 0.9561$ und -0.3491

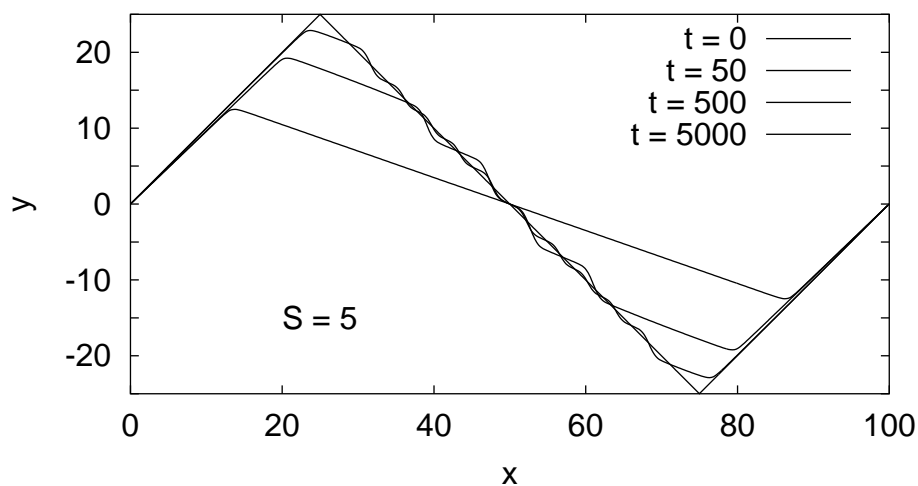


Abbildung 3.6: Auftauchen und Verschwinden der durch die Randbedingungen ausgeschlossenen Orientierungen m_2 . Anfangskonfiguration war ein Kink-Antikink-Paar mit Steigung ± 1 ($n=6$, $S=5$).

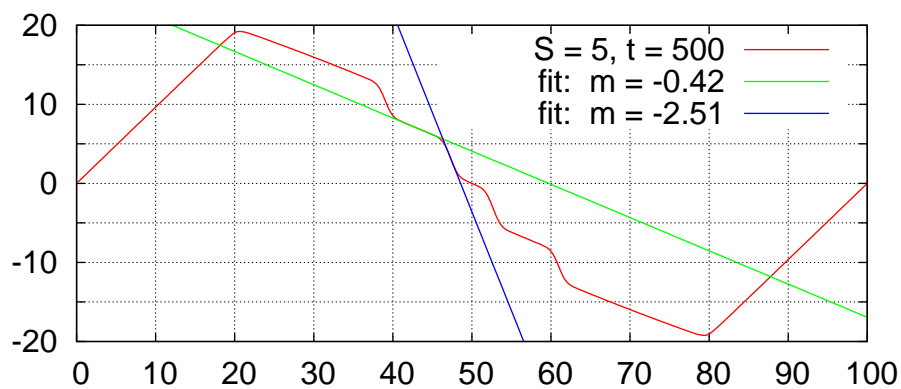


Abbildung 3.7: Linearer Fit an die Facetten. Die vorausgesagten Steigungen sind $m_1 = -0.41$ und -2.79 .

Kapitel 4

Inseldynamik bei isotroper Beweglichkeit

4.1 Die Untersuchungen von Yang, Wang, Suo und Hao

Nun wenden wir die durch (2.3) und (2.4) gegebene Dynamik auf die Inselgeometrie an. Die Steifigkeit wird für den überwiegenden Teil unserer Untersuchung als isotrop angenommen und zunächst werden wir dasselbe von der Beweglichkeit voraussetzen, also

$$\sigma(\theta) \equiv \sigma_0 = \text{const.}$$

Der Strom ist also¹

$$j = \sigma_0 \left(-\tilde{\gamma} \frac{\partial}{\partial s} \kappa + F \cos \theta \right) \quad (4.1)$$

und die Bewegungsgleichung damit

$$v_n = \sigma_0 \frac{\partial}{\partial s} \left(\tilde{\gamma} \frac{\partial}{\partial s} \kappa - F \cos \theta \right) \quad (4.2)$$

Für diesen Fall existieren einige theoretische Untersuchungen aus den Jahren 1994 und 96, die für uns einen Ausgangspunkt bilden und deshalb im Folgenden etwas ausführlicher besprochen werden sollen. Die auf die technologische Relevanz von Elektromigrationsphänomenen ausgerichtete Arbeit von W. Yang, W. Wang und Z. Suo in [70] entstammt dem Kontext der Elektromigration von Poren in einem Festkörper, bzw. in einer dünnen Schicht. Der wesentliche Unterschied zu der von uns behandelten Migration von Oberflächeninseln besteht darin, dass für die Porenbewegung der Effekt des sogenannten „current crowding“

¹Den Winkel ϕ in (2.4) können wir, wie in Abschnitt 2 erläutert gleich Null setzen.

wichtig wird, d.h. dass die Stromverteilung innerhalb des Festkörpers und damit die lokale Elektromigrationskraft durch die Anwesenheit der Pore verändert wird. Deshalb muss das elektrische Potential durch die Laplace-Gleichung $\Delta\Phi = 0$ mit geeigneten Randbedingungen am Porenrand bestimmt werden². Die sich daraus ergebende Dynamik ist im Gegensatz zur Gleichungen (4.2) und (2.4) nichtlokal. Ein Modell, das zwischen den beiden Fällen interpoliert, erhält man, indem man dem Inneren und Äußeren der Pore verschiedene Leitfähigkeiten Σ und Σ' zuordnet. Für $\Sigma'/\Sigma = 0$ erhält man eine nichtleitende Pore in einem leitenden Medium, während für den Fall $\Sigma'/\Sigma = 1$ das Feld von der Pore unbeeinflusst bleibt. Das heißt, das Potential ist das lineare Potential $\Phi(x, y) = Ex$ und das zugehörige Feld konstant, so dass sich dieser Fall als äquivalent zu unserem Modell erweist. Die Autoren von [59] und [70] wendeten dieses Modell auf die Bewegung von Versetzungsschleifen (auch Versetzungsringe genannt) an. Versetzungen sind aus topologischen Gründen immer linienartige Gebilde, die nicht im Kristall aufhören können und deshalb entweder an der Oberfläche enden oder geschlossene Kurven bilden. Da sich hier das Innere und Äußere natürlich physikalisch nicht unterscheiden, haben beide dieselbe Leitfähigkeit und erfüllen so auf natürliche Weise die Modellannahme. Setzt man für die Insel einen Kreis vom Radius R_0 an, d.h. für die Krümmung $\kappa = 1/R_0 \equiv \text{const.}$, so ergibt sich aus (4.2)

$$\begin{aligned} v_n &= -\sigma_0 F \partial_S \cos \theta \stackrel{(4.3)}{=} -\sigma_0 F \kappa \partial_\theta \cos \theta \\ &= \frac{\sigma_0 F}{R_0} \sin \theta. \end{aligned}$$

Dabei haben wir den folgenden, wichtigen Zusammenhang zwischen der Ableitung nach s und der nach θ ausgenutzt:

$$\begin{aligned} d\theta &= \kappa ds \\ \Rightarrow \partial_s &= \kappa \partial_\theta. \end{aligned} \tag{4.3}$$

Nun hat aber die Bedingung für eine stationäre Inselform, d.h. für eine zeitliche Entwicklung, bei der sich die Insel *ohne ihre Form zu ändern* mit Geschwindigkeit V in Richtung des Feldes bewegt, genau die Form (siehe Abbildung 4.1)

$$v_n = V \sin \theta \tag{4.4}$$

Die Kapillarität spielt bei der Kreisform wegen $\kappa = \text{const.}$ keine Rolle und das Feld verschiebt einfach jeden Punkt des Randes um einen konstanten Betrag in

²Die am Festkörper angelegte Spannung, also die eigentliche treibende Kraft, wird durch die Forderung berücksichtigt, dass das Potential Φ in großem Abstand von der Pore asymptotisch gegen das lineare Potential $\Phi(x) = Ex$ gehen soll.

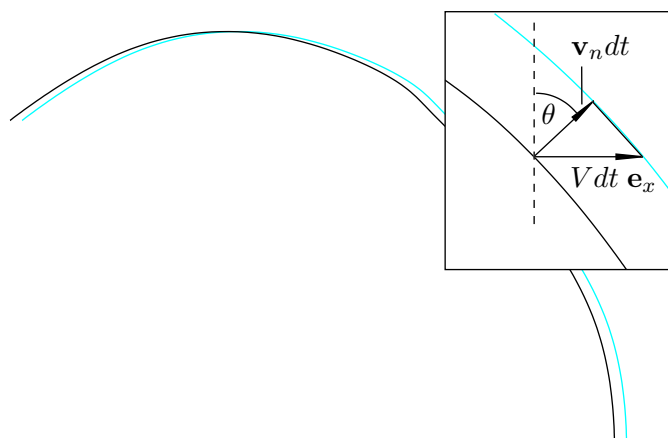


Abbildung 4.1: Geometrische Deutung der Stationaritätsbedingung $v_n = V \sin \theta$. Dargestellt ist ein Teil des Inselrandes zu den Zeitpunkten t und $t + dt$. Die Insel wird während dieses Zeitraumes stationär um $V dt \mathbf{e}_x$ verschoben. Die Bedingung ergibt sich aus der geometrischen Beziehung zwischen $V dt \mathbf{e}_x$ und $\mathbf{v}_n dt$.

Richtung des Feldes, so dass die ganze Insel sich unter Beibehaltung ihrer Form mit der Driftgeschwindigkeit

$$V = \sigma_0 F / R_0 \quad (4.5)$$

bewegt. Diese Tatsache wurde bereits 1970 von P.S. Ho festgestellt [27]. Insbesondere folgt damit die Beziehung

$$V \sim \frac{1}{R_0}$$

zwischen Geschwindigkeit und Inselradius³.

In [70] wurde nun untersucht, inwieweit diese stationäre Lösung instabil werden kann. Zu diesem Zweck wurde nach weiteren stationären Lösungen gesucht und einige heuristische Argumente angeführt, die darauf hindeuten, dass diese Lösungen tatsächlich stabiler als der Kreis sind. Die Gleichung für eine stationäre Form ist zunächst

$$\sigma_0 \frac{\partial}{\partial s} \left(\tilde{\gamma} \frac{\partial}{\partial s} \kappa - F \cos \theta \right) = V \sin \theta \quad (4.6)$$

Unter Verwendung von (4.3) wird dies zu

³D.h. kleinere Inseln driften *schneller*.

$$\begin{aligned}\sigma_0 \tilde{\gamma} \kappa \partial_\theta (\kappa \partial_\theta \kappa) - \sigma_0 F \kappa \partial_\theta \cos \theta - V \sin \theta &= 0 \\ \Rightarrow \partial_\theta (\kappa \partial_\theta \kappa) + \frac{F}{\tilde{\gamma}} \left(1 - \frac{V}{\sigma_0 F} \frac{1}{\kappa} \right) \sin \theta &= 0.\end{aligned}$$

Wählen wir nun $l_S = \sigma_0 F / V$ als Längenskala⁴, so wird daraus

$$\partial_\theta (\kappa \partial_\theta \kappa) + \eta \left(1 - \frac{1}{\kappa} \right) \sin \theta = 0 \quad (4.7)$$

mit der dimensionslosen Größe $\eta = \frac{\sigma_0^2 F^3}{\tilde{\gamma} V^2}$. Der dimensionslose Radius der Insel und ihre Krümmung sind beide gleich Eins (vgl. (4.5)). Die Lösung der Gleichung (4.7) beschreibt stationäre Inseln in der Darstellung $\kappa(\theta)$. Wir suchen nun nach stationären Formen in einer infinitesimalen Umgebung des Kreises, machen also den Ansatz

$$\kappa(\theta) = 1 + \varepsilon \kappa_1(\theta)$$

Einsetzen in (4.7) ergibt in erster Ordnung in ε

$$\partial_\theta^2 \kappa_1 + \eta \kappa_1 \sin \theta = 0, \quad (4.8)$$

und damit eine Form der sogenannten Mathieu-Gleichung, die in Abschnitt A.5 kurz beschrieben wird⁵. Wir suchen nun nach Werten von η , für die nichttriviale, periodische und gerade Lösungen existieren⁶. Dies steht im Zusammenhang mit der Frage, für welche Werte η einer der in Abschnitt A.5 beschriebenen Eigenwerte a_n gleich Null ist⁷. In [70] wurde der kleinste dieser Werte zu $\eta_c = 10.65$ bestimmt. An dieser Stelle entsteht eine Bifurkation in der Lösungsstruktur von (4.7), bei der aus dem Kreis eine weitere Lösung hervorgeht. Genauer gesagt ist natürlich mit κ_1 auch $-\kappa_1$ eine Lösung von (4.8), so dass tatsächlich in der

⁴Das heißt wir reskalieren die Längen mit $l \rightarrow l_S l$ und dementsprechend $\kappa \rightarrow \kappa / l_S$. Der Index S soll daran erinnern, dass diese Längenskala die stationäre Geschwindigkeit V enthält und damit nur für stationäre Bewegungen sinnvoll ist.

⁵Um der allgemeinen Form (A-34) in Abschnitt A.5 zu entsprechen, muss noch die Substitution $\theta = \pi/2 - 2x$ durchgeführt werden.

⁶Die Krümmung muss natürlich notwendigerweise eine 2π -periodische Funktion des Winkels θ sein, um einer geschlossenen Form zu entsprechen (die zugehörige Lösung der Mathieu-Gleichung (A-34) muss dafür π -periodisch sein, da sie, wie weiter oben erwähnt, über $\theta = \pi/2 - 2x$ mit (4.8) zusammenhängt). Eine gerade Lösung entspricht einer Form, die achsensymmetrisch zum Feld ist.

⁷Es stellt sich heraus, dass es der Eigenwert a_2 ist, der die Lösung mit den gewünschten Eigenschaften erzeugt. Der Eigenwert a_1 hat zwar bereits für einen kleineren η -Wert eine Nullstelle, erzeugt aber nicht die gewünschte π -Periodizität.

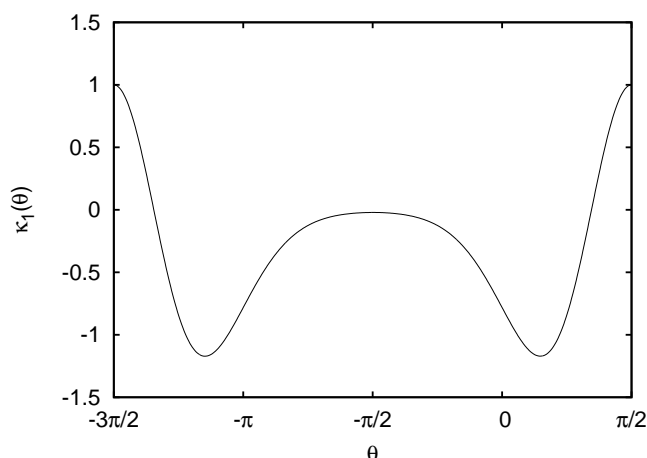


Abbildung 4.2: Lösung der Gleichung (4.8) für $\eta = 10.65$ (Mode I).

Nähe von $\eta = \eta_c$ zwei neue Lösungen existieren, die auf diesem linearen Niveau⁸ symmetrisch zueinander liegen. Folgen wir der Bezeichnung aus [70] und nennen diese beiden Lösungen Mode I und II, so gilt also wegen $\kappa_1^{(I)} = -\kappa_1^{(II)}$ für deren Krümmungen $\kappa_I + \kappa_{II} = 2$. In Abbildung 4.2 ist die Funktion $\kappa_1(\theta)$ der Mode I dargestellt.

Die Zuordnung der θ -Werte zu den Punkten des Inselrandes ist folgendermaßen: Der Punkt zu $\theta = -\pi/2$ befindet sich am hinteren Ende⁹ der Insel, $\theta = \pi/2$ (bzw. $-3\pi/2$) entspricht der vorderen Spitze und die Punkte mit horizontaler Tangente an der Ober- bzw. Unterseite gehören zu $\theta = 0$ bzw. $-\pi$. Aus der Abbildung 4.2 ist demnach zu entnehmen, dass die Krümmung $\kappa + \varepsilon\kappa_1$ sich am hinteren Ende nur wenig von der des Kreises unterscheidet, während die Spitze eine etwas größere Krümmung besitzt. Ober- und Unterseite sind dagegen etwas weniger stark gekrümmt. Für die Mode II gilt entsprechend das Gegenteil. Die Krümmung an Ober- und Unterseite ist stärker und die an der Spitze schwächer. Die Form lässt sich aus der Krümmung wie in Abschnitt 3 durch Integration gewinnen. Der Unterschied zwischen den Formen von Mode I und II besteht im Wesentlichen darin, dass Mode I gegenüber dem Kreis *in Feldrichtung* gestreckt ist, während Mode II *senkrecht zum Feld* verlängert ist (vgl. Abb. 4.3).

Die Autoren von [70] führten neben dieser linearen Analyse auch die numerische Integration der vollen Gleichung (4.7) durch. Die Schwierigkeit besteht hier darin, unter den Lösungen zu verschiedenen Anfangsbedingungen die gesuchte periodische Lösung zu finden. Da die Gleichung (4.7) für $\kappa = 0$ singulär wird, wurden nur Lösungen berechnet, bei denen die Krümmung positiv ist, was ins-

⁸D.h. linear in ε . Bislang beziehen sich alle Aussagen auf einen infinitesimalen Bereich um die Bifurkation herum.

⁹Bezogen auf die Bewegungsrichtung, d.h. das hintere Ende befindet sich links.

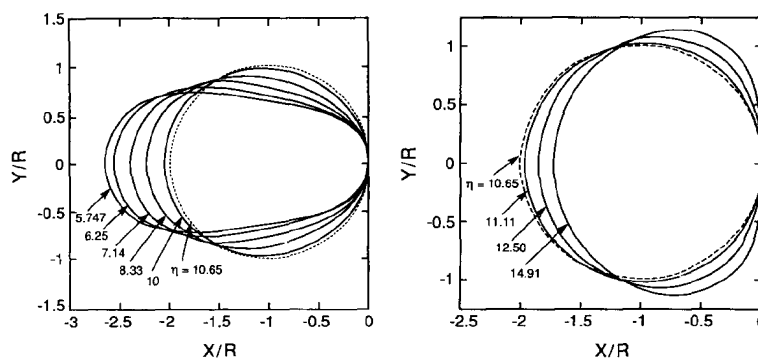


Abbildung 4.3: Die numerisch ermittelten Lösungen von (4.7). Das linke Bild enthält die „Mode I“-Lösungen, die Konfigurationen im rechten Bild gehören zu den Formen von Mode II. Die zugehörigen Werte von η sind eingezeichnet. Die Abbildung ist entnommen aus [70].

besondere heißt, dass diese Formen *konvex* sind. Die so erhaltenen Formen sind in Abbildung 4.3 dargestellt. Wie zu ersehen ist, gehören die Moden I zu Werten $\eta \leq \eta_c$, während die Moden II nur für $\eta \geq \eta_c$ existieren. Der Kreis ist natürlich für jeden Wert von η eine Lösung, so dass zu einem gegebenen η -Wert also jeweils zwei Lösungen existieren. Vergleicht man die Geschwindigkeit von Mode I und II mit der des Kreises, so stellt sich heraus, dass sich Mode I schneller als der Kreis bewegt, während Mode II bei vorgegebenem Feld eine kleinere Driftgeschwindigkeit besitzt. Eine qualitative Erklärung ergibt sich dadurch, dass das Feld am stärksten entlang der horizontalen Orientierung wirkt und diese Orientierung bei Mode I stärker ausgeprägt ist als beim Kreis, während bei Mode II das Gegenteil der Fall ist.

In [68] führten W. Wang, Z. Suo und T.-H. Hao eine lineare Stabilitätsanalyse durch, die in der Tat zeigt, dass der Kreis instabil wird. Dazu wurden die Gleichungen linearisiert und eine Störung des Kreises in Form einer Fourierreihe angesetzt. Das sich daraus ergebende Gleichungssystem für die Fourierkoeffizienten hat die Form eines Eigenwertproblems, wobei ab einem kritischen Wert für η ein positiver Eigenwert entsteht, der die Instabilität anzeigt. In [68] wurde der Bereich für diesen kritischen Wert auf $10 < \eta_c < 11$ eingegrenzt. Eine etwas genauere Behandlung des genannten Gleichungssystems ergibt $\eta_c = 10.649$, in Übereinstimmung mit dem Wert, der durch die Analyse von (4.8) gewonnen wurde. Der Kreis ist also für $\eta < \eta_c$ stabil und verliert seine Stabilität bei η_c zugunsten einer der beiden neuen Lösungen. Die im folgenden Abschnitt dargestellten Ergebnisse der numerischen Behandlung der zeitabhängigen Gleichung zeigen, dass die Stabilität offensichtlich auf den „Mode I“-Typ übergeht, während „Mode II“-Formen auch dann nicht auftraten, wenn als Anfangsbedingung ein senkrecht zur Feldrichtung deformierter Kreis gewählt wurde. Dies ist ein starker Hinweis darauf, dass dieser Fall keiner stabilen Lösung entspricht.

Bevor wir im nächsten Abschnitt weiter auf diese numerischen Simulationen eingehen, werden wir noch einen weiteren Lösungstyp besprechen, der in [60] beschrieben wurde. Es handelt sich nicht um eine Inselform, sondern um einen unendlich langen Finger¹⁰, der stationär in Feldrichtung propagiert. Abbildung 4.4 zeigt eine schematische Darstellung dieses Lösungstyps. Die Seiten des Fingers werden für $x \rightarrow -\infty$ annähernd horizontal, wobei die Krümmung gegen Null geht. In diesen flachen Teilen wird die Wirkung der Kapillarität unwichtig, so dass der Massenstrom nach (4.1) einfach durch $\sigma_0 F$ gegeben ist. Die Massenerhaltung erfordert nun, dass die Geschwindigkeit V , mit der ein Finger der Breite b wächst mit diesem Strom über

$$Vb = 2\sigma_0 F \quad (4.9)$$

zusammenhängt¹¹. Um die Form dieses Fingers zu ermitteln, werden wir zunächst (4.6) etwas umformen. Zuerst bemerken wir, dass aus

$$\left. \begin{array}{l} \frac{dy}{dx} = -\tan \theta \\ ds = \sqrt{(dy)^2 + (dx)^2} \end{array} \right\} \Rightarrow ds = \sqrt{1 + \tan^2 \theta} dx = \frac{dx}{\cos \theta}$$

die Beziehung

$$\frac{dy}{ds} = -\sin \theta \quad (4.10)$$

folgt. Einsetzen in (4.6) ergibt damit die Gleichung¹²

¹⁰Der physikalische Zusammenhang in [60] war allerdings die Dynamik eines Schlitzes innerhalb eines Festkörpers. Die Bildung derartiger Schlitz wurden in Experimenten als Schädigungsmechanismus in dünnen Leiterbahnen identifiziert.

¹¹Die Fläche, um die der Finger an der Spitze pro Zeiteinheit wächst ist Vb . Dieser Zuwachs muss von dem Massenstrom $\sigma_0 F$ (jeweils von Ober- und Unterseite) bereitgestellt werden.

¹²Diese Gleichung gilt natürlich auch bei anisotroper Beweglichkeit

$$Vy = \sigma(\theta) \left(-\frac{\partial}{\partial s}(\tilde{\gamma}\kappa) + F \cos \theta \right).$$

Betrachtet man den (etwas künstlichen) kapillaritätsfreien Fall $\tilde{\gamma} \equiv 0$, so ergibt sich

$$y = \frac{F}{V} \sigma(\theta) \cos \theta.$$

Diese Gleichung kann nun zusammen mit $\frac{dy}{dx} = -\tan \theta \Rightarrow \frac{dx}{d\theta} = -\frac{dy}{d\theta} \frac{1}{\tan \theta}$ dazu verwendet werden, eine Parameterdarstellung für die stationäre Form zu gewinnen. In diesem Fall lässt sich also die *anisotrope* stationäre Form analytisch bestimmen [33].

$$Vy = \sigma_0 \left(-\frac{\partial}{\partial s}(\tilde{\gamma}\kappa) + F \cos \theta \right) \quad (4.11)$$

Aus (4.10) folgt

$$\begin{aligned} -\sin \theta &= \frac{dy}{ds} = \frac{dy}{d \cos \theta} \underbrace{\frac{d \cos \theta}{d\theta}}_{-\sin \theta} \frac{d\theta}{ds} \\ \Rightarrow \kappa &= \frac{d\theta}{ds} = \frac{d \cos \theta}{dy} \\ \Rightarrow \frac{d\kappa}{ds} &\stackrel{(4.10)}{=} -\sin \theta \frac{d\kappa}{dy} = -\sin \theta \frac{d^2 \cos \theta}{dy^2}. \end{aligned}$$

Einsetzen in Gleichung (4.11) liefert zusammen mit (4.1)

$$\frac{V}{\sigma_0} y = \tilde{\gamma} \sin \theta \frac{d^2 \cos \theta}{dy^2} + F \cos \theta$$

Führen wir nun noch die Bezeichnung $f = \cos \theta$ ein und wählen als Längenskala wieder $l_S = \sigma_0 F / V$ ergibt sich

$$\frac{1}{\eta} (1 - f^2)^{\frac{1}{2}} \frac{d^2 f}{dy^2} + f - y = 0. \quad (4.12)$$

In den reskalierten Einheiten hat der Finger nach (4.9) die Breite 2, d.h. y läuft von -1 bis 1 . Als Anfangsbedingung lässt sich deshalb

$$\begin{aligned} f(-1) &= 1 \\ \frac{df}{dy}(-1) &= 0 \end{aligned}$$

ansetzen, d.h. wir fordern horizontale Orientierung ($\cos \theta = 1$) und verschwindende Krümmung $\kappa = \frac{df}{dy}$ an der Unterseite. Mit dieser Anfangsbedingung lässt sich (4.12) von $y = -1$ bis 0 integrieren. Bei $y = 0$, d.h. an der Spitze wird die Gleichung singulär, fordern wir aber dort zusätzlich $f(0) = 0$ (vertikale Orientierung), lässt sich diese Lösung mit einem entsprechenden symmetrischen Gegenstück ergänzen. Diese Forderung ist nur für $\eta = 5.74$ erfüllt, was interessanterweise dem niedrigsten Wert für η entspricht, für die in [70] eine Mode I-Lösung ermittelt werden konnte (siehe Abb. 4.3). Im folgenden Abschnitt wird sich zeigen, dass diese Finger-Lösung auch für die Inselgeometrie eine gewisse Bedeutung besitzt.

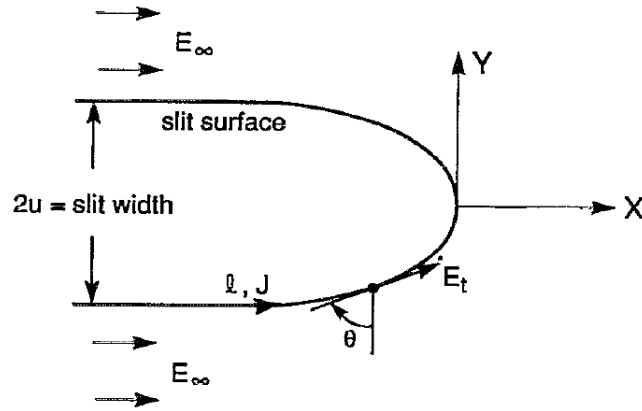


Abbildung 4.4: Schematische Darstellung der Finger-Lösung von (4.12). Die Abbildung ist entnommen aus [60]. Die Bezeichnungen unterscheiden sich etwas von unserer Darstellung und die Definition des Winkels θ ist eine andere (die Umrechnung auf den von uns benutzten Winkel ist $\theta \rightarrow \pi - \theta$).

4.2 Numerische Ergebnisse für das zeitabhängige Problem

Nach dieser Zusammenfassung bereits existierender Untersuchungen zum isotropen Fall, werden wir nun unsere eigenen Arbeiten diskutieren. Statt der stationären Gleichung (4.6) wurde die volle zeitabhängige Gleichung (4.2) numerisch integriert¹³. Die dabei verwendete Längen- und Zeitskala ist (2.6). Der Zusammenhang mit dem im vorherigen Abschnitt verwendeten Größen ist¹⁴

$$\eta = \left(\frac{l_S}{l_E} \right)^2.$$

Für einen Kreis ist der Radius laut (4.5) $R_0 = l_S$, so dass in diesem Fall $\eta = (R_0/l_E)^2$ ist. Für eine beliebige stationäre Form ist in den Einheiten (2.6) $\eta = 1/V^2$. Die Bewegungsgleichung selbst lautet in dieser dimensionslosen Form

$$v_n = \frac{\partial}{\partial s} \left(\frac{\partial}{\partial s} \kappa - \cos \theta \right)$$

Wir erwarten selbstverständlich, dass sich die oben beschriebenen stationären Formen auch in der Simulation zeigen werden und untersuchen deshalb das

¹³Die dabei verwendete Methode wird in Kapitel 5 beschrieben.

¹⁴Wie bereits erwähnt, ist die Längenskala l_S nur für stationäre Bewegungen anwendbar, da sie die stationäre Geschwindigkeit V enthält.

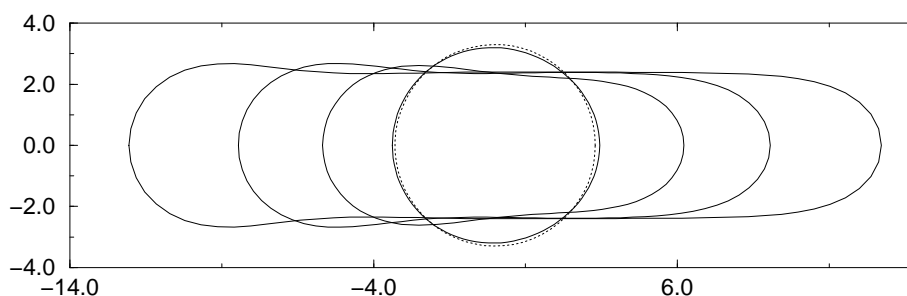


Abbildung 4.5: Stationäre Inseln für den Fall isotroper Beweglichkeit. Die ursprünglichen (dimensionslosen) Radien der Anfangskonfigurationen, die zu den dargestellten Formen gehören, sind $R_0 = 3.3, 4, 5$ und 6 . Die gestrichelte Form ist ein Kreis mit $R_0 = R_c \simeq 3.26$.

Langzeitverhalten der Inselbewegung. Als Anfangskonfiguration wurde in der Regel ein Kreis gewählt, eventuell mit einer Deformation versehen, um eine bestimmte stationäre Form anzusteuern¹⁵. Der Radius dieser Anfangsbedingung ist in den Einheiten (2.6) für das stationäre Verhalten der einzige Kontrollparameter, den wir wie im vorangegangenen Abschnitt mit R_0 bezeichnen werden. Für kleine Radien zeigte sich für alle von uns gewählten Anfangsbedingungen, dass sich die Insel wieder zum Kreis zusammenzieht, was die in [68] durchgeführte Stabilitätsanalyse bestätigt und darüber hinaus darauf hinweist, dass der Kreis vermutlich die einzige stabile Lösung ist. Für größere Radien entsteht ab einem kritischen Wert R_c die vorhergesagte Bifurkation in die gestreckte Mode I-Formen. Wir ermittelten den Wert zu $R_c \simeq 3.26$, was sich auf $\eta_c = R_c^2 \simeq 10.63$ umrechnet, in guter Übereinstimmung mit dem Wert $\eta_c = 10.65$ aus [70]. Mit steigendem Wert von $R > R_c$ wird die Verlängerung in Feldrichtung zunehmend stärker. Die entsprechenden Formen sind Abbildung 4.5 dargestellt.

Die Inseln entsprechen bis etwa $R_0 \simeq 3.5$ den Moden I aus [70]. Ab $R_0 \geq 3.5$ sind die Formen jedoch nicht mehr konvex. Das heißt, die Krümmung wechselt das Vorzeichen und die zugehörigen Formen liegen damit außerhalb des Bereichs, der in [70] durch numerische Integration von (4.7) abgedeckt wurde. Die Streckung der Inseln wird mit steigendem R_0 immer ausgeprägter, bis sich der anfängliche Kreis ab etwa $R_0 \simeq 6.2$ nicht mehr zu einer stationären Form ausprägt, sondern in kleinere Inseln zerfällt. Hier zeigt sich jedoch eine klare Abhängigkeit von der Anfangskonfiguration. Wählt man von Anfang an eine in Feldrichtung langgezogene Form, so entwickelt sich diese auch über $R_0 = 6.2$ hinaus in eine stationäre Form, deren Aspektverhältnis von horizontaler zu vertikaler Ausdehnung mit der Größe der Insel weiter zunimmt. Wählt man umgekehrt eine senkrecht zur Feldrichtung gedehnte Insel, so lässt sich

¹⁵Falls nicht explizit etwas anderes angegeben ist, sind die folgenden Ergebnisse von der genauen Anfangsbedingung unabhängig.

Tabelle 4.1: Die dimensionslose Größe η für einige Werte von R_0 . Diesen Werten liegen Simulationen mit $N = 100$ Knoten zugrunde. Es wurde am Beispiel von $R_0 = 5$ sichergestellt, dass die Ergebnisse für $N = 120$ und 150 dieselben sind.

R_0	η
3.5	6.570
4.5	5.920
5.0	5.778
6.0	5.741
7.0	5.738

das Zerfallen auch für Werte unterhalb von 6.2 herbeiführen. Zwar wurden in diese Richtung keine systematischen Untersuchungen durchgeführt, jedoch gibt es Hinweise, dass sich eventuell beliebig langgestreckte Inseln erzeugen lassen. Berechnet man für einige Fälle die dimensionslose Größe $\eta = 1/V^2$, so ergeben sich die Werte in Tabelle 4.1.

Damit ergibt sich ein gewisser Widerspruch zu den Ergebnissen aus [70]. So ist in Abbildung 4.3 eine Form zum Wert $\eta = 5.747$ dargestellt, was nach unseren Ergebnissen aber schon längst einer nichtkonvexen Form entsprechen sollte¹⁶. Die Behauptung in [70] war, dass sich η in der Umgebung der Stelle, an der die Krümmung Null wird, dem Wert $\eta = 5.743$ der Finger-Lösung aus ([60]) annähert. Auch dies stimmt nicht mit unseren Beobachtungen überein, bei denen sich für $R = 3.5$, wo sich ungefähr der Übergang zu den nichtkonvexen Formen vollzieht, ein Wert von $\eta = 6.570$ ergibt. Unsere Ergebnisse sind jedoch insoweit konsistent, als sich diese Annäherung an $\eta = 5.743$ für die immer stärker langgestreckten Formen zu vollziehen scheint. Insofern ließe sich der unendlich lange Finger als ein Grenzfall der gestreckten Formen auffassen.

Kommen wir auf das Zerfallen in kleinere Inseln zurück. In Abbildung 4.6 ist dieses Zerbrechen oder Zerfließen für einen Kreis mit Radius $R_0 = 7$ dargestellt. In der Anfangsphase bildet sich relativ langsam eine Ausbuchtung an der Spitze der Insel, die sich dann zunehmend schneller in einen langen Finger ausprägt. Wie der Abbildung zu entnehmen ist, bewegt sich die Spitze dieses Fingers wesentlich schneller als der Rest der Insel. Mit wachsendem Finger bildet sich eine Einschnürung, die mit der Zeit immer dünner wird. In der letzten Phase des Einschnürvorganges bilden sich zwei „negative“ Finger ober- und unterhalb des dünner werdenden Halses, die sich in die Insel eingraben. Diese Finger erreichen den hinteren Rand der Insel (und zerteilen damit die Insel in *drei* Teile), sogar noch bevor der Hals unendlich dünn geworden ist¹⁷. Die weitere Entwicklung

¹⁶Es sei daran erinnert, dass die Formen nur bis etwa $R_0 \simeq 3.5$ konvex sind. Der zweite Wert in Tabelle 4.1 entspricht also einer nichtkonvexen Form, gehört aber zu einem größeren η -Wert.

¹⁷In keiner der im Rahmen dieser Arbeit durchgeführten Simulationen wurde die zeitli-

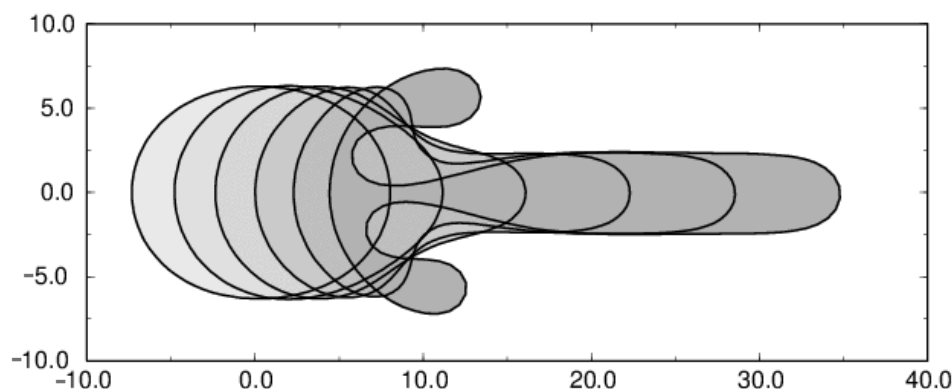


Abbildung 4.6: Zerfließen einer Insel im isotropen Fall. Abgebildet ist die Inselform zu verschiedenen Zeitpunkten, wobei der Grauton, mit dem jede Form eingefärbt ist, die zeitliche Abfolge (von hell nach dunkel) besser sichtbar machen soll. Die Anfangskonfiguration ist ein leicht in Feldrichtung gestreckter Kreis mit $R_0 = 7$ (die Streckung beschleunigt lediglich das Einsetzen der Fingerbildung). Das elektrische Feld erzeugt anfangs eine Ausbuchtung an der Spitze, die sich zu einem langen Finger entwickelt und schließlich zum Zerbrechen der Insel führen wird. Die leichte Asymmetrie in der letzten Phase der Entwicklung ist eine Folge des numerischen Rauschens.

lässt sich daher nicht zweifelsfrei vorhersagen. Entweder wird auch der Hals zerbrechen, so dass am Ende vier Teile entstanden sind, oder die Kapillarität wird die beiden durch den Hals verbundenen Teile wieder zusammenziehen, so dass eine größere Insel in der Mitte entsteht, die von zwei kleineren Satelliteninseln begleitet wird.

In den hier auftretenden Fingern erkennen wir den Lösungstyp aus [60] wieder. Die Vorhersage für die Geschwindigkeit, mit der der Finger propagiert, lautet $V = 1/\eta^2 = 0.417$, was sich sehr gut mit dem aus unseren Simulationsdaten gewonnenen Wert $V = 0.41 \pm 0.01$ deckt¹⁸. Darüber hinaus ergibt sich aus (4.9) in unseren dimensionslosen Einheiten $Vb = 2$, was mit $Vb = 2.03$ aus den Simulationsdaten ebenfalls sehr gut erfüllt ist.

che Entwicklung über eine derartige topologische Änderungen hinaus implementiert. Da die Simulation die Insel einfach als geschlossene Kurve behandelt, entstehen an diesen Stellen unphysikalische Selbstüberschneidungen.

¹⁸Die Geschwindigkeit des Fingers lässt sich nicht mit vergleichbar hoher Präzision bestimmen wie die Schwerpunktschwindigkeit der Insel. Die Spitze enthält auch bei $N = 150$ nur relativ wenige Knotenpunkte, so dass das vordere Ende nur verhältnismäßig schlecht definiert ist (um die vordere Spitze einwandfrei lokalisieren zu können, müssen dort möglichst viele Knotenpunkte vorhanden sein).

Kapitel 5

Inseldynamik bei anisotroper Beweglichkeit

Der wesentliche neue Beitrag, den wir in diesem Kapitel gegenüber den im vorhergehenden Abschnitt besprochenen älteren Untersuchungen leisten wollen, ist es, den Einfluss der Anisotropie der Kristalloberfläche auf die Inselbewegung zu berücksichtigen [34]. Im Prinzip gibt es zwei Einflüsse, die zu einer Anisotropie führen. Zum einen die Orientierungsabhängigkeit der freien Energie der Stufenkante, die in $\tilde{\gamma}(\theta)$ eingeht und zum anderen die Beweglichkeit, die zu einer *kinetischen* Anisotropie führt. Wir sind daran interessiert, diese Einflüsse getrennt zu studieren und werden deshalb zunächst nur die Richtungsabhängigkeit der Beweglichkeit berücksichtigen. Rechnungen, bei denen auch die Steifigkeit von der Orientierung abhängt, weisen jedoch darauf hin, dass sich die meisten Ergebnisse dadurch qualitativ nicht ändern.

Für die Form der Beweglichkeit benutzen wir wieder die in Abschnitt 2 eingeführte Abhängigkeit von θ . Damit wird neben der Größe der Insel als weiterer Kontrollparameter die Stärke S der Anisotropie eingeführt. Wir untersuchen also im Folgenden einen zweidimensionalen Parameterbereich im Bezug auf die sich daraus ergebenden Inselbewegungen. Die Abhängigkeit von der Anfangsbedingung ist hier wieder, falls nicht anders gekennzeichnet, von untergeordneter Bedeutung¹, im Allgemeinen werden wir wieder einen (evtl. deformierten) Kreis benutzen, so dass sich die Ausdehnung der Insel wiederum durch den anfänglichen (dimensionslosen) Radius R_0 ausdrücken lässt. Die meisten Rechnungen wurden für den sechszähligen Fall $n = 6$ und $\alpha = 0$ durchgeführt². Die Bewegungsgleichung lautet in der dimensionslosen Form

¹Ähnlich wie bei dem Zerbrechen im vorangegangenen Kapitel lässt sich die Insel durch die Anfangsbedingung zwar destabilisieren, abgesehen davon führen jedoch verschiedene Anfangskonfigurationen immer zu denselben Bewegungsformen.

²Der sechszählige Fall erscheint uns im Hinblick auf eine mögliche experimentelle Realisierung besonders wichtig. $\alpha = 0$ bedeutet, dass Kanten, die parallel zum Feld orientiert sind, einem Maximum der Beweglichkeit entsprechen und die Kanten senkrecht dazu minimale Beweglichkeit besitzen.

$$v_n = \frac{\partial}{\partial s} \sigma(\theta) \left(\frac{\partial}{\partial s} \kappa - \cos \theta \right). \quad (5.1)$$

Für die Integration wurden zwei verschiedene numerische Verfahren benutzt. Das von uns verwendete ist eine Modifikation des in [53] (und ausführlicher in [51]) beschriebenen Algorithmus für die Berechnung von Poren-Elektromigration. Darin wird der Inselrand durch N äquidistante Knotenpunkte diskretisiert, deren Bewegung dann durch ein System gekoppelter (nichtlinearer) Differentialgleichungen bestimmt ist. Diese werden dann mittels eines Prädiktor-Korrektor-Verfahrens integriert. Da die Verschiebung in Normalenrichtung jedoch ungleichmäßige Abstände zwischen den Knoten erzeugt, wird während des Integrationsschritts eine Korrektur *in Tangentialrichtung* vorgenommen³, die durch ein implizites Gleichungssystem bestimmt wird. Auf diese Weise wird sichergestellt, dass die Knoten während der gesamten Zeit äquidistant verteilt bleiben.

Die Simulationen, die von Frank Haußer durchgeführt wurden, verwenden für die Ortsdiskretisierung der Kurve parametrische lineare finite Elemente. Eine Formulierung der Evolutionsgleichung 4. Ordnung als ein System von Gleichungen 2. Ordnung führt dann zu einem semiimpliziten Verfahren für die Berechnung der Zeitentwicklung der diskreten Kurve mit guten Stabilitätseigenschaften. Darüberhinaus wird das Gitter der Knotenpunkte – falls notwendig – in jedem Zeitschritt adaptiv verfeinert, bzw. vergrößert. Eine detaillierte Darstellung ist in [4] zu finden. Für kleine bis mittlere Inselgrößen ergeben beide Algorithmen konsistente Ergebnisse. Das oben beschriebene Finite-Differenzen-Verfahren stellte sich für kleine Inseln als etwas effizienter heraus, während das Finite-Elemente-Verfahren für größere Inseln wesentlich bessere Resultate liefert.

Wir werden die verschiedenen Bewegungsarten jeweils zuerst phänomenologisch beschreiben und dann Beziehungen zwischen ihnen herstellen und für einige Aspekte darüber hinaus analytische Betrachtungen durchführen.

5.1 Stationäre Bewegungsformen

Für kleine Inseln erhalten wir, wie im isotropen Fall, im Langzeitverhalten stationäre Formen. Durch die anisotrope Beweglichkeit handelt es sich jedoch nun nicht mehr um Kreise, sondern die in Abbildung 5.1 dargestellten Konfigurationen. Für schwache bis mittlere Anisotropie ist die Form noch annähernd kreisförmig, während die Abweichung von der Kreisform mit steigendem S immer deutlicher wird. Für starke Anisotropie bilden sich abgeflachte Bereiche entlang des Randes aus, in denen der Krümmungsradius relativ groß wird. Wie schon im Falle der facettierten Stufenkanten in Abschnitt 3 ist dies ein rein *kinetischer* Effekt; bei Abschalten des elektrischen Feldes würde die Insel wieder zu

³Eine Verschiebung in Tangentialrichtung entspricht einer Umparametrisierung des Inselrandes und verändert die „physikalische“ Bewegung des Randes *nicht*.

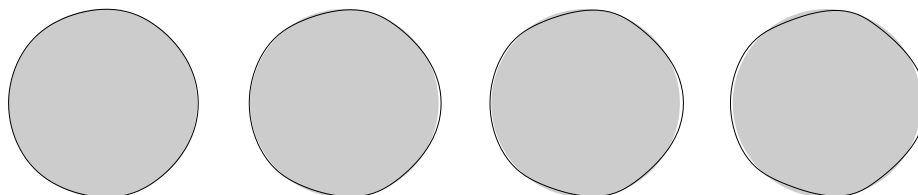


Abbildung 5.1: Stationäre geradeauslaufende Formen für $R_0 = 2$ und $S = 1, 3, 5$ und 10 (von links nach rechts). Die graue Form im Hintergrund ist jeweils der Kreis mit Radius R_0 . Mit steigender Anisotropie bilden sich entlang des Randes abgeflachte Bereiche aus. Die Simulation wurde mit $N = 100$ Knoten durchgeführt.

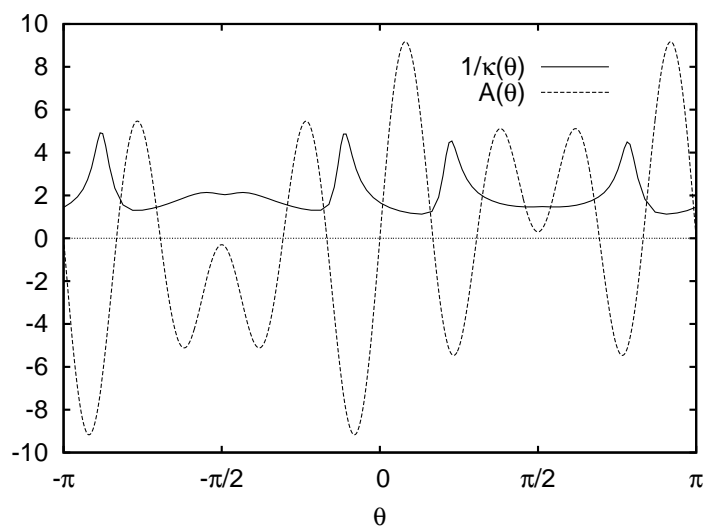


Abbildung 5.2: Lage der bevorzugten Orientierungen der stationär geradeauslaufenden Inselformen für $R = 2$ und $S = 10$ (Abbildung 5.1 ganz rechts). Dargestellt ist der Krümmungsradius $1/\kappa(\theta)$ und der Stabilitätskoeffizient $A(\theta)$ aus Gleichung (3.8). Die Skala an der y -Achse gilt nur für die Größe $1/\kappa(\theta)$.

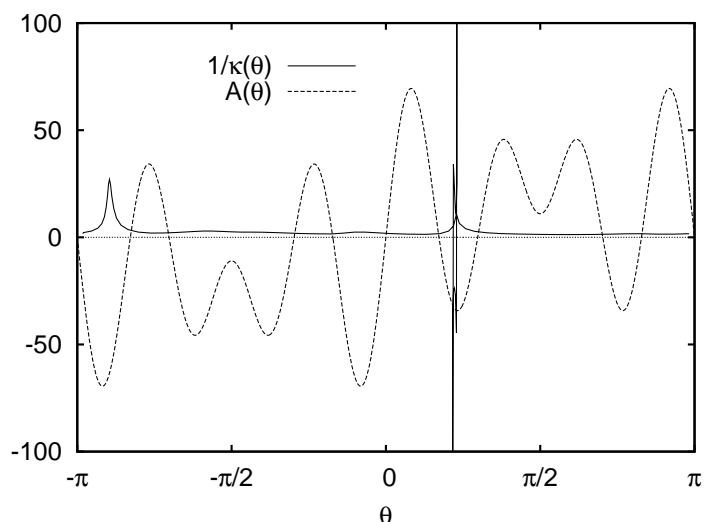


Abbildung 5.3: Lage der bevorzugten Orientierungen der stationär schräglauflenden Inselformen für $R = 2.7$ und $S = 2$ (Abbildung 5.4 Mitte). Die Funktion $1/\kappa(\theta)$ ist aufgrund der Nichtkonvexität der Form nicht eindeutig. Im Bereich der leichten Einbuchtung ergeben sich an jeweils zwei Stellen die gleichen Orientierungswinkel θ . Außerdem entstehen aufgrund des Vorzeichenwechsels der Krümmung Singularitäten, an denen $1/\kappa$ von ∞ nach $-\infty$ springt und umgekehrt.

der thermodynamischen Gleichgewichtsform des Kreises zurückkehren. Obwohl es sich hier keineswegs um statische Strukturen handelt, können wir bis zu einem gewissen Grad die Analyse aus Abschnitt 3 anwenden. Für eine (unendlich ausgedehnte) gerade Stufenkante der Orientierung θ entscheidet das Vorzeichen des Terms $A(\theta)$ in (3.8) darüber, ob die Stufenkante stabil ist oder nicht. Wir wenden dieses Stabilitätskriterium nun auf die Orientierungen der statischen Konfiguration an, für welche die Formel (3.8) wegen $\theta \neq \text{const.}$ zwar streng genommen nicht gilt, aber trotzdem zumindest *lokal* einen gewissen Aussagewert haben sollte. In Abbildung 5.2 ist dazu der Krümmungsradius $1/\kappa(\theta)$ zusammen mit dem „Stabilitätskoeffizienten“ $A(\theta)$ dargestellt⁴. Der Krümmungsradius gibt Aufschluss darüber, welche Orientierungen bei der Form besonders ausgeprägt sind⁵. Im Mittel⁶ ist $1/\kappa$ natürlich gleich R_0 , aber wie der Abbildung zu ent-

⁴Die Skala an der vertikalen Achse gilt nur für den Krümmungsradius (in dimensionslosen Einheiten). Der Koeffizient $A(\theta)$ ist für diese gemeinsame Darstellung in vertikaler Richtung etwas gestaucht worden. Wir interessieren uns hier vornehmlich für die Lage der Maxima und Minima, die Amplitude von $A(\theta)$ spielt hier keine Rolle.

⁵Man kann die Formel $d\theta = 1/\kappa ds$ so interpretieren, dass der Krümmungsradius so etwas wie eine Dichtefunktion der Orientierung darstellt, die angibt, wie stark eine bestimmte Orientierung entlang der Bogenlänge gewichtet ist.

⁶Das Mittel $(\int 1/\kappa ds)/(\int ds)$ bezüglich der Bogenlänge.

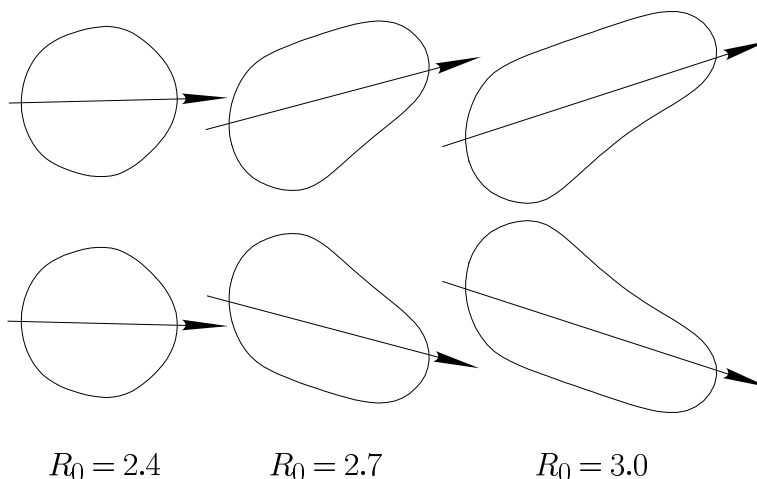


Abbildung 5.4: Stationäre, schräglauflende Inselformen mit ihrer jeweiligen Bewegungsrichtung. Durch die Symmetrie der Gleichung treten diese Formen in spiegelsymmetrischen Paaren auf. Die Stärke der Anisotropie ist $S = 2$. Die Berechnung erfolgte mit $N = 70$ Knotenpunkten.

nehmen ist, sind einige der Orientierungen stärker vertreten als andere. Dies sind genau die abgeflachten Randstücke, deren Richtung jeweils in der Nähe der Richtungen mit minimalem $A(\theta)$ liegt. Die Form stellt sich also so ein, dass die stabilsten Orientierungen bevorzugt und umgekehrt die instabilen Orientierungen vermieden werden. Die Lage der ausgewählten Orientierungen fällt indes nicht exakt mit den Minima von $A(\theta)$ zusammen, die Anwendbarkeit der Formel (3.8) ist hier natürlich begrenzt und erlaubt lediglich eine qualitative Aussage. Wenn wir nun den Radius vergrößern, erfährt der Bewegungsmodus in dem Bereich $2 < R_0 < 3$ eine qualitative Veränderung⁷. Die Bewegung wird zwar weiterhin nach einer gewissen Zeit stationär, doch beobachten wir hier erstmals, dass Bewegungsrichtung und Feldrichtung *nicht* zusammenfallen. Die Insel bewegt sich stattdessen *schräg* zum Feld. Da die Beweglichkeit und damit die gesamte Bewegungsgleichung (5.1) für $\alpha = 0$ spiegelsymmetrisch bezüglich der Feldrichtung ist, handelt es sich hier um eine sogenannte *spontane Symmetriebrechung*⁸. Da jedoch aufgrund der Symmetrie der Gleichung weiterhin mit jeder Lösung auch die gespiegelte Form eine Lösung ist, gibt es zu jeder dieser schräglauflenden Formen einen symmetrischen Gegenpart. Welche der beiden spiegelsymmetrischen Formen angenommen wird, wird durch die Anfangsbedingung festgelegt. In Abb. 5.4 sind einige dieser Formen und ihre jeweilige

⁷Die genaue Lage dieser kritischen Größe hängt von der Stärke S der Anisotropie ab.

⁸D.h. die Lösung der Gleichung besitzt weniger Symmetrien als die Gleichung selbst.

Bewegungsrichtung abgebildet. Für kleine Inselradien ist der Winkel zwischen Bewegungsrichtung und Feld relativ klein und die Form ähnelt noch sehr den weiter oben besprochenen stationär geradeauslaufenden Formen⁹. Für größere Radien ergibt sich eine typische Nierenform, bei der die konvexe Seite eine ausgeprägt facettenartige Form besitzt. Die Orientierung dieser Facette lässt sich wiederum mit der Stabilitätsanalyse für die gerade Stufenkante in Verbindung bringen. In Abbildung 5.3 sind wieder Krümmungsradius und Stabilitätskoeffizient gegenübergestellt. Die Peaks sind hier wegen der stärkeren Facettierung wesentlich ausgeprägter und liegen wieder in der Nähe von Minima der Funktion $A(\theta)$. Es sei hier darauf hingewiesen, dass das System bezüglich der Orientierungen einer zusätzlichen geometrischen Einschränkung unterliegt, da sich die gewählten Orientierungen zu einer geschlossenen Insel zusammensetzen müssen. Insofern lassen sich die stabilen Orientierungen nicht beliebig besetzen. Eine derartige Einschränkung wird weiter unten noch wichtig werden.

