

Freie Universität  Berlin

ASYMPTOTISCH SCHNELLE
ÜBERGÄNGE UND DER
STEIFHEITSBEGRIFF BEI
GEWÖHNLICHEN
DIFFERENTIALGLEICHUNGEN

Inauguraldissertation
zur Erlangung des Grades eines
Doktors der Naturwissenschaften (Dr. rer. nat.)
am Fachbereich Mathematik und Informatik
der Freien Universität Berlin
vorgelegt von

PATRIK MARSCHALIK

Berlin, 2014

Betreuer und erster Gutachter:
Prof. Dr. Rupert Klein
Institut für Mathematik
Fachbereich Mathematik und Informatik
Freie Universität Berlin
Arnimallee 6
14195 Berlin
Telefon: +49 30 838 75414
Fax: +49 30 838 75412
e-mail: rupert.klein@math.fu-berlin.de

Zweite Gutachterin:
Prof. Dr. Barbara Wagner
Institut für Mathematik
Technische Universität Berlin
Straße des 17. Juni 136
10623 Berlin
Telefon: +49 30 314 22246
Fax: +49 30 314 21110
e-mail: bwagner@math.fu-berlin.de

Tag der Disputation: 16. März 2015

Um der geltenden Promotionsordnung zu genügen, findet sich die hier zu erwartende Zusammenfassung im Anhang auf Seite [123](#).

DANKSAGUNG

An erster Stelle geht mein Dank an Prof. Dr. Rupert Klein für eine intensive kontinuierliche Betreuung, viele fruchtbare Diskussionen und sehr viel Verständnis.

Bei Prof. Dr. Barbara Wagner möchte ich mich für Ihre Bereitschaft bedanken, meine Dissertation zu begutachten.

Weiter möchte ich Prof. Dr. Peter Spichtinger danken, der mir die Möglichkeit gegeben hat, vor Ende der Promotion eine PostDoc-Stelle anzutreten, und der sehr geduldig auf die Fertigstellung der Arbeit gewartet hat.

Besonderer Dank geht an Martin Götze für viele wertvolle Stunden Diskussion und kritische Anmerkungen.

Für außerordentlich gründliches Korrekturlesen und hilfreiche Kommentare danke ich Linda Michalk, Maya Gussmann und Manuel Baumgartner.

Für das Finden von Orthographie- und Interpunktionsfehlern danke ich außerdem Franz Marschalik.

INHALTSVERZEICHNIS

i	SINGULÄRE STÖRUNGSTHEORIE EINES ATMOSPHERISCHEN WIRBELS	1
1	EINLEITUNG	3
2	ASYMPTOTIK IN DER ATMOSPHERE	5
2.1	Einheitlicher Ansatz für meteorologische Phänomene	5
2.2	Wahl des Gleichungssystems	6
2.3	Dimensionsanalyse	7
2.4	Gekoppelter Grenzübergang	10
2.5	Wie groß ist ε ?	12
2.6	Asymptotische Ansatzfunktionen	14
3	ASYMPTOTIK EINES ATMOSPHERISCHEN WIRBELS	17
3.1	Ausgangspunkt	18
3.1.1	Beschreibung der quasigeostrophischen Hintergrundströmung	18
3.1.2	Beschreibung des Wirbels	19
3.2	Quasigeostrophische Hintergrundströmung	21
3.3	Wirbelkernstruktur	23
3.4	Matching	25
3.5	Alternative Bestimmung der Stromfunktion	27
3.5.1	Lokales mitbewegtes orthogonales Koordinatensystem	27
3.5.2	Vorgehen zum Lösen der elliptischen Gleichung von ψ	28
ii	STEIFES VERHALTEN VON DIFFERENTIALGLEICHUNGEN	29
1	EINLEITUNG	31
2	MOTIVATION	33
2.1	Steife Differentialgleichungen im klassischen Sinne	33
2.2	Familien von Differentialgleichungen	41
2.3	Formalisierung steifer Differentialgleichungen	55
3	STEIFE LÖSUNGSFAMILIEN	59
3.1	Langsame Mannigfaltigkeiten	60
3.1.1	Langsame Mannigfaltigkeit linearer Probleme	60
3.1.2	Langsame Mannigfaltigkeiten nichtlinearer Probleme	64
3.2	Exponentiell anziehende Mengen	67
3.3	Steife Familien von Lösungen	71
3.4	Steife Lösungsfamilien linearer Probleme	73
3.4.1	Invarianz unitärer Transformationen	75
3.4.2	Steife Lösungen von Schurformen	77
3.4.3	Normale und diagonale Probleme	85

3.4.4	Steife Lösungsfamilien linearer Probleme und explizite numerische Verfahren	87
3.5	Steife Lösungsfamilien nichtlinearer Probleme	88
3.5.1	Probleme beim Einfrieren von Koeffizienten	89
3.5.2	Inhomogene lineare Probleme	91
3.5.3	Allgemeine nichtlineare Probleme	97
4	DISKUSSION	101
4.1	Lineare, nichtlineare und intermediäre-Theorie	102
4.1.1	Lineare Theorie	102
4.1.2	Nichtlineare Theorie	103
4.1.3	Intermediäre Theorie	107
4.1.4	Zusammenhang der drei Theorien	109
4.2	Steifes Verhalten und Kondition	110
4.2.1	Charakterisierung steifer Lösungen durch die Kondition	110
4.2.2	Kondition steifer Probleme	113
4.3	Singulär gestörte Probleme	115
4.3.1	Charakterisierung steifen Verhaltens durch singular gestörte Probleme	115
4.3.2	Wahl des Störungsparameters	116
5	FAZIT UND AUSBLICK	117
iii	ANHANG	121
	LITERATUR	125

Teil I

SINGULÄRE STÖRUNGSTHEORIE EINES
ATMOSPHERISCHEN WIRBELS

1

EINLEITUNG

Physikalische Vorgänge in der Atmosphäre sind von enormer Bedeutung. Sie beeinflussen das Klima und sind damit von großer gesellschaftliche Relevanz. Die Beschreibung dieser Vorgänge in ihrer Allgemeinheit scheint unmöglich. Es gibt jedoch verschiedene Möglichkeiten, diese Vorgänge näherungsweise zu beschreiben.

In Kapitel [2](#) auf Seite [5](#) wird eine solche Möglichkeit vorgestellt. Ziel ist es, einen transparenten Weg aufzuzeigen von den komplizierten Navier-Stokes-Gleichungen zu vereinfachten asymptotischen Näherungen zu gelangen.

In Kapitel [3](#) auf Seite [17](#) wird dieses Vorgehen genutzt, um einen konzentrierten, beinahe achsensymmetrischen, stark geneigten, baroklinen Wirbel im Gradient-Wind-Regime, der in eine quasigeostrophische Hintergrundströmung mit schwacher vertikaler Scherung eingebettet ist, zu analysieren.

2

ASYMPTOTIK IN DER ATMOSPHERE

Inhalt

2.1	Einheitlicher Ansatz für meteorologische Phänomene	5
2.2	Wahl des Gleichungssystems	6
2.3	Dimensionsanalyse	7
2.4	Gekoppelter Grenzübergang	10
2.5	Wie groß ist ε ?	12
2.6	Asymptotische Ansatzfunktionen	14

In diesem Kapitel wird ein einheitlicher asymptotischer Ansatz vorgestellt, atmosphärische Strömung zu beschreiben. Im Wesentlichen wird sich auf die Arbeiten Klein (2010, 2008) bezogen.

Ziel dieses Ansatzes ist es, ein einheitliches Gerüst zur Herleitung meteorologischer Modelle zu schaffen. Die Mächtigkeit dieses Ansatzes wird besonders bei Mehrskalenproblemen deutlich. Bei Klein (2010) wird dies für mehrere gängige Modelle gezeigt. Im nächsten Kapitel wird dieser Ansatz genutzt, Bewegung und Struktur eines atmosphärischen baroklinen trockenen Wirbels zu untersuchen (Päsche u. a., 2012).

2.1 EINHEITLICHER ANSATZ FÜR METEOROLOGISCHE PHÄNOMENE

Werden in der klassischen Meteorologie atmosphärische Phänomene untersucht, startet man selten mit den vollen kompressiblen dreidimensionalen Navier-Stokes-Gleichungen. Stattdessen werden Annahmen getroffen und dazu passende einfachere Gleichungen gewählt. Die Gleichungen werden mit zu den Phänomen passenden charakteristischen Größen dimensionslos gemacht. Lassen sich nach der Entdimensionalisierung kleine dimensionslose Größen identifizieren, können die Gleichungen mit den Methoden der Asymptotik analysiert werden. Dieses Vorgehen wird für jedes einzelne Phänomen durchgeführt.

Bei dem hier vorgestellten Ansatz (Klein, 2010) werden nicht einzelne Phänomene betrachtet, sondern die Atmosphäre als Ganzes. Ausgegangen wird von den vollen kompressiblen dreidimensionalen Navier-Stokes-Gleichungen. Zur Entdimensionalisierung dieser Gleichungen werden allgemeine atmosphärische Referenzgrößen gewählt (Tabel-

le 2.1 auf Seite 9). Man erhält ein dimensionsloses Gleichungssystem, in dem drei kleine Parameter identifiziert werden können, die mittels eines gekoppelten Grenzübergangs (distinguished limit) an einen einzigen kleinen Parameter ε gekoppelt werden. Das daraus resultierende asymptotische Gleichungssystem wird als Ausgangspunkt der Analyse gewählt.

Ziel der Analyse ist, für bestimmte Regime das dominante Verhalten aus den vollen kompressiblen dreidimensionalen Gleichungen zu extrahieren. Durch Skalierung der Koordinaten und geeigneter Ansatzfunktionen lassen sich dann die in dem gewählten Regime geltenden Kausalitäten in der Struktur und insbesondere die relevanten Teilprozesse herausarbeiten.

Die Wahl der zu einem konkreten Problem gehörenden Skalierungen basiert auf Beobachtungen und Erfahrungen. Ein Vorteil dieses Ansatzes ist die Transparenz der getroffenen Annahmen.

Der Ansatz liefert keine rigorose Methode, zu beliebigen Anfangs- und Randwerten asymptotische Näherungen der Zustandsgrößen zu finden. Er ist vielmehr als Generator für Vermutungen zu verstehen, dies aber auf systematische Art und Weise. Es ist eine weitere, zum Teil schwierige Aufgabe zu beweisen, dass die gefundenen Näherungslösungen tatsächlich Näherungen des gestörten Problems sind. In einer formalen Sprache heißt das: Sei $L = L_0 + o(1)$ ein Operator, u eine Lösung von $Lx = 0$ und u_0 eine Lösung von $L_0x = 0$. Dann muss für beliebige Anfangs- und Randwerte keinesfalls $u = u_0 + o(1)$ gelten.

Als Rechtfertigung für diesen Ansatz sei erwähnt, dass sich alle gängigen meteorologischen Modelle daraus ableiten lassen.

Im Folgenden wird das konkrete Vorgehen kurz beschrieben.

2.2 WAHL DES GLEICHUNGSSYSTEMS

Den Ausgangspunkt bilden die vollen kompressiblen, dreidimensionalen Strömungsgleichungen für Masse, Impuls und potentielle Temperatur (Energie), die von Klein (2010, S. 252) in der folgenden Form gegeben werden:

$$\frac{\partial \rho}{\partial t} + \nabla_{\parallel} \cdot (\rho \mathbf{u}) + \frac{\partial \rho w}{\partial z} = 0 \quad (2.1a)$$

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla_{\parallel} \mathbf{u} + w \frac{\partial \mathbf{u}}{\partial z} + (2\boldsymbol{\Omega} \times \mathbf{v})_{\parallel} + \frac{1}{\rho} \nabla_{\parallel} p = \mathbf{Q}_u \quad (2.1b)$$

$$\frac{\partial w}{\partial t} + \mathbf{u} \cdot \nabla_{\parallel} w + w \frac{\partial w}{\partial z} + (2\boldsymbol{\Omega} \times \mathbf{v})_{\perp} + \frac{1}{\rho} \frac{\partial p}{\partial z} = Q_w - g \quad (2.1c)$$

$$\frac{\partial \Theta}{\partial t} + \mathbf{u} \cdot \nabla_{\parallel} \Theta + w \frac{\partial \Theta}{\partial z} = Q_{\Theta}. \quad (2.1d)$$

Hierbei bezeichne ρ die Dichte, Θ die potentielle Temperatur und \mathbf{v} die Strömungsgeschwindigkeit mit der horizontalen und vertikalen

Komponente \mathbf{u} beziehungsweise w . Dies sind die unbekannt Funktionen der horizontalen und vertikalen Koordinaten $\mathbf{x} = x \mathbf{i} + y \mathbf{j}$ und z , und der Zeit t . Die Größen ρ , Θ und p sind durch die Zustandsgleichung

$$\Theta = \frac{p_{\text{ref}}^{\frac{\gamma-1}{\gamma}} p^{\frac{1}{\gamma}}}{R_d \rho} \quad (2.1e)$$

miteinander gekoppelt, wobei γ für den Isentropenexponenten, R_d für die Gaskonstante von trockener Luft und p_{ref} für einen Referenzdruck steht, siehe Tabelle 2.1 auf Seite 9. Die Erdbeschleunigung und der Rotationsvektor der Erde sind durch g und $\boldsymbol{\Omega}$ dargestellt. Die Terme Q_u , Q_w und Q_Θ stehen für allgemeine Effekte von molekularem und turbulentem Transport.

Ein beliebiger Vektor \mathbf{v} wird in seinen horizontalen und vertikalen Anteil $\mathbf{u} = \mathbf{v}_\parallel = (\mathbf{1} - \mathbf{k} \otimes \mathbf{k}) \mathbf{v}$ und $w = v_\perp = \mathbf{v} \cdot \mathbf{k}$ aufgespalten, wobei \otimes das dyadische Produkt und \mathbf{k} der vertikale Einheitsvektor ist. Die Divergenz wird in einen horizontalen und einen vertikalen Anteil $\nabla \cdot \mathbf{v} = \nabla_\parallel \cdot \mathbf{v}_\parallel + \partial \mathbf{v} \cdot \mathbf{k} / \partial z$ gespalten. Mit $\nabla_\parallel \psi = \partial \psi / \partial x \mathbf{i} + \partial \psi / \partial y \mathbf{j}$ analog wird der Gradient in $\nabla \psi = \nabla_\parallel \psi + \partial \psi / \partial z \mathbf{k}$ aufgeteilt.

2.3 DIMENSIONSANALYSE

Betrachtet man Gleichungen zur Modellierung physikalischer Probleme, dann sind die auftretenden Variablen und Parameter physikalische Größen, die sich aus unabhängigen *Basisgrößen* zusammensetzen. Jede physikalische Größe Q hat eine Dimension $[Q]$ zum Ausdrücken der qualitativen Eigenschaften, die sich aus *Basisdimensionen* L_1, \dots, L_r zusammensetzen,

$$[Q] = \prod_{i=1}^r L_i^{\alpha_i}, \quad (2.2)$$

mit $\alpha_i \in \mathbb{Q}$. Es bedarf maximal sieben solcher Basisgrößen mit dazugehörigen Basisdimensionen: Länge $[L]$, Masse $[M]$, Zeit $[T]$, Stromstärke $[I]$, Temperatur $[\Theta]$, Stoffmenge $[N]$ und Lichtstärke $[J]$ (vgl. z.B. Kuchling, 2007). Nach Wahl eines *Einheitensystems* entspricht jeder Dimension eine *Einheit*, und jede physikalische Größe kann durch eine Zahl quantifiziert werden. Kennt man das Einheitensystem und die Dimension einer physikalischen Größe, lässt sich aus dieser Zahl wieder ein physikalischer Wert ermitteln. Für Gleichungen, die aus realen Problemen hervorgehen, kann man stets davon ausgehen, dass Lösungen nicht von der Wahl des Einheitensystems abhängen. Das gängigste Einheitensystem ist das SI-System mit den Einheiten m , kg , s , A , K , mol und cd für die sieben genannten Basisgrößen. Größen, für die in der Darstellung (2.2) sämtliche Exponenten α_i ver-

schwinden, heißen dimensionslos. Dimensionslose Größen bekommen bei der Modellierung physikalischer Probleme eine besondere Bedeutung, die in dem folgenden Satz begründet ist.

Satz 2.1 (Buckingham's Pi-theorem). *Seien Q, Q_1, \dots, Q_n unabhängige physikalische Größen (Variablen und Parameter), die r unabhängige Basisdimensionen umfassen und für die ein funktionaler Zusammenhang*

$$Q = f(Q_1, \dots, Q_n) \quad (2.3)$$

besteht, für eine Funktion f . Dann lässt sich (2.3) durch dimensionslose Größen ausdrücken. Genauer gibt es $k + 1 = n + 1 - r$ dimensionslose Größen Π, Π_1, \dots, Π_k der Form

$$\Pi = Q \cdot Q_1^{\beta_1} \dots Q_n^{\beta_n} \quad (2.4a)$$

$$\Pi_i = Q_1^{\alpha_{i1}} \dots Q_n^{\alpha_{in}}, \quad i = 1, \dots, k \quad (2.4b)$$

für gewisse $\alpha_i, \beta_i \in \mathbb{Q}$ und eine Funktion g mit

$$\Pi = g(\Pi_1, \dots, \Pi_k). \quad (2.5)$$

Insbesondere gilt:

$$Q = Q_1^{-\beta_1} \dots Q_n^{-\beta_n} g(\Pi_1, \dots, \Pi_k), \quad (2.6)$$

bzw.

$$Q = g(\Pi_1, \dots, \Pi_k) \quad (2.7)$$

für dimensionsloses Q .

Beweis. Siehe Bluman und Kumei (1989, §1.2.3). \square

Für den vorliegenden Fall besagt der Satz folgendes: Gibt es einen funktionalen Zusammenhang

$$\begin{pmatrix} \rho \\ \mathbf{v} \\ \Theta \end{pmatrix} = f(\alpha, \boldsymbol{\Omega}, g, p_{\text{ref}}, T_{\text{ref}}, \Delta T|_p^{\text{eq}}, R_d) \quad (2.8)$$

der abhängigen Größen ρ, \mathbf{v} und Θ aus (2.1) und den ersten sieben dimensionsbehafteten charakteristischen Größen aus Tabelle 2.1 mit einer Funktion f , dann gibt es bei den vier auftretenden Basiseinheiten m, s, kg und K drei dimensionslose Größen Π_1, Π_2 und Π_3 und eine Funktion g mit

$$\begin{pmatrix} \rho' \\ \mathbf{v}' \\ \Theta' \end{pmatrix} = g(\Pi_1, \Pi_2, \Pi_3), \quad (2.9)$$

für $\rho' = \rho/\rho_{\text{ref}}, \mathbf{v}' = \mathbf{v}/u_{\text{ref}}$ und $\Theta' = \Theta/T_{\text{ref}}$. Für ρ_{ref} und u_{ref} siehe Tabelle 2.2.

Erdradius	$a \sim 6 \times 10^6 \text{ m}$
Rotationsrate der Erde	$\Omega \sim 1 \times 10^{-4} \text{ s}^{-1}$
Erdbeschleunigung	$g \sim 9.81 \text{ m s}^{-2}$
Druck auf Meeresebene	$p_{\text{ref}} \sim 1 \times 10^5 \text{ kg m}^{-1} \text{ s}^{-2}$
Gefrierpunkt von Wasser	$T_{\text{ref}} \sim 273 \text{ K}$
Pot. Temperaturdiff. (Äquator-Pol)	$\Delta T_{\text{p}}^{\text{eq}} \sim 40 \text{ K}$
Vert. Temperaturdiff. i. d. Troposphäre	
Gaskonstante der trockenen Luft	$R_d = 287 \text{ m}^2 \text{ s}^{-2} \text{ K}^{-1}$
Trockener Isentropenexponent	$\gamma = 1.4$

Tabelle 2.1: Charakteristische Größen in der Atmosphäre (Klein, 2010, Table 1)

Druck auf Meereshöhe	$\rho_{\text{ref}} = p_{\text{ref}} / (R_d T_{\text{ref}}) \sim 1.25 \text{ kg m}^{-3}$
Druckskalenhöhe	$h_{\text{sc}} = \gamma p_{\text{ref}} / (g \rho_{\text{ref}}) \sim 11 \text{ km}$
Schallgeschwindigkeit	$c_{\text{ref}} = \sqrt{\gamma p_{\text{ref}} / \rho_{\text{ref}}} \sim 330 \text{ m s}^{-1}$
Geschwindigkeit innerer Wellen	$c_{\text{int}} = \sqrt{g h_{\text{sc}} \frac{\Delta T_{\text{p}}^{\text{eq}}}{T_{\text{ref}}}} \sim 110 \text{ m s}^{-1}$
Thermische Windgeschwindigkeit	$u_{\text{ref}} = \frac{2}{\pi} \frac{g h_{\text{sc}}}{\Omega a} \frac{\Delta T_{\text{p}}^{\text{eq}}}{T_{\text{ref}}} \sim 12 \text{ m s}^{-1}$

Tabelle 2.2: Abgeleitete Hilfsgrößen (Klein, 2010, Table 3)

Die Existenz solcher dimensionsloser Größen ermöglicht die Klassifizierung sämtlicher Lösungen eines physikalischen Gleichungssystems. Jedes Tupel von dimensionslosen Größen Π_1, \dots, Π_k definiert eine Klasse von Lösungen, die sich nur durch eine Umskalierung der Referenzgrößen unterscheidet. Die Wahl der dimensionslosen Parameter ist dabei nicht eindeutig und Teil des Modellierungsprozesses. Für eine ausführlichere Darstellung, insbesondere für Beispiele, sei auf Barenblatt (1996), Bluman und Kumei (1989), Howison (2005) und Klein, Botta u. a. (2001) verwiesen.

Die in Klein (2008) gewählten und aus den charakteristischen Größen und Hilfsgrößen aus den Tabellen 2.1 und 2.2 gebildeten dimensionslosen Parameter sind durch

$$\Pi_1 = \frac{h_{\text{sc}}}{a} \sim 1.6 \times 10^{-3}, \quad (2.10a)$$

$$\Pi_2 = \frac{\Delta T_{\text{p}}^{\text{eq}}}{T_{\text{ref}}} \sim 1.5 \times 10^{-1}, \quad (2.10b)$$

$$\Pi_3 = \frac{c_{\text{ref}}}{\Omega a} \sim 4.7 \times 10^{-1} \quad (2.10c)$$

gegeben.

Um eine kleine Inkonsistenz in Klein (2010) zu beheben, werden im Folgenden für u_{ref} und c_{ref} die Definitionen aus Klein (2008) verwendet, nämlich

$$u_{\text{ref}} = \frac{g h_{\text{sc}}}{\Omega a} \frac{\Delta T_{\text{p}}^{\text{eq}}}{T_{\text{ref}}} \quad \text{und} \quad c_{\text{ref}} = \sqrt{\frac{p_{\text{ref}}}{\rho_{\text{ref}}}}. \quad (2.11)$$

Dadurch ändert sich an den Resultaten aber nichts. Als nächstes entdimensionalisiert man das Gleichungssystem aus (2.1e) mit den Referenzgrößen h_{sc} für die räumlichen Koordinaten, p_{ref} , T_{ref} , ρ_{ref} für die thermodynamischen Variablen, u_{ref} für die Geschwindigkeit und $t_{ref} = h_{sc}/u_{ref}$ für die Zeit (siehe Tabellen 2.1 und 2.2 und (2.11)). Es sei also

$$\mathbf{x}' = \frac{\mathbf{x}}{h_{sc}}, \quad \mathbf{v}' = \frac{\mathbf{v}}{u_{ref}}, \quad t' = \frac{t}{t_{ref}} \quad (2.12a)$$

$$p' = \frac{p}{p_{ref}}, \quad \rho' = \frac{\rho}{\rho_{ref}}, \quad \Theta' = \frac{\Theta}{T_{ref}}. \quad (2.12b)$$

Man erhält das dimensionslose Gleichungssystem

$$\frac{\partial \rho}{\partial t} + \nabla_{\parallel} \cdot (\rho \mathbf{u}) + \frac{\partial \rho w}{\partial z} = 0 \quad (2.13a)$$

$$\begin{aligned} \frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla_{\parallel} \mathbf{u} + w \frac{\partial \mathbf{u}}{\partial z} + \frac{\Pi_1}{\Pi_2 \Pi_3^2} (2\boldsymbol{\Omega} \times \mathbf{v})_{\parallel} \\ + \frac{1}{(\Pi_2 \Pi_3)^2} \frac{1}{\rho} \nabla_{\parallel} p = \mathbf{Q}_u \end{aligned} \quad (2.13b)$$

$$\begin{aligned} \frac{\partial w}{\partial t} + \mathbf{u} \cdot \nabla_{\parallel} w + w \frac{\partial w}{\partial z} + \frac{\Pi_1}{\Pi_2 \Pi_3^2} (2\boldsymbol{\Omega} \times \mathbf{v})_{\perp} \\ + \frac{1}{(\Pi_2 \Pi_3)^2} \frac{1}{\rho} \frac{\partial p}{\partial z} = -\frac{1}{(\Pi_2 \Pi_3)^2} + Q_w \end{aligned} \quad (2.13c)$$

$$\frac{\partial \Theta}{\partial t} + \mathbf{u} \cdot \nabla_{\parallel} \Theta + w \frac{\partial \Theta}{\partial z} = Q_{\Theta}. \quad (2.13d)$$

Die dimensionslose Form der Zustandsgleichung lautet

$$\Theta = \frac{p^{\frac{1}{\gamma}}}{\rho}. \quad (2.13e)$$

Da nun nur noch die dimensionslosen Gleichungen betrachtet werden, werden die dimensionsbehafteten Variablen aus (2.1) und die dimensionslosen Variablen aus (2.13) mit denselben Symbolen bezeichnet, die Striche aus (2.12) wurden weggelassen. Dasselbe gilt für die Quellterme.

Um im Folgenden eine asymptotische Analyse des Gleichungssystems zu ermöglichen, müssen die drei kleinen dimensionslosen Parameter aus (2.10) an einen kleinen Parameter ε gekoppelt werden. Das geschieht durch die Wahl eines gekoppelten Grenzübergangs im nächsten Abschnitt.

2.4 GEKOPPELTER GRENZÜBERGANG

Sei y eine Lösung einer dimensionslosen Gleichung, die neben den unabhängigen Variablen x_1, \dots, x_n auch von k unabhängigen dimensionslosen „kleinen“ Parametern abhängt, genauer

$$y = y(x_1, \dots, x_n; \Pi_1, \dots, \Pi_k). \quad (2.14)$$

Es stellt sich die Frage, wie sich y für kleine Parameter verhält. Vertauscht man die Rolle von unabhängigen Variablen und Parametern und betrachtet y als Funktion der Parameter, ist man an dem Verhalten von y in einer Umgebung des Ursprungs im Parameterraum interessiert, also in einer Umgebung von

$$\Pi_1 = \dots = \Pi_k = 0. \quad (2.15)$$

Der erste naheliegende Ansatz ist eine Taylor-Entwicklung um den Ursprung. Dies setzt aber die Fréchet-Differenzierbarkeit im Ursprung voraus. Das ist eine starke Annahme. Schwächer ist die Annahme der Gâteaux-Differenzierbarkeit. Das entspricht der Annäherung des Ursprungs entlang beliebiger Geraden. Noch allgemeiner ist die Annäherung entlang einer beliebigen Kurve. Sei also

$$\begin{aligned} \gamma : [0; \varepsilon_0) &\rightarrow \mathbb{R}^k \\ \varepsilon &\mapsto (\hat{\Pi}_1(\varepsilon), \dots, \hat{\Pi}_k(\varepsilon)) \end{aligned} \quad (2.16a)$$

eine Kurve mit

$$\lim_{\varepsilon \rightarrow 0} \gamma(\varepsilon) = 0. \quad (2.16b)$$

Die Kurve γ bezeichnet man auch als *gekoppelten Grenzübergang* (distinguished limit). Man betrachte nun an Stelle von y die Komposition

$$\hat{y} \circ \gamma(\varepsilon) = \hat{y}(x_1, \dots, x_n; \hat{\Pi}_1(\varepsilon), \dots, \hat{\Pi}_1(\varepsilon)). \quad (2.17)$$

Ist $\hat{y} \circ \gamma$ in $\varepsilon = 0$ differenzierbar, dann gilt

$$\hat{y} \circ \gamma(\varepsilon) = y(x_1, \dots, x_n; 0) + \varepsilon (\hat{y} \circ \gamma)'(0) + o(\varepsilon). \quad (2.18)$$

Man beachte, dass diese Anforderungen an $\hat{y} \circ \gamma$ deutlich schwächer sind als die Voraussetzungen für die Existenz von Fréchet- beziehungsweise Gâteaux-Ableitungen von y . Die Wahl einer Kurve wie in (2.16) ist wesentlicher Teil des Modellierungsprozesses. Ein klassisches Beispiel, das zeigt, dass eine andere Wahl des gekoppelten Grenzübergangs ganz andere Prozesse beschreibt, ist der harmonische Oszillator (siehe z.B. Kevorkian und Cole, 1981; Klein, Vater u. a., 2011).

Konkret werden die dimensionslosen Parameter aus Π_1 , Π_2 und Π_3 (2.10) in Klein (2010) an einen kleinen Parameter ε gekoppelt. Für $\varepsilon \rightarrow 0$ sei

$$\Pi_1 = \mathcal{O}(\varepsilon^3), \quad (2.19a)$$

$$\Pi_2 = \mathcal{O}(\varepsilon), \quad (2.19b)$$

$$\Pi_3 = \mathcal{O}(\sqrt{\varepsilon}). \quad (2.19c)$$

Mit diesem gekoppelten Grenzübergang erhält man aus (2.13) die Gleichungen

$$\frac{\partial \rho}{\partial t} + \nabla_{\parallel} \cdot (\rho \mathbf{u}) + \frac{\partial \rho w}{\partial z} = 0 \quad (2.20a)$$

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla_{\parallel} \mathbf{u} + w \frac{\partial \mathbf{u}}{\partial z} + 2\varepsilon (\boldsymbol{\Omega} \times \mathbf{v})_{\parallel} + \frac{1}{\varepsilon^3} \frac{1}{\rho} \nabla_{\parallel} p = \mathbf{Q}_u \quad (2.20b)$$

$$\begin{aligned} \frac{\partial w}{\partial t} + \mathbf{u} \cdot \nabla_{\parallel} w + w \frac{\partial w}{\partial z} + 2\varepsilon (\boldsymbol{\Omega} \times \mathbf{v})_{\perp} \\ + \frac{1}{\varepsilon^3} \frac{1}{\rho} \frac{\partial p}{\partial z} = -\frac{1}{\varepsilon^3} + Q_w \end{aligned} \quad (2.20c)$$

$$\frac{\partial \Theta}{\partial t} + \mathbf{u} \cdot \nabla_{\parallel} \Theta + w \frac{\partial \Theta}{\partial z} = Q_{\Theta} \quad (2.20d)$$

und

$$\Theta = \frac{p^{\frac{1}{\gamma}}}{\rho}. \quad (2.20e)$$

2.5 WIE GROSS IST ε ?

In der Asymptotik ist man grundsätzlich an dem funktionalen Verhalten für $\varepsilon \rightarrow 0$ interessiert. Dennoch ist es manchmal hilfreich für asymptotische Größen, eine Vorstellung zu haben, in welcher Größenordnung der numerische Wert liegt. Die Frage, die sich stellt, ist für welchen Wert von ε die Gleichungen von (2.10) und (2.19) am besten zusammen passen. Es ist hervorzuheben, dass die folgenden Betrachtungen unabhängig von den betrachteten Gleichungen sind.

Gleichung (2.19) besagt, dass es von ε unabhängige Konstanten C_1, C_2 und C_3 mit

$$\Pi_1 = C_1 \cdot \varepsilon^3 \quad (2.21a)$$

$$\Pi_2 = C_2 \cdot \varepsilon \quad (2.21b)$$

$$\Pi_3 = C_3 \cdot \sqrt{\varepsilon} \quad (2.21c)$$

gibt. Als Größen der Ordnung $\mathcal{O}(1)$ sollen die Konstanten C_1, C_2 und C_3 dabei „möglichst nahe“ bei 1 liegen. Eine naheliegende Möglichkeit ist, die Funktion

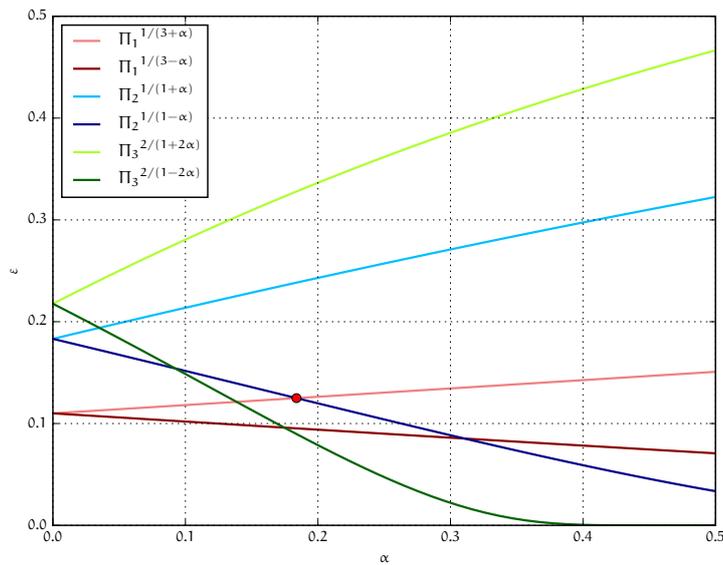
$$C(\varepsilon) = \left\| \begin{pmatrix} C_1 \\ C_2 \\ C_3 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\| \quad (2.22)$$

zu minimieren. Für die p-Norm erhält man beispielsweise

$$C(\varepsilon) = \sqrt[p]{|\varepsilon^{-3} \Pi_1 - 1|^p + |\varepsilon^{-1} \Pi_2 - 1|^p + |\varepsilon^{-1/2} \Pi_3 - 1|^p}. \quad (2.23)$$

Die Werte von ε , für die das Minimum erreicht wird, lassen sich Tabelle 2.3 entnehmen.

p	$1/\varepsilon$
2	7.598 27
4	7.687 72
6	7.689 61
8	7.685 64
10	7.681 46

Tabelle 2.3: Werte für ε bezüglich der p-Norm.Abbildung 2.1: Grafische Darstellung der Abhängigkeit von ε und α . (Referenzgrößen aus Klein (2008))

Ein anderer Ansatz berücksichtigt die Struktur asymptotischer Entwicklungen stärker. Betrachtet werden für $\alpha \in \mathbb{R}$ die folgenden sechs Ungleichungen:

$$\varepsilon^\alpha \leq C_i \leq \varepsilon^{-\alpha}, \quad i = 1, 2, 3. \quad (2.24)$$

Gesucht ist nun dasjenige ε , das bei kleinstmöglichem α alle sechs Ungleichungen erfüllt. Zusammen mit (2.21) führt das zu sechs Ungleichungen für ε :

$${}^{3+\alpha}\sqrt{\Pi_1} \leq \varepsilon \leq {}^{3-\alpha}\sqrt{\Pi_1} \quad (2.25a)$$

$${}^{1+\alpha}\sqrt{\Pi_2} \leq \varepsilon \leq {}^{1-\alpha}\sqrt{\Pi_2} \quad (2.25b)$$

$${}^{1/2+\alpha}\sqrt{\Pi_3} \leq \varepsilon \leq {}^{1/2-\alpha}\sqrt{\Pi_3}. \quad (2.25c)$$

In Abbildung 2.1 ist die α -Abhängigkeit von ε grafisch dargestellt. Liegt ε zwischen allen drei Paaren von Kurven, sind alle sechs Ungleichungen erfüllt. Gesucht ist der minimale Wert von α , für den das möglich ist. Der Abbildung entnimmt man, dass es sich genau

um den Schnittpunkt der dunkelblauen und hellroten Kurve handelt. Damit ist α durch

$$\alpha = \frac{\log \Pi_1 - 3 \log \Pi_2}{\log \Pi_1 + \log \Pi_2} = 0.184 \quad (2.26)$$

bestimmt. Das entspricht den Werten

$$C_1 = 0.68, \quad C_2 = 1.47, \quad C_3 = 1.32 \quad (2.27)$$

und

$$\varepsilon = 1/8.00. \quad (2.28)$$

2.6 ASYMPTOTISCHE ANSATZFUNKTIONEN

Mit dem asymptotischen Gleichungssystem aus (2.20) ist der Grundstein für eine asymptotische Analyse atmosphärischer Phänomene gelegt.

Findet man nun Funktionen $\rho(\mathbf{x}, z, t; \varepsilon)$, $\Theta(\mathbf{x}, z, t; \varepsilon)$, $\mathbf{u}(\mathbf{x}, z, t; \varepsilon)$ und $w(\mathbf{x}, z, t; \varepsilon)$, die das Gleichungssystem (2.20) lösen, hat man eine Beschreibung der atmosphärischen Strömungen, für die die Referenzgrößen aus Tabelle 2.2 auf Seite 9 charakteristische Werte sind. In dieser Allgemeinheit wird das auf Grund der Komplexität der Gleichungen nicht möglich sein. Ein möglicher Ausweg ist es, Lösungen der degenerierten Gleichungen für $\varepsilon \rightarrow 0$ zu suchen. Wie eingangs schon erwähnt, müssen diese Lösungen für beliebige Anfangs- und Randwerte keineswegs Approximationen der gestörten Gleichung sein. Die Lösungen liefern allenfalls eine Vermutung, wie eine Lösung aussehen könnte. Mathematisch muss das in einem rigorosen Beweis gezeigt werden. Das ist jedoch nicht Gegenstand dieser Arbeit.

Zum Finden möglicher Lösungen wird häufig der Ansatz

$$\alpha(\mathbf{x}, z, t; \varepsilon) = \sum_i \phi^{(i)}(\varepsilon) \alpha^{(i)}(\mathbf{x}, z, t) \quad (2.29)$$

gewählt, wobei $\phi^{(i)}$ eine asymptotische Folge ist, für $\varepsilon \rightarrow 0$ gilt also

$$\phi^{(i+1)} = o\left(\phi^{(i)}\right). \quad (2.30)$$

Die asymptotische Größenordnung von α ist durch die asymptotische Größenordnung von $\phi^{(0)}$ gegeben. Betrachtet man zum Beispiel für $\alpha = v$ das Geschwindigkeitsfeld und sei $\phi^{(0)} = \mathcal{O}(1)$, dann werden atmosphärische Phänomene untersucht, deren charakteristische Geschwindigkeit u_{ref} entspricht, denn als Referenzgröße ist u_{ref} von der Ordnung $\mathcal{O}(1)$. Ist man an anderen Größenordnungen interessiert, wird $\phi^{(0)}$ entsprechend gewählt. Möchte man zudem Phäno-

mene studieren, deren räumliche und zeitliche charakteristische Größen nicht denen entsprechen, mit denen dimensionslos gemacht wurde, werden die Argumente in den $\alpha^{(i)}$ entsprechend umskaliert. Ein entsprechend erweiterter Ansatz ist

$$\alpha(x, z, t; \varepsilon) = \sum_i \phi^{(i)}(\varepsilon) \alpha^{(i)}(\varepsilon^{\alpha_x} x, \varepsilon^{\alpha_z} z, \varepsilon^{\alpha_t} t) \quad (2.31)$$

mit geeigneten Werten für die α_x , α_z und α_t . Die Wahl der $\phi^{(i)}$ ergibt sich aus den Gleichungen. Passen die gewählten $\phi^{(i)}$ nicht zu den Gleichungen, erhält man keine Lösungen für kleine ε .

Als Beispiel betrachte man mit abweichender Notation Holmes (1995, (1.20)):

$$\alpha^2 + 2\varepsilon\alpha - 1 = 0. \quad (2.32)$$

Für α nimmt man eine asymptotische Entwicklung der Form

$$\alpha = \alpha^{(0)} + \varepsilon^\beta \alpha^{(1)} + \dots \quad (2.33)$$

an. Einsetzen von (2.33) in (2.32) ergibt

$$\begin{aligned} (\alpha^{(0)})^2 + 2\varepsilon^\beta \alpha^{(0)} \alpha^{(1)} + \dots \\ + 2\varepsilon (\alpha^{(0)} + \varepsilon^\beta \alpha^{(1)} + \dots) - 1 = 0 \end{aligned} \quad (2.34)$$

für $\varepsilon \rightarrow 0$ (ebd., (1.22)). Zunächst folgt unmittelbar $(\alpha^{(0)})^2 = 1$. Nur für $\beta = 1$ bekommt man eine widerspruchsfreie asymptotische Lösung, die von 0 verschiedene Störungen zulässt (vgl. ebd., Exercise 4, S. 24).

Um die später folgende asymptotische Analyse mit solchen Ansatzfunktionen zu erleichtern, bietet es sich an, neue Koordinaten gemäß der α_x , α_z und α_t einzuführen und (2.20) entsprechend umzuschreiben (siehe Klein, 2010, S. 253). Führt man die umskalierten Koordinaten

$$\tau = \varepsilon^{\alpha_t} t \quad \text{und} \quad \xi = \varepsilon^{\alpha_x} x \quad (2.35)$$

ein, entspricht das dem Entdimensionalisieren mit der charakteristischen Zeit $T = t_{\text{ref}}/\varepsilon^{\alpha_t} t$ und der horizontalen charakteristischen Länge $L = h_{\text{sc}}/\varepsilon^{\alpha_x}$.

Nutzt man $\Delta\Theta/T_{\text{ref}} = \mathcal{O}(\varepsilon)$ aus (siehe (2.19)), lässt sich die potentielle Temperatur Θ ähnlich wie in Klein (ebd., S. 253) umschreiben:

$$\Theta(\xi, z, \tau; \varepsilon) = 1 + \varepsilon(\Theta_1(z; \varepsilon) + \tilde{\Theta}(\xi, z, \tau; \varepsilon)), \quad (2.36)$$

mit $\Theta_1, \tilde{\Theta} = \mathcal{O}(1)$.

Da in den meisten betrachteten Regimen das Verhältnis von vertikaler Geschwindigkeit zu horizontaler Geschwindigkeit dem Verhältnis

von h_{sc} zu L entspricht und demnach von der Ordnung $\mathcal{O}(\varepsilon^{\alpha_x})$ ist, führt Klein (2010, S. 253) eine skalierte vertikale Geschwindigkeit \tilde{w} mit

$$w(\xi, z, \tau) = \varepsilon^{\alpha_x} \tilde{w}(\xi, z, \tau) \quad (2.37)$$

ein.

Aus (2.20) wird dann das umskalierte und zu Klein (ebd., (9)) ähnliche Gleichungssystem

$$\frac{\varepsilon^{\alpha_t}}{\varepsilon^{\alpha_x}} \frac{\partial \rho}{\partial t} + \nabla_{\parallel} \cdot (\rho \mathbf{u}) + \frac{\partial \rho \tilde{w}}{\partial z} = 0 \quad (2.38a)$$

$$\frac{\varepsilon^{\alpha_t}}{\varepsilon^{\alpha_x}} \frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla_{\parallel} \mathbf{u} + \tilde{w} \frac{\partial \mathbf{u}}{\partial z} + \frac{(2\boldsymbol{\Omega} \times \mathbf{v})_{\parallel}}{\varepsilon^{\alpha_x - 1}} + \frac{1}{\varepsilon^3} \frac{1}{\rho} \nabla_{\parallel} p = Q_{\mathbf{u}}^{\varepsilon} \quad (2.38b)$$

$$\frac{\varepsilon^{\alpha_t}}{\varepsilon^{\alpha_x}} \frac{\partial \tilde{w}}{\partial t} + \mathbf{u} \cdot \nabla_{\parallel} \tilde{w} + \tilde{w} \frac{\partial \tilde{w}}{\partial z} + \frac{(2\boldsymbol{\Omega} \times \mathbf{v})_{\perp}}{\varepsilon^{2\alpha_x - 1}} + \frac{1}{\varepsilon^{3+\alpha_x}} \frac{1}{\rho} \frac{\partial p}{\partial z} = -\frac{1}{\varepsilon^{3+\alpha_x}} + Q_w^{\varepsilon} \quad (2.38c)$$

$$\frac{\varepsilon^{\alpha_t}}{\varepsilon^{\alpha_x}} \frac{\partial \tilde{\Theta}}{\partial t} + \mathbf{u} \cdot \nabla_{\parallel} \tilde{\Theta} + \tilde{w} \frac{\partial \tilde{\Theta}}{\partial z} + \tilde{w} \frac{\partial \Theta_1}{\partial z} = Q_{\tilde{\Theta}}^{\varepsilon} \quad (2.38d)$$

$$\varepsilon \left(\tilde{\Theta} + \Theta_1 \right) = \frac{p^{\frac{1}{\gamma}}}{\rho} - 1. \quad (2.38e)$$

Hierbei bezeichnen $Q_{\mathbf{u}}^{\varepsilon}$, Q_w^{ε} und $Q_{\tilde{\Theta}}^{\varepsilon}$ an die Umskalierung angepasste Quellterme.

Die volle Mächtigkeit entfaltet der Ansatz bei Mehrskalenproblemen. Dort haben die Ansätze die Form

$$\alpha(\mathbf{x}, z, t; \varepsilon) = \sum_i \phi^{(i)}(\varepsilon) \alpha^{(i)} \left(\dots, \frac{t}{\varepsilon}, t, \varepsilon t, \dots, \frac{\mathbf{x}}{\varepsilon}, \mathbf{x}, \varepsilon \mathbf{x}, \dots, z \right) \quad (2.39)$$

Bei dem hier vorgestellten Vorgehen stellt das Entdimensionalisieren mit globalen Referenzgrößen keine Einschränkung bezüglich der Vielfalt der zu untersuchenden Probleme dar. Ganz im Gegenteil ermöglicht dieser Ansatz, ausgehend von einer einheitlichen Basis, ein transparentes Vorgehen.

