

**FINITE DIFFERENZENVERFAHREN
FÜR HYPERBOLISCHE SYSTEME MIT
ABSORBIERENDEN RANDBEDINGUNGEN**

DIPLOMARBEIT

bei Herrn Prof. Dr. Peter Markowich
am Fachbereich Mathematik
der Technischen Universität Berlin
vorgelegt von

MATTHIAS EHRHARDT

Berlin, Juni 1995

Die selbständige und eigenhändige Anfertigung versichert an Eidesstatt
Berlin, den

Inhaltsverzeichnis

Einleitung	iii
Kapitel 1. Absorbierende Randbedingungen für hyperbolische Systeme	1
1.1. Anfangsrandwertprobleme für lineare hyperbolische Systeme	1
1.1.1. Systeme in einer Ortsdimension	2
1.1.2. Systeme in zwei Ortsdimensionen	5
1.2. Herleitung der absorbierenden Randbedingungen	6
1.3. Sachgemäße Gestelltheit in einer Ortsdimension	9
1.3.1. Die Abschätzung der Lösung	10
1.3.2. Die Existenz und Eindeutigkeit der Lösung	11
Kapitel 2. Numerische Betrachtung des linearen Falls in 1D	14
2.1. Das numerische Schema für 1D	14
2.2. Die Stabilität des Differenzschemas in einer Ortsdimension	16
2.2.1. Abschätzung der Lösung der Differentialgleichung	19
2.2.2. Die Definition der Stabilität	22
2.2.3. Das gestörte Schema	22
2.2.4. Die Stabilität des gestörten Schemas	23
2.3. Numerische Ergebnisse	26
2.3.1. Das Testbeispiel	26
2.3.2. Numerische Verifikation der Stabilität	29
2.3.3. Die Absorptionsqualität der absorbierenden Randbedingungen	33
Kapitel 3. Numerische Betrachtung des linearen Falls in 2D	36
3.1. Das numerische Schema für 2D	36
3.2. Die linearisierten Shallow-Water Gleichungen in 2D	38
3.2.1. Die Diagonalisierung der Systemmatrix	39
3.2.2. Die absorbierenden Randbedingungen	39

3.3. Numerische Ergebnisse	41
3.3.1. Das Testbeispiel	41
3.3.2. Numerische Untersuchung der Stabilität	46
3.3.3. Die Absorptionsqualität der absorbierenden Randbedingungen	49
Kapitel 4. Numerische Betrachtung des nichtlinearen Falls in 1D	54
4.1. Das numerische Schema für 1D	54
4.2. Das hydrodynamische Modell der Halbleitergleichungen	56
4.2.1. Die Diagonalisierung der Systemmatrix	57
4.2.2. Die Randbedingungen	58
4.3. Numerische Ergebnisse	60
4.3.1. Das Testbeispiel	60
Kapitel 5. Numerische Betrachtung des nichtlinearen Falls in 2D	64
5.1. Das numerische Schema für 2D	64
5.2. Die Shallow-Water Gleichungen in zwei Ortsdimensionen	67
5.2.1. Die Randbedingungen	67
5.3. Numerische Ergebnisse	69
5.3.1. Das Testbeispiel	69
5.3.2. Numerische Untersuchung der Stabilität	73
5.3.3. Die Absorptionsqualität der absorbierenden Randbedingungen	75
Anhang A. Glossar	81
1.1. Notation	81
1.2. Mathematische Größen	81
1.3. Physikalische Größen	82
Literaturverzeichnis	83

Einleitung

Diese Arbeit beschäftigt sich mit der numerischen Approximation von hyperbolischen Differentialgleichungen, die auf einem unbeschränkten Gebiet des Ortsraumes (meist \mathbb{R}^N) gestellt sind. Bei der numerischen Diskretisierung dieses *Ganzraumproblems* stößt man auf das Problem, daß nur ein begrenzter Speicherplatz zur Verfügung steht und man somit nur auf einem beschränkten Teilgebiet rechnen kann.

In einigen Fällen kann man sich durch eine Koordinatentransformation mit Hilfe von *konformen Abbildungen* helfen, indem man das ursprüngliche Ganzraumproblem in ein Problem auf einem beschränkten Gebiet überführt, wobei allerdings die Differentialgleichungen meist sehr kompliziert werden [35]. Zudem versagt diese Technik der konformen Abbildungen, wenn die Lösung im Unendlichen oszillierend wird und ist für viele physikalische Probleme ungeeignet [16].