5.2 Oszillatorische Bewegung

Bei weiterer Vergrößerung der Inseln stellen wir fest, dass die Inselbewegung eine weitere, noch drastischere qualitative Änderung zeigt. Anstatt in eine stationäre Bewegung hineinzulaufen oder zu zerbrechen, beginnt die Form in einer periodischen Weise zu *oszillieren*¹⁰. Innerhalb des oszillatorischen Regimes lassen sich verschiedene Typen klassifizieren, so dass wir nun, um Überblick zu gewinnen, eine etwas systematischere Karte dieser Bewegungs-Landschaft entwerfen werden. In Abbildung 5.5 ist ein Phasendiagramm dargestellt, das die Bewegungsformen als Funktion der beiden Parameter R_0 und S darstellt¹¹. Im unteren Bereich, d.h. für kleine Radien finden wir die eben besprochenen stationären Formen, die wir kurz mit SS („straight stationary“) für die geradeauslaufenden und OS („oblique stationary“) für die schräglaufenden Inseln bezeichnen werden. Für mittlere bis starke Anisotropie grenzt daran in Richtung größerer Radien der Bereich der *schräglaufend oszillierenden* Bewegungen an (OO für „oblique oscillatory“). Er stellt für mittlere Radien den vorherrschenden Bewegungstyp dar. Eine typische Bewegungsabfolge ist in Abbildung 5.6 dargestellt. Anfänglich bildet sich eine ähnliche Form, wie bei den stationär schräglaufenden Inseln in Abb. 5.4. Danach durchläuft die Insel eine langgestreckte Nierenform, die eine

⁹Die Abhängigkeit zwischen der Steigung der Inselbewegung und der Inselgröße ist in Grafik 5.9 quantifiziert. Wie man dort sieht, besteht zwischen den geradeaus- und schräglaufenden Formen ein kontinuierlicher Übergang.

¹⁰Periodische Formänderungen wurden auch in [21] bei Simulationen von Poren-Elektromigration beobachtet.

¹¹Die Simulationen für das Phasendiagramm wurden von Frank Haußer durchgeführt. Im unteren Bereich, d.h. für kleine bis mittlere Inseln ergeben die beiden oben beschriebenen numerischen Verfahren übereinstimmende Ergebnisse. Für größere Inseln liefert das adaptive Finite-Elemente-Verfahren jedoch wesentlich stabilere Ergebnisse, d.h. in dem Finite-Differenzen-Verfahren zeigt sich ab einer bestimmten Größe die Tendenz der Inseln zu zerbrechen.

deutlich stärkere Einbuchtung an der Unterseite besitzt als die Ausgangskonfiguration. Diese Einbuchtung bewirkt offenbar einen schnelleren Massentransport, so dass sich die Form im letzten Stadium wieder zusammenzieht, wobei sich dieser Ablauf danach periodisch wiederholt. Im mitbewegten Bezugssystem stellt sich dieser Ablauf so dar, dass die Einbuchtung am vorderen Teil der Insel entsteht, dann am Rand entlangwandert, bis sie die Insel am hinteren Ende wieder verlässt. Diese Einbuchtung entsteht hier nur an einer Seite der Insel, wohingegen die andere während der gesamten Bewegung ihre Form beibehält. Die Orientierung dieser stationären Seite liegt wieder in einer Richtung maximaler Stabilität, im Sinne des Koeffizienten $A(\theta)$. Für den hier behandelten Fall $\alpha = 0$ besitzt $A(\theta)$ die Symmetrie $A(\theta + \pi) = -A(\theta)$, d.h. Orientierungen, die um 180° zueinander gedreht sind, haben entgegengesetzte Stabilitätseigenschaften. Liegt also z.B. die stationäre Oberkante in einem maximal stabilen Bereich, so wäre die symmetrisch dazu liegende Unterkante maximal instabil. Das System vermeidet Orientierungen in dieser Richtung indem es die Einbuchtung ausbildet, deren Orientierungen wiederum im stabilen Bereich liegen. Im zeitlichen Mittel bewegt sich die Insel ebenfalls mit einem gewissen Winkel zur Feldrichtung. Der Verlauf der Steigung m der Schwerpunktsbewegung als Funktion von R_0 beim Übergang von $SS \rightarrow OS \rightarrow OO$ ist in Abbildung 5.9 für $S = 2$ dargestellt. Der Übergang von der geradeaus- zur schräglaufenden Bewegung liegt in der Nähe von $R_0 \simeq 2.4$, das Einsetzen der Oszillationen erfolgt bei etwa $R_0 \simeq 3.0$. Beide Bifurkationen hinterlassen eine Signatur im Verlauf von $m(R_0)$. Der erste Übergang entspricht der Stelle, an der die Steigung das erste Mal einen von Null verschiedenen Wert annimmt, der zweite zeigt sich in Form eines Knicks. Bis dahin vergrößert sich der Steigungswinkel monoton mit R_0 , während er nach dem Knick zunächst wieder kleiner wird.

Als Größe, mit der sich die Oszillationen charakterisieren lassen, bietet sich die Periodendauer T an. Diese läßt sich z.B. aus dem Umfang als Funktion der Zeit ablesen¹². Beim Übergang $OO \rightarrow OS$, divergiert die Periodendauer¹³ in einer charakteristischen Weise. Die genaue Lage und Form dieser Singularität ist numerisch jedoch relativ schwierig zu ermitteln. Wir haben zu diesem Zweck sehr umfangreiche numerische Untersuchungen durchgeführt. Das Problem dabei ist, dass sich in der Nähe der Singularität eine starke Abhängigkeit von der Anzahl der verwendeten Knotenpunkte ergibt. Die Ergebnisse konvergieren jedoch für eine ausreichende Zahl von Knoten und es zeigt sich, dass die Divergenz die Form $T \sim |R_0 - R_0^c|^{-\nu}$ mit $R_0^c \simeq 3.024$ und $\nu \simeq 2.5$ besitzt. In den beiden Abbildungen 5.10 ist $T(R_0)$ und $1/T^{2.5}$ für verschiedene Werte von N dargestellt.

Kommen wir nun auf den mit ZZ (Zickzack) gekennzeichneten Bereich in Abbildung 5.5 zu sprechen. Für schwache Anisotropie und mittlere Inselgrößen neh-

¹²Tatsächlich wurde von uns das Trägheitsmoment des Polygons aus den Knotenpunkten verwendet, was sich ebenfalls als Maß für die Formänderung eignet. Dieser Unterschied ist hier jedoch ohne Bedeutung.

¹³So gesehen entsprechen die stationären Bewegungen einer unendlichen Periodendauer.

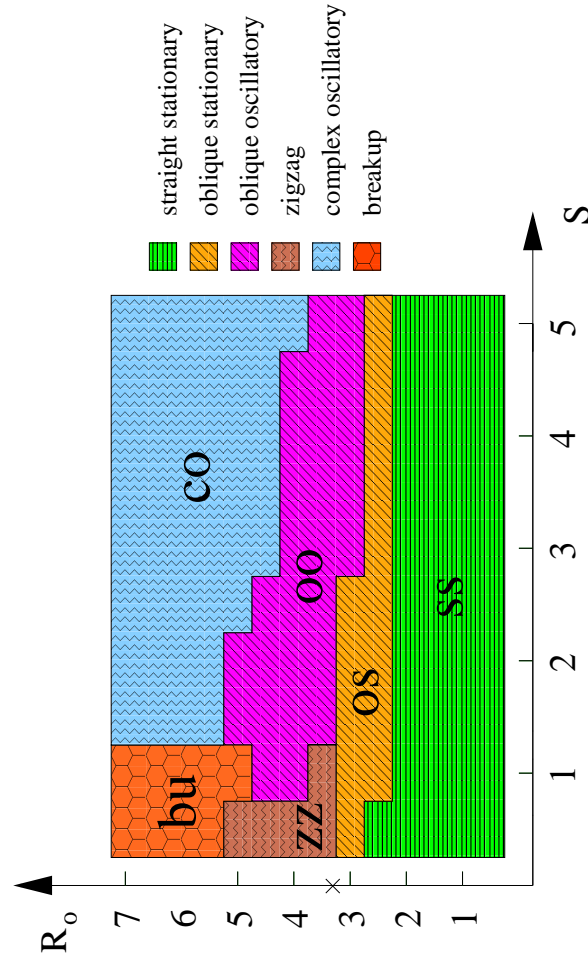


Abbildung 5.5: Phasendiagramm der Migrationsmoden in der Parameter-Ebene von Radius R_0 und Stärke S der Anisotropie. Die Beweglichkeit besitzt sechszählige Symmetrie $n = 6$ und das Feld zeigt in Richtung der Orientierung maximaler Beweglichkeit ($\alpha = 0$). Dem Diagramm liegt ein Raster der Auflösung 0.5×0.5 zugrunde. Für jeden Punkt dieses Rasters wurde die Zeitentwicklung der Insel verfolgt, bis das Langzeitverhalten identifiziert werden konnte. Wir unterscheiden geradeaus- und schräglaufende stationäre Bewegung (SS bzw. OS), schräglaufend oszillierende und Zickzack-Bewegung (OS und ZZ) und schließlich die komplex oszillierende Bewegung (CO). Im oberen, linken Bereich des Diagramms zerbrechen die Inseln (BU). Das Kreuz auf der R_0 -Achse markiert den Punkt, an dem die Kreislösung im isotropen Fall instabil wird (siehe Abschnitt 4).

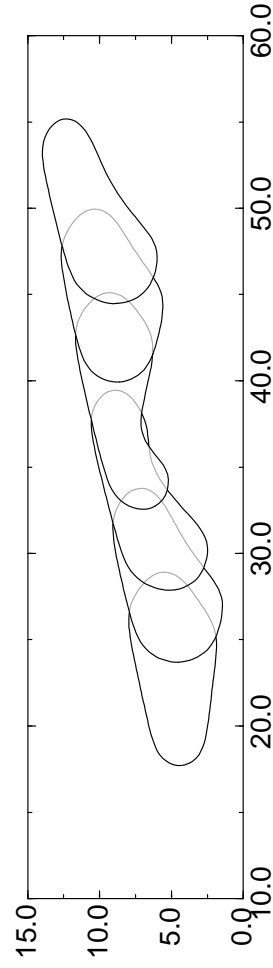


Abbildung 5.6: Typische Oszillation des OO-Regimes für $R_0 = 4$ und $S = 1$. Der zeitliche Abstand zwischen den einzelnen Schnappschüssen ist $\Delta t = 20$. Die Anzahl der Knoten ist $N = 100$.

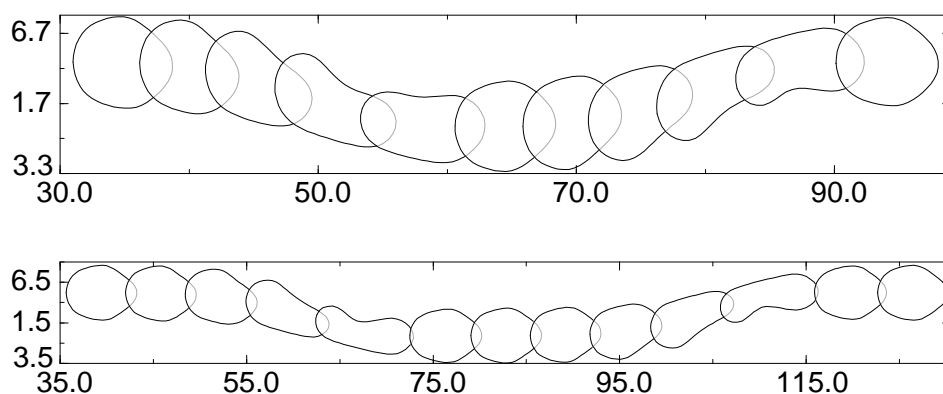


Abbildung 5.7: Bewegungsabfolge der Zickzack-Bewegung (ZZ). Die Parameterwerte für das obere Bild sind $R_0 = 3.5$ und $S = 0.5$, für das untere Bild $R_0 = 3.5$ und $S = 1$. Der Zeitabstand ist $\Delta t = 20$, die Anzahl der Knoten $N = 50$.

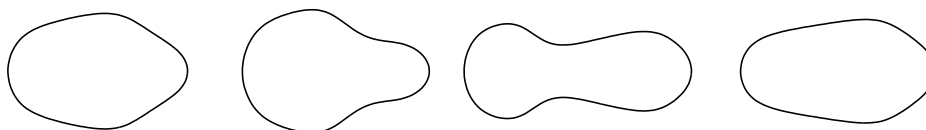


Abbildung 5.8: Grenzfall der Zickzack-Bewegung: Die Einbuchtung erscheint auf *beiden* Seiten der Insel. Die einzelnen Formen sind der Übersichtlichkeit halber auseinandergezogen. Der Zeitabstand ist $\Delta t = 20$. Die Parameterwerte sind $R_0 = 4.5$ und $S = 0.5$ (Simulation von Frank Haußer).

men die Oszillationen einen etwas exotischen Charakter an. Anstatt entweder schräg nach oben oder unten zu laufen, wechseln sich die beiden Bewegungsrichtungen ab, so dass die Insel im Mittel in Feldrichtung driftet (Abbildung 5.7). Der Wechsel zwischen beiden Bewegungsrichtungen kann dabei längere Phasen beinhalten, in denen die Insel eine Zeitlang in horizontale Richtung wandert, bevor sie erneut eine Schrägrichtung einschlägt (siehe Abbildung 5.7 unten). Als Grenzfall der Zickzack-Bewegung ergibt sich in einem sehr kleinen Parameterbereich die in Abb. 5.8 dargestellte Bewegung, bei der *beide* Seiten *gleichzeitig* Einbuchtungen entwickeln, wobei die Insel in Feldrichtung läuft.

Für schwache Anisotropie schließt sich an das ZZ-Regime für größere Radien ein Bereich an, in dem die Inseln zerbrechen. Es konnte anhand der Simulationen nicht mit völliger Sicherheit ausgeschlossen werden, daß sich die Inseln durch eine bestimmte Anfangsbedingung stabilisieren lassen, jedoch scheint die Ten-

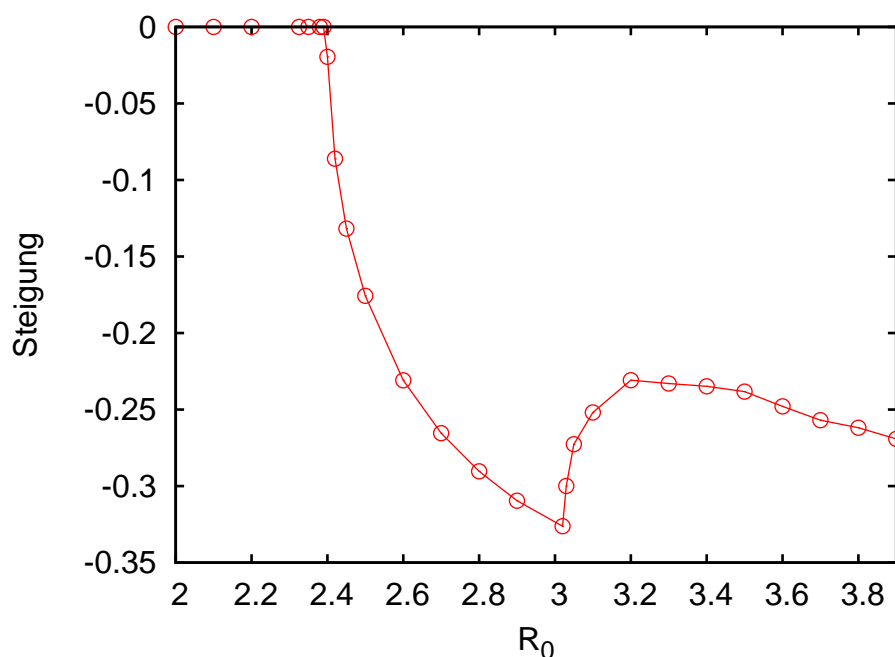


Abbildung 5.9: Steigung der Schwerpunktsbewegung als Funktion von R_0 für $S = 2$. Die Anfangsbedingungen wurde so gewählt, dass die Bewegung schräg nach unten verläuft. Für die entsprechend symmetrisch dazu liegende Anfangsbedingungen ist der Verlauf der Kurve vertikal gespiegelt. Die Oszillationen setzen ab etwa $R_0 \simeq 3.03$ ein. Die Berechnung erfolgte mit $N = 70$ Knotenpunkten.

denz zu zerbrechen sehr stark zu sein.

Für mittlere bis starke Anisotropie beginnt für große Inseln der ausgedehnte Bereich der *komplexen Oszillationen*. An der Grenze zu diesem Bereich erscheint neben der beschriebenen Einbuchtung eine weitere, so dass eine Seite der Insel die Form einer Zickzacklinie annimmt. Sowohl die Ober- als auch die Unterkante sind in diesem Regime stark facettiert wie das obere Bild von Abbildung 5.13 zeigt. Mit dem Auftreten der zweiten Einbuchtung beginnt die Inselbewegung zunehmend unregelmäßig zu werden. In Abbildung 5.11 ist der Umfang als Funktion der Zeit für verschiedene Parameterwerte dargestellt. Bildet man aus einer derartigen Zeitreihe ein Fourierspektrum, so zeigen sich wie in Abbildung 5.12 charakteristische Peaks bei einer Grundperiode und deren Doppelten, bzw. Vierfachen. Damit deutet sich an, dass es sich hier um einen Übergang zu chaotischer Bewegung handelt [42]. Eine weitergehende Untersuchung, wie z.B. eine Charakterisierung des Attraktors wurde im Rahmen dieser Arbeit jedoch nicht durchgeführt.

Die Unregelmäßigkeiten zeigen sich dabei anfangs nur in der Größe der Einbuch-

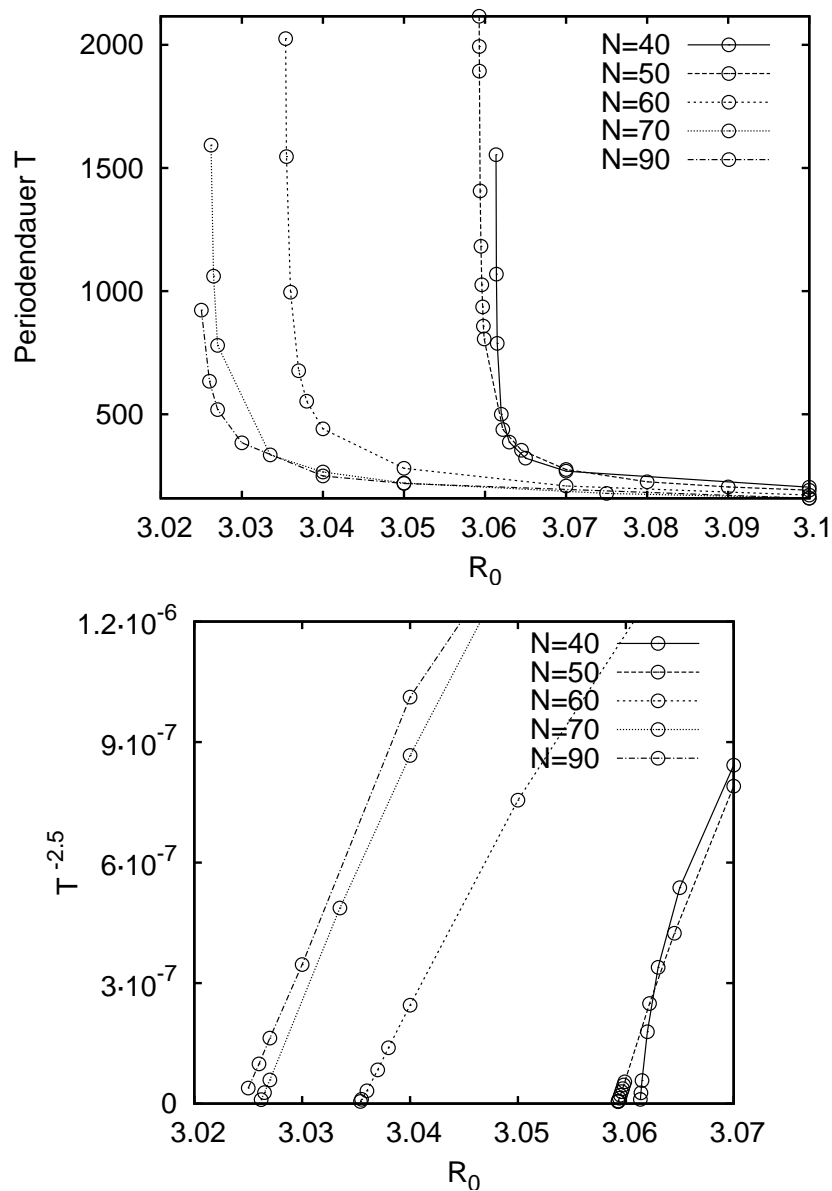


Abbildung 5.10: Divergenz der Periodendauer beim Übergang $OO \rightarrow OS$. Dargestellt ist die Periodendauer T also Funktion des Inselradius R_0 für verschiedene Werte von N .

tungen, für größere Radien ist darüber hinaus jedoch auch die *Bewegungsrichtung* betroffen. Dabei vollzieht die Insel Wechsel zwischen den beiden, spiegelsymmetrisch zueinander liegenden Bewegungsrichtungen. Ein solcher Vorgang ist in Abbildung 5.13 unten dargestellt. Er besteht darin, dass eine der Facetten in der Einbuchtung zu einer langen Kante auswächst und daraufhin die Rolle der stationären Inselseite übernimmt.

5.3 Berechnung der ausgewählten Facettenorientierungen

Es zeigt sich nun, dass für die Facettenorientierungen, die bei diesen Inseln auftauchen, eine analytische Berechnung möglich ist. Der Schlüssel dafür liegt in der Beobachtung, dass die Einbuchtungen, die den Inselrand entlangwandern, *relativ zum Substrat stillstehen*. In Abbildung 5.14 ist eine typische Inselbewegung dargestellt. Die dunkle gestrichelte Linie verdeutlicht, dass es sich bei den Einbuchtungen tatsächlich um statische Elemente handelt¹⁴. Für derartige statische Konfigurationen, genauer gesagt für eine statische Stufenkante mit zwei Facetten-Orientierungen, haben wir nun jedoch in Abschnitt 3 eine Möglichkeit gefunden, die Orientierungen vorherzusagen. Erinnern wir uns noch einmal kurz daran, dass in der dort entwickelten mechanischen Analogie die Stufenkonfiguration einer Teilchen-Bewegung in einer Mulde zwischen zwei entarteten Maxima entspricht. Das Potential ist nach Gleichung (3.3) im Wesentlichen durch die Beweglichkeit (und damit durch den Parameter S) und durch den statischen Strom j^* bestimmt. Bei gegebenem S lassen sich nun Werte für j^* ermitteln, so dass das Potential zwei entartete Maxima enthält. Die Position dieser Maxima ergibt dann ein mögliches Paar von Orientierungen. In Abbildung 5.16 sind diese Orientierungen als Funktion von S dargestellt. Wie zu sehen ist, gibt es für einen vorgegebenen S -Wert im Allgemeinen mehrere Lösungen. Das Potential kann für $n = 6$ höchstens drei Maxima haben, von denen jeweils zwei entartet sein können, so dass sich in bestimmten Bereichen zwei Paare finden lassen¹⁵. Außerdem gibt es, wie in Abschnitt 3 beschrieben, aufgrund der Symmetrie zu jeder Orientierung jeweils eine gespiegelte Orientierung¹⁶. Wir werden im Folgenden zwei zueinander gehörende Orientierungen mit A und B bezeichnen, und die symmetrischen Gegenstücke mit \tilde{A} und \tilde{B} . Die Lösungen zu verschiedenen Werten des Stroms j^* werden wir durch einen Index kennzeichnen. So ergeben sich in Abbildung 5.15 also insgesamt acht Orientierungen. Welche davon

¹⁴Die Einbuchtung ist auch bei den OO-Inseln, wie z.B. der in Abbildung 5.6 näherungsweise statisch. Sie bewegt sich zumindest sehr langsam zum Feld, so dass sich die nachfolgenden Überlegungen mit Einschränkungen (wegen der fehlenden Facettierung der Einbuchtung) auch auf das OO-Regime übertragen lassen.

¹⁵Wir interessieren uns nur für benachbarte entartete Maxima, daher die Zahl zwei.

¹⁶D.h. zu θ^* gibt es entsprechend eine Orientierung $\pi - \theta^*$ (bzw. $-\pi - \theta^*$ falls θ^* negativ ist, da wir die Orientierung durch einen Winkel $-\pi \leq \theta \leq \pi$ darstellen).

nun vom System ausgewählt werden, lässt sich mit der weiter oben angedeuteten geometrischen Bedingung feststellen. Im Wesentlichen besteht die Inselform immer aus einer facettierten Ober- und Unterseite und relativ runder Vorder- und Rückseite. Um aus vorgegebenen Orientierungen eine Ober- und Unterseite zu bilden müssen die Winkel in einer bestimmten Weise zueinander passen. In Abbildung 5.16 sind links Facettenstücke mit den Orientierungen A_1, B_1 (und \tilde{A}_1, \tilde{B}_1) und rechts Stücke mit A_2, B_2 (und \tilde{A}_2, \tilde{B}_2) dargestellt. Wie man leicht sieht, lässt sich der erste Satz nicht zu einer Insel vom geforderten Typ zusammensetzen, womit er für die facettierten Inseln ausscheidet. Die Auswahlregel lässt sich so formulieren, dass für die beiden Orientierungen θ_1 und θ_2 die Beziehung $-\pi/2 < \theta_1 < 0 < \theta_2 < \pi/2$ gelten muss. Wie aus der Abbildung ersichtlich ist, erfüllt das zweite Paar A_2, B_2 diese Bedingung¹⁷, wodurch die Orientierung eindeutig festgelegt ist¹⁸. Es wurde durch Stichproben in den relevanten Bereichen des Phasendiagramms sichergestellt, dass sich auf diese Weise die in der Simulation auftretenden Facetten zuverlässig bestimmen lassen.

Mit diesem Verfahren, die ausgewählten Orientierungen zu bestimmen, lassen sich im Prinzip natürlich auch Vorhersagen für die Fälle machen, in denen die Parameter n und α andere Werte annehmen. Da allerdings die numerischen Simulationen und deren Auswertung mit einem gewissen Aufwand verbunden sind, wurden für diese Fälle nur eher unsystematische Untersuchungen durchgeführt. So ergibt sich zunächst, dass es für $n \leq 3$ keine derartigen Facettenorientierungen gibt¹⁹, was erklären mag, dass in den probeweise durchgeführten Simulationen für $n = 3$ keine Oszillationen beobachtet wurden. Des Weiteren existieren keine geeigneten Orientierungen für $\alpha = \pi/6$, was einem Minimum der Beweglichkeit für Kanten in Feldrichtung entspricht. Auch wurden in der Simulation keine Oszillationen beobachtet²⁰. Für eine größere Zahl an Symmetrieachsen würden sich natürlich auch mehr mögliche Facettenorientierungen ergeben. Was dies für das dynamische Verhalten bedeuten würde, lässt sich schwer vorhersagen. Insgesamt lässt sich noch anmerken, dass es sich bei dem oszillatorischen Verhalten um ein relativ robustes Merkmal zu handeln scheint, das nicht von der präzisen Form der Beweglichkeit abhängt. Wie wir im Ab-

¹⁷Interessanterweise ergibt sich dieses erlaubte Paar erst für $S \geq S_c \simeq 1.77$. Unterhalb dieses Wertes gibt es also streng genommen keine statischen facettierten Einbuchtungen im Sinne der vorangegangenen Überlegungen. Die Oszillationen sind jedoch auch für $S < S_c$ relativ ähnlich zu denen des CO-Regimes (siehe Abbildungen 5.4, 5.7). Die Einbuchtungen sind dort nicht völlig statisch und die Facetten weniger ausgeprägt. Insofern ist die Existenz einer statischen Zickzack-Konfiguration kein „harte“ Bedingung, die dem System keine anderen Möglichkeiten mehr lässt.

¹⁸Natürlich bestehen immer noch zwei symmetrische Möglichkeiten, die einer schräg nach unten oder oben laufenden Insel entsprechen.

¹⁹Die Anzahl der Symmetrieachsen hängt über (3.3) mit der Zahl möglicher Maxima des Potentials zusammen. Weniger Symmetrieachsen führen also zu weniger Maxima.

²⁰Dies scheint ein Sonderfall zu sein. Für $\alpha = \pi/4$ (was keiner irgendwie ausgezeichneten Richtung entspricht, siehe Abbildung 2.1) ergibt sich ein qualitativ ähnliches Phasendiagramm wie im Fall $\alpha = 0$. Zumindest die grobe Topologie, d.h. die Lage der stationären und oszillatorischen Bereiche zueinander bleibt dieselbe.

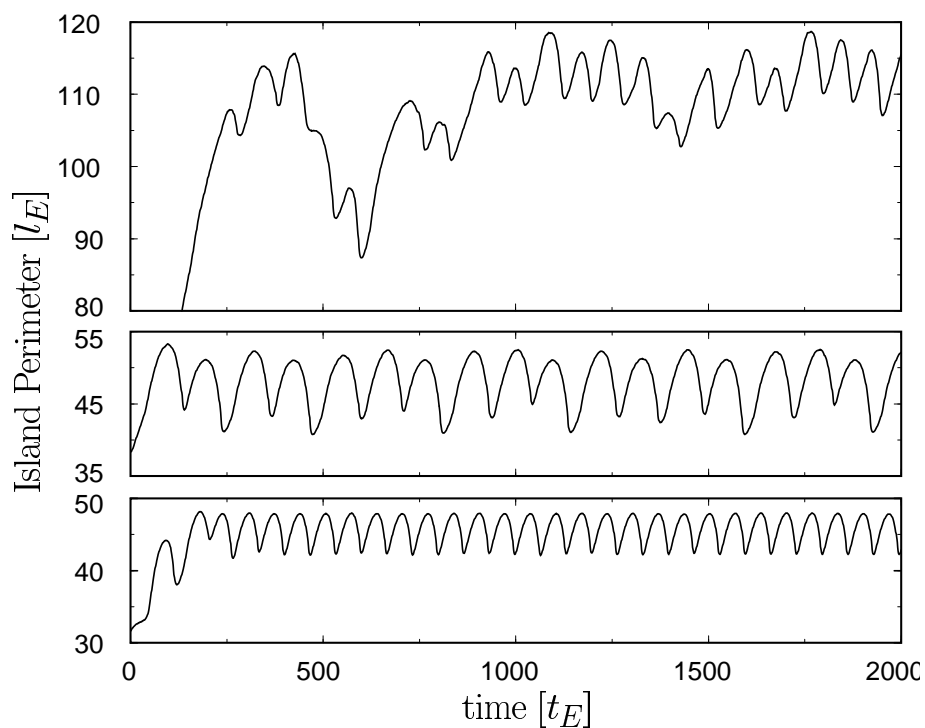


Abbildung 5.11: Inselumfang als Funktion der Zeit. Von unten nach oben sind die Parameter $S = 2, R_0 = 5$, $S = 5, R_0 = 5$ und $S = 5, R_0 = 6.5$. Beim Übergang zum Bereich CO wird der Verlauf zunehmend unregelmäßig. Simulation von Frank Haußer.

schnitt 6 im Vergleich mit dem KMC-Simulationen sehen werden, bleibt das allgemeine Verhalten bei einer anderen Form von $\sigma(\theta)$ und sogar unter anisotroper Steifigkeit $\tilde{\gamma} \neq \text{const.}$ erhalten.

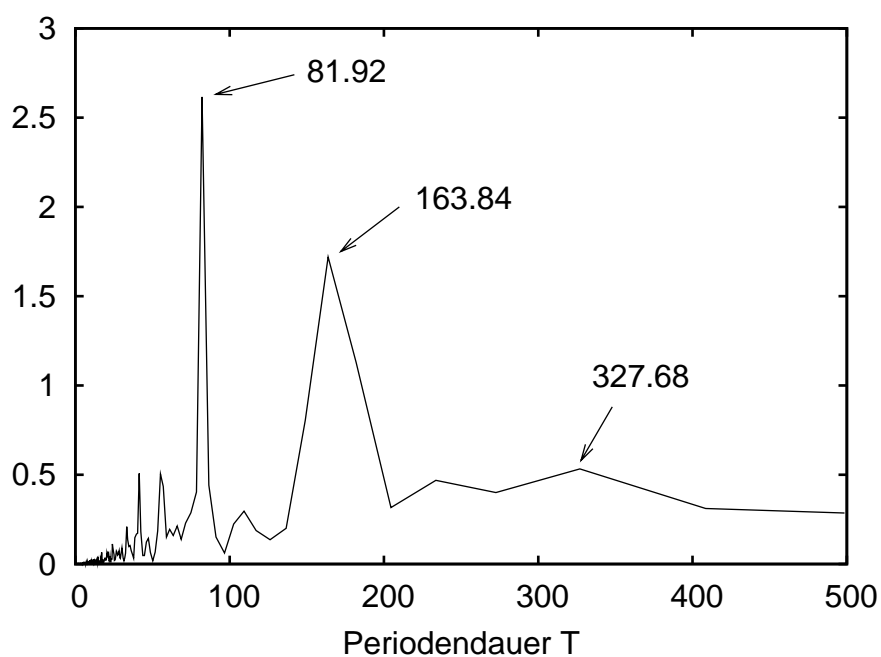


Abbildung 5.12: Fourierspektrum der Zeitreihe des Inselumfangs für $S = 3$ und $R_0 = 6$. Die Werte für die Periodendauer bei den gekennzeichneten Peaks entsprechen exakt einem Verhältnis 1 : 2 : 4. Die zugehörige Simulation wurde von Frank Hauber durchgeführt.

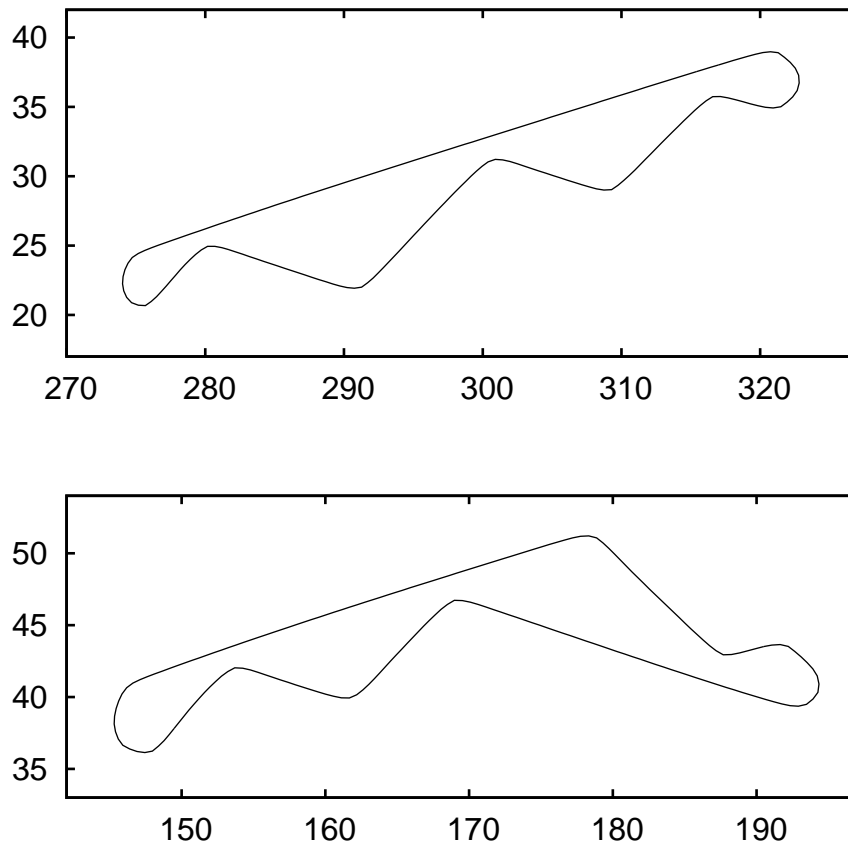


Abbildung 5.13: Schnappschüsse einer Inselbewegung aus dem CO-Regime ($R_0 = 8$, $S = 3$). Dargestellt ist die Insel zum Zeitpunkt $t = 920$ (unten) und $t = 1500$ (oben). Im oberen Bild enthält die Unterseite drei Einbuchtungen, die eine (unregelmäßige) Zickzack-Linie bilden. Im unteren Bild führt die Insel gerade einen Richtungswechsel durch. Ober- und Unterseite zeigen jeweils eine klare Facetten-Struktur. Simulation von Frank Haußer.

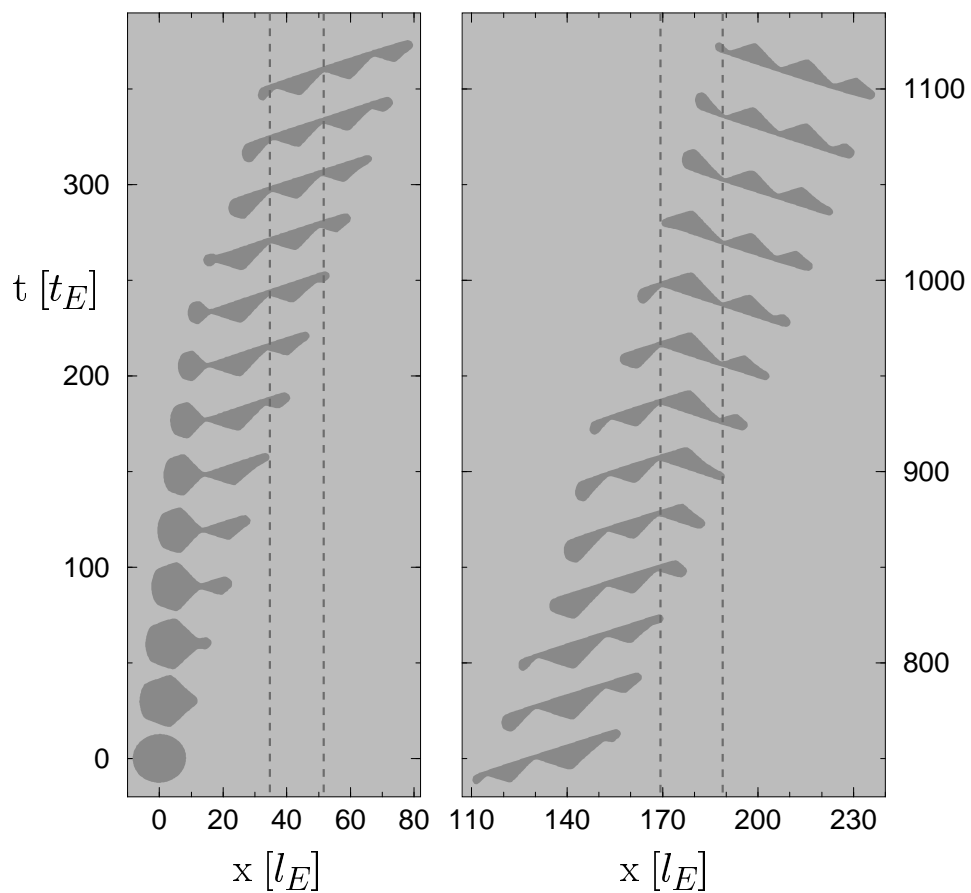


Abbildung 5.14: Inselbewegung aus dem CO-Regime mit $S = 3$ und $R_0 = 8$. Die anfängliche (leicht deformierte) Kreisform entwickelt an beiden Seiten stark facettierte Bereiche. Die gestrichelte Linie zeigt, dass die Einbuchtungen relativ zum Substrat stillstehen. Zeitlich aufeinanderfolgende Schnappschüsse sind vertikal nach oben verschoben. Simulation von Frank Haußer.

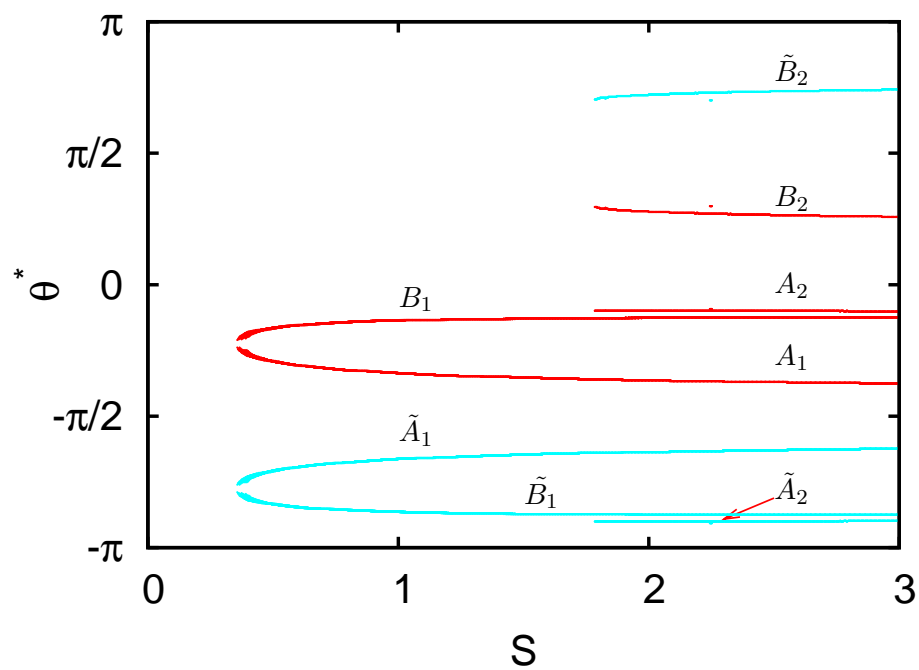


Abbildung 5.15: Die möglichen Paare A und B von Facetten-Orientierungen für die statischen Lösungen. Zu jedem dieser Paare A, B gibt es Gegenstücke \tilde{A} und \tilde{B} , die symmetrisch zur Feldachse gespiegelt liegen. Die Linien von der Orientierungen mit Index 1 weisen bei kleinen S -Werten eine Verdickung auf. Die Ursache ist, dass der Algorithmus, mit dem die Kurven erzeugt wurden, bei den Maxima von $V(\theta)$ nicht exakt zwischen entarteten und beinahe-entarteten Paaren unterscheiden kann.

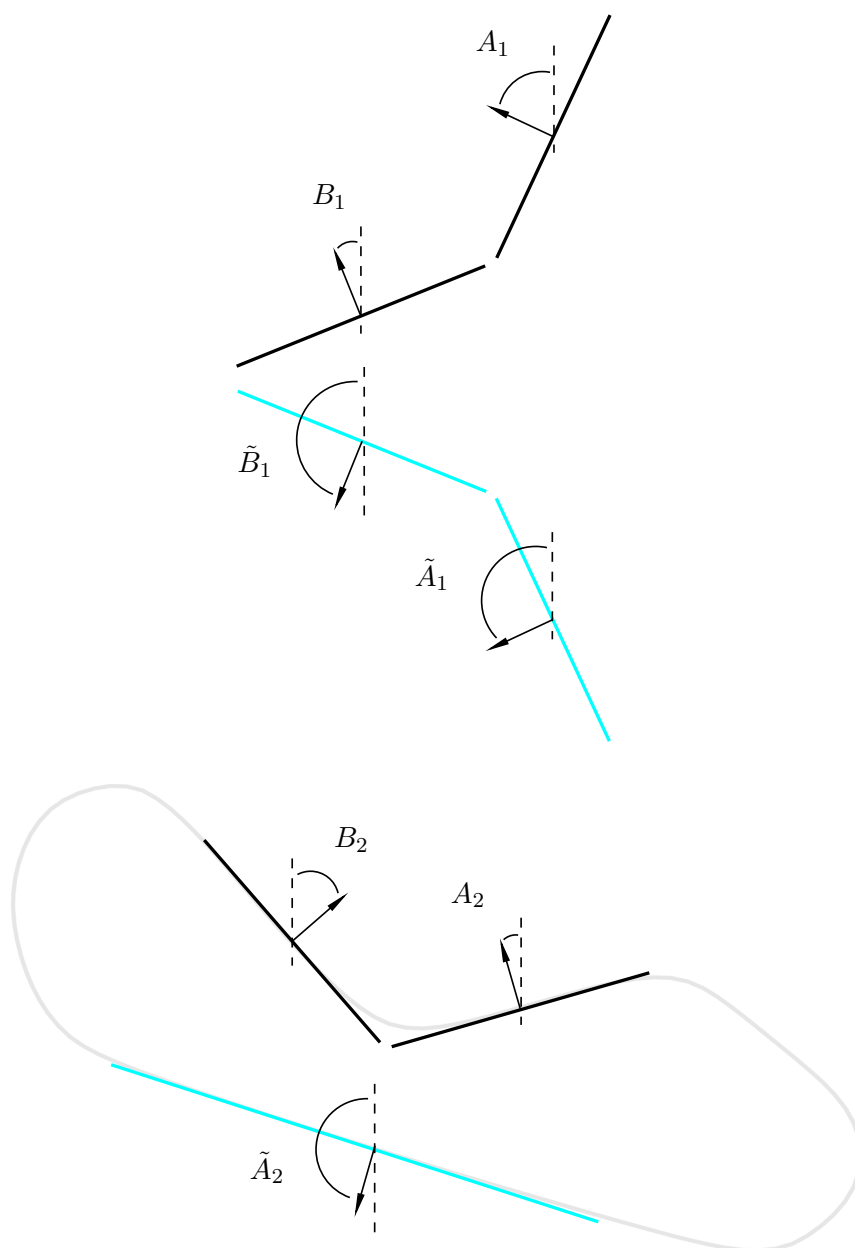


Abbildung 5.16: Geometrische Auswahlregel für die möglichen Facettenorientierungen. Die obere Skizze soll verdeutlichen, dass sich aus den Orientierungen A_1 und B_1 und ihren symmetrischen Gegenstücken \tilde{A}_1 und \tilde{B}_1 keine geschlossene Inselform konstruieren lässt. Im unteren Bild sind die Orientierungen A_2, B_2 und \tilde{A}_2 zu einer schräg nach unten laufenden Form zusammengesetzt. Die Winkel entsprechen den Orientierungen für $S = 2$. Die Inselform im Hintergrund ist das Ergebnis einer Simulation mit $S = 2$ und $R_0 = 4.5$.

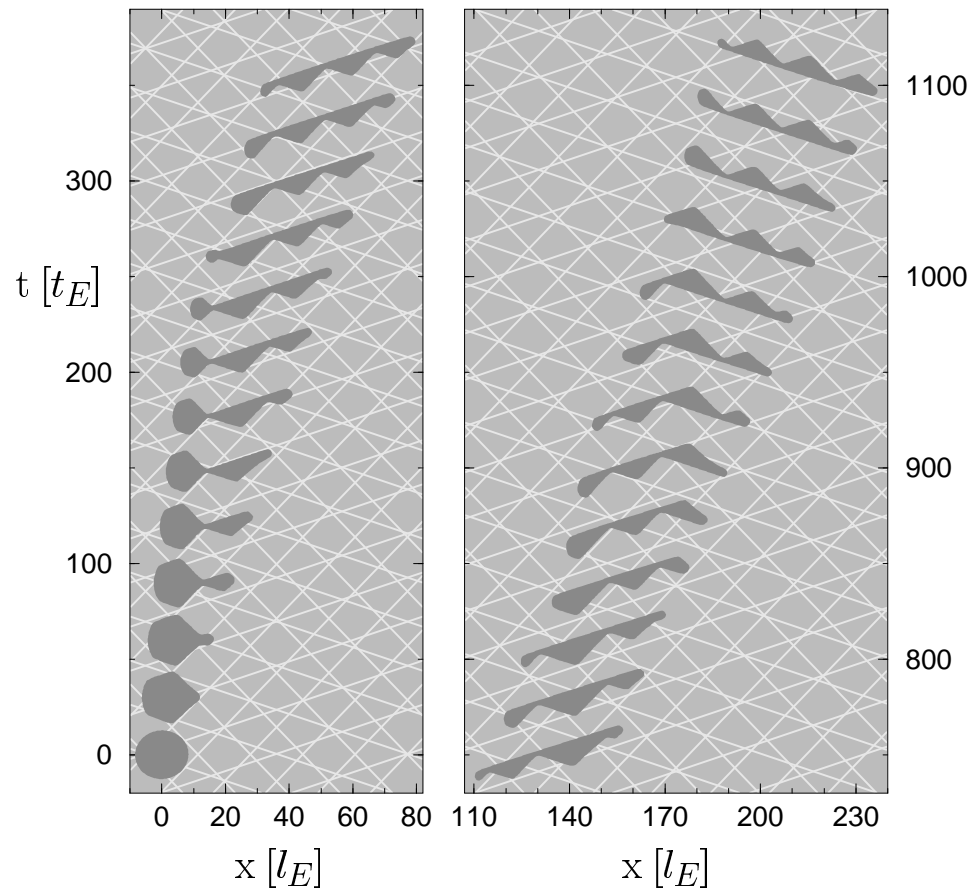


Abbildung 5.17: Die Inselbewegung aus Abbildung 5.14 ($S = 3, R_0 = 8$). Die hellen Linien im Hintergrund entsprechen den vorhergesagten Facettenorientierungen.

Kapitel 6

Vergleich mit KMC-Simulationen

Bei den von uns durchgeführten Untersuchungen behandeln wir die Insel als ein *kontinuierliches* Gebilde, dessen Zeitentwicklung darüber hinaus *deterministischen* Regeln folgt. Reale Kristalloberflächen bestehen dagegen auf mikroskopischer Skala aus diskreten Einheiten, die sich weitgehend stochastisch verhalten. Im Hinblick darauf ist es wichtig, die Frage zu klären, inwieweit die Merkmale der mit den Kontinuumsungleichungen erzeugten Inselbewegung robust gegenüber dieser Diskretheit und Stochastizität sind. Um diese beiden Effekte zu berücksichtigen, hat sich seit langem die Methode der kinetischen Monte-Carlo-Simulation (kurz KMC) etabliert. Wir präsentieren in diesem Abschnitt einige Ergebnisse einer solchen Simulation für die Cu(100)-Oberfläche, die von Marko Rusanen durchgeführt wurde. Der zugrunde liegende Algorithmus ist eine Variante des in [10] vorgeschlagenen Verfahrens¹. Die in der Simulation enthaltenen Raten für die Elementarprozesse² ergeben sich aus Energiebarrieren, die mit Hilfe der „effective-medium“-Theorie [28] berechnet wurden und die in guter Übereinstimmung mit den experimentellen Daten für die Cu(100)-Oberfläche sind [18]. Wir merken hier jedoch an, dass einige dieser Prozesse ausgeschlossen wurden, um die Grundannahmen für das Kontinuumsmodell nicht zu verletzen. So wurde die Ablösung von der Insel und die Leerstellendiffusion im Innern der Insel unterbunden, um dem Grenzfall der reinen Kantendiffusion nahe zu kommen.

Die Elektromigration wurde in das Modell eingefügt, indem die Energiebarriere für einen Sprungprozess gegen die Feldrichtung um einen zusätzlichen Betrag E_{bias} erhöht wurde³.

¹Ein die Rechengeschwindigkeit entscheidend beeinflussender Faktor bei KMC-Simulationen ist die „Verwaltung“ der möglichen Elementarprozesse, aus denen in jedem Schritt ausgewählt wird. In dem von Marko Rusanen verwendeten Algorithmus werden die Prozesse in Form eines binären Baums angeordnet, wodurch sich ein Geschwindigkeitsvorteil ergibt (siehe auch [8]).

²Für die berücksichtigten Prozesse für die Inseldiffusion, siehe [24].

³Für einen Sprung in Feldrichtung ist die Barriere entsprechend um E_{bias} verringert.

Um nun den Einfluss des quadratischen Cu(100)-Gitters in unserem Kontinuumsmodell zu berücksichtigen, wurde aufgrund mikroskopischer Betrachtungen (siehe [49]) folgende Form⁴ für die Steifigkeit und Beweglichkeit angenommen⁵

$$\tilde{\gamma}(\theta) = \frac{k_B T}{a} \frac{(1 + m^2)^{3/2}}{m^2 + \sqrt{m^2 + (a/l_k)^2}}, \quad m = \tan(\theta), \quad (6.1)$$

wobei T die Temperatur, k_B die Boltzmann-Konstante und a die atomare Gitterkonstante bezeichnet. Die Länge $l_k = (a/2)e^{\epsilon/k_B T}$ ist der mittlere Abstand thermisch erzeugter „Kinks“⁶, wobei ϵ die Energie ist, die für die Erzeugung eines Kinks aufgewendet werden muß. Die Beweglichkeit $\sigma(\theta)$ hängt mit der Steifigkeit über

$$\sigma(\theta) = \frac{a\nu_0 e^{-\frac{E_{det}}{k_B T}}}{k_B T} \frac{\tilde{\gamma}(\theta)}{\tilde{\gamma}(\theta) + (k_B T/a)p_{kr}^{-1}} \quad (6.2)$$

zusammen⁷. Darin enthalten sind die Frequenz ν_0 , mit der ein Teilchen Versuche unternimmt, den Gitterplatz zu wechseln, die Barriere E_{det} , die mit dem Ablösen von einem Kink verbunden ist und schließlich die Wahrscheinlichkeit p_{kr} , dass ein Teilchen einen Kink umrundet. Mit dem Umrunden ist die Überwindung einer zusätzlichen Energiebarriere E_{kr} nötig, die in die Wahrscheinlichkeit mit $p_{kr} = e^{-E_{kr}/k_B T}$ eingeht. Die Energiebarrieren ergeben sich durch Vergleich der Bindungsverhältnisse vor und nach dem Sprung. Für eine (horizontale) Bindung setzen wir die Bindungsenergie $E_B = 0.26 \text{ eV}$ an. Für das Ablösen von einer Kink-Position müssen zwei Bindungen aufgebrochen werden, d.h. $E_{det} = 2 \times E_B = 0.52 \text{ eV}$ und für das Umrunden eines Kinks muss eine Bindung gelöst werden, d.h. $E_{kr} = E_B = 0.26 \text{ eV}$. Ein Kink entspricht einer nicht abgesättigten Bindung, wodurch sich die Kinkenergie zu $\epsilon = E_B/2 = 0.13 \text{ eV}$ ergibt.

Neben der dimensionslosen Inselgröße ist also hier die Temperatur T der bestimmende Parameter. Die Temperaturabhängigkeit in (6.1) und (6.2) bestimmt

⁴Die beiden folgenden Ausdrücke gelten nur für $|m| \leq 1$, bzw. $-\pi/4 \leq \theta \leq \pi/4$. Die restlichen Bereiche ergeben sich durch die vierzählige Symmetrie.

⁵Für die Herleitung der folgenden Formeln ist ein detailliertes Verständnis der mikroskopischen Prozesse notwendig. Wir können an dieser Stelle kein allzu genaues Bild vermitteln, für eine ausführlichere Darstellung sei auf [49] verwiesen.

⁶Im Deutschen verwendet man auch die Bezeichnung *Halbkristalllage*. Im Falle des Quadratgitters, das wir hier betrachten, handelt es sich dabei um ein Atom in einer Stufenkante, das nur zwei direkte horizontale Nachbarn besitzt und deshalb relativ schwach gebunden ist. Derartige Konfigurationen sind für die freie Stufenenergie und die Beweglichkeit von entscheidender Bedeutung.