Die wesentliche Annahme ist der gekoppelte Grenzübergang (2.19). Die Wahl der Ansatzfunktionen in (2.39) ist weniger eine Annahme als vielmehr eine Konsequenz des betrachteten Regimes.

3

ASYMPTOTIK EINES ATMOSPHERISCHEN WIRBELS

Inhalt

3.1	Ausgangspunkt	18
3.1.1	Beschreibung der quasigeostrophischen Hintergrundströmung	18
3.1.2	Beschreibung des Wirbels	19
3.2	Quasigeostrophische Hintergrundströmung	21
3.3	Wirbelkernstruktur	23
3.4	Matching	25
3.5	Alternative Bestimmung der Stromfunktion	27
3.5.1	Lokales mitbewegtes orthogonales Koordinatensystem	27
3.5.2	Vorgehen zum Lösen der elliptischen Gleichung von ψ	28

Grundlage dieses Kapitels ist die Arbeit von Päsche u. a. (2012), in der ein konzentrierter, beinahe achsensymmetrischer, stark geneigter, barokliner Wirbel im Gradient-Wind-Regime betrachtet wird, der in eine quasigeostrophische Hintergrundströmung mit schwacher vertikaler Scherung eingebettet ist. Die Betrachtungen beschränken sich auf den Fall mit trockener Luft, wobei aber Effekte asymmetrischer diabatischer Erwärmung berücksichtigt werden.

Ein Ziel (ebd.) ist die Herleitung asymptotischer Gleichungen für die achsensymmetrische Umfangsgeschwindigkeit des Wirbelkerns in führender Ordnung und der Wirbelbewegung unter Berücksichtigung der Wirbelkernstruktur.

Es gibt eine Vielzahl an Arbeiten, die Wirbelbewegungen untersuchen, aber die Kernstruktur außer Acht lassen. Ebenso gibt es viele Arbeiten zur Modellierung der Wirbelkernstruktur, zur Beschreibung der Evolution von achsensymmetrischen, aufrechten Wirbeln im Gradient-Wind-Regime, ohne die Wirbelbewegung zu berücksichtigen. Für den zweidimensionalen Fall gibt es auch Arbeiten, die sowohl die Wirbelkernstruktur, als auch den Einfluss der Hintergrundströmung und die Bewegung des Wirbelzentrums berücksichtigen. Damit lässt sich aber nicht die vertikale Struktur, insbesondere eine Neigung der Wirbelachse beschreiben (siehe Referenzen in ebd.).

Die Herleitung basiert auf dem einheitlichen Mehrskalensatz von Klein (2010, 2008), der in Kapitel 2 auf Seite 5 vorgestellt wurde. Motiviert durch Callegari und Ting (1978) werden dazu Hintergrundströmung und Wirbelkernstruktur zunächst separat analysiert und

dann mit der Methode der angepassten asymptotischen Entwicklungen (siehe Eckhaus, 1979; van Dyke, 1964) aneinander angepasst.

Die vorliegende Arbeit stellt Päschke u. a. (2012) zusammengefasst dar, wobei der Schwerpunkt auf der mathematischen Beschreibung und weniger auf dem meteorologischen Hintergrund liegt. Des Weiteren wird auf einen kleinen Fehler in der Arbeit hingewiesen und diskutiert, wie dieser zu beheben ist, auch wenn die Details nicht bis zum Ende ausgearbeitet werden.

3.1 AUSGANGSPUNKT

3.1.1 Beschreibung der quasigeostrophischen Hintergrundströmung

Ausgangspunkt der asymptotischen Analyse in Päschke u. a. (ebd., (2.7)) ist das Gleichungssystem aus (2.38). Untersucht werden soll ein konzentrierter Wirbel, der in eine quasigeostrophische Hintergrundströmung eingebettet ist. Das quasigeostrophische Modell folgt aus (2.38) für die Wahl von $\varepsilon^{\alpha_x} = \varepsilon^{\alpha_t} = 2$. Das entspricht dem synoptischen Regime (Klein, 2010, §2.2.2). Des Weiteren wird von der β -Ebenen-Approximation

$$(\boldsymbol{\Omega} \times \mathbf{v})_{\parallel} = (f_0 + \varepsilon\beta y) \mathbf{k} \times \mathbf{u} \quad (3.1a)$$

$$(\boldsymbol{\Omega} \times \mathbf{v})_{\perp} = \boldsymbol{\Omega}_{\parallel} \times \mathbf{u} \quad (3.1b)$$

Gebrauch gemacht. Dabei wird von einer linearen Approximation des Coriolis-Parameters $f = f_0 + \beta y$ bei einem Breitengrad $\phi_0 = 30^\circ\text{N}$ ausgegangen. Es ist $f_0 = f(\phi_0)$ und $\beta = (df/dy)(\phi_0)$ und y steht für die lokale meridionale Koordinate in einer Tangentialebene an die Erdkugel bei dem Breitengrad ϕ_0 . Mit den Annahmen $\mathbf{Q}_u^\varepsilon = \mathbf{Q}_w^\varepsilon = 0$ und $Q_\Theta^\varepsilon = \text{Da}/\varepsilon \cdot Q_\Theta$ (Päschke u. a., 2012, (2.7)) erhält man aus (2.38) das Gleichungssystem

$$\frac{\partial \rho}{\partial t} + \nabla_{\parallel} \cdot (\rho \mathbf{u}) + \frac{\partial \rho \tilde{w}}{\partial z} = 0 \quad (3.2a)$$

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla_{\parallel} \mathbf{u} + \tilde{w} \frac{\partial \mathbf{u}}{\partial z} + \frac{f_0 + \varepsilon\beta y}{\varepsilon} \mathbf{k} \times \mathbf{u} + \frac{1}{\varepsilon^3} \frac{1}{\rho} \nabla_{\parallel} p = 0 \quad (3.2b)$$

$$\frac{\partial \tilde{w}}{\partial t} + \mathbf{u} \cdot \nabla_{\parallel} \tilde{w} + \tilde{w} \frac{\partial \tilde{w}}{\partial z} + \frac{2\boldsymbol{\Omega}_{\parallel} \times \mathbf{u}}{\varepsilon^3} + \frac{1}{\varepsilon^5} \frac{1}{\rho} \frac{\partial p}{\partial z} = -\frac{1}{\varepsilon^5} \quad (3.2c)$$

$$\frac{\partial \tilde{\Theta}}{\partial t} + \mathbf{u} \cdot \nabla_{\parallel} \tilde{\Theta} + \tilde{w} \frac{\partial \tilde{\Theta}}{\partial z} + \tilde{w} \frac{d\Theta_1}{dz} = \frac{\text{Da}}{\varepsilon} Q_\Theta \quad (3.2d)$$

und

$$\varepsilon \left(\tilde{\Theta} + \Theta_1 \right) = \frac{p^{\frac{1}{\gamma}}}{\rho} - 1. \quad (3.2e)$$

Diese Beschreibung der quasigeostrophischen Hintergrundströmung entspricht im Wesentlichen Päschke u. a. (ebd., (2.7)). Allerdings wurde dort der Coriolis-Term aus (3.2c) weggelassen. Das ist insofern

vertretbar, als dass er für die betrachteten Ordnungen keinen Beitrag liefert. Auch wurde dort \tilde{w} mit w bezeichnet.

3.1.2 Beschreibung des Wirbels

Betrachtet wird hier ein konzentrierter, beinahe achsensymmetrischer, stark geneigter, barokliner Wirbel, der in eine quasigeostrophische Hintergrundströmung mit schwacher vertikaler Scherung eingebettet ist (ebd., §1.2). Der Wirbelkern hat einen typischen Durchmesser von $L_{\text{mes}} \sim 150$ km. Da von einem konzentrierten Wirbel ausgegangen wird, gilt für das Verhältnis von Wirbeldurchmesser und synoptischer Längenskala

$$\delta = L_{\text{mes}}/L_{\text{syn}} \ll 1. \quad (3.3)$$

Dabei ist die ε -Abhängigkeit von δ noch zu bestimmen. Die als stark angenommene Neigung des Wirbels soll sich dadurch auszeichnen, dass die horizontale Auslenkung des Wirbels in der Höhe z von der Größenordnung des Wirbeldurchmessers ist. Bei einer Höhe des Wirbels von $h_{\text{sc}} \sim 10$ km sind so sehr starke Neigungen möglich.

Im Folgenden bezeichne \mathbf{X}_0 die horizontale Position des Wirbels auf der synoptischen Längenskala L_{syn} , das heißt $\mathbf{X}_0 = \mathcal{O}(L_{\text{syn}})$. Störungen von dieser Position in der Größenordnung des Wirbelkerns $\mathcal{O}(L_{\text{mes}})$ werden mit $\mathbf{X}^{(1)}$ bezeichnet. Um unrealistisch starke Neigungen des Wirbels zu vermeiden und damit der schwachen vertikalen Scherung Rechnung zu tragen, wird \mathbf{X}_0 als z -unabhängig angenommen. Damit lässt sich die horizontale Position des Wirbelzentrums in einer Höhe z auf der synoptischen Skala als

$$\mathbf{X}(t, z) \equiv X(t, z) \mathbf{i} + Y(t, z) \mathbf{j} = \mathbf{X}_0(t) + \delta \mathbf{X}^{(1)}(t, z) \quad (3.4)$$

schreiben (ebd., (1.1)). Das entspricht einer Normalen zur $x - y$ -Ebene, die in der Höhe z eine horizontale Störung von der Größenordnung $\mathcal{O}(L_{\text{mes}})$ erfährt. Siehe Abbildung 3.1 auf der nächsten Seite.

Wegen (3.3) erscheint der Wirbel im synoptischen Regime als Linienvirbel, das heißt die komplette Wirbelstärke ist auf einer Linie konzentriert. Wie in Callegari und Ting (1978) beschrieben, erzeugt der Linienvirbel ein Geschwindigkeitsfeld, das entlang des Wirbels singular ist. Ziel ist es aber, diese Singularität aufzulösen und das komplette endliche Geschwindigkeitsfeld zu beschreiben. Um die Wirbelkernstruktur zu studieren, wird das Gradient-Wind-Regime betrachtet.

Im Gegensatz zum Quasigeostrophischen Regime wo eine Balance zwischen Druckgradient und Coriolis-Termen herrscht, zeichnet sich das Gradient-Wind-Regime durch eine Balance des Druckgradienten, der Zentrifugalbeschleunigung und der Coriolis-Terme aus. Diese Balance liefert letztendlich auch die (asymptotische) Größenordnung von L_{mes} . Es ist also diejenige Umskalierung gesucht, die in

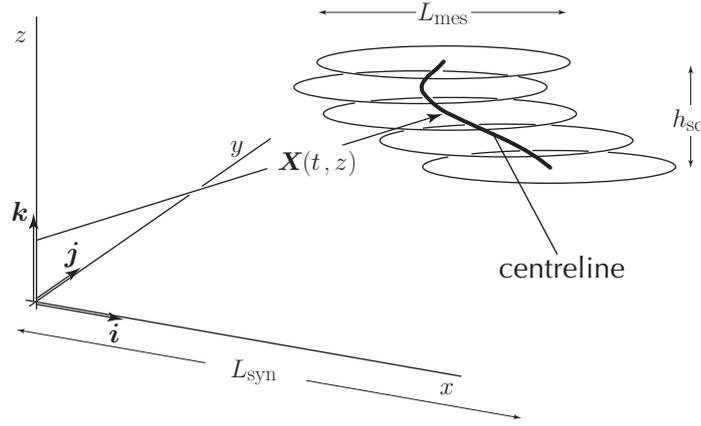


Abbildung 3.1: Geneigter konzentrierter mesoskaliger Wirbel der charakteristischen Höhe h_{sc} mit einem typischen Kerndurchmesser von L_{mes} , eingebettet in eine Umgebung auf der synoptischen Skala (L_{syn}), mit $h_{sc} \ll L_{mes} \ll L_{syn}$. (Päschke u. a., 2012, FIGURE 1)

(3.2b) eine Balance des zweiten, vierten und fünften Terms erzeugt. Sei $\delta = \varepsilon^\alpha$ und $p_{mes} = \varepsilon^\gamma p_{ref}$. Um eine Balance zwischen den drei Termen herzustellen, muss

$$3\alpha = 1 + \alpha = 3 + \alpha - \gamma \quad (3.5)$$

gelten. Es folgt $\alpha = 1/2$ und $\gamma = 2$ und damit $L_{mes} = \varepsilon^{1/2} L_{syn}$ und $p_{mes} = \mathcal{O}(\varepsilon^2 p_{ref})$.

Das Vorgehen der weiteren Analyse ähnelt stark dem von Callegari und Ting (1978), denn mit $L_{mes}/L_{syn} = \delta = \varepsilon^{1/2}$ liegen wie in Callegari und Ting (ebd.) zwei asymptotische Regime vor. Dort werden schlanke Wirbel betrachtet, das heißt, das Verhältnis von Wirbeldurchmesser zu Wirbellänge ist sehr klein. Im Gegensatz dazu folgt in dem hier betrachteten Fall aus (3.3) und $h_{sc} = \varepsilon^2 L_{syn}$ für das Verhältnis von Wirbeldurchmesser zu Wirbellänge

$$\frac{L_{mes}}{h_{sc}} = \varepsilon^{-3/2}. \quad (3.6)$$

Mit L_{mes} und L_{syn} hat man zwei asymptotisch separierte Längenskalen, die in den folgenden beiden Abschnitten zunächst getrennt von einander analysiert werden. In Abschnitt 3.4 auf Seite 25 werden die gewonnenen Resultate dann mit einander in Beziehung gesetzt.

3.2 QUASIGEOSTROPHISCHE HINTERGRUNDSTRÖMUNG

Die asymptotische Analyse der quasigeostrophischen Hintergrundströmung besteht im Wesentlichen aus einer leicht modifizierten Herleitung nach Pedlosky (1987) (siehe Klein, 2008; Päsche u. a., 2012). In diese Hintergrundströmung wird ein geneigter, singulärer barokliner Linienwirbel eingebettet. In führender Ordnung herrscht in der quasigeostrophischen Approximation geostrophisches Gleichgewicht, also ein Gleichgewicht zwischen Corioliskraft und Druckkraft. In höherer Ordnung werden aber auch Divergenzen der horizontalen Strömung berücksichtigt. Wie in der Meteorologie üblich, wird zur Beschreibung des Strömungsfeldes nicht das Geschwindigkeitsfeld \mathbf{v} verwendet, sondern die Wirbelstärke $\zeta = \mathbf{k} \cdot (\nabla \times \mathbf{v})$ und eine Stromfunktion ψ . Für ein horizontales Geschwindigkeitsfeld \mathbf{u} und den horizontalen Gradienten ∇_{\parallel} gilt dann $\mathbf{u} = -\mathbf{k} \times \nabla_{\parallel} \psi$. Die „[...] potentielle Vorticity [ist] eine Erhaltungsgröße für adiabatische Vorgänge“ (Etling, 2008) und steht in direktem Zusammenhang zu der Vorticity.

Wie in Päsche u. a. (2012, §3.1) werden asymptotische Ansatzfunktionen für Druck, Dichte, potentielle Temperatur sowie horizontales und vertikales Geschwindigkeitsfeld gewählt. Insbesondere sei

$$\mathbf{u} = \mathbf{u}^{(0)} + \varepsilon^{1/2} \mathbf{u}^{(1/2)} + \varepsilon \mathbf{u}^{(1)} + \varepsilon^{3/2} \mathbf{u}^{(3/2)} + \mathcal{O}(\varepsilon^{3/2}). \quad (3.7)$$

Das horizontale Geschwindigkeitsfeld und die geostrophische Stromfunktion in den ersten beiden Ordnungen werden durch

$$\tilde{\mathbf{u}} = \mathbf{u}^{(0)} + \delta \mathbf{u}^{(1/2)}, \quad (3.8a)$$

$$\psi = -\frac{1}{\rho_0 f_0} \left(p^{(2)} + \delta p^{(5/2)} \right) \quad (3.8b)$$

abgekürzt. Die Größen ρ_0 , $p^{(2)}$ und $p^{(5/2)}$ sind von der Ordnung $\mathcal{O}(1)$ und aus den asymptotischen Ansätzen von Päsche u. a. (ebd., (3.1)). Aus (3.2b) kann man die beiden Gleichungen

$$\nabla_{\parallel} \cdot \tilde{\mathbf{u}} = 0 \quad \text{und} \quad \tilde{\mathbf{u}} = -\mathbf{k} \times \nabla_{\parallel} \psi \quad (3.9)$$

gewinnen. Wie bei Klein (2008) lässt sich eine elliptische Gleichung zur Beschreibung von ψ herleiten. Es gilt nämlich

$$(\nabla_{\parallel}^2 + \mathcal{L}_z) \psi = -q_{\text{rel}}, \quad (3.10)$$

wobei ∇_{\parallel}^2 der horizontale Laplace-Operator,

$$\mathcal{L}_z[\cdot] = \frac{f_0^2}{\rho_0} \frac{\partial}{\partial z} \left(\frac{\rho_0}{\Theta_1'} \frac{\partial \cdot}{\partial z} \right) \quad (3.11)$$

ein linearer Operator (siehe Päsche u. a., 2012, (3.12)) und q_{rel} die relative potentielle Vorticity ist. Da hier von einem Linienwirbel von

endlicher Zirkulation ausgegangen wird, weist das induzierte Geschwindigkeitsfeld entlang des Linienwirbels eine Singularität auf. Diese Betrachtungen sind analog zu Callegari und Ting (1978) und auch hier wird der Quellterm, die relative Vorticity, in einen regulären q_r und einen singulären q_s Anteil aufgespalten:

$$q_{\text{rel}} = q_s + q_r \quad (3.12)$$

(Päschke u. a., 2012, (3.12)). Bezeichne r den horizontalen Abstand zur Wirbelzentrumslinie und $\mathbf{X}(t, z)$ einen Punkt in der Nähe der Wirbelzentrumslinie, dann lässt sich jeder Punkt $\mathbf{x} = (x, y, z)$ in der Form

$$\mathbf{x} - \mathbf{X}(t, z) = r \mathbf{e}_r = r (\mathbf{i} \cos \theta + \mathbf{j} \sin \theta) \quad (3.13)$$

darstellen (ebd., (3.17)). Das ist zumindest dann der Fall, wenn bei gegebener Krümmung der Abstand zur Zentrumslinie klein genug ist, wovon hier immer ausgegangen wird. Der reguläre Anteil besitzt für $r \rightarrow 0$ dann einen endlichen Grenzwert. Der singuläre Anteil wird in Form einer δ -Distribution dargestellt:

$$q_s(t, \mathbf{x}, z) = \Gamma(t, z) \delta_2(\mathbf{x} - \mathbf{X}(t, z)) \quad (3.14)$$

(ebd., (3.14)). Hier bezeichnet $\delta_k(\cdot)$ die k -dimensionale Delta-Distribution und

$$\Gamma(t, z) = \Gamma^{(0)}(t, z) + \delta \Gamma^{(1)}(t, z) + o(\delta) \quad (3.15)$$

die totale Wirbelzirkulation als Maß für die Wirbelstärke (ebd., (3.15a)).

Um ψ zu bestimmen, wird in Päschke u. a. (ebd., § 3.3.1) sukzessive das dominante singuläre Verhalten eliminiert. Wie sich kürzlich herausstellte, ist dieses Vorgehen nicht ganz korrekt. Weiter unten wird beschrieben, wie stattdessen vorzugehen ist. Eine genaue Ausarbeitung ist in Arbeit und wird als Ergänzung zu Päschke u. a. (ebd.) veröffentlicht werden. Es soll nun kurz erläutert werden, wo das Problem in dem Vorgehen bei Päschke u. a. (ebd.) liegt.

Die Idee des sukzessiven Vorgehens, um das dominante singuläre Verhalten zu eliminieren, geht nicht auf, da das betrachtete Koordinatensystem nicht orthogonal ist, insbesondere steht die vertikale Koordinate nicht senkrecht zu einem horizontalen Schnitt durch den Wirbel. Das erzeugt in jedem Schritt einen singulären Anteil der Größenordnung $o(\delta)$. Für die vertikale Ableitung des horizontalen Abstands r zur Wirbelzentrumslinie gilt dann nach (3.13) und (3.4)

$$\frac{\partial}{\partial z} r(x, y, z, t) = \frac{\partial}{\partial z} \|\mathbf{x}_{\parallel} - \mathbf{X}(t, z)\| = \left(\mathbf{x}_{\parallel} - \mathbf{X}^{(1)}(t, z) \right) \frac{\delta}{r} \frac{\partial \mathbf{X}^{(1)}}{\partial z}. \quad (3.16)$$

Es entsteht also ein weiterer singulärer Anteil. Ausweg wird wie bei Callegari und Ting (1978) ein lokales orthogonales Koordinatensystem sein.

Es ist jedoch nicht zu erwarten, dass sich an dem Resultat Wesentliches ändern wird. Fährt man deshalb wie bei Päsche u. a. (2012) fort, erhält man nach einer aufwendigen Rechnung

$$\psi = -\frac{\Gamma}{2\pi} \ln r + M_0 r^2 (\ln r - 1) + \delta M_1 r \ln r - \delta^2 M_2 + \psi_2^* \quad (3.17)$$

(ebd., (3.25)). Hierbei ist ψ_2^* ein regulärer Anteil der Stromfunktion, der in $r = 0$ insbesondere stetig ist. Die Größen $M_0(t, z)$, $M_1(t, \theta, z) = -\mathbf{M}_1 \cdot \mathbf{e}_r$ und $M_2(t, \theta, z)$ sind wie in Päsche u. a. (ebd., (3.22)) durch

$$M_0 = \frac{1}{8\pi} \frac{f_0^2}{\rho_0} \frac{\partial}{\partial z} \left(\frac{\rho_0}{\Theta_1'} \frac{\partial \Gamma}{\partial z} \right), \quad (3.18a)$$

$$M_1 = -\frac{f_0^2}{4\pi\rho\Gamma} \frac{\partial}{\partial z} \left(\frac{\rho_0 \Gamma^2}{\Theta_1'} \frac{\partial \mathbf{X}^{(1)}}{\partial z} \right) \cdot \mathbf{e}_r \equiv -\mathbf{M}_1 \cdot \mathbf{e}_r, \quad (3.18b)$$

$$M_2 = \frac{f_0^2 \Gamma}{8\pi \Theta_1'} \left(\left[\mathbf{e}_\theta \cdot \frac{\partial \mathbf{X}^{(1)}}{\partial z} \right]^2 - \left[\mathbf{e}_r \cdot \frac{\partial \mathbf{X}^{(1)}}{\partial z} \right]^2 \right). \quad (3.18c)$$

gegeben. Damit hat man eine explizite Darstellung der Singularität auf der Wirbelzentrumslinie. Außerhalb der Wirbelzentrumslinie ist (3.17) eine asymptotische Approximation der geostrophischen Stromfunktion. Aus (3.9) und (3.17) ergibt sich für das horizontale Geschwindigkeitsfeld \mathbf{u} in synoptischen Koordinaten bis zu einer Ordnung $\mathcal{O}(\delta)$

$$\begin{aligned} E^{(1)} \mathbf{u} = \tilde{\mathbf{u}} = & \left(\frac{\Gamma}{2\pi r} - M_0 r (2 \ln r - 1) \right) \mathbf{e}_\theta \\ & \delta (\ln r \mathbf{k} \times \mathbf{M}_1 + (\mathbf{M}_1 \cdot \mathbf{e}_r) \mathbf{e}_\theta) + \mathbf{u}_{\text{reg}}, \end{aligned} \quad (3.19)$$

mit dem regulären Geschwindigkeitsfeld

$$\mathbf{u}_{\text{reg}} = \mathbf{u}_{\text{reg}}^{(0)} + \delta \mathbf{u}_{\text{reg}}^{(1/2)} = -\mathbf{k} \times \nabla_{\parallel} \psi_2^*. \quad (3.20)$$

Dabei bezeichnet $E^{(1)} \mathbf{u}$ die asymptotische Entwicklung von \mathbf{u} bis zur Ordnung $\mathcal{O}(\delta)$.

3.3 WIRBELKERNSTRUKTUR

Um das singuläre Verhalten genauer zu studieren, wird die Wirbelkernstruktur im Gradient-Wind-Regime betrachtet.

Für die Analyse wird ein mitbewegtes Koordinatensystem eingeführt, das es erlaubt, den mesoskaligen Bereich um die Wirbelzentrumslinie aufzulösen. Die Wirbelzentrumslinie aus (3.4) wird dabei um einen Term der Größenordnung des Wirbelkerndurchmessers gestört,

$$\mathbf{x} = \mathbf{X}(t, z) + \delta \hat{\mathbf{x}} = \mathbf{X}_0(t) + \delta \left(\mathbf{X}^{(1)}(t, z) + \hat{\mathbf{x}} \right). \quad (3.21)$$

Um die Achsensymmetrien des Wirbels beschreiben zu können, wird $\widehat{\mathbf{x}} = \widehat{x} \mathbf{i} + \widehat{y} \mathbf{j}$ umgeschrieben

$$\widehat{x} = \widehat{r} \cos \theta, \quad \mathbf{i} = \mathbf{e}_r \cos \theta - \mathbf{e}_\theta \sin \theta, \quad (3.22a)$$

$$\widehat{y} = \widehat{r} \sin \theta, \quad \mathbf{j} = \mathbf{e}_r \sin \theta + \mathbf{e}_\theta \cos \theta, \quad (3.22b)$$

(Päschke u. a., 2012, (4.3)) wobei \mathbf{e}_r und \mathbf{e}_θ die Einheitsvektoren in Polarkoordinaten sind. Zusammen mit (3.13) ergibt sich

$$\widehat{r} = r/\delta. \quad (3.23)$$

Für das Geschwindigkeitsfeld als zeitliche Änderung des Ortes erhält man aus (3.21) und (3.22) eine Zerlegung in die Wirbelbewegung und eine relative Geschwindigkeit

$$\mathbf{u} = \mathbf{U} + \mathbf{u}_{\text{rel}} = \frac{\partial \mathbf{X}}{\partial t} + (u_r \mathbf{e}_r + u_\theta \mathbf{e}_\theta) \quad (3.24)$$

mit der Umfangsgeschwindigkeit u_θ und einer Geschwindigkeitskomponenten in radialer Richtung u_r . Dabei gilt für die horizontalen Ableitungsoperatoren der synoptischen Skala ∇_{\parallel} und der Mesoskala $\widehat{\nabla}$ die Transformationsformel

$$\nabla_{\parallel} = \delta^{-1} \left(\mathbf{e}_r \frac{\partial}{\partial \widehat{r}} + \mathbf{e}_\theta \frac{1}{\widehat{r}} \frac{\partial}{\partial \theta} \right) = \delta^{-1} \widehat{\nabla} \quad (3.25)$$

(siehe ebd., (4.6)). Für u_θ und u_r werden die asymptotischen Ansätze

$$u_\theta(t, \mathbf{x}, z; \varepsilon) = \delta^{-1} u_\theta^{(0)}(t, \widehat{r}, z) + u_\theta^{(1)}(t, \widehat{r}, z) \quad (3.26a)$$

$$+ \delta u_\theta^{(2)}(t, \widehat{r}, \theta, z) + \mathcal{O}(\delta), \quad (3.26b)$$

$$u_r(t, \mathbf{x}, z; \varepsilon) = \delta u_r^{(2)}(t, \widehat{r}, \theta, z) + \mathcal{O}(\delta), \quad (3.26c)$$

gewählt (siehe ebd., (4.12)). Der asymptotische Ansatz für u_θ erlaubt hohe Umfangsgeschwindigkeiten und geht von einer Achsensymmetrie, also einer Winkelunabhängigkeit der Umfangsgeschwindigkeit, aus, die erst ab der Ordnung δ durch eine Winkelabhängigkeit gestört wird. Nach einiger Rechnung gelangt man so zu einer Evolutionsgleichung für die achsensymmetrische Umfangsgeschwindigkeit in führender Ordnung $u_\theta^{(0)}$,

$$\begin{aligned} \frac{\partial u_\theta^{(0)}}{\partial t} + w_0^{(1)} \frac{\partial u_\theta^{(0)}}{\partial z} + u_{r,00}^{(2)} \left(\frac{\partial u_\theta^{(0)}}{\partial \widehat{r}} + \frac{u_\theta^{(0)}}{\widehat{r}} + f_0 \right) \\ = u_{r,*}^{(2)} \left(\frac{u_\theta^{(0)}}{\widehat{r}} \right) \end{aligned} \quad (3.27)$$

(siehe ebd., (4.21)). Hierbei bezeichnen $u_{r,00}^{(2)}$ und $w_0^{(1)}$ die achsensymmetrischen Anteile der radialen und vertikalen Geschwindigkeiten, die direkt durch achsensymmetrische diabatische Erwärmung induziert werden. Der Quellterm auf der rechten Seite kommt durch

nichtlineare Interaktion der Wirbelneigung mit asymmetrischer diabatischer Erwärmung zustande. Die Größe $u_{r,*}^{(2)}$ hängt von der relativen Ausrichtung der Wirbelneigung und der Asymmetrie der diabatischen Erwärmung ab und ist durch

$$u_{r,*}^{(2)} = \frac{1}{2} \left[\frac{\partial X^{(1)}}{\partial z} \frac{Q_{\Theta,12}}{\Theta_1'} + \frac{\partial Y^{(1)}}{\partial z} \frac{Q_{\Theta,11}}{\Theta_1'} \right] \quad (3.28)$$

gegeben (ebd., (1.3d)).

Bei der Wechselwirkung zwischen Wirbelbewegung und Wirbelkernstruktur spielen die asymmetrischen Anteile von $u_{\theta}^{(2)}$ und $u_r^{(2)}$ eine entscheidende Rolle. Für eine genaue Herleitung der asymmetrischen Anteile sei auf Päschke u. a. (ebd., §4) verwiesen. Allgemein wird für winkelabhängige Funktionen F die Fourierzerlegung in einen symmetrischen (winkelunabhängigen) und einen asymmetrischen (winkelabhängigen) Anteil

$$F(\theta) = F_0 + \sum_n (F_{n1} \sin(n\theta) + F_{n2} \cos(n\theta)) \quad (3.29)$$

aufgeteilt. Für die asymmetrischen Anteile von $u_{\theta}^{(2)}$ und $u_r^{(2)}$, ausgedrückt in einem Geschwindigkeitspotential $\phi^{(2)}$ und einer Stromfunktion $\psi^{(2)}$, bedeutet das

$$u_r^{(2)} - u_{r,0}^{(2)} = \frac{1}{\hat{r}} \frac{\partial \psi^{(2)}}{\partial \theta} + \frac{\partial \phi^{(2)}}{\partial \hat{r}} \quad (3.30a)$$

$$u_{\theta}^{(2)} - u_{\theta,0}^{(2)} = -\frac{\partial \psi^{(2)}}{\partial \hat{r}} + \frac{1}{\hat{r}} \frac{\partial \phi^{(2)}}{\partial \theta} \quad (3.30b)$$

(siehe ebd., (4.28)). Für die asymptotische Entwicklung des horizontalen Geschwindigkeitsfeldes bis zu einer Ordnung von $\mathcal{O}(\delta)$ ergibt sich zusammen mit (3.24), (3.25) und (3.26)

$$\begin{aligned} \hat{\mathbf{E}}^{(1)} \mathbf{T} \mathbf{u} = & \frac{d\mathbf{X}_0}{dt} + \delta \frac{\partial \mathbf{X}^{(1)}}{\partial t} + \frac{1}{\delta} \left(u_{\theta}^{(0)} + \delta u_{\theta}^{(1)} + \delta^2 u_{\theta,0}^{(2)} \right) \\ & \mathbf{e}_{\theta} + \delta u_{r,0}^{(2)} \mathbf{e}_r + \delta \left(\hat{\nabla} \phi^{(2)} - \mathbf{k} \times \hat{\nabla} \psi^{(2)} \right). \end{aligned} \quad (3.31)$$

Hierbei steht \mathbf{T} für die Transformation in das mitbewegte mesoskalige Koordinatensystem (3.21). Mit $u_{r,00}^{(2)}$ und $u_{r,*}^{(2)}$ aus (3.27) gilt

$$u_{r,0}^{(2)} = u_{r,00}^{(2)} + u_{r,*}^{(2)}. \quad (3.32)$$

Damit hat man eine asymptotische Darstellung des horizontalen Geschwindigkeitsfeldes auf der Mesoskala.

3.4 MATCHING

Mit (3.19) und (3.31) sind zwei Darstellungen für das horizontale Geschwindigkeitsfeld \mathbf{u} gegeben, einmal auf der synoptischen Skala

und einmal auf der Mesoskala. Ziel ist es nun, daraus Bewegungsgleichungen für die Wirbelzentrumslinie unter Berücksichtigung der Wirbelkernstruktur herzuleiten, die insbesondere die asymmetrische diabatische Erwärmung berücksichtigen. Die dazu verwendete Methode geht auf die Matching-Prinzipien von van Dyke (1975) zurück, in einer Form wie sie in Eckhaus (1979) zu finden ist.

Loosely speaking, van Dyke's matching principle requires "the near-field behaviour of the outer solution to match the farfield behavior of the inner solution" when both are expressed in the inner coordinates. (Päschke u. a., 2012, S. 17)

Dazu wird (3.19) in den mesoskaligen Koordinaten ausgedrückt und bis zu der Ordnung $\mathcal{O}(\delta)$ entwickelt,

$$\widehat{\mathbf{E}}^{(1)} \mathbf{T} \mathbf{E}^{(1)} \mathbf{u} \quad (3.33a)$$

$$= \mathbf{u}_0^{(0)} + \delta \left(\mathbf{u}_0^{(1/2)} + \mathbf{X}^{(1)} \cdot (\nabla_{\parallel} \mathbf{u}^{(0)})_0 \right) \quad (3.33b)$$

$$+ \left(\frac{1}{\delta} \frac{\Gamma}{2\pi\widehat{r}} - \delta M_0 \widehat{r} (2 \ln \widehat{r} - 1 + \ln \delta) \right) \mathbf{e}_{\theta} \quad (3.33c)$$

$$+ \delta \left(\widehat{\mathbf{r}} \mathbf{e}_r \cdot (\nabla_{\parallel} \mathbf{u}^{(0)})_0 \right) \quad (3.33d)$$

$$+ \delta \left(\left[\ln \widehat{r} - \ln \frac{1}{\delta} \right] \mathbf{k} \times \mathbf{M}_1 + (\mathbf{M}_1 \cdot \mathbf{e}_r) \mathbf{e}_{\theta} \right) \quad (3.33e)$$

(ebd., (5.6)). Dabei bezeichnet $\widehat{(\cdot)}$ das mesoskalige Koordinatensystem und das Superscript $(\cdot)^{(1)}$ eine Entwicklung bis zur Ordnung $\mathcal{O}(\delta)$. Der erste $\mathcal{O}(\delta)$ -Term resultiert aus einer Zwei-Term Taylorentwicklung des regulären Geschwindigkeitsfeldes aus (3.20) unter Berücksichtigung von (3.24).

Außerdem drückt man (3.31) in den Koordinaten der synoptischen Skala aus, entwickelt ebenfalls bis zu einer Genauigkeit von $\mathcal{O}(\delta)$ und transformiert dann abermals in das mesoskalige Koordinatensystem und entwickelt wieder bis zur Ordnung $\mathcal{O}(\delta)$

$$\widehat{\mathbf{E}}^{(1)} \mathbf{T} \mathbf{E}^{(1)} \mathbf{T}^{-1} \widehat{\mathbf{E}}^{(1)} \mathbf{T} \mathbf{u} \quad (3.34a)$$

$$= \frac{d\mathbf{X}_0}{dt} + \delta \frac{\partial \mathbf{X}^{(1)}}{\partial t} \quad (3.34b)$$

$$+ \widehat{\mathbf{E}}^{(1)} \mathbf{T} \mathbf{E}^{(1)} \mathbf{T}^{-1} \left[\frac{1}{\delta} \left(\mathbf{u}_{\theta}^{(0)} + \delta \mathbf{u}_{\theta}^{(1)} + \delta^2 \mathbf{u}_{\theta,0}^{(2)} \right) \right] \mathbf{e}_{\theta} \quad (3.34c)$$

$$- \widehat{\mathbf{E}}^{(1)} \mathbf{T} \mathbf{E}^{(1)} \mathbf{T}^{-1} \left[\delta \mathbf{k} \times \widehat{\nabla} \tilde{\psi}_2^{(2)} \right] \quad (3.34d)$$

$$+ \delta (\ln \widehat{r} \mathbf{k} \times \mathbf{M}_1 + (\mathbf{M}_1 \cdot \mathbf{e}_r) \mathbf{e}_{\theta} - \mathbf{k} \times \Psi) \quad (3.34e)$$

(ebd., (5.14)). Hierbei bezeichne \mathbf{T} wieder die Transformation in das mesoskalige Koordinatensystem und \mathbf{T}^{-1} bezeichne die inverse Transformation in das synoptische Koordinatensystem. Weiter ist

$$\tilde{\psi}_2^{(2)} = \psi_{21}^{(2)} \sin(2\theta) + \psi_{22}^{(2)} \cos(2\theta) \quad (3.35)$$

(siehe 3.29) und ψ ein komplexer Ausdruck, der sowohl die Wirbelkernstruktur, als auch die asymmetrische Erwärmung und die Wirbelneigung beinhaltet (siehe ebd., (5.12)). Der Herleitung der letzten Gleichung geht eine sehr komplexe Rechnung voraus (siehe ebd., §4.4, §5.2.2).

Van Dykes *Matching-Prinzip* in der Form von Eckhaus (1979) postuliert die Gleichheit von (3.33e) und (3.34e). Ein Vergleich der einzelnen Terme in den jeweiligen Ordnungen ergibt insbesondere die Bewegungsgleichungen für die Wirbelzentrumslinie unter Berücksichtigung der Wirbelkernstruktur. Für Details und weitere Ergebnisse aus dem Matching sei auf Päsche u. a. (2012, § 5.2.3) verwiesen. Die Bewegungsgleichungen sind nach (3.33e) und (3.34e) durch

$$\frac{d\mathbf{X}_0}{dt} = \mathbf{u}_0^{(0)} \quad (3.36a)$$

$$\frac{\partial \mathbf{X}^{(1)}}{\partial t} = \mathbf{u}_0^{(1/2)} + \mathbf{X}^{(1)} \cdot (\nabla_{\parallel} \mathbf{u}^{(0)})_0 \quad (3.36b)$$

$$- \ln \frac{1}{\delta} (\mathbf{k} \times \mathbf{M}_1) + (\mathbf{k} \times \Psi) \quad (3.36c)$$

gegeben.

Aus den Bewegungsgleichungen für die Wirbelbewegung lassen sich zwei Konsequenzen ableiten. Erstens kann ein Wirbel unter den betrachteten Bedingungen einer zu starken Scherung der Hintergrundströmung nicht standhalten. Zweitens gibt es eine selbstinduzierte Bewegung des geneigten baroklinen Wirbels, die Auslenkungen in der Größenordnung des Wirbelkerns induziert. Vergleichbar einem Schwimmer, der sich aufgrund seiner Körperbewegung im Wasser fortbewegen kann, unabhängig von der Strömung des Gewässers.

3.5 ALTERNATIVE BESTIMMUNG DER STROMFUNKTION

Wie weiter oben schon erwähnt ist das Vorgehen sukzessive die dominanten singulären Anteile der Stromfunktion zu eliminieren, nicht ganz korrekt. Es wurde erläutert, dass das Problem auf ein nicht orthogonales Koordinatensystem zurückzuführen ist. Um das zu umgehen betrachtet man ein an die Wirbelzentrumslinie geheftetes lokales Koordinatensystem, das im Detail von Callegari und Ting (1978, Appendix A) beschrieben wird.

3.5.1 Lokales mitbewegtes orthogonales Koordinatensystem

Sei C eine Kurve in dem orthogonalen Koordinatensystem auf der synoptischen Skala mit den Basisvektoren $\mathbf{i}, \mathbf{j}, \mathbf{k}$ und dem Parameter z zur Zeit $t = 0$. Dann kann C zu jeder Zeit t durch $\mathbf{X}(z, t)$ parametri-

siert werden. Zu C gehören dann der Tangentialvektor $\boldsymbol{\tau}$, der Normalenvektor \mathbf{n} und der Binormalenvektor \mathbf{b} , die durch die Serret-Frenet Formeln gegeben sind.

$$\frac{\partial \mathbf{X}}{\partial z} = \boldsymbol{\tau} \sigma, \quad \frac{\partial \boldsymbol{\tau}}{\partial z} = \kappa \sigma \mathbf{n}, \quad (3.37a)$$

$$\frac{\partial \mathbf{b}}{\partial z} = -T \sigma \mathbf{n}, \quad \frac{\partial \mathbf{n}}{\partial z} = (T \mathbf{b} - \kappa \boldsymbol{\tau}) \sigma, \quad (3.37b)$$

mit

$$\sigma = \frac{\partial \tilde{z}}{\partial z} = \left\| \frac{\partial \mathbf{X}}{\partial z} \right\|, \quad \kappa = \frac{1}{\sigma} \left\| \frac{\partial \boldsymbol{\tau}}{\partial z} \right\|, \quad T = \frac{-1}{\sigma} \frac{\partial \mathbf{b}}{\partial z} \cdot \mathbf{n}, \quad (3.38)$$

wobei $\tilde{z}(z, t)$ die Bogenlänge zur Zeit t gemessen von $z = 0$ ist.

Sei $\mathbf{x} = \xi_1 \mathbf{i} + \xi_2 \mathbf{j} + z \mathbf{k}$ der Ortsvektor zu einem beliebigen Punkt und bezeichne \mathbf{X} wie in (3.4) einen Punkt auf der Wirbelzentrumslinie, dann wird \mathbf{x} in der Form

$$\mathbf{x} = \mathbf{X}(z_{\perp}, t) + r_{\perp} \mathbf{r}_{\perp} + z_{\perp} \mathbf{k} \quad (3.39)$$

geschrieben, wobei \mathbf{r}_{\perp} ein Einheitsvektor in der $\mathbf{n}\mathbf{b}$ -Ebene ist und z_{\perp} so gewählt ist, dass

$$\|\mathbf{X}(z_{\perp}, t) + z_{\perp} \mathbf{k} - \mathbf{x}\| \quad (3.40)$$

minimal ist. Sei φ der Winkel zwischen \mathbf{n} und \mathbf{r}_{\perp} dann sind $(r_{\perp}, \theta_{\perp}, z_{\perp})$ orthogonale Koordinaten, wenn θ_{\perp} durch

$$\theta_{\perp} = \varphi - \theta_0(z_{\perp}, t) \quad (3.41)$$

mit

$$\frac{\partial \theta_0}{\partial z_{\perp}} = -\sigma T \quad (3.42)$$

definiert ist (Callegari und Ting, 1978, Appendix A).

3.5.2 Vorgehen zum Lösen der elliptischen Gleichung von ψ

Der Trick zum Lösen von (3.10) besteht aus zwei Schritten.

Der erste Schritt besteht aus einer Umskalierung der Vertikalkoordinate gemäß der Transformation

$$dz = \frac{1}{\rho^{(0)}} d\zeta. \quad (3.43)$$

Mit $Q(z) = f_0^2 \rho^{(0)} / \Theta_1'(z)$ wird dann aus (3.10) die Gleichung

$$\nabla_{\xi, \zeta} \cdot (D \nabla_{\xi, \zeta} \psi) = -q_{\text{rel}} \quad (3.44)$$

wobei $D = \text{diag}(1, 1, Q(z)\rho^{(0)}(z))$ eine Diagonalmatrix ist.

Der zweite Schritt besteht darin die letzte Gleichung in den oben eingeführten $(r_{\perp}, \theta_{\perp}, z_{\perp})$ -Koordinaten zu schreiben und darin zu lösen.

Eine genaue Ausarbeitung der Schritte wird als Ergänzung zu Päsche u. a. (2012) veröffentlicht werden.

Teil II

STEIFES VERHALTEN VON DIFFERENTIALGLEICHUNGEN

1

EINLEITUNG

Eine Vielzahl mathematischer Modelle aus unterschiedlichen Bereichen, wie den Natur- oder Wirtschaftswissenschaften oder dem Ingenieurwesen, lassen sich durch Differentialgleichungen beschreiben. Häufig ist man an zeitlichen Zustandsänderungen interessiert, die sich durch gewöhnliche Differentialgleichungen beschreiben lassen. Nur wenige dieser Gleichungen lassen sich analytisch lösen. Ist dies nicht möglich, versucht man qualitative Aussagen über das Lösungsverhalten zu treffen oder man sucht Näherungslösungen. Solche Näherungslösungen können durch numerische Verfahren oder auch eine asymptotische Analyse gewonnen werden. Zu den einfachsten numerischen Verfahren gehören die expliziten Verfahren, das einfachste unter ihnen ist das explizite Euler-Verfahren. Anfang der 1950er Jahre hat man spezielles Augenmerk auf gewöhnliche Differentialgleichungen gelegt, die nur mit besonders hohem Aufwand mit expliziten Verfahren numerisch genähert werden konnten. Sie wurden als steife Differentialgleichungen bezeichnet. Der Name kommt daher, dass solche Systeme auf Auslenkungen aus einer Ruhelage durch sehr starkes zurückschnellen antworten, um „steif“ in ihrer Ruhelage zu verharren. Was dies genau bedeutet wird in den folgenden Kapiteln deutlich. Es gibt auch noch einen anderen Typ von Gleichungen die numerisch anspruchsvoll sind, nämlich solche, die zu oszillatorischen Problemen gehören. Derartige Probleme werden in dieser Arbeit aber nicht betrachtet.