⁷Dies ist eine Verallgemeinerung, eines phänomenologischen Ausdrucks aus [30]. Wie in [36] gezeigt wurde, ist die Anisotropie der Beweglichkeit ausschließlich auf die Kink-Umrundungs-Barriere zurückzuführen.

insbesondere die Stärke der Anisotropie, wobei diese bei tiefen Temperaturen ausgeprägter ist. Abbildung 6.1 zeigt die Winkelabhängigkeit der Steifigkeit (6.1) für die verschiedenen Temperaturwerte. Die Beweglichkeit hat praktisch dieselbe Form, so dass wir auf eine Abbildung verzichten. Die leichten Knicke bei $\theta = \pm\pi/4, \pm3\pi/4$ entstehen dadurch, dass das Minimum des Ausdrucks (6.1) nicht bei $\pi/4$ liegt, sondern leicht dazu verschoben ist⁸.

Die Monte-Carlo-Simulationen wurden hauptsächlich bei Temperaturen von 500 K und 700 K durchgeführt. Entsprechend konzentrierten sich die Simulationen des Kontinuumsmodells ebenfalls auf diese beiden Fälle, wobei in geringerem Umfang auch die Temperaturen $T = 600 K, 900 K$ und $1000 K$ untersucht wurden. Abbildung 6.2 zeigt ein ähnliches Phasendiagramm wie Abbildung 5.5, wobei die Felder der weniger gut aufgelösten Temperaturwerte eher Stichprobencharakter haben⁹. Die beste Übereinstimmung mit den KMC-Daten ergaben sich für $T = 700 K$. Die Abfolge der verschiedenen Bewegungs-Regimes ist dabei ähnlich wie im sechszähligen Fall mit isotroper Steifigkeit, der in Abschnitt 5 beschrieben wurde. Der Bereich mit schräglaufenden und oszillierenden Bewegungsformen ist dabei jedoch wesentlich schmaler. Bis $R_0 = 2.16$ laufen die Inseln stationär geradeaus, für $R_0 = 2.2$ wurde eine stationär schräglaufende Bewegung identifiziert und zwischen $R_0 = 2.24$ und 2.36 ergaben sich schräglaufend oszillierende Bewegungen. Ab $R_0 = 2.4$ führte die Bewegung zum Zerschneiden der Insel¹⁰. In der KMC-Simulation ergab sich ein qualitativ ähnliches Bild, bei dem sich mit steigender Inselgröße stationäres und oszillatorisches Verhalten zeigte, gefolgt von einem Bereich, in dem die Inseln zerbrechen. Ob die Bewegungsrichtung bei der stationären und oszillatorischen Bewegung einen Winkel zum Feld aufweist, ließ sich nicht zweifelsfrei feststellen, im letzteren Fall wechselt die Insel von einer Oszillation zur nächsten scheinbar zufällig ihre Richtung, so dass die Schwerpunktsbewegung einer Art Irrpfad folgt. Die Oszillationen selbst sehen denen des Kontinuumsmodells erstaunlich ähnlich. In Abbildung 6.3 ist eine typische Sequenz aus der KMC-Simulation (links) einer Abfolge aus dem Kontinuumsmodell (rechts) gegenübergestellt.

Das Bewegungsmuster ist im Wesentlichen dasselbe wie im sechszähligen Fall,

⁸Der $(a/l_k)^2$ -Term regularisiert gewissermaßen den (rein entropischen) $T = 0$ -Ausdruck

$$\tilde{\gamma}(\theta) = \frac{k_B T (1 + m^2)^{3/2}}{a (m^2 + |m|)},$$

der bei $m = 0$ divergiert. Eine andere analytische Form für die Steifigkeit wurde in [56] vorgeschlagen.

⁹In einem ersten Schritt musste zuerst die Lage der „interessanten“ Phasen OS und OO lokalisiert werden, die hier einen sehr schmalen Streifen zwischen SS und BU bilden. Lediglich für die beiden Temperaturen $T = 500 K$ und $700 K$ wurde in einem zweiten Schritt ein feineres Raster verwendet.

¹⁰Die Simulationsdaten in diesem Abschnitt wurden ausschließlich mit dem Finite-Differenzen-Verfahren gewonnen. Vom sechszähligen Fall her wissen wir, dass mit diesem Algorithmus die Inseln leichter zerbrechen. Insofern lässt sich nicht zweifelsfrei sagen, ob über $R_0 = 2.36$ hinaus stabile Bewegungen möglich sind.

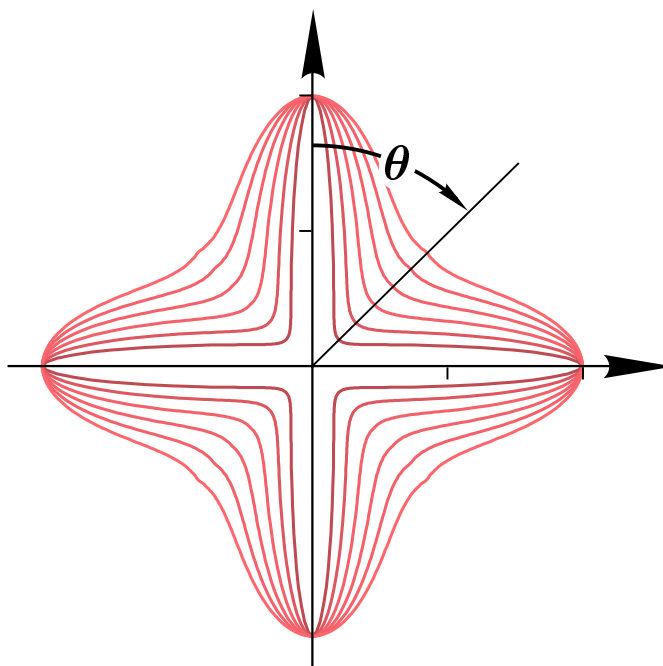


Abbildung 6.1: Orientierungsabhängigkeit der Steifigkeit für den Vergleich mit den KMC-Simulationen (Formel (6.1)). Die Temperaturen sind $T = 500, 600, \dots, 1000 \text{ K}$. Die Anisotropie ist für niedrige Temperaturen ausgeprägter. Das Maximum wurde jeweils auf Eins normiert.

wobei sich die Formen wegen der vierzähligen Symmetrie etwas unterscheiden. Es entsteht wieder eine statische Einbuchtung, durch die die Insel vertikal in die Länge gezogen wird, bevor die Einbuchtung das hintere Ende erreicht und die Form sich wieder zusammenzieht. Rechnet man die Inselgrößen des KMC-Modells (d.h. in Einheiten der Gitterkonstanten a) auf die Skala l_E des Kontinuumsfalls um, so ergibt sich für den in 6.3 links dargestellten Fall ein Radius¹¹ von $R_0 = 1.56$, der etwas niedriger ist als der Wert $R_0 = 2.35$ der rechts abgebildeten Kontinuums-Oszillation. Trägt man den Umfang der Insel als Funktion der Zeit auf (Abbildung 6.4), so zeigt sich auch hier eine sehr ähnliche Form der Oszillation, wobei die KMC-Daten natürlich eine starke Unregelmäßigkeit aufweisen. Rechnet man die Zeit auf die Skala t_E des Kontinuumsmodells um, so zeigt sich, dass sich der im unteren Bild dargestellte Zeitabschnitt um einen Faktor von ungefähr 2.6 von dem des unteren Bildes unterscheidet. Da zwischen räumlicher und zeitlicher Skala der Zusammenhang $t_E \sim l_E^4$ (siehe Gleichungen 2.6) besteht, entspricht das in etwa der oben erwähnten Diskrepanz bei den Inselgrößen.

¹¹Da die Anzahl N der in der Insel vorhandenen Atome der Fläche entspricht, wird der Radius mit $R_0 = \sqrt{N/\pi}(a/l_E)$ berechnet.

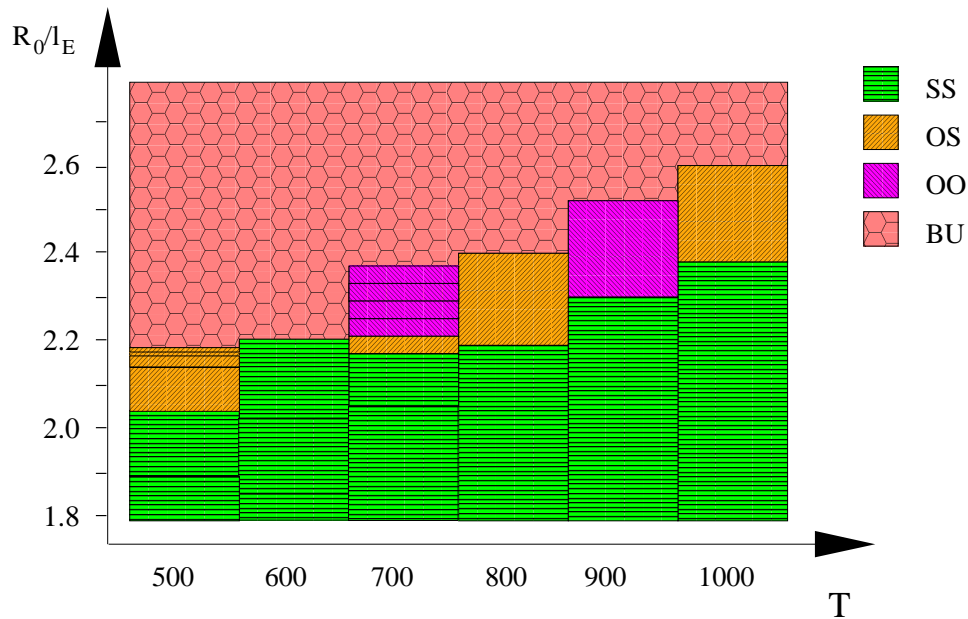


Abbildung 6.2: Phasendiagramm für das Kontinuumsmodell mit Beweglichkeit (6.2) und Steifigkeit (6.1). Näher aufgelöst sind nur die Phasen für die Temperaturen $T = 500$ K und 700 K. Die Rechtecke kennzeichnen jeweils einen Wert für R_0 , für den eine Simulation durchgeführt wurde. Die vertikalen Ränder dieser Rechtecke liegen immer genau *zwischen* zweien solcher Werte. Im Fall $T = 500$ K wurden auch nach einer genaueren Suche keine Oszillationen gefunden, während sich bei $T = 700$ K ein kleiner oszillatorischer Bereich ergibt. Wegen der sehr groben Auflösung sind bei den restlichen Temperaturen entweder nur SS (bei $T = 600$ K), oder entweder SS und OS ($T = 800$ K und 1000 K) oder SS und OO-Phasen ($T = 900$ K) zu erkennen.

Wir hatten in Abschnitt 5 angemerkt, dass im sechszähligen Fall keine Oszillationen auftreten, wenn die Kanten in Feldrichtung einem Minimum der Beweglichkeit entsprechen. Dies scheint im vierzähligen Fall ebenso zu sein, und zwar für KMC- und Kontinuumsmodell gleichermaßen¹². Es ergeben sich offenbar ausnahmelos stationär in Feldrichtung driftende Inseln, deren Formen in den beiden Modellen, wie Abbildung 6.5 zeigt, erstaunlich gut übereinstimmen. Rechnet man den Inselradius auf die Skala l_E um, so ergibt sich für das linke Bild $R_0 = 1.106$ und für das rechte $R_0 = 1.56$, was wiederum etwas kleiner ist als die Radien der dargestellten Kontinuumsformen.

Bei der niedrigeren Temperatur $T = 500$ K ergeben sich dagegen Abweichungen zwischen den beiden Modellen. Wie in Abbildung 6.3 im mittleren Bild zu sehen ist, zeigen die KMC-Simulationen auch hier Oszillationen, während im Konti-

¹²Dazu wurde das Feld um $\pi/4$ gedreht.

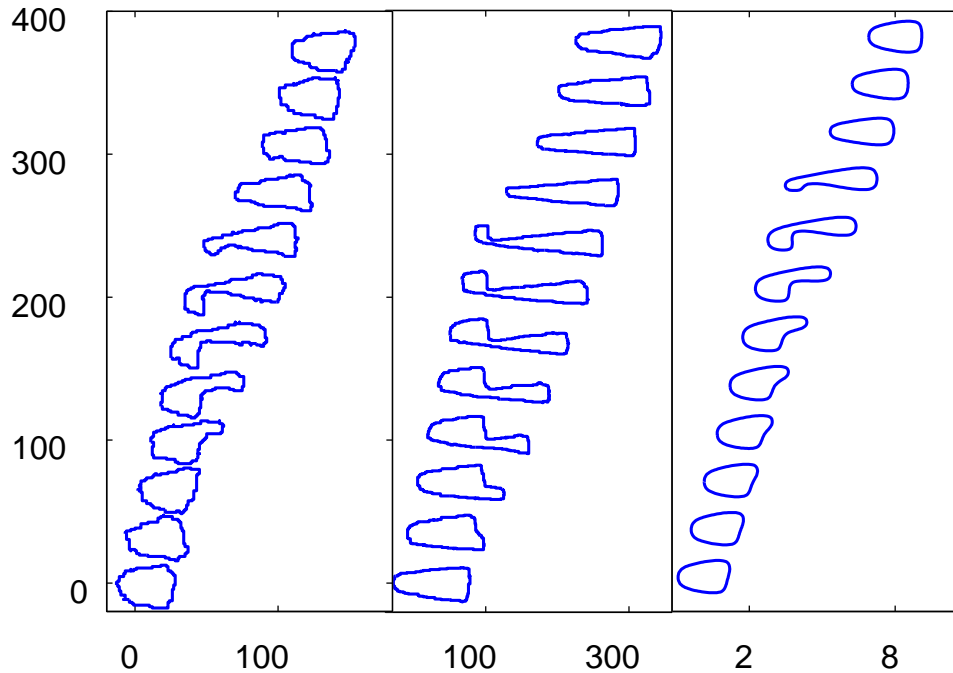


Abbildung 6.3: Zeitliche Abfolge der oszillatorischen Formänderung in den KMC-Simulationen (linkes und mittleres Bild) und im Kontinuumsmodell (rechts). Die Konfigurationen wurden jeweils vertikal zueinander versetzt. Die Parameter sind $T = 700 \text{ K}$, $E_{bias} = 2 \times 10^{-3} \text{ eV}$, bei einer Inselgröße von $N = 1000$ Atomen (links), $T = 500 \text{ K}$, $E_{bias} = 2 \times 10^{-3} \text{ eV}$, $N = 4000$ (Mitte) und $T = 700 \text{ K}$, $R_0 = 2.35$ (rechts).

nummsfall nur stationäre Formen beobachtet wurden. Die Inselform ist in allen Phasen der Oszillationsbewegung gegenüber den Formen bei $T = 700 \text{ K}$ langgestreckt. Diese Dehnung in Feldrichtung ist auch bei den kleineren stationären Inseln vorhanden. Im Kontinuumsfall ist dieses Phänomen nicht zu finden, hier weicht das Aspektverhältnis nicht wesentlich von Eins ab.

Im Kontinuumsmodell geht die Größe der Inseln nur über den dimensionslosen Radius R_0 ein, so dass verschiedene Inselgrößen R_0 und Feldstärken F zu demselben dynamischen Verhalten führen, sofern das Produkt $F \times R_0^2$ konstant bleibt. Die Fläche πR_0^2 der Insel ist proportional zur Anzahl N der darin enthaltenen Atome, so dass sich dies in $N \times E_{bias} = \text{const.}$ übersetzt. Die Bewegung sollte also von E_{bias} auf dieselbe Weise abhängen wie von N . Während sich bei $T = 700 \text{ K}$ tatsächlich für $N \times E_{bias} = \text{const.}$ qualitativ dieselbe Bewegung findet, ist diese Eigenschaft bei $T = 500 \text{ K}$ nicht mehr gegeben. Das Aspektverhältnis sättigt in Abhängigkeit von N bis zum Einsetzen der Oszillationen, während es in Abhängigkeit von E_{bias} monoton anwächst. Bei einer Inselgröße

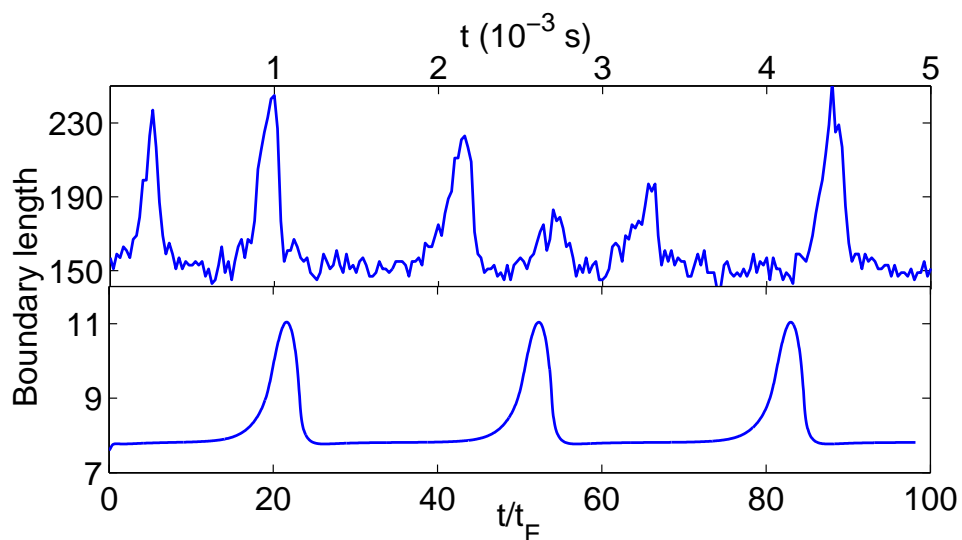


Abbildung 6.4: Oszillationen des Umfangs in der KMC-Simulation (obere Abbildung) und im Kontinuumsmodell (unten). Die Parameter sind die der rechten und linken Bewegung in Abbildung 6.3.

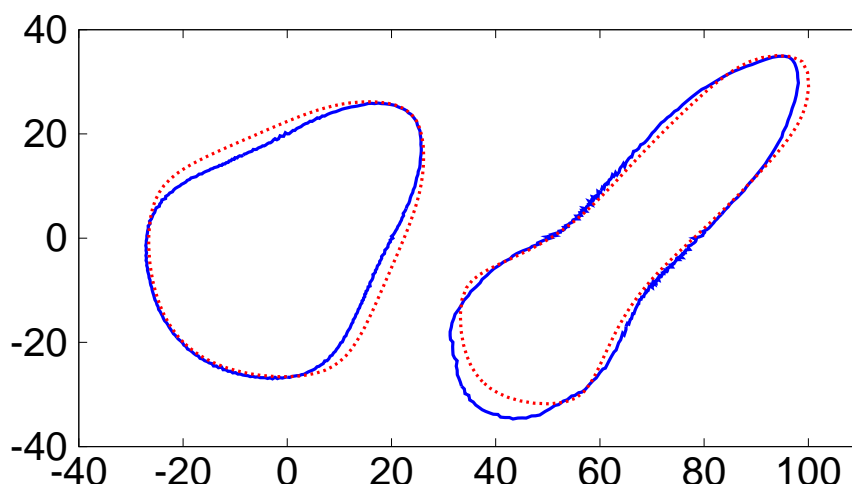


Abbildung 6.5: Stationäre Formen für $T = 700 \text{ K}$ mit Feldrichtung entlang der Diagonalen des Gitters (d.h. der Winkel zwischen Feld und Horizontalen beträgt $\pi/4$). Die durchgezogenen dunklen Linien kennzeichnen die (zeitlich gemittelten) Inselformen im KMC-Modell mit den Parametern $N = 2000$ und $E_{bias} = 5 \times 10^{-4} \text{ eV}$ (links) bzw. $E_{bias} = 10^{-3} \text{ eV}$ (rechts). Die gepunkteten, helleren Linien sind die Ergebnisse des Kontinuumsmodells mit $R_0 = 1.9$ (links) und 2.9 (rechts).

von $N = 1000$ wurden selbst für starke Felder $\sim 10^{-2} eV$ keine Oszillationen beobachtet. Diese Verletzung des Skalenverhaltens zeigt ebenfalls an, dass die Kontinuumstheorie nicht mehr gültig ist.

Die Ursache für diese Abweichungen bei tieferen Temperaturen sind noch unklar. Insgesamt zeigt sich aber, dass die wesentlichen Merkmale des Kontinuumsmodells auch in den KMC-Simulationen vorhanden sind. Insbesondere die Oszillationen stellen sich in dieser Hinsicht als robustes Phänomen heraus. Zudem zeigt sich hier für das Kontinuumsmodell, dass auch im vierzähligen Fall mit Anisotropie der Beweglichkeit *und* der Steifigkeit das Verhalten im Wesentlichen dasselbe ist wie im sechszähligen Fall mit isotroper Steifigkeit.

Abschließend erwähnen wir noch, dass ein ähnliches KMC-Modell von Mehl *et al.* für die Elektromigration von Inseln eingesetzt [37] wurde. Dort wurden jedoch etwas niedrigere Temperaturen und kleinere Inseln untersucht. Die hier beschriebenen Phänomene wurden dort nicht beobachtet.

Teil II

Das nichtlokale Modell

Kapitel 7

Die dynamischen Gleichungen

Wir kommen nun im zweiten Teil unserer Arbeit auf das kinetische Regime zu sprechen, in dem die Adatom-Diffusion auf der Terrasse der dominante Transport-Mechanismus ist. Der entsprechende mathematische Rahmen wurde auch hier in der Arbeit von O. Pierre-Louis und T.L. Einstein [44] aufgestellt. Wir werden zunächst den Satz der dynamischen Gleichungen in voller Allgemeinheit herleiten, uns aber im Rahmen dieser Arbeit auf einen Spezialfall, das sogenannte *innere Modell* konzentrieren.

Da wir von nun an Massenaustausch zwischen Inselkante und Terrasse zulassen, enthält das Modell zwei Größen, die das dynamische Verhalten bestimmen. Zum einen, wie im lokalen Fall, die Inselkante selbst, zum anderen die Konzentration von Adatomen auf der Terrasse. Letztere stellen wir durch eine kontinuierliche Funktion $c(x, y, t)$ von zwei räumlichen Variablen dar, deren Bewegungsgleichung aus der Kontinuitätsgleichung für einen Massenstrom folgt

$$\partial_t c + \nabla \mathbf{j} = 0 \quad (7.1)$$

$$\mathbf{j} = -D \nabla c + \frac{D}{\xi} c \mathbf{e}_x. \quad (7.2)$$

Der Strom beinhaltet also zwei Beiträge. Der erste der beiden Terme auf der rechten Seite von (7.2) entspricht dem diffusiven Teil, d.h. dem Massentransport infolge der thermischen Bewegung der Adatome, der zum Gradienten der Konzentration proportional ist. Der zweite Term entspricht einer konstanten Drift, bei der die Adatome durch die Elektromigrationskraft in Feldrichtung verschoben werden¹. Die Diffusionskonstante D , die beide Terme multipliziert, übernimmt hier die Rolle, die die Beweglichkeit σ im lokalen Modell gespielt hat². Die Konstante ξ hat die Dimension einer Länge und hängt mit der Stärke F des elektrischen Feldes über

¹Wir nehmen also wieder ein konstantes Feld in x -Richtung an.

²Im Gegensatz zum lokalen Modell werden wir uns hier allerdings auf isotrope Diffusion (und Steifigkeit) beschränken.

$$\xi = \frac{k_B T}{F}$$

zusammen³. Dabei bezeichnet T die Temperatur und k_B die Boltzmann-Konstante. Wir behandeln nun gleichzeitig den Fall einer Insel und den Fall einer Leerstelleninsel. Dazu müssen wir zwei voneinander getrennte Konzentrationen berücksichtigen, die Konzentration auf der Insel, bzw. im Innern der Leerstelleninsel, und die Konzentration im Außenraum. Die Kopplung zwischen Konzentration und Inselrand ist in Form einer Randbedingung gegeben. Wir versehen von nun an die Größen auf der unteren Terrasse mit dem Index $+$ und die auf der oberen Terrasse entsprechend mit $-$ ⁴. Des Weiteren vereinbaren wir, dass der Normalenvektor \mathbf{n} in Richtung der unteren Terrasse zeigen soll⁵. Die beiden Ströme⁶ j_{\pm} zur Stufe hin sind damit

$$j_{\pm} = \mp (\mathbf{j}_{\pm} \Big|_{\partial\Omega} \cdot \mathbf{n} - c_{\pm} \Big|_{\partial\Omega} v_n),$$

mit der Normalengeschwindigkeit v_n . Den Bereich in der Ebene, der von der Insel eingenommen wird, bezeichnen wir von nun an mit Ω und den Inselrand mit $\partial\Omega$. Der erste Term ist der durch Diffusion und Feld erzeugte Strom, während der zweite Term die Tatsache berücksichtigt, dass eine Stufe, die über das Substrat wandert, während dieser Bewegung Adatome „aufammelt“. Die Randbedingung besteht nun in der Forderung, dass diese Ströme proportional zur Abweichung der Konzentration an der Stufe von der Gleichgewichtskonzentration sind [31]

$$j_{\pm} = k_{\pm} (c_{\pm} \Big|_{\partial\Omega} - c_{eq}). \quad (7.3)$$

Dabei bezeichnen k_+ und k_- die Anlagerungsraten an die Kante von der unteren, bzw. oberen Terrasse aus. Aus diesen Raten und der Diffusionskonstanten D lassen sich die beiden kinetischen Längen

$$d_{\pm} = D/k_{\pm}$$

³D.h. ξ ist die charakteristische Längenskala, die sich aus der Feldstärke und der thermischen Energie $k_B T$ bilden läßt. Typische experimentelle Werte sind $\xi = 10^8 \times a$ (wobei a die atomare Gitterkonstante bezeichnet) für Silizium und etwa $\xi = 10^5 \times a$ für Metalle [44].

⁴Der Index $+$ bezeichnet also im Falle einer Insel das Äußere und für eine Leerstelleninsel das Innere.

⁵Für eine Insel stimmt diese Konvention also mit der des lokalen Modells überein, wo der Normalenvektor *nach außen* zeigt.

⁶Dabei handelt es sich um skalare Größen. Die Richtung der zugehörigen Vektoren ist die des Normalenvektors.

bilden, deren Größe die Geschwindigkeit widerspiegelt, mit denen die Anlageprozesse im Vergleich zu der Bewegung der Atome auf der Terrasse ablaufen.

Die Gleichgewichtskonzentration ist durch die Gibbs-Thomson Beziehung

$$c_{eq} = c_{eq}^0 \exp(\Gamma\kappa) \quad (7.4)$$

gegeben. Die darin enthaltene Kapillaritätslänge Γ hängt mit der Steifigkeit $\tilde{\gamma}$ über

$$\Gamma = \frac{a^2 \tilde{\gamma}}{k_B T}$$

zusammen. Mit κ bezeichnen wir wieder die Krümmung des Inselrandes⁷. Wir werden im Folgenden auch die linearisierte Form

$$c_{eq} = c_{eq}^0 (1 + \Gamma\kappa) \quad (7.5)$$

benutzen. Voraussetzung für ihre Gültigkeit ist, dass die Kapillaritätslänge klein gegenüber dem Krümmungsradius $1/\kappa$ ist. Schließlich folgt die Bewegungsgleichung für den Inselrand wie im lokalen Fall aus der Massenerhaltung

$$v_n = a^2 (j_+ + j_-). \quad (7.6)$$

Wir werden nun diese allgemeinen Gleichungen schrittweise einschränken, bis wir bei dem System angelangt sind, das den Gegenstand unserer Untersuchung bilden wird. Zunächst nehmen wir an, dass es sich bei der Insel um eine Leerstellen-Insel handelt, deren Stufenkante eine unendlich hohe Stufenrandbarriere (Ehrlich-Schwoebel-Barriere) aufweist⁸, d.h. die Anlagerung von der oberen Terrasse ist damit unterbunden ($k_- = 0, d_- = \infty$) und das Konzentrationsfeld im Außenraum hat keinerlei Einfluß mehr auf die Inselbewegung. Alle Gleichungen beziehen sich daher im Folgenden ausschließlich auf die Konzentration im Inneren, so dass wir den Index wieder weglassen können. Die Randbedingung ist also

$$-(\mathbf{j}|_{\partial\Omega} \cdot \mathbf{n} - c|_{\partial\Omega} v_n) = k(c|_{\partial\Omega} - c_{eq}) \quad (7.7)$$

⁷Im Falle einer konvexen Insel legen wir die Krümmung wie im lokalen Fall als positiv fest. Für eine konvexe Leerstelleninsel ergibt sich damit eine *negative* Krümmung.

⁸Diese Barriere kommt dadurch zustande, dass ein Atom beim Sprung von der oberen zur unteren Terrasse zwischenzeitlich eine nur schwach gebundene Position einnehmen muss. Für eine ausführlichere Darstellung dieses Effekts und seine Bedeutung für die Strukturbildung auf Oberflächen, siehe [39] S. 138 ff..

und die Normalengeschwindigkeit des Inselrandes ist

$$v_n = -a^2 \left(\mathbf{j} \Big|_{\partial\Omega} \cdot \mathbf{n} - c \Big|_{\partial\Omega} v_n \right). \quad (7.8)$$

Die Gleichungen (7.1), (7.2), (7.4), (7.7) und (7.8) definieren das sogenannte „innere Modell“ [44], d.h. die Dynamik einer Leerstelleninsel, die nur mit der Adatomkonzentration im Innern wechselwirkt⁹. Für den Grenzfall der reinen Terrassen-Diffusion, wo die Diffusion auf der Terrasse der langsamste und damit bestimmende Prozess ist, wird die Anlagerung von unten instantan ($k_+ \rightarrow \infty, d_+ \rightarrow 0$) und Gleichung (7.7) reduziert sich auf die sogenannte *thermodynamische Randbedingung*

$$c \Big|_{\partial\Omega} = c_{eq} = c_{eq}^0 \exp(\Gamma\kappa) \simeq c_{eq}^0 (1 + \Gamma\kappa). \quad (7.9)$$

Setzt man eine kreisrunde Inselform $\kappa \equiv \text{const} = -1/R_0$ an¹⁰, so ergibt sich aus (7.4) für die Konzentration der konstante Wert

$$c_{eq} = c_{eq}^0 \exp\left(-\frac{\Gamma}{R_0}\right) \quad (7.10)$$

am Inselrand. Setzt man die Konzentration im Inneren ebenfalls konstant $c \equiv c_{eq}$, so ergibt sich aus (7.2)

$$\mathbf{j} = c_{eq} \frac{D}{\xi} \mathbf{e}_x,$$

was offensichtlich (7.1) erfüllt. Bei einer konstanten Konzentration spielt also die Diffusion keine Rolle und das Feld erzeugt einen Strom konstanter Stärke und Richtung. Mit dem Skalarprodukt¹¹ $\mathbf{e}_x \cdot \mathbf{n} = -\cos\theta$ ergibt sich aus (7.8) die Normalengeschwindigkeit

$$\begin{aligned} v_n &= -a^2 \left(-c_{eq} \frac{D}{\xi} \cos\theta - c_{eq} v_n \right) \\ \Rightarrow v_n &= \frac{D}{\xi} \frac{a^2 c_{eq}}{1 - a^2 c_{eq}} \cos\theta \end{aligned}$$

⁹Äquivalent dazu ist der quasi „inverse“ Fall einer Insel mit Leerstellen-Diffusion im Inneren.

¹⁰Das negative Vorzeichen kommt daher, dass es sich um eine Leerstelleninsel handelt.

¹¹Der Winkel θ ist hier der Winkel zur Feldrichtung. Die Festlegung ist also eine etwas andere als im lokalen Modell.

Das ist wieder die Gleichung einer stationären Bewegung in horizontaler Richtung¹², die Driftgeschwindigkeit lässt sich sofort ablesen und hat den Wert

$$\mathbf{V} = -\frac{D}{\xi} \frac{a^2 c_{eq}}{1 - a^2 c_{eq}} \mathbf{e}_x \quad (7.11)$$

Wir erhalten also wieder eine kreisförmige Insel, die stationär in *negative x*-Richtung driftet¹³

In [44] wurde für diese Lösung eine lineare Stabilitätsanalyse durchgeführt, deren Ergebnis war, dass die Insel für $R_0/\xi > 0.1$ instabil sein sollte. Um diese Instabilität genauer zu untersuchen, wurde eine Monte-Carlo-Simulation durchgeführt, bei der für $R_0/\xi = 10/6 > 0.1$ eine leichte Destabilisierung der Rückseite beobachtet wurde, während die Insel als Ganzes stabil blieb. Die Vermutung der Autoren war, dass es sich dabei um die berechnete lineare Instabilität handelt. Wir werden nun im nächsten Abschnitt ebenfalls eine Stabilitätsanalyse durchführen und dabei zu dem Schluss kommen, dass das System, im Gegensatz zu den Ergebnissen in [44] für beliebige Werte von R_0/ξ stabil ist. Anschließend diskutieren wir die Ergebnisse einer numerischen Simulation, bei der sich eine nichtlineare Instabilität andeutet, die allerdings erst bei einem wesentlich höheren Wert von R_0/ξ einsetzt [23].

¹²Die Stationaritätsbedingung hat eine ähnliche Form wie Gleichung (4.4) im lokalen Fall. Wegen der etwas anderen Definition des Winkels und der nach innen gerichteten Normalen ist sie hier allerdings $v_n = -V \cos \theta$.

¹³Das Feld treibt die Adatome nach rechts, wodurch sich die Leerstelleninsel nach links bewegt.

Kapitel 8

Lineare Stabilitätsanalyse

8.1 Die Eigenwertgleichung

Die Rechnung in diesem Abschnitt stellt gewissermaßen den Kern unserer Überlegungen zum inneren Modell dar und bildet die Grundlage für alle späteren Abschnitte. Für alle der folgenden analytischen Betrachtungen werden wir Gebrauch von der sogenannten *quasistatischen Näherung* machen. Diese ist anwendbar, wenn die Bewegung der Stufenkante langsam gegenüber der Bewegung der Adatome auf der Terrasse ist. In diesem Fall passt sich die Adatomkonzentration quasi instantan an die Form der Insel an. Wir können deshalb in der vorgegebenen Inselgeometrie das *stationäre* Konzentrationsprofil durch

$$\begin{aligned}\partial_t c &\equiv 0 \\ \Rightarrow \nabla \mathbf{j} &\equiv 0\end{aligned}\tag{8.1}$$

bestimmen und darüber hinaus den zweiten Term in Gleichung (7.8) vernachlässigen¹ und erhalten damit

$$v_n = -a^2 \mathbf{n} \cdot \mathbf{j}|_{\partial\Omega}.\tag{8.2}$$

In dieser Näherung berechnet sich die Driftgeschwindigkeit der kreisförmigen Insel zu²

$$\mathbf{V} = -\frac{1}{\xi} D c_{eq} \mathbf{e}_x.\tag{8.3}$$

Wir merken noch an, dass in der quasistatischen Näherung die Fläche A der Leerstelleninsel erhalten bleibt. Die Zeitableitung von A berechnet sich nach

¹Die quasistatische Inselkante „sammelt“ keine Adatome auf.

²Der Faktor $(1 - a^2 c_{eq})^{-1}$ stellt also den Korrekturterm gegenüber dieser Näherung dar. Die quasistatische Näherung erfordert $a^2 c_{eq} \ll 1$, wodurch der Faktor Eins wird.

$$\frac{d}{dt}A = \int_{\partial\Omega} v_n ds \stackrel{(8.2)}{=} -a^2 \int_{\partial\Omega} \mathbf{j} \cdot \mathbf{n} ds = -a^2 \int_{\Omega} \nabla \cdot \mathbf{j} dA \stackrel{(8.1)}{=} 0,$$

wobei wir im vorletzten Schritt den (ebenen) Gaußschen Integralsatz verwendet haben. Die Flächenerhaltung bedeutet, dass eine Art Fließgleichgewicht im Materialaustausch zwischen Konzentration und Inselrand besteht, bei dem insgesamt genauso viel Masse in den Inselrand hineinströmt wie umgekehrt abgegeben wird. In diesem Sinne vermittelt das Konzentrationsfeld den Massentransport von einem Teil des Inselrandes zu einem anderen, ohne dabei im Ganzen Masse aufzunehmen oder zu verlieren.

Um nun die lineare Stabilität dieser Lösung zu untersuchen, setzen wir eine infinitesimale Störung der Inselform und der Konzentration an

$$\begin{aligned} c(r, \theta, t) &= \tilde{c}_{eq} + \varepsilon c_1(r, \theta, t) \\ R(\theta, t) &= R_0 + \varepsilon \rho(\theta, t). \end{aligned}$$

Wir werden diese gestörten Lösungen in die Gleichungen einsetzen und nur Terme berücksichtigen, die höchstens linear in ε sind. Damit erhalten wir lineare Gleichungen, denen wir das zeitliche Verhalten der Störung entnehmen können. Zuerst wollen wir uns jedoch einen Überblick verschaffen, indem wir das Schema der folgenden Rechnung entwerfen. Wir suchen eine gemeinsame Lösung der Gleichungen (8.1), (7.9) und (8.2). Während (7.9) und (8.2) sowohl den Inselrand $R(\theta, t)$ also auch die Konzentration $c(r, \theta, t)$ enthalten, ist (8.1) eine Gleichung für die Konzentration allein. Es ist also sinnvoll, zuerst nach der allgemeinen Lösung dieser Gleichung zu fragen und den Ansatz für c_1 entsprechend zu wählen. Es bleiben dann die Gleichungen (7.9) und (8.2) für die beiden Größen R und c , aus denen wir dann durch Elimination einer der zwei Unbekannten eine Gleichung für den Inselrand oder die Konzentration gewinnen können. Bei allen diesen linearen Beziehungen handelt es sich um Differentialgleichungen. Mit einem geeigneten Reihenansatz für c_1 und ρ lassen sich diese in lineare Gleichungssysteme zwischen den Koeffizienten dieser Reihen übersetzen. Fassen wir die Koeffizienten ρ_n von $\rho(\theta, t)$ zu einem Spaltenvektor $\boldsymbol{\rho}$ zusammen und die Koeffizienten von c_1 entsprechend zu einem Vektor \mathbf{c} , so können wir die linearisierte Form von (8.2) symbolisch als Matrixgleichung $\partial_t \boldsymbol{\rho} = \mathbf{M}_1 \mathbf{c}$ schreiben und die (7.9) entsprechende Gleichung als $\mathbf{c} = \mathbf{M}_2 \boldsymbol{\rho}$. Aus diesen beiden Gleichungen folgt dann direkt $\partial_t \boldsymbol{\rho} = (\mathbf{M}_1 \cdot \mathbf{M}_2) \boldsymbol{\rho}$ und die Zeitentwicklung von $\boldsymbol{\rho}$ ist dann einfach $\boldsymbol{\rho}(t) = \exp(t \mathbf{M}_1 \cdot \mathbf{M}_2) \cdot \boldsymbol{\rho}(0)$.

In [44] wurde dagegen eine andere Richtung eingeschlagen. Zuerst wurde die Konzentration in eine Fourierreihe bezüglich θ entwickelt, deren Koeffizienten wiederum Potenzreihen in r sind

$$c_1(r, \theta, t) = \sum_n \left(\sum_m a_{nm} r^m \right) e^{in\theta}.$$

Der Konzentration c_1 entspricht also damit ein Gebilde aus zwei Indizes. Im Prinzip lassen sich, ähnlich wie eben skizziert, mit den Gleichungen (7.1), (7.2), (7.9) und (8.2) Beziehungen für die Koeffizienten a_{nm} und ρ_n herleiten, aus denen sich die ρ_n eliminieren lassen³. Leider war es uns jedoch nicht möglich, den Weg, der auf die in [44] angegebene Eigenwertgleichung für das Stabilitätsproblem führt, zu rekonstruieren.

Nach dieser Vorrede nun zur eigentlichen Rechnung: Obwohl die einzelnen Schritte nicht weiter kompliziert sind, werden wir die Rechnung der Übersichtlichkeit halber in Abschnitte unterteilen.

Die allgemeine Form des Konzentrationsprofils

In der quasistatischen Näherung reduziert sich die Bewegungsgleichung für die Adatomkonzentration auf

$$\nabla^2 c - \frac{1}{\xi} \partial_x c = 0. \quad (8.4)$$

Das Problem, auf das wir hier zunächst stoßen, ist das folgende: Während der Laplace-Operator in Polarkoordinaten in Radial- und Winkelanteil separiert, bricht die Ableitung nach x die Rotationssymmetrie und die Gleichung als Ganzes kann nicht separiert werden. Wir bilden die Gleichung deshalb zunächst mit dem Ansatz⁴

$$c(x, y) = \exp\left(\frac{x}{2\xi}\right) f\left(\frac{x}{2\xi}, \frac{y}{2\xi}\right) \quad (8.5)$$

auf die Helmholtz-Gleichung

$$\nabla^2 f = f$$

ab. Unter Verwendung von (A-14) wird die Gleichung in Polarkoordinaten zu

$$\frac{1}{r} \partial_r (r \partial_r f) + \frac{1}{r^2} \partial_\theta^2 f = f,$$

³Die Rechnung in [44] unterscheidet sich von der unseren, neben dem Ansatz für das Diffusionsfeld, auch darin, dass auf die quasistatische Näherung verzichtet wurde.

⁴Die Faktoren $\frac{1}{2\xi}$ im Argument reskalieren lediglich die Gleichung und bringen so den Faktor $\frac{1}{\xi}$ zum Verschwinden.

was mit dem Ansatz $f(r, \theta) = F(r)\Phi(\theta)$ in die beiden Gleichungen

$$\begin{aligned} \partial_\theta^2 \Phi + \lambda \Phi &= 0 \\ r^2 \partial_r^2 F + r \partial_r F - (\lambda + r^2) F &= 0 \end{aligned} \quad (8.6)$$

zerfällt. Der Winkelanteil Φ muss eine 2π -periodische Funktion von θ sein, so dass der Separationsparameter λ notwendigerweise gleich n^2 mit einem ganzzahligen n sein muss. Die Lösung ist demnach $\Phi_n(\theta) = e^{in\theta}$.

Mit der Substitution $u = ir$ ($i = \sqrt{-1}$) wird aus (8.6) die Besselsche Differentialgleichung

$$\frac{d^2 F}{du^2} + \frac{1}{u} \frac{dF}{du} + \left(1 - \frac{n^2}{u^2}\right) F = 0.$$

Die Lösungen sind die Bessel-Funktionen J_n , so dass die Lösung von (8.6) die sogenannten modifizierten Bessel-Funktionen⁵ ([19], S. 901) mit

$$I_n(z) = (i)^{-n} J_n(iz)$$

sind. Die allgemeine Lösung von (8.4) ist damit

$$c_1(r, \theta) = \exp\left(\frac{r}{2\xi} \cos \theta\right) \sum_{n=-\infty}^{\infty} c_n I_n\left(\frac{r}{2\xi}\right) e^{in\theta}. \quad (8.7)$$

Da die modifizierten Besselfunktionen laut (A-20) symmetrisch bezüglich des Index sind, ist $c_1(r, \theta)$ genau dann eine reelle Funktion, wenn die Koeffizienten der Relation

$$c_{-n} = c_n^*$$

erfüllen.

Die linearisierte Form der Beziehung (8.2)

Der Exponentialfaktor in (8.7) wird von nun an in jeder Gleichung auftauchen, die die Konzentration enthält. Da er sich beim Ableiten reproduziert, wird er insbesondere auch den Strom multiplizieren, der ja nach (7.2) durch eine Differentialoperation aus der Konzentration gewonnen wird.

Die Störung εc_1 des Konzentrationsprofils führt zu einer Störung $\varepsilon \mathbf{J}_1$ mit

⁵Genauer gesagt handelt es sich hier um modifizierte Bessel-Funktionen *der ersten Art*. Die Funktionen zweiter Art K_n (auch MacDonald-Funktionen oder Basset-Funktionen genannt) sind ebenfalls Lösungen, scheiden hier jedoch aus, weil sie im Nullpunkt singular werden.

$$\begin{aligned}
\mathbf{J}_1 &= -D\nabla c_1 + \frac{D}{\xi} c_1 \mathbf{e}_x \\
&= -D\nabla \left(\exp\left(\frac{x}{2\xi}\right) f\left(\frac{x}{2\xi}, \frac{y}{2\xi}\right) \right) + \frac{D}{\xi} \exp\left(\frac{x}{2\xi}\right) f\left(\frac{x}{2\xi}, \frac{y}{2\xi}\right) \mathbf{e}_x \\
&= -D \exp\left(\frac{x}{2\xi}\right) \left(\nabla f\left(\frac{x}{2\xi}, \frac{y}{2\xi}\right) + \frac{1}{2\xi} f\left(\frac{x}{2\xi}, \frac{y}{2\xi}\right) \mathbf{e}_x \right) + \frac{D}{\xi} \exp\left(\frac{x}{2\xi}\right) f\left(\frac{x}{2\xi}, \frac{y}{2\xi}\right) \mathbf{e}_x \\
&= D \exp\left(\frac{x}{2\xi}\right) \left(-\nabla f\left(\frac{x}{2\xi}, \frac{y}{2\xi}\right) + \frac{1}{2\xi} f\left(\frac{x}{2\xi}, \frac{y}{2\xi}\right) \mathbf{e}_x \right) \\
&= D \exp\left(\frac{r \cos \theta}{2\xi}\right) \left(-\left(\frac{\partial f}{\partial r} \mathbf{e}_r + \frac{1}{r} \frac{\partial f}{\partial \theta} \mathbf{e}_\theta\right) + \frac{1}{2\xi} f(\cos \theta \mathbf{e}_r - \sin \theta \mathbf{e}_\theta) \right),
\end{aligned}$$

woraus sich mit der Formel (A-17) die gestörte Normalengeschwindigkeit ergibt⁶

$$\begin{aligned}
v_n &= -a^2 \mathbf{n} \cdot \mathbf{J} \\
&= -a^2 \left(\frac{\varepsilon}{R_0} \frac{\partial \rho}{\partial \theta} \mathbf{e}_\theta - \mathbf{e}_r \right) \left(\frac{D}{\xi} \tilde{c}_{eq} (\cos \theta \mathbf{e}_r - \sin \theta \mathbf{e}_\theta) \right. \\
&\quad \left. + \varepsilon D \exp\left(\frac{R_0 \cos \theta}{2\xi}\right) \left(-\left(\frac{\partial f}{\partial r} \mathbf{e}_r + \frac{1}{r} \frac{\partial f}{\partial \theta} \mathbf{e}_\theta\right) + \frac{1}{2\xi} f(\cos \theta \mathbf{e}_r - \sin \theta \mathbf{e}_\theta) \right) \right) \\
&= \frac{a^2 D}{\xi} \tilde{c}_{eq} \cos \theta + \varepsilon a^2 D \left(\frac{\tilde{c}_{eq}}{R_0 \xi} \sin \theta \frac{\partial \rho}{\partial \theta} + \exp\left(\frac{R_0 \cos \theta}{2\xi}\right) \left(\frac{1}{2\xi} f \cos \theta - \frac{\partial f}{\partial r} \right) \right).
\end{aligned}$$

Damit erhalten wir für die Zeitentwicklung von ρ

$$\partial_t \rho = a^2 D \exp\left(\frac{R_0 \cos \theta}{2\xi}\right) \left(\frac{\partial f}{\partial r} - \frac{1}{2\xi} f \cos \theta \right). \quad (8.8)$$

Alternativ können wir sie natürlich auch durch c_1 ausdrücken als⁷

$$\partial_t \rho = a^2 D \left(\partial_r c_1 - \frac{1}{\xi} c_1 \cos \theta \right). \quad (8.9)$$

Die Beziehung zwischen der Konzentration und der geometrischen Form des Inselrandes

Die Randbedingung (7.9) ergibt unter Benutzung von (A-18) und Entwicklung der Exponentialfunktion

⁶Im Folgenden werden die Funktion f und ihre Ableitungen immer jeweils an der Stelle $r = R_0$ ausgewertet. Um allzu überfrachtete Formeln zu vermeiden, werden wir statt $f(R_0, \theta)$, $\frac{\partial f}{\partial r}(R_0, \theta)$ usw., lediglich f , $\frac{\partial f}{\partial r}$... schreiben.

⁷Es sei darauf hingewiesen, dass sich die Terme in (8.8) und (8.9) *nicht* jeweils entsprechen. Das heißt, der Term $\partial_r f$ übersetzt sich nicht einfach in $\partial_r c_1$. Da $c_1 = \exp(\dots) f$ nach der Produktregel abgeleitet wird, trägt $\partial_r c_1$ zu beiden Termen in (8.8) bei. Dadurch ergeben sich die unterschiedlichen Vorfaktoren $\frac{1}{\xi}$ und $\frac{1}{2\xi}$.

$$\begin{aligned}
\tilde{c}_{eq} + \varepsilon c_1(R_0 + \varepsilon \rho(\theta), \theta) &= c_{eq}^0 \exp\left(\Gamma\left(-\frac{1}{R_0} + \varepsilon \frac{\rho + \frac{\partial^2 \rho}{\partial \theta^2}}{R_0^2}\right)\right) \\
&= c_{eq}^0 \exp\left(-\frac{\Gamma}{R_0}\right) \left(1 + \varepsilon \frac{\Gamma}{R_0^2} \left(\rho + \frac{\partial^2 \rho}{\partial \theta^2}\right)\right) \\
\Rightarrow c_1(R_0, \theta) &= \tilde{c}_{eq} \frac{\Gamma}{R_0^2} \left(\rho + \frac{\partial^2 \rho}{\partial \theta^2}\right). \tag{8.10}
\end{aligned}$$

Diese Beziehung zwischen c_1 und ρ kann nun dazu verwendet werden, in (8.8) die Konzentration (d.h. die Funktion f) durch ρ auszudrücken und dadurch eine geschlossene Gleichung für die Störung des Inselrandes zu gewinnen. Hier stoßen wir allerdings auf einen wichtigen Punkt. Mit der Beziehung (8.10) lässt sich bei vorgegebener Störung $\rho(\theta)$ die Konzentration *am Inselrand* berechnen. Die Zeitentwicklung des Randes (8.8) enthält aber neben der Konzentration auch deren (Normalen-) Ableitung am Rand. Die Beziehung zwischen Funktionswert und Ableitung ist natürlich implizit in (8.4) enthalten, aber sie ist nicht durch eine Differentialgleichung von der Form $\partial_r c_1 = D_\theta c_1$ ausdrückbar (D_θ bezeichnet hier einen Differentialoperator nach θ)⁸. Das führt aber dazu, dass Gleichung (8.8) (nachdem c_1 , bzw. f durch ρ ausgedrückt wurde) ebenfalls keine Differentialgleichung mehr ist. Nun ist dies nicht weiter verwunderlich, denn eine Differentialgleichung würde bedeuten, dass die Zeitentwicklung des Randes vollständig *lokal* ist, d.h. dass die Bewegung eines Punktes lediglich von der lokalen Geometrie um diesen Punkt herum bestimmt wird. Aber da der Massenstrom durch das Konzentrationsfeld vermittelt wird und dieses von der gesamten Inselform abhängt, ist die Dynamik des Gebildes aus Rand und Adatomkonzentration eben gerade *nichtlokal*. Diese Tatsache wird uns allerdings nicht weiter behindern, wenn wir die Reihendarstellung (8.7) verwenden und einen ähnlichen geeigneten Ansatz für $\rho(\theta)$ machen. Damit zerlegen wir Konzentration und Rand in Moden und alle Gleichungen werden damit zu Beziehungen zwischen den Amplituden dieser Moden, ohne dass wir uns darum kümmern müssten, ob sie Differentialgleichungen entsprechen oder nicht. Um die spezielle Struktur von (8.7) und insbesondere den Exponentialfaktor zu berücksichtigen, wählen wir als Ansatz für die Störung des Inselrandes

$$\rho = \exp\left(\frac{R_0}{2\xi} \cos \theta\right) \sum_{n=-\infty}^{\infty} \rho_n \exp(in\theta). \tag{8.11}$$

Da sich diese Wahl im Weiteren als folgenswer erweisen wird, werfen wir zuerst einen Blick auf die Vorteile, die sich aus ihr ergeben. Setzt man einen beliebigen

⁸Vielmehr entspricht die Beziehung einem Faltungs-Operator. Siehe dazu Fußnote 11 in diesem Abschnitt.

Reihenansatz für ρ in Gleichung (8.10) ein, so ergibt sich durch Koeffizientenvergleich eine Beziehung zwischen den Koeffizienten der Reihen für ρ und c_1 . Dieser Koeffizientenvergleich erfordert aber zunächst, dass die beiden Reihen dieselben Funktionen enthalten. Würde man also z.B. für $\rho(\theta)$ eine Fourierreihe ansetzen, so müsste auch zuerst (8.7) in die Form einer solchen Reihe gebracht werden. So erhalten wir unter Verwendung von (A-28)

$$\begin{aligned} c_1(R_0, \theta) &= \sum_{m=-\infty}^{\infty} I_m \left(\frac{R_0}{2\xi} \right) e^{im\theta} \sum_{n=-\infty}^{\infty} c_n I_n \left(\frac{R_0}{2\xi} \right) e^{in\theta} \\ &= \sum_{n=-\infty}^{\infty} \left(\sum_{m=-\infty}^{\infty} c_m I_m \left(\frac{R_0}{2\xi} \right) I_{n-m} \left(\frac{R_0}{2\xi} \right) \right) e^{in\theta}. \end{aligned}$$

Die Amplitude jeder Fouriermode enthält damit *alle* Koeffizienten c_n aus (8.7). Der Koeffizientenvergleich mittels (8.10) wird also dementsprechend ebenfalls eine Beziehung ergeben, die unendlich viele Koeffizienten beinhaltet. Obwohl dieser Weg durchaus gangbar ist, wird sich der Ansatz (8.11) als wesentlich bequemer erweisen und am Ende auch etwas kompaktere Formeln erzeugen. Aus (8.11) erhalten wir zunächst die erste und zweite Ableitung von ρ als

$$\begin{aligned} \frac{\partial \rho}{\partial \theta} &= \exp \left(\frac{R_0}{2\xi} \cos \theta \right) \sum \left(in - \frac{R_0}{2\xi} \sin \theta \right) \rho_n \exp(in\theta) \\ \frac{\partial^2 \rho}{\partial \theta^2} &= \exp \left(\frac{R_0}{2\xi} \cos \theta \right) \sum \left(in(in - \frac{R_0}{2\xi} \sin \theta) - \frac{R_0}{2\xi} \cos \theta \right. \\ &\quad \left. - \frac{R_0}{2\xi} \sin \theta (in - \frac{R_0}{2\xi} \sin \theta) \right) \rho_n \exp(in\theta) \\ \sin^2(\theta) &= \frac{1}{2}(1 - \cos(2\theta)) \\ &\exp \left(\frac{R_0}{2\xi} \cos \theta \right) \sum \left(\frac{R_0^2}{8\xi^2} (1 - \cos(2\theta)) - 2in \frac{R_0}{2\xi} \sin \theta \right. \\ &\quad \left. - \frac{R_0}{2\xi} \cos \theta - n^2 \right) \rho_n \exp(in\theta). \end{aligned}$$

Einsetzen in (8.10) liefert⁹

⁹Wir verwenden dazu die Formeln

$$\begin{aligned} \sum a_n \sin \theta \exp(in\theta) &= \sum \frac{a_{n-1} - a_{n+1}}{2i} \exp(in\theta) \\ \sum a_n \cos \theta \exp(in\theta) &= \sum \frac{a_{n-1} + a_{n+1}}{2} \exp(in\theta) \end{aligned} \quad (8.12)$$

$$\begin{aligned}
\sum_{n=-\infty}^{\infty} c_n I_n \left(\frac{R_0}{2\xi} \right) \exp(in\theta) &= \frac{\tilde{c}_{eq}\Gamma}{R_0^2} \sum \left(1 + \frac{R_0^2}{8\xi^2} - \frac{R_0^2}{8\xi^2} \cos(2\theta) - 2in \frac{R_0}{2\xi} \sin \theta \right. \\
&\quad \left. - \frac{R_0}{2\xi} \cos \theta - n^2 \right) \rho_n \exp(in\theta) \\
&= \frac{\tilde{c}_{eq}\Gamma}{R_0^2} \sum \left(\left(1 + \frac{R_0^2}{8\xi^2} - n^2 \right) \rho_n - \frac{R_0^2}{16\xi^2} (\rho_{n+2} + \rho_{n-2}) - \frac{R_0}{2\xi} \left((n-1)\rho_{n-1} - (n+1)\rho_{n+1} \right) \right. \\
&\quad \left. - \frac{R_0}{4\xi} (\rho_{n+1} + \rho_{n-1}) \right) \exp(in\theta)
\end{aligned}$$

und der Koeffizientenvergleich ergibt

$$\begin{aligned}
c_n I_n \left(\frac{R_0}{2\xi} \right) &= \frac{\tilde{c}_{eq}\Gamma}{R_0^2} \left(\left(1 + \frac{R_0^2}{8\xi^2} - n^2 \right) \rho_n - \frac{R_0^2}{16\xi^2} (\rho_{n+2} + \rho_{n-2}) \right. \\
&\quad \left. - \frac{R_0}{2\xi} \left((n-1)\rho_{n-1} - (n+1)\rho_{n+1} \right) - \frac{R_0}{4\xi} (\rho_{n+1} + \rho_{n-1}) \right) \\
&= \frac{\tilde{c}_{eq}\Gamma}{R_0^2} \left(\left(1 + \frac{R_0^2}{8\xi^2} - n^2 \right) \rho_n - \frac{R_0^2}{16\xi^2} (\rho_{n+2} + \rho_{n-2}) - \frac{R_0}{4\xi} \left((2n-1)\rho_{n-1} - (2n+1)\rho_{n+1} \right) \right)
\end{aligned} \tag{8.13}$$

Diese Gleichung (bzw. dieses Gleichungssystem) stellt nun die benötigte Beziehung zwischen Konzentration und Inselrand her. Als Matrixgleichung geschrieben enthält sie eine Bandmatrix¹⁰ und diese Bandstruktur wird sich auf alle daraus folgenden Gleichungen vererben. Dies ist der speziellen Wahl (8.11) zu verdanken. Wie oben erläutert würde eine, z.B. der Ansatz einer Fourierreihe für ρ , zu einer vollbesetzten Matrix führen.

Die linearisierte Bewegungsgleichung

Mittels (8.13) lässt sich nun die Konzentration c_1 (bzw. f) durch die ρ_n ausdrücken und damit auch die für (8.8) benötigte Ableitung $\partial_r f$. Für diese benutzen wir Formel (A-26) und erhalten¹¹

¹⁰Der Koeffizient c_n hängt nur von den Koeffizienten $\rho_{n-2}, \rho_{n-1}, \dots, \rho_{n+2}$ ab.

¹¹Kommen wir hier noch einmal kurz auf den Zusammenhang zwischen $f(R_0, \theta)$ und $\frac{\partial f}{\partial r}(R_0, \theta)$ zurück. Aus den Gleichungen

$$\begin{aligned}
f(R_0, \theta) &= \sum c_n I_n \left(\frac{R_0}{2\xi} \right) \exp(in\theta) \\
\frac{\partial f}{\partial r}(R_0, \theta) &= \sum c_n \frac{1}{2\xi} \left(I_{n+1} \left(\frac{R_0}{2\xi} \right) + \frac{2\xi n}{r} I_n \left(\frac{R_0}{2\xi} \right) \right) \exp(in\theta)
\end{aligned}$$

ist zu sehen, dass $\frac{\partial f}{\partial r}(R_0, \theta)$ aus $f(R_0, \theta)$ gewonnen werden kann, indem man den n-ten Fourierkoeffizienten mit $\frac{1}{2\xi} \left(\frac{I_{n+1}}{I_n} + \frac{2\xi n}{R_0} \right)$ multipliziert. Diese Operation entspricht aber einer

$$\begin{aligned}
f(r, \theta) &= \sum c_n I_n\left(\frac{r}{2\xi}\right) \exp(in\theta) \\
\Rightarrow f(R_0, \theta) &= \frac{\tilde{c}_{eq}\Gamma}{R_0^2} \sum \left(\left(1 + \frac{R_0^2}{8\xi^2} - n^2\right) \rho_n - \frac{R_0^2}{16\xi^2} (\rho_{n+2} + \rho_{n-2}) \right. \\
&\quad \left. - \frac{R_0}{4\xi} \left((2n-1)\rho_{n-1} - (2n+1)\rho_{n+1} \right) \right) \exp(in\theta) \\
\frac{\partial f}{\partial r}(r, \theta) &= \sum c_n \frac{1}{2\xi} \left(I_{n+1}\left(\frac{r}{2\xi}\right) + \frac{2\xi n}{r} I_n\left(\frac{r}{2\xi}\right) \right) \exp(in\theta) \\
\Rightarrow \frac{\partial f}{\partial r}(R_0, \theta) &= \frac{\tilde{c}_{eq}\Gamma}{2\xi R_0^2} \sum \left(\left(1 + \frac{R_0^2}{8\xi^2} - n^2\right) \rho_n - \frac{R_0^2}{16\xi^2} (\rho_{n+2} + \rho_{n-2}) \right. \\
&\quad \left. - \frac{R_0}{4\xi} \left((2n-1)\rho_{n-1} - (2n+1)\rho_{n+1} \right) \right) \left(\frac{I_{n+1}}{I_n} + \frac{2\xi n}{R_0} \right) \exp(in\theta).
\end{aligned}$$

Einsetzen in (8.8) ergibt

$$\begin{aligned}
\frac{1}{a^2 D} \partial_t \sum \rho_n \exp(in\theta) &= \frac{\tilde{c}_{eq}\Gamma}{2\xi R_0^2} \sum \left(\left(1 + \frac{R_0^2}{8\xi^2} - n^2\right) \rho_n - \frac{R_0^2}{16\xi^2} (\rho_{n+2} + \rho_{n-2}) \right. \\
&\quad \left. - \frac{R_0}{4\xi} \left((2n-1)\rho_{n-1} - (2n+1)\rho_{n+1} \right) \right) \left(\frac{I_{n+1}}{I_n} + \frac{2\xi n}{R_0} \right) \exp(in\theta) \\
&\quad - \frac{\tilde{c}_{eq}\Gamma}{2\xi R_0^2} \cos \theta \sum \left(\left(1 + \frac{R_0^2}{8\xi^2} - n^2\right) \rho_n - \frac{R_0^2}{16\xi^2} (\rho_{n+2} + \rho_{n-2}) \right. \\
&\quad \left. - \frac{R_0}{4\xi} \left((2n-1)\rho_{n-1} - (2n+1)\rho_{n+1} \right) \right) \exp(in\theta) \\
&= \frac{\tilde{c}_{eq}\Gamma}{2\xi R_0^2} \sum \left(\left(1 + \frac{R_0^2}{8\xi^2} - n^2\right) \rho_n - \frac{R_0^2}{16\xi^2} (\rho_{n+2} + \rho_{n-2}) \right. \\
&\quad \left. - \frac{R_0}{4\xi} \left((2n-1)\rho_{n-1} - (2n+1)\rho_{n+1} \right) \right) \left(\frac{I_{n+1}}{I_n} + \frac{2\xi n}{R_0} \right) \exp(in\theta) \\
&\quad - \frac{\tilde{c}_{eq}\Gamma}{4\xi R_0^2} \sum \left(\left(1 + \frac{R_0^2}{8\xi^2}\right) (\rho_{n-1} + \rho_{n+1}) - \left((n-1)^2 \rho_{n-1} + (n+1)^2 \rho_{n+1} \right) \right. \\
&\quad \left. - \frac{R_0^2}{16\xi^2} (\rho_{n+1} + \rho_{n+3} + \rho_{n-3} + \rho_{n-1}) \right. \\
&\quad \left. - \frac{R_0}{4\xi} \left((2n-3)\rho_{n+2} + 2\rho_n - (2n+3)\rho_{n-2} \right) \right) \exp(in\theta)
\end{aligned}$$

Faltung. Bei einer Faltung zweier Fourierreihen multiplizieren sich die Koeffizienten nach

$$\left(\sum f_n e^{in\theta} \right) * \left(\sum g_n e^{in\theta} \right) = \sum f_n g_n e^{in\theta}.$$

Insofern sind $\frac{\partial f}{\partial r}(R_0, \theta)$ und $f(R_0, \theta)$ also über einen Faltungsoperator miteinander verknüpft. f hängt wiederum über (8.10) mit ρ zusammen, so dass sich also die rechte Seite von (8.8) mittels einer Kombination aus Differentiation und Faltungsintegral durch ρ ausdrücken lässt. Damit entspricht die linearisierte Zeitentwicklung einem Integro-Differential-Operator.

so dass sich am Ende durch Koeffizientenvergleich

$$\begin{aligned}
\frac{1}{a^2 \tilde{c}_{eq} D \Gamma} \partial_t \rho_n &= \left(\frac{1}{8\xi^2 R_0} + \frac{1}{2\xi R_0^2} \left(1 + \frac{R_0^2}{8\xi^2} - n^2 \right) \left(\frac{I_{n+1}}{I_n} + \frac{2\xi n}{R_0} \right) \right) \rho_n \\
&+ \left(-\frac{1}{4\xi R_0^2} \left(\left(1 + \frac{R_0^2}{8\xi^2} \right) - (n+1)^2 - \frac{R_0^2}{16\xi^2} \right) + (2n+1) \frac{1}{8\xi^2 R_0} \left(\frac{I_{n+1}}{I_n} + \frac{2\xi n}{R_0} \right) \right) \rho_{n+1} \\
&+ \left(-\frac{1}{4\xi R_0^2} \left(\left(1 + \frac{R_0^2}{8\xi^2} \right) - (n-1)^2 - \frac{R_0^2}{16\xi^2} \right) - (2n-1) \frac{1}{8\xi^2 R_0} \left(\frac{I_{n+1}}{I_n} + \frac{2\xi n}{R_0} \right) \right) \rho_{n-1} \\
&+ \left(-\frac{1}{16\xi^2 R_0} (2n+3) - \frac{1}{32\xi^3} \left(\frac{I_{n+1}}{I_n} + \frac{2\xi n}{R_0} \right) \right) \rho_{n+2} \\
&+ \left(+\frac{1}{16\xi^2 R_0} (2n-3) - \frac{1}{32\xi^3} \left(\frac{I_{n+1}}{I_n} + \frac{2\xi n}{R_0} \right) \right) \rho_{n-2} \\
&+ \frac{1}{64\xi^3} (\rho_{n+3} + \rho_{n-3}) \\
&= \left(\frac{1}{2\xi R_0^2} \left(1 + \frac{R_0^2}{8\xi^2} - n^2 \right) \left(\frac{I_{n+1}}{I_n} + \frac{2\xi n}{R_0} \right) + \frac{1}{8\xi^2 R_0} \right) \rho_n \\
&+ \left(-\frac{1}{64\xi^3} + \frac{1}{4\xi R_0^2} n(n+2) + (2n+1) \frac{1}{8\xi^2 R_0} \left(\frac{I_{n+1}}{I_n} + \frac{2\xi n}{R_0} \right) \right) \rho_{n+1} \\
&+ \left(-\frac{1}{64\xi^3} + \frac{1}{4\xi R_0^2} n(n-2) - (2n-1) \frac{1}{8\xi^2 R_0} \left(\frac{I_{n+1}}{I_n} + \frac{2\xi n}{R_0} \right) \right) \rho_{n-1} \quad (8.14) \\
&+ \left(-\frac{1}{16\xi^2 R_0} (2n+3) - \frac{1}{32\xi^3} \left(\frac{I_{n+1}}{I_n} + \frac{2\xi n}{R_0} \right) \right) \rho_{n+2} \\
&+ \left(\frac{1}{16\xi^2 R_0} (2n-3) - \frac{1}{32\xi^3} \left(\frac{I_{n+1}}{I_n} + \frac{2\xi n}{R_0} \right) \right) \rho_{n-2} \\
&+ \frac{1}{64\xi^3} (\rho_{n+3} + \rho_{n-3})
\end{aligned}$$

ergibt. Diese Gleichung stellt nun die linearisierte Zeitentwicklung der Störung dar, die der folgenden Stabilitätsanalyse zugrunde liegt. Das Gleichungssystem kann wieder durch eine Matrixgleichung dargestellt werden, wobei die Matrix, die der rechten Seite von (8.14) entspricht, im Folgenden mit \mathbf{A} bezeichnet wird. Da die Matrixeinträge A_{ij} nur für $|i-j| \leq 3$ von Null verschieden sind, handelt es sich um eine Bandmatrix, bei der das Band in jeder Zeile sieben von Null verschiedene Einträge umfasst. Es sei ausdrücklich darauf hingewiesen, dass beide Indizes die ganzen (also auch die negativen) Zahlen durchlaufen, die Matrix sich also vertikal und horizontal jeweils „in beide Richtungen“ ins Unendliche erstreckt. Diese Matrixdarstellung, die durch die Reihenentwicklungen (8.7) und (8.11) entstanden ist, ist natürlich eine diskrete Fassung eines Operators $\hat{\mathbf{A}}$, der auf die kontinuierliche Funktion $\rho(\theta)$ wirkt. Dieser Operator ist jedoch für alle praktischen Belange für uns unbedeutend, da hier die diskretisierte Fassung wesentlich besser handhabbar ist.

Abschließend können wir noch eine Reskalierung durchführen, die schon durch

den Ansatz (8.5) nahegelegt wird. Dazu multiplizieren wir (8.14) mit R_0^3 und führen die Bezeichnung z für das Verhältnis R_0/ξ ein und erhalten

$$\begin{aligned}
\frac{R_0^3}{a^2 \tilde{c}_{eq} D\Gamma} \partial_t \rho_n &= \left(\frac{z}{2} \left(1 + \frac{z^2}{8} - n^2 \right) C_n + \frac{z^2}{8} \right) \rho_n \\
&+ \left(\frac{z}{4} n(n+2) + (2n+1) \frac{z^2}{8} C_n - \frac{z^3}{64} \right) \rho_{n+1} \\
&+ \left(\frac{z}{4} n(n-2) + (-2n+1) \frac{z^2}{8} C_n - \frac{z^3}{64} \right) \rho_{n-1} \\
&- \left(\frac{z^2}{16} (2n+3) + \frac{z^3}{32} C_n \right) \rho_{n+2} \\
&+ \left(\frac{z^2}{16} (2n-3) - \frac{z^3}{32} C_n \right) \rho_{n-2} \\
&+ \frac{z^3}{64} (\rho_{n+3} + \rho_{n-3}),
\end{aligned} \tag{8.15}$$

wobei wir außerdem die Abkürzung C_n für den Term $\frac{I_{n+1}}{I_n} + \frac{2n}{z}$ ($= \frac{I_{n+1} + I_{n-1}}{2I_n}$) benutzen. Damit ist deutlich zu ersehen, dass der einzige relevante Parameter für die linearisierte Theorie das Verhältnis $z = R_0/\xi$ ist (die Konstanten auf der linken Seite bestimmen lediglich die Zeitskala). Kleine Werte von z entsprechen kleinen Inseln, bzw. schwachen Feldern und große Werte entsprechen umgekehrt großen Inseln, bzw. starken Feldern.

Wir merken abschließend noch an, dass eine Entwicklung bis zur ersten Ordnung in z (unter Benutzung von (A-27)) das System auf

$$\frac{R_0^3}{a^2 \tilde{c}_{eq} D\Gamma} \partial_t \rho_n = (1 - n^2) |n| \rho_n \tag{8.16}$$

$$+ \frac{z}{4} (n^2 + 2n|n| + 2n + |n|) \rho_{n+1} \tag{8.17}$$

$$+ \frac{z}{4} (n^2 - 2n|n| - 2n + |n|) \rho_{n-1} \tag{8.18}$$

reduziert.

8.2 Symmetriebetrachtungen

Um über die lineare Stabilität unseres Systems zu entscheiden, werden wir das Gleichungssystem (8.15) dazu verwenden, folgendes Eigenwert-Problem zu formulieren: Um das Verhalten von beliebigen Störungen $\rho(\theta, t)$ zu untersuchen, genügt es im Allgemeinen, den zeitlichen Verlauf der sogenannten *Eigenmoden* zu behandeln. Darunter verstehen wir Störungen, die exponentiell mit der Zeit anwachsen oder zerfallen, also die Gleichung

$$\partial_t \rho = \tilde{\lambda} \rho$$

erfüllen. Unter Verwendung dieser Beziehung wird (8.15) zu der Eigenwertgleichung¹² $\hat{\mathbf{A}}\rho = \lambda\rho$, bzw. in der Matrix-Schreibweise

$$\mathbf{A}\rho = \lambda\rho. \quad (8.19)$$

Die Frage nach der Stabilität wird dann durch die Betrachtung der Eigenwerte beantwortet. Existiert ein positiver Eigenwert, so ist die zugehörige Eigenmode instabil, andernfalls stabil. Da wir jede beliebige Störung durch Überlagerung von Eigenmoden darstellen können, ist somit das System genau dann stabil, wenn alle Eigenwerte negativ sind¹³.

Wir werden das Gleichungssystem numerisch behandeln. Allerdings wollen wir vorher einige Betrachtungen über die Symmetrien des Problems anstellen, die uns im Voraus einige analytische Ergebnisse über die Eigenmoden liefern werden. Kommen wir zuerst auf die diskrete Reflektionssymmetrie bezüglich der x -Achse (d.h. der Feldrichtung) zu sprechen. Das Feld ist der einzige Einfluss, der eine bestimmte Richtung auszeichnet und eine Spiegelung bezüglich dieser Richtung lässt die Dynamik unverändert. Das bedeutet zunächst, dass eine Insel, deren Form spiegelsymmetrisch zur Feld-Achse ist, diese Symmetrie während ihrer zeitlichen Entwicklung beibehält. Für die linearisierte Bewegung gemäß (8.15) bedeutet es aber noch etwas mehr. Reflektion an der x -Achse bedeutet für die Störung die Transformation

$$\rho(\theta) \rightarrow \rho(-\theta)$$

wendet man diese Operation auf die Reihendarstellung (8.11) an, so erkennt man, dass dies wegen

$$\begin{aligned} \rho(\theta) &= \exp\left(\frac{R_0}{2\xi} \cos \theta\right) \sum_{n=-\infty}^{\infty} \rho_n \exp(in\theta) \\ \Rightarrow \rho(-\theta) &= \exp\left(\frac{R_0}{2\xi} \cos \theta\right) \sum_{n=-\infty}^{\infty} \rho_n \exp(-in\theta) \\ &= \exp\left(\frac{R_0}{2\xi} \cos \theta\right) \sum_{n=-\infty}^{\infty} \rho_{-n} \exp(in\theta) \end{aligned}$$

¹²Es ist sinnvoll, die Konstanten, die die linke Seite von (8.15) multiplizieren, in den Eigenwert λ aufzunehmen. Der Ausdruck $\frac{R_0^3}{a^2 c_{eq} D \Gamma}$ hat die Dimension einer Zeit, so dass der Eigenwert $\lambda = \frac{R_0^3}{a^2 c_{eq} D \Gamma} \tilde{\lambda}$ dimensionslos ist.

¹³Es gibt gewisse Überlegungen, die die Gültigkeit dieser Schlussweise etwas einschränken. Siehe dazu den Abschnitt 9.4.

gleichbedeutend mit

$$\rho_n \rightarrow \rho_{-n}$$

ist. Nun ist der Matrix in (8.15) aufgrund der oben beschriebenen Reflektionsinvarianz eine besondere Symmetrie zu eigen. Es ist unmittelbar ersichtlich, dass die Koeffizienten z.B. von ρ_{n+1} und ρ_{n-1} durch $n \rightarrow -n$ aufeinander abgebildet werden¹⁴, d.h. für die entsprechenden Matrixelemente gilt $A_{n,n+1} = A_{-n,n-1}$, was wiederum nichts anderes als $A_{-n,-(n+1)}$ ist. Analoges gilt für die restlichen Matrixelemente, so dass die Matrix die Symmetrie $A_{n,m} = A_{-n,-m}$ besitzt¹⁵. Daraus folgt nun, dass zwischen einem Vektor \mathbf{x} und seinem Bild $\mathbf{y} = \mathbf{A}\mathbf{x}$ die Beziehung

$$\begin{aligned} y_n &= \sum_m A_{n,m} x_m \\ \Rightarrow y_{-n} &= \sum_m A_{-n,m} x_m \\ &= \sum_m A_{-n,-m} x_{-m} \\ &\stackrel{\text{Symmetrie}}{=} \sum_m A_{n,m} x_{-m} \end{aligned} \quad (8.20)$$

besteht. Für eine symmetrische Störung $\rho_n = \rho_{-n}$ heißt das, dass die Symmetrie durch die von \mathbf{A} erzeugte Zeitentwicklung erhalten bleibt, wie es schon für die volle nichtlineare Zeitentwicklung der Fall ist. Darüberhinaus zeigt (8.20) aber auch, dass die Antisymmetrie $\rho_n = -\rho_{-n}$ einer Störung erhalten bleibt. Wir haben also hier eine Situation, die der Paritätserhaltung in der Quantenmechanik ähnelt. Dort folgt im Falle eines symmetrischen Potentials, dass die Lösungen der Schrödingergleichung in zwei Klassen zerfallen, von denen eine aus symmetrischen, die andere aus antisymmetrischen Wellenfunktionen besteht und deren Zeitentwicklung niemals einen Übergang zwischen den beiden Klassen erzeugt. Damit zerfallen dort, wie auch in unserem Fall, insbesondere die Eigenmoden in diese zwei Klassen, ein Sachverhalt, von dem wir weiter unten bei der Handhabung des Gleichungssystems noch Gebrauch machen werden.

Eine noch wichtigere Symmetrie ist mit der Tatsache verbunden, dass die Dynamik gegenüber der Lage der Insel auf dem Substrat gewissermaßen blind ist. Die Position z.B. des Schwerpunktes geht an keiner Stelle in die Bewegungsgleichungen ein, wodurch sich die *Translationsinvarianz* des Systems ergibt. Es ist eine Invarianz gegenüber Verschiebungen in beliebige Richtungen. Das heißt, Störungen von der Form einer Translation \mathcal{T}_x in x -Richtung, bzw. \mathcal{T}_y in y -Richtung

¹⁴Für den Term $C_n = \frac{I_{n+1} + I_{n-1}}{2I_n}$ gilt wegen (A-20) $C_n = C_{-n}$.

¹⁵Das heißt, die Matrixeinträge gehen gewissermaßen durch Punktspiegelung an $A_{0,0}$ in sich selbst über.

erfüllen $\partial_t \mathcal{T}_{x,y} = 0$ und liefern uns damit zwei Eigenmoden zum Eigenwert $\lambda = 0$.

Wie sehen nun diese beiden Eigenmoden in der Darstellung (8.11) aus? Zunächst sind die entsprechenden Funktionen $\rho(\theta) = \cos \theta$ für die horizontale Translation und $\rho(\theta) = \sin \theta$ für die vertikale¹⁶. Wir erhalten mit

$$\begin{aligned} \cos \theta &= \exp\left(\frac{z}{2} \cos \theta\right) \exp\left(-\frac{z}{2} \cos \theta\right) \cos \theta \\ &\stackrel{(A-28)}{=} \exp\left(\frac{z}{2} \cos \theta\right) \cos \theta \sum_n I_n\left(-\frac{z}{2}\right) e^{in\theta} \\ &\stackrel{(8.12)}{=} \exp\left(\frac{z}{2} \cos \theta\right) \sum_n \frac{1}{2} \left(I_{n-1}\left(-\frac{z}{2}\right) + I_{n+1}\left(-\frac{z}{2}\right)\right) e^{in\theta} \end{aligned}$$

die Darstellung

$$\mathcal{T}_x : \rho_n = \frac{1}{2} \left(I_{n-1}\left(-\frac{z}{2}\right) + I_{n+1}\left(-\frac{z}{2}\right)\right) \stackrel{(A-21)}{=} \frac{(-1)^{n+1}}{2} \left(I_{n-1}\left(\frac{z}{2}\right) + I_{n+1}\left(\frac{z}{2}\right)\right) \quad (8.21)$$

und mit einer analogen Rechnung

$$\mathcal{T}_y : \rho_n = \frac{1}{2i} \left(I_{n-1}\left(-\frac{z}{2}\right) - I_{n+1}\left(-\frac{z}{2}\right)\right) = \frac{(-1)^{n+1}}{2i} \left(I_{n-1}\left(\frac{z}{2}\right) - I_{n+1}\left(\frac{z}{2}\right)\right). \quad (8.22)$$

Diese beiden Formeln werden wir bei der numerischen Behandlung des Eigenwertproblems dazu verwenden, die Qualität der Ergebnisse zu überprüfen, indem wir die numerisch ermittelten Eigenvektoren zum Eigenwert Null mit den analytischen Ausdrücken (8.21) und (8.22) vergleichen.

Wir sehen also, dass der Nulleigenwert mindestens zweifach entartet sein muss. Abgesehen davon enthält er jedoch noch eine Subtilität. Zunächst bemerken wir, dass die Dynamik für $z = 0$ eine weitere Symmetrie enthält. Da ja kreisförmige Inseln mit einem beliebigen Radius stationäre Lösungen sind, kann man durch Vergrößerung des Radius und gleichzeitiger Erhöhung der konstanten Adatomkonzentration von einer Lösung zur nächsten gelangen. Ist diese Vergrößerung infinitesimal, so haben wir also eine Störung \mathcal{D} , für die ebenfalls $\partial_t \mathcal{D} = 0$ gilt. Diese Störung ist aber für $z > 0$ keine Eigenmode mehr. Wir betrachten hier immer Störungen einer kreisrunden Insel *mit einer bestimmten Driftgeschwindigkeit*. Da die infinitesimale Vergrößerung des Radius (und damit der Konzentration) eine Geschwindigkeitsänderung nach sich zieht (siehe Gleichung

¹⁶Wie man leicht sieht, wird mit diesem Ansatz die rechte Seite von (8.10) zu Null, durch eine Störung dieser Art bleibt also die Konzentration ungestört ($c_1 \equiv 0$). Damit bleibt die Störung des Randes aber auch für den Strom und für die Normalgeschwindigkeit ohne Wirkung. So folgt aus (8.9) direkt $\partial_t \rho \equiv 0$.

(8.3)), gilt hier nicht mehr $\partial_t \mathcal{D} = 0$. Vielmehr driften die gestörte und die ungestörte Form aufgrund der unterschiedlichen Geschwindigkeiten auseinander, die Störung *wächst* in diesem Sinne sogar an! Versuchen wir diesen Sachverhalt anhand der Gleichungen nachzuvollziehen! Zunächst folgt aus Gleichung (8.10), dass die Störung $\rho(\theta) \equiv 1$ (bei der sich also der Radius um ε vergrößert) zu einer Störung $c_1 = \tilde{c}_{eq} \frac{\Gamma}{R_0^2}$ der Konzentration führt. Setzt man dies wiederum in (8.9) ein, so ergibt sich $\partial_t \rho = \frac{a^2 \tilde{c}_{eq} \Gamma D}{\xi R_0^2} \cos \theta$, und damit $\partial_t \mathcal{D} = \frac{a^2 \tilde{c}_{eq} \Gamma D}{\xi R_0^2} \mathcal{T}_x$. Das System erwidert also gewissermaßen die Dilatation mit einer Translation. Dies ist genau die infinitesimale Beschleunigung, von der weiter oben die Rede war. Dieser interessante Umstand hat nun Konsequenzen für die Beschaffenheit des Nullleigenwerts. Wir haben die folgenden beiden Gleichungen

$$\begin{aligned}\hat{\mathbf{A}}\mathcal{D} &= \mu \mathcal{T}_x \\ \hat{\mathbf{A}}\mathcal{T}_x &= 0\end{aligned}$$

mit der Bezeichnung $\mu := \frac{a^2 \tilde{c}_{eq} \Gamma D}{\xi R_0^2}$ und damit

$$\hat{\mathbf{A}}^2 \mathcal{D} = 0.$$

Diese Gleichung besagt, dass es sich bei der Dilatation um einen *verallgemeinerten Eigenvektor* handelt. Der Null-Eigenwert ist damit dreifach entartet, wobei die geometrische Vielfachheit (d.h. die Anzahl an linear unabhängigen Eigenvektoren) zwei beträgt. Die Matrix lässt sich in diesem Fall nicht vollständig diagonalisieren. Beschränken wir uns für einen Moment auf den Raum, der von dem System $\{\mathcal{T}_x, \mathcal{D}, \mathcal{T}_y\}$ aufgespannt wird. Mit diesen drei Moden als Basis nimmt die Matrix die Jordan-Form

$$\mathbf{J} = \begin{pmatrix} 0 & \mu & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

an¹⁷. Das zugehörige charakteristische Polynom hat natürlich die dreifache Nullstelle $\lambda = 0$, d.h. die *algebraische Vielfachheit* und damit der Entartungsgrad des Eigenwertes ist drei, bei nur zwei (echten) Eigenvektoren.

8.3 Das Spektrum

Um das Gleichungssystem (8.19) numerisch zu behandeln, müssen wir es geeignet trunkieren. Dazu machen wir uns die oben angeführte Spiegelsymmetrie

¹⁷Genau genommen sind bei der Jordan-Form die Nichtdiagonalelemente gleich Eins, was sich hier jedoch durch eine triviale Reskalierung von \mathcal{D} erreichen lässt.

zunutze. Wir werden die beiden Eigenräume der symmetrischen und antisymmetrischen Moden von nun an getrennt behandeln. Für den symmetrischen Teil haben wir

$$\rho_n = \rho_{-n} \quad (8.23)$$

$$A_{n,m} = A_{-n,-m} \quad (8.24)$$

so dass wir die Eigenwertgleichung folgendermaßen umformen können

$$\begin{aligned} \lambda \rho_n &= \sum_{m=-\infty}^{\infty} A_{n,m} \rho_m \\ &= \underbrace{\sum_{m=-\infty}^{-1} A_{n,m} \rho_m}_{\sum_{m=1}^{\infty} A_{n,-m} \rho_{-m} \stackrel{(8.24)}{=} \sum_{m=1}^{\infty} A_{-n,m} \rho_m} + A_{n,0} \rho_0 + \sum_{m=1}^{\infty} A_{n,m} \rho_m \\ &= A_{n,0} \rho_0 + \sum_{m=1}^{\infty} (A_{n,m} + A_{-n,m}) \rho_m. \end{aligned}$$

Im Prinzip durchläuft der Index n immer noch alle ganzen Zahlen. Wir können aber den negativen Teil davon außer Acht lassen, da durch (8.23) der gesamte Vektor ρ bereits durch den halben Satz der Koeffizienten bestimmt wird. Analog führt dieses Verfahren für die antisymmetrischen Moden auf¹⁸¹⁹

$$\lambda \rho_n = \sum_{m=1}^{\infty} (A_{n,m} - A_{-n,m}) \rho_m.$$

Das Gleichungssystem erstreckt sich nun nur noch „in eine Richtung“ ins Unendliche. Um es zu einem endlichen System zu machen, müssen wir es noch zu den großen Indizes hin abschneiden, um damit eine Matrix endlicher Dimension N zu erhalten. Dies entspricht im Wesentlichen der Vernachlässigung von sehr kurzwelligen Störungen, ähnlich dem Trunkieren von Fourierreihen. Wie wir im Abschnitt (9.3) sehen werden, ist diese Näherung für die ersten Eigenmoden durchaus gerechtfertigt. Will man zu höheren Eigenmoden übergehen, muss entsprechend der cutoff N zu größeren Werten hin verschoben werden. Im Prinzip ist diese Näherung jedoch problemlos und gut kontrollierbar.

Nach dieser Vorarbeit besteht die eigentliche Eigenwertbestimmung in der Anwendung eines numerischen Verfahrens. Dazu haben wir im Wesentlichen die

¹⁸Der Term $A_{n,0} \rho_0$ verschwindet, da bei den antisymmetrischen Moden $\rho_0 = 0$ ist.

¹⁹Wir haben es also im Folgenden mit zwei Gleichungssystemen und entsprechend mit zwei Matrizen zu tun. Da jedoch keine Verwirrung zu befürchten ist, werden wir in beiden Fällen von der Matrix \mathbf{A} sprechen.

Routinen des LAPACK-Programmpakets verwendet, entweder in der Form der üblichen Bibliotheken zur Einbindung in C-Programme, oder durch die entsprechenden vorgefertigten Funktionen in Octave bzw. Matlab. Auf die numerischen Eigenheiten, die mit der Lösung des Eigenwertproblems zusammenhängen, werden wir später in den Abschnitten 9.1 und 9.5 eingehen. Hier präsentieren wir lediglich die Ergebnisse in Form der Abbildung 8.1.

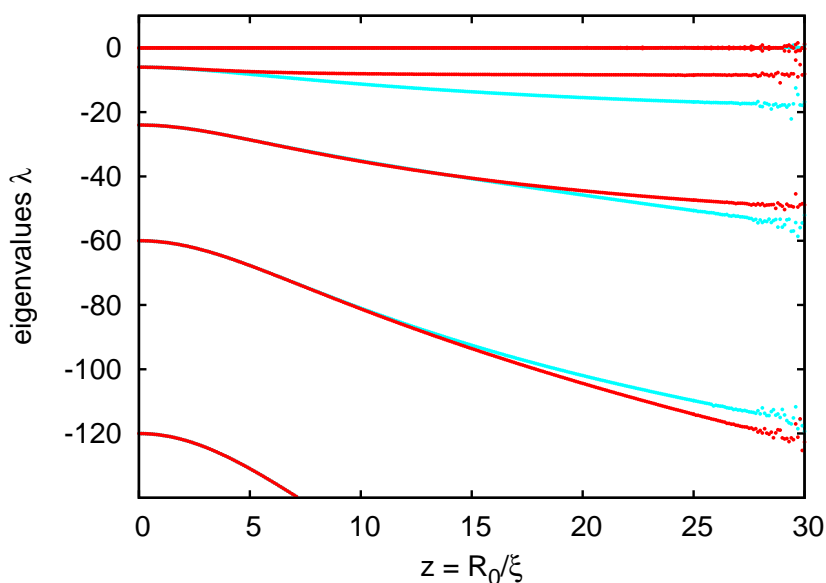


Abbildung 8.1: Die numerisch bestimmten Eigenwerte der Matrix **A**. Dargestellt sind die Eigenwerte zu den symmetrischen (schwarz) und antisymmetrischen (grau) Eigenmoden als Funktion des Parameters z . Zu den Unregelmäßigkeiten der Linien bei den höheren z -Werten, siehe Abschnitt (9.5).

Die wichtigste Aussage, die diese grafische Darstellung bereithält, ist, dass offensichtlich (bis auf die vorhergesagten Null-Eigenwerte) alle Eigenwerte negativ sind. Die Folgerung daraus ist also, dass das System linear stabil ist und zwar offensichtlich für alle Werte von z . Im Prinzip lassen sich die Eigenwerte ab $z \geq 30$ nicht mehr zuverlässig bestimmen, aber wie wir in der Zusammenfassung am Ende der Arbeit sehen werden, entspricht bereits dies Inselgrößen, bzw. Feldstärken, die keinen realistischen experimentellen Bedingungen mehr entsprechen. Des Weiteren besteht die allgemeine Tendenz, dass die Eigenwerte mit wachsendem z immer stärker in den negativen Bereich hineinwachsen. Das heißt, die dimensionslosen Zerfallskonstanten der Störungen werden immer größer und das System findet bezüglich der dimensionslosen Zeit immer schneller zu der Kreisform zurück. Diese Eigenwerte rechnen sich mit $\tilde{\lambda} = \frac{a^2 \tilde{c}_{eq} D \Gamma}{R_0^3} \lambda$ auf die physikalischen, d.h. dimensionsbehafteten Zerfallsraten um. Der Umrechnungsfaktor enthält unter anderem den Inselradius R_0 , so dass größere Inseln

bei festgehaltenem Verhältnis R_0/ξ *langsamer* zerfallen. Hält man dagegen R_0 fest und geht zu immer größeren Feldstärken über, so zerfallen die Störungen schneller und das System wird in diesem Sinne *stabiler*.

Wir werden im Abschnitt 10 die Frage nach der Stabilität erneut stellen, indem wir die numerischen Lösungen der vollen nichtlinearen Gleichungen daraufhin untersuchen, inwieweit *endliche* Störungen zu einer Instabilität führen könnten. Gehen wir jedoch noch einmal kurz auf Abbildung 8.1 und den Grenzwert $z \rightarrow 0$ ein. Aus (8.16) können wir die Matrix \mathbf{A} für $z = 0$ bestimmen. In diesem Fall verschwindet das Feld und die Dynamik der Insel wird alleine von der Diffusion bestimmt. Die Matrix \mathbf{A} wird diagonal und die Eigenwerte sind

$$\lambda_n = |n|(n^2 - 1). \quad (8.25)$$

Die zugehörigen Eigenvektoren sind einfach die Einheitsvektoren²⁰. Die Reihe (8.11) reduziert sich für $z = 0$ auf eine reine Fourierreihe. Wir erkennen damit in dem nullten Eigenvektor $\rho_n = \delta_{n0} \Rightarrow \rho(\theta) = 1$ die Dilatation wieder, die in diesem Fall ein echter Eigenvektor ist. Die Eigenvektoren mit Index ± 1 entsprechen den Moden $\rho(\theta) = e^{\pm i\theta}$ was wir natürlich durch Linearkombination in die reellen Moden $\rho(\theta) = \cos(\theta) = \mathcal{T}_x$ und $\sin(\theta) = \mathcal{T}_y$ verwandeln können. Für $z = 0$ erhalten wir also die vorhergesagte dreifache Entartung. Für $z > 0$ ist die Dilatation keine Eigenmode mehr, der Null-Eigenwert bleibt jedoch dreifach entartet. Das numerische Verfahren bestimmt die Eigenwerte in der richtigen algebraischen Vielfachheit (kann jedoch nur die echten Eigenvektoren bestimmen), so dass die Linie $\lambda = 0$ in Abbildung (8.1) tatsächlich aus drei exakt übereinanderliegenden Linien besteht. Da für $z = 0$ kein Feld die Rotationssymmetrie bricht, ist mit jeder Eigenmode $\rho(\theta)$ auch die gedrehte Störung $\rho(\theta + \theta_0)$ eine Eigenmode. Da die (reellen) Eigenmoden $\cos(n\theta)$ und $\sin(n\theta)$ durch eine Drehung auseinander hervorgehen, sind daher die symmetrischen und antisymmetrischen Moden über die Rotationssymmetrie miteinander verbunden und jeder Eigenwert deshalb zweifach entartet²¹. Diese Entartung wird durch das Feld aufgehoben und die Linien in Abbildung 8.1 spalten sich dementsprechend für $z > 0$ in Paare auf²².

²⁰Hier durchlaufen die Indizes wieder alle ganzen Zahlen. Der m -te Einheitsvektor hat die Darstellung $\rho_n = \delta_{nm}$.

²¹Wie man auch aus (8.25), bei der jeweils $\pm n$ denselben λ -Wert ergeben.

²²Die Aufspaltung der ($n = 5$) -Linie in Abbildung 8.1, die bei $z = 0$ den Wert $\lambda = -120$ hat, liegt außerhalb des abgebildeten Bereichs. Die Grafik zeigt also insgesamt die ersten elf Eigenwerte ($n = -5 \dots 5$), die jedoch teilweise übereinander liegen.

Kapitel 9

Pseudospektren und nichtnormale Matrizen

Grundsätzlich zeigen die Ergebnisse aus Abschnitt 8.3 durch die Negativität der Eigenwerte die lineare Stabilität unseres Systems. Trotzdem werden wir uns in diesem Abschnitt etwas eingehender mit dem System der Eigenmoden befassen und mit Fragen, die über eine bloße Betrachtung des Spektrums hinausgehen. Grund hierfür ist zum einen, dass diese Fragen einige Eigentümlichkeiten der Matrix \mathbf{A} und des zugehörigen Eigenwertproblems (8.19) beleuchten, zum anderen kündigt sich gerade in der Struktur der Eigenmoden ein Verhalten an, das mit der noch zu besprechenden nichtlinearen Instabilität in Zusammenhang zu stehen scheint.

9.1 Störanfälligkeit des Eigenwertproblems

Eine Besonderheit des Eigenwertproblems lässt sich anhand eines Punktes darstellen, an dem unsere Untersuchung eine Zeitlang in die Irre geführt wurde. Die numerische Behandlung des Systems (8.19) zerfällt gewissermaßen in zwei Teile. Zum einen die Implementierung der Matrix und zum anderen die eigentliche Eigenwertbestimmung. Für die Eigenwertbestimmung wurden die Routinen der LAPACK-Bibliothek verwendet, die einen seit langem etablierten und gut getesteten Standard für Eigenwertprobleme darstellen [1]. Was das Aufstellen der Matrix angeht, so enthalten die Matrixeinträge die modifizierten Besselfunktionen $I_n(x)$ und es stellt sich damit also zunächst die Frage nach einem geeigneten numerischen Verfahren, um diese zu berechnen. Zuerst fiel unsere Wahl auf die Numerical Recipes-Routine `bessi`, die im Wesentlichen $I_0(x)$ und $I_1(x)$ polynomial approximiert und alle höheren Besselfunktionen rekursiv aus diesen beiden berechnet (Siehe auch [47], Abschnitt 6.6). Die Eigenwerte, die sich damit ergeben sind in den Abbildungen 9.1 und 9.2 dargestellt. Der Unterschied zu Abb. 8.1 ist frappierend. Nicht nur, dass das Spektrum komplex ist, ab $z \simeq 20$ entstehen Eigenwerte mit positivem Realteil, wie es charakteristisch

für eine lineare Instabilität ist. Der einzige Hinweis, dass das Spektrum nicht ganz richtig sein kann, ist die Tatsache, dass die Entartung des Nulleigenwerts in der Nähe von $z \simeq 16$ aufgehoben wird. Aus den Überlegungen des vorangegangenen Abschnitts wissen wir natürlich, dass die Entartung für beliebige z -Werte bestehen bleiben muss. Auf der Suche nach der Fehlerquelle griffen wir schließlich für die Berechnung der Besselfunktionen auf die Routine `bessik` der Numerical Recipes zurück. Hier beruht die Berechnung auf Darstellungen durch Kettenbrüche, die dann mit der sogenannten Steed-Methode ausgewertet werden ([47], Abschnitt 6.7 und [61]). Diese Implementierung der Matrix erzeugt nun das richtige Spektrum (9.3). Das Verblüffende ist nun der gewaltige Unterschied der beiden Spektren, obwohl sich die numerischen Werte, die beide Routinen für die Besselfunktionen liefern im Vergleich dazu nur relativ geringfügig unterscheiden¹. So ist für $z = 25$, an einem Punkt also, an dem die Abweichungen der Spektren bereits sehr deutlich sind, die relative Abweichung von der Größenordnung 10^{-7} , die daraus resultierende Abweichung der Matrixeinträge ist von derselben Größenordnung².

Die Änderung, die dies im Spektrum hervorruft, ist nicht einfach nur die einer Deformation, sondern eine strukturelle, d.h. topologische. Die Linien sind nun nicht mehr voneinander getrennt, sondern laufen zusammen und auseinander, wobei sich nichtverschwindende Imaginärteile ergeben³. Ein weiterer interessanter Unterschied zwischen den beiden Spektren sind die Fluktuationen im `bessik`-Spektrum ab etwa $z = 26$ (siehe Abbildung 9.3). Wir werden auf die Ursache noch zu sprechen kommen. Zu erwähnen ist an dieser Stelle, dass sie ein Merkmal des echten Spektrums sind, während die Linien des verfälschten Spektrums glatt sind. Auch dies ist ein Punkt, der bei der Untersuchung eine Zeitlang für Verwirrung sorgte. Unser Vorurteil war selbstverständlich, dass eine Routine, die glatte Linien erzeugt, vertrauenswürdiger sein sollte. Die dargestellten Veränderungen sind in einem gewissen Sinne generisch. Auch eine absichtliche, relativ beliebige Störung der Matrix hat ähnliche Auswirkungen und erzeugt zusammenlaufende Linien, Glättung und positive Eigenwerte.

Um nun diese offensichtliche Empfindlichkeit gegenüber kleinen Änderungen der Matrix etwas genauer zu analysieren, bietet es sich zunächst an, anstatt lediglich ein Spektrum darzustellen, die Spektren von sehr vielen gestörten Matrizen zu betrachten. Dies ist auch die Grundidee der Theorie der Pseudospektren, die wir etwas weiter unten benutzen werden. Wir untersuchen exemplarisch den Fall $z = 25$. Das Verfahren ist das Folgende: Zu den von Null verschiedenen Einträgen der Matrix \mathbf{A} werden Zufallszahlen addiert, die gleichverteilt zwischen

¹Die Differenz hängt allerdings vom Argument z ab und wird bei sehr großen Werten durchaus beträchtlich.

²Tatsächlich ist der Unterschied bereits im ersten Eintrag $A_{1,1}$ am größten, die Einträge unterscheiden sich immer weniger, je höher die Indizes sind. Dieser Matrixeintrag ist selbst von der Größenordnung 10^3 .

³D.h. die Trajektorien $(z, \lambda(z))$ in $\mathbb{R}^+ \times \mathbb{C}$ schneiden sich untereinander.

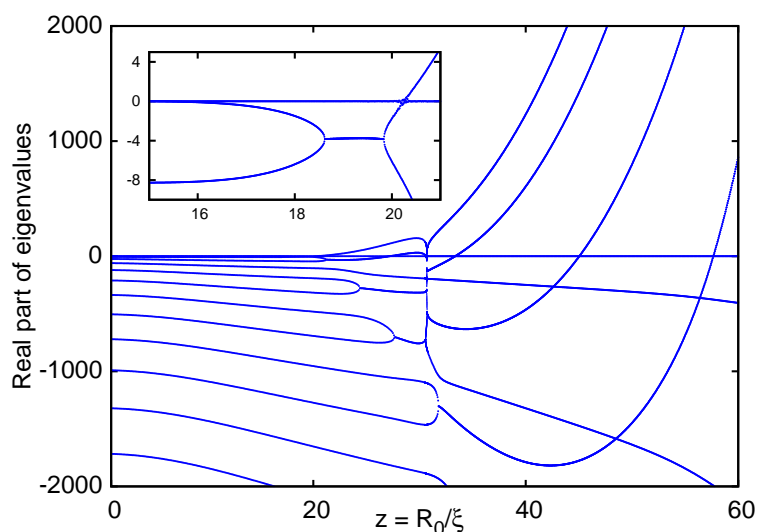


Abbildung 9.1: Realteil der Eigenwerte von \mathbf{A} . Die Matrixeinträge wurden unter Benutzung der Routine `bessi` berechnet. Im Gegensatz zum Spektrum in Abbildung 8.1 werden die meisten Eigenwerte mit steigendem z komplex. An diesen Stellen laufen jeweils zwei Linien zu einer zusammen, d.h. aus zwei reellen Werten entsteht ein komplexes Paar mit gleichem Realteil (siehe auch Abbildung unten). Die Ausschnittvergrößerung zeigt den Bereich, in dem der erste Eigenwert positiv wird.

0 und ε liegen⁴. Die Lage der Eigenwerte dieser Matrizen wird dann als Punkt in der komplexen Ebene dargestellt. Würde man tatsächlich *alle* Matrizen mit Störungen zwischen 0 und ε berücksichtigen, ergäbe sich ein zusammenhängender Bereich⁵, mit unseren zufälligen Störungen wählen wir also gewissermaßen Stichproben aus diesem (kontinuierlichen) Ensemble aus und gewinnen dadurch einen Überblick über die Ausdehnung und Struktur dieses Bereichs und damit über die Auswirkung der Störung. Die Eigenwerte von 10^4 solcher Matrizen sind in Abbildung 9.4 für verschiedene Werte von ε dargestellt. Zunächst sticht ins Auge, dass die Eigenwerte durch die Störungen im Allgemeinen vom Reellen ins Komplexe getragen werden. Nur ungefähr ein Prozent der Punkte liegt auf der reellen Achse. Die komplexen Werte bilden Punktwolken, die rippenförmig senkrecht von der reellen Achse abstehen, wobei das Maximum der Punktdichte mit steigender Amplitude ε der Störung immer weiter von der Achse abrückt⁶.

⁴Wir betrachten hier ausschließlich *reelle* Störungen *der nichtverschwindenden Einträge*, da es derartige Störungen sind, die zu der oben dargestellten Verfälschung des Spektrums führen.

⁵Die Eigenwerte sind als Nullstellen des charakteristischen Polynoms stetige Funktionen der Polynomkoeffizienten (siehe z.B. [25], S. 281). Die Koeffizienten sind wiederum stetige Funktionen der Matrixeinträge.

⁶Insofern bilden die Punkte für ein gegebenes ε also eher Inseln als Rippen. Aber natürlich sind die Störungen zu z.B. $\varepsilon = 10^{-7}$ in denen zu $\varepsilon = 10^{-6}$ und 10^{-5} enthalten und der

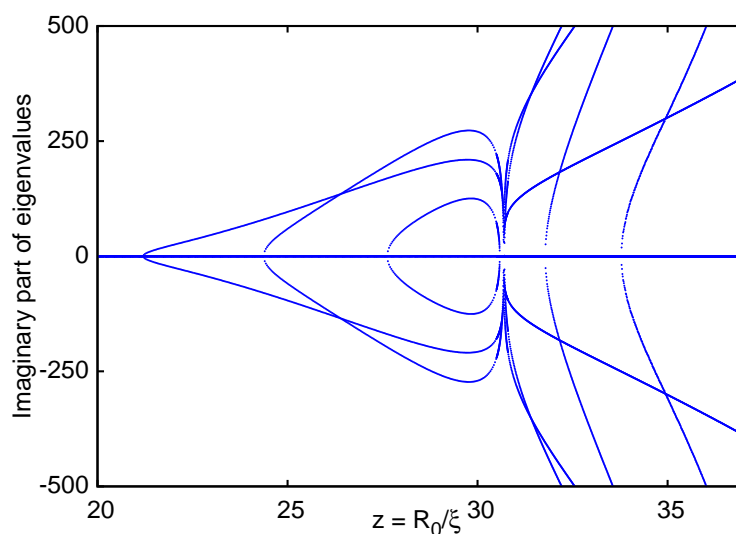


Abbildung 9.2: Imaginärteil der Eigenwerte von \mathbf{A} . Da die Matrix \mathbf{A} reell ist, kommen komplexe Eigenwerte nur in konjugierten Paaren vor, d.h. die Daten liegen spiegelsymmetrisch zur z -Achse. Die Stellen, an denen nichtverschwindende Imaginärteile entstehen, korrespondieren mit den Stellen im oberen Bild, an denen Linien zusammenlaufen. Eigenwerte können also in bestimmten Bereichen gleichen Real- oder Imaginärteil haben, sind aber im Allgemeinen nicht entartet (außer in einzelnen Punkten). Ausnahme ist natürlich der entartete Nulleigenwert, wobei hier jedoch die Entartung bei ungefähr $z \simeq 16$ aufgehoben wird.

Die Eigenwerte erhalten also einen nichtverschwindenden Imaginärteil, der auch keineswegs klein ist, sondern durchaus (abhängig von ε) von derselben Größenordnung wie der Realteil sein kann. Interessanterweise liegen diese Realteile (d.h. die horizontale Lage der Rippen) nicht bei den ungestörten Eigenwerten, sondern dazwischen. Dies hängt eng mit dem oben erwähnten Zusammenlaufen der Eigenwert-Linien zusammen. Der Grund für das Zusammenlaufen ist der Folgende: Die Dimension der Matrix und die Tatsache, dass es sich um eine rein reelle Matrix handelt, bleiben von den Störungen unberührt. Damit auch die feste Anzahl der Eigenwerte und der Umstand, dass sie immer in komplex konjugierten Paaren auftreten. Ein reeller Eigenwert (und für kleine Werte von z sind die Eigenwerte auch bei Störung reell) kann also nicht einfach komplex werden, sondern muss sich in ein konjugiertes Paar aufspalten. Da die Gesamtzahl der Eigenwerte sich nicht verändert, muss er sich dazu allerdings mit einem anderen (reellen) Eigenwert vereinigen. Anders ausgedrückt wird sozusagen die Entartung bezüglich des Imaginärteils aufgehoben und der Realteil entartet. Bildlich

scheinbar geringe Überlapp der entsprechenden Dichteverteilungen ist durch die unterschiedliche Häufigkeit bedingt, mit der solche Störungen durch den Zufallsgenerator erzeugt werden. Größere Störungen erzeugen also häufiger größere Imaginärteile.

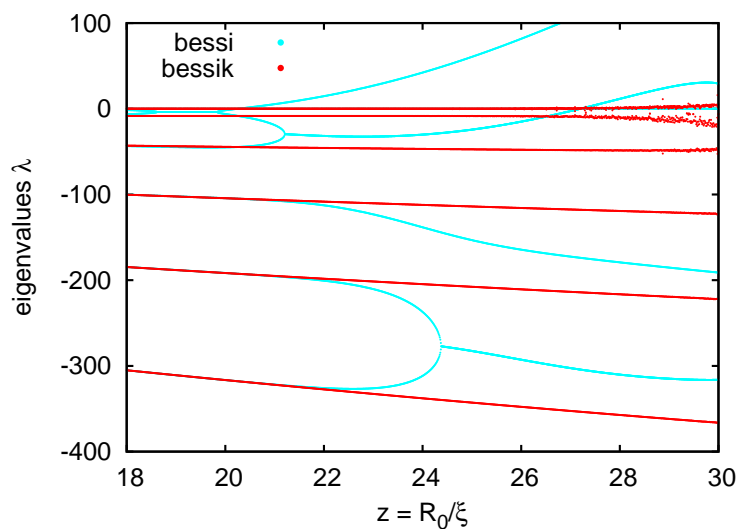


Abbildung 9.3: Eigenwerte als Funktion von z . Die Matrix wurde einmal mit der `bessi`- und einmal mit der `bessik`-Routine erstellt. Letztere ergibt das korrekte Spektrum aus Abschnitt 8.3. Zu den Fluktuationen der größten `bessik`-Eigenwerte ab $z = 26$ siehe Abschnitt 9.5.

gesprochen wandern also zwei benachbarte ungestörte Eigenwerte auf der reellen Achse zusammen, treffen sich etwa in der Mitte und laufen als konjugiertes Paar wieder in imaginärer Richtung auseinander. Was aus Abbildung 9.4 nicht zu ersehen ist, ist die Verteilung der *reellen* Eigenwerte. Diese ist in Abbildung 9.5 dargestellt⁷. Wie zu erwarten vergrößern sich bei Einschalten der Störung die Schwankungen um den exakten Wert, bis sich die Werte für $\varepsilon = 10^{-5}$ praktisch gleichmäßig über den gesamten Bereich verteilen. Die Schwankungsbreite hängt von der Ordnung des Eigenwertes ab, wobei die Eigenwerte mit hohen Indizes weniger stark schwanken als die Niedrigeren. Dies ist in Übereinstimmung damit, dass sich die Verfälschungen zuerst im oberen Teil des Spektrums bemerkbar machen (vgl. Abb. 9.3). Am bedeutsamsten im gegenwärtigen Zusammenhang ist aber der Peak um den größten Eigenwert (d.h. derjenige, der durch die Störung aus dem Nulleigenwert hervorgeht). Selbst für die kleinste dargestellte Störampplitude liegt das Maximum bei einem positiven Wert und es wandert mit steigendem ε immer weiter in den positiven Bereich hinein. Das bedeutet, dass eine gestörte Matrix typischerweise einen positiven und damit instabilen Eigenwert besitzt.