Zur selben Zeit etablierte sich eine ganz andere Methode um Näherungslösungen zu finden, die Störungsrechnung bzw. die asymptotische Analyse – die Anfänge der asymptotischen Analyse gehen auf Poincaré und Lindstedt zurück und sind älter. Diese beiden Ansätze sind vollkommen verschieden und wurden lange Zeit völlig getrennt voneinander genutzt. Erst in den letzten Jahren wird versucht beide Ansätze parallel zu nutzen um das Lösungsverhalten von Differentialgleichungen besser beschreiben zu können.

Es gibt Bemühungen diese Ansätze zu verbinden, wie zum Beispiel in der „Computational Singular Perturbation“-Methode (Lam und Goussis, 1989). In dieser Arbeit wird aber ein anderer Weg gegangen. Steifes Verhalten von Differentialgleichungen soll ausschließlich anhand der Struktur der zugrunde liegenden Gleichungen beschrieben werden, ohne Begriffe aus der Numerik zu verwenden, wie beispielsweise Diskretisierung, Anzahl von Rechenschritten oder ähnliches.

In Abschnitt 2 geht es zunächst um steife Differentialgleichungen im klassischen Sinne, bevor der Bezug zur Asymptotik hergestellt

wird. Abschnitt 3 liefert eine mathematische Definition für steife Lösungen von Familien von Differentialgleichungen, die in Abschnitt 4 zusammen mit Arbeiten aus der einschlägigen Literatur diskutiert wird.

2

MOTIVATION

Inhalt

2.1	Steife Differentialgleichungen im klassischen Sinne	33
2.2	Familien von Differentialgleichungen	41
2.3	Formalisierung steifer Differentialgleichungen	55

In Abschnitt 2.1 wird geklärt was man im klassischen Sinne unter steifen Differentialgleichungen versteht. Dabei geht es zunächst vor allem um die historisch erste Sichtweise aus dem Blickwinkel der Numerik, die an Beispielen verdeutlicht wird. In Abschnitt 2.2 wird durch die Parametrisierung von gewöhnlichen Differentialgleichungen die Brücke zu einer asymptotischen Sichtweise geschlagen, in der statt einzelner Differentialgleichungen ganze Familien von Differentialgleichungen betrachtet werden. In Abschnitt 2.3 wird die Schwierigkeit einer von der Numerik losgelösten mathematischen Formalisierung steifer Differentialgleichungen an einer Vielzahl von Zitaten dargelegt.

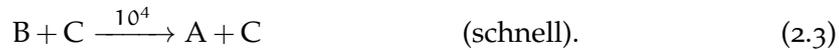
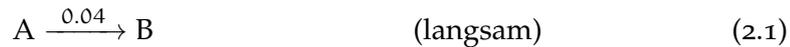
2.1 STEIFE DIFFERENTIALGLEICHUNGEN IM KLAS- SISCHEN SINNE

Historisch geht der Begriff der steifen Differentialgleichung auf Curtiss und Hirschfelder (1952) zurück.

Sie [Curtiss und Hirschfelder] führten den Begriff der steifen Differentialgleichungen (engl. stiff differential equations) für solche chemischen Reaktionen ein, bei denen sich die schnell reagierenden Komponenten in sehr kurzer Zeit einem Gleichgewichtszustand nähern, wenn die langsam veränderlichen Komponenten eingefroren, d. h. „steif“ gemacht werden. (Strehmel, Weiner und Podhaisky, 2012, S. 203).

Als Beispiel für solche chemischen Reaktionen wird das Reaktionsmodell nach Robertson (Hairer und Wanner, 2010, S. 3) betrachtet.

Beispiel 2.1. Das Reaktionsmodell nach Robertson kann durch drei Reaktionsgleichungen beschrieben werden.



Die erste Reaktionsgleichung beschreibt die Stoffumwandlung eines Stoffes A mit einer zeitlichen Rate von 0.04 in einen Stoff B. Dies bedeutet genauer, dass Stoff A mit einer Rate von 0.04 abgebaut, während mit der selben Rate Stoff B gewonnen wird. Aus der dritten Reaktionsgleichung geht hervor, dass die Stoffe B und C mit der Rate 10^4 zu den Stoffen A und C umgewandelt werden. An der Bilanz von Stoff C ändert sich durch diese Reaktion nichts. Die zweite Reaktionsgleichung gibt an, dass zwei Anteile des Stoffes B in einen Anteil des Stoffes C und einen Anteil des Stoffes B umgewandelt werden, und das mit einer Rate von $3 \cdot 10^7$. Das Entscheidende an dieser Stelle ist, dass hier drei Reaktionen mit sehr unterschiedlichen Reaktionsgeschwindigkeiten ablaufen. Mit entsprechenden Startwerten ergibt sich das Anfangswertproblem

$$\dot{x}_1 = -0.04 x_1 + 10^4 x_2 x_3, \quad x_1(0) = 1, \quad (2.4)$$

$$\dot{x}_2 = 0.04 x_1 - 10^4 x_2 x_3 - 3 \cdot 10^7 x_2^2, \quad x_2(0) = 0, \quad (2.5)$$

$$\dot{x}_3 = 3 \cdot 10^7 x_2^2, \quad x_3(0) = 0. \quad (2.6)$$

Für eine formale Beschreibung wie man von einem System von Reaktionsgleichungen auf ein System von Differentialgleichungen gelangt, sei auf Aiken (1985, S. 44) verwiesen.

Für sämtliche in dieser Arbeit gerechneten Beispiele wurde das Assimulo Framework (Andersson, Führer und Åkesson, 2014) unter Python verwendet. Ohne näher auf die numerische Lösung einzugehen (siehe dazu auch Hairer und Wanner (2010, S. 3)), betrachte man die zeitliche Entwicklung der drei Komponenten in den Abbildungen 2.1 bis 2.2 auf den Seiten 35–36. Abbildung 2.1 zeigt das Verhalten über einen „langen“ Zeitraum bis zur Zeit $t = 1$. Alle drei Komponenten verhalten sich nahezu linear. Die ersten beiden fallen, die dritte wächst. In Abbildung 2.2 ist der Verlauf bis zur Zeit $t = 0.01$ dargestellt. Das Verhalten der ersten und dritten Komponente in Abbildung 2.2 ist fast identisch zu dem Verhalten der jeweiligen Komponente in Abbildung 2.1. Ganz anders verhält sich die zweite Komponente. In Abbildung 2.2 gibt es einen rapiden Anstieg in kurzer Zeit, danach verläuft der Graph wieder so wie in Abbildung 2.1. Im Gegensatz zu dem rapiden Anstieg der zweiten Komponente sind die erste und dritte Komponente über die ganze Integrationszeit hinweg „eingefroren“ und ändern ihr nahezu lineares Verhalten nicht. Diese Beobachtungen gelten als charakteristisch für steife Differentialgleichungen.

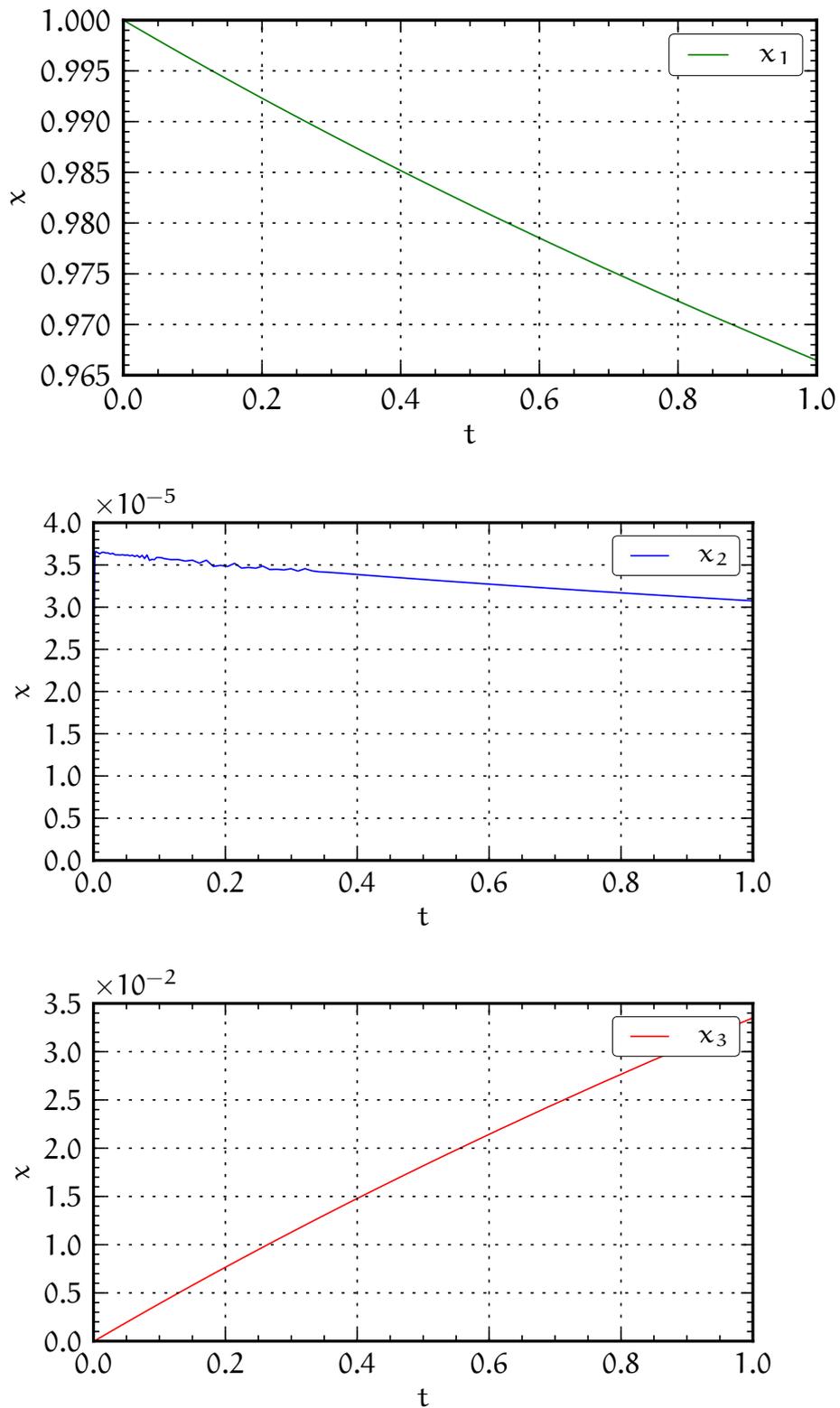


Abbildung 2.1: Chemische Reaktion nach Robertson (langsame Dynamik).

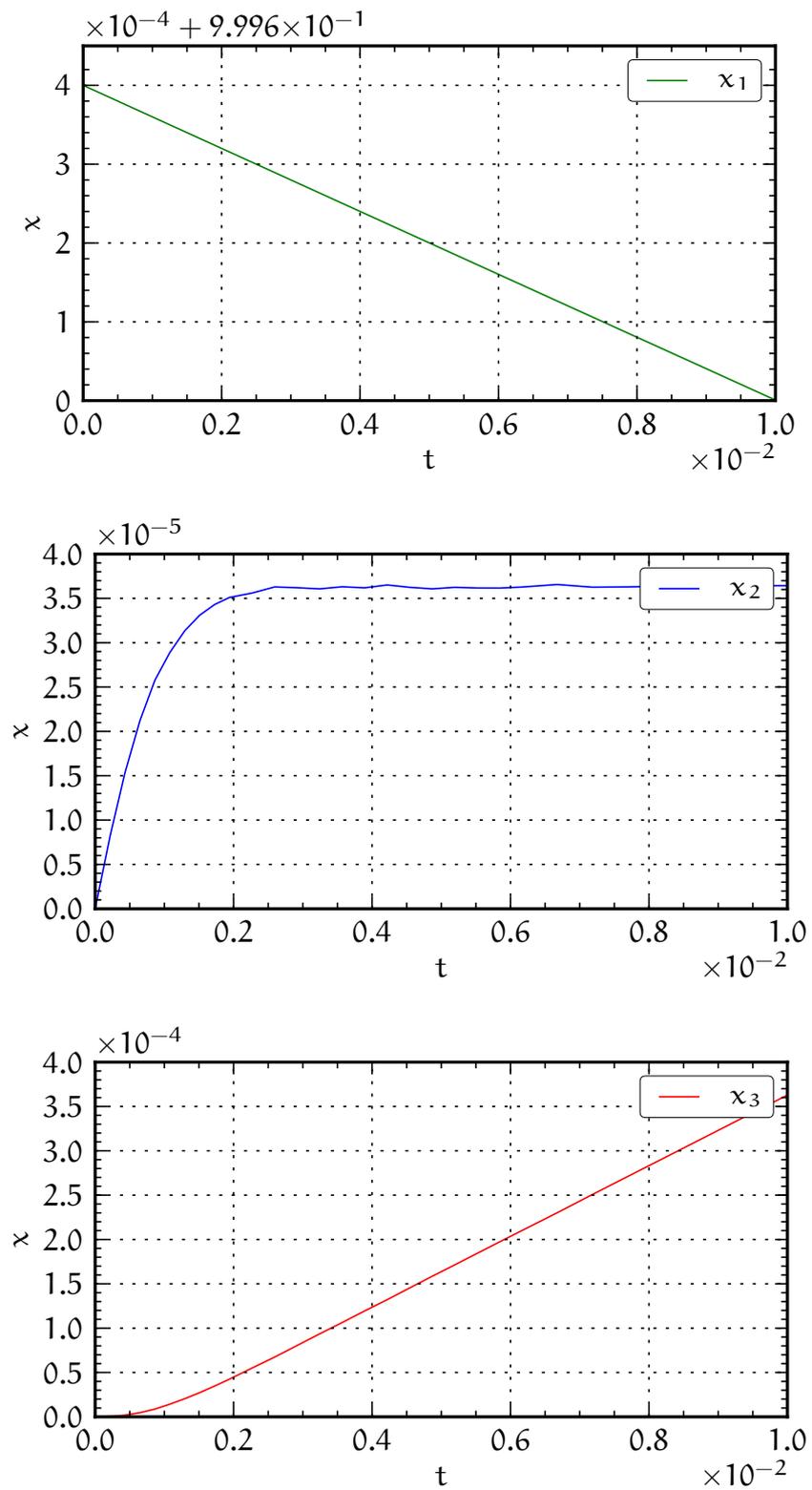


Abbildung 2.2: Chemische Reaktion nach Robertson (schnelle Dynamik).

Um solches Verhalten bei gewöhnlichen Differentialgleichungen zu beschreiben, gaben Curtiss und Hirschfelder (1952) die historisch erste Definition für steife Differentialgleichungen.

Curtiss und Hirschfelder (ebd.) betrachten Differentialgleichungen vom Typ

$$\frac{dy}{dx} = [y - G(x)]/a(x, y), \quad (2.7)$$

wobei die rechte Seite für eine allgemeine Funktion von x und y steht, die für jeden Wert von x die Nullstelle $y = G(x)$ besitzt. Weiterhin soll $G(x)$ deutlich schwächer in x variieren als $\exp(x/a(x, G(x)))$. Was genau das bedeuten soll wird nicht klar. Für eine vorgegebene gewünschte numerische Auflösung Δx von x nennen sie die Gleichung steif, falls die folgende Bedingung erfüllt ist

$$\left| \frac{a(x, y)}{\Delta x} \right| \ll 1. \quad (2.8)$$

Diese Definition hängt also auch von einer vorgegebenen numerischen Auflösung ab und nicht allein von der mathematischen Struktur der Gleichung.

Um den Zusammenhang der historischen Definition nach Curtiss und Hirschfelder (ebd.) und Beispiel 2.1 zu verstehen, werden zunächst einfachere Beispiele betrachtet, die das wesentliche Verhalten steifer Differentialgleichungen herausstellen.

Beispiel 2.2. Betrachtet wird ein einfacher Zerfallsprozess, bei dem die Abnahme einer bestimmten Stoffmenge $-x$ proportional zu der Stoffmenge x selber ist. Dieser Zerfallsprozess genügt dem Anfangswertproblem

$$\dot{x} = -\frac{1}{\varepsilon}x, \quad x(0) = x_0, \quad (2.9)$$

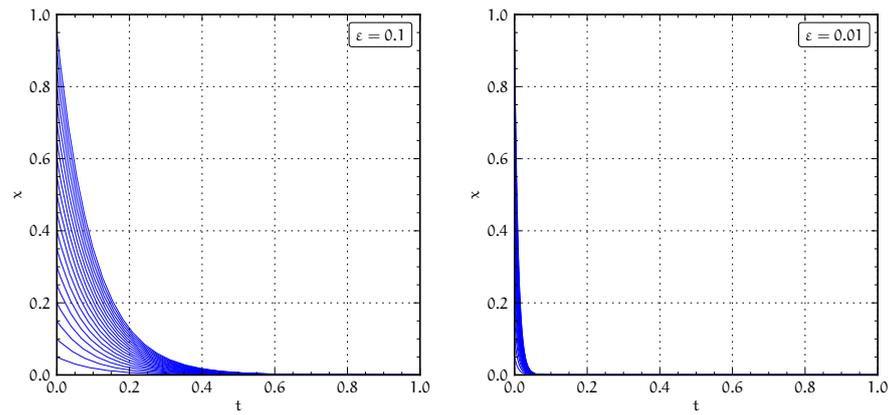
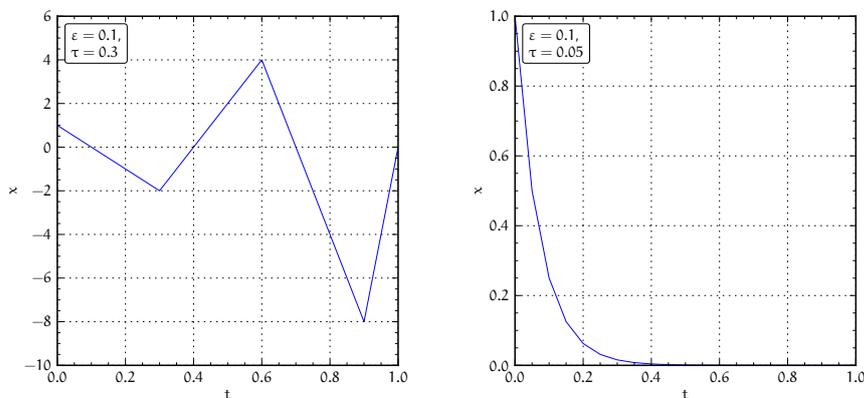
mit $\varepsilon > 0$. Diese Gleichung kann als Spezialfall von (2.7) angesehen werden. Die exakte Lösung

$$x(t) = x_0 \exp\left(-\frac{t}{\varepsilon}\right) \quad (2.10)$$

fällt für $x_0 \neq 0$ exponentiell und zwar umso schneller, je kleiner ε ist. Der Punkt $x_0 = 0$ ist ein Fixpunkt. Abbildung 2.3 auf der nächsten Seite zeigt für zwei Werte von ε Lösungskurven für verschiedene Startwerte. Selbst für kleine Startwerte schmiegen sich die Lösungskurven, anders als in Abbildung 2.5 auf Seite 39, rasch aneinander.

Wählt man nun für eine numerische Näherungslösung das explizite Euler-Verfahren auf einem äquidistanten Gitter der Schrittweite $\tau > 0$, berechnen sich die approximierten Werte an den Gitterpunkten für $x_0 = 1$ durch

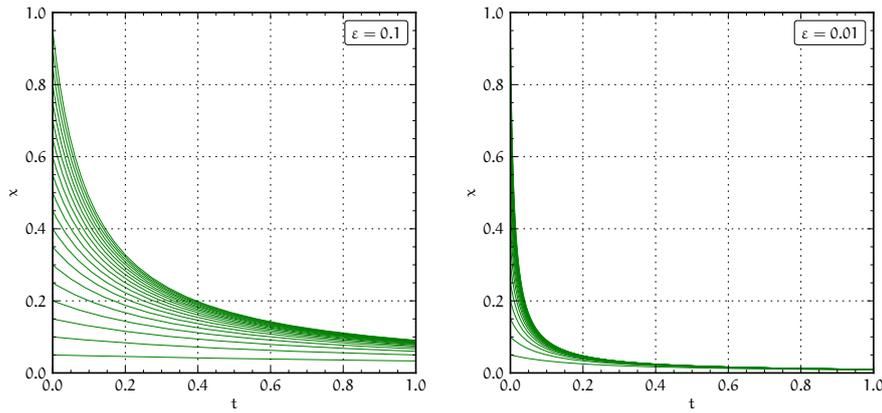
$$x_{k+1} = \left(1 - \frac{\tau}{\varepsilon}\right)^k, \quad k \in \mathbb{N}_0. \quad (2.11)$$

Abbildung 2.3: $\dot{x} = -\frac{1}{\epsilon}x$.Abbildung 2.4: Explizites Euler-Verfahren für $\dot{x} = -\frac{1}{\epsilon}x$.

Für $\tau > \epsilon$ spiegelt die Näherung das monotone Fallen der exakten Lösung nicht wider und x_k oszilliert um 0. Die Oszillation ist umso stärker, je kleiner ϵ ist. Für $\tau > 2\epsilon$ oszilliert die Näherungslösung nicht nur, sie explodiert auch, siehe Abbildung 2.4. Übersetzt in die Notation von Curtiss und Hirschfelder (1952) bedeutet $\tau \gg \epsilon$, dass (2.8) erfüllt ist. Nach der Definition von Curtiss und Hirschfelder (ebd.) ist (2.9) also steif.

Es ist überraschend, dass das numerische Verfahren bei einer exakten Lösung, die nach kurzer Zeit fast identisch der Nullfunktion ist, solche Probleme hat. Das ist ein typisches Verhalten steifer Differentialgleichungen und passt zu einer Charakterisierung von Simeon (2003).

Ein System gewöhnlicher Differentialgleichungen ist steif, wenn explizite Verfahren aus Stabilitätsgründen extrem kleine Schrittweiten verwenden, obwohl sich die Lösung kaum ändert; implizite Verfahren dagegen große Schrittweiten einsetzen können. (ebd.)

Abbildung 2.5: $\dot{x} = -\frac{1}{\epsilon}x^2$.

Für (2.10) ergibt sich aus der Definition der Exponentialfunktion als Grenzwert einer Folge und der Anzahl der Iterationen $j = t/\tau$

$$\exp\left(-\frac{t}{\epsilon}\right) = \lim_{j \rightarrow \infty} \left(1 - \frac{t}{\epsilon j}\right)^j = \lim_{\tau \rightarrow 0} \left(1 - \frac{\tau}{\epsilon}\right)^{\frac{t}{\tau}}. \quad (2.12)$$

Diese Darstellung lässt erkennen, dass eine gute Näherung ein kleines τ erzwingt, das um so kleiner sein muss, je kleiner ϵ ist. Integriert man in die andere Richtung und ist $t < 0$ ist der numerische Aufwand zu erwarten, da sich die Lösung in sehr kurzer Zeit stark ändert. Im heute üblichen Terminus wird die Gleichung im Falle $t < 0$ deshalb nicht als steif bezeichnet. Curtiss und Hirschfelder (1952) haben noch nicht zwischen Integration in positiver bzw. negativer Richtung unterschieden.

Beispiel 2.3. Als Verallgemeinerung von (2.9) betrachte man nun für $\alpha > 1$ und $\epsilon > 0$ Anfangswertprobleme vom Typ

$$\dot{x} = -\frac{1}{\epsilon}x^\alpha, \quad x(0) = x_0, \quad (2.13)$$

die durch

$$x(t) = \frac{1}{\left(\frac{t}{\epsilon}(\alpha - 1) + x_0^{1-\alpha}\right)^{\frac{1}{\alpha-1}}}, \quad t > 0, \quad (2.14)$$

gelöst werden. Die Lösung fällt monoton auf 0. Für festes $\alpha > 1$ fällt die Lösung mit wachsendem t umso schwächer, je kleiner der Startwert ist (siehe Abbildung 2.5 für $\alpha = 2$). Legt man die Abbildungen 2.3 und 2.5 übereinander wird der Unterschied zum exponentiellen Abfall bei $\alpha = 1$ deutlich (siehe Abbildung 2.6 auf der nächsten Seite). Selbst bei sehr kleinen Startwerten gibt es im Falle $\alpha = 1$ noch deutlich negative Steigungen, während im Fall $\alpha = 2$ die Lösungskurven beinahe parallel zur t -Achse verlaufen.

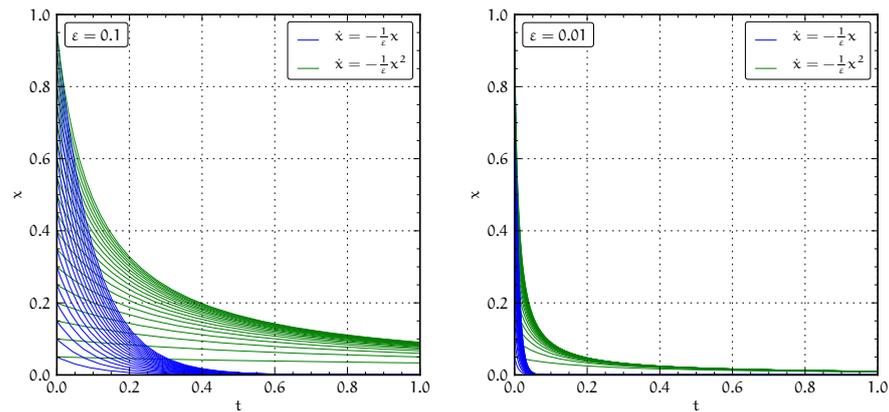


Abbildung 2.6: $\dot{x} = -\frac{1}{\varepsilon}x$ und $\dot{x} = -\frac{1}{\varepsilon}x^2$.

Für einen festen Startwert fallen die Lösungen um so stärker ab, je kleiner $\alpha - 1$ ist. Bei $\alpha = 1$ liegt eine Singularität vor. Die Lösung fällt dann exponentiell, so schnell wie für kein anderes $\alpha > 1$, für die die Lösung schwächer fällt. Der Fall $\alpha = 1$ (siehe (2.10)) kann somit als Grenzverhalten betrachtet werden. Warum das Verhalten beim expliziten Euler-Verfahren im Grenzfall einen solch anderen Charakter vorweist lässt sich im Folgenden leicht einsehen.

Nach dem expliziten Euler-Verfahren berechnet sich der $(k + 1)$ -te Zeitschritt für das Anfangswertproblem aus (2.13) mit $x_0 = 1$ durch

$$x_{k+1} = \left(1 - \frac{\tau}{\varepsilon} x_k^{\alpha-1}\right) x_k. \quad (2.15)$$

Aus dem monoton abfallenden Verhalten der exakten Lösung geht hervor, dass für eine brauchbare Näherungslösung der Faktor vor x_k im offenen Einheitsintervall $(-1, 1)$ liegen muss. Die obere Schranke 1 wird wegen $\tau, \frac{1}{\varepsilon} > 0$ und $x(0) = 1$ trivialerweise eingehalten. Um die untere Schranke einzuhalten, muss

$$\tau < \frac{2\varepsilon}{x_n^{\alpha-1}} \quad (2.16)$$

gefordert werden. Zu Beginn ist $x_k \approx 1$ und die Schrittweite τ muss umso kleiner sein, je kleiner ε ist. Ist $\alpha > 1$ und $x_k \approx 0$ ist die Schrittweitenbeschränkung wesentlich schwächer. Bei $\alpha = 1$ bleibt die Forderung nach einer kleinen Schrittweite unabhängig von dem Wert von x_k bestehen. Das macht die Sonderstellung für den Fall $\alpha = 1$ noch einmal deutlich und erklärt warum der numerische Aufwand mit dem expliziten Euler-Verfahren so viel größer ist als in dem Fall $\alpha > 1$. [Abbildung 2.7 auf der nächsten Seite](#) vergleicht den Aufwand (Anzahl der Zeitschritte) von einem expliziten und einem impliziten numerischen Verfahren in Abhängigkeit von ε , einmal im Fall $\alpha = 1$ und einmal im Fall $\alpha = 2$. Während bei (2.9) das implizite Verfahren deutlich weniger Rechenschritte benötigt, ist bei (2.13) mit $\alpha = 2$ das explizite Verfahren die bessere Wahl. Die Unterschiede fallen umso

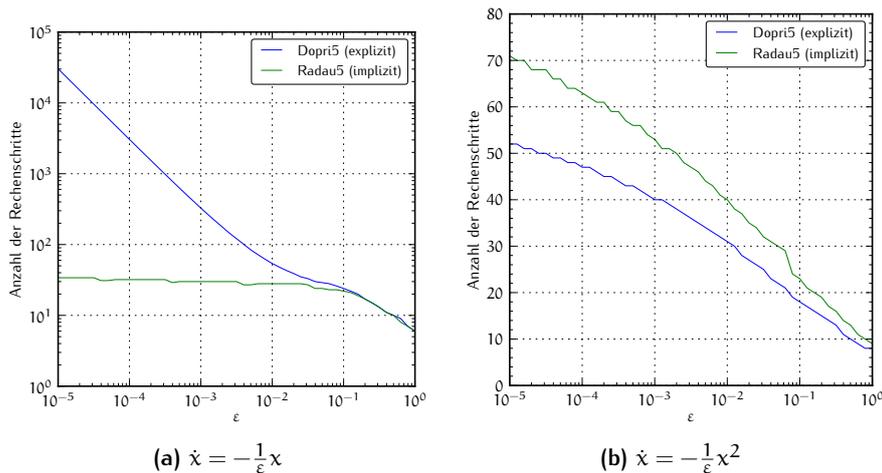


Abbildung 2.7: Explizites Verfahren vs. implizites Verfahren.

deutlicher aus, je kleiner ε ist. Diese Beobachtungen werden bestätigt von einer Äußerung von Curtiss und Hirschfelder:

Stiff equations are equations where certain implicit methods, in particular BDF, perform better, usually tremendously better, than explicit ones. (Hairer und Wanner, 2010, S. 1)

In diesem Sinne ist (2.13) für $\alpha = 2$ nicht steif, während (2.9) steif ist. Dieses Phänomen wird im Folgenden *numerische Steifheit* genannt. Wenn im restlichen Teil dieses Kapitels von Steifheit die Rede ist, ist damit immer numerische Steifheit gemeint.

Damit wurden die klassischen Eigenschaften steifer Differentialgleichungen aus numerischer Sicht erläutert. Es sei noch einmal darauf hingewiesen, dass hier der gewählten Zeitschrittweite eine besondere Bedeutung zukam. Für eine nicht-numerische, rein strukturelle Beschreibung steifer Differentialgleichungen ist dies sicherlich nicht wünschenswert.

2.2 FAMILIEN VON DIFFERENTIALGLEICHUNGEN

Vereinfacht man die Aussage aus dem letzten Zitat, ist eine Gleichung steif, falls explizite Verfahren nicht effizient sind. Ein explizites Verfahren heie effizient, falls es fur ein gegebenes Problem zu vorgegebener Genauigkeit und Rechenzeit eine Naherungslosung findet. Die Frage, ob ein explizites Verfahren fur eine gegebene Gleichung effizient ist, hangt demnach stark von der verwendeten Hardware ab aber noch starker von der geforderten Genauigkeit und Rechenzeit. Mit hinreichend hoher Rechenleistung wird ein explizites Verfahren fur jedes feste ε zu einer Losung kommen. Probleme, deren Losung vor

ein paar Jahren mit einem expliziten Verfahren undenkbar war, weil schlichtweg die Leistung fehlte, bereiten heutzutage möglicherweise keine Schwierigkeiten mehr.

Für das explizite Euler-Verfahren kann man idealisiert die Effizienz mit der Anzahl der möglichen Zeitschritte identifizieren, die in vorgegebener Rechenzeit und Genauigkeit durchlaufen werden. Je größer die Anzahl der möglichen Zeitschritte ist, desto besser ist die Laufzeiteffizienz. Anders ausgedrückt ist die Effizienz umso größer, je kleiner die Zeitschrittweite τ gewählt werden kann.

Für das Anfangswertproblem aus (2.9) wurde gezeigt, dass τ höchstens so groß wie ε sein darf, um eine brauchbare, nicht oszillierende Näherungslösung zu erhalten. Ist τ diejenige Schrittweite, wie laufezeit-effizient das Verfahren auch sein mag und wie klein die damit mögliche Zeitschrittweite τ auch sein darf, es kann immer ein ε gefunden werden, so dass die Bedingung $\tau \leq \varepsilon$ nicht erfüllt ist. Bei (2.13) hängt die Schrittweitenbeschränkung nach (2.16) zusätzlich davon ab, wie klein $\chi_n^{\alpha-1}$ ist. Für kleineres $\chi_n^{\alpha-1}$ ist die Schrittweitenbeschränkung bei festem ε schwächer. Das ist ein wesentlicher Unterschied.

Durch diese Betrachtungsweise wird statt eines einzelnen Anfangswertproblems eine ganze Klasse von Anfangswertproblemen studiert, was aus mathematischer Sicht die interessantere Aufgabe ist. Betrachte man also statt eines einzelnen Anfangswertproblems eine durch ε parametrisierte Familie von Anfangswertproblemen. Gibt es ein $\varepsilon > 0$, so dass die zu wählende Zeitschrittweite eine höhere Effizienz als die gegebene erzwingen würde, so kann die Familie von Anfangswertproblemen als steif angesehen werden. Dies soll dazu dienen eine Intuition für eine familienbasierte Definition steifer Differentialgleichungen zu bekommen.

Im Folgenden werden gewöhnliche Differentialgleichungen erster Ordnung der Form

$$\dot{x} = f(t, x(t, \varepsilon), \varepsilon) \quad (2.17)$$

mit einer rechten Seite

$$f : \mathbb{R} \times D \times (0; \varepsilon_0] \rightarrow \mathbb{R}^n \quad (2.18)$$

betrachtet, wobei $D \subset \mathbb{R}^n$ ein offenes Gebiet ist. Es wird davon ausgegangen, dass die Funktionen $f(t, x, \cdot)$ mit $x \in D$ auf dem Intervall $(0, \varepsilon_0]$ stetig sind. Durch den Parameter ε wird (2.17) zu einer Familie von Differentialgleichungen:

$$\dot{x}_\varepsilon = f_\varepsilon(x, t). \quad (2.19)$$

Die Betrachtung von Differentialgleichungsfamilien ist in diesem Kontext ungewöhnlich, hat aber einige Vorteile. Bei dem Studium von Differentialgleichungen werden häufig das Langzeitverhalten und das

schnelle Relaxationsverhalten von Lösungen analysiert. In der Numerik nutzt man ersteres insbesondere um Stabilitätsfragen zu beantworten. Beides spielt auch bei der Untersuchung steifer Differentialgleichungen eine Rolle und lässt sich allgemeiner auch durch den familienbasierten Ansatz beschreiben.

Zunächst wird auf das Langzeitverhalten eingegangen.

Definition 2.1 (Deuflhard und Bornemann, 2008, Definition 3.19.). Sei $(t_0, x_0) \in \Omega$, so dass die Lösung $\Phi^{t, t_0} x_0$ für alle $t \geq t_0$ existiert. Hierbei bezeichne Ω den Phasenraum und Φ^{t, t_0} den Phasenfluss. Die Integralkurve durch (t_0, x_0) heißt in

- *Vorwärtsrichtung stabil* (im Sinne von Ljapunov), falls zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass

$$\Phi^{t, t_0} x \in B_\varepsilon(\Phi^{t, t_0} x_0) \quad (2.20)$$

für alle $t \geq t_0$ und für alle gestörten Anfangswerte $x \in B_\delta(x_0)$ gilt.

- *Vorwärtsrichtung asymptotisch stabil*, falls es zusätzlich ein $\delta_0 > 0$ gibt, so dass

$$\lim_{t \rightarrow \infty} |\Phi^{t, t_0} x_0 - \Phi^{t, t_0} x| = 0 \quad (2.21)$$

für alle gestörten Anfangswerte $x \in B_{\delta_0}(x_0)$ gilt.

Da in der Numerik nur über endliche Zeitintervalle integriert wird, ist es nicht intuitiv, Lösungen für $t \rightarrow \infty$ zu betrachten. In dem man parametrisierte Familien von gewöhnlichen Differentialgleichungen betrachtet, kann dieses Langzeitverhalten anders interpretiert werden.

Satz 2.1. Sei $f : D \rightarrow \mathbb{R}^n$ eine Lipschitz-stetige Funktion auf der offenen Teilmenge $D \subset \mathbb{R}^n$ und $(t_0, x_0) \in \Omega$, so dass die Lösung $\Phi^{t, t_0} x_0$ von

$$\frac{dx}{dt} = f(t, x) \quad (2.22)$$

für alle $t \geq t_0$ existiert. Weiter sei die Familie $y_\varepsilon(\tau) = x(t)$ mit $\tau = \phi(\varepsilon) t$ und $\phi = o(1)$ gegeben. Aus Gleichung (2.22) erhält man durch Koordinatentransformation eine Familie von Differentialgleichungen:

$$\frac{dy_\varepsilon}{d\tau} = \phi(\varepsilon) f_\varepsilon(\tau, y_\varepsilon). \quad (2.23)$$

Die Integralkurve durch (x_0, t_0) ist genau dann in Vorwärtsrichtung asymptotisch stabil, wenn für $\varepsilon \rightarrow 0$ der folgende Grenzwert verschwindet:

$$\lim_{\varepsilon \rightarrow 0} |y_\varepsilon(\tau/\phi(\varepsilon); x_0) - y_\varepsilon(\tau/\phi(\varepsilon); x)| = 0 \quad (2.24)$$

Beweis. In Definition 2.1 substituiere man t durch $\tau/\phi(\varepsilon)$. \square

Auf den ersten Blick mag der Nutzen dieses nahezu trivialen Zusammenhangs nicht erkennbar sein. Das Langzeitverhalten lässt sich auf diese Weise auf das Grenzverhalten für $\varepsilon \rightarrow 0$ für einen fest gehaltenen Zeitpunkt, im Sinne des Grenzverhaltens der Lösungen einer Familie von Differentialgleichungen, übertragen. Um dies zu verdeutlichen, betrachte man die Differentialgleichung

$$\frac{dx}{dt} = -x. \quad (2.25)$$

Wird diese Gleichung auf dem Intervall $[0, 1]$ integriert, werden die diskutierten Probleme bei expliziten numerischen Verfahren nicht beobachtet, wohl aber wenn das Integrationsintervall groß wird. Mit zunehmender Intervallgröße, muss auch τ entsprechend gewählt werden, damit das Verfahren „zeitnah“ terminiert. Für $\tau > 2$ explodiert im expliziten Euler-Verfahren x_k für wachsendes k , siehe (2.11). Die Länge des Integrationsintervalls ist hier von entscheidender Bedeutung, was durch ein Zitat von Aiken (1985) belegt wird.

Here and later we say the Lipschitz constant is large as an abbreviation for the statement that $(b - a)L$ is large; the reader should remember that the length of the interval is also important. [Hier stehen a und b für die Intervallgrenzen und L für die Lipschitzkonstante.] (ebd.)

Durch den Übergang zu einer Familie von Problemen kann man sich auf Zeitintervalle der Größenordnung $\mathcal{O}(1)$ beschränken. Durch die Transformation $\tau = \varepsilon t, \varepsilon > 0$ wird das Intervall $[0, 1/\varepsilon)$ auf das Intervall $[0, 1)$ abgebildet; das Intervall $[0, \infty)$ wird von den Intervallen $[0, 1/\varepsilon)$ überdeckt. Aus (2.25) erhält man damit

$$\frac{dy}{d\tau} = -\frac{1}{\varepsilon}y. \quad (2.26)$$

Dadurch kann das steife Verhalten von (2.25), das für hinreichend große Zeiten auftritt, durch den Faktor $1/\varepsilon$ direkt in Gleichung (2.26) „sichtbar“ gemacht werden, insbesondere auf einem ε -unabhängigen endlichen Intervall. Diese Flexibilität ist gerade bei Mehrskalenproblemen nicht zu unterschätzen, in denen mehrere Zeitskalen mittels eines *distinguished limit* an einen kleinen Parameter ε gekoppelt sind. An dieser Stelle sei auf Teil i verwiesen, wo genau das gemacht wurde. Durch Umskalierung in ε auf die Integrationslänge erhält man die zu studierende Familie von Gleichungen.

Als nächstes wird darauf eingegangen was schnelles Relaxationsverhalten für Differentialgleichungsfamilien bedeutet.

Gleichung (2.9) kann als Prototyp für steife Differentialgleichungen betrachtet werden und wird nach dem schwedischen Mathematiker Germund Dahlquist (1925 - 2005) als dahlquistsches Testproblem benannt. Wesentlich ist das Auftreten einer exponentiell abklingenden und einer langsamen Dynamik in der Lösung. Beides wird von dem

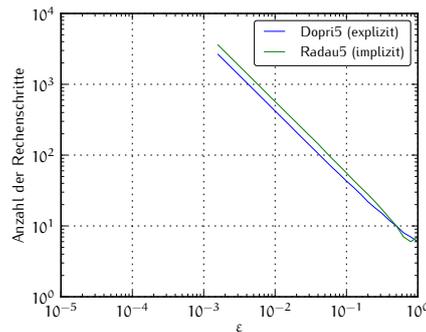


Abbildung 2.8: $\dot{x} = \frac{1}{\varepsilon}(x - t^2)$, $x(0) = 1$.

dahlquistschen Testproblem erfüllt, auch wenn die langsame Dynamik trivial ist.

Redet man von schneller und langsamer Dynamik, so ist das ein Indiz für Mehrskalenprobleme. Die langsame Dynamik beschreibt das Verhalten der Lösung über einen langen Zeitraum, während die schnelle Dynamik Änderungen beschreibt, die in kurzer Zeit ablaufen. Diese Betrachtungsweise kann genutzt werden, um das wesentliche Verhalten von Lösungen steifer Differentialgleichungen hervorzuheben, nämlich die Überlagerung von exponentiell abklingender und langsamer Dynamik.

Man betrachte eine Familie von Funktionen x der Form

$$x(t) = g(t) + \exp(-t/\varepsilon) h(t), \quad h \neq 0. \quad (2.27)$$

Je kleiner ε ist, desto früher fällt der zweite Term unter einen bestimmten Schwellwert. Der zweite Term beschreibt die schnelle, der erste Term die langsame Dynamik. Lösungsfamilien dieser Gestalt sind klassischerweise steif – wie bereits gesehen bereitet dieses schnelle exponentielle Abklingen bei expliziten numerischen Verfahren Probleme.

Als Beispiel betrachte man ein modifiziertes Problem aus Curtiss und Hirschfelder (1952).

Beispiel 2.4. Für $\varepsilon = 1/5$ nennen Curtiss und Hirschfelder (ebd.)

$$\dot{x} = \frac{1}{\varepsilon}(x - t^2) \quad (2.28)$$

eine typische steife Gleichung. In Abbildung 2.8 sieht man die Anzahl der benötigten Rechenschritte in Abhängigkeit von ε für ein explizites und ein implizites Verfahren. Beide brauchen quasi gleichermaßen viele Rechenschritte. Nach dem Zitat am Ende von Abschnitt 2.1 ist das kein Zeichen für steifes Verhalten. Modifiziert man (2.28) und betrachtet statt dessen die Familie

$$\dot{x} = -\frac{1}{\varepsilon}(x - t^2) \quad (2.29)$$

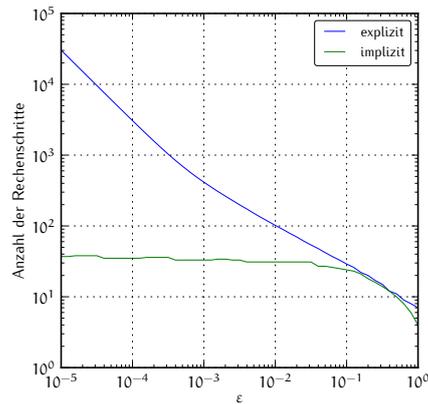


Abbildung 2.9: $\dot{x} = -\frac{1}{\varepsilon}(x - t^2)$, $x(0) = 1$.

mit den Lösungen

$$x(t) = C \exp(-t/\varepsilon) + t^2 - 2\varepsilon t + 2\varepsilon^2, \quad C \in \mathbb{R}, \quad (2.30)$$

erkennt man einerseits, dass die Lösungen der Form aus (2.27) entsprechen, andererseits erkennt man an Abbildung 2.9 auch aus numerischer Sicht das steife Verhalten.

Bei dem Versuch solche Probleme mit einem typischen asymptotischen Einskalenansatz zu analysieren stößt man schnell an Grenzen. Betrachtet man nur die Zeitkoordinate t erhält man lediglich das Verhalten der langsamen Dynamik, $g(t)$ in (2.27) bzw. t^2 (führende Ordnung) in (2.30). Das entspricht dem klassischen Langzeitverhalten. Betrachtet man stattdessen ein transformiertes Problem mit der Zeitkoordinate $\tau = t/\varepsilon$ (hier wird τ als feste Koordinate interpretiert), verliert man die nicht-triviale langsame Dynamik und erkennt nur das schnelle Verhalten. Um dies zu verdeutlichen folgt ein Beispiel.

Beispiel 2.5. Betrachtet wird die Familie von zweidimensionalen Anfangswertproblemen

$$\dot{x} = \begin{pmatrix} 1 & 0 \\ 0 & -\frac{1}{\varepsilon} \end{pmatrix} x, \quad x(0) = (1, 1)^T \quad (2.31)$$

mit den Lösungen

$$x(t) = \begin{pmatrix} \exp(t) \\ \exp(-t/\varepsilon) \end{pmatrix}. \quad (2.32)$$

Die zweite Komponente ist gerade das steife dahlquistsche Testproblem. Für $t \rightarrow \infty$ existiert keine Lösung, denn die erste Komponente explodiert. Betrachtet man die Lösung nun für festes t , existiert der Grenzwert der Lösung für $\varepsilon \rightarrow 0$. An der exakten Lösung ist zu erkennen, dass die exponentiell schnelle Dynamik sich für $\varepsilon \rightarrow 0$ einer nichttrivialen langsamen Dynamik annähert. Die zweite Lösungskomponente x_2 verschwindet exponentiell. Übrig bleibt das „Langzeitverhalten“, das nur durch die erste Komponente gegeben ist.

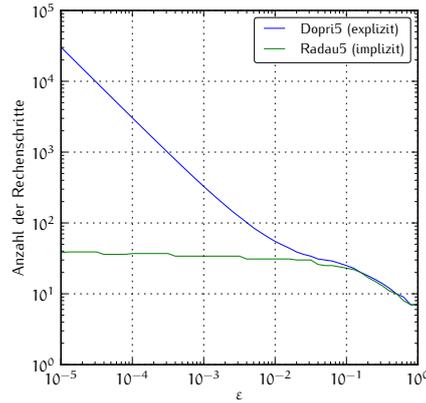


Abbildung 2.10: Numerisches Verhalten von (2.31)

Die Lösung für $\varepsilon \rightarrow 0$ lebt gerade in dem Eigenraum des asymptotisch kleineren Eigenwerts. Betrachtet man lediglich das Verhalten für $t \rightarrow \infty$, lässt sich diese Struktur so nicht identifizieren.

Aus numerischer Sicht ist das System steif, siehe Abbildung 2.10. Asymptotisch stabil ist das Problem nicht, da die Spektralabszisse positiv ist (siehe Deuffhard und Bornemann, 2008, Satz 3.23.).

Ein klassischer asymptotischer Einskalenansatz

$$x(t; \varepsilon) = x^{(0)}(t) + \varepsilon x^{(1)}(t) + o(\varepsilon) \quad (2.33)$$

liefert

$$x(t) = \begin{pmatrix} \exp(t) \\ 0 \end{pmatrix} + o(1). \quad (2.34)$$

Die Anfangsbedingung ist hier nicht erfüllt. Der gewählte Ansatz ist also ungeeignet.

Führt man stattdessen eine „schnelle“ Zeitkoordinate $\tau = t/\varepsilon$ ein und betrachtet das transformierte Problem

$$\frac{dx}{d\tau} = \begin{pmatrix} \varepsilon & 0 \\ 0 & -1 \end{pmatrix} x, \quad x(0) = (1, 1)^T. \quad (2.35)$$

so führt ein analoger Einskalenansatz wie eben zu

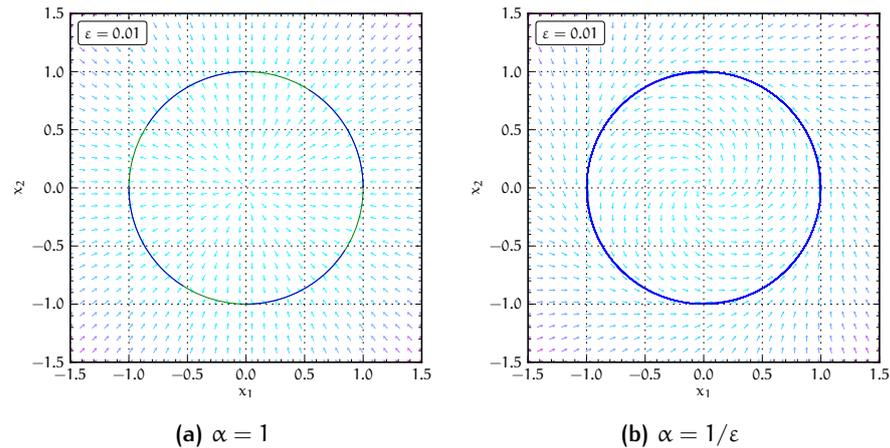
$$x(\tau) = \begin{pmatrix} 1 \\ \exp(-\tau) \end{pmatrix} + o(1). \quad (2.36)$$

Diesmal wird zwar das asymptotische Abklingen der zweiten Komponente richtig abgebildet, nicht jedoch das Verhalten in der ersten Komponente. Durch klassische Einskalenansätze lassen sich also nicht beide Prozesse beschreiben.

Mit einem Multiskalenansatz

$$x(t; \varepsilon) = x^{(0)}\left(\frac{t}{\varepsilon}, t\right) + \varepsilon x^{(1)}\left(\frac{t}{\varepsilon}, t\right) + o(\varepsilon) \quad (2.37a)$$

$$= x^{(0)}(\tau, t) + \varepsilon x^{(1)}(\tau, t) + o(\varepsilon) \quad (2.37b)$$

Abbildung 2.11: Nichtlinearer Strudel ($x(0) = (2, 2)$).

erhält man hingegen

$$x(\tau, t) = \begin{pmatrix} \exp(t) \\ 0 \end{pmatrix} + \exp(-\tau) \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (2.38)$$

Das entspricht nicht nur der exakten Lösung, sondern ist auch von der Gestalt wie die Funktion in (2.27). Das Langzeitverhalten der Lösung wird alleine durch die erste Komponente beschrieben, die zweite Komponente dagegen beschreibt das exponentielle Abklingen am Anfang. Durch den Mehrskalenansatz kann die Anfangsbedingung vollständig erfüllt werden.

Das nächste Beispiel zeigt einerseits, dass die Form der Lösung aus (2.27) keine notwendige Bedingung für steifes Verhalten ist, und andererseits, dass steifes Verhalten nicht invariant unter nichtlinearen Koordinatentransformationen ist. In den folgenden Kapiteln wird noch häufiger darauf zurück gegriffen.

Beispiel 2.6 (Nichtlinearer Strudel). Für $\alpha \in \{1, 1/\varepsilon\}$ betrachte man die Familie nichtlinearer Systeme gewöhnlicher Differentialgleichungen, die durch

$$\dot{x} = \frac{1}{\varepsilon} (1 - \|x\|^2) x + \alpha \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} x \quad (2.39)$$

gegeben ist. Der erste Term auf der rechten Seite kontrolliert den Abstand zum Einheitskreis, der zweite Term beschreibt eine Drehung um den Ursprung. Das Phasenportrait ist in Abbildung 2.11 zu sehen. Für $\alpha = 1$ ist die Bewegung in Umfangsrichtung langsamer als das radiale Relaxieren auf den Einheitskreis. Für $\alpha = 1/\varepsilon$ ist die Bewegung in beiden Richtungen gleich schnell. Abbildung 2.12 auf der nächsten Seite zeigt das numerische Verhalten bei einem Anfangswert von $x_0 = (2, 2)$. Für $\alpha = 1$ ist das typische steife Verhalten zu beobachten, für $\alpha = 1/\varepsilon$ hingegen nicht.

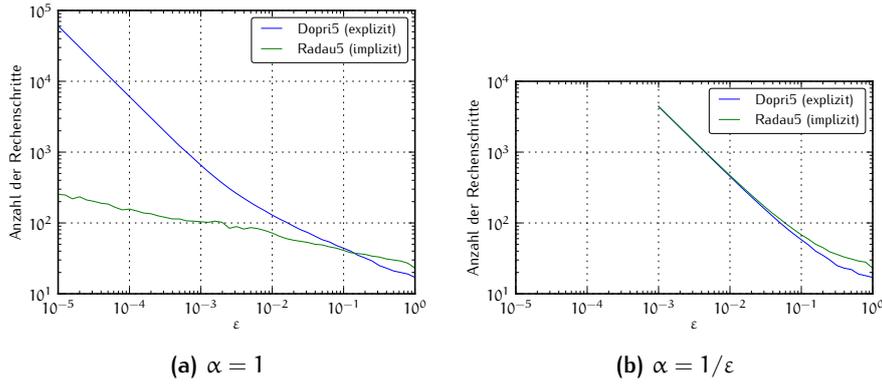
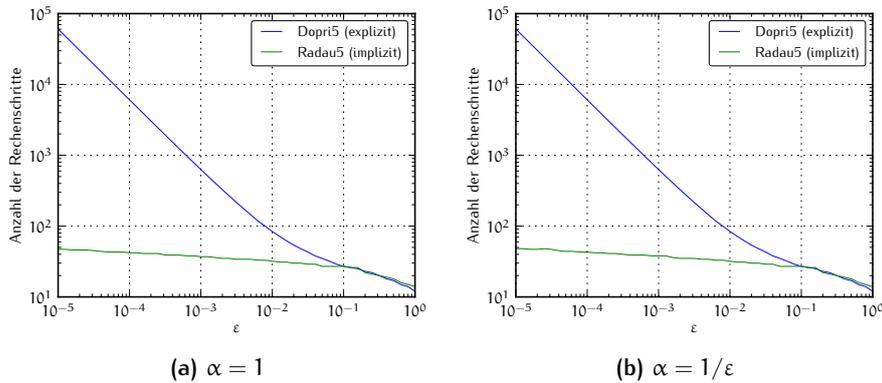


Abbildung 2.12: Nichtlinearer Strudel (kartesische Koordinaten).


 Abbildung 2.13: Nichtlinearer Strudel ((u, ϕ) -Koordinaten).

Um die Symmetrieeigenschaften des Systems voll ausnutzen zu können, betrachte man das System für $x \neq 0$ in einem den Polarkoordinaten ähnlichen Koordinatensystem (vgl. Schmitt, 2012, S. 39), in dem r der Radius und ϕ der Winkel in Polarkoordinaten sowie $u = r^2 = \langle x, x \rangle$ gegeben sei. Für $x = (x_1, x_2)^T$, $x_1 \neq 0$ wird nun das transformierte System

$$\dot{u} = 2\langle x, \dot{x} \rangle = \frac{2}{\epsilon} (1 - u) u \quad (2.40a)$$

$$\dot{\phi} = \arctan\left(\frac{x_2}{x_1}\right) = \alpha \frac{x_1 \dot{x}_2 - \dot{x}_1 x_2}{x_1^2 + x_2^2} = \alpha \quad (2.40b)$$

studiert. Abbildung 2.13 zeigt das steife Verhalten des Strudels in dem (ϕ, u) -Koordinatensystem, unabhängig von der Wahl von α . Das ist ein bedeutsamer und hervorzuhebender Unterschied zu Abbildung 2.12, in der das Verhalten in kartesischen Koordinaten dargestellt ist.

Für ein tieferes analytisches Verständnis wird eine asymptotische Zweiskalenlösung in führender Ordnung für die Zeitskalen t und $\tau = t/\epsilon$ gesucht. Die Umfangsgeschwindigkeit ist nach der letzten Gleichung $\dot{\phi} = \alpha$. Für $\alpha = 1$ ist die Umfangsbewegung Teil der lang-

samen Dynamik auf der langsamen Zeitskala t , für $\alpha = 1/\varepsilon$ ist die Drehung Teil der schnellen Dynamik auf der schnellen Zeitskala τ . Um die radiale Bewegung zu verstehen, ist mehr Arbeit nötig. Ausgangspunkt ist ein Zweizeitskalenansatz:

$$x(t; \varepsilon) = x^{(0)}\left(\frac{t}{\varepsilon}, t\right) + \varepsilon x^{(1)}\left(\frac{t}{\varepsilon}, t\right) + o(\varepsilon) \quad (2.41a)$$

$$= x^{(0)}(\tau, t) + \varepsilon x^{(1)}(\tau, t) + o(\varepsilon). \quad (2.41b)$$

Nach diesem Ansatz ergeben sich

$$\begin{aligned} \frac{dx}{dt} = \frac{1}{\varepsilon} \frac{\partial x^{(0)}}{\partial \tau} + \left(\frac{\partial x^{(0)}}{\partial t} + \frac{\partial x^{(1)}}{\partial \tau} \right) \\ + \varepsilon \left(\frac{\partial x^{(1)}}{\partial t} + \frac{\partial x^{(2)}}{\partial \tau} \right) + \mathcal{O}(\varepsilon^2), \end{aligned} \quad (2.42)$$

und

$$u = \|x\|^2 = \langle x, x \rangle = \|x^{(0)}\|^2 + 2\varepsilon \langle x^{(0)}, x^{(1)} \rangle + \mathcal{O}(\varepsilon^2), \quad (2.43)$$

für die zeitliche Änderung und die Norm. Wertet man die Gleichungen in den einzelnen ε -Potenzen aus erhält man:

$$\mathcal{O}\left(\frac{1}{\varepsilon}\right): \frac{\partial u^{(0)}}{\partial \tau} = 2(1 - u^{(0)})u^{(0)} \quad (2.44)$$

$$\implies u^{(0)}(\tau) = \frac{\exp(2\tau)}{C_0(t) + \exp(2\tau)},$$

$$\mathcal{O}(1): \frac{\partial u^{(1)}}{\partial \tau} = 2(1 - 2u^{(0)})u^{(1)} - \frac{\partial u^{(0)}}{\partial \tau} \quad (2.45)$$

Daraus folgt

$$u^{(1)}(\tau) = \tau \frac{C'_0(t) \exp(2\tau)}{(C_0(t) + \exp(2\tau))^2} + \frac{C_1(t) \exp(2\tau)}{(C_0(t) + \exp(2\tau))^2}. \quad (2.46)$$

Zusammengesetzt ergibt das

$$u(\tau, t) = u^{(0)}(\tau, t) + \varepsilon u^{(1)}(\tau, t) + \mathcal{O}(\varepsilon^2) \quad (2.47a)$$

$$= \frac{\exp(2\tau)}{C_0(t) + \exp(2\tau)} \quad (2.47b)$$

$$+ \varepsilon \tau \left(\frac{C'_0(t) \exp(2\tau)}{(C_0(t) + \exp(2\tau))^2} \right) \quad (2.47c)$$

$$+ \varepsilon \left(\frac{C_1(t) \exp(2\tau)}{(C_0(t) + \exp(2\tau))^2} \right) \quad (2.47d)$$

$$+ \mathcal{O}(\varepsilon^2). \quad (2.47e)$$

Der zweite Term ist sekulär, da er mit $\varepsilon\tau = t$ nicht asymptotisch kleiner als der erste ist. Es folgt $C'_0(t) = 0$ (Sub-linear-growth-condition)

und daraus $C_0(t) = \tilde{C}_0$, wobei \tilde{C}_0 eine Konstante ist, die durch die Anfangsbedingung $u(0) = \|x(0)\|^2$ gegeben ist. In führender Ordnung gilt

$$u^{(0)}(\tau, t) = \frac{1}{1 + \tilde{C}_0 \exp(-2\tau)} = 1 - \frac{\exp(-2\tau)}{1 + \tilde{C}_0 \exp(-2\tau)}, \quad (2.48a)$$

$$\tilde{C}_0 = \frac{1 - \|x(0)\|^2}{\|x(0)\|^2}. \quad (2.48b)$$

In dem Koordinatensystem (ϕ, u) lautet die Lösung führender Ordnung:

$$\begin{pmatrix} \phi^{(0)} \\ u^{(0)} \end{pmatrix} = \alpha t \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \exp(-2t/\varepsilon) \begin{pmatrix} 0 \\ \frac{1}{1 + \tilde{C}_0 \exp(-2\tau)} \end{pmatrix}. \quad (2.49)$$

Für $\alpha = 1$ ist der Ausdruck von der Form wie in (2.27), nicht so für $\alpha = 1/\varepsilon$. Nach Abbildung 2.13 zeigt sich aber unabhängig von α steifes Verhalten. Damit ist die Struktur aus (2.27) kein notwendiges Kriterium für Steifheit. Später wird genauer darauf eingegangen von welcher Art die $(\tau = t/\varepsilon)$ -Abhängigkeit von g in (2.27) sein darf.

Das Beispiel des nichtlinearen Strudels zeigt auch, dass steifes Verhalten nicht invariant unter Koordinatenwechsel sein muss. Man vergleiche dazu nochmals die Abbildungen 2.12 und 2.13. Eine Beobachtung, die in dieser Form in der Literatur so nicht zu finden ist. Bewusst ist man sich darüber, dass die Implementierung eines mathematischen Problems von Bedeutung ist, wie das folgende Zitat belegt.

Furthermore, whether a given code exhibits stiffness depends on the particular formula, how it is implemented, and how it relates to the particular mathematical problem.
(Aiken, 1985)

Für Familien linearer Gleichungen wird später, in Abschnitt 3.4.1, gezeigt, dass steife Familien von Lösungen invariant unter orthogonalen Transformationen sind. Für nichtlineare Transformationen trifft dies im Allgemeinen nicht zu. Neben dem nichtlinearen Strudel sieht man das an dem folgenden einfachen Beispiel.

Beispiel 2.7. Gegeben sei die Familie von Anfangswertproblemen

$$\dot{x} = -\frac{1}{\varepsilon}, \quad x(0) = 0, \quad (2.50)$$

mit der Lösung $x(t) = -t/\varepsilon$. Hier liegt offensichtlich kein steifes Verhalten vor, da das wesentliche exponentielle Abklingen fehlt. Mittels der Transformation $x = \ln y$ gelangt man zu dem Problem

$$\dot{y} = -\frac{1}{\varepsilon}y, \quad y(0) = 1. \quad (2.51)$$

Das ist das dahlquistsche Testproblem, zeigt also steifes Verhalten. Gerade für numerische Berechnungen bedeutet das, dass die Wahl des Koordinatensystems von entscheidender Bedeutung sein kann.

Es ist sicherlich nicht immer so einfach möglich die Lösung, sei es in exakter Form oder als asymptotische Näherung, zu finden. In so einem Fall lässt sich natürlich die Gestalt aus (2.27) auch nicht erkennen und als Indiz für steifes Verhalten verwenden. Außerdem muss selbst eine steife Lösung nicht immer von dieser Gestalt sein, (siehe Beispiel 3.25 auf Seite 92).

Ein klassisches Beispiel, in dem die Gestalt aus (2.27) nicht so einfach wie in den vorangegangenen Beispielen gefunden werden kann, ist der Van-der-Pol-Oszillator:

Die van der Pol Gleichung ist eines der am häufigsten verwendeten Testbeispiele für steife Differentialgleichungen, da sie für $\mu \gg 1$ sehr steif wird und sich in der Lösung Intervalle, in denen die Lösung sich nur langsam ändert, mit Intervallen abwechseln, in denen eine sehr starke Änderung der Lösung vonstattengeht. (Strehmel, Weiner und Podhaisky, 2012, S. 13)

Beispiel 2.8. Der Van-der-Pol-Oszillator ist ein schwingendes System mit nichtlinearer Dämpfung:

$$\varepsilon \ddot{x} = (1 - x^2)\dot{x} + \varepsilon x = 0. \quad (2.52)$$

Der Parameter μ aus dem Zitat entspricht dem Faktor $1/\varepsilon$. Im Folgenden wird der Fall betrachtet, in dem Dämpfung und Rückstellkraft von gleicher asymptotischer Größenordnung sind:

$$\varepsilon \ddot{x} = (1 - x^2)\dot{x} + x = 0. \quad (2.53)$$

In Liénhard Koordinaten lässt sich (2.53) als zweidimensionales System schreiben ($x = (x_1, x_2)$):

$$\dot{x}_1 = -x_2, \quad (2.54a)$$

$$\dot{x}_2 = \frac{1}{\varepsilon} \left(x_1 - \frac{1}{3} x_2^3 + x_2 \right). \quad (2.54b)$$

Abbildung 2.14 auf der nächsten Seite zeigt, dass es sich um ein steifes Problem im klassischen Sinne handelt. In Abbildung 2.15 auf der nächsten Seite sieht man, dass für kleinere Werte von ε die Lösungskurven im Ort-Zeit-Diagramm „zackiger“ werden. In den leicht gerundeten Bereiche des grünen Graphen ist die Lösung steif. Das Phasenportrait in Abbildung 2.16 auf Seite 54 ermöglicht weitere Einblicke. Die rote Linie ist die Nullkline der zweiten Komponente, entlang dieser gilt $\dot{x}_2 = 0$. Die x_2 -Komponente ist durch den Faktor $1/\varepsilon$ in Gleichung (2.54b) für die großen Geschwindigkeiten verantwortlich. Die rote Nullkline ist eine Art Ruhelage, außerhalb der roten

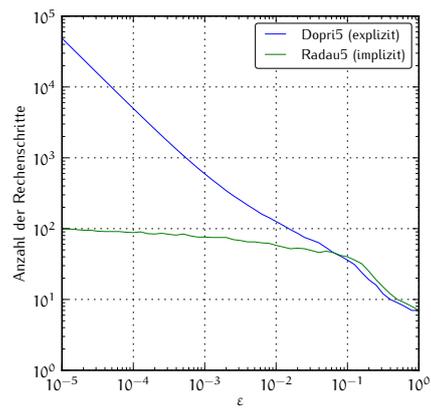


Abbildung 2.14: Van-der-Pol-Oszillators (numerisches Verhalten)

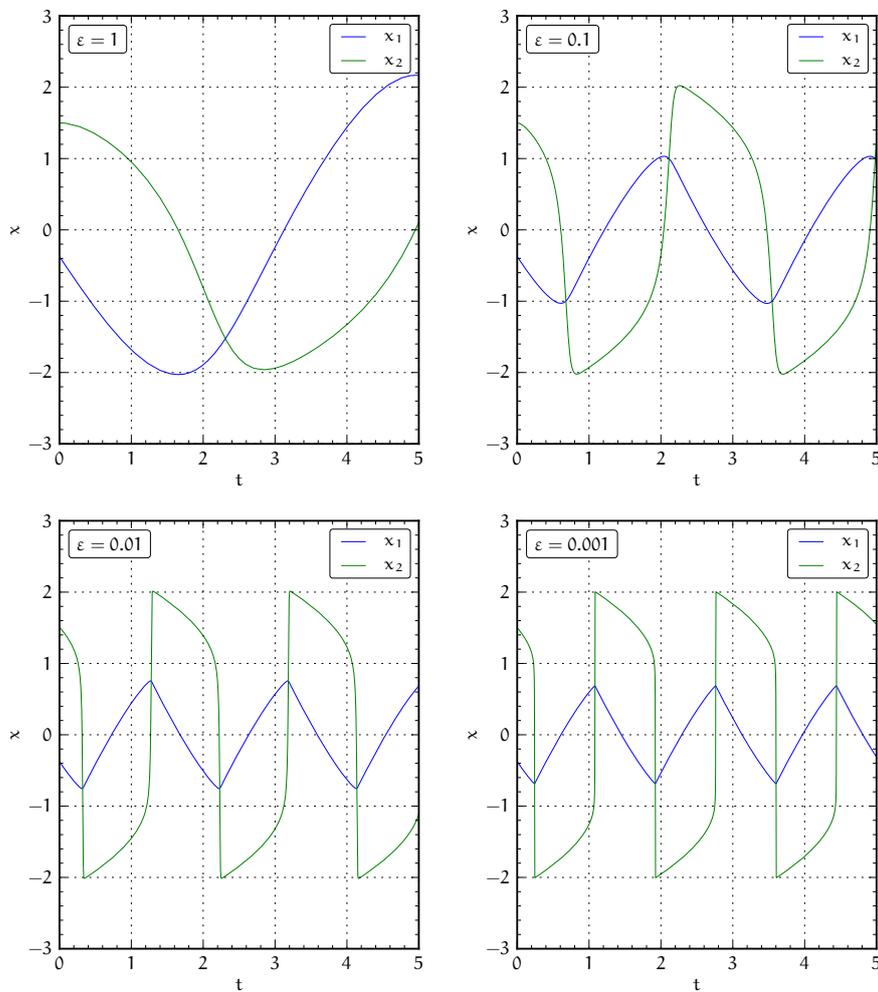


Abbildung 2.15: Ort-Zeit-Diagramm des Van-der-Pol-Oszillators.

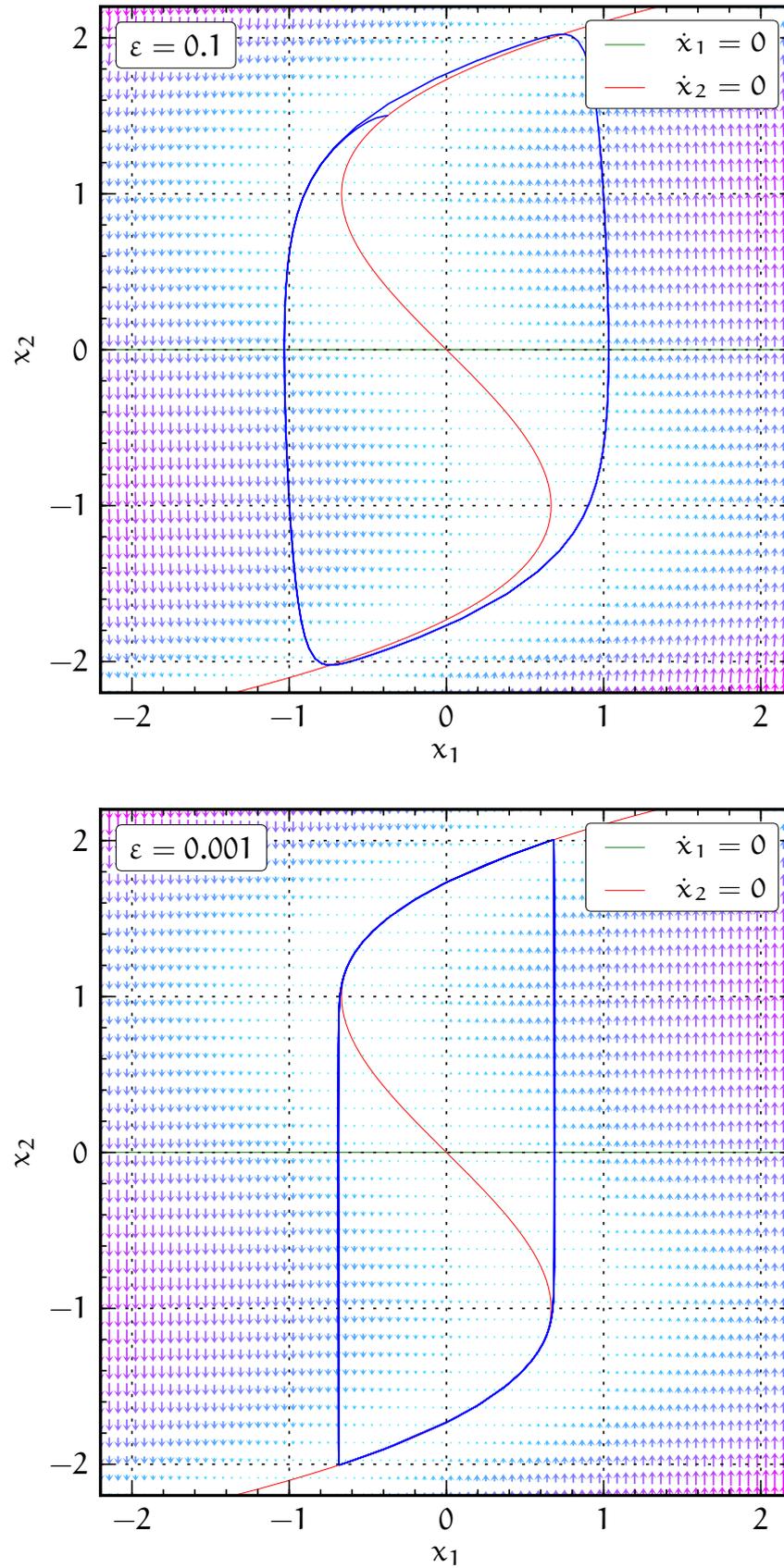


Abbildung 2.16: Phasenportrait des Van-der-Pol-Oszillators.

Linie treten wesentlich größere Geschwindigkeiten auf. Je kleiner ε ist, desto dichter schmiegt sich das Phasenportrait in gewissen Bereichen (blaue Linie) an die rote Linie an. Für $\varepsilon = 0.001$ ist deutlich zu erkennen, dass das Vektorfeld quasi parallel zur x_2 -Achse verläuft. In dem Bereich, in dem sich blaue und rote Linie nahezu schneiden, ist außerdem ein abrupter Geschwindigkeitswechsel, in Form eines Richtungswechsels zu sehen. All dies sind typische Eigenschaften steifer Probleme, denn dort knallt die Lösung exponentiell auf eine Art Ruhelage und dieser Geschwindigkeitswechsel lässt sich mit expliziten Verfahren nur schwer auflösen.

Eine asymptotische Multiskalenanalyse, ähnlich dem Beispiel des nichtlinearen Strudels, ist bei dem Van-der-Pol-Oszillator nicht so einfach möglich. Das liegt im Wesentlichen an der Struktur der Nichtlinearität. Kevorkian und Cole (1996) führen eine aufwendige Analyse des Van-der-Pol-Oszillators mit der Methode der Matched Asymptotic Expansions durch. Ein entsprechender Multiskalenansatz wäre noch aufwendiger. Werden die Probleme noch schwieriger ist es nicht zu erwarten, dass sich die Frage nach Steifheit immer durch einen Multiskalenansatz beantworten lässt.

2.3 FORMALISIERUNG STEIFER DIFFERENTIALGLEICHUNGEN

In den beiden vorangegangenen Abschnitten wurden die wesentlichen Eigenschaften steifer Differentialgleichungen diskutiert. Allen voran gehört das numerische Verhalten expliziter Verfahren, aber auch das Mehrskalenverhalten mit langsamer und exponentiell schnell abklingender Dynamik der Lösung dazu. Der Begriff der steifen Differentialgleichung wurde dabei stets im klassischen Sinne verwendet. Eine exakte Definition anhand der mathematischen Struktur und unabhängig von der numerischen Auflösung gab es nicht. Die Suche in der Literatur nach einer solchen, die ganze Komplexität erfassenden Definition dieses Begriffs bleibt ergebnislos. Häufig werden nur Beispiele gegeben, um die charakteristischen Eigenschaften herauszustellen oder es werden „Definitionsversuche“ (Simeon, 2003, S. 89) unternommen. Dieser Zustand wird allgemein als unbefriedigend empfunden. Versuche, eine solche Definition zu finden, werden seit langem als schwierig oder teilweise sogar als unmöglich angesehen. So schreibt zum Beispiel Aiken (1985)

Many authors have sought a definition of stiffness involving only the mathematical problem. Unfortunately the situation is far more complex than that. (ebd., S. 1)

oder Strehmel, Weiner und Podhaisky (2012)

Es gibt keine zufriedenstellende Definition für den Begriff „Steifheit einer Differentialgleichung“, da das Problem der Steifheit sehr vielschichtig sein kann. (Strehmel, Weiner und Podhaisky, 2012)

oder Quarteroni, Sacco und Saleri (2002)

[...] verschiedene Autoren [...] stimmen aber andererseits der Tatsache zu, dass es nicht möglich ist, exakt zu beschreiben, was mit einem *steifen* Problem gemeint ist. (ebd., S. 204)

Es geht sogar so weit den Sinn eines Terminus für solche Probleme in Frage zu stellen. Auf die Frage

QUESTION 1: *Should the term stiff be reserved for the initial-value ODE problem with sharp initial transients or could it be used in a wider sense (differential-algebraic, discontinuous, large sets, highly oscillatory, TPBVPs, PDE/BVPs with sharp transients, etc.)? Should another general term be defined? Should a number of terms, one for each specific type of difficulty, be defined? How important is terminology?* (Aiken, 1985, S. 360),

die auf der *International Conference on Stiff Computation* vom 12.-14. April 1982, in Park City, Utah, USA, gestellt wurde, antwortet Cellier

Cellier: ... What do you need the terminology for? I am rather pragmatic. I feel that the term should help people in solving their problems. Therefore, if I use the term freely to denote almost any integration problem which is numerically difficult to solve, the only thing a potential user can know is the fact that he has a difficult problem to solve and may be forced to consult a specialist. This may be useful information, but it may then be more straightforward to call this term “difficult problem” rather than “stiff problem”. ... (ebd., S. 361).

Es gibt aber auch Befürworter einer solchen Terminologie:

Shampine: I believe it is important to develop a reasonable terminology. (ebd., S. 363)

Die Idee der Steifheit ist nach Higham und Trefethen (1993) aber einfach

It is generally agreed that the essence of stiffness is a simple idea,

Stability is more of a constraint than accuracy,

with its familiar consequence as expressed in the words of Hairer and Wanner [5, p. 2],

Stiff equations are problems for which explicit methods don't work.

As soon as one tries to turn these ideas into a mathematical criterion for stiffness, however, disagreements set in. What makes a stiff problem stiff? No single answer seems right for all problems. In the face of this confusion some authors propose multiple criteria for stiffness, and others, none at all. (ebd., S. 1).

In den folgenden Abschnitten wird detailliert darauf eingegangen welche Bedeutung die Eigenwerte linearer Probleme bei der Beschreibung steifen Verhaltens haben. Häufig wird nämlich versucht steifes Verhalten linearer Probleme durch deren Eigenwerte zu erklären, aber auch das ist nicht ohne Einschränkung möglich:

It is argued that even for a linear system of ODEs with constant coefficients, stiffness cannot properly be characterized in terms of the eigenvalues of the Jacobian, because stiffness is a transient phenomenon whereas the significance of eigenvalues is asymptotic. (ebd., S. 1)

All diese Zitate belegen einerseits die Schwierigkeiten steife Probleme mathematisch zu fassen, zeigen aber andererseits auch das enorme Interesse daran.

In dem folgenden Kapitel wird ein neuer Blickwinkel auf steife Differentialgleichungen geschaffen. Grundlegend dafür ist die Betrachtung von Familien von Differentialgleichungen, die in Abschnitt 2.2 motiviert und eingeführt wurde.

3

STEIFE LÖSUNGSFAMILIEN

Inhalt

3.1	Langsame Mannigfaltigkeiten	60
3.1.1	Langsame Mannigfaltigkeit linearer Probleme	60
3.1.2	Langsame Mannigfaltigkeiten nichtlinearer Probleme	64
3.2	Exponentiell anziehende Mengen	67
3.3	Steife Familien von Lösungen	71
3.4	Steife Lösungsfamilien linearer Probleme	73
3.4.1	Invarianz unitärer Transformationen	75
3.4.2	Steife Lösungen von Schurformen	77
3.4.3	Normale und diagonale Probleme	85
3.4.4	Steife Lösungsfamilien linearer Probleme und explizite numerische Verfahren	87
3.5	Steife Lösungsfamilien nichtlinearer Probleme	88
3.5.1	Probleme beim Einfrieren von Koeffizienten	89
3.5.2	Inhomogene lineare Probleme	91
3.5.3	Allgemeine nichtlineare Probleme	97

In diesem Abschnitt wird ein familienbasierter, asymptotischer Ansatz zur Beschreibung der mathematischen Struktur steifen Verhaltens von Lösungen von Differentialgleichungen entwickelt. Wie in Kapitel 2 schon gesehen, ist ein wesentliches Merkmal steifer Lösungen das schnelle Relaxationsverhalten. In Abschnitt 3.1 werden durch die langsamen Mannigfaltigkeiten diejenigen Punkte des Phasenraum identifiziert, in denen keine exponentiell schnelle Bewegung stattfindet. Vielmehr relaxieren steife Lösungen auf die langsamen Mannigfaltigkeiten. In Abschnitt 3.2 wird das Relaxationsverhalten von Lösungsfamilien studiert. Dazu werden asymptotisch exponentiell anziehende Mengen betrachtet. Insbesondere wird auf Unterschiede zu dem Langzeitverhalten aus der Theorie der dynamischen Systeme eingegangen. In Abschnitt 3.3 werden die Konzepte aus den ersten beiden Abschnitten dieses Kapitels verwendet, um zu definieren was in dieser Arbeit unter einer steifen Familie von Lösungen verstanden wird. Konkret wird eine Lösungsfamilie steif heißen, wenn es eine exponentiell anziehende langsame Mannigfaltigkeit gibt. Für lineare Probleme werden die Implikationen der Definition in Abschnitt 3.4 diskutiert. Schon im linearen Fall wird die mathematische Struktur steifer Lösungen deutlich. In Abschnitt 3.4.4 wird gezeigt, dass für steife Lösungsfamilien linearer Probleme im Sinne dieser

Arbeit explizite Runge-Kutta-Verfahren tatsächlich ungeeignet sind numerische Näherungslösungen zu finden. Der nichtlineare Fall, der in Abschnitt 3.5 behandelt wird, erweist sich als deutlich schwieriger, was unter anderem daran liegt, dass die langsamen Mannigfaltigkeiten in diesem Fall zeitabhängig sein können. In Abschnitt 3.5.2 wird auf linear inhomogene Probleme eingegangen, welche einen wichtigen Spezialfall nichtlinearer Probleme darstellen. Für den allgemeinen nichtlinearen Fall wird in Abschnitt 3.5.3 eine Vermutung geäußert, wie sich steifes Verhalten vom linearen Fall überträgt.

3.1 LANGSAME MANNIGFALTIGKEITEN

Bei der Analyse dynamischer Systeme spielen Fixpunkte eine besondere Rolle; sie beschreiben die Ruhelagen des Systems. Für das System $\dot{x} = f(x)$ sind die Fixpunkte gerade die Nullstellen von f . Ruhelagen in diesem Sinne sind also jene Punkte, in denen das Geschwindigkeitsfeld verschwindet. Auch ohne Kenntnis der exakten Lösung kann eine Analyse der Fixpunkte zu qualitativen Aussagen des Lösungsverhaltens führen.

Wie schon in Abschnitt 2 diskutiert, bereitet das exponentielle Abklingen aus numerischer Sicht bei expliziten Verfahren Probleme. Dabei sind Punkte des Phasenraums, in denen die Geschwindigkeit nicht asymptotisch exponentiell ist, unproblematisch.

In den *langsamen Mannigfaltigkeiten* sollen nun gerade diejenigen Punkte des Phasenraums liegen, in denen die Geschwindigkeit auf eine gewisse Weise langsam, nämlich nicht exponentiell ist. Die Betonung liegt dabei ausdrücklich auf „gewisse Weise“. Die *langsamen Mannigfaltigkeiten* verallgemeinern damit den Begriff der Ruhelage. Sie werden sich als wesentlich für die strukturelle Beschreibung steifen Verhaltens erweisen.

Zunächst werden *langsame Mannigfaltigkeiten* für Familien linearer autonomer Systeme definiert (Abschnitt 3.1.1). Im nichtlinearen Fall (Abschnitt 3.1.2) wird die *langsame Mannigfaltigkeit* durch die *langsamen Mannigfaltigkeiten* der Ableitungen lokal approximiert.

3.1.1 Langsame Mannigfaltigkeit linearer Probleme

Zu Beginn des Abschnittes soll an die O-Notation erinnert werden.

Definition 3.1 (Eckhaus (1979, S. 5-6)). Seien f, g zwei reelle stetige Funktionen auf einem halboffenem Intervall $(0, \varepsilon_0]$. Dann sei

- (i) $f = \mathcal{O}(g)$ für $\varepsilon \rightarrow 0$, falls positive Zahlen k und C existieren mit $|f(\varepsilon)| \leq k|g(\varepsilon)|$ für $0 < \varepsilon < C$.
- (ii) $f = \mathcal{o}(g)$ für $\varepsilon \rightarrow 0$, falls $\lim_{\varepsilon \rightarrow 0} f(\varepsilon)/g(\varepsilon) = 0$.

$$(iii) f \preceq g \Leftrightarrow f = \mathcal{O}(g).$$

$$(iv) f \prec g \Leftrightarrow f = \mathcal{o}(g).$$

$$(v) f \approx g \Leftrightarrow f \preceq g \text{ und } g \preceq f.$$

$$(vi) f = \mathcal{O}_S(g) \Leftrightarrow f = \mathcal{O}(g) \text{ und } f \neq \mathcal{o}(g).$$

Im Folgenden wird der Begriff der *langsamen Mannigfaltigkeit* für Familien von autonomen linearen Differentialgleichungen definiert. Betrachtet wird eine Familie $A_\varepsilon \in \text{Mat}(n, \mathbb{R})$ von Endomorphismen des \mathbb{R}^n und die dadurch gegebene Familie autonomer, linearer, homogener gewöhnlicher Differentialgleichungen

$$\dot{x}_\varepsilon(t) = A_\varepsilon x(t). \quad (3.1)$$

Definition 3.2. Die Familie von Systemen (3.1) heißt in dem Punkt $x_0 \in \mathbb{R}^n$

(i) *asymptotisch langsam*, falls

$$\|A_\varepsilon x_0\|_\infty = \mathcal{O}(1) \quad (3.2)$$

(ii) und *asymptotisch schnell*, falls

$$\|A_\varepsilon x_0\|_\infty = \mathcal{O}_S(1/\varepsilon). \quad (3.3)$$

Bemerkung 3.1. Streng genommen ist asymptotisch gesehen eine Geschwindigkeit der Ordnung $\mathcal{O}_S(1)$ nicht langsam, sondern „normal“ schnell. Dennoch soll an diesem Terminus festgehalten werden.

Für das Spektrum $\sigma(A_\varepsilon)$ von A_ε betrachte man die Zerlegung

$$\sigma(A_\varepsilon) = \sigma_1(A_\varepsilon) \cup \sigma_{1/\varepsilon}(A_\varepsilon) \quad (3.4a)$$

mit

$$\sigma_1(A_\varepsilon) = \left\{ \lambda \in \sigma(A_\varepsilon) \mid \text{Re } \lambda = \mathcal{O}(1) \right\} \quad (3.4b)$$

und

$$\sigma_{1/\varepsilon}(A_\varepsilon) = \{ \lambda \in \sigma(A_\varepsilon) \mid 1 = \mathcal{O}(\text{Re } \lambda) \}. \quad (3.4c)$$

Dabei können die Eigenwerte λ von ε abhängen. In fast allen nachfolgenden Beispielen sind die Eigenwerte aus $\sigma_{1/\varepsilon}(A_\varepsilon)$ von der Ordnung $\mathcal{O}_S(1/\varepsilon)$.

Definition 3.3. Der Eigenwert $\lambda \in \sigma(A_\varepsilon)$ heißt

- *klein* oder *langsam*, falls $\lambda \in \sigma_1(A_\varepsilon)$ gilt und
- *groß* oder *schnell*, falls $\lambda \in \sigma_{1/\varepsilon}(A_\varepsilon)$ gilt.

Für den Phasenraum wird eine Zerlegung in *langsame* und *schnelle* Haupträume betrachtet. Dabei bezieht sich *langsam* und *schnell* auf die die Haupträume definierenden Eigenwerte. Es sei

$$\mathbb{R}^n = \text{Hau}_1(A_\varepsilon) \oplus \text{Hau}_{1/\varepsilon}(A_\varepsilon) \quad (3.5a)$$

mit

$$\text{Hau}_1(A_\varepsilon) = \bigoplus_{\lambda \in \sigma_1(A_\varepsilon)} \text{Hau}(A_\varepsilon; \lambda) \quad (3.5b)$$

und

$$\text{Hau}_{1/\varepsilon}(A_\varepsilon) = \bigoplus_{\lambda \in \sigma_{1/\varepsilon}(A_\varepsilon)} \text{Hau}(A_\varepsilon; \lambda). \quad (3.5c)$$

Zur Erinnerung: Für einen Eigenwert $\lambda \in \sigma(A_\varepsilon)$ mit algebraischer Vielfachheit r ist der Hauptraum zu λ durch

$$\text{Hau}(A_\varepsilon, \lambda) = \{v \in \mathbb{R}^n \mid (A_\varepsilon - \lambda)^r v = 0\} \quad (3.6)$$

gegeben. Damit lässt sich für den linearen Fall die *langsamen Mannigfaltigkeit* definieren.

Definition 3.4. Gegeben sei das lineare autonome System aus (3.1). Dann heißt

$$\mu_\varepsilon = \mu_\varepsilon(A_\varepsilon) = \text{Hau}_1(A_\varepsilon) \quad (3.7)$$

die zu A_ε gehörende Familie *langsamer Mannigfaltigkeiten*, oder abkürzend die zu A_ε gehörende *langsame Mannigfaltigkeit*. Hängt μ_ε nicht von ε ab, wird einfach μ geschrieben.

Die Basisvektoren des Hauptraums, werden dabei so gewählt, dass die einzelnen Komponenten von der Ordnung $\mathcal{O}(1)$ sind. Zum Beispiel wähle man $(1, \varepsilon)$ anstatt $(1/\varepsilon, 1)^T$.

Bemerkung 3.2. Als direkte Summe von Haupträumen ist μ_ε invariant unter A_ε .

Der Begriff der langsamen Mannigfaltigkeit soll an einigen Beispielen verdeutlicht werden.

Beispiel 3.1. Das einfachste Beispiel ist wieder das dahlquistsche Testproblem

$$\dot{x} = A_\varepsilon x = -\frac{1}{\varepsilon}x. \quad (3.8)$$

Da der einzige Eigenwert von A_ε von der Ordnung $\mathcal{O}_S(1/\varepsilon)$ ist, gilt für die langsame Mannigfaltigkeit $\mu = \{0\}$.

Beispiel 3.2. Für das Problem

$$\dot{x} = \begin{pmatrix} -1 & 0 \\ 0 & -\frac{1}{\varepsilon} \end{pmatrix} x \quad (3.9)$$

ist die langsame Mannigfaltigkeit durch $\mu = \text{span}(e_1)$ gegeben, wobei e_1 den ersten Einheitsvektor bezeichne. Das System ist auf μ asymptotisch langsam.

Bemerkung 3.3. Nach dem Satz über die Hauptraumzerlegung lässt sich jeder Vektorraum V über einem algebraisch abgeschlossenen Körper stets in die Haupträume eines Endomorphismus über V zerlegen. Ist der Endomorphismus diagonalisierbar, zerfällt V sogar in dessen Eigenräume. Ist A_ε für alle $\varepsilon > 0$ diagonalisierbar, dann ist die Familie langsamer Mannigfaltigkeiten also gerade direkte Summe der Eigenräume zu den Eigenwerten mit Realteilen der Ordnung $\mathcal{O}(1)$. Das System ist dann auf der langsamen Mannigfaltigkeit asymptotisch langsam.