⁷Die Verteilungen spiegeln die relative Häufigkeit derjenigen Punkte aus Abbildung 9.4 wieder, die exakt auf der reellen Achse liegen. Wie oben angemerkt handelt es sich hierbei nur um einen kleinen Bruchteil (etwa ein Prozent).

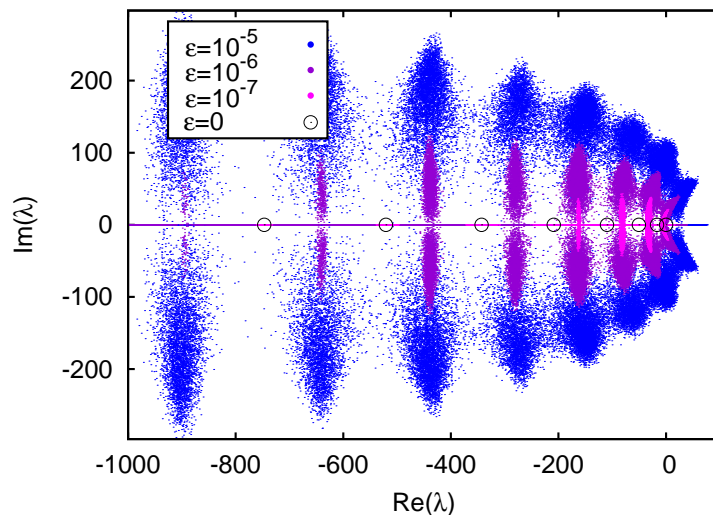


Abbildung 9.4: Eigenwerte der gestörten Matrix für $z = 25$: Die nichtverschwindenden Matrixeinträge wurden mit einer zufälligen Störung zwischen 0 und ε versehen. Dargestellt sind die Eigenwerte von 10^4 derart gestörter Matrizen in der komplexen Ebene. Die schwarzen Kreise kennzeichnen das ungestörte Spektrum. Da es sich um reelle Matrizen handelt, liegen die Punkte symmetrisch zur reellen Achse.

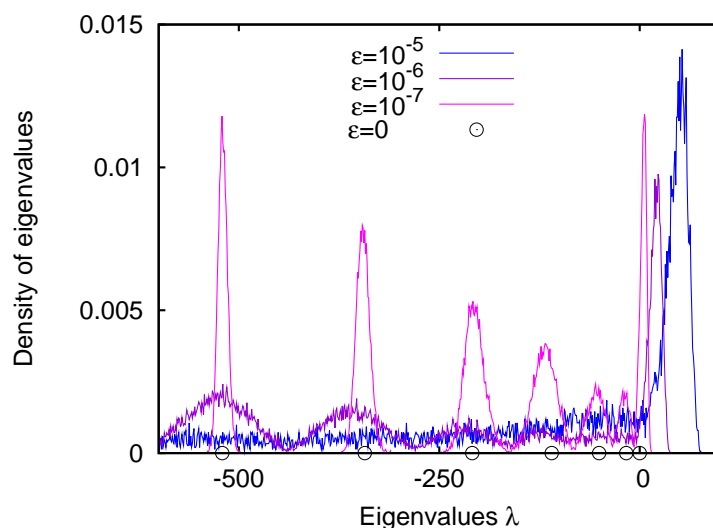


Abbildung 9.5: Verteilung der gestörten Eigenwerte entlang der reellen Achse: Dargestellt ist die Dichteverteilung der Punkte aus Abb. 9.4, die auf der reellen Achse liegen. Das starke Rauschen ist eine Folge der relativ geringen Stichprobenanzahl. Für jeden ε -Wert fließen etwa 10^3 Werte ein.

9.2 Theorie der Pseudospektren

Wir werden nun im Folgenden die Ursache für die Empfindlichkeit der Eigenwerte etwas eingehender untersuchen. Dazu führen wir den zentralen Begriff des *Pseudospektrums* ein⁸. Für die Definition benötigen wir die *Matrixnorm*, die, ausgehend von einer Norm⁹ $\|\mathbf{x}\|$ für Vektoren \mathbf{x} aus dem \mathbb{R}^N bzw. \mathbb{C}^N durch

$$\|\mathbf{M}\| = \max_{\mathbf{x}} \frac{\|\mathbf{M}\mathbf{x}\|}{\|\mathbf{x}\|}$$

definiert ist. Anders ausgedrückt ist sie das Maximum von $\|\mathbf{M}\mathbf{x}\|$ auf der Einheitskugel $\|\mathbf{x}\| = 1$ (für diagonalisierbare Matrizen entspricht das einfach dem Betrag des (betragsmäßig) größten Eigenwertes). Durch sie wird der Raum der Matrizen zu einem normierten und damit metrischen Raum, was es insbesondere erlaubt, zu präzisieren, was unter einer „kleinen“ Störung einer Matrix zu verstehen ist. Wir notieren noch drei Folgerungen aus der Definition. So ergibt sich sofort die Ungleichung

$$\|\mathbf{M}\mathbf{x}\| \leq \|\mathbf{M}\| \|\mathbf{x}\|, \quad (9.1)$$

woraus für ein Matrixprodukt $\mathbf{M}_1\mathbf{M}_2$ folgt

$$\begin{aligned} \|\mathbf{M}_1\mathbf{M}_2\mathbf{x}\| &\leq \|\mathbf{M}_1\| \|\mathbf{M}_2\mathbf{x}\| \leq \|\mathbf{M}_1\| \|\mathbf{M}_2\| \|\mathbf{x}\| \\ \Rightarrow \|\mathbf{M}_1\mathbf{M}_2\| &= \max_{\mathbf{x}} \frac{\|\mathbf{M}_1\mathbf{M}_2\mathbf{x}\|}{\|\mathbf{x}\|} \leq \|\mathbf{M}_1\| \|\mathbf{M}_2\| \end{aligned} \quad (9.2)$$

und schließlich für die Inverse

$$\begin{aligned} 1 &= \|\mathbf{M}\mathbf{M}^{-1}\| \leq \|\mathbf{M}\| \|\mathbf{M}^{-1}\| \\ \Rightarrow \|\mathbf{M}^{-1}\| &\geq \frac{1}{\|\mathbf{M}\|} \end{aligned} \quad (9.3)$$

Mit der Bezeichnung $\sigma(\mathbf{M})$ für das Spektrum der Matrix \mathbf{M} formulieren wir die folgende¹⁰

⁸Die nachfolgende Darstellung von Elementen der Theorie der Pseudospektren beruht auf [64]. Die eingeführten Begriffe erweitern sich auf Operatoren, wir werden uns jedoch der Einfachheit halber auf (endlichdimensionale) Matrizen beschränken. Des Weiteren werden wir die Sätze dieses Abschnitts nicht streng beweisen, sondern lediglich andeuten, wie sich die Zusammenhänge ergeben. Für eine strengere Herleitung sei auf [64] verwiesen.

⁹In den meisten Fällen benutzen wir die euklidische Norm $\|\mathbf{x}_2\|_2 = (\sum |x_i|^2)^{\frac{1}{2}}$.

¹⁰ x bezeichnet hier eine *komplexe* Zahl. Wir benutzen diese Notation, um Verwechslungen mit der Größe $z = \frac{R_0}{\xi}$ zu vermeiden.

Definition 1. Das ε -Pseudospektrum ist die Menge der $x \in \mathbb{C}$, so dass

$$x \in \sigma(\mathbf{M} + \mathbf{E})$$

für eine Matrix $\mathbf{E} \in \mathbb{C}^{N \times N}$ mit $\|\mathbf{E}\| < \varepsilon$.

Das heißt, das ε -Pseudospektrum besteht aus den Eigenwerten aller ε -gestörten Matrizen. Interessant ist nun nicht das Pseudospektrum zu einem einzelnen ε -Wert, sondern vielmehr die Gesamtheit aller σ_ε für verschiedene Werte von ε . Aus der Definition ist ersichtlich, dass das Pseudospektrum von der Wahl der Norm abhängt¹¹. Im physikalischen Zusammenhang gibt es häufig eine „natürliche“ Norm, die sich aus der Physik des jeweiligen Problems ergibt, z.B. in Form einer Energie. Für unseren Fall ist allerdings nicht unmittelbar klar, welche Norm zu bevorzugen wäre. Diese Abhängigkeit von der gewählten Norm wird uns im Folgenden noch beschäftigen.

So wie das Spektrum essenzielle Informationen über die zugrundeliegende Matrix enthält, können Pseudospektren eine gewisse Charakterisierung erlauben, die in bestimmten Fällen eine stärkere Aussagekraft hat als das Spektrum allein. Ein großer Vorzug dieser Begriffsbildung ist es zunächst, dass sie ähnlich wie im vorangegangenen Abschnitt dazu verwendet werden kann, die Störanfälligkeit einer Matrix zu visualisieren. Die σ_ε bilden ineinandergeschachtelte Punktmen- gen $\sigma_{\varepsilon_1} \subset \sigma_{\varepsilon_2} \subset \dots$, deren Gestalt die Auswirkung von Störungen abbildet. Eine alternative Definition des Pseudospektrums ergibt sich daraus, dass die Eigenwerte diejenigen x -Werte sind, für welche die Matrix $x - \mathbf{M}$ singular wird. Das heißt, es gibt einen Vektor \mathbf{v} so dass $(x - \mathbf{M})\mathbf{v} = \mathbf{0}$. Diesen Vektor können wir ohne Einschränkung der Allgemeinheit als normiert annehmen. Für einen Eigenwert x_ε der mit \mathbf{E} , $\|\mathbf{E}\| < \varepsilon$ gestörten Matrix gibt es also demzufolge einen normierten Vektor \mathbf{v}_ε mit

$$\begin{aligned} (x_\varepsilon - (\mathbf{M} + \mathbf{E}))\mathbf{v}_\varepsilon &= \mathbf{0} \\ \Leftrightarrow (x_\varepsilon - \mathbf{M})\mathbf{v}_\varepsilon &= \mathbf{E}\mathbf{v}_\varepsilon \end{aligned}$$

Nehmen wir auf beiden Seiten die Norm, ergibt sich

$$\|(x_\varepsilon - \mathbf{M})\mathbf{v}_\varepsilon\| = \|\mathbf{E}\mathbf{v}_\varepsilon\| \leq \|\mathbf{E}\|\|\mathbf{v}_\varepsilon\| < \varepsilon.$$

Damit folgt unter Ausnutzung von (9.3) für die *Resolvente* $(x - \mathbf{M})^{-1}$

¹¹Im Gegensatz zum Spektrum, für das Konzepte wie ε -Umgebungen etc. nicht notwendig sind. Auf den Punkt gebracht: „Pseudospectra are an idea of analysis; eigenvalues belong to algebra“ [64], S. 13.

$$\|(x_\varepsilon - \mathbf{M})^{-1}\| \geq \frac{1}{\varepsilon}.$$

Für einen exakten Eigenwert ist $x - \mathbf{M}$ nicht invertierbar, d.h. die Resolvente existiert nicht, was wir formal als $\|(x - \mathbf{M})^{-1}\| = \infty$ schreiben können. Im Falle eines ε -Pseudoeigenwertes wird dagegen $\|(x - \mathbf{M})^{-1}\|$ lediglich sehr groß, was die folgende Definition motiviert, die sich als äquivalent zu der vorhergehenden erweist:

Definition 2. Das ε -Pseudospektrum ist die Menge der $x \in \mathbb{C}$, so dass

$$\|(x - \mathbf{M})^{-1}\| > \varepsilon^{-1}$$

Damit lassen sich die Pseudospektren grafisch als Konturlinien der Funktion $\|(x - \mathbf{M})^{-1}\|$ in der komplexen Ebene darstellen.

In welchen Fällen ist das Pseudospektrum nun nützlich? Dazu ist erst einmal zu bemerken, dass es für eine wichtige Klasse von Matrizen tatsächlich keine wesentlichen Informationen enthält, die über das bloße Spektrum hinausgehen. Diese Klasse ist die der *normalen* Matrizen. Eine Matrix ist normal, wenn ihre Eigenvektoren senkrecht aufeinander stehen, etwas präziser

Definition 3. Eine Matrix $\mathbf{M} \in \mathbb{C}^{N \times N}$ heißt normal, wenn sie ein vollständiges System von orthogonalen Eigenvektoren besitzt, das heißt, wenn sie unitär diagonalisierbar ist:

$$\mathbf{M} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^*.$$

Hierbei ist \mathbf{U} die unitäre Matrix der Eigenvektoren und $\mathbf{\Lambda}$ die Diagonalmatrix der Eigenwerte. Der Stern bedeutet hier Transposition und Konjugation. Alternativ lässt sich eine normale Matrix dadurch charakterisieren, dass sie mit ihrer Konjugierten kommutiert, d.h. $\mathbf{M}\mathbf{M}^* = \mathbf{M}^*\mathbf{M}$.¹² Wichtigstes Beispiel für normale Matrizen sind natürlich die hermiteschen Matrizen. Für normale Matrizen lässt sich nun der Betrag der Resolventen sehr leicht bestimmen. Zunächst ist für *Diagonalmatrizen* $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_N)$ die Resolvente einfach

$$(x - \mathbf{\Lambda})^{-1} = (\text{diag}(x - \lambda_1, \dots, x - \lambda_N))^{-1} = \text{diag}\left(\frac{1}{x - \lambda_1}, \dots, \frac{1}{x - \lambda_N}\right)$$

und ihre Norm damit

¹²Die beiden Definitionen, die auf den ersten Blick wenig miteinander zu tun zu haben scheinen, sind beide äquivalent zu der Aussage, dass \mathbf{M} und \mathbf{M}^* dieselben Eigenvektoren besitzen.

$$\begin{aligned}\|(x - \mathbf{\Lambda})^{-1}\| &= \max\left(\left|\frac{1}{x - \lambda_1}\right|, \dots, \left|\frac{1}{x - \lambda_N}\right|\right) \\ &= \frac{1}{\min(|x - \lambda_1|, \dots, |x - \lambda_N|)} \\ &= \frac{1}{\text{dist}(x, \sigma(\mathbf{\Lambda}))},\end{aligned}$$

wobei $\text{dist}(x, \sigma(\mathbf{\Lambda}))$ den üblichen Abstand eines Punktes von einer Menge (hier des Spektrums) in der komplexen Ebene bezeichnet. Für normale Matrizen, die per Definition unitär äquivalent zu einer Diagonalmatrix sind, überträgt sich diese Eigenschaft wegen

$$\begin{aligned}\|(x - \mathbf{M})^{-1}\| &= \|(x - \mathbf{U}\mathbf{\Lambda}\mathbf{U}^*)^{-1}\| \\ &= \|(\mathbf{U}(x - \mathbf{\Lambda})\mathbf{U}^*)^{-1}\| \\ &= \|\mathbf{U}(x - \mathbf{\Lambda})^{-1}\mathbf{U}^*\| \\ &= \|(x - \mathbf{\Lambda})^{-1}\|.\end{aligned}$$

Im letzten Schritt haben wir ausgenutzt, dass unitäre Transformationen die Matrixnorm invariant lassen. Damit ist für eine normale Matrix das Pseudospektrum

$$\sigma_\varepsilon(\mathbf{M}) = \{x : \text{dist}(x, \sigma(\mathbf{M})) < \varepsilon\},$$

also die Vereinigung von kreisförmigen Gebieten mit Radius ε , wobei die Mittelpunkte jeweils die Eigenwerte von \mathbf{M} sind. Offensichtlich sind also die Fälle, für die das Pseudospektrum von Interesse ist, unter den *nichtnormalen* Matrizen zu suchen. Zur Illustration untersuchen wir die zwei folgenden einfachen Beispiele:

$$\mathbf{M}_1 = \begin{pmatrix} -1 & 5 \\ 0 & -2 \end{pmatrix} \quad \mathbf{M}_2 = \begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix}.$$

Beide Matrizen besitzen die Eigenwerte $\lambda_1 = -1$ und $\lambda_2 = -2$. Die Eigenvektoren von \mathbf{M}_1 sind

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \mathbf{v}_2 = \frac{1}{\sqrt{26}} \begin{pmatrix} 5 \\ -1 \end{pmatrix}.$$

Wir fassen \mathbf{v}_1 und \mathbf{v}_2 zu der Matrix $\mathbf{V} = (\mathbf{v}_1, \mathbf{v}_2)$ zusammen und berechnen $\mathbf{V}\mathbf{V}^*$, deren Einträge einfach die Skalarprodukte zwischen \mathbf{v}_1 und \mathbf{v}_2 sind¹³

¹³Letztlich sind wir natürlich nur an diesen Skalarprodukten interessiert. Den hier etwas umständlich erscheinenden Weg über das Matrixprodukt $\mathbf{V}\mathbf{V}^*$ schlagen wir nur ein, um diese für große Matrizen bequeme Darstellungsweise einzuführen, die wir später noch benutzen werden.

$$\mathbf{V}\mathbf{V}^* = \begin{pmatrix} \mathbf{v}_1 \cdot \mathbf{v}_1 & \mathbf{v}_1 \cdot \mathbf{v}_2 \\ \mathbf{v}_2 \cdot \mathbf{v}_1 & \mathbf{v}_2 \cdot \mathbf{v}_2 \end{pmatrix} = \begin{pmatrix} 1 & \frac{5}{\sqrt{26}} \\ \frac{5}{\sqrt{26}} & 1 \end{pmatrix}.$$

Die Diagonale besteht wegen der Normierung von $\mathbf{v}_{1/2}$ aus Einsen und die Matrix ist natürlich symmetrisch. Interessant für uns sind nun die Nichtdiagonalelemente, die das Skalarprodukt zwischen verschiedenen Eigenvektoren angeben. Der Wert $\frac{5}{\sqrt{26}} \simeq 0.98$ zeigt an, dass die beiden Vektoren keineswegs orthogonal sind, sondern ganz im Gegenteil einen relativ kleinen Winkel von 11.3° einschließen. Es handelt sich also um ein Beispiel einer nichtnormalen Matrix. Ihre Pseudospektren sind in Abbildung 9.6, 9.7 den Pseudospektren der Matrix \mathbf{M}_2 gegenübergestellt^{14 15}.

Die Pseudospektren $\sigma_\varepsilon(\mathbf{M}_2)$ sind kreisförmig mit Radius ε , wobei sich die Kreise ab $\varepsilon = \frac{1}{2}$ überlappen. Ab hier besteht jedes Pseudospektrum aus einem zusammenhängenden Gebiet, das beide Eigenwerte umschließt. Das heißt, bei einer Veränderung von \mathbf{M}_2 mit einer Störung der Amplitude $> \frac{1}{2}$ (d.h. praktisch von der Größenordnung der Matrixeinträge) können die beiden Eigenwerte nicht mehr sauber getrennt werden. Bei der Matrix \mathbf{M}_1 tritt dieses Phänomen schon ab etwa $\varepsilon \simeq 0.05$ auf und auch insgesamt sind die Pseudospektren deutlich größer¹⁶. Dieses Verhalten ist im Allgemeinen umso stärker, je kleiner der Winkel zwischen den Eigenvektoren ist.

Wir werden nun auf einen Punkt zu sprechen kommen, der bei allen Aussagen über Pseudospektren berücksichtigt werden muss. Die (Nicht-) Orthogonalität von Vektoren ist natürlich immer nur im Bezug auf ein bestimmtes Skalarprodukt formulierbar. In unserem Fall sind die Eigenvektoren immer als N -dimensionale Spaltenvektoren gegeben, die numerisch aus der Matrix \mathbf{A} bestimmt werden. Es liegt nun nahe, Skalarprodukte mit dem kanonischen inneren Produkt des \mathbb{R}^N , d.h. mit $\mathbf{x} \cdot \mathbf{y} = \sum x_i y_i$ auszurechnen. Doch welches Skalarprodukt damit wirklich berechnet wird, hängt von der Basis ab, da die Komponenten x_i und y_i eben die Entwicklungskoeffizienten bezüglich *eines bestimmten* Satzes von Basisvektoren sind¹⁷. Tatsächlich hängen ja auch die Einträge der Matrix selbst davon ab, welche Basis gewählt wurde. Die Matrix \mathbf{A} wirkt auf die Koeffizienten ρ_n der Entwicklung $\rho(\theta) = \exp(\frac{z}{2} \cos(\theta)) \sum \rho_n \exp(in\theta)$. Das heißt, die zugrundeliegende Basis besteht aus den Funktionen $\mathbf{e}_n := \exp(\frac{z}{2} \cos(\theta)) \exp(in\theta)$

¹⁴Wir benutzen der Einfachheit halber eine Diagonalmatrix, da wir ja bereits wissen, dass eine beliebige normale Matrix mit diesen Eigenwerten die gleichen Pseudospektren besitzt.

¹⁵Für die Berechnung der Pseudospektren benutzen wir das Matlab-Paket EigTool, das von Thomas G. Wright im Rahmen seiner Doktorarbeit am Oxford University Computing Laboratory entwickelt wurde. Das Paket ist frei verfügbar unter <http://web.comlab.ox.ac.uk/projects/pseudospectra/eigtool/>

¹⁶Tatsächlich sind die ε -Pseudospektren immer größer als die ε -Kreise. Nur bei normalen Matrizen fallen beide zusammen.

¹⁷Dieser Umstand kann leicht zu Irrtümern führen, besonders dann, wenn sich die Basis, wie in unserem Fall, eher implizit aus der Rechnung ergibt als durch eine explizite Wahl.

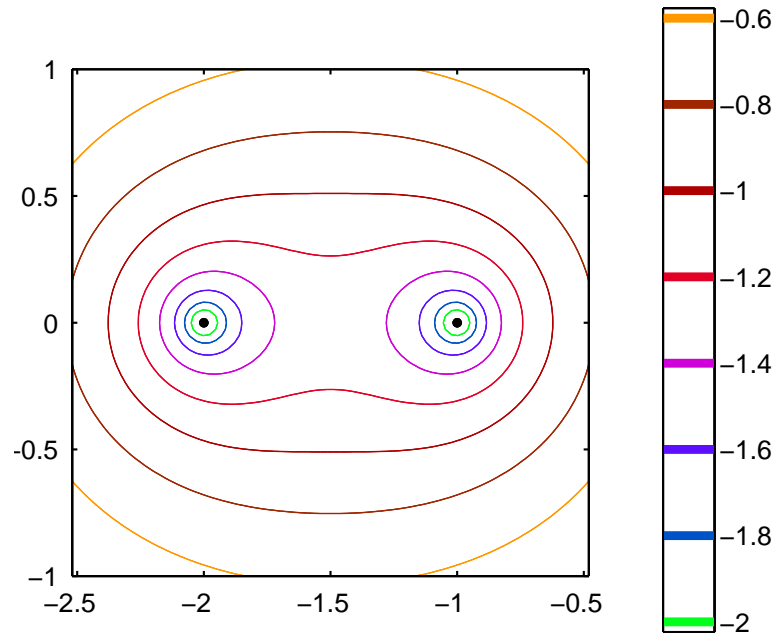


Abbildung 9.6: Pseudospektren der Matrix M_1 . Die schwarzen Punkte bei -1 und -2 sind die (echten) Eigenwerte. Die Farbskala rechts gibt $\log_{10}(\varepsilon)$ an.

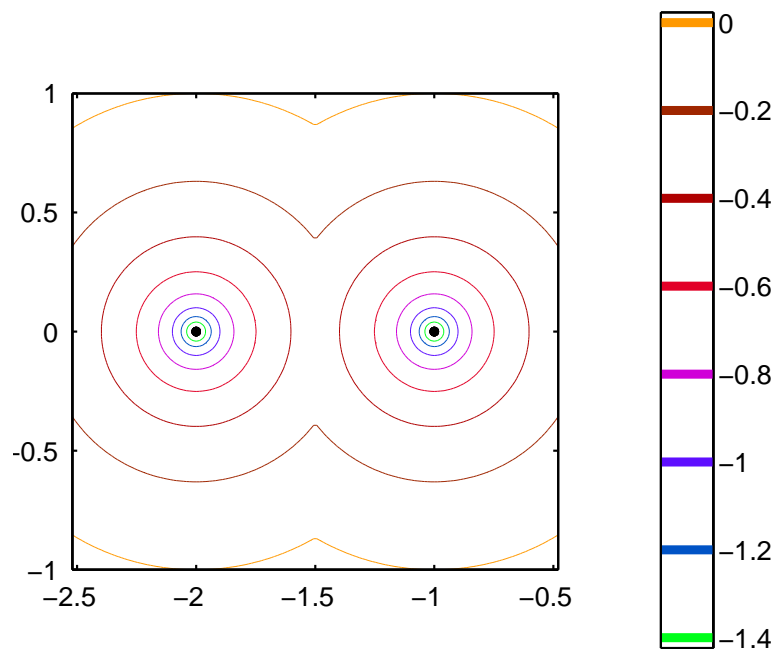


Abbildung 9.7: Pseudospektren der Matrix M_2 . Zu beachten ist, dass die ε -Skala eine andere ist als im Bild darüber.

bzw. deren Real- und Imaginärteil. Diese Basis ergibt sich in natürlicher Weise aus unserer Rechnung¹⁸, ist aber durch den Exponentialfaktor für große z in anderer Hinsicht etwas „unnatürlich“: Durch den Faktor erhalten die Funktionswerte eine starke, von θ abhängige Gewichtung, die in der Nähe von $\theta = 0$ am größten ist, bei $\theta = \pi$ dagegen exponentiell klein wird. Die Skalarprodukte der (numerischen) Eigenvektoren bilden wir mit $\mathbf{x} \cdot \mathbf{y} = \sum x_i y_i$, wobei x_i und y_i die Komponenten der Vektoren *bezüglich der Basis* \mathbf{e}_n sind. Das heißt, wir verwenden ein Skalarprodukt, bezüglich dessen die \mathbf{e}_n orthogonal sind. Anders ausgedrückt induziert die Wahl der Basis ein Skalarprodukt, das den exponentiellen Gewichtungsfaktor enthält.

Unsere Aussagen über Nichtnormalität und Pseudospektren gelten ebenfalls immer nur bezogen auf dieses Skalarprodukt und die damit verbundene Norm. Für die mathematische Struktur des Eigenwertproblems ist unsere Basis durchaus angepasst, aus physikalischer Sicht gibt es jedoch keinen Grund, sie zu bevorzugen. Wir werden deshalb einen Vergleich mit dem üblichen L^2 -Skalarprodukt anstellen, um den Aspekt der Basisabhängigkeit deutlicher herauszuarbeiten und ein Verständnis dafür zu gewinnen, welche Aussagekraft den pseudospektralen Eigenschaften zukommt. Dies wird insbesondere für das nachfolgende Kapitel 9.4 von Bedeutung sein, in dem wir das transiente Verhalten der linearisierten Dynamik untersuchen werden.

Für das L^2 -Skalarprodukt bietet sich die übliche Basis der trigonometrischen Funktionen $\cos(n\theta)$ und $\sin(n\theta)$ an, für die wir die Bezeichnung Fourierbasis verwenden werden. Im Anhang A.6 werden die Gleichungssysteme für den Wechsel zu Fourierbasis hergeleitet. Sie entsprechen einer Matrix \mathbf{T} , mit der sich die Transformation des Basiswechsels einfach als eine Matrixmultiplikation der Form $\tilde{\mathbf{A}} = \mathbf{T}\mathbf{A}\mathbf{T}^{-1}$ schreiben lässt.

Analysieren wir nun unsere Matrix \mathbf{A} unter dem Gesichtspunkt der Nichtnormalität! Zuerst berechnen wir die oben eingeführte Matrix $\mathbf{V}\mathbf{V}^*$ der Skalarprodukte zwischen den (normierten) Eigenvektoren. Es bietet sich hier an, die Matrix als eine zweidimensionale Anordnung aus kleinen Quadraten darzustellen, wobei jedes Quadrat mit einem Grauwert versehen ist, der den Betrag des jeweiligen Skalarprodukts kennzeichnet¹⁹. Abbildung 9.8 zeigt $\mathbf{V}\mathbf{V}^*$ für $z = 1, 5, 10$ und 20 bei einer Matrixgröße²⁰ von $N = 100$. Die Diagonalelemente $\mathbf{v}_i \cdot \mathbf{v}_i$ sind aufgrund der Normierung Eins und die Diagonale entsprechend weiß. Für $z = 0$ ist \mathbf{A} diagonal und die Eigenvektoren bilden ein Orthogonalsystem. In diesem Fall wäre also alles außerhalb der Diagonale schwarz. Wie in der Abbildung zu

¹⁸Dieser Basiswahl ist es zu verdanken, dass die Matrix \mathbf{A} ihre bequeme Gestalt, d.h. die Form einer Bandmatrix erhält.

¹⁹Das Vorzeichen des Skalarprodukts ist bedeutungslos, da wir anstatt eines Eigenvektors \mathbf{v} ebenso gut dessen Negatives $-\mathbf{v}$ benutzen können.

²⁰Tatsächlich wurde für die Berechnung von \mathbf{V} eine Matrix der Größe 150 angesetzt. Durch den „Rand“ (d.h. dort, wo das Gleichungssystem trunziert wird) ergeben sich Verfälschungen, die sich aber nur auf Eigenmoden und -Werte mit einem Index in der Nähe der Matrixdimension auswirken. Durch die Verwendung einer größeren Matrix wird dieser Bereich vermieden.

erkennen ist, entsteht bereits für $z = 5$ ein grauer Rand an der Diagonale und für größere Werte ein immer breiter werdender heller Bereich, der die zunehmende Parallelität von (bezüglich des Index) benachbarten Eigenvektoren anzeigt. Eigenvektoren mit stark unterschiedlichen Indizes sind dagegen weiterhin orthogonal.

In den Abbildungen 9.9 sind nun die L^2 -Skalarprodukte der Eigenvektoren dargestellt. Auch hier ist die Nichtnormalität vorhanden, allerdings in einem geringeren Ausmaß. Die Nichtnormalität ist also nicht lediglich ein Artefakt unserer Rechnung, nichtsdestoweniger zeigt sich hier, dass die Basiswahl einen gewissen Einfluss hat und insofern ein Vergleich von verschiedenen Basen sinnvoll ist.

In der linken unteren Ecke in jeder der Abbildungen 9.8 und 9.9 ist ein kleiner weißer 2×2 -Block zu sehen (im symmetrischen Fall geht er für $z \neq 1$ in dem hellen Diagonalband auf). Er entsteht dadurch, dass die numerische Routine als Eigenvektoren zum doppelten Eigenwert Null zweimal die Translation ausgibt²¹, so dass das Skalarprodukt $\mathbf{v}_1 \cdot \mathbf{v}_2 = \mathbf{v}_1 \cdot \mathbf{v}_1 = 1$ ergibt. In der Fourierbasis (Abb. 9.9) ist dieser Block deutlich von dem Rest abgetrennt, da die restlichen Eigenvektoren kaum Translationsanteil aufweisen und die Projektion $\mathbf{v}_1 \cdot \mathbf{v}_j$ bezüglich des L^2 -Skalarprodukts dementsprechend gering ist. Ausnahme ist hier \mathbf{v}_3 für größere Werte von z , wo der Translationsanteil durchaus bemerkbar wird (siehe auch Bemerkung in der Diskussion der Eigenmoden in Abschnitt 9.3).

Die Abbildungen 9.10, 9.11, 9.12 und 9.13 zeigen die Pseudospektren für $z = 10$ und 25 jeweils im symmetrischen und antisymmetrischen Fall. Zunächst fällt auf, dass sich die topologische Struktur der Niveaulinien-Gebilde für die beiden z -Werte nicht wesentlich unterscheidet, dass aber die zu den Linien gehörigen ε -Werte bei $z = 25$ sehr viel kleiner sind. So teilen sich im symmetrischen Fall der Nulleigenwert und der erste negative Eigenwert für $z = 10$ ab etwa $\varepsilon \leq 10^{-3.5}$ ein gemeinsames Gebiet, während dies bei $z = 25$ bereits ab etwa $\varepsilon \leq 10^{-10}$ passiert²². Tatsächlich bestimmt die numerische Routine hier bereits den Nulleigenwert nicht mehr sehr präzise, in der Abbildung ist er in zwei beieinanderliegende Werte aufgespalten. Diese liegen ab $\varepsilon \leq 10^{10.5}$ in einem gemeinsamen Bereich, der natürlich auch den wahren Wert Null beinhaltet.

Kommen wir noch einmal kurz auf die in Abbildung 9.4 dargestellten Pseudoeigenwerte zurück, die wir durch Störung der Matrix gewonnen hatten. Grundsätzlich liegen die dort abgebildeten Punktwolken natürlich innerhalb der entsprechenden ε -Bereiche des Pseudospektrums, jedoch ist eine Umrechnung der ε -Werte notwendig. Unsere Störung war die Addition einer reellen Zufallszahl zu den (nichtverschwindenden) Matrixeinträgen, also die Addition einer Matrix mit zufälligen Einträgen zwischen 0 und ε . Die Pseudospektren entsprechen dagegen Störungen durch Addition einer Zufallsmatrix *mit der Norm* ε . Die Umrechnung ergibt sich nun folgendermaßen: Eine Matrix der Dimension N , deren Einträge

²¹Verallgemeinerte Eigenvektoren wie die Dilatation können mit dem Verfahren nicht bestimmt werden.

²²Der erste negative Eigenwert liegt für $z = 10$ bei 8.04, bei $z = 25$ bei etwa 8.5. Die Lage ist also nur geringfügig verschoben.

alle Eins sind besitzt die Norm $\|\mathbf{M}\| = N^{23}$, liegen die Einträge nun zufällig und gleichverteilt zwischen 0 und ε ist die Norm dieser Zufallsmatrix im Mittel $\frac{1}{2}\varepsilon N$, woraus sich der Umrechnungsfaktor $\frac{1}{2}N$ ergibt. Mit dieser Umrechnung liegen die Punktwolken aus Abb. 9.4 tatsächlich im entsprechenden Pseudospektrum. Die zusätzliche Struktur in 9.4, z.B. die rippenförmige Verteilung der Punkte in der komplexen Ebene, kommt daher, dass die Störungen eben nur das Band betrafen und darüber hinaus reell waren. Wir haben also, statt alle möglichen Störungen zu berücksichtigen, ein Subensemble benutzt, das besser an die eigentliche Fragestellung angepasst war und uns somit etwas detailliertere Informationen geliefert hat.

Ein weiteres wichtiges Merkmal der Niveaulinien in 9.10 - 9.13 ist es, dass die Pseudospektren im Allgemeinen in den Bereich mit positivem Realteil hineinragen. Wir hatten bereits am Anfang von Abschnitt 9.1 mit positiven Pseudoeigenwerten zu tun, wo sie uns die Untersuchung der Stabilität erschwerten hatten. Wir werden ein weiteres Mal darauf zu sprechen kommen, wenn wir uns in Abschnitt 9.4 mit dem transienten Verhalten der Zeitentwicklung beschäftigen werden.

Abschließend werfen wir noch einen Blick auf die Pseudospektren der Matrix $\tilde{\mathbf{A}}$ in den Abbildungen 9.14 und 9.15. Es ist unmittelbar ersichtlich, dass die Pseudospektren sehr stark dem normalen Fall ähneln, wie z.B. dem in Abbildung 9.7. Die Ränder der Pseudospektren sind annähernd kreisförmig und die ε -Werte, ab denen Gebiete zu den einzelnen Eigenwerten überlappen, sind sehr groß²⁴. Darüber hinaus ist der Unterschied zwischen $z = 10$ (Abb. 9.14) und $z = 25$ (Abb. 9.15) eher gering. Bei etwas genauerem Hinsehen kann man erkennen, dass die ε -Bereiche für $z = 25$ etwas größer sind und sich die Durchmesser der Bereiche um die einzelnen Eigenwerte in Abb. 9.15 etwas stärker voneinander unterscheiden. Bei normalen Matrizen haben die ε -Kreise um die einzelnen Eigenwerte alle den Radius ε . Dies ist hier nur für den Translations-Eigenwert der Fall, die anderen Bereiche sind etwas größer. Dass der Nulleigenwert eine Sonderrolle einnimmt, ist nicht weiter verwunderlich. In der Fourierbasis ist die Translationseigenmode von den anderen separiert, die Matrix $\tilde{\mathbf{A}}$ enthält also einen 1×1 -Block. Die Translation steht damit auf allen anderen Eigenmoden senkrecht, ist also in diesem Sinne „normal“.

9.3 Geometrische Form der Eigenmoden

Wie sehen diese Eigenvektoren nun aus? In den Abbildungen 9.17 und 9.18 sind die ersten sechs Eigenmoden für verschiedene Werte von z abgebildet. Die erste Zeile enthält die erste Eigenmode, d.h. unabhängig von z die Translation

²³Dies sieht man zum Beispiel dadurch, dass eine derartige Matrix die Gleichung $\mathbf{M}^2 = N\mathbf{M}$, bzw. $\mathbf{M}(\mathbf{M} - N) = 0$ erfüllt. In der Eigenbasis ist $\mathbf{M} = \text{diag}(\lambda_1, \dots, \lambda_N)$ und wir erhalten damit $\lambda_i(\lambda_i - N) = 0$. Das heißt $\lambda_i = 0$ oder N (tatsächlich ist der Nulleigenwert $N - 1$ -fach entartet) und damit $\|\mathbf{M}\| = N$.

²⁴Man vergleiche die Farbskala mit den von den Abbildungen 9.12 und 9.13.

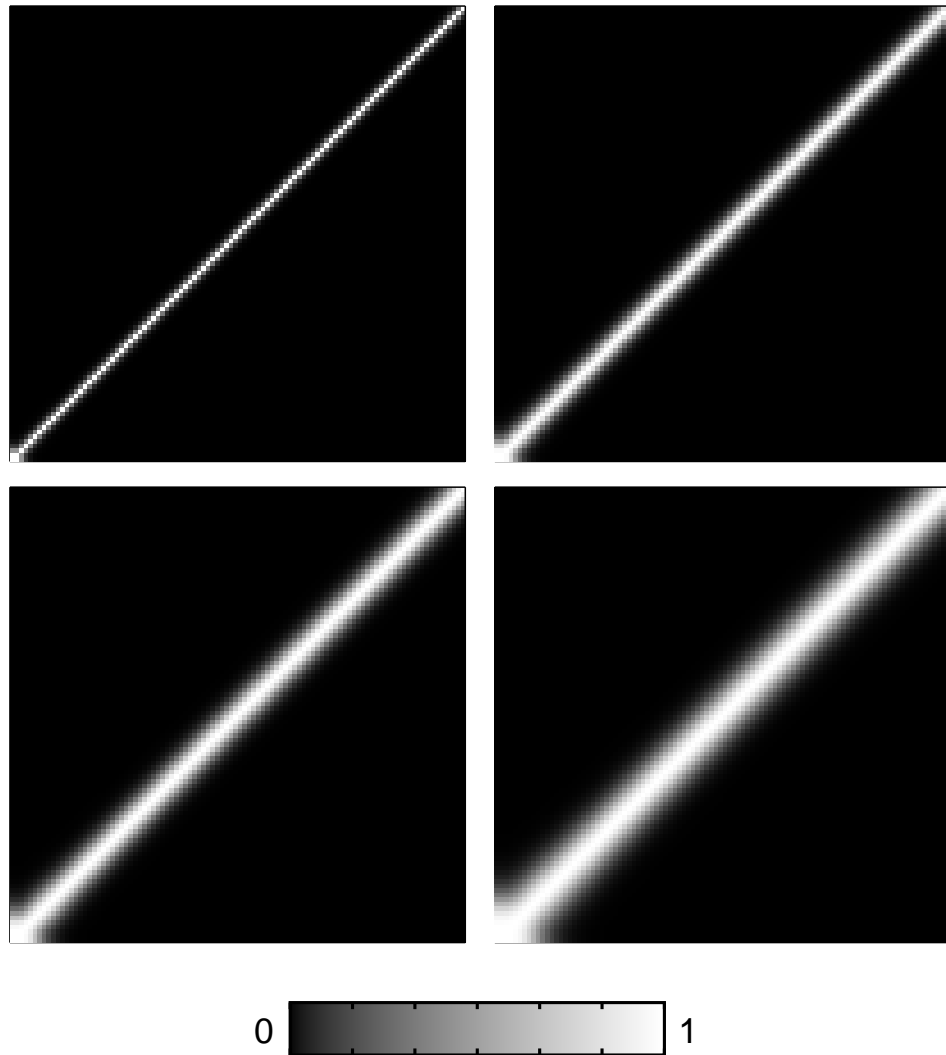


Abbildung 9.8: Skalarprodukte zwischen den Eigenvektoren für verschiedene Werte von z . Jedes quadratische Farbfeld (i, j) trägt einen Grauwert, der dem Betrag des Skalarprodukts $\mathbf{v}_i \cdot \mathbf{v}_j$ entspricht. Weiße Felder bedeuten Parallelität, schwarze dagegen Orthogonalität. Von links oben nach rechts unten ist $z = 1, 5, 10$ und 20 . Der mit wachsendem z immer breiter werdende Streifen um die Diagonale zeigt an, dass benachbarte Eigenvektoren zunehmend parallel werden.

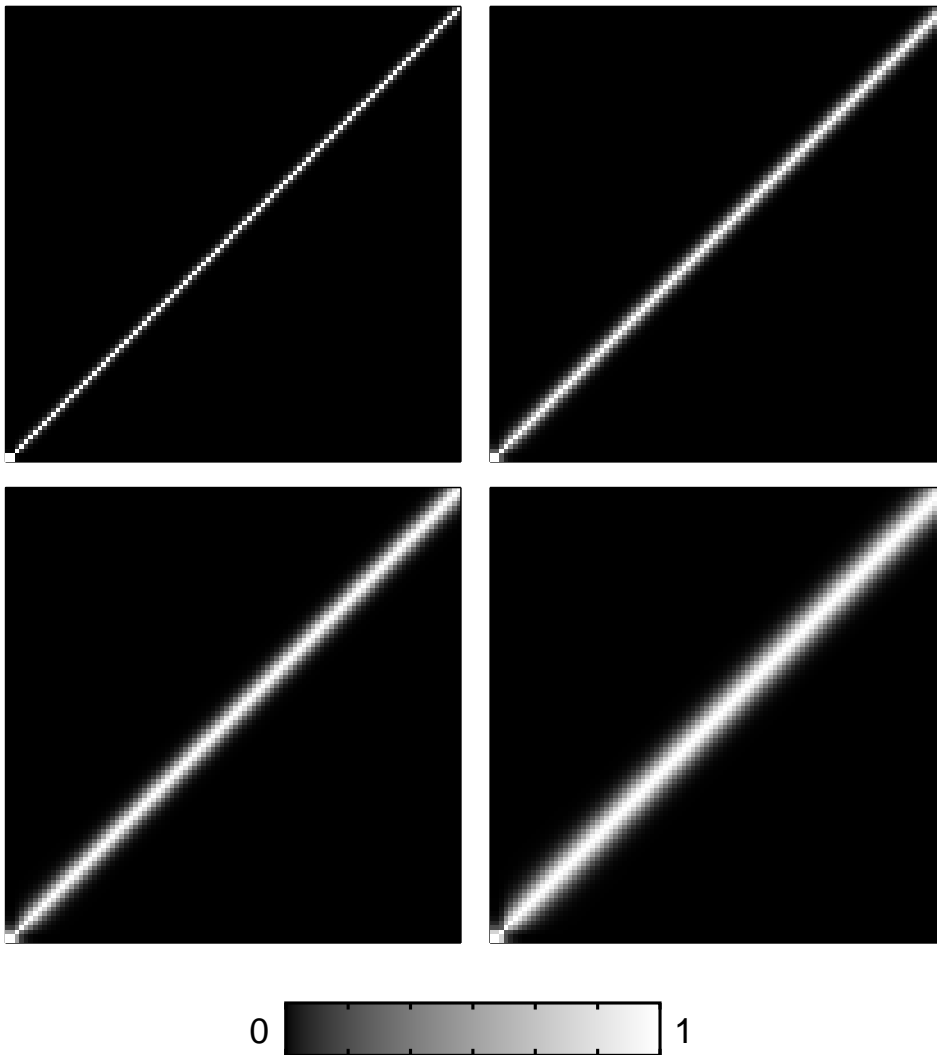


Abbildung 9.9: Skalarprodukte zwischen den Eigenvektoren der transformierten Matrix $\tilde{\mathbf{A}}$ für $z = 1, 5, 10$ und 20 .

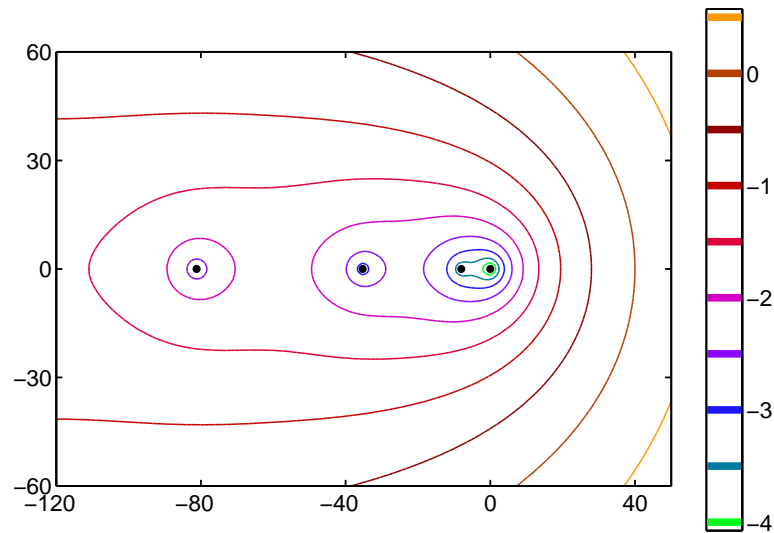


Abbildung 9.10: Pseudospektren der Matrix A (symmetrischer Fall) für $z = 10$. Die Farbskala rechts gibt $\log_{10}(\epsilon)$ an.

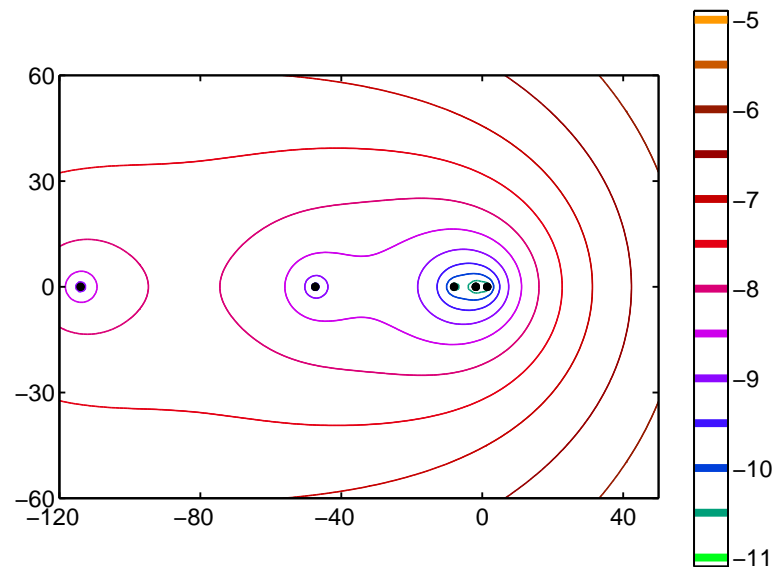


Abbildung 9.11: Pseudospektren der Matrix A für $z = 25$ (symmetrischer Fall).

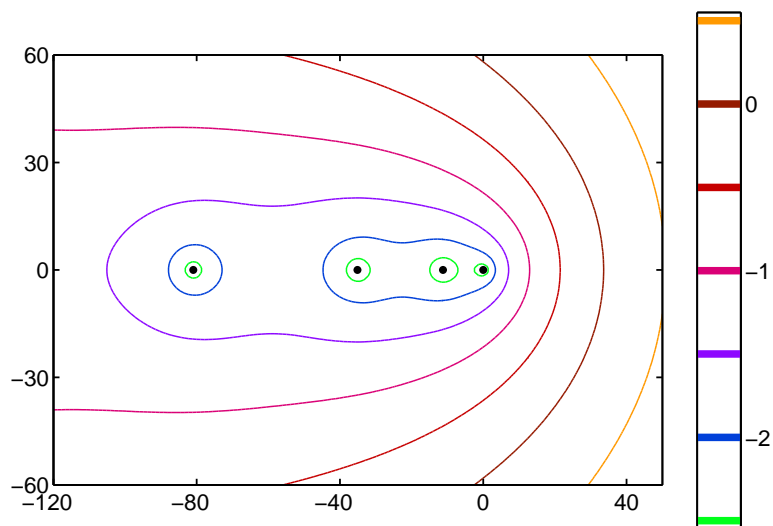


Abbildung 9.12: Pseudospektren der Matrix A (antisymmetrischer Fall) für $z = 10$. Die Farbskala rechts gibt $\log_{10}(\epsilon)$ an.

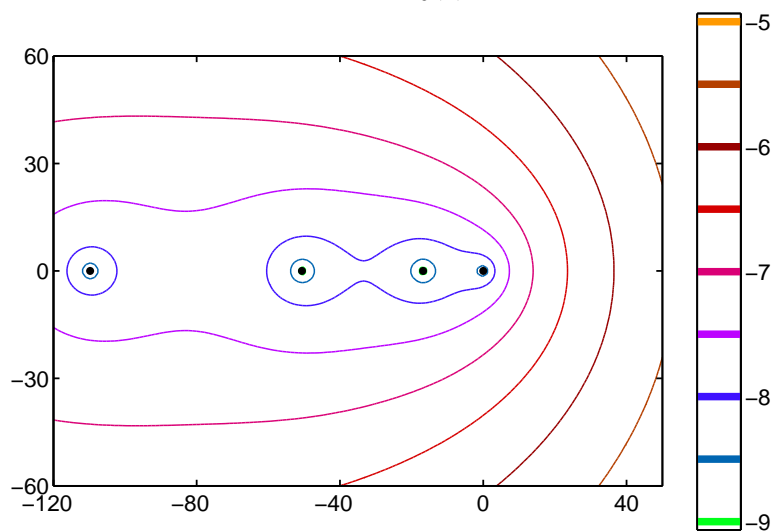


Abbildung 9.13: Pseudospektren der Matrix A für $z = 25$ (antisymmetrischer Fall).

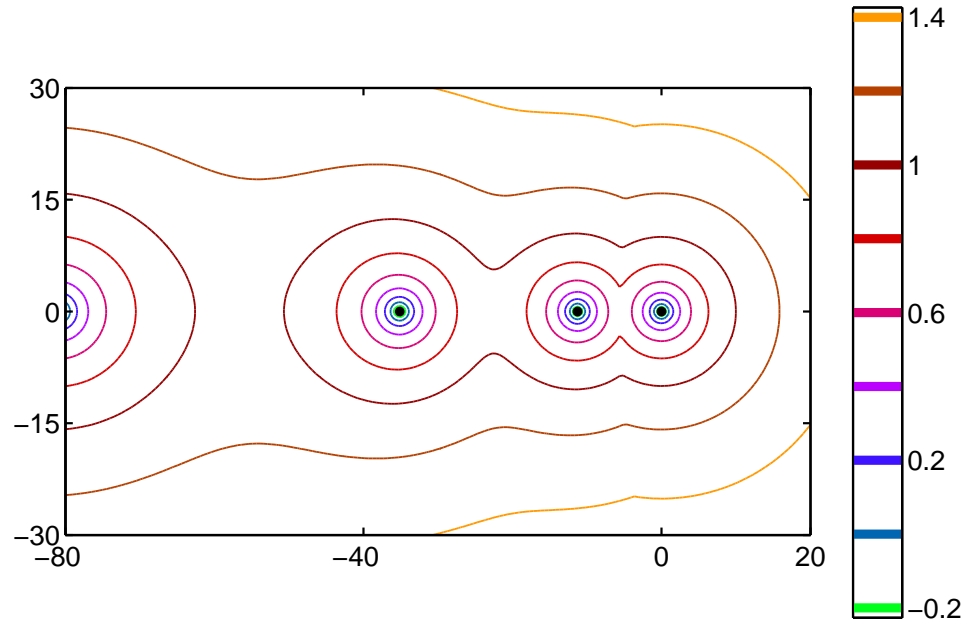


Abbildung 9.14: Pseudospektren der Matrix $\tilde{\mathbf{A}}$ (antisymmetrischer Fall) für $z = 10$.

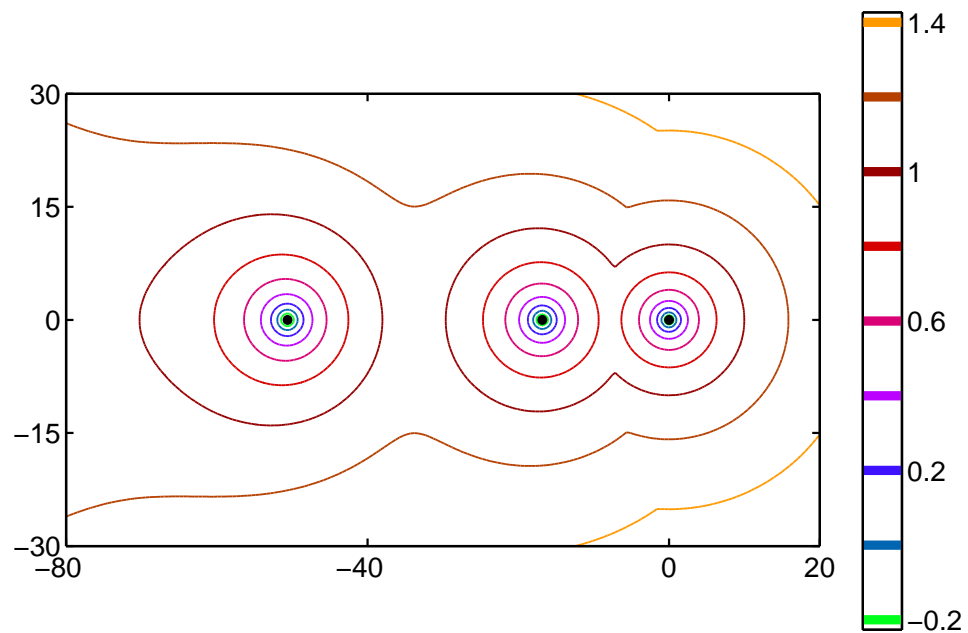


Abbildung 9.15: Pseudospektren der Matrix $\tilde{\mathbf{A}}$ für $z = 25$ (antisymmetrischer Fall).

in x (symmetrisch) bzw y -Richtung (antisymmetrisch)²⁵. Im Fall $z = 0$ ist die erste²⁶ Eigenmode eigentlich die Dilatation, d.h. die Reihenfolge ist die in Abb. 9.16 gezeigte. Die Dilatation wird für $z \neq 0$ zu einer „Pseudoeigenmode“, d.h. zu einem Hauptvektor zweiter Stufe (siehe Abschnitt 8.2), der von dem von uns verwendeten numerischen Verfahren nicht bestimmt wird.

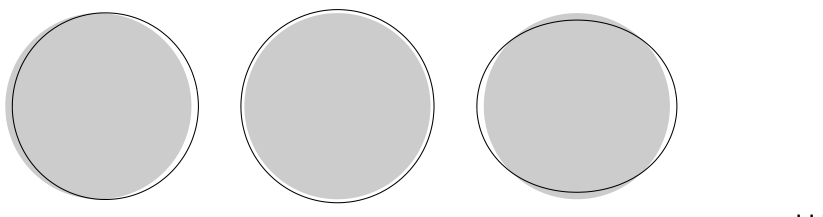


Abbildung 9.16: Die ersten drei symmetrischen Eigenmoden für $z = 0$. Hier ist neben der Translation (erste von links) auch die Dilatation (zweite von links) eine Symmetrietransformation.