Ist die Familie der Endomorphismen nicht für alle $\varepsilon > 0$ diagonalisierbar, kann es Punkte der langsamen Mannigfaltigkeit geben, in denen das System asymptotisch schnell ist. Das wird durch das nächste Beispiel gezeigt.

Beispiel 3.3. Für das System

$$\dot{x} = A_\varepsilon x = \begin{pmatrix} 1 & \frac{1}{\varepsilon} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{\varepsilon} \end{pmatrix} x \quad (3.10)$$

ist die langsame Mannigfaltigkeit durch $\mu = \text{span}(e_1, e_2)$ gegeben. In $x_0 \in \text{span}(0, 1, 0)^\top$ ist das System asymptotisch schnell.

Wie sich später zeigen wird, ist dies für steifes Verhalten aber irrelevant. Entscheidend für steifes Verhalten ist nur, dass das System in Richtung der langsamen Eigenvektoren asymptotisch langsam ist.

Die Eigenwerte dürfen asymptotisch große Imaginärteile haben; das „langsam“ in langsame Mannigfaltigkeit bezieht sich lediglich auf die Größenordnung der Realteile der Eigenwerte, wie das nächste Beispiel deutlich macht.

Beispiel 3.4. In rein oszillatorischen Problemen verschwinden die Realteile der Eigenwerte gänzlich. Zum Beispiel ist das Spektrum des Systems

$$\dot{x} = A_\varepsilon x = \begin{pmatrix} 0 & -1/\varepsilon \\ 1/\varepsilon & 0 \end{pmatrix} x \quad (3.11)$$

durch $\sigma(A_\varepsilon) = \{i/\varepsilon, -i/\varepsilon\}$ gegeben. Die langsame Mannigfaltigkeit ist dann gerade der ganze Raum $\mu = \mathbb{R}^2$.

Die Kernidee bei der Definition der langsamen Mannigfaltigkeit ist den Teil des Phasenraumes zu identifizieren, in dem die Lösung des Systems für $\varepsilon \rightarrow 0$ keine exponentiell schnelle reelle Geschwindigkeitskomponente hat. Für Startwerte außerhalb der langsamen Mannigfaltigkeit kann es zu einem exponentiellen Abklingen von Lösungskomponenten und zu steifem Verhalten kommen. Dies wird später noch ausführlich diskutiert.

3.1.2 Langsame Mannigfaltigkeiten nichtlinearer Probleme

Als Verallgemeinerung zu dem linearen System aus (3.1) wird nun der nichtlineare Fall

$$\dot{x}(t) = f_\varepsilon(x(t)) \quad (3.12)$$

betrachtet. Es wird davon ausgegangen, dass die rechte Seite hinreichend glatt ist. Für einen beliebigen Punkt des Phasenraums x_0 sei $df_\varepsilon(x_0)$ die Ableitung in x_0 . Langsame Mannigfaltigkeiten werden jetzt durch die langsamen Mannigfaltigkeiten der Ableitungen $\mu(df_\varepsilon(x_0))$ lokal approximiert. Im Folgenden werden nur solche Probleme betrachtet, für die nachfolgende Definition sinnvoll ist.

Definition 3.5. Die Mannigfaltigkeit M ist eine *langsame Mannigfaltigkeit* von f_ε , falls für alle $x_0 \in M$ folgendes gilt:

$$T_{x_0}M = \mu_\varepsilon(df_\varepsilon(x_0)), \quad \text{für } \varepsilon \rightarrow 0. \quad (3.13)$$

Hierbei bezeichne $T_{x_0}M$ den geometrischen Tangentialraum an M im Punkt x_0 .

Die folgenden Beispiele sollen die Definition verdeutlichen.

Beispiel 3.5. Für das Beispiel 2.3 auf Seite 39 mit $\alpha = 2$

$$\dot{x} = f_\varepsilon(x) = -\frac{1}{\varepsilon}x^2 \quad (3.14)$$

ist jede diskrete Teilmenge $M \subset \mathbb{R} \setminus \{0\}$ eine langsame Mannigfaltigkeit. Für alle $x_0 \neq 0$ ist $\mu(df_\varepsilon(x_0)) = \{0\}$. Die Untermannigfaltigkeiten von \mathbb{R} sind entweder offene oder diskrete Teilmengen. Jede diskrete Teilmenge M hat für alle $x_0 \in M$ den Tangentialraum $T_{x_0}M = \{0\}$. Also ist jede diskrete Teilmenge $M \subset \mathbb{R} \setminus \{0\}$ eine langsame Mannigfaltigkeit. Weitere langsame Mannigfaltigkeiten gibt es nicht: Für jede offene Menge $N \subset \mathbb{R}$ und alle $x_0 \in N$ ist $T_{x_0}N = \mathbb{R}$.

Nachfolgend wird das Beispiel des nichtlinearen Strudels (Beispiel 2.6 auf Seite 48) aufgegriffen und in einer allgemeineren Form betrachtet.

Beispiel 3.6. Als Verallgemeinerung zu (2.39) werden die Differentialgleichungen

$$\dot{x} = f_\varepsilon(x) = \beta(1 - \|x\|^2)x + \alpha \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} x, \quad (3.15a)$$

mit

$$\beta \in \left\{ \pm \frac{1}{\varepsilon} \right\}, \quad \alpha \in \{1, \beta\}, \quad \varepsilon > 0 \quad (3.15b)$$

betrachtet. In Abbildung 2.11 auf Seite 48 ist der Phasenraum für $\beta = 1/\varepsilon$ graphisch dargestellt.

Es sollen nun alle langsamen Mannigfaltigkeiten gefunden werden. Zunächst werden dazu die langsamen Mannigfaltigkeiten der Ableitungen bestimmt. In $x_0 = (x_1, x_2)$ ist die Ableitungen von f_ε durch

$$df_\varepsilon(x_0) = \begin{pmatrix} \beta(1 - 3x_1^2 - x_2^2) & -\alpha - 2\beta x_1 x_2 \\ \alpha - 2\beta x_1 x_2 & \beta(1 - x_1^2 - 3x_2^2) \end{pmatrix} \quad (3.16)$$

mit den dazugehörigen Eigenwerten

$$\lambda_{1/2} = \beta \left(1 - 2\|x_0\|^2 \pm \sqrt{\|x_0\|^4 - \left(\frac{\alpha}{\beta}\right)^2} \right) \quad (3.17)$$

und den entsprechenden Eigenvektoren

$$v_{1/2} = \left(-\beta \left(x_1^2 - x_2^2 \mp \sqrt{\|x_0\|^4 - \left(\frac{\alpha}{\beta}\right)^2} \right), \alpha - 2\beta x_1 x_2 \right). \quad (3.18)$$

gegeben.

Für $\alpha = \beta$ hat der Ausdruck auf der rechten Seite von (3.17) keine reellen Nullstellen in $\|x_0\|$. Das bedeutet, dass die Realteile der Eigenwerten in jedem Punkt höchstens von der Größenordnung $\mathcal{O}_S(1)$ sind. Die langsame Mannigfaltigkeit ist immer der Nullraum $\mu_\varepsilon = \{0\}$.

Für $\alpha = 1$ sind die Eigenwerte aus (3.17) nach asymptotischer Entwicklung durch

$$\lambda_{1/2} = -\beta \left(\|x_0\|^2 (2 \mp 1) - 1 \right) + \mathcal{O}(\varepsilon) \quad (3.19)$$

gegeben.

Für $\|x_0\| \notin \{1, 1/\sqrt{3}\}$ gibt es wieder keinen Eigenwert mit asymptotisch kleinem Realteil und die langsame Mannigfaltigkeit ist wiederum der Nullraum $\mu = \{0\}$.

Ist $\|x_0\| \in \{1, 1/\sqrt{3}\}$ verschwindet für jeweils einen Eigenwert der asymptotisch große Realteil. Nach (3.18) ist die langsame Mannigfaltigkeit für $\|x_0\| = 1$ durch $\mu = \text{span}((x_2, -x_1)^T)$ und für $\|x_0\| = 1/\sqrt{3}$ durch $\mu = \text{span}((1/3 - x_2^2, x_1 x_2)^T)$ gegeben.

Damit sind die langsamen Mannigfaltigkeiten der Ableitungen in allen Punkten des Phasenraums bestimmt. Es werden nun die langsamen Mannigfaltigkeiten im nichtlinearen Falle diskutiert.

Für $\alpha = \beta$ sind genau die diskreten Teilmengen des \mathbb{R}^2 langsame Mannigfaltigkeiten, denn die diskreten Teilmengen sind die einzigen, die als Tangentialraum den Nullraum haben.

Für $\alpha = 1$ ist die S^1 langsame Mannigfaltigkeit: Für $x_0 \in S^1$ mit $x_0 = (x_1, x_2)$ ist $T_{x_0} S^1 = \text{span}((x_2, -x_1)^\top) = \mu_\varepsilon(\text{df}_\varepsilon(x_0))$ für $\varepsilon \rightarrow 0$. Damit sind auch alle eindimensionalen Untermannigfaltigkeiten der S^1 langsame Mannigfaltigkeiten.

Es sei bemerkt, dass für $\beta > 0$ die S^1 eine attraktive ω -Limesmenge ist. Für $\beta < 0$ ist dies nicht der Fall, dennoch ist die S^1 auch dann eine langsame Mannigfaltigkeit.

Ebenso sind alle nulldimensionalen Mannigfaltigkeiten, die weder auf der S^1 noch auf dem Kreis mit Radius $1/\sqrt{3}$ liegen, langsame Mannigfaltigkeiten.

Die Vermutung liegt nahe, dass analog zur S^1 auch der Kreis mit Radius $1/\sqrt{3}$ eine langsame Mannigfaltigkeit liefert das ist allerdings nicht der Fall, denn in allen Punkten auf diesem Kreis ist Bedingung (3.13) verletzt.

Als letztes Beispiel zu den langsamen Mannigfaltigkeiten nichtlinearer Probleme wird der Van-der-Pol-Oszillator betrachtet.

Beispiel 3.7. In diesem Beispiel soll lediglich verifiziert werden, dass die Menge

$$M = \left\{ x = (x_1, x_2) \mid x_1 - \frac{1}{3} x_2^3 + x_2 = 0, x_2^2 > 1 \right\} \quad (3.20)$$

eine langsame Mannigfaltigkeit des Van-der-Pol-Oszillators aus Beispiel 2.8 auf Seite 52 ist. Der Van-der-Pol-Oszillator ist zur Erinnerung durch die Gleichung

$$\dot{x} = f_\varepsilon(x) = \begin{pmatrix} -x_2 \\ 1/\varepsilon (x_1 - \frac{1}{3} x_2^3 + x_2) \end{pmatrix} \quad (3.21)$$

gegeben. Wie man leicht sieht, gilt für $x_0 = (x_1, x_2) \in M$

$$T_{x_0} M = \text{span} \left((x_2^2 - 1, 1)^\top \right). \quad (3.22)$$

Die Ableitung von f_ε in $x_0 = (x_1, x_2)$ ist durch

$$\text{df}_\varepsilon(x_0) = \begin{pmatrix} 0 & -1 \\ 1/\varepsilon & (1 - x_2^2)/\varepsilon \end{pmatrix} \quad (3.23)$$

gegeben. Diese hat die Eigenwerte

$$\lambda_{1/2} = \frac{(1 - x_2^2) \pm \sqrt{(x_2^2 - 1)^2 - 4\varepsilon}}{2\varepsilon} \quad (3.24a)$$

$$= \frac{(1 - x_2^2) \pm |x_2^2 - 1|}{2\varepsilon} \mp \frac{1}{|x_2^2 - 1|} + \mathcal{O}(\varepsilon) \quad (3.24b)$$

und die Eigenvektoren

$$v_{1/2} = \left(\frac{1}{2}(x_2^2 - 1 \pm |x_2^2 - 1|), 1 \right)^T + \mathcal{O}(\varepsilon). \quad (3.25)$$

Für $x_2^2 > 1$ ist $\lambda_1 = (1 - x_2^2)^{-1} + \mathcal{O}(1)$ langsamer Eigenwert mit dazugehörigem Eigenvektor $v_1 = (x_2^2 - 1, 1)^T + \mathcal{O}(1)$. Damit gilt für alle $x_0 \in M$

$$T_{x_0}M = \mu(df_\varepsilon(x_0)) \quad (3.26)$$

für $\varepsilon \rightarrow 0$. Das zeigt, dass M eine langsame Mannigfaltigkeit des Van-der-Pol-Oszillators ist.

3.2 EXPONENTIELL ANZIEHENDE MENGEN

Im klassischen Sinn zeichnen sich steife Differentialgleichungen dadurch aus, dass sie schnell abklingende Lösungskomponenten besitzen. In den Beispielen des dahlquistschen Testproblems und der Verallgemeinerung mit $\alpha > 1$ (Beispiel 2.2 auf Seite 37 und 2.3 auf Seite 39) wurde gezeigt, dass für steifes Verhalten exponentielles Abklingen nötig ist und dass einfaches Abklingen nicht ausreicht.

In diesem Kapitel wird geklärt, was in dieser Arbeit unter *schneller* Relaxation verstanden wird. Das führt zu dem Begriff der (asymptotisch) exponentiell anziehenden Menge. Es wird einige Parallelen zu Attraktoren geben, wie man sie aus der Theorie der dynamischen Systeme kennt, aber auch entscheidende Unterschiede.

In Beispiel 3.1 auf Seite 62 wurde die langsame Mannigfaltigkeit des dahlquistschen Testproblems mit $\mu(f_\varepsilon) = \{0\}$ bestimmt. An der exakten Lösung (2.10) erkennt man, dass die Lösungskurven für $x_0 \neq 0$ exponentiell auf den einzigen Punkt der linearen Mannigfaltigkeit $x = 0$ abklingen.

Allgemeiner formuliert wird der Abstand der Lösung $\Phi_\varepsilon^t x_0$ zu $\mu(f_\varepsilon)$ für hinreichend kleines ε durch eine abklingende Exponentialfunktion dominiert

$$\text{dist}(\Phi_\varepsilon^t x_0, \mu_\varepsilon) \leq \text{dist}(x_0, \mu_\varepsilon) \exp\left(-\frac{1}{\varepsilon}t\right). \quad (3.27)$$

Für $t = 0$ ist diese Abschätzung scharf. Für die weiteren Betrachtungen ist nur das asymptotische Verhalten für $\varepsilon \rightarrow 0$ bei positiven Zeiten $t > 0$ relevant. Das führt zu der nächsten Definition.

Definition 3.6. Sei M eine Teilmenge des Phasenraums der Differentialgleichung

$$\dot{x}(t) = f_\varepsilon(x(t)) \quad (3.28)$$

und $x(0) = x_0 \notin M$ ein Anfangswert. Die Menge M heißt (*asymptotisch*) *exponentiell anziehend*, falls für den Phasenfluss des Anfangswertproblems Φ_ε^t mit x_0 aus einer hinreichend kleinen Umgebung um M zu jedem fest gewählten Zeitpunkt $t > 0$

$$0 \neq \text{dist}(\Phi_\varepsilon^t x_0, M) = \mathcal{O}\left(\exp\left(-\frac{Ct}{\varepsilon}\right)\right) \quad \text{für } \varepsilon \rightarrow 0 \quad (3.29)$$

gilt. Dabei ist C eine problemabhängige von ε unabhängige Konstante. M heißt (*asymptotisch*) *anziehend*, falls zu jedem fest gewählten $t > 0$

$$0 \neq \text{dist}(\Phi_\varepsilon^t x_0, M) = \mathcal{O}(1) \quad \text{für } \varepsilon \rightarrow 0 \quad (3.30)$$

gilt.

Bemerkung 3.4. In (3.28) wurde der Einfachheit halber das exponentielle Abklingen durch die Größenordnung $\mathcal{O}(\exp(-\frac{t}{\varepsilon}))$ beschrieben. In allen in dieser Arbeit betrachteten Beispielen ist das auch ausreichend. Allgemeiner würde man statt $1/\varepsilon$ diejenige Ordnungsfunktion wählen, die in (3.4b) die Größenordnung von $\text{Re } \lambda$, $\lambda \in \sigma_{1/\varepsilon}(A_\varepsilon)$ beschreibt. Damit ist aber auch klar, dass die Realteile der Eigenwerte in $\sigma_{1/\varepsilon}(A_\varepsilon)$ nicht *nur* von der Ordnung $\mathcal{O}_S(\ln(1/\varepsilon))$ sein dürfen. Andernfalls gäbe es in (3.28) kein exponentielles Abklingen. In Definition 3.4 auf Seite 62 wurde das stillschweigend vorausgesetzt.

In Definition 3.6 auf der vorherigen Seite muss das Existenzintervall der Lösung genauer betrachtet werden. Während bei linearen Problemen das Existenzintervall unbeschränkt ist, muss dies bei nichtlinearen Problemen im Allgemeinen nicht der Fall sein. Bedingung (3.29) ergibt also nur auf dem ε -abhängigen Existenzintervall Sinn und genau so ist sie auch zu verstehen.

Ist M ein linearer Raum und π_{M^\perp} die orthogonale Projektion auf das orthogonale Komplement von M , lässt sich (3.29) in der Form

$$\|\pi_{M^\perp} \Phi_\varepsilon^t x_0\| = \mathcal{O}\left(\exp\left(-\frac{Ct}{\varepsilon}\right)\right) \quad (3.31)$$

schreiben. Da hier nur das asymptotische Verhalten von Interesse ist und auf endlichdimensionalen Vektorräumen alle Normen äquivalent sind, ist (3.31) unabhängig von der Wahl einer Norm.

Der Vollständigkeit halber seien hier nochmals die Beispiele 2.2 und 2.3 aufgeführt.

Beispiel 3.8. Die Lösung des dahlquistischen Testproblems

$$\dot{x} = -\frac{1}{\varepsilon}x, \quad x(0) = x_0, \quad (3.32)$$

ist durch $x(t) = x_0 \exp(-1/\varepsilon t)$ gegeben. Für alle x_0 ist die Menge $M = \{0\}$ asymptotisch exponentiell anziehend. Für alle $t < 0$ gilt

$$\text{dist}(\Phi_\varepsilon^t x_0, M) = \mathcal{O}\left(\exp\left(-\frac{Ct}{\varepsilon}\right)\right). \quad (3.33)$$

Beispiel 3.9. Die Lösung von

$$\dot{x} = -\frac{1}{\varepsilon}x^2, \quad x(0) = x_0 \quad (3.34)$$

ist durch

$$x(t) = \frac{x_0}{\frac{x_0}{\varepsilon}t + 1} \quad (3.35)$$

gegeben. Die Lösung klingt für $x_0, t > 0$ und $\varepsilon \rightarrow 0$ ab, also ist die Menge $M = \{0\}$ asymptotisch anziehend, aber offensichtlich nicht exponentiell. Weiter muss jede anziehende Menge von (3.34) notwendigerweise die 0 enthalten.

Beispiel 3.10. Für den nichtlinearen Strudel aus den Beispielen 2.6 auf Seite 48 und 3.6 auf Seite 65,

$$\dot{x} = \beta \left(1 - \|x\|^2\right) + \alpha \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} x, \quad (3.36a)$$

mit

$$\beta \in \left\{ \pm \frac{1}{\varepsilon} \right\}, \quad \alpha \in \{1, \beta\}, \quad (3.36b)$$

werden die Mengen $M_0 = \{0\}$ und $M_1 = S^1$ betrachtet. Wie in Beispiel 2.6 auf Seite 48 wird die Abstandsfunktion $u = r^2 = \langle x, x \rangle$ betrachtet. Die analytische Lösung mit dem Anfangswert $u(0) = u_0 = \|x_0\|_2^2 > 0$ ist berechenbar und durch

$$u(t) = 1 - \frac{1 - u_0}{u_0(\exp(2\beta t) - 1) + 1} \quad (3.37)$$

gegeben. Hieran ist zu erkennen, dass für $\beta = 1/\varepsilon$ die Menge M_1 exponentiell anziehend ist. Für $\beta = -1/\varepsilon$ ist die Menge M_0 exponentiell anziehend.

Betrachtet man für einen Startwert $u_0 > 1$ die Integralkurve für ein festes $\varepsilon > 0$ und $\beta < 0$, lässt sich leicht überprüfen, dass das Existenzintervall $[0, t_{\text{end}})$ mit

$$t_{\text{end}} = \frac{\log\left(1 - \frac{1}{u_0}\right)}{2\beta} > 0$$

endlich ist. Für $t \rightarrow t_{\text{end}}$ gilt dann $u(t) \rightarrow \infty$. Asymptotisch gesehen verschwindet die Länge des Existenzintervalls für $|\beta| \rightarrow \infty$ bzw. $\varepsilon \rightarrow 0$ unabhängig von $u_0 > 0$. Bedingung (3.29) ergibt hier also keinen Sinn.

Man beachte, dass in den Beispielen Wissen über das Lösungsverhalten verwendet wurde, um zu entscheiden, ob eine Menge (exponentiell) anziehend ist.

Auf den ersten Blick mag es nicht einleuchten, warum ein besonderes Augenmerk auf die exponentiell anziehenden Mengen geworfen wird. Der Begriff der exponentiell anziehenden Menge ähnelt dem Begriff des Attraktors oder ω -Limesmenge aus der Theorie der dynamischen Systeme. Dort betrachtet man allerdings das Verhalten für $t \rightarrow \infty$ (bzw. $t \rightarrow t_{\text{end}}$). Wie sich später zeigen wird, führen diese Begriffe bei der Untersuchung auf steifes Verhalten von Mehrskalenproblemen aber nicht weiter. Bei Mehrskalenproblemen kann das einen entscheidenden und für steifes Verhalten wichtigen Unterschied machen.

Beispiel 3.11. Man betrachte das autonome homogene lineare Problem $\dot{x} = Ax$ der Dimension $n = 2$. In der Theorie der dynamischen Systeme lässt sich das Lösungsverhalten für $t \rightarrow \infty$ vollständig anhand der Eigenwerte klassifizieren. Dazu betrachtet man wie z.B. in Schmitt (2012, S. 25) das charakteristische Polynom von A

$$\det(\lambda - A) = \lambda^2 - \sigma\lambda + \delta, \quad \sigma = \text{tr}(A), \quad \delta = \det(A). \quad (3.38)$$

Mit der Diskriminante $\Delta = \sigma^2 - 4\delta$ sind die Eigenwerte durch

$$\lambda_{1/2} = \frac{1}{2}(\sigma \pm \sqrt{\Delta}) = \frac{\sigma}{2} \pm \sqrt{\frac{\sigma^2}{4} - \delta} \quad (3.39)$$

gegeben. Es werden dann die Fälle $\Delta \neq 0$ und $\Delta = 0$ bei den verschiedenen Vorzeichen der Realteile der Eigenwerte betrachtet.

Ist $\Delta \neq 0$ gibt es zwei verschiedene Eigenwerte λ_1 und λ_2 mit zwei dazugehörigen linear unabhängigen Eigenvektoren v_1 und v_2 . Schmitt (ebd., S. 26) gibt in diesem Fall für die beiden Konstanten c_1 und c_2 die allgemeine Lösung mit

$$x(t) = c_1 e^{\lambda_1 t} v_1 + c_2 e^{\lambda_2 t} v_2 \quad (3.40)$$

an. In Beispiel 3.2 auf Seite 63 sind die beiden Eigenwerte $\lambda_1 = -1$ und $\lambda_2 = -1/\varepsilon$. Für $\varepsilon > 0$ ist $x = 0$ anziehender Attraktor des Systems. Die Menge $M = \{0\}$ ist aber nicht asymptotisch anziehend. Asymptotisch gesehen relaxiert das System auf den Eigenraum zum Eigenwert λ_1 . Dieser erzeugt in Beispiel 3.2 die langsame Mannigfaltigkeit des Problems. In der Lösung $x(t) = (\exp(-t), \exp(-t/\varepsilon))^T$ erkennt man deutlich die beiden Zeitskalen t und t/ε . Bei solchen Mehrskalenproblemen ist es also nicht ausreichend nur das Verhalten für $t \rightarrow \infty$ zu untersuchen.

Für $\Delta = 0$ ist $\lambda_1 = \lambda_2 = \lambda$ mit dem Eigenvektor v_1 und einem Hauptvektor v_2 . Die Lösung wird in diesem Fall mit

$$x(t) = e^{\lambda t}((c_1 + c_2 t)v_1 + c_2 v_2) \quad (3.41)$$

angegeben. Im Kontext der dynamischen Systeme ist der Fixpunkt $x = 0$ für $\lambda < 0$ attraktiv, da dann die Lösung für $t \rightarrow \infty$ auf den

Nullpunkt fällt. Für $\lambda = -1$, ist die Menge $M = \{0\}$ aber nicht asymptotisch anziehend. An der Lösung erkennt man, dass es sich um ein Einskalenproblem handelt. Für $\lambda = -1/\varepsilon$ spielt es deshalb keine Rolle, ob man den Grenzwert für $\varepsilon \rightarrow 0$ bei festem t oder aber den Grenzwert $t \rightarrow \infty$ bei festem $\varepsilon > 0$ betrachtet.

Untersucht man ein System auf Attraktoren, betrachtet man das Langzeitverhalten für eine Zeitskala $t \rightarrow \infty$. Betrachtet man hingegen Familien von Systemen und stellt die Frage nach exponentiell anziehenden Mengen, betrachtet man das Grenzverhalten für $\varepsilon \rightarrow 0$ bei festem t . Wie man an (3.40) und (3.41) sieht steckt die ε -Abhängigkeit in den Eigenwerten und damit in der Familie von Problemen. Ebenfalls an (3.40) und (3.41) erkennt man, dass die Eigenwerte zusammen mit der Zeit auftreten. In (3.41) spielt es keine Rolle ob $\lambda \rightarrow \infty$ oder $t \rightarrow \infty$ betrachtet. In (3.40) ist das anders, da spielt es nur dann keine Rolle, wenn sich λ_1 und λ_2 asymptotisch gleich verhalten. Das ist ein wesentlicher Unterschied. Einmal hat man es mit einem Einskalen- das andere Mal mit einem Zweiskalenproblem zu tun.

3.3 STEIFE FAMILIEN VON LÖSUNGEN

Mit Hilfe der Begriffe der langsamen Mannigfaltigkeiten und der exponentiell anziehenden Mengen lassen sich nun *steife Lösungen* definieren.

Definition 3.7. Gegeben sei eine Familie von Anfangswertproblemen, zu einem ε -unabhängigen Startwert.

$$\dot{x}(t) = f_\varepsilon(x(t)), \quad x(0) = x_0. \quad (3.42)$$

Die Lösungsfamilie durch x_0 heißt *steif*, falls es eine langsame Mannigfaltigkeit M von f_ε gibt, die exponentiell anziehend ist:

$$0 \neq \text{dist}(\Phi_\varepsilon^t x_0, M) = \mathcal{O}\left(\exp\left(-\frac{Ct}{\varepsilon}\right)\right). \quad (3.43)$$

Hierbei ist C wieder eine problemabhängige, ε -unabhängige Konstante.

Die Menge aller Anfangswerte x_0 mit steifer Lösungsfamilie bekommt einen eigenen Namen.

Definition 3.8. Gegeben sei die Differentialgleichung

$$\dot{x} = f_\varepsilon(x) \quad (3.44)$$

und bezeichne $\Phi_\varepsilon^t x_0$ die zu dem Anfangswert $x_0 = x(0)$ gehörende Lösung. Dann heißt die Menge

$$S(f_\varepsilon) = \left\{ x_0 \mid \Phi_\varepsilon^t x_0 \text{ ist steif.} \right\} \quad (3.45)$$

die *Steifheitsmenge* der Differentialgleichung.

Nicht selten ist es üblich, allgemein von steifen Differentialgleichungen zu sprechen. Die letzten beiden Definitionen machen aber deutlich, dass es vielmehr Lösungen zu konkreten Anfangswerten sind, die steif sind. Das passt zu einer Aussage von C. W. Gear im Jahre 1982:

Although it is common to talk about “stiff differential equations,” an equation per se is not stiff, a particular initial value problem for that equation may be stiff, in some regions, but the sizes of these regions depend on the initial values and the error tolerance. (Hairer und Wanner, 2010)

Beispiel 3.12. In Beispiel 3.1 auf Seite 62 wurde gezeigt, dass die langsame Mannigfaltigkeit des dahlquistschen Testproblems durch $\mu = \{0\}$ gegeben ist. In Beispiel 3.8 auf Seite 68 konnte man sehen, dass diese langsame Mannigfaltigkeit für alle Startwerte $x_0 \neq 0$ exponentiell anziehend ist. Die Steifheitsmenge S des dahlquistschen Testproblems ist demnach durch $S = \mathbb{R} \setminus \{0\}$ gegeben.

Beispiel 3.13. Für das Problem $\dot{x} = -1/\varepsilon x^2$ aus den Beispielen 3.5 auf Seite 64 und 3.9 auf Seite 69 existieren keine steifen Lösungen. Es wurde gezeigt, dass jede diskrete Teilmenge $M \subset \mathbb{R} \setminus \{0\}$ eine langsame Mannigfaltigkeit ist und dass es keine exponentiell anziehenden Mengen gibt. Deshalb existierten keine steifen Lösungen, die Steifheitsmenge ist leer. Das passt zu Abbildung 2.5 auf Seite 39, in der deutlich zu erkennen war, dass für $\varepsilon \rightarrow 0$ die Lösungskurven immer flacher verlaufen und in einer Umgebung um 0 damit fast keine Bewegung mehr in dem Sinne statt findet, dass die rechte Seite fast verschwindet.

Beispiel 3.14. Für den nichtlinearen Strudel aus den Beispielen 3.6 auf Seite 65 und 3.10 auf Seite 69 kann auch ohne Kenntnis der exakten Lösung die Steifheitsmenge angegeben werden.

Für $\alpha = \beta$ sind nach Beispiel 3.6 nur die diskreten Teilmengen des \mathbb{R}^2 langsame Mannigfaltigkeiten.

Für $\alpha = 1$ sind alle nulldimensionalen Mannigfaltigkeiten, die nicht auf einem der Kreise mit Radius $r = 1$ oder $r = 1/\sqrt{3}$ liegen, langsam, insbesondere der Nullpunkt.

Ist $\beta > 0$, so ist die S^1 nach Beispiel 3.10 für alle Startwerte x_0 mit $0 < u_0 = \|x_0\|^2 \neq 1$ eine exponentiell anziehende Menge. Alle Lösungen laufen in Spiralen zum Einheitskreis, insbesondere ist das Existenzintervall unbeschränkt. Für $u_0 = \|x_0\|^2 = 1$ erhält man periodische Lösungen und die S^1 ist nicht exponentiell anziehend.

Ist $\beta < 0$, so ist die Menge $M_0 = \{0\}$ exponentiell anziehend für Startwerte mit $0 < u_0 = \|x_0\|^2 < 1$. Für $u_0 = \|x_0\|^2 = 1$ erhält man wieder periodische nicht steife Lösungsfamilien. Für $u_0 = \|x_0\|^2 > 1$ verschwindet das Existenzintervall für $\varepsilon \rightarrow 0$, siehe Beispiel 3.10. Insbesondere ist M_0 nicht exponentiell anziehend.

Bezeichne $B_{\mathbb{R}^2} \setminus \{0\}$ den offenen Einheitsball des \mathbb{R}^2 erhält man zusammenfassend für die Steifheitsmenge:

	$\alpha = \beta$	$\alpha = 1$
$\beta > 0$	$S = \emptyset$	$S = \{x \in \mathbb{R}^2 : \ x\ \neq 0, 1\}$
$\beta < 0$	$S = B_{\mathbb{R}^2} \setminus \{0\}$	$S = B_{\mathbb{R}^2} \setminus \{0\}$

Wie man in der Tabelle sieht ist für $\alpha = \beta > 0$ die Steifheitsmenge leer. Die Bewegung auf dem Einheitskreis ist asymptotisch schnell, es findet kein abruptes „abbremsen“ statt und die Lösungen sind nicht steif.

Der Van-der-Pol-Oszillator aus Beispiel 3.7 auf Seite 66 ist weitaus schwieriger zu analysieren, insbesondere die exponentiell anziehenden Mengen gestalten sich kompliziert, was im Wesentlichen daran liegt, dass sich für den Van-der-Pol-Oszillator keine geschlossene analytische Lösung angeben lässt. Auf das steife Verhalten des Van-der-Pol-Oszillators wird deshalb erst später eingegangen.

3.4 STEIFE LÖSUNGSFAMILIEN LINEARER PROBLEME

Der Fall autonomer linearer Anfangswertprobleme bietet einen tiefen Einblick in die mathematische Struktur steifer Lösungen. Jedes nichtautonome System ist zwar äquivalent zu einem autonomen System, das aber im Allgemeinen nichtlinear ist. Der nichtlineare Fall ist deutlich schwerer zu behandeln und wird in dieser Arbeit auch nicht vollständig diskutiert. Der autonome lineare Fall verdient deshalb besonderes Augenmerk.

Für $A_\varepsilon \in \text{Mat}(n, \mathbb{R})$ betrachte man

$$\dot{x} = A_\varepsilon x, \quad x_0 = x(0). \quad (3.46)$$

Es gibt genau eine langsame Mannigfaltigkeit $\mu = \mu_\varepsilon$ und diese ist ein linearer Raum. Der Phasenfluss des linearen Problems ist bekanntlich durch $\Phi_\varepsilon^t = \exp(tA_\varepsilon)$ gegeben. Ist π_{μ^\perp} die orthogonale Projektion auf das orthogonale Komplement von μ , lässt sich Bedingung (3.43) aus Definition 3.7 auf Seite 71 in der Form

$$0 \neq \|\pi_{\mu^\perp} \Phi_\varepsilon^t x_0\| = \|\pi_{\mu^\perp} \exp(tA_\varepsilon) x_0\| = \mathcal{O}\left(\exp\left(-\frac{C}{\varepsilon}t\right)\right) \quad (3.47)$$

wie in (3.31) schreiben. An der orthogonalen Projektion π_{μ^\perp} wird deutlich, dass nur das exponentielle Abklingen gewisser Lösungskomponenten Relevanz hat.

In (3.47) wird explizit gefordert, dass die Norm positiv ist. Da μ nach Bemerkung 3.2 als Summe von Haupträumen invariant unter A_ε ist, ist es auch invariant unter $\exp(tA_\varepsilon)$. Insbesondere gilt

$$x_0 \in \mu \iff \|\pi_{\mu^\perp} \Phi_\varepsilon^t x_0\| = \|\pi_{\mu^\perp} \exp(tA_\varepsilon) x_0\| = 0. \quad (3.48)$$

Das bedeutet, dass die Lösung zu dem Anfangswert $x_0 \in \mu$ komplett in μ liegt. Es findet keine Relaxation auf μ statt. Hier ist es wichtig zu bemerken, dass eine numerische Lösung aufgrund von Rundungsfehlern die langsame Mannigfaltigkeit sehr wohl verlassen kann. In führender Ordnung ($\varepsilon = 0$) tritt dieses Phänomen analytisch nicht auf, da verlässt die Lösung die langsame Mannigfaltigkeit nicht. Nur in höheren Ordnungen kann es zu Abweichungen kommen.

In dem sehr einfachen Fall, dass das lineare System gerade durch eine Diagonalmatrix gegeben ist, lässt sich auch der Phasenfluss sehr einfach bestimmen. Das liegt daran, dass die Matrixexponentielle einer Diagonalmatrix wieder diagonal ist, wobei die Diagonaleinträge der ursprünglichen Matrix als Exponenten der Exponentialfunktion auftreten. Im Allgemeinen ist die Berechnung der Matrixexponentiellen jedoch nicht so einfach möglich.

Im Allgemeinen Fall wird die Matrixexponentielle üblicherweise über die Jordansche Normalform berechnet. Das ist hier jedoch keine geeignete Option, wie sich später zeigen wird. Es ist deshalb besonders wichtig, das Transformationsverhalten steifer Lösungen linearer Probleme zu verstehen.

Es folgen zwei Beispiele, die die Diskussion zu der Theorie dynamischer Systeme in Beispiel 3.11 auf Seite 70 am Ende von Abschnitt 3.2 noch einmal untermauern.

Beispiel 3.15. Für das System

$$\dot{x} = A_\varepsilon x = \begin{pmatrix} 1 & 0 \\ 0 & -\frac{1}{\varepsilon} \end{pmatrix} x, \quad (3.49)$$

ist $\mu_\varepsilon = \text{span}(e_1)$. In diesem einfachen Fall einer Diagonalmatrix ergibt sich für $0 \neq x_0 = (x_1, x_2)$

$$\text{dist}(\Phi_\varepsilon^t x_0, \mu_\varepsilon) = \left\| \pi_{\mu_\varepsilon^\perp} \Phi_\varepsilon^t x_0 \right\| \quad (3.50a)$$

$$= \left\| \pi_{\mu_\varepsilon^\perp} \exp(tA_\varepsilon) x_0 \right\| \quad (3.50b)$$

$$= \left\| x_2 \exp\left(-\frac{t}{\varepsilon}\right) \right\| \quad (3.50c)$$

$$= \mathcal{O}\left(\exp\left(-\frac{t}{\varepsilon}\right)\right), \quad (3.50d)$$

wobei $\text{dist}(\Phi_\varepsilon^t x_0, \mu_\varepsilon) \neq 0$ genau dann gilt, wenn $x_0 \notin \mu$. Für die Steifheitsmenge heißt das $S = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_2 \neq 0\}$.

Damit erhält man auch gleich ein Beispiel für ein Problem, dass steife Lösungen besitzt jedoch nicht asymptotisch stabil ist im Sinne dynamischer Systeme, da die Spektralabszisse von f_ε positiv ist (Strehmel, Weiner und Podhaisky, 2012, S. 192).

Beispiel 3.16. Modifiziert man Beispiel 3.15 dahingehend, dass auch die erste Komponente der rechten Seite asymptotisch groß ist

$$\dot{x} = A_\varepsilon x = \begin{pmatrix} \frac{1}{\varepsilon} & 0 \\ 0 & -\frac{1}{\varepsilon} \end{pmatrix} x, \quad (3.51)$$

dann gilt $\mu_\varepsilon = \{0\}$, $\pi_{\mu^\perp} = \text{Id}$ und

$$\text{dist}(\Phi_\varepsilon^t x_0, \mu) = \|\pi_{\mu^\perp} \Phi_\varepsilon^t x_0\| \quad (3.52a)$$

$$= \|\pi_{\mu^\perp} \exp(tA_\varepsilon) x_0\| \quad (3.52b)$$

$$= \left\| \begin{pmatrix} x_1 \exp\left(\frac{t}{\varepsilon}\right) \\ x_2 \exp\left(-\frac{t}{\varepsilon}\right) \end{pmatrix} \right\|. \quad (3.52c)$$

Die Steifheitsmenge ist demnach durch

$$S = \{0 \neq x_0 = (x_1, x_2) \in \mathbb{R}^2 \mid x_1 = 0\} \quad (3.53)$$

gegeben.

Zur Veranschaulichung der letzten beiden Beispiele betrachte man in Abbildung 3.1 auf der nächsten Seite die Phasenportraits. Die rote Gerade in Abbildung 3.1(a) kennzeichnet die langsame Mannigfaltigkeit aus Beispiel 3.15. Für $\varepsilon = 1$ verläuft das Phasenportrait in einer Umgebung um $x_2 = 0$ nahezu parallel zur roten Linie. Für $\varepsilon = 0.001$ hingegen ist es fast orthogonal und der Relaxationsprozeß auf μ ist deutlich erkennbar.

In Beispiel 3.16 ist die langsame Mannigfaltigkeit durch $\mu_\varepsilon = \{0\}$ gegeben. Die beiden Phasenportraits für $\varepsilon = 1$ und $\varepsilon = 0.001$ in Abbildung 3.1(b) sehen beinahe identisch aus. Es ist zu erkennen, dass quasi nur für $x_1 = 0$ eine schnelle Relaxation auf μ stattfindet. Diese Beobachtung passt zu der Steifheitsmenge.

3.4.1 Invarianz unitärer Transformationen

Die Darstellung einer linearen Abbildung als Matrix hängt von der Wahl einer Basis ab. Eine geeignete Basis kann damit die Betrachtung eines linearen Problems erheblich erleichtern. Aus diesem Grund wird im Folgenden auf das Transformationsverhalten steifer Lösungen eingegangen. Die Definition 3.7 steifer Lösungsfamilien ist invariant unter der Gruppe der unitären Transformationen:

Sei $S_\varepsilon \in \text{U}(n)$ für alle $\varepsilon > 0$ eine unitäre Matrix. Zur besseren Unterscheidung werden wie in Deuffhard und Bornemann (2008) Bild und Urbildraum verschieden bezeichnet. Es wird deshalb S_ε als Familie von Abbildungen zwischen zwei zu \mathbb{R}^n isomorphen Räumen angesehen,

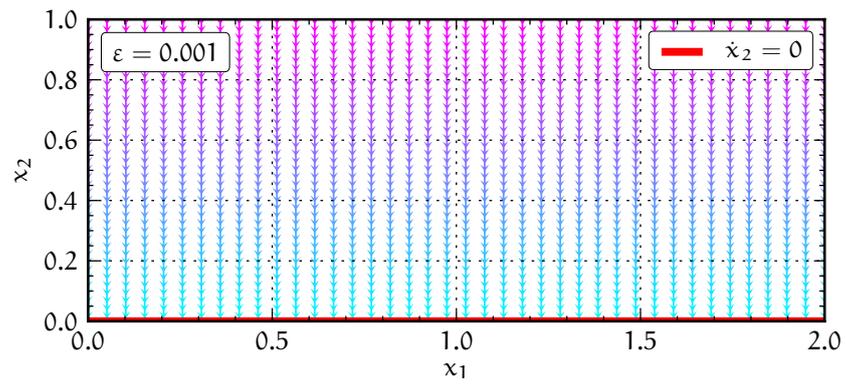
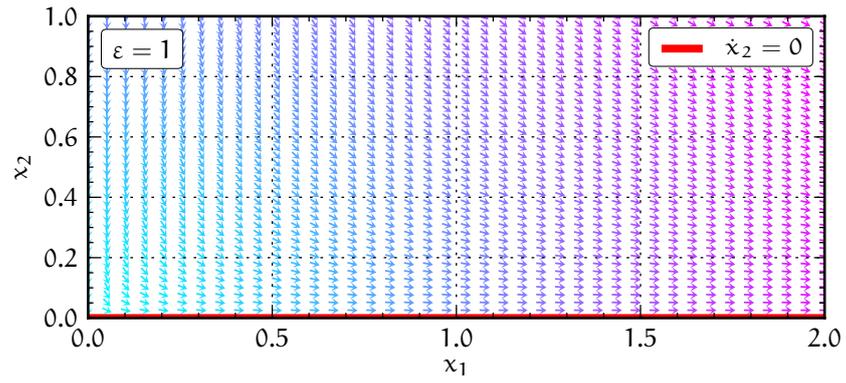
$$S_\varepsilon : X \rightarrow \hat{X}. \quad (3.54)$$

Die Abbildung transformiert die Familie von Phasenflüssen Φ_ε^t zu

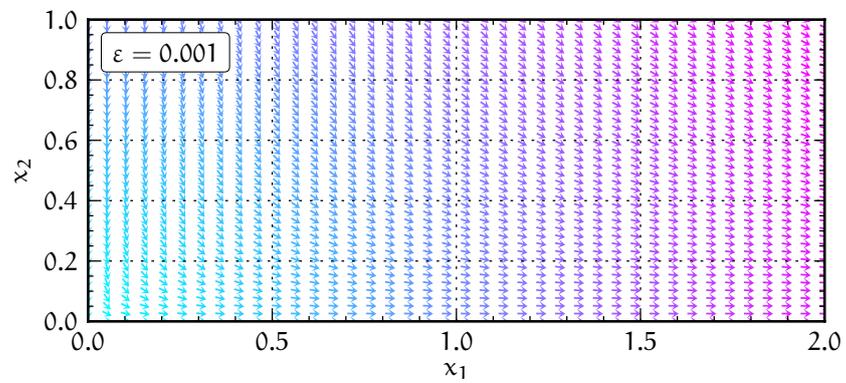
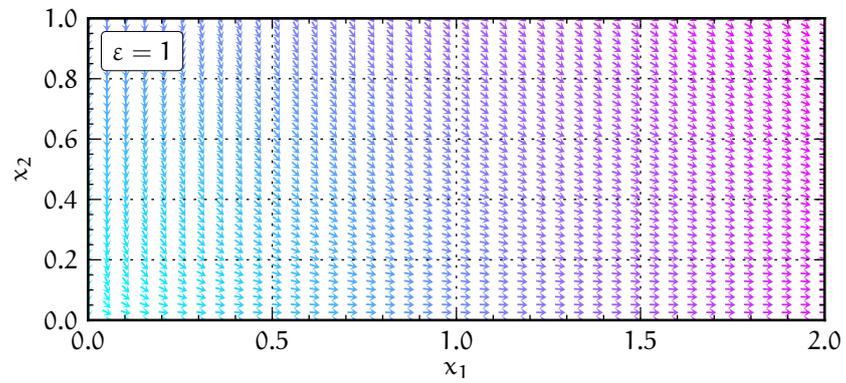
$$\hat{\Phi}_\varepsilon^t \hat{x} = S_\varepsilon \Phi_\varepsilon^t S_\varepsilon^{-1} \hat{x}, \quad (3.55)$$

einer Familie von Abbildungen auf dem transformierten Phasenraum

$$\hat{\Omega} = \{S_\varepsilon x \mid x \in \Omega\}. \quad (3.56)$$



(a) Beispiel 3.15.



(b) Beispiel 3.16.

Abbildung 3.1

Der folgende Satz wird nützlich dafür sein, einen geeigneten Repräsentanten des linearen Problems für die Analyse auf steifes Verhalten zu finden.

Satz 3.1. Sei $S_\varepsilon \in U(n)$ eine unitäre Familie von Abbildungen. Die Lösung in Ω durch x_0 ist genau dann steif, wenn es ihr Bild in $\hat{\Omega}$ durch $\hat{x}_0 = S_\varepsilon x_0$ ist.

Beweis. Sei die durch x_0 gehende Lösung steif und $S_\varepsilon \in U(n)$ eine unitäre Transformation. Ist μ die zu A_ε gehörige langsame Mannigfaltigkeit, dann ist wegen der Invarianz der Haupträume unter Transformation $\hat{\mu} = S_\varepsilon \mu$ die zu dem transformierten Problem gehörige langsame Mannigfaltigkeit. Da S_ε unitär ist, gilt weiter

$$\hat{\mu}^\perp = (S_\varepsilon \mu)^\perp = S_\varepsilon \mu^\perp. \quad (3.57)$$

Damit ist die durch $\hat{\pi}_{\hat{\mu}^\perp} = S_\varepsilon \pi_{\mu^\perp} S_\varepsilon^{-1}$ definierte Abbildung genau dann die orthogonale Projektion auf $\hat{\mu}^\perp$, wenn π_{μ^\perp} die orthogonale Projektion auf μ^\perp ist. Da die Lösung durch x_0 steif ist, gilt

$$\text{dist}(\hat{\Phi}_\varepsilon^t x_0, \hat{\mu}) = \|\hat{\pi}_{\hat{\mu}^\perp} \hat{\Phi}_\varepsilon^t x_0\| \quad (3.58a)$$

$$= \|S_\varepsilon \pi_{\mu^\perp} \Phi_\varepsilon^t x_0\| \quad (3.58b)$$

$$= \|\pi_{\mu^\perp} \Phi_\varepsilon^t x_0\| \quad (3.58c)$$

$$= \mathcal{O}\left(\exp\left(-\frac{C}{\varepsilon} t\right)\right). \quad (3.58d)$$

Damit ist die durch $\hat{x}_0 = S_\varepsilon x_0$ gehende Lösung des transformierten Problems $\hat{\dot{x}} = S_\varepsilon A_\varepsilon S_\varepsilon^{-1} \hat{x}$ steif. Die Umkehrung folgt da mit S_ε auch S_ε^{-1} unitär ist. \square

3.4.2 Steife Lösungen von Schurformen

Satz 3.1 rechtfertigt, die Untersuchung steifer Lösungen im linearen Fall auf spezielle Matrizen zu beschränken. Jede Matrix ist unter einer unitären Transformation ähnlich zu einer oberen Dreiecksmatrix. Da steife Lösungskurven invariant unter unitären Transformationen sind, genügt es, sich auf die Klasse dieser Matrizen zu beschränken.

Damit ist auch klar, dass die Jordansche Normalform bei der Untersuchung steifer Lösungen linearer Probleme im Allgemeinen kein geeigneter Repräsentant ist. Die Jordansche Normalform ist zwar ein Repräsentant der Äquivalenzklasse der ähnlichen Matrizen, nicht jedoch der unitär-ähnlichen Matrizen.

Beispiel 3.17. Betrachte man die beiden Matrizen

$$A_\varepsilon = \begin{pmatrix} 1/\varepsilon & 1/\varepsilon \\ 0 & -1/\varepsilon \end{pmatrix} \quad \text{und} \quad B_\varepsilon = \begin{pmatrix} 1/\varepsilon & 1 \\ 0 & -1/\varepsilon \end{pmatrix} \quad (3.59)$$

und die dazugehörigen linearen Anfangswertprobleme

$$\dot{x} = A_\varepsilon x, \quad x(0) = x_0, \quad \text{und} \quad \dot{y} = B_\varepsilon y, \quad y(0) = y_0. \quad (3.60)$$

Da A_ε und B_ε die gleichen Eigenwerte haben, liegen sie in der selben Ähnlichkeitsklasse. Wegen $\operatorname{tr}(A_\varepsilon A_\varepsilon^*) \neq \operatorname{tr}(B_\varepsilon B_\varepsilon^*)$ sind sie aber nicht unitär-ähnlich, dabei bezeichne $\operatorname{tr}(A_\varepsilon)$ die Spur von A_ε und A_ε^* die Adjungierte zu A_ε . Mit nicht unitärem

$$S = \begin{pmatrix} 1 & \varepsilon/2 - 1/2 \\ 0 & 1 \end{pmatrix} \quad (3.61)$$

gilt

$$A_\varepsilon = S_\varepsilon B_\varepsilon S_\varepsilon^{-1}. \quad (3.62)$$

Sei $x_0 = (1, -2)^T$ und $\mu_\varepsilon = \{0\}$ die langsame Mannigfaltigkeit von A_ε , dann erhält man

$$\left\| \pi_{\mu_\varepsilon^\perp} \exp(tA_\varepsilon) x_0 \right\| = \mathcal{O} \left(\exp \left(-\frac{t}{\varepsilon} \right) \right) \quad (3.63)$$

und somit eine steife Lösung. Der zu x_0 gehörige Anfangswert des transformierten Problems ist $y_0 = (\varepsilon, -2)^T$. Dieser ist nicht unabhängig von ε und das transformierte Problem liefert nur für Anfangswerte $y_0 = (0, y_2)$ mit $y_2 \neq 0$ steife Lösungen.

Nichtunitäre Transformationen sind also ungeeignet dafür, Familien von Anfangswertproblemen mit ε -unabhängigem Anfangswert zu studieren.

Es ist demnach nötig, einen Repräsentanten der Äquivalenzklassen für unitär-ähnliche Matrizen zu bestimmen. Hierfür bietet sich die Schurform oder Schur-Zerlegung an. Über einem algebraisch abgeschlossenen Körper gibt es zu jeder Matrix A eine unitäre Matrix U mit der zu U adjungierten Matrix U^* , so dass

$$R = U^* A U \quad (3.64)$$

eine obere Dreiecksmatrix der Form $R = D + N$ ist. Hierbei ist D eine Diagonalmatrix, die bis auf die Reihenfolge der Einträge eindeutig ist, und N ist nilpotent. Im Allgemeinen ist N nur bezüglich der Frobeniusnorm eindeutig.

Zur Wahl eines geeigneten Repräsentanten werden nun Schurzerlegungen betrachtet, in denen die Eigenwerte auf der Diagonalen nach asymptotischer Größenordnung und Vorzeichen sortiert sind.

Satz 3.2. Sei A_ε eine Schurform der Gestalt

$$A_\varepsilon = \begin{pmatrix} U_1 & N \\ 0 & U_s \end{pmatrix} \quad \text{mit} \quad U_s = \begin{pmatrix} U_s^+ & B \\ 0 & U_s^- \end{pmatrix}. \quad (3.65)$$

Dabei seien $U_1 \in \operatorname{Mat}(\mathbb{C}; r)$, $U_s \in \operatorname{Mat}(\mathbb{C}; t)$, $U_s^+ \in \operatorname{Mat}(\mathbb{C}; q)$ und $U_s^- \in \operatorname{Mat}(\mathbb{C}; t - q)$ obere Dreiecksmatrizen und N und B nilpotente Anteile. Die Eigenwerte von U_1 seien aus $\sigma_1(A_\varepsilon)$, die von U_s aus $\sigma_{1/\varepsilon}(A_\varepsilon)$,

wobei U_s^+ Eigenwerte mit positiven und U_s^- mit negativen Realteilen habe. Die zu A_ε passende Basis sei

$$\mathcal{B} = (l_1, \dots, l_r, s_{r+1}, \dots, s_q, s_{r+q+1}, \dots, s_t). \quad (3.66)$$

Dann gilt

$$\mu_\varepsilon = \text{span}(l_1, \dots, l_r) \quad (3.67a)$$

und

$$\mu_\varepsilon^\perp = \mu_\varepsilon^{\perp+} \oplus \mu_\varepsilon^{\perp-} \quad (3.67b)$$

$$= \text{span}(s_{r+1}, \dots, s_{r+q}, s_{r+q+1}, \dots, s_t), \quad (3.67c)$$

mit

$$\mu_\varepsilon^{\perp+} = \text{span}(s_{r+1}, \dots, s_{r+q}) \quad (3.67d)$$

und

$$\mu_\varepsilon^{\perp-} = \text{span}(s_{r+q+1}, \dots, s_t). \quad (3.67e)$$

Beweis. Sei $\lambda \in \sigma_1(A_\varepsilon)$ und $v \in \text{Hau}(A_\varepsilon; \lambda)$, dann gibt es nach Definition 3.4 von μ_ε ein k mit $(A_\varepsilon - \lambda)^k v = 0$. In der Basis \mathcal{B} kann v in der Form $v = (v_l, v_s)^T$ geschrieben werden, wobei v_l für die ersten r Komponenten, also die Koeffizienten der Basisvektoren l_1, \dots, l_r und v_s für die restlichen Komponenten steht. Dann ist

$$0 = (A_\varepsilon - \lambda)^k v = \begin{pmatrix} U_l - \lambda & N \\ 0 & U_s - \lambda \end{pmatrix}^k \begin{pmatrix} v_l \\ v_s \end{pmatrix} \quad (3.68a)$$

$$= \begin{pmatrix} (U_l - \lambda)^k & * \\ 0 & (U_s - \lambda)^k \end{pmatrix} \begin{pmatrix} v_l \\ v_s \end{pmatrix}. \quad (3.68b)$$

Insbesondere ist $(U_s - \lambda)^k v_s = 0$. Einerseits ist $\lambda \in \sigma_1(A_\varepsilon)$, andererseits gilt wenn $\tilde{\lambda}$ ein Eigenwert von U_s ist $\tilde{\lambda} \in \sigma_{1/\varepsilon}(A_\varepsilon)$. Damit sind $U_s - \lambda$ und $(U_s - \lambda)^k$ regulär und es folgt $v_s = 0$. Da $\lambda \in \sigma_1(A_\varepsilon)$ beliebig gewählt war und μ_ε direkte Summe der Haupträume dieser Eigenwerte ist, ergibt das $\mu_\varepsilon \subset \text{span}(l_1, \dots, l_r)$. Mit $\dim \mu_\varepsilon = r$ gilt sogar Gleichheit. \square

Korollar 3.1. Nach der Konstruktion der Schurzerlegung aus Satz 3.2 gilt für die orthogonale Projektion auf μ_ε^\perp

$$\pi_{\mu_\varepsilon^\perp} \begin{pmatrix} U_l & N \\ 0 & U_s \end{pmatrix} = \begin{pmatrix} 0 & U_s \end{pmatrix}. \quad (3.69)$$

Es werden im Folgenden nur noch Repräsentanten der Äquivalenzklassen der unitär-ähnlichen Matrizen betrachtet, die von der Gestalt (3.65) sind.

Ein wesentlicher Schritt beim Finden steifer Lösungsfamilien ist die Auswertung von $\|\pi_{\mu_\varepsilon} \exp(tA_\varepsilon) x_0\|$. Zur Berechnung der Matrixexponentiellen einer Matrix A_ε betrachtet man üblicherweise die Jordansche Normalform von A_ε . Wie bereits erwähnt ist diese aber für die Untersuchung auf steifes Verhalten ungeeignet, da sie im Allgemeinen nicht durch einen unitären Basiswechsel erreicht werden kann. Der folgende Satz hilft weiter.

Satz 3.3 (Davis (1973, S. 5)). *Sei*

$$A = \begin{pmatrix} a_1 & b_{12} & \cdots & b_{1n} \\ 0 & a_2 & \cdots & \vdots \\ 0 & 0 & \ddots & b_{n-1,n} \\ 0 & 0 & \cdots & a_n \end{pmatrix} \quad (3.70)$$

und f ein Polynom, dann ist

$$f(A) = \begin{pmatrix} f(a_1) & c_{12} & \cdots & c_{1n} \\ 0 & f(a_2) & \cdots & \vdots \\ 0 & 0 & \ddots & c_{n-1,n} \\ 0 & 0 & \cdots & f(a_n) \end{pmatrix} \quad (3.71)$$

mit

$$c_{ij} = \sum f^{[m]}(a_{k_0}, \dots, a_{k_m}) b_{k_0 k_1} b_{k_1 k_2} \cdots b_{k_{m-1} k_m}, \quad (3.72)$$

für $1 \leq i < j \leq n$, wobei die Summe über alle verschiedenen Tupel (k_0, \dots, k_m) gebildet wird, die

$$i = k_0 < k_1 < \cdots < k_m = j, \quad 1 \leq m \leq j - i \quad (3.73)$$

erfüllen. Hierbei sind die geteilten Differenzen $f^{[m]}$ für Polynome nach Davis (ebd.) wie folgt definiert. Für ein Polynom f in s Unbestimmten ist $f^{[1]}$ ein Polynom in $s + 1$ Unbestimmten, das für $m_1 \neq m_2$ durch

$$\begin{aligned} f^{[1]}(\lambda_1, \dots, \lambda_{s-1}, m_1, m_2) \\ = \frac{f(\lambda_1, \dots, \lambda_{s-1}, m_1) - f(\lambda_1, \dots, \lambda_{s-1}, m_2)}{m_1 - m_2} \end{aligned} \quad (3.74)$$

gegeben ist. Für ein Polynom in einer Unbestimmten ist $f^{[0]} = f$ und induktiv $f^{[s]} = (f^{[s-1]})^{[1]}$.

Für $m_1 = m_2$ ist nach Higham (2008)

$$f^{[1]}(\lambda_1, \dots, \lambda_{s-1}, m_1, m_1) = f'(\lambda_1, \dots, \lambda_{s-1}, m_1), \quad (3.75)$$

wobei f' die Ableitung von f nach m_1 ist.

Bemerkung 3.5. Anschaulich gesprochen sagt Bedingung (3.73), dass zur Berechnung von c_{ij} nur jene Einträge aus der Matrix A herangezogen werden, die in dem Dreieck (a_i, c_{i-1j-1}, a_j) liegen.

Zu bemerken ist, dass Satz 3.3 nur für Polynome formuliert ist. Die Aussage lässt sich aber auf die Matrixexponentielle übertragen:

[...] and f a function. For well-known reasons, we may assume (as long as only finitely many matrices are in consideration as arguments at once) that each function f considered is a polynomial. (Davis, 1973, S. 1).

Dabei steht „function“ für eine Funktion, die durch eine Potenzreihe entwickelbar ist. Die folgende Argumentation ist Higham (2008, Remark 1.8) entnommen. Betrachtet wird zu einer $(n \times n)$ -Matrix A das dazugehörige charakteristische Polynom $q(t) = \det(tI - A)$. Nach dem Satz von Cayley-Hamilton ist $q(A) = 0$. Damit kann jede Potenz A^x mit $x \geq n$ als Linearkombination von $I, A, A^2, \dots, A^{n-1}$ geschrieben werden. Jede Potenzreihe in A kann also zu einem Polynom in A von höchstens Grad $n - 1$ reduziert werden. Es gibt somit ein Polynom P mit $\exp(tA) = P(tA)$. Das rechtfertigt die Annahme, dass f ein Polynom ist und Satz 3.3 gilt auch für $f = \exp$. In diesem Fall haben nur die Diagonalelemente a_i einen exponentiellen Einfluss auf die Koeffizienten von $\exp(A)$ wie das nächste Korollar zeigt.

Korollar 3.2. Sind in Satz 3.3 die a_i und b_j rationale Funktionen in ε und ist weiter $f = \exp$ dann sind die c_{ij} von der Form

$$c_{ij} = \sum_{l=0}^m \exp(a_{k_l}) r_l, \quad 1 \leq i < j \leq n, \quad i \leq k_l \leq j, \quad (3.76)$$

für in ε rationale Funktionen r_l .

Beweis. Für den Beweis wird die Struktur von $f^{[m]}$ betrachtet. Mittels Induktion über m wird gezeigt, dass für in ε rationale Funktionen \tilde{r}_{l_m}

$$\exp^{[m]}(a_{k_0}, \dots, a_{k_m}) = \sum_{l=0}^m \exp(a_{k_l}) \tilde{r}_{l_m} \quad (3.77)$$

gilt. Es wird nur der Fall für $m_1 \neq m_2$ aus (3.74) betrachtet, der Fall $m_1 = m_2$ aus (3.75) folgt analog.

Für $m = 1$ ist die Behauptung mit

$$\begin{aligned} & \exp^{[1]}(a_{k_0}, a_{k_1}) \\ &= \exp(a_{k_0}) \frac{1}{a_{k_0} - a_{k_1}} - \exp(a_{k_1}) \frac{1}{a_{k_0} - a_{k_1}} \end{aligned} \quad (3.78)$$

richtig. Nach der induktiven Definition von $f^{[m+1]} = (f^{[m]})^{[1]}$ und der Induktionsvoraussetzung gilt

$$\exp^{[m+1]}(a_{k_0}, \dots, a_{k_{m-1}}, a_{k_m}, a_{k_{m+1}}) \quad (3.79a)$$

$$= \frac{1}{a_{k_m} - a_{k_{m+1}}} \exp^{[m]}(a_{k_0}, \dots, a_{k_{m-1}}, a_{k_m}) \quad (3.79b)$$

$$- \frac{1}{a_{k_m} - a_{k_{m+1}}} \exp^{[m]}(a_{k_0}, \dots, a_{k_{m-1}}, a_{k_{m+1}}) \quad (3.79c)$$

$$= \sum_{l=0}^{m+1} \exp(a_{k_l}) \tilde{r}_{l_{m+1}}. \quad (3.79d)$$

Berücksichtigt man, dass sich die r_l aus (3.76) und die \tilde{r}_{l_m} aus (3.77) durch geeignete Produkte der b_j nach der Definition der c_{ij} aus (3.72) unterscheiden, folgt die Behauptung. \square

Bemerkung 3.6. Satz 3.3 gilt noch etwas allgemeiner, siehe Davis (1973, Theorem 2). Etwas verkürzt dargestellt sagt das Theorem folgendes. Sei A eine Blockdiagonalmatrix

$$A = \begin{pmatrix} A_1 & \cdots & * \\ 0 & \ddots & \vdots \\ 0 & 0 & A_n \end{pmatrix} \quad (3.80)$$

mit Untermatrizen A_1, \dots, A_n und f wie in Satz 3.3, dann gilt

$$f(A) = \begin{pmatrix} f(A_1) & \cdots & * \\ 0 & \ddots & \vdots \\ 0 & 0 & f(A_n) \end{pmatrix}. \quad (3.81)$$

Die Sterne in A und $f(A)$ stehen dabei für verschiedene Einträge.

Lemma 3.1. Sei f eine auf dem Ring der Matrizen definierte Funktion, A_ε wie in Satz 3.2 und a_i und c_{ij} wie in Satz 3.3, dann gilt für die 1-Norm

$$\|\pi_{\mu^\perp} f(tA_\varepsilon) x_0\|_1 \quad (3.82a)$$

$$= \|f(tU_s) \pi_{\mu^\perp} x_0\|_1 \quad (3.82b)$$

$$= \sum_{i=1}^s \left| f(ta_i)(x_0)_{r+i} + \sum_{j=i+1}^s c_{ij} (x_0)_{r+j} \right|. \quad (3.82c)$$

Beweis. Nach Korollar 3.1, Satz 3.3 und Bemerkung 3.6 gilt:

$$\|\pi_{\mu^\perp} f(tA_\varepsilon)x_0\|_1 \quad (3.83a)$$

$$= \left\| \pi_{\mu^\perp} f \left(t \begin{pmatrix} U_l & * \\ 0 & U_s \end{pmatrix} \right) x_0 \right\|_1 \quad (3.83b)$$

$$= \left\| \pi_{\mu^\perp} \begin{pmatrix} f(tU_l) & * \\ 0 & f(tU_s) \end{pmatrix} x_0 \right\|_1 \quad (3.83c)$$

$$= \|(0 \ f(tU_s)) x_0\|_1 \quad (3.83d)$$

$$= \|f(tU_s) \pi_{\mu^\perp} x_0\|_1 \quad (3.83e)$$

$$= \sum_{i=1}^t |(f(tU_s) \pi_{\mu^\perp} x_0) \cdot \mathbf{e}_{i+r}| \quad (3.83f)$$

$$= \sum_{i=1}^t \left| \sum_{j=i}^t (f(tU_s))_{ij} (x_0)_{r+j} \right| \quad (3.83g)$$

$$= \sum_{i=1}^t \left| f(ta_i)(x_0)_{r+i} \sum_{j=i+1}^t c_{ij} (x_0)_{r+j} \right|. \quad (3.83h)$$

□

Damit sind die Vorbereitungen getroffen, um steife Lösungsfamilien zu identifizieren.

Satz 3.4. Sei A_ε wie in Satz 3.2 mit $B = 0$ und die a_i und c_{ij} wie in Satz 3.3, dann ist die Steifheitsmenge durch $S = \text{span}(\mu \cup \mu^{\perp-}) \setminus \mu$ gegeben. Insbesondere sind die Lösungsfamilien für jedes $x_0 \notin \mu$ steif, wenn $\mu^{\perp+} = \{0\}$ gilt. Anders ausgedrückt ist die Lösungsfamilie durch $x_0 \notin \mu$ steif, falls A_ε Eigenwerte mit asymptotisch großen Realteilen hat, die alle negativ sind.

Beweis. Sei A_ε von der geforderten Gestalt. Nach Lemma 3.1 gilt dann

$$\|\pi_{\mu^\perp} \exp(tA_\varepsilon)x_0\|_1 \quad (3.84a)$$

$$= \sum_{i=r+1}^t \left| \exp(ta_i)(x_0)_i + \sum_{j=i+1}^t c_{ij}(x_0)_j \right| \quad (3.84b)$$

$$= \sum_{i=r+1}^{r+q} \left| \exp(ta_i)(x_0)_i \right. \quad (3.84c)$$

$$\left. + \sum_{j=i+1}^{r+q} c_{ij}(x_0)_j + \sum_{j=r+q+1}^t c_{ij}(x_0)_j \right|$$

$$+ \sum_{i=r+q+1}^{r+t} \left| \exp(ta_i)x_{0,i} + \sum_{j=i+1}^{r+t} c_{ij}(x_0)_j \right|. \quad (3.84d)$$

Wegen $B = 0$ ist $c_{ij} = 0$ für alle $r+1 \leq i \leq r+q$ und $r+q+1 \leq j \leq t$. Ist $x_0 \in S$, dann ist $\pi_{\mu^\perp} x_0 = 0$ und damit ist auch $(x_0)_i = 0$ für alle $r+1 \leq i \leq r+q$. Die Zeile (3.84c) verschwindet deshalb komplett. Nach Konstruktion in (3.65) ist $\operatorname{Re} \alpha_i < 0$ für alle $r+q+1 \leq i \leq r+t$. Für $r+q+1 \leq i \leq r+t$ und $i \leq j \leq r+t$ hängen die c_{ij} nach Korollar 3.2 von denjenigen α_k ab, für die $i \leq k \leq j$ gilt. Für diese c_{ij} gilt nach Korollar 3.2 $c_{ij} = \mathcal{O}(\exp(-t/\varepsilon))$. Zudem ist mit $\pi_{\mu^\perp} x_0 \neq 0$ Gleichung (3.84b) von 0 verschieden. Insgesamt ergibt das:

$$\|\pi_{\mu^\perp} \exp(tA_\varepsilon) x_0\| = \mathcal{O}\left(\exp\left(-\frac{t}{\varepsilon}\right)\right). \quad (3.85)$$

Damit ist die Lösungsfamilie durch x_0 steif.

Sei andersherum $x \in \mathbb{R}^n$ mit $x \notin \mu_\varepsilon$ und $x \notin S$, dann gibt es einen Index i , $r+1 \leq i \leq q$, mit $(x_0)_i \neq 0$. Sei k der größte solche Index, dann sorgt der Summand $\exp(t\alpha_k) (x_0)_k$ in (3.84b) für asymptotisch exponentielles Wachstum und die Lösungsfamilie durch x_0 ist nicht steif. \square

Satz 3.4 liefert ein hinreichendes Kriterium für steife Lösungsfamilien durch x_0 . Notwendig ist dieses Kriterium aber nicht, wie das folgende Beispiel zeigt.

Beispiel 3.18. Gegeben sei das Anfangswertproblem

$$\dot{x} = A_\varepsilon x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\varepsilon} & -\frac{2}{\varepsilon} \\ 0 & 0 & -\frac{1}{\varepsilon} \end{pmatrix} x, \quad x_0 = x(0) = (1, 1, 1)^\top. \quad (3.86)$$

Für die Matrixexponentielle gilt dann

$$\exp(tA_\varepsilon) = \begin{pmatrix} e^t & 0 & 0 \\ 0 & e^{t/\varepsilon} & -e^{-t/\varepsilon}(-1 + e^{2t/\varepsilon}) \\ 0 & 0 & e^{-t/\varepsilon} \end{pmatrix}. \quad (3.87)$$

Für die langsame Mannigfaltigkeit ergibt das nach Satz 3.2

$$\mu_\varepsilon = \operatorname{span}(e_1), \quad \mu_\varepsilon^\perp = \operatorname{span}(e_2, e_3), \quad (3.88a)$$

$$\mu_\varepsilon^{\perp+} = \operatorname{span}(e_2), \quad \mu_\varepsilon^{\perp-} = \operatorname{span}(e_3). \quad (3.88b)$$

Einerseits ist nun $\pi_{\mu_\varepsilon^{\perp+}} x_0 = 1 \neq 0$. Andererseits ist die Lösungsfamilie durch x_0 steif, denn es gilt

$$0 \neq \|\pi_{\mu_\varepsilon^\perp} \exp(tA_\varepsilon) x_0\| = \mathcal{O}\left(\exp\left(-\frac{t}{\varepsilon}\right)\right). \quad (3.89)$$

Grund hierfür ist eine Auslöschung der exponentiell großen Anteile in der 1-Norm der zweiten Komponenten von $\pi_{\mu^\perp} \exp(tA_\varepsilon) x_0$. Diese Auslöschung tritt genau dann auf, wenn für $x_0 = (x_1, x_2, x_3)^\top$ die Gleichheit $x_2 = x_3$ gilt. Die Steifheitsmenge ist demnach durch $S = \{(x_1, x_2, x_3)^\top \notin \mu \mid x_2 = x_3\}$ gegeben.

Im Allgemeinen gibt es im Fall $\mu^{\perp+} \neq \{0\}$ keine steifen Lösungsfamilien. Unter gewissen Bedingungen an den nilpotenten Anteil existieren solche steifen Lösungsfamilien dennoch.

Bemerkung 3.7. Sei U_s aus Satz 3.2 von der Gestalt

$$U_s = \begin{pmatrix} D_s^+ & B \\ 0 & D_s^- \end{pmatrix} \quad (3.90)$$

mit Diagonalmatrizen $D_s^+ \neq 0$ und D_s^- mit den positiven bzw. negativen Eigenwerten und einem nilpotenten Anteil $B = (b_{ij})$. Sei weiter $b_{ij} = \mathcal{O}_S(1/\varepsilon)$ für alle i, j , dann gibt es steife Lösungsfamilien.

Beweis. Der Beweis folgt, wie der Beweis von Satz 3.4, wieder durch genaues Auswerten der b_{ij} nach Satz 3.3. \square

Zum Abschluss wird noch zu einigen bereits bekannten Beispielen die Steifheitsmenge angegeben. Die Aussage folgt jeweils aus Satz 3.4.

Beispiel 3.19. Die Steifheitsmenge des Problems aus Beispiel 3.3 auf Seite 63 ist durch $S = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_3 \neq 0\}$ gegeben.

Es sei nochmals betont, dass für steifes Verhalten nur das exponentielle Abklingen auf die langsame Mannigfaltigkeit von Bedeutung ist. Die komplette Lösung kann sehr wohl explodieren, solange das Explodieren nicht exponentiell ist. Die exakte Lösung des Problems aus Beispiel 3.3 ist durch

$$x(t) = \begin{pmatrix} \exp(t) x_1 + t/\varepsilon \exp(t) x_2 \\ \exp(t) x_2 \\ \exp(-t/\varepsilon) x_3 \end{pmatrix} \quad (3.91)$$

gegeben. Für $x_3 \neq 0$ erkennt man sofort das exponentielle Abklingen für $\varepsilon \rightarrow 0$, das wesentlich für das steife Verhalten ist. Ist zudem $x_2 \neq 0$ explodiert die Lösung für $\varepsilon \rightarrow 0$, jedoch nicht exponentiell.

An diesem Beispiel zeigt sich der mathematische Charakter steifer Lösungsfamilien besonders deutlich.

Beispiel 3.20. Für die Differentialgleichung aus Beispiel 3.16 auf Seite 74 ist die Steifheitsmenge durch $S = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_2 \neq 0\}$ gegeben.

Beispiel 3.21. Mit $\mu = \mathbb{R}^2$ ist die Steifheitsmenge des oszillatorischen Problems aus Beispiel 3.4 auf Seite 63 leer.

3.4.3 Normale und diagonale Probleme

In der Literatur werden nicht selten nur Diagonalmatrizen betrachtet, um steifes Verhalten zu untersuchen, deshalb legen Higham und Trefethen (1993) ein großes Augenmerk auf den Übergang von einem

linear konstanten Problem zu skalaren Problemen durch Diagonalisierung. Häufig wird allerdings stillschweigend davon ausgegangen, dass es ausreichend ist, sich auf Diagonalmatrizen und somit auf skalare Probleme vom Typ des dahlquistischen Testproblems zu beschränken, und dass die größte Herausforderung bei den nichtlinearen Problemen liegt. Dem ist aber nicht so, wie das nächste Zitat unterstreichen soll.

We wish to argue that the diagonalizations [...] is also a process that may change the nature of an ODE significantly, and in fact, that some effects which are customarily attributed to linearization or freezing of coefficients can with greater justice be blamed on diagonalization. (Higham und Trefethen, 1993, S. 287)

Normale Matrizen lassen sich unitär diagonalisieren und da nach Satz 3.1 steifes Verhalten invariant unter unitären Transformationen ist, genügt es tatsächlich Diagonalmatrizen zu betrachten. Für nichtnormale Matrizen gilt das aber im Allgemeinen nicht, sie müssen sich nicht einmal unitär in Jordansche Normalform bringen lassen.

Wie in Abschnitt 3.4.2 gesehen, muss steifes Verhalten von Lösungen zu einem fest gewählten Anfangswert für nichtunitäre Transformationen nicht erhalten bleiben. Deshalb wurden in Satz 3.2 Schurformen der Gestalt

$$A_\varepsilon = \begin{pmatrix} U_l & N \\ 0 & U_s \end{pmatrix}, \quad U_s = \begin{pmatrix} U_s^+ & B \\ 0 & U_s^- \end{pmatrix}, \quad (3.92)$$

mit oberen Dreiecksmatrizen U_l , U_s , U_s^+ und U_s^- und nilpotenten Anteilen N und B betrachtet. Solche Schurformen sind Repräsentanten der Klasse der unitär-ähnlichen Matrizen. Die Realteile der Eigenwerte von U_l sind dabei von der Größenordnung $\mathcal{O}(1)$, während die von U_s asymptotisch groß sind und positives respektive negatives Vorzeichen haben. Für normale Matrizen sind U_l und U_s ebenfalls diagonal und N und B verschwinden.

Schreibt man den Startwert x_0 in der zu (3.92) gehörigen Basis in der Form $x_0 = (x_l, x_s^+, x_s^-)^T$, ist die Lösungskurve durch x_0 nach Satz 3.4 für Matrizen mit $B = 0$, insbesondere also auch für normale Matrizen, genau dann steif, falls $x_s^+ = 0$ und $x_s^- \neq 0$ gelten. Während die Untermatrix N keinen Einfluss auf steifes Verhalten von Lösungskurven hat, können selbst für $B \neq 0$ steife Lösungskurven existieren, für die $x_s^+ \neq 0$ gilt (Beispiel 3.18 auf Seite 84). Damit kann man einen Schritt weiter gehen als Higham und Trefethen (ebd.) und genau sagen, welcher nichtnormale Anteil für „Überraschungen“ (ebd., S. 286) sorgen kann.

3.4.4 Steife Lösungsfamilien linearer Probleme und explizite numerische Verfahren

Im klassischen Sinne gelten Differentialgleichungen als steif, wenn explizite Verfahren eine kleine Schrittweite erfordern damit das Verfahren überhaupt stabil ist, während implizite Verfahren damit keine Schwierigkeiten haben (siehe Abschnitt 2).

Im Folgenden wird gezeigt, dass die Schrittweite eines expliziten Runge-Kutta-Verfahrens bei der Approximation steifer Lösungsfamilien für $\varepsilon \rightarrow 0$ verschwinden muss. Damit sind explizite Runge-Kutta-Verfahren zur Approximation steifer Lösungsfamilien linearer Probleme ungeeignet.

Satz 3.5. Sei A_ε wie in Satz 3.2 auf Seite 78 und die Lösungsfamilie des durch

$$\dot{x} = A_\varepsilon x, \quad x(0) = x_0, \quad (3.93)$$

gegebenen Anfangswertproblems steif. Weiter sei $\Psi_\varepsilon^\tau x_0$ die diskrete Approximation durch ein explizites Runge-Kutta-Verfahren mit Zeitschrittweite τ . Fordert man für den diskreten Phasenfluss $\Psi_\varepsilon^\tau x_0$

$$0 \neq \text{dist}(\Psi_\varepsilon^\tau x_0, \mu_\varepsilon) = o(1), \quad (3.94)$$

dann gilt für die Zeitschrittweite τ

$$\tau = o(1). \quad (3.95)$$

Insbesondere sind somit explizite Runge-Kutta-Verfahren zur numerischen Näherung steifer Lösungen ungeeignet.

Beweis. Bei expliziten Runge-Kutta-Verfahren ist der diskrete Phasenfluss für lineare Differentialgleichungen durch ein Polynom P gegeben, $\Psi_\varepsilon^\tau = P(\tau A_\varepsilon)$ (siehe z.B. Deuflhard und Bornemann, 2008, Lemma 4.15). Nach Lemma 3.1 ist mit den Bezeichnungen aus Satz 3.2

$$\text{dist}(\Psi_\varepsilon^\tau x_0, \mu_\varepsilon) = \left\| \pi_{\mu_\varepsilon^\perp} P(\tau A_\varepsilon) x_0 \right\|_1 \quad (3.96a)$$

$$= \sum_{i=r+1}^t \left| P(\tau a_i)(x_0)_i + \sum_{j=i+1}^t c_{ij}(x_0)_j \right|. \quad (3.96b)$$

Sei $r+1 \leq m \leq t$ der größte Index mit $(x_0)_m \neq 0$. Da die Lösungsfamilie nach Voraussetzung steif ist, gilt $x_0 \notin \mu_\varepsilon$ und es gibt einen solchen Index. Es folgt

$$\text{dist}(\Psi_\varepsilon^\tau x_0, \mu_\varepsilon) = \left\| \pi_{\mu_\varepsilon^\perp} P(\tau A_\varepsilon) x_0 \right\|_1 \geq |P(\tau a_m)(x_0)_m|. \quad (3.97)$$

Für $a_m \in \sigma_{1/\varepsilon}(A_\varepsilon)$ ist $1 = o(\text{Re } a_m)$. Da Polynome am Rand explodieren, also $|P(z)| \rightarrow \infty$ für $|z| \rightarrow \infty$ gilt, muss nach Voraussetzung (3.94) somit für einen stabilen Zeitschritt $\tau = o(1)$ gelten. \square

Explizite Verfahren fordern also bei der Approximation steifer Lösungen Schrittweiten, die asymptotisch verschwinden. Das stimmt mit der klassischen Vorstellung steifer Probleme überein.

Bemerkung 3.8. Bedingung (3.94) stellt sicherlich eine sinnvolle Bedingung an die Näherungslösung dar, die, wie gesehen, bei expliziten Verfahren aber zwingend eine asymptotisch kleine Zeitschrittweite erfordert. Für implizite Verfahren muss das nicht zutreffen, wie das folgende Beispiel zeigt.

Beispiel 3.22. Betrachtet wird wieder das dahlquistsche Testproblem, diesmal mit dem impliziten Euler-Verfahren. Der diskrete Phasenfluss bezüglich des impliziten Euler-Verfahrens ist durch

$$\Psi_\varepsilon^\tau x_0 = x_0 \left(1 + \frac{\tau}{\varepsilon}\right)^{-1}, \quad (3.98)$$

gegeben. Wie in Beispiel 3.1 auf Seite 62 gesehen, ist $\mu_\varepsilon = \{0\}$ die zu dem dahlquistschen Testproblem gehörende langsame Mannigfaltigkeit. Betrachtet man nun $\text{dist}(\Psi_\varepsilon^\tau x_0, 0)$, lässt sich leicht erkennen, dass (3.94) für jedes τ mit $\tau = \mathcal{O}(1)$ erfüllt ist.

3.5 STEIFE LÖSUNGSFAMILIEN NICHTLINEARER PROBLEME

Für nichtlineare Probleme ist die Charakterisierung steifer Lösungsfamilien deutlich schwieriger als im linearen oder sogar skalaren Fall des dahlquistschen Testproblems. Zum einen kann das Finden der Lösungsstruktur und damit auch der exponentiell anziehenden Mengen nichtlinearer Probleme sehr viel komplexer sein, zum anderen lassen sich langsame Mannigfaltigkeiten nicht so einfach wie im linearen Fall bestimmen.

Higham und Trefethen (1993, Fig. 1.) beschreiben den klassischen Weg um von einem allgemeinen Problem $y' = f(t, y)$ zu einem entkoppelten linearen System $u' = \lambda u$, der dahlquistschen Testgleichung, zu kommen, dies ist in Abbildung 3.2 auf der nächsten Seite dargestellt. Darin bezeichne u die Störung einer speziellen Lösung y_0 von $y' = f(t, y)$ mit $y(t) = y_0(t) + u(t)$. Die Hoffnung dabei ist, dass es ausreicht die einfach zu analysierenden Gleichung (3.99d) zu studieren und dann auf die Ausgangsgleichung (3.99a) zurück zu schließen. Dem ist im Allgemeinen nicht so.

Auf die Schwierigkeiten, die beim Übergang von (3.99c) nach (3.99d) auftreten können wurde bereits in Abschnitt 3.4.3 eingegangen. Die Probleme beim Einfrieren von Koeffizienten, dem Schritt von (3.99c) nach (3.99d), werden in Abschnitt 3.5.1 anhand eines Beispiels diskutiert.

$$y' = f(t, y) \quad (3.99a)$$

↓ Linearisierung

$$u' = A(t)u \quad (3.99b)$$

↓ Einfrieren von Koeffizienten

$$u' = Au \quad (3.99c)$$

↓ Diagonalisierung

$$u' = \lambda u \quad (3.99d)$$

Abbildung 3.2: Higham und Trefethen (1993, Fig. 1.)

In dieser Arbeit werden in erster Linie autonome und lineare Probleme betrachtet. Mit den Problemen vom Typ (2.13), dem nichtlinearen Strudel und dem Van-der-Pol-Oszillator wurden aber auch schon typische nichtlineare, aber autonome, Probleme diskutiert. Die Beschränkung auf autonome Probleme ist, zumindest aus theoretischer Sicht, gerechtfertigt, da jedes nichtautonome Problem zu einem autonomen Problem höherer Dimension äquivalent ist. Allerdings wird aus einem zeitabhängigen linearen Problem durch Autonomisierung ein nichtlineares Problem.

Ein Spezialfall der nichtlinearen Probleme sind diejenigen Probleme, die durch Autonomisierung aus inhomogenen linearen Problemen der Gestalt $\dot{x} = A_\varepsilon x + g_\varepsilon(t)$ entstehen. Für diese lässt sich in Abschnitt 3.5.2 etwas mehr sagen als im allgemeinen nichtlinearen Fall, wie er in Abschnitt 3.5.3 betrachtet wird. Dort wird eine Vermutung für ein hinreichendes Kriterium für steife Lösungsfamilien formuliert. Da dieses Kriterium in dieser Arbeit als Vermutung offen bleibt, wird auf die Effekte, die durch Linearisierung verloren gehen können und auf die Higham und Trefethen (ebd.) hinweisen, an dieser Stelle nicht näher eingegangen. Es ist jedoch eine spannende Aufgabe diese Frage zu beantworten.

3.5.1 Probleme beim Einfrieren von Koeffizienten

Zum Einfrieren von Koeffizienten bemerken Higham und Trefethen (ebd., S. 288): „freezing coefficients may change the behavior of an ODE significantly“. Dies wird an einem Beispiel gezeigt, das auf Kreiss (1978) zurückgeht (Higham und Trefethen, 1993, S. 299).

Beispiel 3.23. Gegeben sei die Differentialgleichung

$$y'(t) = \varepsilon^{-1} U^T(t) \begin{pmatrix} -1 & \eta \\ 0 & -1 \end{pmatrix} U(t) y(t) = A(t) y(t) \quad (3.100)$$

mit der zeitabhängigen unitären Matrix

$$U(t) = \begin{pmatrix} \cos \alpha t & -\sin \alpha t \\ \sin \alpha t & \cos \alpha t \end{pmatrix}. \quad (3.101)$$

Das Beispiel ist akademisch, denn das Problem ist nur scheinbar zeitabhängig. Es ermöglicht aber eine umfassende Analyse, die zeigt, dass das Einfrieren von Koeffizienten Einfluss auf das steife Verhalten von Lösungen hat.

Der einzige Eigenwert des Systems aus (3.100) ist $-\varepsilon^{-1}$. Die Transformation $v(t) = U(t)y(t)$ ist für alle t normerhaltend und mit (3.100) bekommt man

$$v' = U'y + Uy' \quad (3.102a)$$

$$= \begin{pmatrix} 0 & -\alpha \\ \alpha & 0 \end{pmatrix} v + \begin{pmatrix} -\varepsilon^{-1} & \eta\varepsilon^{-1} \\ 0 & -\varepsilon^{-1} \end{pmatrix} v \quad (3.102b)$$

$$= \begin{pmatrix} -\varepsilon^{-1} & \eta\varepsilon^{-1} - \alpha \\ \alpha & -\varepsilon^{-1} \end{pmatrix} v. \quad (3.102c)$$

Eine Besonderheit die hervorzuheben ist, ist dass das transformierte Problem zeitunabhängig ist. Die Eigenwerte der Matrix auf der rechten Seite geben Higham und Trefethen (1993, S. 299) mit $-\varepsilon^{-1}(1 \pm \sqrt{\alpha\varepsilon(\eta - \alpha\varepsilon)})$ an. Da für alle t die langsame Mannigfaltigkeit invariant unter $U(t)$ und $U(t)$ orthogonal für alle t ist, bleiben steife Lösungen nach Satz 3.1 durch die Transformation erhalten. Ob alle von der Nulllösung verschiedenen Lösungen steif sind, hängt von den Vorzeichen der beiden asymptotisch großen Eigenwerte und damit von η ab. Da Eigenwerte für $\eta = 2/(\alpha\varepsilon)$ unterschiedliches Vorzeichen haben existieren nichtsteife Lösungskurven. Für „kleine“ η sind beide Eigenwerte negativ und alle von der Nulllösung verschiedenen Lösungen sind steif.

Friert man dagegen $A(t)$ in (3.100) für ein t_0 ein, ist die orthogonale Matrix $U(t_0)$ zeitlich konstant und die Ableitung U' in (3.102a) verschwindet. Die Eigenwerte bleiben unter der zeitlich konstanten Transformation $v(t) = U(t_0)y(t)$ dann erhalten. Insbesondere sind beide Eigenwerte negativ und alle von der Nulllösung verschiedenen Lösungen sind steif. Orthogonale, bzw. unitäre Transformationen ändern steifes Verhalten nicht. Es spielt demnach keine Rolle ob man (3.100) oder (3.102a) auf steifes Verhalten untersucht. Friert man aber $A(t)$ für ein t_0 ein, verschwindet U' in (3.102a) und wie gesehen unterscheiden sich die Gleichungen $v' = U'y + Uy'$ und $v' = Uy'$ in Bezug auf steifes Verhalten.

Higham und Trefethen (ebd.) schreiben, dass für kleine Werte von η die Matrix $A(t)$ fast normal ist und für festes α die Eigenwerte deutlich von den Eigenwerten von $A(t)$ abweichen. Nach einer konkreten Wahl von ε und η schreiben sie:

Thus we may regard the variable coefficient problem as nearly 5/3 times as stiff as the constant coefficient problem for this particular computation. (ebd., S. 300).

Aus asymptotischer Sicht ist ein Faktor von 5/3 nicht sonderlich groß. Auch die Aussage „[...] 5/3 times as stiff as [...]“ passt nicht zu der Auffassung steifer Lösungen wie sie hier vertreten wird.

Übereinstimmend lässt sich aber festhalten, dass das Einfrieren von Koeffizienten Einfluss darauf hat, welche Lösungen steif sind.

3.5.2 Inhomogene lineare Probleme

Zu Beginn dieses Abschnitts werden zwei Beispiele für inhomogene lineare Probleme betrachtet, die in Satz 3.7 eine Verallgemeinerung finden.