Für $z = 0$ sind die Eigenmoden einfach die Fouriermoden. Die Anzahl der Knoten steigt mit dem Index an, eine Eigenschaft, die auch für $z > 0$ erhalten bleibt. Mit steigendem z ist die Modulation des Kreises nicht mehr periodisch. Insgesamt ist sie an der (im Bezug auf die Bewegungsrichtung) hinteren rechten Seite wesentlich ausgeprägter, während die vordere linke Seite für größere z -Werte praktisch kreisförmig bleibt. Diese immer stärkere Konzentration am hinteren Ende spiegelt die Wirkung des Exponentialfaktors $\exp(\frac{z}{2} \cos(\theta))$ in (8.11) wider, der für große z -Werte am hinteren Ende ($\theta = 0$) sehr groß wird, am vorderen Ende ($\theta = \pi$) die Modulation dagegen stark unterdrückt. Diese Konzentration um $\theta = 0$ ist gewissermaßen die sichtbare Folge der oben diskutierten Nichtnormalität. In dem Maße wie sich Eigenvektoren in dem N -dimensionalen Raum in eine bestimmte Richtung bündeln, zieht sich die Modulation der Kreisform auf einen immer kleiner werdenden Bereich um die hintere Spitze zusammen. Wie wir noch sehen werden ist das interessanterweise auch ein hervorstechendes Merkmal der nichtlinearen Zeitentwicklung.

Die horizontale Lage der Eigenmoden relativ zum ungestörten Kreis in Abb. 9.17 enthält eine kleine Subtilität. Am ausgeprägtesten findet man sie bei den $n=2$ -Moden in der zweiten Zeile. Es ist deutlich sichtbar, dass die Vorderseite zwar wie bei allen anderen kreisförmig wird, jedoch gegenüber dem Kreis ein wenig verschoben ist. Der Grund dafür ist folgender: Da sich die gestörten Formen mit einer etwas anderen Geschwindigkeit als der Kreis bewegen, enthält

²⁵Die Tatsache, dass wir mit den Translationen jeweils eine Eigenmode exakt kennen, lässt sich auch hier ausnutzen, um die Güte des numerischen Verfahrens zu überprüfen

²⁶Die Reihenfolge von Dilatation und Translation kann nicht durch den zugehörigen Eigenwert festgelegt werden, da beide zum Eigenwert Null gehören. In unserem Fall nehmen wir die Dilatation zuerst, da sie keine Knoten besitzt und in allen anderen Fällen die Knotenzahl mit dem Index korrespondiert.

jede Eigenmode im Prinzip einen kleinen Translationsanteil, der die Beschleunigung beschreibt, die mit der Störung einhergeht. In dem genannten $n=2$ -Fall ist dies eine *Reduzierung* der Geschwindigkeit, obwohl das Bild suggeriert, daß die gestörte Form dem Kreis vorausseilt (es sei daran erinnert, daß sich die Insel nach links bewegt). Dies ist folgendermaßen einzusehen: Die Eigenmode gehört zu einem negativen Eigenwert, wird also mit der Zeit wieder zum Kreis zusammenschrumpfen. Da der Schwerpunkt der gestörten Form im Bild vor dem Kreis liegt, heißt das, dass der Kreis die gestörte Form einholt und sich folglich schneller bewegt. Dieser Effekt der Geschwindigkeitsänderung ist im Prinzip auch in allen der anderen dargestellten Fälle vorhanden, ist jedoch weitaus weniger ausgeprägt. Bei den antisymmetrischen Moden tritt dieses Phänomen dagegen nicht auf, die Moden können aufgrund ihrer Symmetrie keine horizontale Translation enthalten. Zum Schluss sei noch erwähnt, dass es zu jeder der dargestellten Störungen ein Gegenstück gibt, bei der die Modulation ein anderes Vorzeichen besitzt, das heißt, dass aus jeder Aus- eine Einbuchtung wird und umgekehrt²⁷. Für die antisymmetrischen Moden bedeutet diese Inversion lediglich eine Reflexion an der x -Achse.

9.4 Transientes Verhalten

Eine weitere Folge der Nichtnormalität einer Matrix \mathbf{M} ist die Auswirkung auf das *transiente Verhalten* der Zeitentwicklung $\exp(t\mathbf{M}) \cdot \mathbf{x}_0$ eines Anfangszustandes \mathbf{x}_0 . Während das asymptotische Verhalten in jedem Fall durch das Spektrum bestimmt ist²⁸, kann die Nichtnormalität einen anderen Verlauf erzeugen, als durch die Eigenwerte nahegelegt wird. Der zugrunde liegende Mechanismus ist in Abbildung 9.19 dargestellt. Die Anfangsbedingung ist hier die Überlagerung zweier Eigenvektoren \mathbf{v}_1 und \mathbf{v}_2 . Deren Entwicklung ist einfach $\exp(\lambda_1 t)\mathbf{v}_1$ bzw. $\exp(\lambda_2 t)\mathbf{v}_2$, wobei $\lambda_{1,2}$ die zugehörigen Eigenwerte sind, von denen wir $0 > \lambda_1 > \lambda_2$ annehmen. Während im Falle zweier orthogonaler Eigenvektoren dieses gleichzeitige Schrumpfen von \mathbf{v}_1 und \mathbf{v}_2 zu einer monotonen Abnahme der Überlagerung führt, ist es im nichtnormalen Fall mit der dargestellten Anfangsbedingung möglich, dass, obwohl beide Vektoren kleiner werden, die Überlagerung anwächst. Dieses Phänomen ist besonders ausgeprägt, wenn die anfängliche Überlagerung, wie in der Zeichnung nahezu senkrecht auf den Eigenvektoren steht. Die Zeitentwicklung wird dann wie in der Abbildung dazu tendieren, den Vektor in Richtung der Eigenvektoren zu drehen. Es ist anschaulich klar, dass das transiente Anwachsen umso stärker sein wird, je parallel die Eigenvektoren sind und je stärker sich die Zerfallsraten λ_1 und λ_2 unterscheiden. Rechnet man in dem genannten einfachen Beispiel von zwei Eigenvektoren

²⁷Hier kehrt sich auch der eben beschriebene Effekt der Geschwindigkeitsänderung um.

²⁸D.h. wenn wie in unserem Fall alle Eigenwerte negativ sind, muss jede anfängliche Störung auf lange Sicht gegen Null gehen.

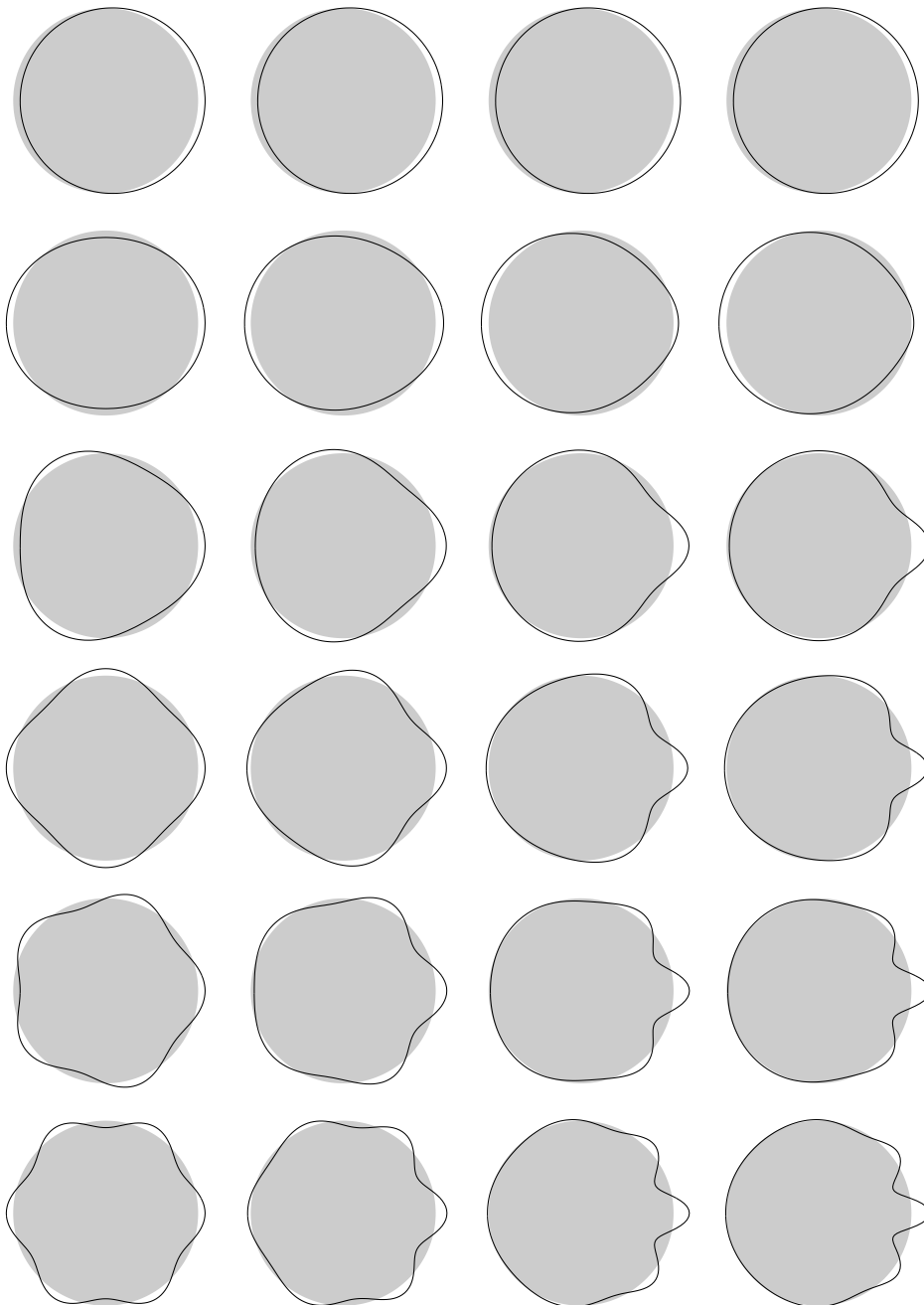


Abbildung 9.17: Eigenmoden im symmetrischen Fall. In jeder Spalte sind jeweils von oben nach unten die ersten sechs Eigenmoden abgebildet (die Dilatation für $z = 0$ wurde ausgelassen. Die zu den Spalten gehörigen z -Werte sind 0, 5, 15 und 25 (von links nach rechts). Die Bewegungsrichtung der Inseln ist von rechts nach links.

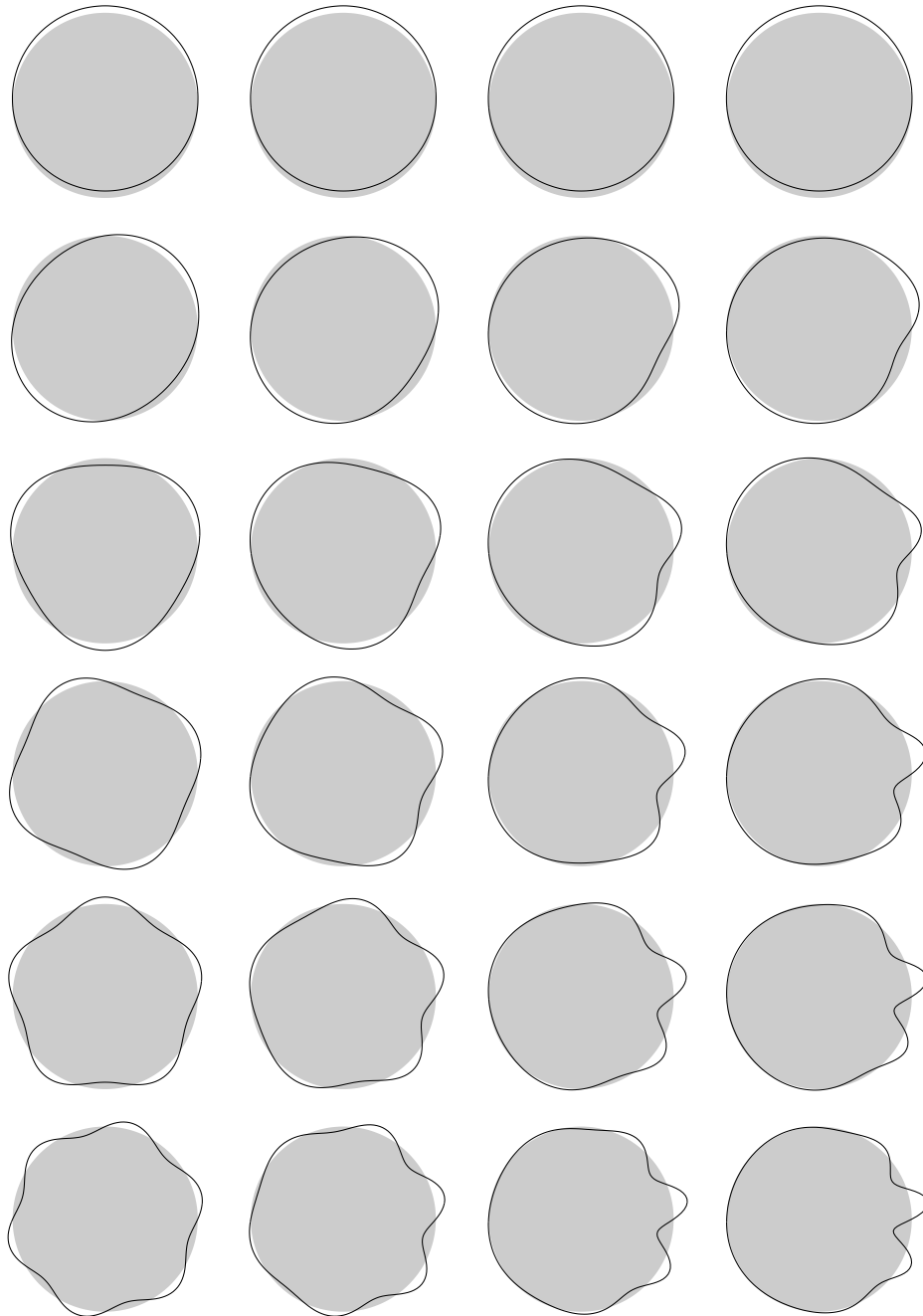


Abbildung 9.18: Die ersten sechs Eigenmoden im asymmetrischen Fall. Die zu den Spalten gehörigen z -Werte sind 0, 5, 15 und 25 (von links nach rechts). Die Bewegungsrichtung der Inseln ist von rechts nach links.

das Betragsquadrat der Überlagerung aus und berechnet die Zeitableitung bei $t = 0$, so ergibt sich

$$\|e^{\lambda_1 t} \mathbf{v}_1 + e^{\lambda_2 t} \mathbf{v}_2\|^2 = e^{2\lambda_1 t} \|\mathbf{v}_1\|^2 + e^{2\lambda_2 t} \|\mathbf{v}_2\|^2 + 2e^{(\lambda_1 + \lambda_2)t} \mathbf{v}_1 \cdot \mathbf{v}_2$$

Die beiden ersten Terme auf der rechten Seite sind positiv und zeitlich monoton fallend. Der dritte Term zerfällt zwar ebenfalls mit der Zeit, aber da das Skalarprodukt $\mathbf{v}_1 \cdot \mathbf{v}_2$ negativ sein kann (wie in der Zeichnung), kann eben genau dieses Zerfallen zu einem Anstieg der Summe führen. Es kommt also nicht nur auf den Wert von $\mathbf{v}_1 \cdot \mathbf{v}_2$ an, d.h. auf den Grad der Nichtorthogonalität, sondern auch auf das Verhältnis der Zerfallsraten. Das Zusammenspiel aus beiden ist es, was transientes Wachstum erzeugen kann.

Die Idee, dass Nichtnormalität die Ursache für die Instabilität eines Systems sein könnte, wurde bereits seit längerem in der Hydrodynamik diskutiert. Ein zentrales Thema in diesem Gebiet ist das sogenannte Problem der hydrodynamischen Stabilität [13]. Dabei handelt es sich um das Phänomen, dass ein anfänglich laminares Strömungsprofil bei Änderung der Systemparameter instabil wird und den Übergang zu einem turbulenten Zustand vollzieht. Die Dynamik der Flüssigkeit wird dabei durch die Navier-Stokes-Gleichung beschrieben und die bestimmende Kenngröße ist hier die sogenannte Reynolds-Zahl Re , die im Wesentlichen das Verhältnis von Trägheit (d.h. Dichte und Geschwindigkeit) zu Viskosität beschreibt. Anders ausgedrückt gibt sie die relative Stärke von Konvektion und Diffusion innerhalb der Flüssigkeit wieder und hat daher eine gewisse Ähnlichkeit mit unserem Parameter z . Der Übergang von laminarer zu turbulenter Strömung vollzieht sich im Allgemeinen, wenn die Reynoldszahl einen bestimmten kritischen Schwellwert Re_c überschreitet²⁹. Für einige Fälle, insbesondere diejenigen, die durch thermische Einflüsse (Rayleigh-Bénard-Konvektion³⁰) oder Zentrifugalkräfte (Taylor-Couette-Strömung³¹) getrieben werden, lässt sich der kritische Wert Re_c durch eine lineare Stabilitätsanalyse bestimmen. Für andere Strömungen versagt dieser Zugang allerdings. Dabei handelt es sich insbesondere um die ebene Poiseuille-Strömung, bei der die Flüssigkeit mittels eines Druckgradienten durch den Raum zwischen zwei (unendlich ausgedehnten) Platten gepresst wird und um die ebene Couette-Strömung zwischen zwei Platten, die sich mit einer konstanten Geschwindigkeit gegeneinander bewegen. Im ersten Fall ergeben Experimente einen Übergang bei einer wesentlich geringeren Reynoldszahl, als die lineare Stabilitätsanalyse vorhersagt, während sich im zweiten Fall Turbulenz entwickelt, obwohl das System für beliebig hohe Reynoldszahlen linear stabil ist.

²⁹Wenn z.B. die Strömungsgeschwindigkeit hohe Werte annimmt.

³⁰Die Ausbildung von Konvektionsrollen zwischen zwei horizontalen Platten verschiedener Temperatur.

³¹Scherströmung zwischen zwei Zylinderflächen, die sich mit verschiedener Winkelgeschwindigkeit um eine gemeinsame Achse drehen.

In den neunziger Jahren begannen einige Autoren die Idee zu propagieren, dass sich diese „Instabilität trotz linearer Stabilität“ aus der Nichtnormalität der mit diesen Systemen verbundenen Eigenmoden erklären ließe [63, 16, 20]. Die Vorstellung dabei ist, dass eine anfängliche Störung, die nahezu senkrecht auf den Eigenmoden steht, zunächst durch den oben dargestellten Mechanismus anwächst, wobei sie in Richtung der Eigenmoden gedreht wird. Ab einer bestimmten Amplitude setzt nun die Nichtlinearität ein, die dazu tendiert, die Störung wieder in die senkrechte Richtung zurückzudrehen. Dadurch ergibt sich eine Art Rückkopplungs-Mechanismus aus linearem Anwachsen und nichtlinearem „Zurückdrehen“, der das System instabil werden lässt. Da sich dieser Ablauf weder im Experiment noch anhand der Navier-Stokes-Gleichung isoliert nachweisen lässt, wurden in [2] und [16] einfache nichtlineare Modelle vorgeschlagen, bei denen dieser Mechanismus realisiert ist. Inwieweit sich dies auf den hydrodynamischen Kontext übertragen lässt, ist jedoch etwas unklar. In [67] wurde eine detaillierte Kritik formuliert, die insbesondere herausarbeitet, dass diese vereinfachten dynamischen Gleichungen einige grundlegende physikalische Eigenschaften der Navier-Stokes-Gleichung verletzen und daher als Modellsysteme für die Hydrodynamik von fraglichem Wert sind. Darüberhinaus wird in diesen Modellen die Rolle der Nichtlinearität mehr oder weniger darauf beschränkt, die anwachsende Störung kontinuierlich in die senkrechte Richtung zurückzudrehen, ohne dass dabei der spezifischen Form und Wirkung der Navier-Stokes-Nichtlinearität Rechnung getragen wird.

Während also die Rolle der Nichtnormalität im hydrodynamischen Zusammenhang strittig ist, weisen die Ergebnisse aus [52] deutlicher auf einen Zusammenhang mit transientem Anwachsen hin. Die Untersuchung betrifft hier die Instabilität einer kreisförmigen Pore³² unter Elektromigrationseinfluss, die für große Radien in kleinere Teile zerbricht. Eine lineare Stabilitätsanalyse zeigt, dass die Kreisform für beliebige Radien linear stabil ist. Die damit verbundenen Eigenmoden stellen sich allerdings als nicht-orthogonal heraus. In Abbildung 9.20 ist der Umfang der Poren als Funktion der Zeit für verschiedene Porenradien in der linearisierten und nichtlinearen Zeitentwicklung dargestellt. Wie zu sehen ist, kommt es sowohl in der linearen als auch in der nichtlinearen Entwicklung zu einem beträchtlichen Ansteigen des Umfangs eines leicht gestörten Kreises. Für kleinere Radien zieht sich die Form schließlich wieder zum Kreis zusammen, während die größeren Poren zerbrechen. Vor dem Hintergrund des oben Gesagten suggeriert dieses Bild natürlich, dass die Nichtnormalität das System durch transientes Anwachsen aus dem linearen Bereich hinaustreibt, wo die nichtlineare Entwicklung schließlich zum Zerbrechen der Poren führt.

Wenden wir uns nun wieder der linearen Entwicklung der Inseln zu. Um Aussagen über das transiente Anwachsen von Störungen zu treffen, benötigen wir ein Maß

³²Das Modell für die Elektromigration von Poren hat starke Ähnlichkeit mit dem im ersten Teil beschriebenen lokalen Modell. Allerdings wird dort das Feld am Inselrand durch die Lösung der Laplace-Gleichung im Außenraum bestimmt.

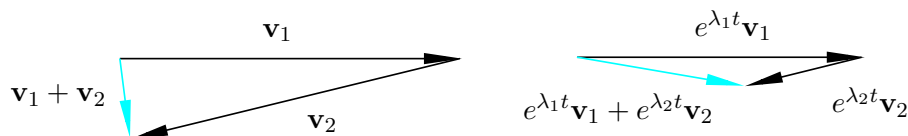


Abbildung 9.19: Mechanismus des transienten Wachstums bei nichtorthogonalen Eigenvektoren. Die Anfangsbedingung links besteht aus zwei Eigenvektoren, die mit verschiedener Geschwindigkeit kleiner werden, so dass die Überlagerung nach der Zeit t (rechts) *größer* geworden ist.

für die „Größe“ dieser Störung. Mathematisch bedeutet das eine Bezugnahme auf eine bestimmte Norm des der Störung entsprechenden Vektors und deren zeitlichen Verlauf. Wir haben bereits angemerkt, dass es für unser Problem nicht von vorneherein eine ausgezeichnete Norm gibt. Wir kommen nun auf diese fehlende Eindeutigkeit zurück, indem wir das transiente Verhalten einmal bezüglich des $\langle \cdot | \cdot \rangle_{\text{exp}}$ -Skalarprodukts und einmal bezüglich des L^2 -Skalarprodukts behandeln. Das heißt, wir werden einmal die Matrix \mathbf{A} und einmal die Matrix $\tilde{\mathbf{A}}$ behandeln. Es wird sich herausstellen, dass die „naheliegende“ Möglichkeit, d.h. die Norm von \mathbf{A} zu berechnen hier zu einer deutlichen Verzerrung führt. Tatsächlich wird die Norm durchaus sehr große Werte annehmen, während dieser Effekt für $\tilde{\mathbf{A}}$ vollständig verschwindet. Ziel dieses Vergleichs ist es, noch einmal ausdrücklich den Einfluß der Basiswahl (die in unserem Fall auch die Norm bestimmt) hervorzuheben.

Anstatt die Norm von einzelnen Vektoren zeitlich zu verfolgen, greifen wir wieder auf den am Anfang von Abschnitt 9.2 eingeführten Begriff der Matrixnorm zurück. Da sie, wie bereits erwähnt, dem Maximum von $\|\mathbf{M}\mathbf{x}\|$ auf der Einheitskugel $\|\mathbf{x}\|$ entspricht, gibt sie uns ein geeignetes Maß dafür, wie groß $\|\mathbf{M}\mathbf{x}\|$ im ungünstigsten Fall werden kann. Bilden wir die Matrixnorm $\|\exp(t\mathbf{M})\|$ der Zeitentwicklungsmatrix, so erhalten wir eine obere Schranke für den zeitlichen Verlauf einer beliebigen Anfangsbedingung.

Im Folgenden sind wir also am Verhalten von $\|\exp(t\mathbf{A})\|$ als Funktion von t interessiert. Hierbei ergibt sich für den symmetrischen Fall durch den doppelten Nulleigenwert eine Besonderheit: Wählen wir als Basis die Translation und die Dilatation, so spannen diese einen invarianten Unterraum auf und die auf diesen Raum eingeschränkte Matrix ist durch

$$\mathbf{A}_s = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

gegeben³³. Exponentieren liefert $\exp(t\mathbf{A}_s) = \sum \frac{1}{n!} (t\mathbf{A}_s)^n$, wobei \mathbf{A}_s^n für $n > 1$ Null ist. Wir erhalten also

³³Die Eins in der rechten oberen Ecke entsteht, wenn \mathcal{D} und \mathcal{T}_x relativ zueinander so skaliert werden, dass $\hat{\mathbf{A}}\mathcal{D} = \mathcal{T}_x$. Wir vermeiden auf diese Weise nur unnötige zusätzliche Faktoren.

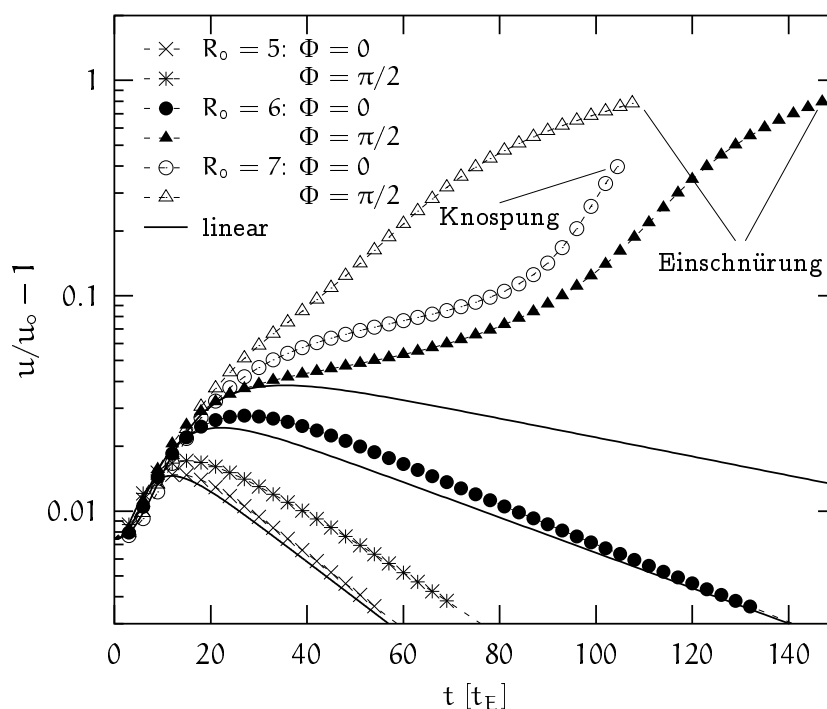


Abbildung 9.20: Zeitliche Entwicklung des Umfangs u einer Pore unter Elektromigrationseinfluss für verschiedene Porenradien R_0 (aus [52]). Dabei ist u_0 der Umfang eines Kreises mit derselben Fläche. Die durchgezogenen Linien kennzeichnen den Verlauf für die lineare Theorie, die Symbole den Verlauf der nicht-linearen Entwicklung. Die für größere Radien einsetzende Instabilität führt zum Zerbrechen der Insel (in der Zeichnung mit „Knospung“ und „Einschnürung“ gekennzeichnet). Die Anfangskonfiguration ist ein in Feldrichtung ($\Phi = 0$), bzw. senkrecht dazu ($\Phi = \pi/2$) gestreckter Kreis.

$$\exp(t\mathbf{A}_s) = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix},$$

so dass die Norm dieser Matrix asymptotisch $\sim t$ ist^{34 35}. Da die Gesamtma-

³⁴Der exakte Ausdruck für die Norm ist

$$\sqrt{\max(1 + \frac{1}{2}t^2 + \frac{1}{2}\sqrt{4t^2 + t^4}, 1 + \frac{1}{2}t^2 - \frac{1}{2}\sqrt{4t^2 + t^4})} \text{ (Maple)}$$

³⁵Zur physikalischen Interpretation: Der Vektor $(1, 0)$, d.h. die Translation wird durch die Matrix auf sich selbst abgebildet (Translationssymmetrie). Die Dilatation $(0, 1)$ wird dagegen auf eine Überlagerung der Dilatation und einer linear anwachsenden Translation abgebildet. Dies spiegelt die Tatsache wieder, dass eine infinitesimal vergrößerte Insel eine Beschleunigung erfährt und die ungestörte und gestörte Form daraufhin linear in der Zeit auseinanderdriften.

trix in die direkte Summe aus \mathbf{A}_s und dem (ebenfalls reduzierten) Rest zerfällt, sorgt dieser Term immer für einen (asymptotisch) linearen Anstieg, auch wenn die Eigenwerte, wie in unserem Fall, negativ sind³⁶. Aus diesem Grund werden wir uns zuerst mit den antisymmetrischen Eigenmoden beschäftigen, wo keine doppelten Nulleigenwerte auftreten.

9.4.1 Asymmetrischer Fall

Überlegen wir zunächst, was bei einer normalen Matrix mit einem Spektrum der Form $\{0, -\lambda_1, -\lambda_2, \dots\}$ für den Verlauf von $\|\exp(t\mathbf{M})\|$ zu erwarten wäre. Das Spektrum von $\exp(t\mathbf{M})$ ist $\{1, \exp(-\lambda_1 t), \exp(-\lambda_2 t), \dots\}$ und die Norm dementsprechend konstant 1. In Abb. (9.21) ist der Verlauf von $\|\exp(t\mathbf{A})\|$ als Funktion von t dargestellt³⁷. Wie zu erwarten, streben die Kurven mit wachsendem t gegen einen konstanten Wert, der allerdings mit steigendem z -Wert (exponentiell) groß wird. Bei $t = 0$ ist die Exponentialmatrix einfach die Identität und die Norm deshalb Eins. Die numerisch ermittelten Kurven werden für $z > 16$ immer unregelmäßiger und schlagen ab $z \simeq 19$ in einen exponentiellen Anstieg um. Da die Numerik in diesem Bereich offensichtlich keine zuverlässigen Ergebnisse mehr liefert, werden wir diese Fälle im Folgenden nicht weiter berücksichtigen.

Berechnen wir nun dieselben Kurven für die Matrix $\tilde{\mathbf{A}}$, d.h. die Norm in der Fourierbasis, so ist $\|\exp(t\tilde{\mathbf{A}})\| \equiv 1$, völlig unabhängig von z . Das ist das Ergebnis, das wir für normale Matrizen erwarten würden. Darüber hinaus steht uns in der Fourierbasis auch die Möglichkeit zur Verfügung, den Translationsanteil der Störungen zu eliminieren. Das heißt, da die (in diesem Fall vertikale) Lage der gestörten Insel für unsere Fragestellung keine Rolle spielt, können wir uns auf den Unterraum einschränken, der die Translationen nicht enthält. Für die Matrix $\tilde{\mathbf{A}}$ bedeutet das lediglich eine Streichung von erster Spalte und Zeile, wodurch wir die reduzierte Matrix $\tilde{\mathbf{A}}_{\text{red}}$ erhalten³⁸. Der Verlauf von $\|\exp(t\tilde{\mathbf{A}}_{\text{red}})\|$ ist in Abbildung 9.22 dargestellt. Für alle z -Werte ergeben sich exponentiell abfallende Kurven und die Abklingkonstante entspricht genau dem ersten negativen Eigenwert λ_1 . Würden wir statt der Matrixnorm die Amplitude einer beliebigen Anfangsbedingung zeitlich verfolgen, würde sich ein ähnliches exponentielles Abfallen ergeben. Die L^2 -Norm gibt also getreu das „wirkliche“ Verhalten wieder, während die $\langle | \rangle_{\text{exp}}$ -Norm offensichtlich ein verzerrtes Bild erzeugt. Was ist die

³⁶Die Matrixnorm einer direkten Summe ist einfach $\|\mathbf{M}_1 \oplus \mathbf{M}_2\| = \max(\|\mathbf{M}_1\|, \|\mathbf{M}_2\|)$

³⁷Technische Anmerkung: Die Matrizen mit $1 \leq z \leq 9$ haben Dimension 100, die mit $10 \leq z \leq 20$ Dimension 200. Die kleinere Matrixdimension für $z \leq 9$ wurde gewählt, weil die in den Matrix-Einträgen enthaltenen Brüche $\frac{I_{n+1} + I_{n-1}}{I_n}$ von Besselfunktionen für große n und kleine z nicht zuverlässig numerisch bestimmt werden können. Das Problem dabei ist, dass in diesem Fall die Funktionswerte sehr klein werden (für $z \ll n$ ist $I_n(z) \approx \frac{1}{n!} \left(\frac{z}{2}\right)^n$), so dass die numerische Routine den Funktionswert Null ausgibt. Der Bruch enthält damit eine Division durch Null.

³⁸Eine ähnliche Einschränkung ist bei \mathbf{A} nicht ohne weiteres möglich.

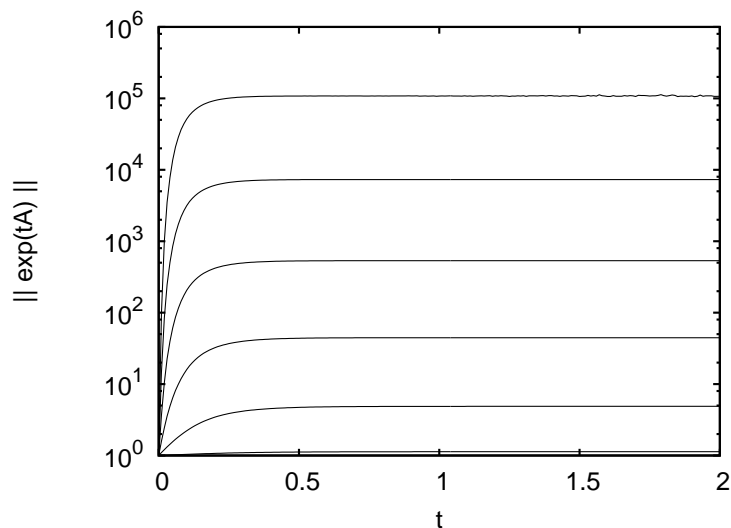


Abbildung 9.21: Verlauf von $\|\exp(t\mathbf{A})\|$ als Funktion von t für $z = 1, 4, 7 \dots 16$ (entspricht den Kurven von unten nach oben). Der Auftrag ist halblogarithmisch. In den leichten Unregelmäßigkeiten in der obersten Kurve ($z = 16$) deutet sich bereits an, dass die numerischen Ergebnisse für $z > 16$ zunehmend unbrauchbar werden.

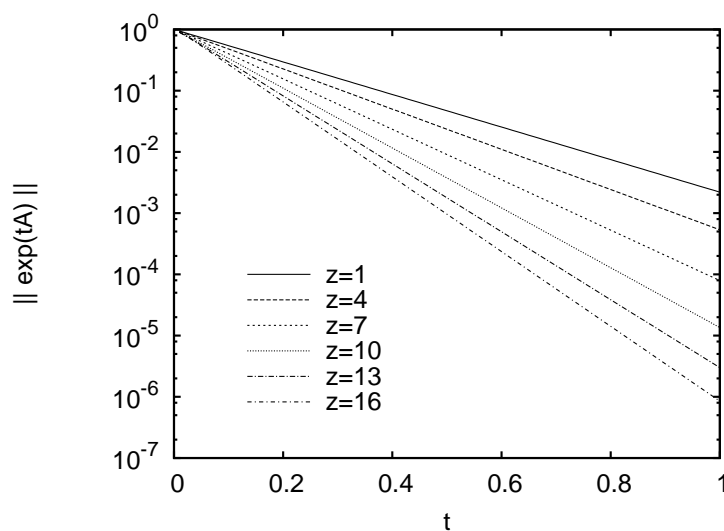


Abbildung 9.22: Verlauf von $\|\exp(t\tilde{\mathbf{A}}_{\text{red}})\|$ als Funktion von t für $z = 1, 4, 7 \dots 16$. Der Auftrag ist halblogarithmisch.

Ursache dafür? Um diese Frage zu klären, sei noch einmal daran erinnert, dass das Skalarprodukt einen exponentiellen Gewichtungsfaktor enthält. Es entspricht also der Integraldarstellung

$$\langle f_1(\theta) | f_2(\theta) \rangle_{\text{exp}} = \int_0^{2\pi} f_1(\theta) f_2(\theta) \exp\left(\frac{z}{2} \cos(\theta)\right) d\theta$$

Je größer z ist, desto stärker wird der Exponentialfaktor das Produkt der Funktionswerte um $\theta = 0$ herum gewichten, während alle anderen Werte stark unterdrückt werden. Zwei Funktionen, die sich bei $\theta = 0$ auch nur geringfügig unterscheiden, werden demnach in dieser Norm einen großen Abstand haben, wobei der Abstand exponentiell mit z wächst. Dies ist nun genau der exponentiell größer werdende Sättigungswert in Abb. (9.21). Der Sättigungswert selbst ist auf die Translationseigenmode zurückzuführen, die einzige Mode, die nicht zerfällt, weswegen die Kurve gegen einen konstanten Wert strebt.

Auch wenn es sich bei dem Anwachsen also um ein Artefakt handelt, werden wir im Folgenden dieses Verhalten mit den pseudospektralen Eigenschaften der Matrix in Verbindung bringen³⁹. Die Grundidee ist dabei, dass, wenn die Pseudoeigenwerte genügend weit in der Halbebene mit positivem Realteil liegen, ein transientes Anwachsen zu erwarten ist. Um zu präzisieren, was „genügend weit“ bedeutet, müssen wir einen Zusammenhang zwischen der Exponentialmatrix $\exp(t\mathbf{M})$ und dem Pseudospektrum, bzw. der Resolvente herstellen. Der Schlüssel dazu liegt in der folgenden Beziehung⁴⁰ für die Laplace-Transformierte von $e^{t\mathbf{M}}$

$$(x - \mathbf{M})^{-1} = \int_0^{\infty} e^{-xt} e^{t\mathbf{M}} dt.$$

Nehmen wir auf beiden Seiten die Norm, so erhalten wir die Abschätzung

$$\begin{aligned} \|(x - \mathbf{M})^{-1}\| &\leq \int_0^{\infty} \|e^{t\mathbf{M}}\| \cdot |e^{-xt}| dt \\ &\leq \left(\sup_{t \geq 0} \|e^{t\mathbf{M}}\| \right) \int_0^{\infty} |e^{-xt}| dt = \frac{\sup_{t \geq 0} \|e^{t\mathbf{M}}\|}{\operatorname{Re} x}. \end{aligned}$$

Transientes Anwachsen bedeutet, dass $\sup_{t \geq 0} \|e^{t\mathbf{M}}\|$ größer als Eins ist. Wir können also den folgenden Satz formulieren

Hilfssatz. *Erfüllt die Resolvente $(x - \mathbf{M})^{-1}$ einer Matrix die Ungleichung $\|(x_0 - \mathbf{M})^{-1}\| \geq K / \operatorname{Re} x_0$ für ein x_0 mit $\operatorname{Re} x_0 > 0$ und $K > 1$, so ergibt sich für das transiente Anwachsen die untere Schranke*

³⁹Alle nachfolgenden Ergebnisse stammen aus [64]. Wir werden auch hier so verfahren, dass wir die Herleitung jeweils nur skizzieren.

⁴⁰Für einen Beweis siehe z.B. [43]

$$\sup_{t \geq 0} \|e^{t\mathbf{M}}\| \geq K$$

Erinnern wir uns nun daran, dass das ε -Pseudospektrum aus den Punkten mit $\|(x - \mathbf{M})^{-1}\| \geq \frac{1}{\varepsilon}$ besteht, und definieren die *pseudospektrale Abszisse* als das Supremum des Realteils des ε -Pseudospektrums

$$\alpha_\varepsilon(\mathbf{M}) = \sup_{x \in \sigma_\varepsilon(\mathbf{M})} \operatorname{Re} x,$$

so erhalten wir das zentrale Ergebnis

$$\sup_{t \geq 0} \|e^{tA}\| \geq \alpha_\varepsilon / \varepsilon. \quad (9.4)$$

Stellen wir nun kurz einen Vergleich mit dem Fall normaler Matrizen an. Wir hatten bereits in Abschnitt 9.2 angemerkt, dass in diesem Fall der Betrag der Resolvente mit dem Abstand vom Spektrum abfällt

$$\|(x - \mathbf{M})^{-1}\| = \frac{1}{\operatorname{dist}(x, \sigma(\mathbf{M}))}.$$

Nehmen wir nun an, der größte Eigenwert sei $\lambda_0 = 0$ (wie es bei uns der Fall ist), die Matrix liege also genau an der Grenze zwischen Abfallen und Anwachsen von $\|\exp(t\mathbf{M})\|$. Der Betrag der Resolvente fällt also mit

$$\|(x - \mathbf{M})^{-1}\| = \frac{1}{|x|}$$

in der Halbebene $\operatorname{Re} x > 0$ ab. Die Pseudospektren ragen somit genau um den Abstand ε in diese Halbebene hinein und die pseudospektrale Abszisse ist demzufolge $\alpha_\varepsilon = \varepsilon$. In (9.4) ist die rechte Seite dann Eins und die Abschätzung trivial. In diesem Sinne bedeutet die Konstante $K > 1$ im Hilfssatz gerade, dass die Norm der Resolvente langsamer abfällt als im normalen Fall und die Pseudospektren dadurch „weit genug“ in die $\operatorname{Re} x > 0$ -Halbebene hineinreichen. Vergleichen wir nun die theoretische untere Schranke mit dem tatsächlichen Anwachsen, d.h. mit dem Sättigungswert in Abbildung 9.21. Die folgenden beiden Tabellen enthalten α_ε für verschiedene Werte von ε für die Fälle $z = 10$ und $z = 15$. Für $z = 10$ liegt das Maximum von $\alpha_\varepsilon / \varepsilon$ in der Nähe von 5.36×10^2 , die zugehörige Kurve sättigt bei 5.34×10^2 . Für $z = 15$ liegt das Maximum bei 4.37×10^4 , der zugehörige Sättigungswert ist 4.38×10^4 . Die Abschätzung ist in unserem Fall also außerordentlich scharf. Interessant ist, dass sich das

$\log_{10}(\varepsilon)$	α_ε	$\alpha_\varepsilon/\varepsilon$	$\log_{10}(\varepsilon)$	α_ε	$\alpha_\varepsilon/\varepsilon$
0	3.36×10^1	3.36×10^1	0	1.27×10^2	1.27×10^2
-1	1.29×10^1	1.29×10^2	-1	6.78×10^1	6.78×10^2
-2	3.39×10^0	3.39×10^2	-2	3.26×10^1	3.26×10^3
-3	4.97×10^{-1}	4.97×10^2	-3	1.25×10^1	1.25×10^4
-4	5.30×10^{-2}	5.30×10^2	-4	3.05×10^0	3.05×10^4
-5	5.34×10^{-3}	5.34×10^2	-5	4.15×10^{-1}	4.15×10^4
-6	5.34×10^{-4}	5.34×10^2	-6	4.35×10^{-2}	4.35×10^4
-7	5.36×10^{-5}	5.36×10^2	-7	4.37×10^{-3}	4.37×10^4

Tabelle 9.1: Die pseudospektrale Abszisse α_ε für verschiedene Werte von ε . Die Tabelle links enthält die Werte für $z = 10$, die rechte Tabelle entspricht $z = 15$. Die Berechnung erfolgte mit der Matlab-Routine `Eigtool`. Für sehr kleine Werte $\varepsilon < 10^{-7}$ ließ sich α_ε nicht zuverlässig bestimmen.

Verhältnis für kleine ε einem konstanten Wert anzunähern scheint. Das scheint darauf hinzuweisen, dass die Resolvente bei dem Nulleigenwert wie im normalen Fall eine $\frac{1}{|x|}$ -Singularität besitzt, wobei der Koeffizient jedoch gegenüber dem normalen Fall von Eins verschieden ist.

Führen wir die selben Betrachtungen für $\tilde{\mathbf{A}}$ durch, so verschwinden alle diese Merkmale des Pseudospektrums. Wie bereits in Abschnitt 9.2 erwähnt sind die Ränder der Pseudospektren um den Nulleigenwert kreisförmig mit Radius ε . Die pseudospektrale Abszisse ist also $\alpha_\varepsilon = \varepsilon$ und die untere Schranke in (9.4) Eins. Beschäftigen wir uns nun mit einer weiteren Ungleichung, die eine obere Schranke für das Anwachsen angibt. Da $\|\mathbf{M}\| = \max_{\|x\|=1} \|\mathbf{M}x\|$ größer oder gleich dem Realteil des größten Eigenwerts ist (Gleichheit im Falle normaler Matrizen), gilt zunächst

$$e^{t\alpha(\mathbf{M})} \leq \|e^{t\mathbf{M}}\|,$$

wobei $\alpha(\mathbf{M})$ die spektrale Abszisse, also das Supremum des Realteils der Eigenwerte bezeichnet. Lässt sich \mathbf{M} mit $\mathbf{M} = \mathbf{V}\mathbf{D}\mathbf{V}^{-1}$ diagonalisieren (\mathbf{V} ist die Matrix der Eigenvektoren), gilt wegen $\|\mathbf{M}_1\mathbf{M}_2\| \leq \|\mathbf{M}_1\| \|\mathbf{M}_2\|$

$$\|e^{t\mathbf{M}}\| = \|\mathbf{V}e^{t\mathbf{D}}\mathbf{V}^{-1}\| \leq \|\mathbf{V}\| \|e^{t\mathbf{D}}\| \|\mathbf{V}^{-1}\|$$

und mit $\|e^{t\mathbf{D}}\| = e^{t\alpha(\mathbf{D})} = e^{t\alpha(\mathbf{M})}$ und der Bezeichnung $\kappa(\mathbf{V}) = \|\mathbf{V}\| \|\mathbf{V}^{-1}\|$ folgt

$$e^{t\alpha(\mathbf{M})} \leq \|e^{t\mathbf{M}}\| \leq \kappa(\mathbf{V}) e^{t\alpha(\mathbf{M})}. \quad (9.5)$$

Falls \mathbf{M} normal ist, ist $\kappa(\mathbf{V}) = 1$, woraus direkt die Gleichung $\|e^{t\mathbf{M}}\| = e^{t\alpha(\mathbf{M})}$ folgt. Für nichtnormale Matrizen lässt sich κ als Maß für die Nichtnormalität

verwenden⁴¹. In Abb. (9.23) ist die Güte der oberen Schranke in (9.5) dargestellt.

Insgesamt lässt sich also feststellen, dass sich das transiente Anwachsen gut durch die Theorie der Pseudospektren beschreiben lässt. Die Stärke dieses Anwachsens geht ebenso wie die Nichtnormalität exponentiell mit dem Parameter⁴² z und es lassen sich durch geeignete Ungleichungen obere und untere Schranken gewinnen.

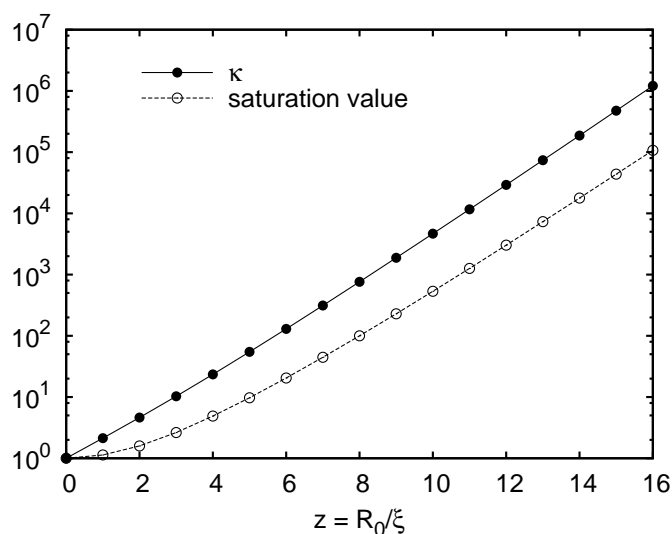


Abbildung 9.23: κ und der Sättigungswert von $\|\exp(t\mathbf{A})\|$ als Funktion von z . Der Auftrag ist halblogarithmisch

9.4.2 Symmetrischer Fall

Wir wiederholen nun die Untersuchung für den symmetrischen Fall. In Abb. 9.24 ist wieder $\|\exp(t\mathbf{A})\|$ als Funktion von t dargestellt. Wie erwartet ergibt sich diesmal asymptotisch ein linearer Anstieg. Die Steigung dieser Geraden wächst, wie aus Abb. 9.25 ersichtlich ist, exponentiell an. Dieses Anwachsen quantitativ mit der wachsenden Nichtnormalität in Verbindung zu bringen, ist jetzt jedoch schwieriger. Da die Matrix einen Jordanblock enthält, ist der Satz \mathbf{V} der Eigenvektoren, den das numerische Verfahren liefert, nicht mehr linear

⁴¹ κ ist die sogenannte Kondition der Matrix \mathbf{M} . Sie gibt an, wie empfindlich die Lösung des Gleichungssystems $\mathbf{Ax} = \mathbf{b}$ auf Störungen der rechten Seite reagiert. Zur Konditionierung des Eigenwertproblems und den Zusammenhang mit der Nichtnormalität, siehe auch das nächste Kapitel 9.5.

⁴²Die exponentielle Parameterabhängigkeit tritt häufig auf, siehe auch [64] S.8.

unabhängig⁴³, \mathbf{V} nicht invertierbar und die Größe κ nicht definiert⁴⁴. Eine Größe, die in einer zu (9.5) analogen Ungleichung die Rolle von κ übernehmen würde, wäre $\|\mathbf{T}\| \|\mathbf{T}^{-1}\|$, wobei \mathbf{T} die Matrix der verallgemeinerten Eigenvektoren ist, d.h. die Matrix, mit der sich \mathbf{A} auf die Jordan-Form \mathbf{J} transformiert $\mathbf{A} = \mathbf{T}\mathbf{J}\mathbf{T}^{-1}$. Das Auffinden der Jordan-Form ist allerdings numerisch schwierig⁴⁵. Die Matrix \mathbf{T} lässt sich in unserem Fall folgendermaßen bestimmen: \mathbf{T} und \mathbf{V} enthalten beide die echten Eigenvektoren und sind deshalb bis auf eine Spalte identisch. Der fehlende verallgemeinerte Eigenvektor erfüllt $\hat{\mathbf{A}}\mathcal{D} = \mathcal{T}_x$, d.h. da wir \mathcal{T}_x , einen echten Eigenvektor, kennen, lässt sich \mathcal{D} aus dieser Gleichung bestimmen. Die so erhaltene Matrix \mathbf{T} transformiert zwar \mathbf{A} in eine Jordan-Form, jedoch ergeben sich für $\|\mathbf{T}\| \|\mathbf{T}^{-1}\|$ enorm hohe Werte. Dies ist jedoch eher Ausdruck der ungünstigen numerischen Eigenschaften der Matrix \mathbf{A} , als dass es uns ein präzises Maß für die Nichtnormalität geben könnte.

9.5 Kondition des Eigenwertproblems

In diesem Abschnitt wenden wir uns wiederum der extremen Sensitivität gegenüber Störungen zu. Anders als im vorangegangenen Abschnitt, wo wir diesen Sachverhalt mittels der Theorie der Pseudospektren analysiert haben, werden wir uns hier dem Problem über eine Art von Störungstheorie für Matrizen nähern. Die folgenden Überlegungen sind äußerst wichtig für die numerische Behandlung von Eigenwertproblemen und werden uns, komplementär zu der „neueren“ Theorie der Pseudospektren, einen etwas anderen Blickwinkel auf das Problem erlauben. Dazu führen wir zunächst einige Begrifflichkeiten ein. Dabei beschränken wir uns der Einfachheit halber auf Matrizen mit nichtentarteten Eigenwerten. Diese Einschränkung ist im Fall des antisymmetrischen Teils des Spektrums erfüllt. Für den symmetrischen Teil mit dem zweifach entarteten Nulleigenwert wäre eine etwas differenziertere Untersuchung notwendig, die aber keinen weiteren Erkenntnisgewinn verspricht. Führen wir zunächst zusätzlich zu den üblichen (rechten) Eigenvektoren \mathbf{x}_n^R die *linken* Eigenvektoren \mathbf{x}_n^L

⁴³Der Algorithmus sucht nach einer Lösung zu $\mathbf{A}\mathbf{V} = \mathbf{V}\mathbf{D}$, wobei \mathbf{D} die Diagonalmatrix mit den Eigenwerten ist und \mathbf{V} im diagonalisierbaren Fall die Eigenvektoren enthält. Bei einem zweifachen Eigenwert, zu dem es nur einen Eigenvektor gibt, enthält \mathbf{V} diesen Eigenvektor einfach doppelt, wodurch die Gleichung erfüllt ist

⁴⁴Auch wenn sie sich numerisch bestimmen lässt. \mathbf{V} ist aufgrund der endlichen Genauigkeit numerisch invertierbar (die Entartung des Nulleigenwertes wird gewissermaßen numerisch aufgehoben und es entstehen zwei dicht beieinander liegende Eigenwerte), es ist aber nicht klar wieviel von $\|\mathbf{V}^{-1}\|$ von der Nichtnormalität und wieviel von der „Beinahe-Singularität“ kommt. So ist z.B. für $z = 10$ $\kappa = 5.75 \times 10^8$ im Vergleich zum asymmetrischen Fall $\kappa = 4.66 \times 10^3$ zum selben z -Wert

⁴⁵Dazu sei zitiert: „Interestingly, neither Matlab nor Octave seem to have a numerical function for computing the Jordan canonical form of a matrix. Matlab will try to do it symbolically when the matrix entries are given as exact rational numbers (ratios of integers) by the jordan function, which requires the Maple symbolic mathematics tool box. Numerically, it is generally difficult to distinguish between poles that are repeated exactly, and poles that are merely close together.“ unter http://ccrma.stanford.edu/jos/filters/Jordan_Canonical_Form.html

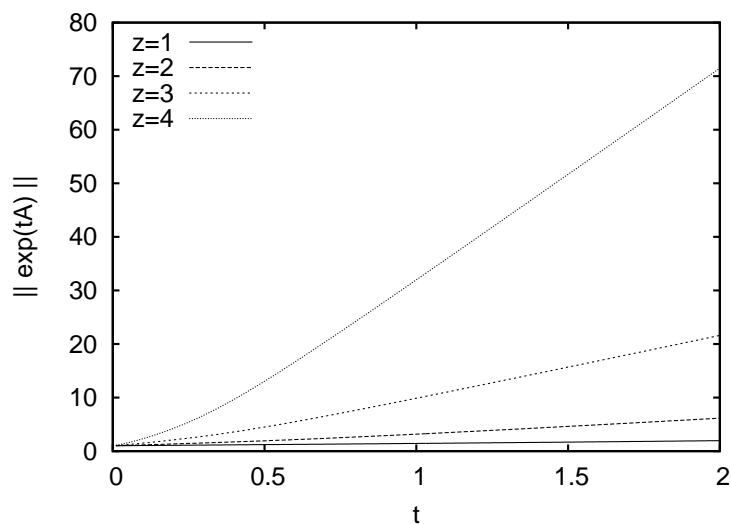


Abbildung 9.24: Verlauf von $\|\exp(t\mathbf{A})\|$ als Funktion von t für $z = 1, 2, 3, 4$.