Beispiel 3.24. Aus Gleichung (2.29) sei die inhomogene lineare Differentialgleichung

$$\dot{x} = -\frac{1}{\varepsilon}(x - t^2), \quad x(0) = x_0 \quad (3.103)$$

mit der Lösungsfamilie

$$x(t) = (x_0 - 2\varepsilon^2) \exp(-t/\varepsilon) + t^2 - 2\varepsilon t + 2\varepsilon^2 \quad (3.104)$$

gegeben. Die Lösungsfamilie nähert sich für $\varepsilon \rightarrow 0$ exponentiell der zeitabhängigen Menge

$$\mu_t = \{x \in \mathbb{R} \mid x = t^2 - 2\varepsilon t + 2\varepsilon^2\} \quad (3.105)$$

an. Für das Problem aus (3.103) ist μ_t also eine exponentiell anziehende Menge. Ob μ_t auch eine langsame Mannigfaltigkeit im Sinne von Definition 3.5 auf Seite 64 ist, lässt sich erst einmal nicht sagen, da (3.103) nicht autonom ist. Gleichung (3.103) ist aber äquivalent zu dem zweidimensionalen autonomen System

$$\dot{x} = -\frac{1}{\varepsilon}(x - t^2), \quad (3.106a)$$

$$\dot{t} = 1. \quad (3.106b)$$

Damit kann μ_t als Teilmenge des Phasenraums dieses zweidimensionalen Systems aufgefasst werden. Es sei

$$M_\varepsilon = \{(x, t) \in \mathbb{R}^2 \mid x = t^2 - 2\varepsilon t + 2\varepsilon^2\}. \quad (3.107)$$

Es stellt sich nun die Frage, ob M_ε im Sinne von Definition 3.5 eine langsame Mannigfaltigkeit ist. Dazu muss der Tangentialraum von M_ε für $\varepsilon \rightarrow 0$ mit der Ableitung der rechten Seite von (3.106) für alle Punkte von M_ε übereinstimmen.

Für den Tangentialraum von M_ε im Punkt x_0 gilt

$$T_{x_0}M_\varepsilon = \text{span} \left((2(t - \varepsilon), 1)^T \right). \quad (3.108)$$

Bezeichne f_ε die rechte Seite von (3.106). Die langsame Mannigfaltigkeit der Ableitung im Punkt $x_0 = (x, t)$, die durch

$$df_\varepsilon(x_0) = \begin{pmatrix} -\frac{1}{\varepsilon} & \frac{2}{\varepsilon}t \\ 0 & 0 \end{pmatrix} \quad (3.109)$$

gegeben ist, wird durch einen Eigenvektor zum Eigenwert $\lambda = 0$ erzeugt. Es gilt

$$\mu_\varepsilon(df_\varepsilon(x_0)) = \text{span} \left((2t, 1)^T \right). \quad (3.110)$$

Damit gilt für alle $x_0 \in M_\varepsilon$

$$T_{x_0}M_\varepsilon = \mu_\varepsilon(df_\varepsilon(x_0)) \quad \text{für } \varepsilon \rightarrow 0 \quad (3.111)$$

und M_ε ist nach Definition 3.5 auf Seite 64 tatsächlich eine langsame Mannigfaltigkeit von f_ε .

Die Lösungsfamilien durch $x_0 \notin M_\varepsilon$ nähern sich μ_t für alle $t > 0$ exponentiell an, denn es gilt

$$0 \neq \text{dist}(x(t), \mu_t) = \inf_{q \in M_\varepsilon} \|x(t) - q\| = \mathcal{O} \left(\exp \left(-\frac{t}{\varepsilon} \right) \right). \quad (3.112)$$

In diesem Sinne sind diese Lösungsfamilien steif.

Das nächste Beispiel ist zweidimensional.

Beispiel 3.25. Betrachtet werden zwei zu den Anfangswertproblemen aus Strehmel, Weiner und Podhaisky (2012, Beispiel 7.4.1) ähnliche Probleme, die für $\varepsilon = 1/2$ übereinstimmen:

$$\dot{y} = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix} x(t) + \begin{pmatrix} 2 \sin(t) \\ 2(\cos(t) - \sin(t)) \end{pmatrix}, \quad (3.113a)$$

$$y(0) = \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \quad (3.113b)$$

$$\dot{y} = \begin{pmatrix} -2 & 1 \\ \frac{1}{\varepsilon} - 1 & -\frac{1}{\varepsilon} \end{pmatrix} x(t) + \begin{pmatrix} 2 \sin(t) \\ \frac{1}{\varepsilon} (\cos(t) - \sin(t)) \end{pmatrix}, \quad (3.114a)$$

$$y(0) = \begin{pmatrix} 2 \\ 3 \end{pmatrix}. \quad (3.114b)$$

Nach Strehmel, Weiner und Podhaisky (ebd., S. 203) haben beide die gleiche exakte, ε -unabhängige Lösung

$$x(t) = 2 \exp(-t) \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} \sin(t) \\ \cos(t) \end{pmatrix}. \quad (3.115)$$

Dennoch beobachten sie im Fall $\varepsilon = 1/999$ steifes Verhalten beim numerischen Lösen.

Um die bisher entwickelte Theorie anwenden zu können, wird der homogene Anteil mit der orthogonalen Transformationsmatrix

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \quad (3.116)$$

in Schurform gebracht

$$\dot{x} = \begin{pmatrix} -1 & 2 - \frac{1}{\varepsilon} \\ 0 & -1 + \frac{1}{\varepsilon} \end{pmatrix} x \quad (3.117a)$$

$$+ \frac{1}{\sqrt{2}} \begin{pmatrix} 2 \sin(t) + \frac{1}{\varepsilon}(\cos(t) - \sin(t)) \\ -2 \sin(t) + \frac{1}{\varepsilon}(\cos(t) - \sin(t)) \end{pmatrix}, \quad (3.117b)$$

$$x(0) = \frac{1}{\sqrt{2}} \begin{pmatrix} 5 \\ 1 \end{pmatrix}. \quad (3.117c)$$

Die Lösung lautet dann

$$x(t) = 2\sqrt{2} \exp(-t) \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{1}{\sqrt{2}} \begin{pmatrix} \cos(t) + \sin(t) \\ \cos(t) - \sin(t) \end{pmatrix}. \quad (3.118)$$

Für den Anfangswert $x_0 = (x_1, x_2)$ ist die allgemeine Lösung durch

$$x(t) = \left(x_1 - x_2 + 2\varepsilon \left(x_2 - \frac{1}{\sqrt{2}} \right) \right) \exp(-t) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (3.119a)$$

$$+ \left(x_2 - \frac{1}{\sqrt{2}} \right) \exp\left(-\left(\frac{1}{\varepsilon} + 1\right)t\right) \begin{pmatrix} 1 - 2\varepsilon \\ 1 \end{pmatrix} \quad (3.119b)$$

$$+ \frac{1}{\sqrt{2}} \begin{pmatrix} \cos(t) + \sin(t) \\ \cos(t) - \sin(t) \end{pmatrix} \quad (3.119c)$$

gegeben. Für $x_2 \neq 1/\sqrt{2}$ relaxiert die Lösung exponentiell auf die zeitabhängige Mannigfaltigkeit

$$\mu_t = \left(x_1 - x_2 + 2\varepsilon \left(x_2 - \frac{1}{\sqrt{2}} \right) \right) \exp(-t) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (3.120)$$

$$+ \frac{1}{\sqrt{2}} \begin{pmatrix} \cos(t) + \sin(t) \\ \cos(t) - \sin(t) \end{pmatrix}.$$

Satz 3.6 wird zeigen, dass μ_t bei richtiger Interpretation tatsächlich eine langsame Mannigfaltigkeit ist.

Für den Anfangswert aus (3.117c) ist mit Blick auf den mittleren Term der Lösung (3.119b) kein steifes Verhalten zu erwarten, denn für $x_2 = 1/\sqrt{2}$ gibt es kein asymptotisch exponentielles Abklingen für $\varepsilon \rightarrow 0$. Dennoch beobachten Strehmel, Weiner und Podhaisky (ebd.) für die dazu gehörende Lösung des transformierten Problems steifes Verhalten. Das bedeutet, dass die numerische Lösung nach einer gewissen Anzahl von Schritten einen Punkt erreicht, für den $x_2 \neq 1/\sqrt{2}$ gilt. Für solche Punkte tritt asymptotisch exponentielles Abklingen für $\varepsilon \rightarrow 0$ auf.

Im Folgenden werden nun allgemeine inhomogene Anfangswertprobleme der Form

$$\dot{x}(t) = A_\varepsilon x(t) + g_\varepsilon(t), \quad x(0) = x_0 \quad (3.121)$$

betrachtet, wobei A_ε wieder von der Form wie in Satz 3.2 auf Seite 78 ist und $\pi_{\mu_\varepsilon^\perp} g_\varepsilon \in C^\infty(\mathbb{R})$. Bezeichne μ_ε die langsame Mannigfaltigkeit von A_ε , dann betrachtet man für jedes $t \geq 0$ den affinen Raum

$$\mu_t = \mu_\varepsilon(A_\varepsilon) + \begin{pmatrix} 0 \\ h_\varepsilon(t) \end{pmatrix} \quad (3.122a)$$

mit

$$h_\varepsilon(t) = -U_s^{-1} \pi_{\mu_\varepsilon^\perp} g_\varepsilon(t) + \int_0^t U_s^{-1} \exp((t-z)U_s) \pi_{\mu_\varepsilon^\perp} g'_\varepsilon(z) dz \quad (3.122b)$$

und $t \in \mathbb{R}$ als Teilmenge des Phasenraums Ω . Die Definition von h_ε ergibt Sinn, da U_s regulär und damit invertierbar ist. Betrachtet man nun den um die Zeitkoordinate erweiterten Phasenraum $\Omega \times \{t\}$ und die zu μ_t korrespondierende Menge

$$M_\varepsilon = \{(x, t) \mid x \in \mu_t\} \quad (3.123a)$$

$$= \{(x, t) \mid \pi_{\mu_\varepsilon^\perp} x - h_\varepsilon(t) = 0\} \quad (3.123b)$$

$$= H^{-1}\{0\}, \quad (3.123c)$$

wobei H mit $s = \dim(\mu_\varepsilon^\perp)$ wie folgt definiert ist:

$$H: \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^s \quad (3.124a)$$

$$(x, t) \mapsto \pi_{\mu_\varepsilon^\perp} x - h_\varepsilon(t). \quad (3.124b)$$

H ist differenzierbar und für die Ableitung im Punkt (x_0, t_0) gilt

$$dH(x_0, t_0) = \begin{pmatrix} \pi_{\mu_\varepsilon^\perp} & -h'_\varepsilon(t_0) \end{pmatrix}. \quad (3.125)$$

Damit ist M_ε als Urbild einer differenzierbaren Abbildung eine Mannigfaltigkeit.

Satz 3.6. Gegeben sei das zu (3.121) äquivalente autonome, nichtlineare System

$$\dot{y} = \begin{pmatrix} \dot{x} \\ 1 \end{pmatrix} = \begin{pmatrix} A_\varepsilon x + g_\varepsilon \\ 1 \end{pmatrix} \quad (3.126a)$$

$$= \begin{pmatrix} A_\varepsilon & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix} + \begin{pmatrix} g_\varepsilon \\ 1 \end{pmatrix} \quad (3.126b)$$

$$= F(y) \quad (3.126c)$$

mit $y = (x, t)^T \in \mathbb{R}^{n+1}$. Wenn

$$U_s^{-1} \pi_{\mu_\varepsilon^\perp} g'_\varepsilon(t) \succ \frac{d}{dt} \int_0^t U_s^{-1} \exp((t-z)U_s) \pi_{\mu_\varepsilon^\perp} g'_\varepsilon(z) dz \quad (3.127)$$

gilt, dann ist M_ε aus (3.123) nach Definition 3.5 auf Seite 64 eine langsame Mannigfaltigkeit des autonomen, nichtlinearen Systems aus (3.126a).

Beweis. Für $y_0 \in M_\varepsilon$ ist zu zeigen, dass für $\varepsilon \rightarrow 0$

$$\mu_\varepsilon((dF)(y_0)) = T_{y_0}M_\varepsilon \quad (3.128)$$

gilt.

Einerseits ist der Tangentialraum von M_ε im Punkt $y_0 = (x_0, t_0)$ durch

$$T_{y_0}M_\varepsilon = \ker(dH(x_0, t_0)) \quad (3.129a)$$

$$= \{(x, t) \mid \pi_{\mu_\varepsilon^\perp} x = h'_\varepsilon(t_0)t\} \quad (3.129b)$$

gegeben und hat die Dimension $\dim(T_{y_0}M_\varepsilon) = \dim(\mu_\varepsilon(A_\varepsilon)) + 1$. Für die Ableitung von F an der Stelle $y_0 = (x_0, t_0)$ gilt andererseits

$$(dF)(y_0) = \begin{pmatrix} A_\varepsilon & g'_\varepsilon(t_0) \\ 0 & 0 \end{pmatrix}. \quad (3.130)$$

Ist A_ε invertierbar, dann ist $\lambda = 0$ Eigenwert mit algebraischer und geometrischer Vielfachheit 1 und für die langsame Mannigfaltigkeit der Ableitung gilt

$$\mu_\varepsilon((dF)(y_0)) = \begin{pmatrix} \mu_\varepsilon(A_\varepsilon) \\ 0 \end{pmatrix} \oplus \text{Eig}((dF)(y_0); 0) \quad (3.131a)$$

$$= \begin{pmatrix} \mu_\varepsilon(A_\varepsilon) \\ 0 \end{pmatrix} \oplus \ker((dF)(y_0)). \quad (3.131b)$$

Es ist $\dim(\mu_\varepsilon((dF)(y_0))) = \dim(\mu_\varepsilon(A_\varepsilon)) + 1$.

Ist A_ε singular, muss statt des Eigenraums zum Eigenwert 0 der entsprechende Hauptraum betrachtet werden. Der Einfachheit halber wird aber davon ausgegangen, dass A_ε regulär ist.

Es bleibt $T_{y_0}M_\varepsilon = \mu_\varepsilon((dF)(y_0))$ zu zeigen. Da beide Räume von der gleichen Dimension sind, genügt es $\mu_\varepsilon((dF)(y_0)) \subseteq T_{y_0}M_\varepsilon$ für $\varepsilon \rightarrow 0$ zu zeigen.

Sei $(x, t) \in \mu_\varepsilon((dF)(y_0))$, dann kann (x, t) als Summe von Basisvektoren der direkten Summanden aus (3.131a) geschrieben werden und es genügt die Summanden einzeln zu betrachten.

Sei zunächst $(x, t) \in (\mu_\varepsilon(A_\varepsilon), 0)$, dann ist $t = 0$ und (3.129b) ist erfüllt. Damit ist $(x, t) \in T_{y_0}M_\varepsilon$.

Für $(x, t) \in \ker((dF)(y_0))$ folgt aus $t = 0$ auch $x = 0$, da A_ε als invertierbar angenommen wurde. Der Nullvektor liegt aber schon im ersten direkten Summanden. Sei also o.B.d.A. $t \neq 0$ und darüber hinaus so gewählt, dass

$$-U_s^{-1} \pi_{\mu_\varepsilon^\perp} g'_\varepsilon(\tau_0) t = \mathcal{O}_S(1) \quad (3.132)$$

gilt. Da $(x, t) \in \ker((dF)(y_0))$ ist, gilt für x

$$x = -A_\varepsilon^{-1} g'_\varepsilon(\tau_0) t. \quad (3.133)$$

Mit (3.127) und (3.132) gilt für h_ε aus (3.122b)

$$h'_\varepsilon(t_0) t = -U_s^{-1} \pi_{\mu_\varepsilon^\perp} g'_\varepsilon(t_0) t + o(1) \quad (3.134)$$

und es folgt

$$\pi_{\mu_\varepsilon^\perp} x = -\pi_{\mu_\varepsilon^\perp} A^{-1} g'_\varepsilon(t_0) t \quad (3.135a)$$

$$= -U_s^{-1} \pi_{\mu_\varepsilon^\perp} g'_\varepsilon(t_0) t \quad (3.135b)$$

$$= h'_\varepsilon(t_0) t + o(1). \quad (3.135c)$$

Mit (3.129b) ist damit $(x, t) \in T_{y_0} M_\varepsilon$ für $\varepsilon \rightarrow 0$ gezeigt.

Insgesamt gilt für $\varepsilon \rightarrow 0$ wegen der Gleichheit der Dimensionen von $\mu_\varepsilon((dF)(y_0))$ und $T_{y_0} M_\varepsilon$ damit $\mu_\varepsilon((dF)(y_0)) = T_{y_0} M_\varepsilon$. Nach Definition 3.5 auf Seite 64 ist M_ε also tatsächlich eine langsame Mannigfaltigkeit von (3.126a). \square

Bemerkung 3.9. Die Bedingung (3.127) ist für alle in dieser Arbeit betrachteten Funktionen von vorne herein erfüllt und stellt keine wirklich Bedingung dar. Auf die Suche nach Funktionen, welche diese Bedingung nicht erfüllen, soll hier verzichtet werden.

Der Satz rechtfertigt die folgenden beiden Definitionen.

Definition 3.9. Gegeben sei das inhomogene Anfangswertproblem aus (3.121). Dann ist der affine Unterraum μ_t für jedes $t \in \mathbb{R}$ eine langsame Mannigfaltigkeit des Problems, falls das zu μ_t korrespondierende M_ε eine langsame Mannigfaltigkeit des zu (3.121) äquivalenten autonomen Problems (3.126a) ist.

Definition 3.10. Eine Lösung $x(t)$ des inhomogenen Anfangswertproblems aus (3.121) heißt steif, falls μ_τ im Sinne von Definition 3.6 exponentiell anziehend ist.

Wie der folgende Satz zeigt, gibt es tatsächlich solche steifen Lösungen.

Satz 3.7. Gegeben sei das inhomogene lineare Anfangswertproblem

$$\dot{x}(t) = A_\varepsilon x(t) + g_\varepsilon(t), \quad x(0) = x_0. \quad (3.136)$$

Dabei sei A_ε von der Gestalt wie in Satz 3.2 auf Seite 78 mit $U_s^+ = 0$ und μ_ε die langsame Mannigfaltigkeit des homogenen Anteils A_ε . Der inhomogene Anteil g_ε hänge nur rational von ε ab. Sei $x_0 \notin \mu_0$, dann ist die Lösungsfamilie durch x_0 steif und es gilt

$$0 \neq \text{dist}(\Phi_\varepsilon^t x_0, \mu_t) = o(\exp(-t/\varepsilon)). \quad (3.137)$$

Beweis. Für die Lösungsfamilie der inhomogenen Gleichung erhält man durch Variation der Konstanten

$$x(t) = \exp(tA_\varepsilon) x_0 + \int_0^t \exp((t-z)A_\varepsilon) g_\varepsilon(z) dz. \quad (3.138)$$

Für die schnellen Komponenten $\pi_{\mu_\varepsilon^\perp} x$ erhält man nach partieller Integration

$$\pi_{\mu_\varepsilon^\perp} x(t) = \exp(tU_s) \pi_{\mu_\varepsilon^\perp} x_0 \quad (3.139a)$$

$$+ U_s^{-1} \exp(tU_s) \pi_{\mu_\varepsilon^\perp} g(0) \quad (3.139b)$$

$$+ \int_0^t U_s^{-1} \exp((t-z)U_s) \pi_{\mu_\varepsilon^\perp} g'(z) dz \quad (3.139c)$$

$$- U_s^{-1} \pi_{\mu_\varepsilon^\perp} g(t) \quad (3.139d)$$

$$= \exp(tU_s) \left(\pi_{\mu_\varepsilon^\perp} x_0 + U_s^{-1} \pi_{\mu_\varepsilon^\perp} g(0) \right) \quad (3.139e)$$

$$+ h_\varepsilon(t) \quad (3.139f)$$

$$= \exp(tU_s) \left(\pi_{\mu_\varepsilon^\perp} x_0 - h_\varepsilon(0) \right) + h_\varepsilon(t). \quad (3.139g)$$

Diese Gleichung ergibt Sinn, da U_s regulär und damit invertierbar ist. Außerdem wurde die Identität $\exp(A)A = A \exp(A)$ verwendet. O.B.d.A. sei $U_s = \mathcal{O}_S(1/\varepsilon)$. Da mit g_ε auch $h_\varepsilon(0)$ nach Definition nur rational von ε abhängt, ist μ_t für alle $t \geq 0$ exponentiell anziehend, denn nach Voraussetzung ist $x_0 \notin \mu_0$ und damit auch $\pi_{\mu_\varepsilon^\perp} x_0 \neq h_\varepsilon(0)$ und mit (3.139g) gilt

$$\text{dist}(x(t), \mu_t) = \text{dist} \left(x(t), \mu_\varepsilon(A) + (0, h_\varepsilon(t))^T \right) \quad (3.140a)$$

$$= \text{dist} \left(x(t) - (0, h_\varepsilon(t))^T, \mu_\varepsilon(A) \right) \quad (3.140b)$$

$$= \left\| \pi_{\mu_\varepsilon^\perp} \left(x(t) - (0, h_\varepsilon(t))^T \right) \right\| \quad (3.140c)$$

$$= \left\| \pi_{\mu_\varepsilon^\perp} x(t) - h_\varepsilon(t) \right\| \quad (3.140d)$$

$$= \left\| \exp(tU_s) \left(\pi_{\mu_\varepsilon^\perp} x_0 - h_\varepsilon(0) \right) \right\| \quad (3.140e)$$

$$= \mathcal{O}(\exp(-t/\varepsilon)). \quad (3.140f)$$

□

Damit kann Satz 3.7 auf die Beispiele 3.24 und 3.25 angewendet werden. Wie man überprüfen kann ist (3.127) erfüllt. Satz 3.7 kann für $U_s^+ = 0$ als Verallgemeinerung von Satz 3.4 auf Seite 83 angesehen werden.

3.5.3 Allgemeine nichtlineare Probleme

In den Beispielen 2.8 auf Seite 52 und 3.7 auf Seite 66 wurde der Van-der-Pol-Oszillator diskutiert. Es wurde gezeigt, dass die Menge

$$M = \left\{ x = (x_1, x_2) \mid x_1 - \frac{1}{3}x_2^3 + x_2 = 0, x_2^2 > 1 \right\} \quad (3.141)$$

eine langsame Mannigfaltigkeit ist. Auf die exponentiell anziehenden Mengen des Van-der-Pol-Oszillators wurde noch nicht eingegangen.

Das Finden exponentiell anziehender Mengen erweist sich im nicht-linearen Fall häufig als schwierig oder gar unmöglich, da dazu sehr genaues Wissen über die Lösung nötig ist. Es lässt sich keine geschlossene analytische Lösung angeben und auch eine asymptotische Näherungslösung ist nicht einfach zu finden. Weiter fehlen dem Van-der-Pol-Oszillator Symmetrieeigenschaften wie sie der nichtlineare Strudel in den Beispielen 3.6 auf Seite 65, 3.10 auf Seite 69 und 3.14 auf Seite 72 hatte, die die Analyse deutlich vereinfacht haben.

Betrachte man in (3.23) die Ableitung des Van-der-Pol-Oszillators, dann sind die Eigenwerte in jedem Punkt $x \in M$ durch

$$\lambda_{1/2} = \frac{1-x_2^2}{2\varepsilon}(1 \pm 1) \pm \frac{1}{1-x_2^2} + \mathcal{O}(\varepsilon) \quad (3.142)$$

gegeben. Insbesondere gibt es einen schnellen und einen langsamen Eigenwert. Der asymptotisch schnelle Eigenwert ist negativ und nach Satz 3.4 sind die Lösungen des durch die Ableitungen in den Punkten $x_0 \notin \mu(df_\varepsilon(x_0)) = \text{span}((x_2^2 - 1, 1)^T)$ induzierten Probleme steif. Betrachtet man nochmals Abbildung 2.16 auf Seite 54. Die langsame Mannigfaltigkeit M ist für $\varepsilon \rightarrow 0$ genau die Schnittmenge von blauer und roter Linie. In Abbildung 2.15 auf Seite 53 ist das dort, wo die Geschwindigkeit der zweiten Komponenten (grüne Linie) für $\varepsilon \rightarrow 0$ nicht explodiert, also asymptotisch kleiner ist als außerhalb.

Das Beispiel des Van-der-Pol-Oszillators motiviert die folgende Vermutung.

Vermutung 3.1. Gegeben sei eine langsame Mannigfaltigkeit M der autonomen nichtlinearen Differentialgleichung $\dot{x} = f_\varepsilon(x)$ derart, dass

$$\mu^{\perp+}(df_\varepsilon(x)) = 0 \quad (3.143)$$

für alle $x \in M$ gilt. Dann ist für einen Startwert $x_0 \notin M$ aus einer hinreichend kleinen Umgebung von M die Lösungsfamilie (sofern sie existiert) steif.

Die Situation erinnert an das Stabilitätsverhalten von Lösungskurven nichtlinearer Probleme. Dort lässt sich das Problem abschließend nur für Fixpunkte lösen. Ist x_0 ein Fixpunkt, so ist $f(x_0) = 0$. Damit ist es möglich, Terme höherer Ordnung abzuschätzen und sich auf den linearen Fall zurückzuziehen. Hier kann man etwas ähnliches versuchen. Statt Fixpunkten müssen jedoch langsame Mannigfaltigkeiten betrachtet werden. Wie genau das zu tun ist, kann in dieser Arbeit nicht geklärt werden. Hierzu ist weitere Forschungsarbeit nötig.

Bemerkung 3.10. Man beachte, dass Vermutung 3.1 lediglich ein hinreichendes Kriterium für steife Lösungsfamilien gibt. Eine vollständige Beschreibung der Steifheitsmenge ist damit nicht möglich.

Beispiel 3.26. Für die Fälle $\|x_0\| = 0, 1$ und $\alpha = 1$ wird der nichtlineare Strudel ein weiteres Mal betrachtet. In Beispiel 3.6 auf Seite 65

wurden $M = \{0\}$ und $M = S^1$ als langsame Mannigfaltigkeiten identifiziert. Im ersten Fall ist $\mu^{\perp+}(df_\varepsilon(x_0)) = 0$ für alle $x_0 \in \{0\}$ genau dann, wenn $\beta \in \{\pm 1/\varepsilon\}$ negativ ist. Im zweiten Fall ist diese Bedingung für alle $x_0 \in S^1$ genau dann erfüllt, wenn β positiv ist, siehe Beispiel 3.6 auf Seite 65. Nach Vermutung 3.1 existieren in hinreichend kleinen Umgebungen um die langsamen Mannigfaltigkeiten also Anfangswerte, so dass die dazugehörigen Lösungen steif sind. In Beispiel 3.14 auf Seite 72 wurde das ausführlich diskutiert und die Steifheitsmengen wurden angegeben.

4

DISKUSSION

Inhalt

4.1	Lineare, nichtlineare und intermediäre-Theorie	102
4.1.1	Lineare Theorie	102
4.1.2	Nichtlineare Theorie	103
4.1.3	Intermediäre Theorie	107
4.1.4	Zusammenhang der drei Theorien	109
4.2	Steifes Verhalten und Kondition	110
4.2.1	Charakterisierung steifer Lösungen durch die Kondition	110
4.2.2	Kondition steifer Probleme	113
4.3	Singulär gestörte Probleme	115
4.3.1	Charakterisierung steifen Verhaltens durch singulär gestörte Probleme	115
4.3.2	Wahl des Störungsparameters	116

In diesem Kapitel werden einige klassischen Aspekte steifen Verhaltens wie man sie in der Literatur findet aufgegriffen und diskutiert. Es wird versucht Ähnlichkeiten und Unterschiede zu der familienbasierten Sichtweise aus dieser Arbeit herauszuarbeiten. Das ist nicht immer ganz einfach, da die verfolgte Intention oft weit auseinander liegt. Ziel dieser Arbeit ist es eine von der Numerik losgelöste mathematische Charakterisierung steifer Lösungen zu geben, während der Fokus in der Literatur fast ausschließlich auf die Numerik gerichtet ist und damit auf das konkrete Lösen von Problemen. Durch den familienbasierten Ansatz wird eine ganze Familie von Probleme betrachtet und es wird studiert, wie sich diese Probleme für $\varepsilon \rightarrow 0$ verhalten. Klassischerweise betrachtet man ein Problem und analysiert das Relaxations- bzw. das Langzeitverhalten. Für Einskalenprobleme gibt es da eine direkte Korrespondenz, siehe Satz 2.1.

In der Literatur finden sich verschiedene Herangehensweisen steife Lösungen von Differentialgleichungen zu charakterisieren. Bei allen Ansätzen werden zwei Größen in Bezug gesetzt, wovon die eine „groß“ und die andere „klein“ ist. Bei keinem der Ansätze wird dabei näher spezifiziert wann etwas „groß“ und wann „klein“ ist. So schreiben beispielsweise Deuffhard und Bornemann (2008):

Die *Definition* muß notwendigerweise vage bleiben, da Einschätzungen wie zu *klein* nur pragmatisch erfolgen können [...] (ebd.).

Durch die Betrachtung des Grenzverhaltens von Familien umgeht man diese „vage Einschätzung“, da der Begriff „klein“ nun nicht

mehr „pragmatisch“, sondern präzise im Sinne der Asymptotik zu verstehen ist.

Im Wesentlichen gibt es vier Ansätze steife Differentialgleichungen im klassischen Sinne zu studieren. Dazu gehören die lineare, die nichtlineare und die intermediäre Theorie, die in Abschnitt 4.1 betrachtet werden. In Abschnitt 4.2 wird der vierte Ansatz beschrieben, der steife Differentialgleichungen über die diskrete Kondition charakterisiert. Zum Schluss dieses Kapitels werden in Abschnitt 4.3 singular gestörte Probleme betrachtet, da solche Probleme häufig mit steifem Verhalten in Verbindung gebracht werden.

4.1 LINEARE, NICHTLINEARE UND INTERMEDIÄRE-THEORIE

Die beiden geläufigsten Ansätze zur Beschreibung steifer Lösungen von Differentialgleichungen der Form $y' = f(t, y)$ bezeichnen Higham und Trefethen (1993) als „lineare und nichtlineare Theorie“ (Abschnitt 4.1.1 und Abschnitt 4.1.2). Sie führen eine dritte Theorie ein, die „intermediäre Theorie“, die in Abschnitt 4.1.3 diskutiert wird.

4.1.1 Lineare Theorie

In der linearen Theorie erfolgt die Klassifizierung über das Spektrum einer konstanten Jacobi-Matrix A von f .

Es ist eine übliche Sichtweise steife lineare Probleme nach den Eigenwerten von A zu charakterisieren. Higham und Trefethen (ebd.) schreiben dazu bei steifen Lösungen: „ A has a large spectral radius but a small spectral abscissa“. Formal bedeutet das für den Spektralradius $\rho(A) = \max_{\lambda \in \sigma(A)} |\lambda|$ von A und für die Spektralabszisse $\nu(A) = \max_{\lambda \in \sigma(A)} \operatorname{Re} \lambda$ von A

$$\nu(A) \ll \rho(A). \quad (4.1)$$

Gleich im Anschluss kritisieren sie diese verbreitete Sichtweisen,

[...] the proper definition of stiffness involves a comparison with the time scale of the exact solution [...].

Hierzu sind zwei Punkte zu bemerken. Erstens ist die Kritik von Higham und Trefethen (ebd.) daran, dass die Zeitskala nicht in die Definition eingeht, im Einklang mit den Betrachtungen in dieser Arbeit. Hier hat die relevante Zeitskala immer die Ordnung $\mathcal{O}_S(1)$; „klein“ ($\mathcal{O}(\varepsilon)$) und „groß“ ($\mathcal{O}_S(1/\varepsilon)$) beziehen sich auf diese Zeitskala. Zweitens wird bei dieser Charakterisierung nicht auf eine Lösung und damit nicht auf einen bestimmten Anfangswert eingegangen. Wie an mehreren Beispielen gesehen wurde, ist das jedoch wesentlich dafür,

ob steifes Verhalten tatsächlich zu beobachten ist. Steifheit ist keine Eigenschaft der Differentialgleichung, sondern der Lösung zu einem Anfangswert.

Für reelle Matrizen in Schurform bedeutet ein asymptotisch großer Spektralradius $\mu_{\varepsilon}^{\perp} \neq \{0\}$. Eine kleine Spektralabszisse ist nur möglich, wenn die großen Realteile der Eigenwerte negativ sind. Für entsprechende Anfangswerte führt das zu steifen Lösungskurven. Wie Beispiel 3.18 auf Seite 84 zeigt, kann aber auch ein Anfangswertproblem mit großer Spektralabszisse steife Lösungskurven haben.

Eine weitere Charakterisierung im Sinne der linearen Theorie geht auf Quarteroni, Sacco und Saleri (2002, S. 204) zurück. Sie definieren den Steifheitsquotienten t einer linearen Differentialgleichung $\dot{x} = Ax$ durch

$$t(A) = \max_{\lambda \in \sigma(A)} |\operatorname{Re} \lambda| / \min_{\lambda \in \sigma(A)} |\operatorname{Re} \lambda|$$

und nennen die Gleichung steif, wenn alle Eigenwerte von A negativ sind und der Steifheitsquotient $t(A)$ groß ist. Es wird jedoch bemerkt, dass

verschiedene Autoren die vorangegangene Definition eines *steifen* Problems ungeeignet (ebd.)

finden. Warum das so ist, wird an dem Problem

$$\dot{x} = - \begin{pmatrix} 1 & 0 \\ 0 & \varepsilon \end{pmatrix} x. \quad (4.2)$$

unmittelbar klar. Das System hat zwar einen großen Steifheitsquotienten, $t(A) = 1/\varepsilon$, und negative Eigenwerte, ist jedoch für keinen Anfangswert steif. Die erste Komponente ist auf der Zeitskala des Problems, der Referenzzeitskala, die zweite Komponente ist im Vergleich dazu langsam.

4.1.2 Nichtlineare Theorie

In der nichtlinearen Theorie wird f selbst studiert. Hier spielen nach Higham und Trefethen (1993) Ideen von Liapunov und der Begriff der Kontraktivität eine wesentliche Rolle.

Definition 4.1 (Strehmel, Weiner und Podhaisky, 2012, Definition 7.2.1). Sei $\langle \cdot, \cdot \rangle$ ein Skalarprodukt im \mathbb{R}^n und $\|y\| = \sqrt{\langle y, y \rangle}$ die dadurch induzierte Norm. Die Funktion $f : [0, \infty) \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ genügt einer *einseitigen Lipschitz-Bedingung*, falls es ein $L \in \mathbb{R}$ gibt, derart, dass

$$\langle f(t, y) - f(t, v), y - v \rangle \leq L \|y - v\|^2 \quad (4.3)$$

für alle $t \geq 0$ und alle $y, v \in \mathbb{R}^n$ gilt. $L \in \mathbb{R}$ heißt dann *einseitige Lipschitz-Konstante*.

Definition 4.2 (Strehmel, Weiner und Podhaisky, 2012, Definition 7.2.2). Erfüllt die Differentialgleichung $y'(t) = f(t, y(t))$ die einseitige Lipschitz-Bedingung (4.3) mit $L < 0$ ($L \leq 0$), so heißt die Differentialgleichung *exponentiell kontraktiv* (*schwach kontraktiv, dissipativ*).

Higham und Trefethen (1993) und bemerken, dass bei Stabilitäts- und Steifheitsbetrachtungen der Kontraktivität eine besondere Bedeutung beigemessen wird, was aber häufig zu Bedingungen führt, die zwar hinreichend aber nicht notwendig für Stabilität sind.

[...] because of the emphasis on contractivity, most of the results involve rather stringent conditions that are sufficient but not always necessary for stability or convergence [...] and thus the nonlinear theory is too conservative to provide a sharp characterization of the phenomenon of stiffness [...] (ebd.)

Aus dem Blickwinkel dieser Arbeit ist Kontraktivität weder ein notwendiges noch hinreichendes Kriterium für steife Lösungen.

Wie man sieht, erfüllt die Differentialgleichung

$$\dot{x} = f_\varepsilon(x) = -x \quad (4.4)$$

für $L = -1$ die einseitige Lipschitz-Bedingung. Für keinen Anfangswert $x_0 = x(0) \neq 0$ ist die dazugehörige Lösung steif im Sinne von Definition 3.7 auf Seite 71, es gibt nicht einmal eine schnelle Dynamik. Andersherum ist in Beispiel 3.15 auf Seite 74 jede Lösung zu einem Anfangswert $x_0 = (x_1, x_2)$ mit $x_2 \neq 0$ steif, kontraktiv ist diese Gleichung für $\varepsilon \rightarrow 0$ jedoch nicht. Seien nämlich $y = (y_1, y_2)^T$, $v = (v_1, v_2)^T \in \mathbb{R}^2$, dann gilt für $f = A$:

$$\frac{\langle f(t, y) - f(t, v), y - v \rangle}{\|y - v\|^2} = \frac{(y_1 - v_1)^2 \varepsilon - (y_2 - v_2)^2}{\varepsilon((y_1 - v_1)^2 + (y_2 - v_2)^2)}. \quad (4.5)$$

Für hinreichend kleines ε erfüllt jedes negative L die Bedingung einer einseitigen Lipschitz-Konstante aus (4.3). Nach Definition 4.2 ist die Differentialgleichung aus Beispiel 3.15 auf Seite 74 demnach kontraktiv.

Kontraktivität ist demnach kein geeignetes Kriterium für die Charakterisierung steifen Verhaltens. Wendet man die nichtlineare Theorie auf lineare Probleme mit konstanten Koeffizienten an, spielt der Begriff der logarithmischen Norm eine zentrale Rolle.

Definition 4.3 (Strehmel, Weiner und Podhaisky, 2012, Definition 7.2.3). Sei $\|\cdot\|$ eine beliebige Vektornorm im \mathbb{R}^n . Für die zugeordnete Matrixnorm heißt der Grenzwert

$$\mu[A] = \lim_{\delta \rightarrow +0} \frac{\|I + \delta A\| - 1}{\delta} \quad (4.6)$$

die zugeordnete *logarithmische Norm* der Matrix A .

Aus Schmitt (2012, S. 25) und Ström (1975, Corollary 1) lässt sich der folgende Zusammenhang zwischen logarithmischer Norm und Spektralabszisse ableiten.

Satz 4.1. Für die Spektralabszisse eines Endomorphismus A gilt

$$\nu(A) = \inf_{\|\cdot\|} \mu[A], \quad (4.7)$$

wobei $\mu[A]$ die zur Vektornorm $\|\cdot\|$ gehörige logarithmische Norm bezeichne. Ist A normal, wird das Infimum mit $\|\cdot\| = \|\cdot\|_2$ angenommen.

Konkret werden im folgenden Satz zu einigen Matrixnormen die dazugehörigen logarithmischen Normen angegeben.

Satz 4.2 (Strehmel, Weiner und Podhaisky, 2012, Satz 7.2.4). Die den Matrixnormen $\|\cdot\|_p$ für $p = 1, 2, \infty$ zugeordneten logarithmischen Matrixnormen $\mu_p[\cdot]$ sind für eine Matrix $A \in \mathbb{R}^{n,n}$ gegeben durch

$$\|A\|_1 = \max_{j=1}^n \sum_{i=1}^n |a_{ij}|, \quad \mu_1[A] = \max_{j=1}^n \left(a_{jj} + \sum_{\substack{i=1 \\ i \neq j}}^n |a_{ij}| \right) \quad (4.8)$$

$$\|A\|_\infty = \max_{i=1}^n \sum_{j=1}^n |a_{ij}|, \quad \mu_\infty[A] = \max_{i=1}^n \left(a_{ii} + \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \right) \quad (4.9)$$

$$\|A\|_2 = \sqrt{\lambda_{\max}(A^T A)}, \quad \mu_2[A] = \lambda_{\max}\left(\frac{1}{2}(A^T + A)\right), \quad (4.10)$$

wobei $\lambda_{\max}(\cdot)$ den maximalen Eigenwert der Matrix bezeichnet.

In der nichtlinearen Theorie fassen Higham und Trefethen (1993) die Bedingung für steifes Verhalten ähnlich zur linearen Theorie wie folgt zusammen: „ A has a large norm but a small logarithmic norm“. Formal nicht ganz präzise lässt sich das durch

$$\mu[A] \ll \|A\| \quad (4.11)$$

ausdrücken. Auch hier steht A wieder für eine „eingefrorene“, konstante Jacobi-Matrix von f . Wie in der linearen Theorie kritisieren sie allerdings den fehlenden Bezug zu der Zeitskala. Den stellen Strehmel, Weiner und Podhaisky (2012) in ihrer Charakterisierung steifer Differentialgleichungen her. Allerdings verzichten sie darauf, diese Charakterisierung Definition zu nennen. Wie oben schon zitiert, finden sie „das Problem der Steifheit sehr vielschichtig“ und kennen „keine zufriedenstellende Definition für den Begriff ‚Steifheit einer Differentialgleichung‘“.

Ein Anfangswertproblem $y' = f(t, y)$, $y(t_0) = y_0$ ist auf dem Intervall $[t_0, t_e]$ steif, wenn eine logarithmische Ma-

trixnorm existiert, so dass für $v \in \mathbb{R}^n$ aus einer Umgebung der exakten Lösung $y(t)$ gilt

$$(t_e - t_0) \sup \|f_y(t, v)\| \gg 1 \quad \text{und} \quad (4.12a)$$

$$\mu_0 = \sup \mu[f_y(t, v)] \ll \sup \|f_y(t, v)\|. \quad (4.12b)$$

(Strehmel, Weiner und Podhaisky, 2012, S. 208)

Sie stellen zwei Eigenschaften steifer Systeme heraus:

- (i) Ein System gewöhnlicher Differentialgleichungen ist steif, wenn gewisse Komponenten der Lösung sehr viel schneller abklingen als andere.
- (ii) Ein System gewöhnlicher Differentialgleichungen ist steif, wenn explizite Verfahren aus Stabilitätsgründen sehr kleine Schrittweiten verwenden, obwohl sich die Lösung kaum ändert, implizite Verfahren dagegen große Schrittweiten zulassen. Das heißt, die Schrittweite wird durch die Stabilität des Verfahrens und nicht durch Genauigkeitsanforderungen bestimmt.

(ebd., S. 207)

Auf beide Eigenschaften wurde schon eingegangen, Definition 3.7 und Satz 3.5.

Es gibt aber Fälle in denen sich die Charakterisierung aus (4.12) und die beiden Eigenschaften (i) und (ii) widersprechen.

Das System im nachfolgenden Beispiel 4.1 erfüllt (4.12), nicht aber die beiden Eigenschaften (i) und (ii). Darüber hinaus ist es auch nach Definition 3.7 nicht steif.

Für autonome lineare Gleichungen mit $f(t, y) = Ay$ reduzieren sich (4.12) nach Strehmel, Weiner und Podhaisky (ebd.) zu

$$\|A\| \gg (t_e - t_0)^{-1}, \quad (4.13a)$$

$$\mu[A] \ll \|A\|. \quad (4.13b)$$

Für solche Systeme ist die Charakterisierung demnach unabhängig vom Anfangswert. In zahlreichen der betrachteten Beispiele konnte man jedoch sehen, dass dem i. A. nicht so ist, nicht nur nach Definition 3.7, sondern auch aus numerischer Sicht.

Beispiel 4.1. Betrachtet wird auf dem Einheitsintervall das oszillatorische System

$$\dot{x} = Ax = \begin{pmatrix} 0 & 1/\varepsilon \\ -1/\varepsilon & 0 \end{pmatrix} x \quad (4.14)$$

zusammen mit der euklidischen Matrixnorm $\|\cdot\|_2$ und der davon induzierten logarithmischen Matrixnorm $\mu_2[\cdot]$. Mit $\sigma(A) = \{i/\varepsilon, -i/\varepsilon\}$ hat A zwei rein imaginäre Eigenwerte und es gilt

$$\|A\|_2 = 1/\varepsilon, \quad (4.15a)$$

$$\mu_2[A] = 0. \quad (4.15b)$$

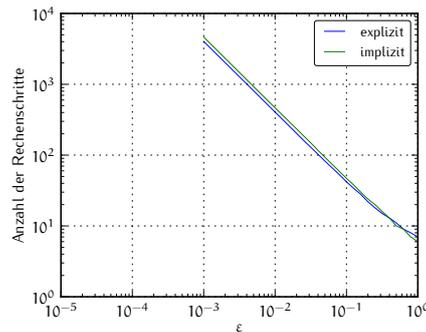


Abbildung 4.1: Numerisches Verhalten von (4.14).

Für $\varepsilon \rightarrow 0$ ist (4.13) erfüllt. Dennoch hat das System nach Definition 3.7 keine steifen Lösungen, denn die langsame Mannigfaltigkeit $\mu_\varepsilon(A)$ hat volle Dimension. In Abbildung 4.1 erkennt man, dass sowohl das explizite, als auch das implizite Verfahren gleichermaßen aufwendig sind. Das Problem hat also keine steifen Lösungen. Trotzdem ist das Problem auf Grund der hohen oszillatorischen Frequenz numerisch aufwendig.

Strehmel, Weiner und Podhaisky (ebd.) verwenden die logarithmische Matrixnorm als Maß für das Relaxationsverhalten. Die logarithmische Matrixnorm ist aber nicht invariant unter Ähnlichkeit, hängt also von der Wahl einer Basis ab. In den Beispielen 2.6 und 2.7 wurde gezeigt, dass steifes Verhalten im Allgemeinen nicht invariant unter Koordinatentransformationen ist. Für orthogonale lineare Transformationen jedoch bleibt steifes Verhalten nach Satz 3.1 jedoch erhalten. Darüber hinaus ist „[...] die Bestimmung einer geeigneten logarithmischen Matrixnorm für nichtlineare Probleme kompliziert [...]“ (ebd.). Damit ist die logarithmische Matrixnorm nur bedingt zur Charakterisierung steifer Lösungen geeignet.