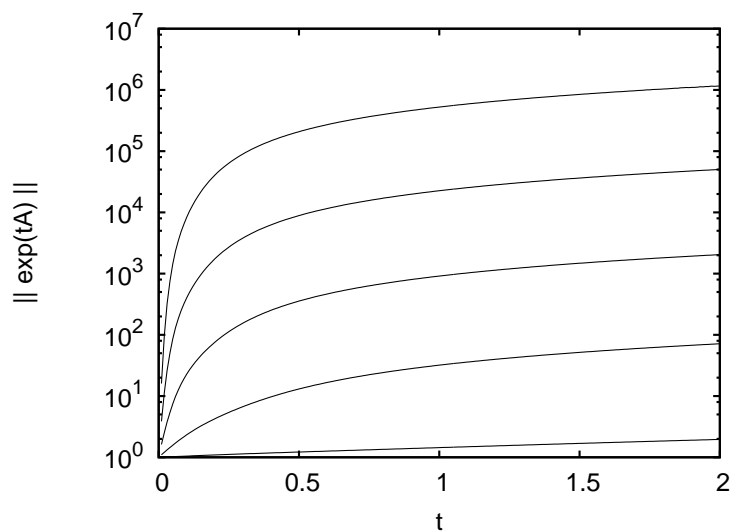


Abbildung 9.25: Verlauf von $\|\exp(t\mathbf{A})\|$ als Funktion von t für $z = 1, 4, 7, 10$ und 13 (entspricht den Kurven von unten nach oben). Der Auftrag ist halb-logarithmisch. Wie zu sehen ist, steigt z.B. bei $t = 2.0$ der Wert von $\|\exp(t\mathbf{A})\|$ exponentiell mit z an. Daraus folgt, dass die Steigung des asymptotisch linearen Verlaufs (siehe Abb. 9.24) exponentiell mit z wächst.

ein. Definiert sind diese beiden Größen durch

$$\mathbf{A}\mathbf{x}_n^R = \lambda_n \mathbf{x}_n^R \quad (9.6)$$

$$\mathbf{x}_n^L \mathbf{A} = \lambda_n \mathbf{x}_n^L. \quad (9.7)$$

Bei \mathbf{x}_n^R handelt es sich hierbei um Spaltenvektoren und bei \mathbf{x}_n^L um Zeilenvektoren. Aus der Gleichung ist ersichtlich, dass jeder rechte Eigenvektor von \mathbf{A} ein linker Eigenvektor von \mathbf{A}^\dagger ist. Damit folgt direkt aus der charakteristischen Gleichung

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0$$

für die Eigenwerte und der Tatsache, dass die Determinante unter Transposition invariant ist, dass linke und rechte Eigenwerte übereinstimmen (so wie es (9.7) bereits andeutet). In diesem Sinne hat jeder linke Eigenvektor einen rechten Partner, nämlich den, der zum selben Eigenwert gehört.

Untersuchen wir nun, wie sich ein Eigenwert bei einer kleinen Störung verändert. Dazu addieren wir zu der Matrix \mathbf{M} eine Matrix $\varepsilon \mathbf{E}$ mit $\|\mathbf{E}\| = 1$ ⁴⁶ hinzu. Die Eigenwertgleichung lautet nun

$$(\mathbf{M} + \varepsilon \mathbf{E})\mathbf{x}^R(\varepsilon) = \lambda(\varepsilon)\mathbf{x}^R(\varepsilon).$$

Differenzieren wir diese Gleichung nach ε und setzen $\varepsilon = 0$ erhalten wir

$$\mathbf{M} \frac{\partial \mathbf{x}^R}{\partial \varepsilon}(0) + \mathbf{E}\mathbf{x}^R = \frac{\partial \lambda}{\partial \varepsilon}(0)\mathbf{x}^R + \lambda \frac{\partial \mathbf{x}^R}{\partial \varepsilon}(0).$$

Multiplizieren wir beide Seiten von links mit \mathbf{x}^L und benutzen $\mathbf{x}^L \mathbf{M} = \lambda \mathbf{x}^L$ ergibt sich

$$\frac{\partial \lambda}{\partial \varepsilon}(0) = \frac{\mathbf{x}^L \mathbf{E} \mathbf{x}^R}{\mathbf{x}^L \mathbf{x}^R}.$$

Wählen wir nun \mathbf{x}^R und \mathbf{x}^L als normiert, können wir den Betrag des Zählers auf der rechten Seite mit der Cauchy-Schwarzschen Ungleichung durch $|\mathbf{x}^L \mathbf{E} \mathbf{x}^R| \leq \|\mathbf{x}^L\| \cdot \|\mathbf{E} \mathbf{x}^R\| \leq \|\mathbf{x}^L\| \cdot \|\mathbf{E}\| \cdot \|\mathbf{x}^R\| = 1$ abschätzen und erhalten damit zusammenfassend den

⁴⁶Im Folgenden bezeichnen die Doppelstriche $\|\dots\|$ immer die übliche euklidische Norm. Zur Definition der Matrixnorm siehe Abschnitt 9.2.

Satz 1. Sei $\mathbf{M} \in \mathbb{C}^{n \times n}$ eine diagonalisierbare Matrix, λ ein einfacher Eigenwert und $\mathbf{x}^R, \mathbf{x}^L$ die zugehörigen linken und rechten normierten Eigenvektoren. Ist nun $\mathbf{A}(\varepsilon) = \mathbf{A} + \varepsilon \mathbf{E}$ mit $\mathbf{E} \in \mathbb{C}^{n \times n}$ und $\|\mathbf{E}\| = 1$, dann gilt für den Eigenwert $\lambda(\varepsilon)$ von $\mathbf{A}(\varepsilon)$ mit $\lambda(0) = \lambda$

$$\left| \frac{\partial \lambda}{\partial \varepsilon}(0) \right| \leq \frac{1}{|\mathbf{x}^L \mathbf{x}^R|}.$$

(Für einen Beweis siehe [48], S. 189⁴⁷). Die Ableitung $\frac{\partial \lambda}{\partial \varepsilon}(0)$ ist nun nichts anderes als der Koeffizient des linearen Terms in der Störungsentwicklung $\lambda(\varepsilon) = \lambda(0) + \frac{\partial \lambda}{\partial \varepsilon}(0)\varepsilon + o(\varepsilon)$. Die Größe $\frac{1}{|\mathbf{x}^L \mathbf{x}^R|} = \frac{1}{\cos(\theta_\lambda)}$, wobei θ_λ der Winkel zwischen rechtem und linkem Eigenvektor ist, ist damit die sogenannte *Konditionszahl* des Eigenwertes. Allgemein gibt die Kondition die Abhängigkeit einer Lösung von der Störung der Eingangsdaten an. Die Konditionszahl quantifiziert diese Abhängigkeit, indem sie den Faktor angibt, um den Eingangsfehler im ungünstigsten Fall verstärkt werden (in unserem Fall haben wir natürlich die Terme höherer Ordnung vernachlässigt). Ausschlaggebend für die Kondition eines Eigenwertes ist also der Winkel zwischen rechtem und zugehörigem linken Eigenvektor. Im Falle normaler Matrizen sind rechter und linker Eigenvektor identisch und das Eigenwerteproblem damit gut konditioniert. Für beliebige Matrizen gilt nun folgende Überlegung ([47]): Sei \mathbf{V}_R die Matrix, deren Spalten aus den rechten Eigenvektoren gebildet werden und \mathbf{V}_L die Matrix, deren Zeilen die linken Eigenvektoren sind. Dann lassen sich die Eigenwertgleichungen (9.7) schreiben als

$$\begin{aligned} \mathbf{A} \cdot \mathbf{V}_R &= \mathbf{V}_R \cdot \text{diag}(\lambda_1 \dots \lambda_N) \\ \mathbf{V}_L \cdot \mathbf{A} &= \text{diag}(\lambda_1 \dots \lambda_N) \cdot \mathbf{V}_L \end{aligned}$$

Multipliziert man die erste Gleichung von links mit \mathbf{V}_L und die zweite von rechts mit \mathbf{V}_R erhält man durch Subtraktion

$$(\mathbf{V}_L \cdot \mathbf{V}_R) \cdot \text{diag}(\lambda_1 \dots \lambda_N) = \text{diag}(\lambda_1 \dots \lambda_N) \cdot (\mathbf{V}_L \cdot \mathbf{V}_R),$$

d.h. das Matrixprodukt $(\mathbf{V}_L \cdot \mathbf{V}_R)$ kommutiert mit der Diagonalmatrix der Eigenwerte. Da die Diagonalelemente verschieden sind, folgt daraus direkt, dass auch das Matrixprodukt Diagonalgestalt haben muss. Mit anderen Worten, *jeder linke Eigenvektor steht auf allen rechten Eigenvektoren senkrecht, mit Ausnahme seines Partners und umgekehrt*. Bei nichtnormalen Matrizen kann nun das Skalarprodukt zwischen den Partnern unter Umständen sehr klein werden. Wie in Abbildung 9.26 skizziert steht die Parallelität zwischen den Partnern in engem Zusammenhang mit der Orthogonalität der rechten (und umgekehrt der linken) Eigenvektoren untereinander.

⁴⁷Aus unserer Herleitung ist z.B. nicht ersichtlich, ob das Skalarprodukt $\mathbf{x}^L \mathbf{x}^R$ überhaupt von Null verschieden ist.

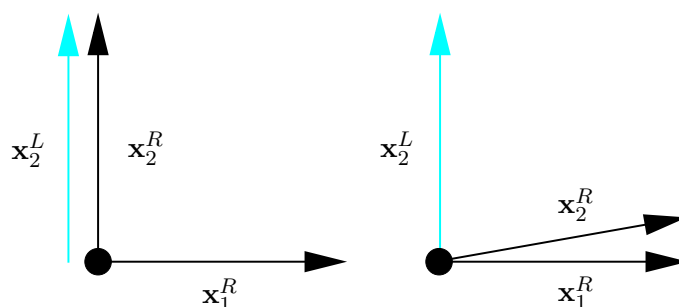


Abbildung 9.26: Veranschaulichung der geometrischen Beziehung zwischen linken und rechten Eigenvektoren: Im linken Bild stehen die rechten Eigenvektoren x_1^R und x_2^R aufeinander senkrecht. Da x_2^L auf x_1^R senkrecht steht (siehe Text), muss er parallel zu x_2^R sein. Im rechten Bild sind $x_{1,2}^R$ nicht orthogonal. Wegen $x_2^L \perp x_1^R$ kann x_2^L deshalb nicht parallel zu x_2^R sein.

Da die rechten Eigenvektoren bei unserer Matrix \mathbf{A} für große Werte von z annähernd parallel sind (siehe Abschnitt 9.2), muss also gleichzeitig die Kondition der Eigenwerte entsprechend schlecht sein. Dies führt dazu, dass die Linien der Eigenwerte als Funktion von z in der Nähe von $z = 30$ stark zu fluktuieren anfangen und sich schließlich in diffuse Punktwolken auflösen (siehe Abbildung 9.27).

Ein auf den ersten Blick seltsamer Umstand ergibt sich nun, wenn man dies dem Spektrum einer leicht gestörten Matrix gegenüberstellt⁴⁸ (Abbildung 9.28). Das Rauschen der Linien verschwindet durch die Störung vollständig, wenn auch das Spektrum, wie wir ja bereits wissen, vollkommen verändert wird. Warum nun bewirkt die leichte Störung eine Glättung der Linien? In Abbildung 9.29 ist die Kondition der größten Eigenwerte halblogarithmisch gegen z aufgetragen. Wie zu ersehen ist, geht die Kondition exponentiell mit z , ein Umstand, der die korrekte Bestimmung der Eigenwerte für große z unmöglich macht. Zusätzlich ist in der Abbildung die Kondition der gestörten Matrix aufgetragen. Wie man sieht, verbessert sich die Kondition durch die Störung dramatisch. Warum eine relativ beliebige Störung praktisch immer zu einer Verbesserung der Kondition führt, wird durch das Bild der nicht-orthogonalen Eigenvektoren verständlich. Dass zwei Vektoren in einem hochdimensionalen Raum annähernd zueinander parallel sind, stellt eine sehr spezielle Konfiguration da. Dementsprechend wird eine Veränderung eines der beiden Vektoren in fast jedem Fall den Winkel zwischen beiden vergrößern. In diesem Sinne entsprechen die Matrix \mathbf{A} und das zugehörige Spektrum einem marginalen Fall und die Nichtnormalität ist gewissermaßen selbst eine empfindliche Eigenschaft.

⁴⁸Die Störung ist $A_{5,5} \rightarrow A_{5,5} + 0.1$. Diese Störung ist relativ willkürlich gewählt, jedoch scheint die Veränderung des Spektrums von der genauen Form der Störung weitgehend unabhängig zu sein.

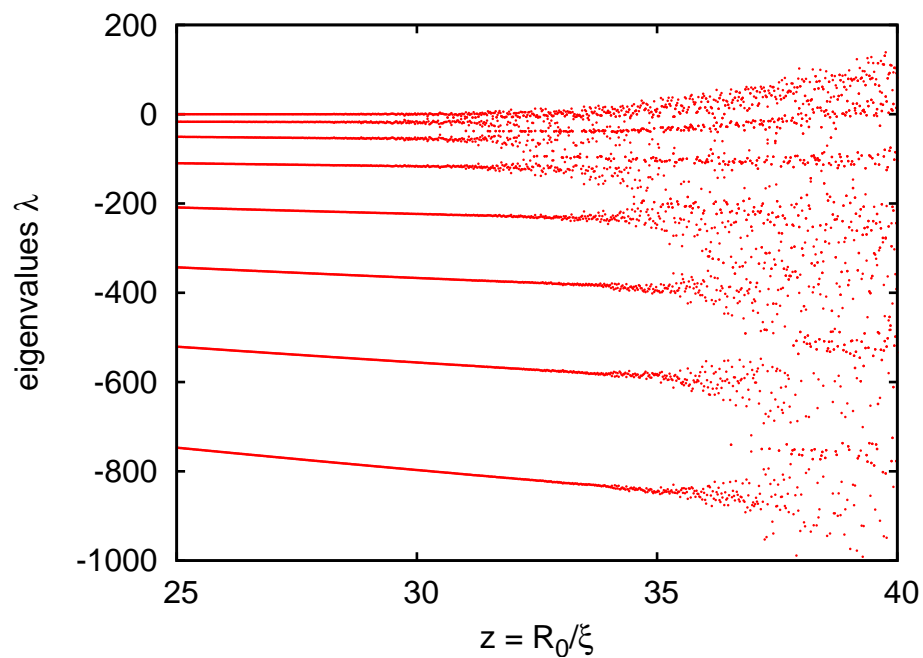


Abbildung 9.27: Die Eigenwerte $\lambda_1 \dots \lambda_8$ als Funktion von z .

Zuletzt sei noch angemerkt, dass der im Abschnitt 9.2 besprochene Basiswechsel vom Standpunkt der Konditionierung keine Verbesserung bringt. Die Matrix wird durch den Basiswechsel zwar weniger nichtnormal, da jedoch nun auch die Transformation des Basiswechsels zur Konditionierung beiträgt bedeutet dies für das Eigenwertproblem keinen Gewinn. Insbesondere das Rauschen der Linien kann damit nicht vermindert werden.

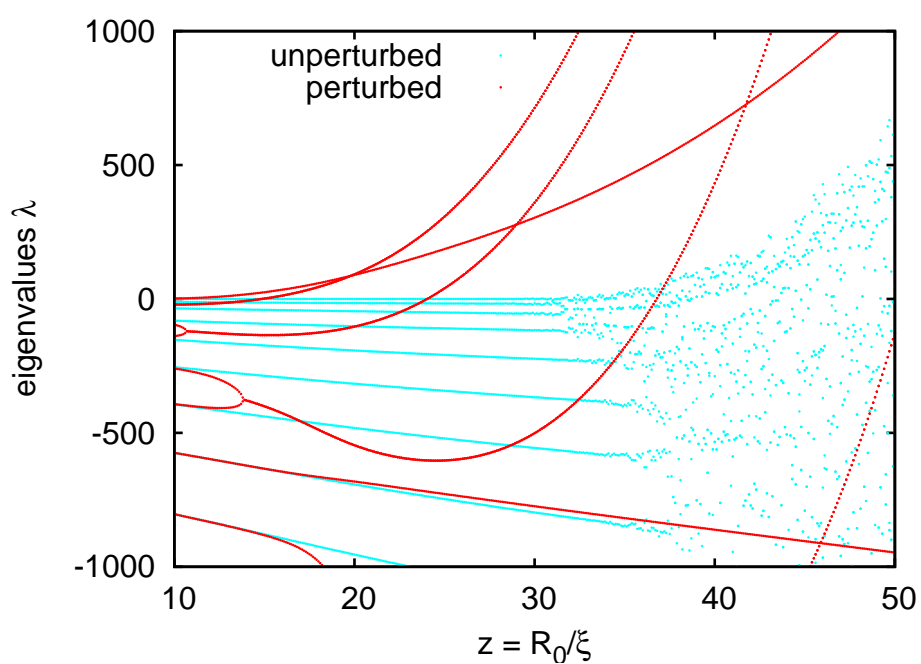


Abbildung 9.28: Die Eigenwerte $\lambda_1 \dots \lambda_8$ als Funktion von z im ungestörten und gestörten Fall. Im gestörten Fall handelt es sich eigentlich um ein komplexes Spektrum, von dem nur der Realteil dargestellt ist. Die zusammenlaufenden Linien entsprechen zwei Eigenwerten mit gleichem Realteil, deren Imaginärteile sich jedoch unterscheiden.

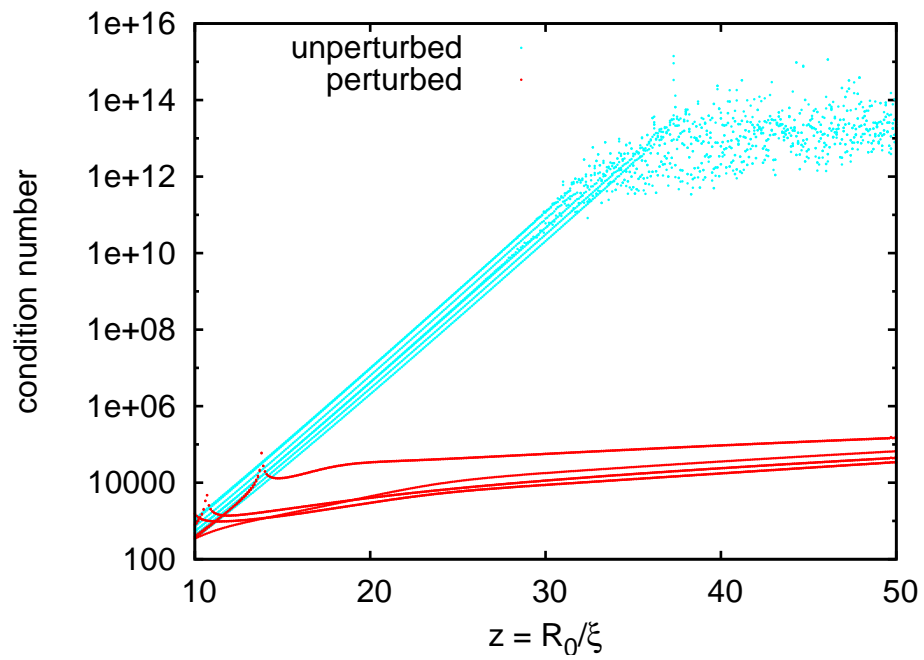


Abbildung 9.29: Kondition der gestörten und ungestörten Eigenwerte $\lambda_1 \dots \lambda_7$ im halblogarithmischen Auftrag.

Kapitel 10

Numerische Ergebnisse für die nichtlineare Zeitentwicklung

Obwohl wir in unserer Darstellung die Stabilitätsanalyse und die Beschreibung der nichtnormalen Eigenmoden an den Anfang gestellt haben, waren tatsächlich die analytischen und numerischen Untersuchungen aufs Engste miteinander verwoben. So war die Motivation, die Stabilitätsanalyse aus [44] zu reproduzieren, bzw. zu korrigieren zunächst die Tatsache, dass sich die dort vorhergesagte lineare Instabilität numerisch nicht nachweisen ließ. In diesem Abschnitt beschreiben wir nun die Simulationsergebnisse von Frank Haußer im Bezug auf die Frage nach der Stabilität der Kreislösung.

Die numerische Problemstellung zerfällt ebenso wie die physikalische in zwei (miteinander gekoppelte) Teilprobleme. Zum einen muss die Zeitentwicklung des Konzentrationsfeldes berechnet werden, zum anderen die Bewegung des Inselrandes. Für eine Diskretisierung dieser beiden Größen werden zwei unabhängige Gitter erzeugt – ein eindimensionales für den Inselrand und ein zweidimensionales für das Konzentrationsfeld. Beide Gitter werden – bei Bedarf – nach jedem Zeitschritt adaptiv verfeinert bzw. vergrößert¹. In jedem diskreten Zeitschritt erfolgt jetzt die Lösung in zwei Schritten:

- Lösung der Diffusionsgleichung auf dem 2-dimensionalen Gitter bei vorgegebenem Inselrand: Die dynamische Gleichung der Adatomkonzentration wird bezüglich des räumlichen Teils mit einem (linearen) Finite-Elemente-Verfahren diskretisiert und dann mittels eines impliziten Verfahrens erster Ordnung in der Zeit gelöst.
- Berechnung der Geschwindigkeit des Inselrandes bei gegebener Adatomkonzentration. Die Geschwindigkeit wird mittels eines impliziten Front-Tracking-Verfahrens berechnet, wobei wiederum finite Elemente – jetzt auf dem eindimensionalen parametrischen Gitter – verwendet werden. Aus

¹Wie wir im Folgenden sehen werden, entwickeln sich am Inselrand Bereiche mit beträchtlicher Krümmung, so dass eine Anpassung der Gitter hier von entscheidender Bedeutung ist.

der berechneten Geschwindigkeit ergibt sich dann der neue Inselrand (d.h. das neue eindimensionale Gitter).

Für eine detaillierte Beschreibung dieser Methode, siehe [3].

Für die Simulation wird also die volle zeitabhängige Gleichung (7.1), (7.2), (7.5), (7.7) und (7.8) integriert². Die quasistatische Gleichung (8.4), die wir für die analytische Rechnung verwendet haben, ist als Näherung gerechtfertigt, wenn die Bewegung des Inselrandes wesentlich langsamer verläuft als die der Adatomkonzentration. Dies ist wiederum der Fall, wenn die Konzentration gering ist. Es wurde sichergestellt, dass dies für die folgenden Ergebnisse der Fall ist.

Als Anfangsbedingung wurden Ellipsen mit Aspektverhältnis a/b von horizontaler zu vertikaler Ausdehnung gewählt. Für diesen Fall ist die Normalengeschwindigkeit einer analytischen Berechnung zugänglich. Die Details dieser Rechnung sind im Anhang A.7 zu finden. Das Ergebnis hat die Form einer Reihe von Mathieu-Funktionen und lässt sich nur für $z \leq 5$ und $a/b \leq 2$ verlässlich auswerten. Für diesen Bereich ist ein Vergleich mit der numerisch gewonnenen Normalengeschwindigkeit in Abbildung 10.1 dargestellt. Wie zu sehen ist, sind die beiden Ergebnisse in guter Übereinstimmung. Des Weiteren zeigt sich, dass die Geschwindigkeit am vorderen ($\theta = \pi$) Ende am größten und positiv, d.h. nach innen gerichtet ist. In der Nähe der Spitze, bei ungefähr $\theta \simeq \pi \pm \pi/4$ bewegt sich der Inselrand dagegen nach außen. Dies deutet darauf hin, dass sich die Ellipse an der Vorderseite wieder zur Kreisform zusammenzieht. Am hinteren Ende ist die Geschwindigkeit kleiner und immer nach innen gerichtet, so dass sich eine Entwicklung in Richtung einer Eiform erwarten lässt. Wie Abbildung 10.2 zeigt, trifft dies auch tatsächlich zu. Diese eiförmige Asymmetrie der Inselform wird mit steigenden Werten für z immer ausgeprägter, was vermutlich letztlich zur weiter unten diskutierten Finger-Instabilität führt, sobald an der Hinterseite ein kritischer Krümmungsradius erreicht ist. Diese Instabilität ist im Fall $z = 25$ in Abbildung 10.3 dargestellt. Bei einer anfänglichen Streckung in Feldrichtung bildet sich während einer längeren Anfangsphase die beschriebene Eiform aus, die sich am hinteren Ende zu einer relativ lokalisierten Ausbuchtung entwickelt. In der darauffolgenden schnelleren Phase wächst daraus ein Finger, der am Hals einschnürt und sich schließlich ablöst. Für die senkrecht zum Feld gestreckte Anfangskonfiguration bilden sich symmetrisch zueinander³ zwei Finger aus, die sich ebenfalls ablösen. Die zeitliche Entwicklung über diese topologische Änderung hinaus wurde nicht simuliert. Es lässt sich absehen, dass sich der abgelöste Teil wegen seiner kleinen Größe vermutlich wieder zu einer Kreisform zusammenzieht. Voraussichtlich wird die Insel so lange kleine Tochterinseln absondern, bis sie eine stabile Größe erreicht hat.

Für große Inseln ist diese Fingerbildung numerisch nicht mehr von einer linearen

²In (7.7) wurde bei der numerischen Integration der zweite Term vernachlässigt. Um der thermodynamischen Randbedingung (7.9) möglichst nahe zu kommen wurde die Rate k_+ entsprechend groß gewählt.

³Die Symmetrie wird allerdings durch das numerische Rauschen zerstört.

Instabilität zu unterscheiden. In Abbildung 10.4 sind einige Werte des Aspektverhältnisses und des Parameter z dargestellt, für die sich eine stabile, bzw. instabile Dynamik ergibt. Während für $z = 8$ auch stark gestreckte Anfangskonfigurationen stabil zur Kreisform zurückkehren, ergibt sich ab $z = 9$ ein Schwellwert, oberhalb dessen die Insel instabil wird. Ab $z = 11$ führen auch sehr kleine Aspektverhältnisse zu einer Fingerbildung.

Obwohl wir die Ursache dieser Instabilität nicht auf analytischer Grundlage klären können, wollen wir zum Abschluss noch ein qualitatives Bild entwerfen, wie die Fingerbildung unserem Verständnis nach zustande kommt. Erinnern wir uns zunächst, dass die Dynamik von zwei Massenströmen bestimmt wird, dem diffusiven Strom j_d mit

$$j_d = -D\nabla c \cdot \mathbf{n}$$

und dem Strom, der durch die Elektromigration verursacht wird

$$j_e = \frac{D}{\xi} c_{eq} \mathbf{n} \cdot \mathbf{e}_x.$$

Zuerst betrachten wir zwei einfache Grenzfälle, in denen jeweils einer der beiden Ströme verschwindet. Versetzen wir den Inselrand mit einer Störung in Form einer Ausbuchtung, so verringert sich dort nach (7.5) die lokale Konzentration. Die beiden Fälle ergeben nun die folgenden Szenarien:

(a) *Kein elektrisches Feld.* Die verringerte Konzentration innerhalb der Ausbuchtung führt zu einem Konzentrationsgradienten und damit zu einem diffusiven Fluss j_d , der den Massenstrom in die Ausbuchtung beschleunigt. Die Ausbuchtung wird damit aufgefüllt und die Störung geht zurück.

(b) *Keine Diffusion.* Der Massenstrom j_e ist proportional zur Adatomkonzentration. Da diese durch die Störung verringert ist, wird auch der Massentransport durch j_e vermindert. Die Ausbuchtung driftet damit gegenüber der ungestörten Vorderseite der Insel langsamer und bleibt hinter ihr zurück. Die Ausbuchtung *wächst* also. An der Vorderseite würde ein analoges Argument dagegen dazu führen, dass die langsamer driftende Ausbuchtung *schrumpft*.

Betrachten wir nun den Fall mit beiden Massenströmen j_d und j_e . Während j_d am Rand im Wesentlichen durch die lokale Geometrie bestimmt ist, ist j_e darüber hinaus proportional zu $1/\xi$. Halten wir alle übrigen Parameter fest, so ist zu erwarten, dass die Verminderung des Stroms j_e durch den diffusiven Strom j_d nicht mehr ausgeglichen werden kann, wenn das Feld sehr stark wird. Die Ausbuchtung würde dann weiter wachsen, was der beobachteten Finger-Instabilität entspricht.

Durch dieses Argument wird auch die Rolle der Kapillarität erhellt, die im Wesentlichen Änderungen der lokalen Geometrie in Änderungen in der Adatomkonzentration übersetzt. Diese Konzentrationsänderung beeinflusst sowohl den

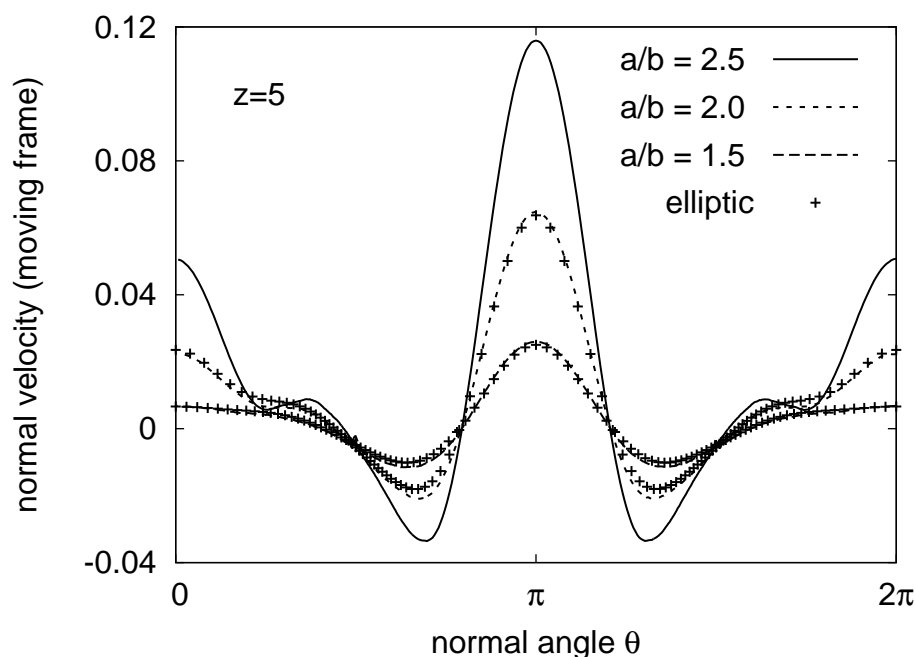


Abbildung 10.1: Normalengeschwindigkeit einer ellipsenförmigen Insel als Funktion des Normalenwinkels θ im Schwerpunktssystem (d.h. nachdem die Driftgeschwindigkeit abgezogen wurde). Dargestellt ist der Verlauf für $z = 5$ und verschiedene Werte des Aspektverhältnisses. Positive Werte bedeuten, dass sich der Inselrand an dieser Stelle nach innen bewegt. Der Winkel $\theta = \pi$ entspricht der Bewegungsrichtung. Die Kurven entsprechen Simulationsdaten, wobei die Normalengeschwindigkeit über die ersten 1000 Zeitschritte gemittelt wurde. Die Kreuze kennzeichnen die analytisch berechneten Werte.

stabilisierenden Strom j_d als auch den destabilisierenden Einfluss von j_e . Während, wie wir im ersten Teil dieser Arbeit gesehen haben, im lokalen Modell Kapillarität und Elektromigration im Widerstreit liegen, ist hier die Auswirkung der Kapillarität vergleichsweise neutral, was auch (im Nachhinein) erklärt, warum die linearen Stabilitätseigenschaften unabhängig von Γ sind.

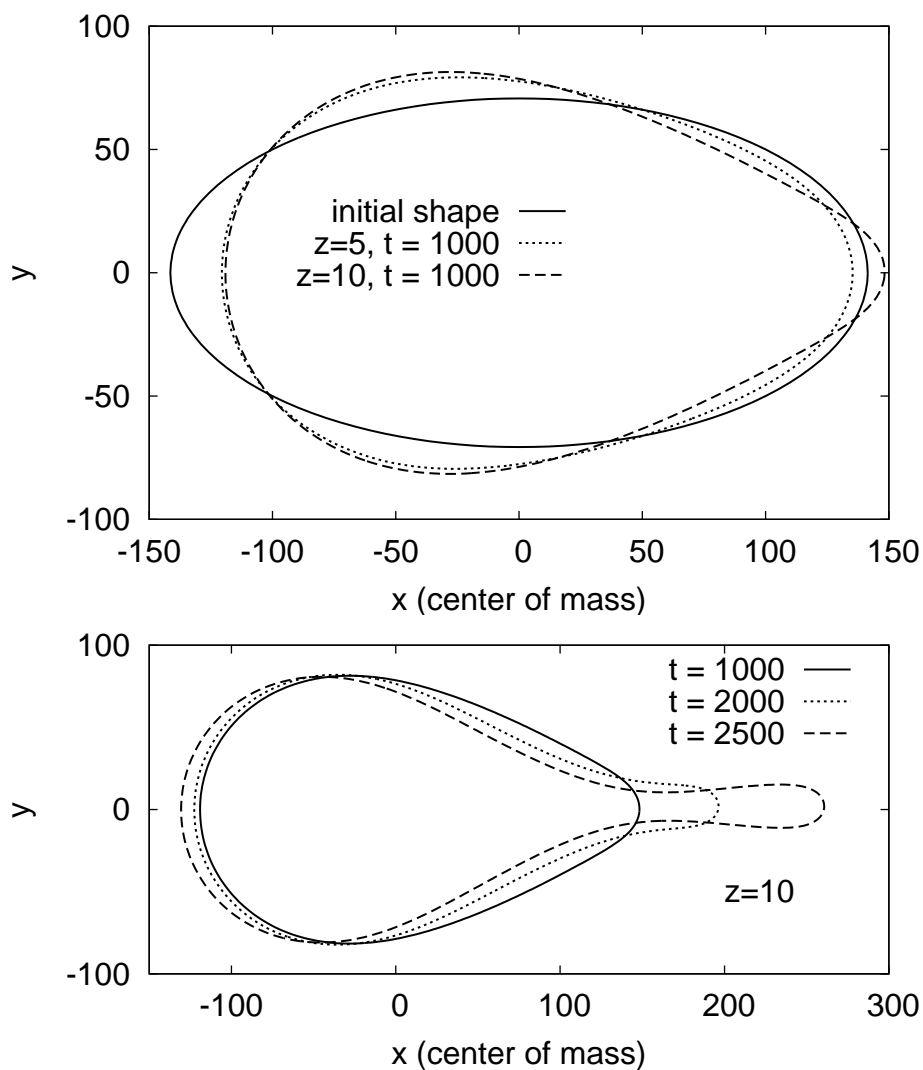


Abbildung 10.2: Numerische Simulation einer ellipsenförmigen Anfangskonfiguration mit Aspektverhältnis $a/b = 2$. Das obere Bild zeigt die Anfangsphase der Entwicklung, in der sich eine Eiform ausbildet. Diese Form wächst in der unteren Abbildung zu einem Finger an der Rückseite der Insel aus. Die Längenskala ist in Einheiten der atomaren Gitterkonstanten a gemessen und die Zeit in Einheiten von $t_0 = R_0/V$, mit der Driftgeschwindigkeit V einer Insel mit $z = 10$. Die Simulationsparameter sind $R_0 = 100a$, $a^2 c_{eq}^0 = 10^{-5}$ und $\Gamma/R_0 = 0.05$.

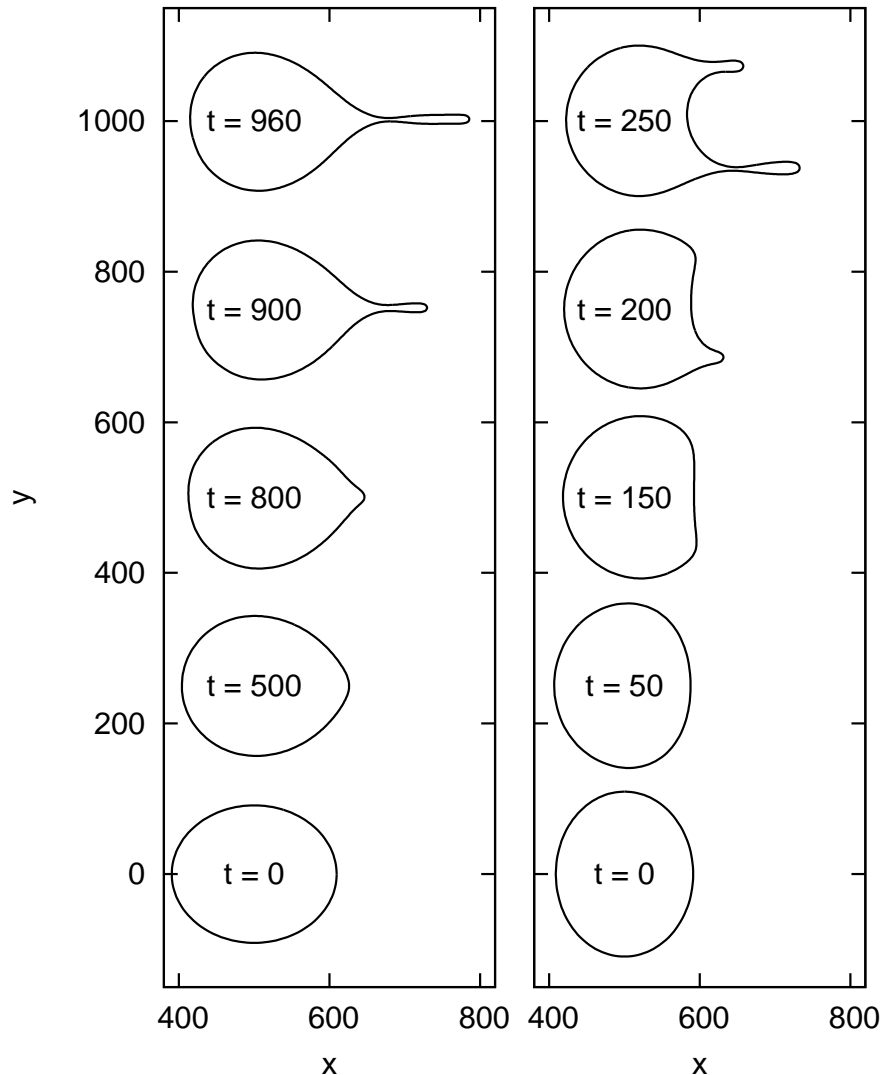


Abbildung 10.3: Nichtlineare Entwicklung der Leerstelleninsel für starken Elektromigrationseinfluss ($z = 25$). Die Anfangskonfiguration ist eine Ellipse, die mit einem Verhältnis von 1.2 in Feldrichtung (links), bzw. senkrecht dazu (rechts) gestreckt wurde. Die einzelnen Schnappschüsse sind der besseren Sichtbarkeit halber vertikal und um $\Delta x = tV$ horizontal verschoben, wobei V die stationäre Driftgeschwindigkeit eines Kreises derselben Fläche ist. Die räumliche und zeitliche Skala ist wieder in Einheiten von a , bzw. $t_0 = R_0/V$ gemessen. Die Simulationsparameter sind $R_0 = 100a$, $a^2 c_{eq}^0 = 10^{-5}$ und $\Gamma/R_0 = 0.02$. Die Asymmetrie zwischen oberer und unterer Hälfte der Insel in der rechten Abbildung ist eine Folge des numerischen Rauschens.

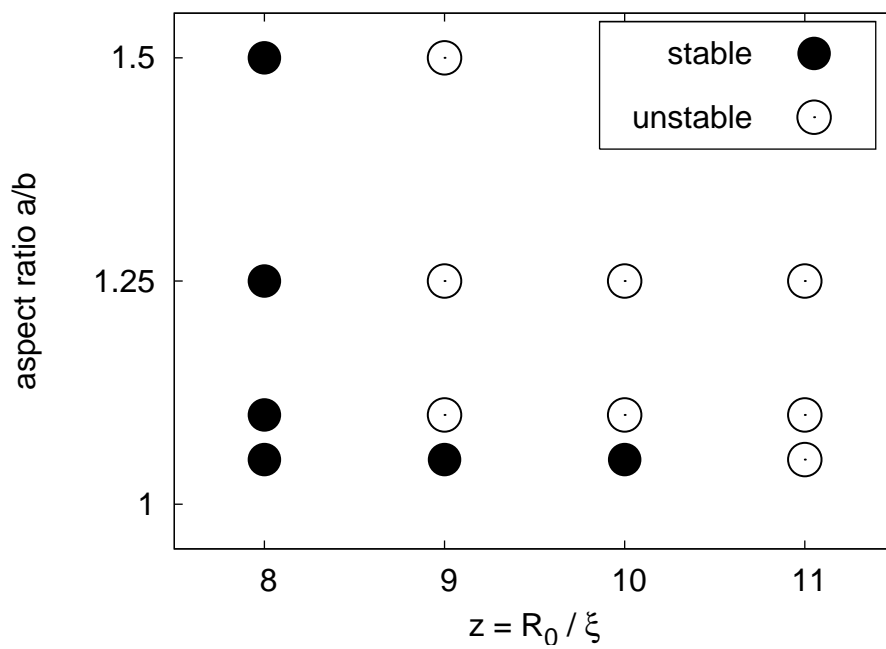


Abbildung 10.4: Stabilitätsdiagramm in der Ebene der Parameter $z = R_0/\xi$ und a/b (Aspektverhältnis). Für $z = 8$ ziehen sich auch stark in Feldrichtung gestreckte Inseln wieder zur Kreisform zusammen, während sich ab $z = 9$ ein Schwellwert ergibt, für den die Form instabil wird. Ab $z \geq 11$ führen auch leichte Abweichungen von der Kreisform zur in Abbildung 10.3 dargestellten Finger-Instabilität.

Kapitel 11

Zusammenfassung

In dieser Arbeit wurde die Dynamik von Oberflächeninseln unter Elektromigrationseinfluss untersucht. Dabei betrachten wir zwei Grenzfälle, bei denen der Massentransport entweder ausschließlich entlang des Inselrandes stattfindet oder durch Materialaustausch zwischen dem Rand und der Adatom-Konzentration auf der Terrasse. Die Modelle für die Inselbewegung gehen auf die Arbeiten [44] von Olivier Pierre-Louis und Theodore L. Einstein zurück. In diesem Modell werden sowohl die Stufenkante als auch die Konzentration als ein kontinuierliches Gebilde behandelt, dessen Dynamik in Form von gewöhnlichen oder partiellen Differentialgleichungen mit geeigneten Randbedingungen beschrieben wird. Die grundlegende Modellannahme ist, dass der Stromfluss, der die Elektromigration erzeugt, nicht von der Anwesenheit der Insel beeinflusst wird und wir damit eine Kraftwirkung von konstanter Stärke und Richtung annehmen können.

Im ersten Teil wurde der Grenzfall der Kantendiffusion behandelt. Die Dynamik ist hier räumlich eindimensional und durch ein lokales Gesetz gegeben. Die wesentliche Erweiterung, die wir hier gegenüber der früheren Untersuchung einbringen, ist die Berücksichtigung der Anisotropie der Kristalloberfläche, die im Rahmen unserer theoretischen Beschreibung in Form einer orientierungsabhängigen Adatom-Beweglichkeit und Stufen-Steifigkeit eingeht. Wir konzentrierten uns dabei größtenteils auf den Fall einer Beweglichkeit mit sechszähliger Symmetrie, bei der die parallel zum Feld orientierten Kanten eine maximale Beweglichkeit besitzen. Darüber hinaus wurde in den meisten Fällen eine isotrope Steifigkeit angenommen, um den Einfluss der kinetischen Anisotropie in Reinform studieren zu können. Die Fälle mit einer geringeren Zähligkeit der Anisotropie und anderen Winkeln zwischen der Richtung maximaler Beweglichkeit und Feld wurden ebenfalls, wenngleich auch weniger systematisch behandelt.

In Kapitel 3 wurde vorbereitend der Fall einer unendlich ausgedehnten Stufenkante betrachtet. In der früheren Arbeit [32] wurde für ein solches System ein Facettierungsübergang vorhergesagt, bei dem sich die Kante in eine Berg- und Tal-Struktur entwickelt, in der sich jeweils zwei ausgewählte Facetten-Orientierungen abwechseln. Im Kapitel 4 wurden numerische Simulationen bei

isotroper Beweglichkeit mit Ergebnissen früherer Arbeiten aus dem Kontext der Elektromigration von Poren im Inneren eines Materials verglichen.

Die Untersuchung für den anisotropen Fall ist in Kapitel 5 dargestellt. Hier stützten wir uns zunächst auf numerische Simulationen, die zum einen mit einem eigenen Finite-Differenzen-Algorithmus und zum anderen von Frank Haußer mit einem stabileren Finite-Elemente-Verfahren berechnet wurden. Wir fanden in Abhängigkeit von der Inselgröße, bzw. Feldstärke und der Stärke der Anisotropie ein reiches dynamisches Verhalten, das neben Inseln, die stationär in Feldrichtung oder *schräg* dazu driften, ein Regime einschließt, in dem die Inseln *oszillatorische* Formänderungen durchlaufen. Diese Oszillationen nehmen mit steigender Inselgröße immer komplexere Formen an. Sie stehen mit dem Auftauchen von Einbuchtungen am Inselrand in Zusammenhang, die relativ zum Substrat stillstehen und für größere Inseln Facetten-Charakter entwickeln. Für diese statischen Elemente ist die in [32] entwickelte analytische Theorie anwendbar, die es uns ermöglicht, aus den damit vorhergesagten Facetten-Orientierungen mittels einer geometrischen Auswahlregel die vom System ausgewählten Facetten-Orientierungen zu bestimmen, die das dynamische Verhalten der oszillierenden Inseln entscheidend bestimmen.

in Kapitel 6 wurden die numerischen Simulationen des Kontinuumsmodells mit den Ergebnissen kinetischer Monte-Carlo-Simulationen verglichen, die von Marko Rusanen durchgeführt wurden. Wir fanden dabei eine sehr gute Übereinstimmung bei einer Temperatur von 700 K , insbesondere entstehen auch im KMC-Modell Oszillationen, die denen des Kontinuumsmodells erstaunlich ähnlich sind. Bei der tieferen Temperatur $T = 500\text{ K}$ ergaben sich Abweichungen, die darauf hindeuten, dass die Kontinuumsgleichungen das KMC-Modell nicht mehr ausreichend gut beschreiben.

Im zweiten Teil der Arbeit untersuchten wir die nichtlokale Dynamik, die sich durch Kopplung zwischen Inselrand und Konzentrationsfeld ergibt. Wir beschränken uns hier auf das sogenannte „innere Modell“, d.h. eine Leerstelleninsel, die nur mit der Konzentration im Inneren wechselwirkt. In Kapitel 8 konnten wir zeigen, dass die Lösung eines, mit konstanter Geschwindigkeit driftenden Kreises für alle Inselgrößen und Feldstärken linear stabil ist. Es ergeben sich jedoch einige Besonderheiten, die mit der Nichtnormalität des mit der Stabilitätsfrage verbundenen Eigenwertproblems zusammenhängen. In Kapitel 9.2 analysierten wir diese Nichtnormalität mittels der Theorie der Pseudospektren. Insbesondere überprüften wir sorgfältig, ob sich für die lineare Theorie ein transientes Anwachsen von Störungen ergibt, die eventuell zu einer nicht-linearen Instabilität führen könnten. In Kapitel 10 diskutierten wir schließlich die numerischen Simulationen von Frank Haußer für diesen Fall, in denen sich das Einsetzen einer Instabilität zeigt, bei der am hinteren Ende der Insel eine Ausbuchtung entsteht, die in einen Finger auswächst und sich schließlich von der Insel ablöst.

Wir wollen zum Schluss noch auf mögliche experimentelle Realisierungen zu sprechen kommen, mit denen sich die Vorhersagen der beiden Modelle testen

lassen. Für das lokale Modell ist es sinnvoll, von den Ergebnissen der KMC-Simulationen auszugehen. Wenn wir annehmen, dass sich experimentell ein Wert von $E_{bias} = 10^{-5} \text{ eV}$ erzielen lässt¹, ergeben sich in der Simulation bei einer Temperatur von 700 K Oszillationen für eine Insel aus ungefähr 2×10^5 Atomen. Der entsprechende Inselradius wäre dann 90 nm und die Oszillationsperiode größenordnungsmäßig etwa 10 Sekunden.

Im nichtlokalen Fall schätzen wir die charakteristische Länge, die mit der Elektromigration verbunden ist, auf $\xi/a = k_B T / E_{bias} \simeq 2.5 \times 10^3$. Die Fingerinstabilität, die bei etwa $z = 10$ einsetzt, wäre demnach für eine Insel zu erwarten, die 2×10^9 Atome enthält, was vier Größenordnungen über der geschätzten Inselgröße für den lokalen Fall liegt.

¹Die Stromstärke, die notwendig ist, um Elektromigrationseffekte zu erzeugen, hängt natürlich davon ab, wie stark die vom Strom erzeugte Windkraft ist, d.h. wie stark ein Oberflächenatom an den Elektronenfluss koppelt. Kürzlich durchgeführte Messungen an dünnen Silber-Schichten ergaben hier Werte, die um etwa eine Größenordnung höher liegen, als üblicherweise angenommen wird [9].

Anhang A

Anhang

A.1 Wichtige geometrische Beziehungen

Wir leiten im Folgenden einige wichtige Formeln her, die die Zeitentwicklung einer Kurve $\mathbf{x}(u, t)$ mit Parameter u mit Normalengeschwindigkeit $v_n(u, t)$ betreffen. Diese Zeitentwicklung schreibt sich mit dem Einheitsvektor \mathbf{N} in Normalenrichtung als

$$\partial_t \mathbf{x} = v_n \mathbf{N}. \quad (\text{A-1})$$

Der Vektor \mathbf{N} ist mit dem Einheitsvektor in Tangentialrichtung \mathbf{T} über die Frenet-Formeln verknüpft

$$\partial_s \mathbf{T} = \kappa \mathbf{N} \quad (\text{A-2})$$

$$\partial_s \mathbf{N} = -\kappa \mathbf{T}. \quad (\text{A-3})$$

bzw.

$$\partial_\theta \mathbf{T} = \mathbf{N} \quad (\text{A-4})$$

$$\partial_\theta \mathbf{N} = -\mathbf{T}. \quad (\text{A-5})$$

Das Bogenelement ds ist mit du über

$$ds = |\partial_u \mathbf{x}| du \quad (\text{A-6})$$

verknüpft, woraus direkt

$$\partial_u = |\partial_u \mathbf{x}| \partial_s$$

folgt. Damit lässt sich nun folgender Ausdruck herleiten

$$(\partial_u \mathbf{x})(\partial_u \partial_t \mathbf{x}) = \underbrace{(\partial_u \mathbf{x})}_{|\partial_u \mathbf{x}| \mathbf{T}} \left(|\partial_u \mathbf{x}| \underbrace{\partial_s \partial_t \mathbf{x}}_{v_n \mathbf{N}} \right) = -|\partial_u \mathbf{x}|^2 \kappa v_n$$

$$\underbrace{(\partial_s v_n) \mathbf{N} - v_n \kappa \mathbf{T}}$$

und damit ist

$$\partial_t |\partial_u \mathbf{x}| = \frac{\partial_u \mathbf{x}}{|\partial_u \mathbf{x}|} \partial_u \partial_t \mathbf{x} = -|\partial_u \mathbf{x}| \kappa v_n.$$

Multipliziert man nun beide Seiten mit du erhält man mit (A-6) die Zeitentwicklung des Bogenelements

$$\partial_t ds = -\kappa v_n ds. \quad (\text{A-7})$$

Eine unmittelbare Folgerung ist die Zeitentwicklung der Bogenlänge L der Kurve, die man durch Integration erhält

$$\partial_t L + \int \kappa v_n ds = 0$$

$$\partial_t L + \int v_n d\theta = 0.$$

Des Weiteren folgt durch

$$\partial_t \partial_s f = \frac{\partial_t df}{ds} - \frac{df}{ds^2} \partial_t ds = \partial_s \partial_t f + \kappa v_n \partial_s f$$

der Kommutator

$$[\partial_t, \partial_s] = \partial_t \partial_s - \partial_s \partial_t = \kappa v_n \partial_s. \quad (\text{A-8})$$

Wenden wir diesen nun auf \mathbf{x} an, erhalten wir unter Ausnutzung von (A-2)

$$\begin{aligned} \partial_t \partial_s \mathbf{x} - \partial_s \partial_t \mathbf{x} &= \kappa v_n \partial_s \mathbf{x} \\ \Leftrightarrow \partial_t \mathbf{T} - \underbrace{\partial_s (v_n \mathbf{N})}_{(\partial_s v_n) \mathbf{N} - \kappa v_n \mathbf{T}} &= -\kappa v_n \mathbf{T} \\ \Rightarrow \partial_t \mathbf{T} &= (\partial_s v_n) \mathbf{N} \\ \Rightarrow \underbrace{\partial_t \mathbf{T}}_{\underbrace{\partial_\theta \mathbf{T}(\partial_t \theta)}_{-\mathbf{N}}} &= (\partial_s v_n) \mathbf{N} \end{aligned}$$

und damit

$$\partial_t \theta = \partial_s v_n \quad (\text{A-9})$$

A.2 Einige nützliche geometrische Beziehungen in Polarkoordinaten 155

woraus folgt

$$\partial_s \partial_t \theta = (\partial_t \partial_s - \kappa v_n \partial_s) \theta = \partial_s^2 v_n$$

d.h.

$$\partial_t \kappa = \partial_s^2 v_n + \kappa^2 v_n \quad (\text{A-10})$$

A.2 Einige nützliche geometrische Beziehungen in Polarkoordinaten

Für die Darstellung in Polarkoordinaten ist es sinnvoll, die Basis $\{\mathbf{e}_r, \mathbf{e}_\theta\}$ mit

$$\mathbf{e}_r = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} \quad \text{und} \quad \mathbf{e}_\theta = \begin{pmatrix} -\sin \theta \\ \cos \theta \end{pmatrix} \quad (\text{A-11})$$

zu verwenden. Es gelten die folgenden Beziehungen für die Ableitungen

$$\begin{aligned} \partial_\theta \mathbf{e}_r &= \mathbf{e}_\theta \\ \partial_\theta \mathbf{e}_\theta &= -\mathbf{e}_r \\ \partial_r \mathbf{e}_{r,\theta} &= 0. \end{aligned}$$

Der Vektor \mathbf{e}_x hat in dieser Basis die Darstellung

$$\mathbf{e}_x = \cos \theta \mathbf{e}_r - \sin \theta \mathbf{e}_\theta, \quad (\text{A-12})$$

und der **Gradient** von $c(r, \theta)$ ist

$$\nabla c = (\partial_r c) \mathbf{e}_r + \frac{1}{r} (\partial_\theta c) \mathbf{e}_\theta. \quad (\text{A-13})$$

Der **Laplace-Operator** hat die Darstellung

$$\nabla^2 = \frac{1}{r} \partial_r (r \partial_r) + \frac{1}{r^2} \partial_\theta^2 \quad (\text{A-14})$$

und die **Ableitung nach x** ist

$$\partial_x = \cos \theta \partial_r - \frac{1}{r} \sin \theta \partial_\theta. \quad (\text{A-15})$$

Der **Normalenvektor** des Inselrandes $\mathbf{R}(\theta) = \mathbf{R}_S + (R_0 + \varepsilon \rho(\theta)) \mathbf{e}_r$ (\mathbf{R}_S ist die Position des Schwerpunktes) ist

$$\mathbf{n} = \varepsilon \frac{1}{R_0} (\partial_\theta \rho) \mathbf{e}_\theta - \mathbf{e}_r. \quad (\text{A-16})$$

Damit ist die **Normalengeschwindigkeit**

$$v_n = \mathbf{n} \cdot \partial_t \mathbf{R} = \frac{a^2 D}{\xi} \tilde{c}_{eq} \cos \theta + \varepsilon \left(\frac{a^2 D \tilde{c}_{eq}}{\xi R_0} \sin \theta \partial_\theta \rho - \partial_t \rho \right). \quad (\text{A-17})$$

Die **Krümmung**¹ ist in Polarkoordinaten

$$\kappa = - \frac{r^2 + 2(\partial_\theta r)^2 - r \partial_\theta^2 r}{(r^2 + (\partial_\theta r)^2)^{\frac{3}{2}}}.$$

Für den gestörten Kreis $r = R_0 + \varepsilon \rho$ ergibt sich damit

$$\kappa = - \frac{1}{R_0} + \varepsilon \frac{\rho + \partial_\theta^2 \rho}{R_0^2}. \quad (\text{A-18})$$

A.3 Einige Beziehungen für die modifizierten Besselfunktionen

Als Ausgangspunkt wählen wir die **erzeugende Funktion**²

$$\exp \left(\frac{x}{2} \left(t + \frac{1}{t} \right) \right) = \sum_{n=-\infty}^{\infty} I_n(x) t^n, \quad (\text{A-19})$$

das heißt, die modifizierten Besselfunktionen tauchen hier als Koeffizienten der Laurent-Reihe auf. Da die linke Seite sich unter $t \rightarrow \frac{1}{t}$ nicht ändert, folgt zunächst die **Symmetrie bezüglich des Index**

$$I_n(x) = I_{-n}(x). \quad (\text{A-20})$$

Aus (A-19) ist ersichtlich, dass auf der linken Seite die Ersetzung $x \rightarrow -x$ dieselbe Wirkung hat wie $t \rightarrow -t$. Damit gilt für die linke Seite

¹Im Falle einer Leerstelleninsel ist die Krümmung für eine konvexe Insel *negativ*.