4.1.3 Intermediäre Theorie

Für Higham und Trefethen (1993) ist Steifheit ein „vorrübergehendes Phänomen“:

The idea here is that instability and stiffness are fundamentally transient phenomena, which may appear near some times t_0 and not others. (ebd., S. 287)

Sie betrachten für das linear konstante Problem

$$u' = Au, \quad u(0) = u_0 \quad (4.16)$$

mit der exakten Lösung

$$u(t) = \exp(tA)u_0 \quad (4.17)$$

die Rekursionsgleichung für ein explizites Runge-Kutta-Verfahren der Ordnung m

$$v^{(n)} = P(A\Delta t)^n v^{(0)} \quad (4.18)$$

mit einem Polynom P der Ordnung 4 und der konstanten Zeitschrittweite Δt als diskretes Analogon zu (4.17). Um die Güte der Approximation zu messen, betrachten Higham und Trefethen (1993) mit Blick auf (4.18) die Größe $\|P(A\Delta t)^n\|$, für die die Abschätzung

$$\rho(P(A\Delta t))^n \leq \|P(A\Delta t)^n\| \leq \|P(A\Delta t)\|^n \quad (4.19)$$

gilt (ebd., (3.1)). Hierbei bezeichne $\rho(A)$ den Spektralradius der Matrix A . Dieser lässt sich als Grenzwert einer durch eine Vektornorm induzierten Matrixnorm schreiben:

$$\lim_{n \rightarrow \infty} \|P(A\Delta t)^n\|^{1/n} = \rho(P(A\Delta t)). \quad (4.20)$$

Für große n verhält sich $\|P(A\Delta t)^n\|$ ungefähr wie die n -te Potenz des Spektralradius von $P(A\Delta t)$. Für $t \rightarrow \infty$ gilt bei konstanter Zeitschrittweite $t/\Delta t = n \rightarrow \infty$. In diesem Sinne wird das Verhalten von $\|P(A\Delta t)^n\|$ für $t \rightarrow \infty$ durch das Spektrum von $P(A\Delta t)$ bzw. $A\Delta t$ bestimmt.

Weiter gilt die Identität

$$\lim_{n \rightarrow 0} \frac{\|P(A\Delta t)^{n+1}\|}{\|P(A\Delta t)^n\|} = \|P(A\Delta t)\|, \quad (4.21)$$

so dass das Verhalten von $\|P(A\Delta t)^n\|$ für $t = \Delta t n \rightarrow 0$ durch die Norm von $P(A\Delta t)$ bestimmt wird.

Higham und Trefethen (ebd.) weisen darauf hin, dass für nichtnormale Matrizen die Abstände der Größen aus (4.19) weit auseinander liegen können und die beiden Schranken aus numerischer Sicht ungeeignet sind etwas über Stabilität oder Steifheit zu sagen (vgl. ebd., Fig. 2). Als dritte Theorie führen sie deshalb die intermediäre Theorie ein, die ähnlich zur linearen Theorie ist, wobei aber das Spektrum durch das Pseudospektrum ersetzt wird:

A problem is stiff for $t \approx t_0$ if the pseudospectra of this linear approximation extend far into the left half-plane as compared with the time scale of the solution for $t \approx t_0$. (ebd.)

Das Pseudospektrum ist dabei wie folgt gegeben.

Definition 4.4 (Trefethen und Embree (2005, S. 13)). Sei $A \in M(n, \mathbb{C})$ und $\delta > 0$. Das δ -Pseudospektrum $\sigma_\delta(A)$ von A ist die Menge der $\lambda \in \mathbb{C}$ mit

$$\|(\lambda - A)^{-1}\| > \delta^{-1}. \quad (4.22)$$

Für normale Matrizen stimmen Spektrum und Pseudospektrum überein.

Satz 4.3 (Trefethen und Embree (ebd., Theorem 2.3) (Bauer-Fike Theorem)). Sei $A \in M(n, \mathbb{C})$. Ist $A = \Lambda D \Lambda^{-1}$ diagonalisierbar mit einer Diagonalmatrix D und einer Matrix Λ aus Eigenvektoren von A und bezeichne $\kappa(\Lambda) = \|\Lambda\|_2 \|\Lambda^{-1}\|_2$ die Kondition von Λ und B_δ den Einheitsball mit Radius δ , dann gilt für jedes $\delta > 0$ und $\|\cdot\| = \|\cdot\|_2$:

$$\sigma(A) + B_\delta \subseteq \sigma_\delta(A) \subseteq \sigma(A) + B_{\delta\kappa(\Lambda)}. \quad (4.23)$$

Für normales A kann Λ unitär gewählt werden; damit ist $\kappa(\Lambda) = 1$. Da die Inklusionskette für beliebiges δ gilt, fällt die intermediäre Theorie von Higham und Trefethen (1993) in diesem Fall mit der linearen und damit auch mit der nichtlinearen Theorie zusammen.

Steifes Verhalten von Lösungen als vorübergehendes Phänomen zu betrachten passt nicht unmittelbar zu der Sichtweise in dieser Arbeit. Das liegt im Wesentlichen an dem Familienansatz, der hier betrachtet wird. Dennoch kann man eine Analogie finden. In dieser Arbeit hat die Integrationszeitskala stets die Ordnung $\mathcal{O}_S(1)$ während die schnelle Relaxation auf der Zeitskala $\mathcal{O}(1)$ (in den Beispielen immer $\mathcal{O}(\varepsilon)$) abläuft. Betrachtet man die Integrationszeitskala inmitten einer Kaskade von asymptotischen Zeitskalen, ist auch die $\mathcal{O}_S(1)$ -Zeitskala „vorübergehend“.

4.1.4 Zusammenhang der drei Theorien

Ist A normal, dann gilt für den Spektralradius und jede von dem zugehörigen Skalarprodukt induzierte Norm $\rho(A) = \|A\|$. Nach Satz 4.1 ist $\nu(A) = \mu[A]$. Zusammen mit Satz 4.3 folgt, dass im Falle normaler, linear konstanter Probleme alle drei Theorien aus Abschnitt 4.1 zusammen fallen. Im Allgemeinen muss das nicht gelten. Higham und Trefethen (ebd.) schreiben dazu:

[...] the “nonlinear” theory differs from the “linear” theory even for linear problems with constant coefficients, where it amounts to analysis of (2.3) [(3.99c)] by means of such quantities as the one-sided Lipschitz constant of f , the logarithmic norm of A , and the norm of the discrete solution operator associated with A . (ebd.)

Ein einfaches Beispiel dafür ist durch

$$\dot{x} = Ax, \text{ mit } A = \begin{pmatrix} -1/\varepsilon & 1 \\ 0 & -1/\varepsilon \end{pmatrix} \quad (4.24)$$

gegeben. Für dieses Problem gelten

$$\rho(A) = 1/\varepsilon, \quad \|A\|_1 = 1/\varepsilon + 1, \quad (4.25a)$$

$$\nu(A) = -1/\varepsilon, \quad \mu_1[A] = -1/\varepsilon + 1. \quad (4.25b)$$

Asymptotisch gesehen können lineare und nichtlineare Theorie also auch für nichtnormale Matrizen zusammen fallen

$$\rho(A) \approx \|A\|_1 \approx 1/\varepsilon, \quad \nu(A) \approx \mu_1[A] \approx -1/\varepsilon. \quad (4.26)$$

Angenommen für die $n \times n$ Schurform A gelte $\dim(\mu_\varepsilon^\perp) = n$, dann sind die Realteile aller Eigenwerte asymptotisch groß. Sind weiter alle Koeffizienten von A von der Größenordnung $\mathcal{O}(1/\varepsilon)$ und ist zudem $\rho(A) = \mathcal{O}_S(1/\varepsilon)$, dann gilt

$$\|A\| \approx \rho(A). \quad (4.27)$$

Diese Aussage folgt unmittelbar aus

$$\mathcal{O}_S(1/\varepsilon) = \rho(A) \leq \|A\| \leq C \max_{ij} |a_{ij}| = \mathcal{O}_S(1/\varepsilon), \quad (4.28)$$

für eine von der Norm abhängige Konstante C .

Ob logarithmische Norm und Spektralabszisse asymptotisch zusammen fallen, lässt sich im Allgemeinen nicht sagen. Dennoch sieht man einmal mehr, dass es wesentlich ist, das Teilproblem der schnellen Dynamik zu studieren. Es genügt nicht, sich auf das Verhalten für $t \rightarrow \infty$ zu beschränken.

4.2 STEIFES VERHALTEN UND KONDITION

4.2.1 Charakterisierung steifer Lösungen durch die Kondition

Eine weitere Möglichkeit der Charakterisierung steifer Anfangswertprobleme geht auf Deuffhard und Bornemann (2008) zurück. Sie vergleichen die *Kondition* und die *diskrete Kondition* eines Problems. Hierbei spielt die gewählte Diskretisierung eine essentielle Rolle.

Definition 4.5 (ebd., Abschnitt 3.1.2). Gegeben sei das Anfangswertproblem $\dot{x} = f(x, t)$, $x(t_0) = x_0$ mit der Lösung $x(t)$, dann beschreibt $x(t) + \delta x(t)$ einen von $x(t)$ gestörten Punkt im Phasenraum. Die Störung des Zustandes zum Zeitpunkt t wird $\delta x(t)$ bezeichnet. Dann ist die *intervallweise Kondition* $\kappa[t_0, t]$ die kleinste Zahl, für die

$$\max_{s \in [t_0, t]} |\delta x(s)| \leq \kappa[t_0, t] \cdot |\delta x_0| \quad (4.29)$$

gilt.

Mit $t_0 < t_1 < \dots < t_n = T$ sei das Gitter $\Delta = \{t_0, t_1, \dots, t_n\}$ auf $[t_0, T]$ gegeben. Die Schrittweite zwischen zwei Gitterpunkten wird mit $\tau_j = t_{j+1} - t_j$, $j = 0, \dots, n-1$ bezeichnet und $\tau_\Delta = \max_{0 \leq j \leq n} \tau_j$ heißt die maximale Schrittweite. Weiter sei $x_\Delta : \Delta \rightarrow \mathbb{R}^n$ eine Gitterfunktion, die die exakte Lösung x eines Anfangswertproblems an den Gitterpunkten approximiert:

$$x_\Delta(t) \approx x(t) \quad \text{für alle } t \in \Delta. \quad (4.30)$$

Definition 4.6 (ebd., Abschnitt 4.1.3). Sei \bar{x}_0 ein gestörter Anfangswert und \bar{x}_Δ die dazugehörige Gitterfunktion, die das gewählte numerische Verfahren erzeugt. Die *diskrete Kondition* κ_Δ ist die kleinste Zahl für

$$\max_{t \in \Delta} |x_\Delta(t) - \bar{x}_\Delta(t)| \leq \kappa_\Delta |x_0 - \bar{x}_0| \quad \text{für } \bar{x}_0 \rightarrow x_0. \quad (4.31)$$

gilt.

Stellen die beiden Gitterfunktionen x_Δ und \bar{x}_Δ für kleine Störungen $|x_0 - \bar{x}_0|$ vernünftige Approximationen der kontinuierlichen Lösungen x und \bar{x} dar, dann erwartet man nach Deuffhard und Bornemann (ebd., (4.5))

$$\kappa_\Delta \approx \kappa[t_0, T]. \quad (4.32)$$

Das entspricht einer numerischen Abweichung, die vergleichbar ist mit der exakten Abweichung. Weiter argumentieren sie, dass für konvergente Verfahren

$$\kappa_\Delta \rightarrow \kappa[t_0, T] \quad \text{für } \tau_\Delta \rightarrow 0 \quad (4.33)$$

gilt. Die intervallweise Kondition hängt nur vom Problem, die diskrete Kondition zusätzlich vom gewählten Verfahren ab. Ist $\kappa_\Delta \gg \kappa[t_0, T]$, so ist die Zeitschrittweite τ_Δ zu groß. Um (4.32) zu erreichen, muss eine kleinere Zeitschrittweite τ_Δ gewählt werden. Darauf basiert die Definition steifer Anfangswertprobleme nach Deuffhard und Bornemann (ebd.):

Es gibt nun Anfangswertprobleme, für welche die im vorliegenden Kapitel vorgestellten expliziten Einschrittverfahren „zu kleine“ τ_Δ und damit „zu großen“ Aufwand benötigen, um dieser Forderung zu genügen. Solche Anfangswertprobleme werden in der Literatur *steif* genannt, alle anderen Anfangswertprobleme heißen *nichtsteif*. (ebd., S. 13)

Durch diese Definition werden Probleme, deren Lösung mit expliziten Verfahren nur mit großem Aufwand möglich ist, nicht aufgrund der mathematischen Struktur klassifiziert, vielmehr ist es genau andersherum (vgl. Kapitel 2).

Hier führt also nicht eine Problemklasse zur Wahl von Verfahren, sondern eine Verfahrensklasse klassifiziert die Probleme! (ebd., S. 138)

Wie zu Beginn des Abschnittes schon erwähnt, muss die Definition nach Deuffhard und Bornemann (ebd.) [...]

[...] notwendigerweise vage bleiben, da Einschätzungen wie zu *klein* nur pragmatisch erfolgen können, indem der Rechenaufwand bewertet wird und sodann eine Entscheidung für oder wider die Verfahrensklasse getroffen wird. (Deuffhard und Bornemann, 2008)

Hieran wird die Stärke einer asymptotischen Betrachtungsweise, die eine genaue Einteilung von Größenordnungen ermöglicht, deutlich.

Beispiel 4.2. Als Beispiel betrachten Deuffhard und Bornemann (ebd.) das dahlquistsche Testproblem

$$\dot{x} = \lambda x, \quad x(0) = 1 \quad (4.34)$$

auf einem äquidistanten Gitter ($t_j = \tau_\Delta$ für $j = 0, \dots, n-1$) mit dem expliziten Euler-Verfahren. Für die beiden Fälle $\lambda \geq 0$ und $\lambda \leq 0$ berechnen sie die intervallweise bzw. die diskrete Kondition:

$$\lambda \geq 0: \quad \kappa[0, T] = e^{\lambda T}, \quad \kappa_\Delta \leq e^{\lambda T}, \quad (4.35a)$$

$$\lambda \leq 0: \quad \kappa[0, T] = 1, \quad \kappa_\Delta = \max_{1 \leq k \leq n} |1 - \tau_\Delta \lambda|^k. \quad (4.35b)$$

Für $\lambda \geq 0$ ist das Anfangswertproblem nicht steif, da die Fehlerverstärkung durch die Diskretisierung κ_Δ stets durch die intervallweise Kondition beschränkt ist. Für $\lambda \leq 0$ hingegen ist die diskrete Kondition nur für $\tau_\Delta < 2/|\lambda|$ durch die intervallweise Kondition beschränkt. Je größer $|\lambda|$ ist, desto schärfer ist die Schrittweitenbeschränkung. Für $\tau_\Delta \gg 2/|\lambda|$ ist die Verstärkung des Eingabefehlers durch die Diskretisierung deutlich größer als in dem Fall der exakten Lösung

$$\kappa_\Delta \geq |1 - \tau_\Delta \lambda| \gg 1 = \kappa[0, T]. \quad (4.36)$$

Um die vertraute Notation mit Bezug auf eine Familie von Problemen zu erhalten, setzt man $\lambda = 1/\varepsilon$.

Die Definition nach Deuffhard und Bornemann (ebd.) ist aber im Allgemeinen kaum anwendbar. Die Schwierigkeit liegt in der Bestimmung der diskreten Kondition.

Betrachte man das explizite Euler-Verfahren und das durch

$$\dot{x} = -\frac{1}{\varepsilon} x^2, \quad x(0) = 1 \quad (4.37)$$

gegebene Anfangswertproblem. Über die Propagationsmatrix (ebd., S. 90)

$$W(t, s) = d_\xi \Phi^{t,s} \xi|_{\xi=\Phi^{s,t_0} x_0} \quad (4.38)$$

lässt sich die intervallweise Kondition einfach berechnen

$$\kappa[0, T] = \max_{s \in [0, T]} \|W(s, 0)\| = \max_{s \in [0, T]} \left| \frac{\varepsilon^2}{(\varepsilon + s x_0)^2} \right| = 1. \quad (4.39)$$

Analog berechnet sich für den diskreten Phasenfluss Ψ , der für ein äquidistantes Gitter Δ durch

$$x_{\Delta}(t_{j+1}) = \Psi^{\tau} x_{\Delta}(t_j), \quad j = 0, \dots, n-1 \quad (4.40)$$

gegeben ist, die diskrete Kondition

$$\kappa_{\Delta} = \max_{1 \leq i \leq n} |d_{\xi} \Psi^{\tau i} \xi|_{\xi=x_0}|. \quad (4.41)$$

Hierzu ist aber die Rekursion (4.40) zu lösen, was sich schon für dieses einfache Beispiel als außerordentlich schwierig erweist. Für das explizite Euler-Verfahren ist die Rekursion

$$x_{\Delta}(j+1) = x_{\Delta}(j) \left(1 - \frac{\tau}{\varepsilon} x_{\Delta}(j)\right) \quad (4.42)$$

zu lösen. Eine Substitution $y_{\Delta} = -\tau/\varepsilon x_{\Delta}$ ergibt

$$y_{j+1} = y_j (1 + y_j). \quad (4.43)$$

Das ist ein Spezialfall der logistischen Gleichung für die keine Lösung bekannt ist (<http://mathworld.wolfram.com/LogisticMap.html>).

4.2.2 Kondition steifer Probleme

Nennt man ein Problem (eine Familie von Problemen) steif, wenn steife Lösungskurven (Familien von Lösungskurven) existieren, gelten steife Probleme im Allgemeinen als gut konditioniert. Hier ist Vorsicht geboten, wie Beispiel 4.2 zeigt.

Satz 4.4. *Gegeben sei die autonome Familie von Differentialgleichungen*

$$\dot{x}_{\varepsilon}(t) = f_{\varepsilon}(x_{\varepsilon}(t)) \quad (4.44)$$

mit einer Familie von langsamen Mannigfaltigkeiten μ_{ε} . Die Lösung $x_{\varepsilon}(t)$ zu dem Anfangswert $x_{\varepsilon}(0) = x_0$ sei steif. Sei weiter $\tilde{x}_{\varepsilon}(t)$ die gestörte Lösung zu dem gestörten Anfangswert $\tilde{x}_{\varepsilon}(0) = \tilde{x}_0$. Die Projektionen auf die langsame Mannigfaltigkeit werden wie folgt abgekürzt:

$$y_{\varepsilon} = \pi_{\mu_{\varepsilon}} x_{\varepsilon}, \quad y_0 = \pi_{\mu_{\varepsilon}} x_0, \quad \tilde{y}_{\varepsilon} = \pi_{\mu_{\varepsilon}} \tilde{x}_{\varepsilon}, \quad \tilde{y}_0 = \pi_{\mu_{\varepsilon}} \tilde{x}_0. \quad (4.45)$$

Ist die gestörte Lösung \tilde{x}_{ε} ebenfalls steif, dann gilt bei hinreichend kleinem ε für alle t

$$\|x_{\varepsilon} - \tilde{x}_{\varepsilon}\| \leq \kappa_1 \kappa_{\pi_{\mu_{\varepsilon}}} \|x_0 - \tilde{x}_0\|. \quad (4.46)$$

Hierbei bezeichne κ_1 die Kondition des auf die langsame Mannigfaltigkeit projizierten Problems und $\kappa_{\pi_{\mu_{\varepsilon}}}$ die Kondition der Projektion auf die langsame Mannigfaltigkeit μ_{ε} . Die Kondition des gesamten Problems κ ist also durch $\kappa_1 \kappa_{\pi_{\mu_{\varepsilon}}}$ nach oben beschränkt.

Beweis. Gelten die Voraussetzungen aus dem Satz, dann folgt die Behauptung für $\varepsilon \rightarrow 0$ aus der folgenden Kette von Ungleichungen:

$$\|x_\varepsilon - \tilde{x}_\varepsilon\| = \|(y_\varepsilon - \tilde{y}_\varepsilon) + (x_\varepsilon - y_\varepsilon) + (\tilde{y}_\varepsilon - \tilde{x}_\varepsilon)\| \quad (4.47a)$$

$$\leq \|y_\varepsilon - \tilde{y}_\varepsilon\| + \text{dist}(x_\varepsilon, \mu_\varepsilon) + \text{dist}(\tilde{x}_\varepsilon, \mu_\varepsilon) \quad (4.47b)$$

$$\leq \kappa_1 \|y_0 - \tilde{y}_0\| + \mathcal{O}(\exp(-t/\varepsilon)) \quad (4.47c)$$

$$\leq \kappa_1 \kappa_{\pi_{\mu_\varepsilon}} \|x_0 - \tilde{x}_0\| + \mathcal{O}(\exp(-t/\varepsilon)). \quad (4.47d)$$

Nach Addition einer „geschickten“ Null entspricht die erste Ungleichung der Dreiecksungleichung. Da die beiden Lösungen steif sind, sind die Abstände zur langsamen Mannigfaltigkeit von der Größenordnung $\mathcal{O}(\exp(-t/\varepsilon))$. \square

Bemerkung 4.1. Ist das Anfangswertproblem aus Satz 4.4 linear und ist in der Notation von Satz 3.2 auf Seite 78 zudem $U_s^+ = 0$, dann ist nach Satz 3.4 auf Seite 83 die Vereinigung der Steifheitsmenge und der langsamen Mannigfaltigkeit der gesamte Raum. In diesem Fall ist jede gestörte Lösung \tilde{x} einer steifen Lösung x steif. Andernfalls muss das nicht gelten.

Ist die langsame Mannigfaltigkeit ein linearer Raum, was für lineare Probleme immer der Fall ist, gilt für die Kondition $\kappa_{\pi_{\mu_\varepsilon}}$ der Orthogonalprojektion π_{μ_ε} gerade $\kappa_{\pi_{\mu_\varepsilon}} = \mathcal{O}(1)$.

Wie man zum Beispiel mit der Zeilensummennorm einsieht, sind steife Probleme dann gut konditioniert, wenn das langsame Teilproblem ($\dot{x} = U_1 x$, siehe (3.92)) keine schnellen Komponenten aufweist.

Dass die Kondition κ_1 des „langsamen“ Teilproblems keinesfalls asymptotisch klein sein muss, zeigt das nächste Beispiel. Grund dafür ist die erlaubte schnelle Dynamik in der langsamen Mannigfaltigkeit.

Beispiel 4.3. Für die Kondition des langsamen Teilproblems aus Beispiel 3.3 auf Seite 63 gilt:

$$\kappa_1[0, t] = \left\| \begin{pmatrix} e^t & e^{t/\varepsilon} \\ 0 & e^t \end{pmatrix} \right\| = \mathcal{O}_s(1/\varepsilon). \quad (4.48)$$

Das gesamte Problem ist nach Satz 4.4 also schlecht konditioniert obwohl sämtliche Lösungskurven mit Startwert außerhalb der langsamen Mannigfaltigkeit steif sind.

4.3 SINGULÄR GESTÖRTE PROBLEME

4.3.1 Charakterisierung steifen Verhaltens durch singular gestörte Probleme

In der Literatur (siehe z.B. Deuffhard und Bornemann, 2008; Strehmel, Weiner und Podhaisky, 2012) werden Probleme vom Typ

$$\dot{x}_1 = f_1(x_1, x_2) \quad (4.49a)$$

$$\varepsilon \dot{x}_2 = f_2(x_1, x_2) \quad (4.49b)$$

oft als singular gestört bezeichnet. Zu diesem Problemtyp gehören z.B. das dahlquistsche Testproblem, der Van-der-Pol-Oszillator, das Problem

$$\dot{x} = \begin{pmatrix} 0 & 1 \\ -\frac{1}{\varepsilon} & 0 \end{pmatrix} x. \quad (4.50)$$

und Probleme vom Typ wie in (2.13).

Aus numerischer Sicht können die ersten drei genannten Probleme Schwierigkeiten bereiten, aber ihr Charakter unterscheidet sich grundlegend. Wie bereits gesehen weisen sowohl das dahlquistsche Testproblem als auch der Van-der-Pol-Oszillator steifes Verhalten auf. Beide sind vom Grenzschichttyp, ebenso wie die Probleme aus (2.13), deren Lösungen aber nie steif sind. Die Realteile der Eigenwerte des Systems aus (4.50) verschwinden gänzlich. Das System ist rein oszillatorisch, hat damit eine leere Steifheitsmenge und relaxiert nicht.

Während oszillatorisch gestörte Probleme kein steifes Verhalten aufweisen, können singular Grenzschichtprobleme steife Lösungen haben, müssen es aber nicht. Es ist deshalb angebracht, strikt zwischen steifen und oszillatorischen Systemen zu unterscheiden, was man auch als Antwort auf „QUESTION 1“ aus Abschnitt 2.3 verstehen kann.

Schließt man die oszillatorischen Probleme aus, gelten singular gestörte Probleme häufig als „Klasse steifer Probleme“ (siehe z.B. Strehmel, Weiner und Podhaisky, 2012, S. 209). Strehmel, Weiner und Podhaisky (ebd.) schließen oszillatorische Probleme implizit aus:

Je kleiner ε , umso größer wird die Norm der Jacobi-Matrix $1/\varepsilon g_v$ [hier: $1/\varepsilon \partial f_2/\partial x_2$] und damit die Steifheit. (ebd.).

Diese Aussage stimmt nur, wenn $\partial f_2/\partial x_2 \neq 0$ ist und auch dann können Lösungskomponenten noch explodieren anstatt schnell abzuklingen.

Deuffhard und Bornemann (2008) nutzen singular gestörte Probleme nicht zur Charakterisierung steifen Verhaltens. Vielmehr untersuchen sie in Deuffhard und Bornemann (ebd., § 6.4.3) für solche singular gestörten Probleme die „dynamische Elimination schneller Freiheitsgrade“ mit linear semi-impliziten Einschrittverfahren, also

das Langzeitverhalten nach Abklingen der schnellen Dynamik. Dabei verwenden sie teilweise ähnliche Begriffe wie in dieser Arbeit. Sie gehen einen Schritt weiter als Strehmel, Weiner und Podhaisky (2012) und fordern in (4.49) mit abweichender Notation $y = x_1$, $z = x_2$, $f = f_1$ und $g = f_2$ für „lokale Eigenwerte $[\lambda(g_z)]$ der Ableitungsmatrix g_z “, dass $\operatorname{Re} \lambda(g_z) < 0$ gilt. Hierbei gilt $g = f_2$ und $z = x_2$. „[D]ie Komponenten [...] [münden] dann ‚rasch‘ in die Mannigfaltigkeit $\mathcal{M} = \{(y, z) \in \Omega_0 : g(y, z) = 0\}$, [...] worin wir den Phasenraum mit Ω_0 bezeichnet haben.“ (Deuffhard und Bornemann, 2008, S. 66). Die Mannigfaltig \mathcal{M} nach Deuffhard und Bornemann (ebd.) hat gewisse Ähnlichkeiten mit der langsamen Mannigfaltigkeit μ_ε aus Definition 3.4. Im Unterschied zu μ_ε darf in \mathcal{M} aber keine schnelle Dynamik auftreten, so wie z.B. in Beispiel 3.18 auf Seite 84.

4.3.2 Wahl des Störungsparameters

Es wurde stets davon ausgegangen, dass die zu betrachtenden Probleme als Familie mit einem Parameter ε vorliegen. Startet man mit einem physikalischen Problem ist die Wahl eines geeigneten *gekoppelten Grenzübergangs (distinguished limit)* und damit eines Störungsparameters ε keinesfalls klar. Man denke z.B. an den harmonischen Oszillator, bei dem je nach Kopplung der beteiligten Parameter unterschiedliches physikalisches Verhalten beobachtet werden kann. Vielmehr ist es Aufgabe des Modellierungsprozesses, einen solchen geeigneten Parameter zu finden. Dies ist nicht immer einfach, wie Deuffhard und Bornemann (ebd.) bemerken:

Singuläre Störungsmethoden sind sowohl für das Verständnis als auch für die analytische Untersuchung dynamischer Systeme schön und oft nützlich. In einem praktischen Problem kann es allerdings recht schwierig sein, schnelle und langsame Freiheitsgrade lediglich auf der Basis von Einsicht in das gestellte natur- oder ingenieurwissenschaftliche Problem zu identifizieren und algorithmisch wirksam zu separieren. Damit steht und fällt natürlich auch die Frage nach einem brauchbaren Störungsparameter. Nicht umsonst spricht man in Fachkreisen vom „goldenen“ ε , das es im konkreten Beispiel immer erst zu finden gilt. (ebd., §2.5)

Der familienbasierten Ansatz erfordert also Zusatzwissen um das „goldenen ε “ zu finden und die mathematische Struktur steifen Verhaltens auf die hier vorgestellte Weise zu analysieren. Woher man das im konkreten Fall bekommt kann hier nicht beantwortet werden.

5

FAZIT UND AUSBLICK

Der neue Ansatz, Familien von Differentialgleichungen zu betrachten, ermöglicht es, die mathematische Struktur steifen Verhaltens zu analysieren. Man verlässt damit den Standpunkt, einzelne konkrete Differentialgleichungen zu analysieren. Das jedoch ermöglicht eine qualitative Klassifizierung steifer Lösungsfamilien. In Abschnitt 2.2 wurde diskutiert, dass das Betrachten einer einzelnen Gleichung einer quantitativen Analyse gleich kommt, die eng mit der zur Verfügung stehenden Rechenleistung verbunden ist und die zugrunde liegende mathematische Struktur nur unzureichend berücksichtigt.

Der Familienparameter ε entspricht einem *distinguished limit*, der durch die Kopplung der beteiligten Zeitskalen entsteht. Die auftretenden Zeitskalen sind deutlich sichtbar. Ist für ein lineares Problem λ ein schneller Eigenwert mit $\operatorname{Re} \lambda = -1/\varepsilon$ gegeben, dann ist die Relaxationszeitskala $\tau = t/\varepsilon$. Bei geeignetem Anfangswert gibt es eine Lösungskomponente, die einen Anteil der Form $\exp(-t/\varepsilon)$ hat. Die Integrationszeitskala ist von der asymptotischen Größenordnung $\mathcal{O}_s(1)$.

Steifes Verhalten ist dann zu beobachten, wenn gewisse Lösungskomponenten für $\varepsilon \rightarrow 0$ exponentiell abklingen. Diese Erkenntnis ist nicht neu. Auch in der Literatur wird steifes Verhalten mit dem Abklingen gewisser Lösungskomponenten in Verbindung gebracht. Dort wird Steifheit nicht selten als Eigenschaft eines Problems bezeichnet. Sowohl nach den hier durchgeführten Betrachtungen, als auch nach numerischen Beobachtungen ist das unzureichend. Wichtig ist vor allem, dass Lösungskomponenten exponentiell abklingen. Man könnte dann eine Familie von Differentialgleichungen steif nennen, falls es eine steife Lösungsfamilie gibt. Doch auch das würde der mathematischen Struktur nicht gerecht werden. Neu ist die Betonung darauf, dass es sehr wohl asymptotisch unbeschränkte Lösungskomponenten geben darf, so lange diese nicht exponentiell unbeschränkt sind (Beispiel 3.19 auf Seite 85).

Nach Definition 3.7 wird eine Lösungsfamilie steif genannt, wenn der Abstand der Lösungsfamilie zu einer langsamen Mannigfaltigkeit für $\varepsilon \rightarrow 0$ exponentiell verschwindet. Dabei dient die langsame Mannigfaltigkeit zur strikten Trennung von (exponentiell) schneller und langsamer Dynamik. In der Notation von Satz 3.2 sind die schnellen und langsamen Teilprobleme durch U_s bzw. U_l gegeben. Die langsame Mannigfaltigkeit ist im Allgemeinen kein Fixpunkt des Problems, sondern repräsentiert nur den Teil des Phasenraums, in dem keine exponentiell schnelle Dynamik stattfindet. Die Schnelle Dynamik wie in

Beispiel 3.19 ist erlaubt und mag wegen des in t exponentiell wachsenden Terms überraschen. Das steht in engem Zusammenhang mit der „Vielschichtigkeit“, die ein steifes Problem nach Strehmel, Weiner und Podhaisky (2012) haben kann und die die Charakterisierung so „schwierig“ macht. Durch den asymptotischen Familienansatz lässt sich diese „Vielschichtigkeit“ strukturell besser verstehen. Es ist möglich, die relevanten exponentiell schnellen Lösungskomponenten zu identifizieren. Durch die bloße Betrachtung des Langzeitverhaltens geht das nicht.

Hauptaugenmerk dieser Arbeit liegt auf autonomen linearen Problemen. Hier lassen sich Charakteristika steifen Verhaltens zeigen, die bisher noch nicht beschrieben wurden oder aber nicht präzise von einer Definition abgeleitet werden konnten. Es wird sehr deutlich, dass der Begriff der steifen Differentialgleichung unzureichend ist, steifes Verhalten genau zu beschreiben. Vielmehr ist steifes Verhalten Eigenschaft einer Lösung, oder genauer, eine Eigenschaft einer Lösungsfamilie von einer Familie von Differentialgleichungen zu einem Anfangswert. Aus numerischer Sicht war schon immer klar, dass die Wahl eines Anfangswerts darüber entscheidet, ob die typischen numerischen Schwierigkeiten „steifer Differentialgleichungen“ auftreten.

Aus mathematischer Sicht ist steifes Verhalten von Lösungen eine Eigenschaft, die eng mit dem asymptotischen Verhalten der Matrixexponentiellen verbunden ist. Durch genaues Studium vom $\exp(tA_\varepsilon)$ (nach Davis, 1973) lassen sich die Koeffizienten von A_ε identifizieren, die zu einer exponentiellen Relaxation auf die langsame Mannigfaltigkeit führen, oder diese verhindern. Möglich ist das durch eine geschickte Basiswahl und Sortierung der Basisvektoren, die zu einer Schurform führt. Schurformen sind deshalb eine gute Wahl, da sie durch unitären Basiswechsel erreicht werden können und sich die Eigenschaft steifen Verhaltens unter unitären Basiswechsel erhält. Schurformen wurden in diesem Kontext bisher so noch nicht betrachtet. Higham und Trefethen (1993) betonen schon, dass Diagonalisierung die „Natur einer Differentialgleichung“ signifikant verändern kann. Neu ist aber das explizite Herausarbeiten, dass steifes Verhalten invariant unter unitären Transformationen ist. Für allgemeine Transformationen trifft dies nicht zu (Beispiel 2.7 und 2.6), weshalb es auch nicht ausreicht, sich bei der Untersuchung von steifem Verhalten auf Diagonalmatrizen oder jordansche Normalformen zu beschränken.

Eine Verallgemeinerung von Definition 3.7 auf nichtlineare System bleibt in Vermutung 3.1 offen. Hier ist weitere Forschungsarbeit nötig, um den nichtlinearen Einfluss auf steifes Verhalten zu verstehen. Nach Higham und Trefethen (ebd.) gibt es einen solchen. Die Vermutung liegt nahe, dass auch hier die Mehrskalenasymptotik weiterhelfen kann. Für spezielle nichtlineare Probleme, nämlich die inhomogenen linearen Probleme mit Konstanten Koeffizienten, konnte das

exponentielle Relaxationsverhalten auf eine langsame Mannigfaltigkeit herausgearbeitet werden.

Aus theoretischer Sicht ist es auch interessant, ob Definition 3.7 geeignet ist, Mehrskalenprobleme mit mehr als zwei Zeitskalen sinnvoll auf steifes Verhalten zu untersuchen. Möglicherweise müssen dazu manche Begriffe etwas weiter gefasst werden.

Offen ist außerdem die Frage, ob Definition 3.7 irgendwelche Artefakte versteckt, also Eigenschaften, die nichts mit steifem Verhalten zu tun haben, bzw. die man bisher nicht mit steifem Verhalten in Verbindung gebracht hat. Hierbei stellt sich natürlich die Frage, was man als Referenz für steifes Verhalten ansieht.

Die Erkenntnisse aus dieser Arbeit sind vorwiegend von theoretischem Interesse. Inwiefern der Familienansatz dazu beitragen kann, neue numerische Verfahren für Differentialgleichungen mit steifen Lösungen zu finden, lässt sich zum jetzigen Zeitpunkt nicht sagen. Auch die Frage nach einem eventuellen Zusammenhang zwischen dem Familienparameter ε und Begriffen wie Fehlertoleranz, Genauigkeit oder Rechenleistung muss vorerst unbeantwortet bleiben.

Teil III

ANHANG

ZUSAMMENFASSUNG

Die vorliegende Arbeit besteht aus zwei Teilen, die unabhängig voneinander gelesen werden können. Beide Teile beschäftigen sich mit asymptotisch schnellen Übergängen. Einmal wird das Geschwindigkeitsfeld eines konzentrierten atmosphärischen Wirbels, das auf der synoptischen Längenskala in der Nähe der Wirbelzentrumslinie eine Singularität aufweist, betrachtet und analysiert. Das andere Mal wird das exponentielle Abklingen von Lösungen gewöhnlicher Differentialgleichungen untersucht und eine neue Sichtweise vorgestellt, um steifes Lösungsverhalten zu betrachten.

In Teil *i* wird eine Möglichkeit dargestellt, die komplexen und sehr schwierig zu analysierenden atmosphärischen Vorgänge einheitlich zu vereinfachen. Ziel ist es, reduzierte Gleichungen für unterschiedliche Regime in der Atmosphäre aus den Navier-Stokes-Gleichungen abzuleiten. Dieses Vorgehen wird dann auf einen konzentrierten, beinahe achsensymmetrischen, stark geneigten, baroklinen Wirbel im Gradient-Wind-Regime, der in eine quasigeostrophische Hintergrundströmung mit schwacher vertikaler Scherung eingebettet ist, angewandt.

In Teil *ii* werden gewöhnliche Differentialgleichungen betrachtet. Motiviert durch die Mehrskalenasymptotik, die es durch einen Grenzprozess ermöglicht, Prozesse auf unterschiedlichen Zeitskalen zu erkennen und im Rahmen einer asymptotischen Analyse zu beschreiben, werden nicht einzelne Differentialgleichungen, sondern durch ε parametrisierte Familien von Differentialgleichungen studiert. Dies eröffnet eine neue Sichtweise auf die mathematische Struktur steifen Verhaltens, und zwar unabhängig von numerischen Verfahren oder Diskretisierungen. Steifes Verhalten liegt dann vor, wenn gewisse Lösungskomponenten für $\varepsilon \rightarrow 0$ exponentiell abklingen. Es wird gezeigt von welcher Gestalt diejenigen Mannigfaltigkeiten sind, auf die die Lösungskomponenten relaxieren, und wie diese Komponenten aussehen.

LITERATUR

- Aiken, Richard C., Hrsg. (1985). *Stiff Computation*. Oxford University Press (siehe S. 34, 44, 51, 55, 56).
- Andersson, Christian, Claus Führer und Johan Åkesson (2014). *Assimulo: A Unified Framework for ODE Solvers*. eng. Techn. Ber. LUTFNA-5005-2014. Centre for Mathematical Sciences, Lund University, S. 24 (siehe S. 34).
- Barenblatt, G.I. (1996). *Scaling, Self-similarity, and Intermediate Asymptotics: Dimensional Analysis and Intermediate Asymptotics*. Cambridge Texts in Applied Mathematics. Cambridge University Press. ISBN: 9780521435222. URL: <http://books.google.de/books?id=r-Az53e-MTYC> (siehe S. 9).
- Bluman, G.W. und S. Kumei (1989). *Symmetries and Differential Equations*. Applied Mathematical Sciences Series Bd. 81. Springer-Verlag. ISBN: 9780387969961. URL: <http://books.google.de/books?id=LbfxMQEACAAJ> (siehe S. 8, 9).
- Callegari, A. J. und Lu Ting (1978). „Motion of a Curved Vortex Filament with Decaying Vortical Core and Axial Velocity“. In: *SIAM Journal on Applied Mathematics* 35.1, S. 148–175. URL: <http://www.jstor.org/stable/2101038> (siehe S. 17, 19, 20, 22, 27, 28).
- Curtiss, C.F. und J.O. Hirschfelder (1952). „Integration of stiff equations“. In: *Proc. Natl. Acad. Sci. USA* (siehe S. 33, 37–39, 45).
- Davis, Chandler (1973). „Explicit Functional Calculus“. In: *Linear Algebra And Its Applications* (siehe S. 80–82, 118).
- Deuflhard, Peter und Folkmar Bornemann (2008). *Numerische Mathematik II. Integration gewöhnlicher Differentialgleichungen*. 3. Aufl. de Gruyter (siehe S. 43, 47, 75, 87, 101, 110–112, 115, 116).
- Eckhaus, Wiktor (1979). *Asymptotic analysis of singular perturbations*. Bd. 9. North-Holland Pub. Co. (siehe S. 18, 26, 27, 60).
- Etling, D. (2008). *Theoretische Meteorologie. Eine Einführung*. 3. Aufl. Springer. ISBN: 9783540759799 (siehe S. 21).
- Hairer, Ernst und Gerhard Wanner (2010). *Solving Ordinary Differential Equations II*. Springer (siehe S. 33, 34, 41, 72).
- Higham, Desmond J. (2008). *Functions of Matrices: Theory and Computation*. Philadelphia, PA, USA: Society for Industrial und Applied Mathematics, S. xx+425. ISBN: 978-0-898716-46-7 (siehe S. 80, 81).
- Higham, Desmond J. und Lloyd N. Trefethen (1993). „Stiffness of ODEs“. In: *BIT Numerical Mathematics* 33 (2). 10.1007/BF01989751, S. 285–303. ISSN: 0006-3835. URL: <http://dx.doi.org/10.1007/BF01989751> (siehe S. 56, 57, 85, 86, 88–91, 102–105, 107–109, 118).
- Holmes, M. H. (1995). *Introduction to Perturbation Methods*. Springer (siehe S. 15).

- Howison, Sam (2005). *Practical applied mathematics: modelling, analysis, approximation*. Cambridge texts in applied mathematics. Cambridge University Press. ISBN: 9780521842747 (siehe S. 9).
- Kevorkian, J. und J.D. Cole (1981). *Perturbation Methods in Applied Mathematics*. 1st. Bd. 34. Applied Mathematical Science. Springer (siehe S. 11).
- (1996). *Multiple Scale and Singular Perturbation Methods*. Springer. ISBN: 0-387-94202-5 (siehe S. 55).
- Klein, Rupert (2010). „Scale-Dependent Models for Atmospheric Flows“. English. In: *ANNUAL REVIEW OF FLUID MECHANICS* 42, 249–274. ISSN: 0066-4189. DOI: {10.1146/annurev-fluid-121108-145537} (siehe S. 5, 6, 9, 11, 15–18).
- (2008). „An unified approach to meteorological modelling based on multiple-scales asymptotics“. In: *Advances in Geosciences* 15, S. 23–33. DOI: 10.5194/adgeo-15-23-2008. URL: <http://www.adv-geosci.net/15/23/2008/> (siehe S. 5, 9, 13, 17, 21).
- Klein, Rupert, N. Botta u. a. (2001). „Asymptotic adaptive methods for multi-scale problems in fluid mechanics“. In: *Journal of Engineering Mathematics* 39.1, S. 261–343. URL: <http://publications.mi.fu-berlin.de/505/> (siehe S. 9).
- Klein, Rupert, Stefan Vater u. a. (2011). „Multiple scales methods in meteorology“. In: *Asymptotic Methods in Fluid Mechanics: Survey and Recent Advances*. Hrsg. von H. Steinrück. Bd. 523. CISM International Centre for Mechanical Sciences. Springer Verlag, S. 127–196. URL: <http://publications.mi.fu-berlin.de/803/> (siehe S. 11).
- Kreiss, Heinz-Otto (1978). „Difference methods for stiff ordinary differential equations“. In: *SIAM Journal on Numerical Analysis* 15.1, S. 21–58 (siehe S. 89).
- Kuchling, H. (2007). *Taschenbuch der Physik*. Fachbuchverl. Leipzig im Carl-Hanser-Verlag. ISBN: 9783446410282 (siehe S. 7).
- Lam, S.H. und D.A. Goussis (1989). „Understanding complex chemical kinetics with computational singular perturbation“. In: *Symposium (International) on Combustion* 22.1, S. 931–941. ISSN: 0082-0784. DOI: [http://dx.doi.org/10.1016/S0082-0784\(89\)80102-X](http://dx.doi.org/10.1016/S0082-0784(89)80102-X). URL: <http://www.sciencedirect.com/science/article/pii/S008207848980102X> (siehe S. 31).
- Päschke, Eileen u. a. (2012). „Motion and structure of atmospheric mesoscale baroclinic vortices: dry air and weak environmental shear“. In: *Journal of Fluid Mechanics* 701, S. 137–170. URL: <http://publications.mi.fu-berlin.de/936/> (siehe S. 5, 17–28).
- Pedlosky, Joseph (1987). *Geophysical Fluid Dynamics*. 2. Aufl. Springer-Verlag New York Inc. (siehe S. 21).
- Quarteroni, A., R. Sacco und F. Saleri (2002). *Numerische Mathematik 2*. Springer (siehe S. 56, 103).

- Schmitt, Bernhard (2012). „Dynamische Systeme“. URL: <http://www.mathematik.uni-marburg.de/~schmitt/dyn/dynsys.pdf> (siehe S. 49, 70, 105).
- Simeon, Bernd (2003). „Numerik Gewöhnlicher Differentialgleichungen“. Skriptum zur Vorlesung im Wintersemester 2003/04, TU München (siehe S. 38, 55).
- Strehmel, Karl, Rüdiger Weiner und Helmut Podhaisky (2012). *Numerik gewöhnlicher Differentialgleichungen*. Springer (siehe S. 33, 52, 55, 56, 74, 92, 93, 103–107, 115, 116, 118).
- Ström, Torsten (1975). „On logarithmic norms“. In: *SIAM Journal on Numerical Analysis* 12.5, S. 741–753 (siehe S. 105).
- Trefethen, Lloyd N. und Mark Embree (2005). *Spectra and pseudospectra: the behavior of nonnormal matrices and operators*. Princeton University Press (siehe S. 108, 109).
- van Dyke, Milton (1964). *Perturbation Methods in Fluid Mechanics*. New York: Academic Press (siehe S. 18).
- (1975). *Perturbation Methods in Fluid Mechanics*. Second Printing 1978. The Parabolic Press (siehe S. 26).