²Wir beschäftigen uns ausschließlich mit Funktionen zu *ganzzahligem* Index. Die Gleichung (8.6) kann auch für beliebiges positives $\lambda = \nu^2$ gelöst werden, wodurch sich die Besselfunktionen I_ν mit $\nu \in \mathbb{R}$ ergeben (für $\lambda < 0$ ergeben sich entsprechend imaginäre Indizes).

$$\sum_{n=-\infty}^{\infty} I_n(-x)t^n = \sum_{n=-\infty}^{\infty} I_n(x)(-t)^n = \sum_{n=-\infty}^{\infty} (-1)^n I_n(x)(t)^n,$$

woraus direkt die Beziehung

$$I_n(-x) = (-1)^n I_n(x) \tag{A-21}$$

folgt.

Aus (A-19) lässt sich eine Reihendarstellung gewinnen, indem man zuerst die Exponentialreihe benutzt³

$$\begin{aligned} \exp\left(\frac{x}{2}\left(t + \frac{1}{t}\right)\right) &= \exp\left(\frac{xt}{2}\right) \exp\left(\frac{x}{2t}\right) \\ &= \left(\sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{xt}{2}\right)^k\right) \left(\sum_{m=0}^{\infty} \frac{1}{m!} \left(\frac{x}{2t}\right)^m\right) = \sum_{n=-\infty}^{\infty} \left(\sum_{k-m=n} \frac{1}{m! k!} \left(\frac{x}{2}\right)^{m+k}\right) t^n \\ &= \sum_{n=-\infty}^{\infty} \left(\sum_{m=0}^{\infty} \frac{1}{m! (m+n)!} \left(\frac{x}{2}\right)^{2m+n}\right) t^n \end{aligned}$$

und damit die **Potenzreihe**

$$\begin{aligned} I_n(x) &= \sum_{m=0}^{\infty} \frac{1}{m! (m+n)!} \left(\frac{x}{2}\right)^{2m+n} \tag{A-22} \\ &\stackrel{\text{falls } n \geq 0}{=} \frac{1}{n!} \left(\frac{x}{2}\right)^n \left(1 + \frac{1}{n+1} \left(\frac{x}{2}\right)^2 + \frac{1}{2(n+1)(n+2)} \left(\frac{x}{2}\right)^4 + \dots\right). \end{aligned}$$

erhält. Daraus folgt für positives n und $x \ll n$ die **führende Ordnung**

$$I_n(x) \approx \frac{1}{n!} \left(\frac{x}{2}\right)^n. \tag{A-23}$$

Ohne Herleitung erwähnen wir noch, dass das **asymptotische Verhalten** für $x \gg n$ durch

$$I_n \approx \frac{1}{\sqrt{2\pi x}} \exp(x) \tag{A-24}$$

³Da in den den folgenden Formeln auch Fakultäten von negativen Zahlen auftauchen, benutzen wir die Konvention $\frac{1}{n!} = 0$ falls $n < 0$ (analog zu $\frac{1}{\Gamma(x+1)} = 0$ für $x = -1, -2, \dots$).

gegeben ist. Die Funktionen I_n haben also, unabhängig vom Index alle dasselbe asymptotische Verhalten.

Leiten wir die linke Seite von (A-19) nach t ab, so erhalten wir

$$\frac{d}{dt} \exp\left(\frac{x}{2}\left(t + \frac{1}{t}\right)\right) = \frac{x}{2}\left(1 - \frac{1}{t^2}\right) \exp\left(\frac{x}{2}\left(t + \frac{1}{t}\right)\right) = \frac{x}{2t}\left(t - \frac{1}{t}\right) \exp\left(\frac{x}{2}\left(t + \frac{1}{t}\right)\right)$$

und ersetzen wir nun die Exponentialfunktionen wiederum durch die rechte Seite von (A-19), ergibt sich

$$\begin{aligned} \frac{d}{dt} \sum_{n=-\infty}^{\infty} I_n(x)t^n &= \frac{x}{2t}\left(t - \frac{1}{t}\right) \sum_{n=-\infty}^{\infty} I_n(x)t^n \\ \Rightarrow \sum_{n=-\infty}^{\infty} n I_n(x)t^{n-1} &= \frac{x}{2t} \left(\sum_{n=-\infty}^{\infty} I_n(x)t^{n+1} - \sum_{n=-\infty}^{\infty} I_n(x)t^{n-1} \right) \\ \Leftrightarrow \sum_{n=-\infty}^{\infty} \frac{2n}{x} I_n(x)t^n &= \sum_{n=-\infty}^{\infty} I_n(x)t^{n+1} - \sum_{n=-\infty}^{\infty} I_n(x)t^{n-1} \end{aligned}$$

und damit durch Koeffizientenvergleich die **Rekursionsbeziehung**

$$\frac{2n}{x} I_n(x) = I_{n-1}(x) - I_{n+1}(x) \quad (\text{A-25})$$

Leiten wir dagegen (A-19) nach x ab, so entsteht durch

$$\begin{aligned} \frac{1}{2}\left(t + \frac{1}{t}\right) \exp\left(\frac{x}{2}\left(t + \frac{1}{t}\right)\right) &= \sum_{n=-\infty}^{\infty} \frac{d}{dx} I_n(x)t^n \\ \Leftrightarrow \frac{1}{2} \left(\sum_{n=-\infty}^{\infty} I_n(x)t^{n+1} + \sum_{n=-\infty}^{\infty} I_n(x)t^{n-1} \right) &= \sum_{n=-\infty}^{\infty} \frac{d}{dx} I_n(x)t^n \end{aligned}$$

die Formel für die **Ableitung** (die beiden äquivalenten Formeln erhält man unter Benutzung der Rekursion (A-25))

$$\begin{aligned} \frac{d}{dx} I_n(x) &= \frac{1}{2}(I_{n-1}(x) + I_{n+1}(x)) \\ &= I_{n-1}(z) - \frac{n}{x} I_n(x) \\ &= I_{n+1}(z) + \frac{n}{x} I_n(x) \end{aligned} \quad (\text{A-26})$$

Zur Berechnung der führenden Ordnung von $\frac{I_{n+1}}{I_n} + \frac{n}{x}$ nehmen wir zuerst $n \geq 0$ an und bekommen

$$\begin{aligned}
 \frac{I_{n+1}}{I_n} + \frac{n}{x} &= \frac{\frac{d}{dx} I_n}{I_n} = \frac{d}{dx} \ln(I_n) \\
 &= \frac{d}{dx} \ln \left(\frac{1}{n!} \left(\frac{x}{2} \right)^2 \left(1 + \frac{1}{n+1} \left(\frac{x}{2} \right)^2 + \mathcal{O}(x^4) \right) \right) \\
 &= \frac{d}{dx} \left(\ln \left(\frac{1}{n!} \right) + n(\ln(x) - \ln(2)) + \underbrace{\ln \left(1 + \frac{1}{n+1} \left(\frac{x}{2} \right)^2 + \mathcal{O}(x^4) \right)}_{\frac{1}{n+1} \left(\frac{x}{2} \right)^2 + \mathcal{O}(x^4)} \right) \\
 &= \frac{n}{x} + \frac{x}{n+1} + \mathcal{O}(x^3)
 \end{aligned}$$

Da wir bereits wissen, dass $I_n(x)$ (und damit $\frac{d}{dx} \ln(I_n)$) symmetrisch unter $n \rightarrow -n$ ist, folgern wir daraus, dass

$$\frac{I_{n+1}}{I_n} + \frac{n}{x} = \frac{|n|}{x} + \frac{z}{|n|+1} + \mathcal{O}(x^3) \tag{A-27}$$

Schließlich erhalten wir noch mit der Ersetzung $t = e^{i\theta}$ aus (A-19) die **Fourierreihe**⁴

$$\exp(x \cos \theta) = \sum_{n=-\infty}^{\infty} I_n(x) e^{in\theta}. \tag{A-28}$$

In Abbildung A.1 sind die ersten sechs der modifizierten Bessel-Funktionen dargestellt. Während die Bessel-Funktionen $J_n(x)$ oszillieren und für große Argumente asymptotisch wie $\sim 1/\sqrt{x}$ gegen Null gehen, sind die modifizierten Besselfunktionen monoton und gehen für große Argumente exponentiell gegen Unendlich.

A.4 Zusammenhang mit der Höendarstellung

Die analytischen Rechnungen in Abschnitt 3 sind im Wesentlichen Ergebnisse, die J. Krug and H.T. Dobbs 1994 in [32] im Bezug auf eindimensionale Oberflächen erhalten haben. Dort wurde allerdings eine etwas andere Darstellung verwendet, die wir im Folgenden in die unsere übersetzen werden. Der Strom j nimmt in der in [32] verwendeten Höendarstellung $h(x)$ die folgende Form an:

$$j = \frac{\sigma(h_x)}{\sqrt{g}} \left[\tilde{\gamma} \left(\frac{h_x}{\sqrt{g}} \right)_{xx} + F \right], \tag{A-29}$$

⁴In der Form $\exp(ir \cos \theta) = \sum I_n(ir) e^{in\theta} = \sum i^n J_n(r) e^{in\theta}$ (Jacobi-Anger-Identität) hat diese Formel die physikalische Interpretation der Entwicklung einer ebenen Welle nach Zylinderwellen.

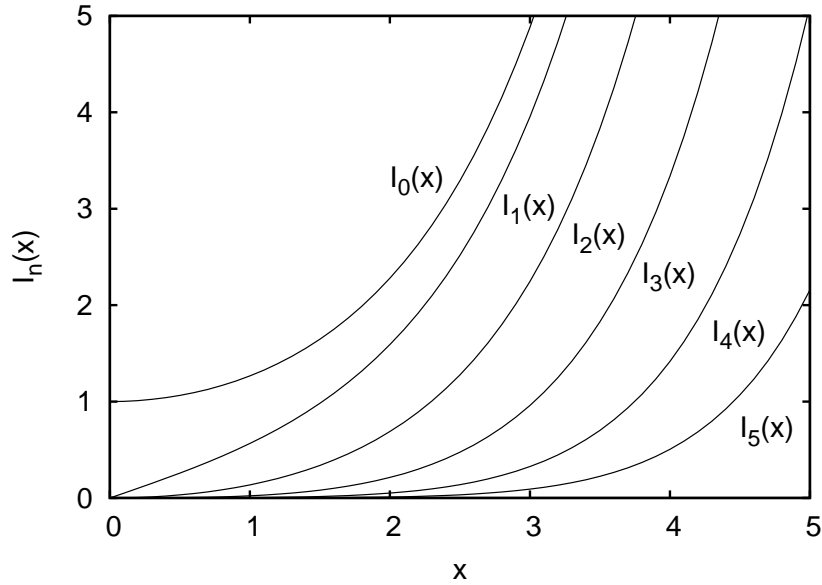


Abbildung A.1: Die modifizierten Besselfunktionen $I_n(x)$ für $n = 0 \dots 5$.

wobei der Index x die Ableitung nach x bezeichnet, g für den Ausdruck $g = 1 + (h_x)^2$ steht und die Steifigkeit $\tilde{\gamma}$ wie oben als isotrop angenommen wird. Das Feld zeigt auch hier in positive x -Richtung.

Führt man nun das Feld $\phi = h_x/\sqrt{g} = \sin\theta$ ein, wird die Gleichung $j = j^*$ wiederum äquivalent zu einer Newtonschen Bewegungsgleichung für eine "Teilchenkoordinate" ϕ , die sich in der "Zeit" x in einem "Potential" $V(\phi)$ bewegt, das mit der Beweglichkeit über

$$\tilde{\gamma}\phi_{xx} = \frac{j^*}{\sigma(\phi)\sqrt{1-\phi^2}} - F \equiv -V'(\phi) \quad (\text{A-30})$$

zusammenhängt.

Die Verbindung zu (3.2) wird nun anhand der folgenden beiden Relationen hergestellt:

$$\phi = \sin\theta \quad (\text{A-31})$$

$$\frac{d}{dx} = \frac{1}{\cos\theta} \frac{\partial}{\partial s} \quad (\text{A-32})$$

Damit wird (A-30) zu

$$\tilde{\gamma} \frac{1}{\cos\theta} \frac{\partial}{\partial s} \left(\frac{1}{\cos\theta} \frac{\partial}{\partial s} \sin\theta \right) = \frac{j^*}{\sigma(\phi)\cos\theta} - F \quad (\text{A-33})$$

$\frac{\partial}{\partial s} \sin\theta$ ist nun aber $\cos\theta \frac{\partial\theta}{\partial s}$ und damit ergibt sich nach Multiplikation mit $\cos\theta$ die Gleichung (3.2).

A.5 Die Mathieufunktionen

Die Lösungen der Mathieu-Gleichung

$$\frac{d^2 y}{dx^2} + (a - 2q \cos(2x)) y = 0 \quad (\text{A-34})$$

werden Mathieu-Funktionen genannt ([19], Kapitel 8.6). Sie wurden 1868 von Émile Léonard Mathieu im Zusammenhang mit der Untersuchung schwingender elliptischer Membranen eingeführt. Sie treten in einer Vielzahl physikalischer Probleme auf, die eine elliptische Geometrie beinhalten. Leider sind sie relativ schwer zu handhaben, da es an geeigneten analytischen Darstellungen, wie z.B. durch eine Potenzreihe fehlt. Wir werden uns hier auf die Beschreibung einiger weniger Eigenschaften beschränken, die für unsere Untersuchungen wichtig sind. Wir sind ausschließlich an *periodischen* Lösungen von (A-34) interessiert. Bei gegebenem q gibt es unendlich viele *Eigenwerte* a_n der Gleichung, für die 2π -periodische Lösungen existieren⁵. Zu jedem dieser Eigenwerte gibt es zwei linear unabhängige Lösungen, von denen jeweils eine gerade und eine ungerade in x ist⁶. Wir bezeichnen sie im Folgenden mit

$ce_n(z, q)$	$n = 0, 1, 2, \dots$	gerade
$se_n(z, q)$	$n = 1, 2, \dots$	ungerade.

Für $q = 0$ reduziert sich (A-34) auf eine einfache Schwingungsgleichung (bzw. auf $y'' = 0$ für $a = 0$). Die Grenzwerte der Funktionen sind demnach

$$\begin{aligned} \lim_{q \rightarrow 0} ce_0 &= \frac{1}{\sqrt{2}} \\ \lim_{q \rightarrow 0} ce_n &= \cos(nx) \\ \lim_{q \rightarrow 0} se_n &= \sin(nx). \end{aligned}$$

Die Mathieufunktionen sind orthogonal

⁵Diese periodischen Lösungen werden manchmal Mathieu-Funktionen erster Art genannt. Da wir uns jedoch nicht mit nichtperiodischen Lösungen beschäftigen, werden wir sie einfach Mathieu-Funktionen nennen.

⁶Ausnahme ist der Eigenwert a_0 , für den nur eine gerade Lösung existiert.

$$\int_0^{2\pi} c_n c_m dx = 0 \quad \text{falls } n \neq m$$

$$\int_0^{2\pi} s_n s_m dx = 0 \quad \text{falls } n \neq m$$

$$\int_0^{2\pi} c_n s_m dx = 0$$

und die Normierung ist⁷

$$\int_0^{2\pi} y_n^2 dx = \pi$$

A.6 Wechsel zur Fourierbasis

Die Rechnung der Stabilitätsanalyse erfolgt in der Basis^{8 9}

$$\{\mathbf{e}_n\}_{n \in \mathbb{N}_0} = \left\{ \frac{1}{2} \exp\left(\frac{z}{2} \cos(\theta)\right), \cos(\theta) \exp\left(\frac{z}{2} \cos(\theta)\right), \cos(2\theta) \exp\left(\frac{z}{2} \cos(\theta)\right), \dots \right\}.$$

Eine in vieler Hinsicht passendere Basis ist dagegen die Basis der Fouriermoden

$$\{\tilde{\mathbf{e}}_n\}_{n \in \mathbb{N}_0} = \left\{ \frac{1}{2}, \cos(\theta), \cos(2\theta), \dots \right\},$$

die auch bezüglich des L^2 -Skalarprodukts eine Orthonormalbasis bildet¹⁰. Wir werden jetzt einen Zusammenhang zwischen beiden Basen herstellen, d.h. die Transformationsmatrix T bestimmen, mit der sich die Koeffizienten der beiden Darstellungen in einander umrechnen lassen. Dazu berechnen wir die Entwicklung von \mathbf{e}_m in der Basis $\{\tilde{\mathbf{e}}_n\}_{n \in \mathbb{N}_0}$ ¹¹.

⁷Wir folgen hier der Konvention aus [19]. Nach einer anderen Konvention, die z.B. von Maple benutzt wird, ist $\lim_{q \rightarrow 0} c_{e_0} = 1$ und dementsprechend $\int_0^{2\pi} c_{e_0}^2(x, q) dx = 2\pi$.

⁸Es ist bequemer, den Index der Basisvektoren bei Null beginnen zu lassen (auf diese Weise korrespondiert der Index mit der Wellenzahl des Kosinus). Die Indexmenge ist demnach $\mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$.

⁹Der Faktor $\frac{1}{2}$ von \mathbf{e}_0 (und $\tilde{\mathbf{e}}_0$) ist ebenfalls bequem, wie aus den nachfolgenden Rechnungen klar werden wird.

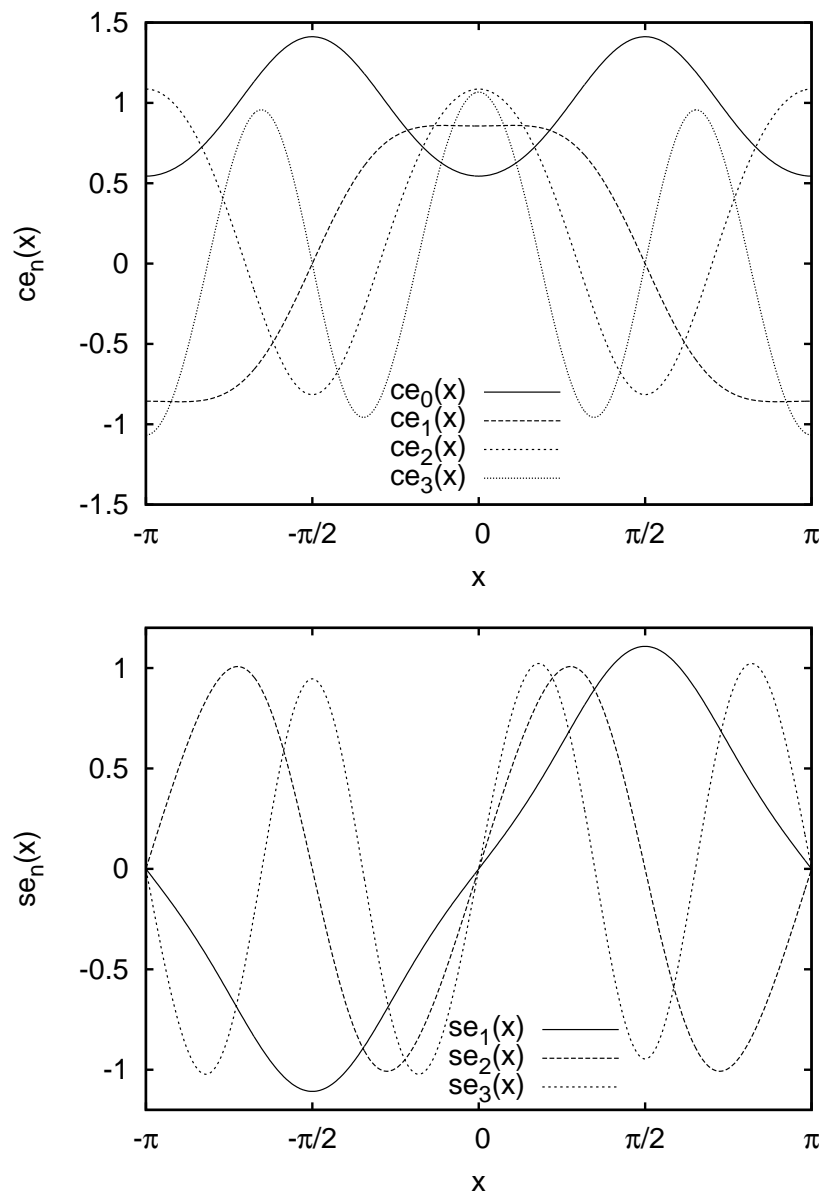
¹⁰d.h. in dieser Basis lässt sich das L^2 -Skalarprodukt auch einfach mit $\mathbf{x} \cdot \mathbf{y} = \sum x_i y_i$ berechnen

¹¹Wir benutzen

$$\exp(\alpha \cos(\theta)) = \sum I_n(\alpha) e^{in\theta}$$

und

$$\cos(m\theta) \sum a_n e^{in\theta} = \frac{1}{2} \sum (a_{n-m} + a_{n+m}) e^{in\theta}$$

Abbildung A.2: Die Mathieufunktionen ce_n und se_n für $q = 1$ und $n \leq 3$

$$\begin{aligned}
\mathbf{e}_{m \neq 0} &= \cos(m\theta) \exp\left(\frac{z}{2} \cos(\theta)\right) \\
&= \cos(m\theta) \sum I_n\left(\frac{z}{2}\right) e^{in\theta} \\
&= \frac{1}{2} \sum (I_{n-m}\left(\frac{z}{2}\right) + I_{n+m}\left(\frac{z}{2}\right)) e^{in\theta} \\
&= I_0\left(\frac{z}{2}\right) + \sum_{n=1}^{\infty} (I_{n-m}\left(\frac{z}{2}\right) + I_{n+m}\left(\frac{z}{2}\right)) \cos(n\theta) \\
&= \sum_{n=0}^{\infty} (I_{n-m}\left(\frac{z}{2}\right) + I_{n+m}\left(\frac{z}{2}\right)) e_n.
\end{aligned}$$

Umgekehrt lässt sich auch e_m in der Basis $\{\tilde{e}_n\}_{n \in \mathbb{N}_0}$ entwickeln

$$\begin{aligned}
\tilde{\mathbf{e}}_{m \neq 0} &= \cos(m\theta) = \exp\left(\frac{z}{2} \cos(\theta)\right) \cos(m\theta) \exp\left(-\frac{z}{2} \cos(\theta)\right) \\
&= \exp\left(\frac{z}{2} \cos(\theta)\right) \cos(m\theta) \sum I_n\left(-\frac{z}{2}\right) e^{in\theta} \\
&= \frac{1}{2} \exp\left(\frac{z}{2} \cos(\theta)\right) \sum (I_{n-m}\left(-\frac{z}{2}\right) + I_{n+m}\left(-\frac{z}{2}\right)) e^{in\theta} \\
&= I_m\left(-\frac{z}{2}\right) \exp\left(\frac{z}{2} \cos(\theta)\right) + \sum_{n=1}^{\infty} (I_{n-m}\left(-\frac{z}{2}\right) + I_{n+m}\left(-\frac{z}{2}\right)) \cos(n\theta) \exp\left(\frac{z}{2} \cos(\theta)\right) \\
&= \sum_{n=0}^{\infty} (I_{n-m}\left(-\frac{z}{2}\right) + I_{n+m}\left(-\frac{z}{2}\right)) \tilde{e}_n
\end{aligned}$$

Für den jeweils ersten Basisvektor \mathbf{e}_0 und $\tilde{\mathbf{e}}_0$ erhält das Ergebnis einen zusätzlichen Faktor $\frac{1}{2}$.

Die Basis $\{\tilde{\mathbf{e}}_n\}_{n \in \mathbb{N}_0}$ hat neben der Orthonormalität bezüglich des L^2 -Skalarprodukts noch einen weiteren Vorteil. Die beiden ersten Basisvektoren $\tilde{\mathbf{e}}_0$ und $\tilde{\mathbf{e}}_1$ sind die Dilatation und die Translation. Rechnet man also die Matrix $\tilde{\mathbf{A}}$ auf diese Basis um, hat sie die folgende Gestalt:

$$\tilde{\mathbf{A}} = \left(\begin{array}{cc|cc} 0 & 0 & 0 & 0 & \cdots \\ c & 0 & B_{23} & B_{24} & \cdots \\ 0 & 0 & B_{33} & B_{34} & \\ \vdots & \vdots & \vdots & & \ddots \end{array} \right) \quad (\text{A-35})$$

Die ersten beiden Spalten ergeben sich aus $\tilde{\mathbf{A}}\tilde{\mathbf{e}}_0 = c\tilde{\mathbf{e}}_1$ und $\tilde{\mathbf{A}}\tilde{\mathbf{e}}_1 = 0$, die Nullen in der ersten Zeile sind eine Folge davon, dass die Zeitentwicklung wie bereits weiter oben angemerkt, flächenerhaltend ist. Die Dilatation $\tilde{\mathbf{e}}_0$ tritt also nicht im Bild von $\tilde{\mathbf{A}}$ auf. Durch Streichen der ersten Zeile und zweiten Spalte ließe sich die Dimension der Matrix also um Eins reduzieren. Dies entspräche

einer Einschränkung auf den Raum der flächenerhaltenden Störungen. Dagegen ist die Translation unentbehrlich. Sie hängt mit der Geschwindigkeitsänderung zusammen, die im Allgemeinen mit einer Störung einhergeht. Näher beschrieben ist dieser Sachverhalt im vorangegangenen Abschnitt 9.3. Im Prinzip ließe sich die Matrix natürlich auf Blockgestalt bringen, d.h. es lässt sich ein 4×4 -Block erzeugen, der dem invarianten Unterraum entspricht, der von der Dilatation und der Translation aufgespannt wird und der vom Rest der Matrix separiert ist. Dazu wäre ein Übergang zu Basisvektoren notwendig, die die Translation nicht mehr enthalten.

Die Konstante c in (A-35) bestimmt sich wie folgt: Wir benutzen die beiden Gleichungen

$$c_1 = c_{eq}^0 \frac{\Gamma}{R_0^2} (\rho + \rho'')$$

$$\partial_t \rho = a^2 D \frac{c_1}{\xi} \mathbf{e}_x \cdot \mathbf{n}_0$$

und setzen die Dilatation $\rho \equiv \frac{1}{2}$ in die erste Gleichung ein. Das Ergebnis für c_1 verwenden wir wiederum in der zweiten und erhalten mit $\mathbf{e}_x \cdot \mathbf{n}_0 = -\cos(\theta)$

$$\partial_t \rho = -a^2 D \frac{c_{eq}^0 \Gamma}{2R_0^2 \xi} \cos(\theta).$$

Dies schreibt sich in der Basis $\{\tilde{\mathbf{e}}_n\}_{n \in \mathbb{N}_0}$ als $\tilde{\mathbf{A}}\tilde{\mathbf{e}}_0 = \frac{z}{2}\tilde{\mathbf{e}}_1$ und wir erhalten damit

$$c = \frac{z}{2}.$$

A.7 Elliptische Anfangsbedingung

Um die Normalengeschwindigkeit einer Insel in der Form einer Ellipse zu berechnen, verwenden wir die elliptischen Koordinaten

$$x = \alpha \cosh(u) \cos(v) \tag{A-36}$$

$$y = \alpha \sinh(u) \sin(v). \tag{A-37}$$

u durchläuft dabei die positiven reellen Zahlen (Radialkoordinate) und v die Winkel von 0 bis 2π (Winkelkoordinate). Die Koordinatenlinien $v = \text{const.}$ sind Parabeln und die dazu senkrechten Koordinatenlinien $u = u_0 = \text{const.}$ Ellipsen, die der Gleichung

$$\left(\frac{x}{\alpha \cosh(u_0)} \right)^2 + \left(\frac{y}{\alpha \sinh(u_0)} \right)^2 = 1$$

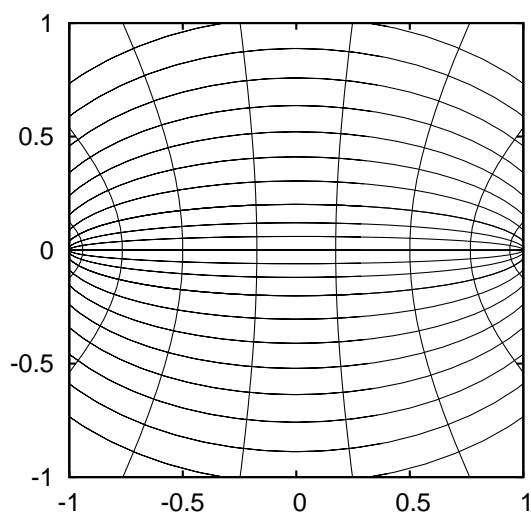


Abbildung A.3: Elliptische Koordinatenlinien.

genügen. Das heißt, das Aspektverhältnis beträgt $\tanh(u_0)$ und läuft damit von 0 (Fall einer zu einer Linie entarteten Ellipse) bis 1 (Kreis), je weiter sich die Radialkoordinate vom Ursprung entfernt. Der Parameter α skaliert das Koordinatensystem und damit die Größe der Ellipse.

Die Koordinatenvektoren sind

$$\begin{aligned}\vec{\xi} &= \partial_u \vec{x} = \alpha \begin{pmatrix} \sinh(u) \cos(v) \\ \cosh(u) \sin(v) \end{pmatrix} \\ \vec{\eta} &= \partial_v \vec{x} = \alpha \begin{pmatrix} -\cosh(u) \sin(v) \\ \sinh(u) \cos(v) \end{pmatrix}\end{aligned}$$

beide haben Norm $\alpha\sqrt{g}$ wobei g

$$\begin{aligned}g &= \sinh^2(u) \cos^2(v) + \cosh^2(u) \sin^2(v) \\ &= \sinh^2(u) + \sin^2(v) \\ &= \cosh^2(u) - \cos^2(v)\end{aligned}$$

ist.

Lösung der Helmholtz Gleichung

Die Gleichung lautet in den Koordinaten u, v

$$(\partial_u^2 + \partial_v^2)f = \alpha^2(\sinh^2(u) + \sin^2(v))f$$

Der Separationsansatz $f(u, v) = f_u(u)f_v(v)$ ergibt

$$f_v'' - \alpha^2 \sin^2(v)f_v = \lambda f_v \quad (\text{A-38})$$

$$-f_u'' + \alpha^2 \sinh^2(u)f_u = \lambda f_u \quad (\text{A-39})$$

mit dem Separationsparameter λ . Die beiden Gleichungen werden durch $u = iv$ ineinander transformiert. Es genügt also die Lösung für $f_u(u)$ zu bestimmen. Mit Hilfe der Beziehung $\sin^2(v) = \frac{1}{2}(1 - \cos(2v))$ wird aus Gleichung (A-39)

$$f_u'' + \underbrace{\left(-\lambda - \frac{\alpha^2}{2}\right)}_{=:a} + \underbrace{\frac{\alpha^2}{2} \cos(2v)}_{=: -2q} f_u = 0$$

Das ist die Mathieugleichung (siehe Abschnitt A.5). Periodische Lösungen existieren nur für bestimmte *Eigenwerte* a_n der Gleichung. Wir benötigen hier nur die symmetrischen Funktionen ce_n .

Die allgemeine Lösung für das Konzentrationsfeld lautet also¹²

$$c(u, v) = \exp\left(\frac{\alpha}{2\xi} \cosh(u) \cos(v)\right) \underbrace{\sum_{n=0}^{\infty} b_n ce_n(v) ce_n(iu)}_{f(u,v)} \quad (\text{A-40})$$

Die Krümmung einer Koordinatenlinie $u = u_0 = \text{const.}$ ist¹³

$$\kappa = -\frac{\cosh(u_0) \sinh(u_0)}{\alpha g^{\frac{3}{2}}}$$

Die (linearisierte) Randbedingung ist

$$c|_{\partial\Omega} = c(u_0, v) = c_{eq}^0 (1 + \Gamma\kappa)$$

und damit wird aus (A-40)

$$\sum_{n=0}^{\infty} b_n ce_n(v) ce_n(iu) = c_{eq}^0 \exp\left(-\frac{\alpha}{2\xi} \cosh(u_0) \cos(v)\right) (1 + \Gamma\kappa).$$

Bildet man nun auf beiden Seiten das Skalarprodukt mit ce_n ergibt sich

¹²Wir benutzen $ce_n(-iu) = ce(iu)$

¹³Das negative Vorzeichen kommt wieder daher, dass es sich um eine Leerstelleninsel handelt.

$$b_n = \frac{c_{eq}^0}{\pi c e_n(iu)} \int_0^{2\pi} c e_n(v) \exp\left(-\frac{\alpha}{2\xi} \cosh(u) \cos(v)\right) (1 + \Gamma\kappa) dv$$

Damit ist das Konzentrationsfeld festgelegt. Die Normalengeschwindigkeit ergibt sich nun durch den Strom zum Inselrand. Für den Normalenvektor des Randes gilt

$$\begin{aligned}\vec{n} &= -\frac{1}{\alpha\sqrt{g}}\vec{\xi} \\ \vec{n} \cdot \nabla &= \partial_{\vec{n}} = -\frac{1}{\alpha\sqrt{g}}\partial_u \\ \vec{n} \cdot \vec{e}_x &= -\frac{1}{\sqrt{g}}\sinh(u_0)\cos(v).\end{aligned}$$

Damit ist die Normalengeschwindigkeit¹⁴

$$\begin{aligned}v_n &= -a^2\vec{n} \cdot \vec{j} \\ &= a^2 D \exp\left(\frac{\alpha}{2\xi} \cosh(u_0) \cos(v)\right) \vec{n} \cdot \left(\nabla f - \frac{1}{2\xi} f \vec{e}_x\right) \\ &= -\frac{a^2 D}{\sqrt{g}} \exp\left(\frac{\alpha}{2\xi} \cosh(u_0) \cos(v)\right) \left(\frac{1}{\alpha} \partial_u f - \frac{1}{2\xi} \sinh(u_0) \cos(v) f\right)\end{aligned}$$

Die Ableitung von f berechnet sich nach

$$\begin{aligned}f(u, v) &= \sum b_n c e_n(v) c e_n(iu) \\ \partial_u f(u, v) &= \sum b_n c e_n(v) \underbrace{ic e'_n(iu)}_{-\text{Im}(c e'_n(iu))}\end{aligned}$$

demnach ist die Normalengeschwindigkeit

$$\begin{aligned}v_n &= \frac{a^2 D}{\sqrt{g}} \exp\left(\frac{\alpha}{2\xi} \cosh(u_0) \cos(v)\right) \sum b_n c e_n(v) \left(\frac{1}{\alpha} \text{Im}(c e'_n(iu))\right) \\ &\quad + \frac{1}{2\xi} \sinh(u_0) \cos(v) c e_n(iu).\end{aligned}$$

¹⁴Wir benutzen für $\nabla c - \frac{1}{\xi} c \vec{e}_x$

$$\begin{aligned}\nabla \left(\exp\left(\frac{x}{2\xi}\right) f\right) - \frac{1}{\xi} \exp\left(\frac{x}{2\xi}\right) f \vec{e}_x &= f \frac{1}{2\xi} \exp\left(\frac{x}{2\xi}\right) \vec{e}_x + \exp\left(\frac{x}{2\xi}\right) \nabla f - \frac{1}{\xi} f \exp\left(\frac{x}{2\xi}\right) \vec{e}_x \\ &= \exp\left(\frac{x}{2\xi}\right) \left(\nabla f - \frac{1}{2\xi} f \vec{e}_x\right)\end{aligned}$$

Sie enthält natürlich noch die Drift der gesamten Insel. Um diesen Anteil herauszuprojizieren, benutzen wir

$$\cos(\theta) = \vec{n} \cdot \vec{e}_x = -\frac{1}{\sqrt{g}} \sinh(u_0) \cos(v)$$

und

$$\begin{aligned} ds &= \|\vec{\eta}\| dv = \alpha \sqrt{g} dv \\ \kappa &= \frac{d\theta}{ds} = -\frac{\cosh(u_0) \sinh(u_0)}{\alpha g^{\frac{3}{2}}} \\ \Rightarrow d\theta &= \kappa ds = -\frac{\cosh(u_0) \sinh(u_0)}{g} dv \end{aligned}$$

Der Schwerpunkt bestimmt sich nach

$$\vec{r}_s = \frac{\int_{\partial\Omega} \vec{r} ds}{\int_{\partial\Omega} ds}$$

Die Driftgeschwindigkeit ist nun

$$\begin{aligned} \vec{v} &= \frac{d}{dt} \vec{r}_s = \frac{\left(\int_{\partial\Omega} ds \right) \left(\int_{\partial\Omega} \frac{d\vec{r}}{dt} ds + \int_{\partial\Omega} \vec{r} \partial_t ds \right) - \left(\frac{d}{dt} \int_{\partial\Omega} ds \right) \left(\int_{\partial\Omega} \vec{r} ds \right)}{\left(\int_{\partial\Omega} ds \right)^2} \\ \Rightarrow v_x &= \vec{v} \cdot \vec{e}_x = \frac{1}{\int_{\partial\Omega} ds} \left(\int_{\partial\Omega} \frac{d\vec{r}}{dt} \cdot \vec{e}_x ds + \int_{\partial\Omega} x \partial_t ds \right) - \frac{1}{\left(\int_{\partial\Omega} ds \right)^2} \left(\frac{d}{dt} \int_{\partial\Omega} ds \right) \left(\int_{\partial\Omega} x ds \right) \end{aligned}$$

mit $\frac{d}{dt} \vec{r} = \vec{v}_n \Rightarrow \frac{d}{dt} \vec{r} \cdot \vec{e}_x = v_n \vec{n} \cdot \vec{e}_x = v_n \cos(\theta)$ und $\partial_t ds = v_n d\theta = v_n \kappa ds$ ist das

$$v_x = \frac{1}{\int_{\partial\Omega} ds} \left(\int_{\partial\Omega} v_n \cos(\theta) ds + \int_{\partial\Omega} x v_n \kappa ds \right) - \frac{1}{\left(\int_{\partial\Omega} ds \right)^2} \left(\int_{\partial\Omega} v_n \kappa ds \right) \left(\int_{\partial\Omega} x ds \right).$$

Es gilt $\int_{\partial\Omega} x ds = 0$, da bei der Anfangskonfiguration der Schwerpunkt im Ursprung liegt und damit ist schließlich

$$v_x = \frac{1}{\int_{\partial\Omega} \sqrt{g} dv} \left(-\sinh(u_0) \int_{\partial\Omega} v_n \cos(v) dv + \alpha \cosh(u_0) \int_{\partial\Omega} \cos(v) v_n \kappa \sqrt{g} dv \right).$$

Danksagung

Die Fertigstellung der vorliegenden Dissertation gibt mir reichlich Grund und Gelegenheit, den Menschen zu danken, die direkt oder indirekt daran beteiligt waren.

An erster Stelle steht hier natürlich Joachim Krug, von dem ich in den Jahren, in denen er meine Arbeit betreut hat, so viel lernen konnte. Des Weiteren gilt mein Dank Axel Voigt und Frank Haüßer, die am größten Teil der hier dargestellten Ergebnisse beteiligt waren. Insbesondere die enge Zusammenarbeit mit Frank war ein unschätzbare Glücksfall, der die Form dieser Arbeit entscheidend geprägt hat.

Ich danke Marko Rusanen, dessen Fertigkeiten im Umgang mit KMC-Simulationen das Projekt in diese Richtung abgerundet und bereichert haben und Martin Schimschak der die ursprüngliche Version des Algorithmus entwickelt hat und mir in der Anfangsphase behilflich war, ihn auf meinem Rechner zum Laufen zu bringen.

Su-Chan Park bin ich zu tiefem Dank verpflichtet für seine stetige Diskussionsbereitschaft und seine Hilfestellungen bei allen möglichen physikalischen und mathematischen Fragen. Zu danken habe ich des Weiteren Kavita Jain, mit der ich viele anregende Gespräche geführt habe. Wertvolle Diskussionspartner waren auch Slava Popkov, Olivier Pierre-Louis, Gerhard Wolf, Stephen Watson und Ute Ebert. Stefan Thoma habe ich die Literaturhinweise aus dem Gebiet der Hydrodynamik zu verdanken und Vakhtang Putkaradze die Anregung, es einmal mit einer Ellipse zu versuchen.

Herzlichen Dank schulde ich auch Barbara Sandhöfer, die einer der Hauptgründe war, dass ich mich hier am Institut so wohlfühlt habe und Hamed Alaei, der immer da war, wenn ich ihn gebraucht habe.

Dieses Projekt war bis 2005 Teil des Sonderforschungsbereichs 616 „Energie-dissipation an Oberflächen“, dessen Mitgliedern ich danken möchte, ebenso wie der Deutschen Forschungsgemeinschaft für die finanzielle Förderung.

Menschen, die nicht aus meinem wissenschaftlichen Umfeld stammen, mich aber doch tief beeinflusst haben, sind Johannes Kopp, dem ich so viel von dem verdanke, das sich schwer in Worte fassen läßt und mein Großvater, der unbeabsichtigt mein Interesse an der Wissenschaft ausgelöst hat.

Und schließlich und am Allerwichtigsten sind da noch meine Eltern, denen ich buchstäblich alles zu verdanken habe.

Literaturverzeichnis

- [1] E. Anderson, Z. Bai, C. Bischof, J. W. Demmel, J. J. Dongarra, J. J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, S. Ostrouchov, and D. C. Sorensen. *LAPACK Users' Guide*. SIAM, Philadelphia, 1999.
- [2] J.S. Baggett, T.A. Driscoll, and L.N. Trefethen. A mostly linear model of transition to turbulence. *Phys. Fluids*, 7:833–838, 1995.
- [3] E. Bänsch, F. Haußer, O. Lakkis, B. Li, and A. Voigt. Finite element method for epitaxial growth with attachment-detachment kinetics. *J. Comput. Phys.*, 194(2):409–434, 2004.
- [4] E. Bänsch, F. Haußer, and A. Voigt. Finite element method for epitaxial growth with thermodynamic boundary conditions. *SIAM Journal on Scientific Computing*, 26(6):2029–2046, 2005.
- [5] J.R. Black. Mass transport of aluminum by momentum exchange with conducting electrons. In *Proceedings of the 6th Annual International Reliability Physics Symposium*, pages 148–159. IEEE, New York, 1967.
- [6] I.A. Blech and E.S. Meieran. Electromigration in thin Al films. *J. Appl. Phys.*, 40:485–491, 1969.
- [7] I.A. Blech and E.S. Meieran. Direct transmission electron microscope observation of electrotransport in aluminum. *Appl. Phys. Lett.*, 47:1203–1208, 1976.
- [8] J.L. Blue, I. Beichl, and F. Sullivan. Faster Monte Carlo simulations. *Phys. Rev. E*, 51(2):R867–R868, Feb 1995.
- [9] O. Bondarchuk, W. G. Cullen, M. Degawa, and Ellen D. Williams. Biased structural fluctuations due to electron wind force (pre-print), 2007.
- [10] A.B. Bortz, M.H. Kalos, and J.L. Lebowitz. A new algorithm for Monte Carlo simulation of ising spin systems. *J. Comput. Phys.*, 17(10):10–18, 1975.

- [11] W.K. Burton, N. Cabrera, and F. Frank. The growth of crystals and the equilibrium structure of their surfaces. *Phil. Trans. R. Soc. London*, 243:299, 1951.
- [12] M.C. Cross and P.C. Hohenberg. Pattern formation outside of equilibrium. *Rev. Mod. Phys.*, 65(3):851–1112, 1993.
- [13] P.G. Drazin and W. H. Reid. *Hydrodynamic Stability*. Cambridge University Press, New York, 1981.
- [14] M. Dufay, J.-M. Debierre, and T. Frisch. Electromigration-induced step meandering on vicinal surfaces: Nonlinear evolution equation. *Physical Review B (Condensed Matter and Materials Physics)*, 75(4):045413, 2007.
- [15] K. Fujita, M. Ichikawa, and S.S. Stoyanov. Size-scaling exponents of current-induced step bunching on silicon surfaces. *Phys. Rev. B*, 60(23):16006–160012, 1999.
- [16] T. Gebhardt and S. Grossmann. Chaos transition despite linear stability. *Phys. Rev. E*, 50(5):3705–3711, Nov 1994.
- [17] M. Gerardin. De l'action de la pile sur les sels de potasse et de soude et sur les alliages soumis á la fusion igne. *C.R. Acad. Sci. Paris*, 53:727–731, 1861.
- [18] M. Giesen. Step and island dynamics at solid/vacuum and solid/liquid interfaces. *Prog. Surf. Sci.*, 68:1–154, 2001.
- [19] I.S. Gradshteyn and I.M. Ryshik. *Table of Integrals, Series and Products*. Academic Press, San Diego, 2000.
- [20] S. Grossmann. The onset of shear flow turbulence. *Rev. Mod. Phys.*, 72:603–618, 2000.
- [21] M.R. Gungor and D. Maroudas. Current-induced non-linear dynamics of voids in metallic thin films: morphological transition and surface wave propagation. *Surf. Sci.*, 461:550–556, 2000.
- [22] H. Haken. *Synergetics, an introduction. Nonequilibrium phase transitions and self-organization in physics, chemistry and biology*. Springer, Berlin, 1983.
- [23] F. Haußer, P. Kuhn, J. Krug, and A. Voigt. Morphological stability of electromigration-driven vacancy islands. *Phys. Rev. E*, 75:046210, 2007.
- [24] J. Heinonen, I. Koponen, J. Merikoski, and T. Ala-Nissila. Island diffusion on metal fcc (100) surfaces. *Phys. Rev. Lett.*, 82(13):2733–2736, Mar 1999.

- [25] P. Henrici. *Applied and Computational Complex Analysis, volume 1*. Wiley, New York, 1974.
- [26] C. Herring. Effect of change of scale on sintering phenomena. *J. Appl. Phys.*, 21:301–303, 1950.
- [27] P.S. Ho. Motion of an inclusion induced by a direct current and a temperature gradient. *J. Appl. Phys.*, 41:64–68, 1970.
- [28] K.W. Jacobsen, J.K. Norskov, and M.J. Puska. Interatomic interactions in the effective-medium theory. *Phys. Rev. B*, 35(14):7423–7442, May 1987.
- [29] R.P. Johnson. Construction of filament surfaces. *Phys. Rev.*, 54:459–467, 1938.
- [30] J. Kallunki and J. Krug. Effect of kink-rounding barriers on step edge fluctuations. *Surf. Sci. L.*, 523:L53–L58, 2003.
- [31] J. Krug. Introduction to step dynamics and step instabilities. In A. Voigt, editor, *Multiscale modeling of epitaxial growth*, pages 69–95. Birkhäuser, Basel, 2005.
- [32] J. Krug and H.T. Dobbs. Current-induced faceting of crystal surfaces. *Phys. Rev. Lett.*, 73:1947–1950, 1994.
- [33] P. Kuhn and J. Krug. Islands in the stream: Electromigration-driven shape evolution with crystal anisotropy. In A. Voigt, editor, *Multiscale modeling of epitaxial growth*, pages 159–173. Birkhäuser, Basel, 2005.
- [34] P. Kuhn, J. Krug, F. Haußer, and A. Voigt. Complex shape evolution of electromigration-driven single-layer-islands. *Phys. Rev. Lett.*, 94:166105, 2005.
- [35] A.V. Latyshev, A.L. Aseev, A.B. Krasilnikov, and S.I. Stenin. Transformations of clean si(111) stepped surface during sublimation. *Surf. Sci.*, 213:157–169, 1989.
- [36] D.-J. Liu and J. W. Evans. Sintering of two-dimensional nanoclusters in metal(100) homoepitaxial systems: Deviations from predictions of mullins continuum theory. *Phys. Rev. B*, 66(16):165407, Oct 2002.
- [37] H. Mehl, O. Biham, O. Millo, and M. Karimi. Electromigration-induced flow of islands and voids on the Cu(001) surface. *Phys. Rev. B*, 61(7):4975–4982, Feb 2000.
- [38] J.-J. Métois, J. c. Heyraud, and A. Pimpinelli. Steady-state motion of silicon islands driven by a DC current. *Surf. Sci.*, 420:075409, 1999.

- [39] T. Michely and J. Krug. *Islands, Mounds and Atoms*. Springer, Berlin, 2004.
- [40] H. Minoda. Direct current heating effect on si(111) vicinal surfaces. *J. Phys.: Condens. Matter*, 15:3255–3280, 2003.
- [41] G. Nicolis and I. Prigogine. *Self-organization in Nonequilibrium Systems, from dissipative structures to order through fluctuations*. Wiley, New York, 1977.
- [42] E. Ott. *Chaos in Dynamical Systems*. Cambridge University Press, Cambridge, 2002.
- [43] A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Springer, New York, 1983.
- [44] O. Pierre-Louis and T. L. Einstein. Electromigration of single-layer clusters. *Phys. Rev. B*, 62(20):13697–13706, Nov 2000.
- [45] A. Pimpinelli and J. Villain. *Physics of crystal growth*. Cambridge University Press, Cambridge, 1998.
- [46] V. Popkov and J. Krug. Dynamic phase transitions in electromigration-induced step bunching. *Phys. Rev. B*, 73:235430, 2006.
- [47] W. Press, B. Flannery, S. Teukolsky, and W. Vetterling. *Numerical Recipes*. Cambridge University Press, New York, 1986.
- [48] A. Quarteroni, R. Sacco, and F. Saleri. *Numerical Mathematics*. Springer, New York, 2000.
- [49] M. Rusanen, P. Kuhn, and J. Krug. Kinetic Monte Carlo simulations of oscillatory shape evolution for electromigration-driven islands. *Physical Review B (Condensed Matter and Materials Physics)*, 74(24):245423, 2006.
- [50] A. Saúl, J.-J. Métois, and A. Ranguis. Experimental evidence for an ehrlich-schwoebel effect on si(111). *Phys. Rev. B*, 65:157–169, 2002.
- [51] M. Schimschak. *Numerische Untersuchungen zur Elektromigration auf metallischen Oberflächen*. PhD thesis, Universität Essen, 1999.
- [52] M. Schimschak and J. Krug. Electromigration-induced breakup of two-dimensional voids. *Phys. Rev. Lett.*, 80:1674–1677, 1998.
- [53] M. Schimschak and J. Krug. Electromigration-driven shape evolution of two-dimensional voids. *J. Appl. Phys.*, 87:695–703, 2000.
- [54] W. Seith and H. Wever. Neue Ergebnisse bei der Elektrolyse fester metallischer Phasen. *Z. Elektrochem.*, 59(10):942–946, 1955.

- [55] F. Skaupy. Die Elektrizitätsleitung in flüssigem Quecksilber. *Verh. Deut. Phys. Ges.*, 16:156–167, 1914.
- [56] T.J. Stasevich and T.L. Einstein. Analytic formulas for the orientation dependence of step stiffness and line tension: Key ingredients for numerical modeling. *Multiscale Model. Simul.*, 6(1):90–104, 2007.
- [57] S.S. Stoyanov. Electromigration induced step bunching on si surfaces - how does it depend on the temperature and heating current direction? *Jap. J. Appl. Phys.*, 30:1–6, 1991.
- [58] A. Straub. *Einflußfaktoren auf das kritische Produkt in der Elektromigration*. PhD thesis, Universität Stuttgart, 2000.
- [59] Z. Suo. Electromigration-induced dislocation climb and multiplication in conducting lines. *Acta metall. mater.*, 42(11):3581–3588, 1994.
- [60] Z. Suo, W. Wang, and M. Yang. Electromigration instability: Transgranular slits in interconnects. *Appl. Phys. Lett.*, 64(15):1944–1946, 1994.
- [61] I.J. Thompson and A.R. Barnett. Coulomb and Bessel functions of complex arguments and order. *Journal of Computational Physics*, 64(2):490–509, 1986.
- [62] A. Torcini and P. Politi. Coarsening process in one-dimensional surface growth models. *Eur. Phys. J. B*, 25:519–529, 2002.
- [63] L. N. Trefethen, A. E. Trefethen, S. C. Reddy, and T. A. Driscoll. Hydrodynamic stability without eigenvalues. *Science*, 261:578–585, 1993.
- [64] L.N. Trefethen and M. Embree. *Spectra and Pseudospectra*. Princeton University Press, 2005.
- [65] A.M. Turing. The chemical basis of morphogenesis. *Phil. Trans. R. Soc. London*, 237:37, 1952.
- [66] Ed. V.I. Krinsky. *Self-organization, autowaves and structures far from equilibrium*. Springer, Berlin, 1984.
- [67] F. Waleffe. Transition in shear flows: non-linear normality versus non-normal linearity. *Phys. Fluids*, 2:3060–3066, 1995.
- [68] W. Wang, Z. Suo, and T.-H. Hao. A simulation of electromigration-induced transgranular slits. *J. Appl. Phys.*, 79:2394–2403, 1996.
- [69] K. Yagi, H. Minoda, and M. Degawa. Step bunching, step wandering and faceting: self-organization at si surfaces. *Surf. Sci. Rep.*, 43:45–126, 2001.
- [70] W. Yang, W. Wang, and Z. Suo. Cavity and dislocation instability due to electric current. *J. Mech. Phys. Solids*, 42(6):897–911, 1994.

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 der Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Herrn Prof. Dr. Joachim Krug betreut worden.

Philipp Kuhn

Köln, den 15. Mai 2007

Teilpublikationen

- P. Kuhn und J. Krug
Islands in the stream: Electromigration-driven shape evolution with crystal anisotropy.
In *Multiscale Modeling of Epitaxial Growth*, ed. by A. Voigt, International Series of Numerical Mathematics Vol. 149 (Birkhäuser 2005), pp. 159-173.
- P. Kuhn, J. Krug, F. Haußer und A. Voigt
Complex shape evolution of electromigration-driven single-layer islands.
Physical Review Letters **94**, 166105 (2005).
- M. Rusanen, P. Kuhn und J. Krug
Kinetic Monte Carlo simulations of oscillatory shape evolution for electromigration-driven islands.
Physical Review B **74**, 245423 (2006).
- F. Haußer, P. Kuhn, J. Krug und A. Voigt
Morphological stability of electromigration-driven vacancy islands.
Physical Review E **75**, 145703 (2007).