Eine anderer möglicher Weg ist der, daß man das Gebiet durch künstlich eingeführte Ränder beschränkt, ohne Änderungen in den Differentialgleichungen zu machen. Auf diesen künstlichen Rändern schreibt man sog. *absorbierende Randbedingungen* vor, die derart bestimmt werden müssen, daß sie die Lösung auf dem beschränkten Gebiet so beeinflussen, daß die Ganzraumlösung möglichst gut approximiert wird. Dabei wird die Approximationsgüte umso höher sein, je weniger die aus dem Innern des beschränkten Gebietes herauslaufenden Komponenten (*Ausströmkomponenten*) am künstlichen Rand reflektiert werden. Denn, anschaulich gesprochen, sollen die ausströmenden Wellen möglichst wenig vom künstlichen Rand „spüren“. Insbesondere sollen die Amplituden der von den künstlichen Rändern reflektierten Wellen möglichst klein sein [10].

Eine notwendige Bedingung, um mit dieser Vorgehensweise die Ganzraumlösung numerisch annähern zu können, ist sicherlich, daß die partielle Differentialgleichung außerhalb des beschränkten Gebietes die Lösung innerhalb dieses Teilgebietes nur vernachlässigbar gering beeinflußt [17]. Auch müssen die künstlichen Ränder sinnvoller Weise so gewählt werden, daß der Träger der Anfangsverteilung vollständig im beschränkten Gebiet liegt.

Es lassen sich in der Literatur zahlreiche Beispiele für die Anwendung von absorbierenden Randbedingungen finden, z. B. in der Elektrodynamik [4], in der Fluidodynamik [3],[32] und in der Geologie [5], [6], [21], [30].

In der Meteorologie werden sie bei örtlichen Wettervorhersagen verwandt [9], weil das zugrundeliegende Gebiet, man denke etwa an die Erdoberfläche, gemessen an den Rechenkapazitäten zu groß ist, um das Problem in der vorgegebenen Zeit zu lösen. Hierbei senkt man den Aufwand nicht durch eine Vergrößerung des Gitters, weil damit die Genauigkeit der Vorhersage sinken würde, sondern verkleinert das Gebiet durch künstliche Ränder.

Die Vorgehensweise zur Lösung eines solchen Problems mit absorbierenden Randbedingungen besteht aus zwei wesentlichen Schritten:

- (1) Herleitung und Untersuchung der absorbierenden Randbedingungen
- (2) Diskretisierung und numerische Lösung des Problems

Demzufolge gliedert sich die vorliegende Arbeit in zwei Teile, einem analytischen und einem numerischen.

Im analytischen Abschnitt wird zunächst mit Hilfe der Pseudodifferentialoperator-Theorie [36] eine *Hierarchie von absorbierenden Randbedingungen* für lineare hyperbolische Systeme erster Ordnung konstruiert. Unsere Vorgehensweise wird sich dabei an dem grundlegenden Artikel von Engquist und Majda [10] orientieren. Anschließend werden wir im eindimensionalen Fall untersuchen, inwiefern die so erhaltenen Anfangsrandwertprobleme *sachgemäß gestellt* sind, d. h. ob eine eindeutige Lösung des Problems existiert und diese stetig von den Ausgangsdaten abhängt. Für hyperbolische Systeme in zwei Ortsdimensionen wurde bereits gezeigt, daß absorbierende Randbedingungen teilweise auf nicht-sachgemäß gestellte Probleme führen [10].

Im numerischen Teil der Arbeit werden die absorbierenden Randbedingungen geeignet diskretisiert und gezeigt, daß das resultierende *Lax-Wendroff-Differenzschema* für das Anfangsrandwertproblem in 1D *stabil* ist. Die Stabilität eines numerischen Verfahrens ist das diskrete Analogon zur sachgemäßen Gestelltheit eines Problems im Kontinuierlichen, was im wesentlichen die Unempfindlichkeit bezüglich Rundungs- und Eingabefehler bedeutet. Bei der Untersuchung des numerischen Verhaltens der hergeleiteten Randbedingungen stehen die *diskreten Absorptionseigenschaften*, d. h. die Absorptionsqualität der Diskretisierung der kontinuierlichen absorbierenden Randbedingungen, im Vordergrund. Dabei werden die absorbierenden Randbedingungen mit homogenen Dirichlet-Randbedingungen verglichen. Die auftretenden *numerischen Reflektionen* werden durch ein Vergleich mit einer Lösung auf einem wesentlich größeren Gitter ermittelt. In den letzten zwei Kapiteln werden *nichtlineare* hyperbolische Systeme in 1 und 2 Dimensionen numerisch untersucht. Die Programme wurden in MATLAB geschrieben und die Testläufe auf einer IBM RS/6000, 590 ausgeführt.

An dieser Stelle möchte ich mich für das Zustandekommen der Diplomarbeit bei Herrn Dr. Anton Arnold für die sehr gute Betreuung bedanken.

KAPITEL 1

Absorbierende Randbedingungen für lineare hyperbolische Systeme

1.1. Anfangsrandwertprobleme für lineare hyperbolische Systeme

Wir wollen hier die wichtigsten Begriffe für lineare hyperbolische Systeme erster Ordnung einführen. Differentialgleichungen höherer Ordnung können mit Hilfe von Standardmethoden auf Systeme erster Ordnung reduziert werden [20], [40]. Beispiele für hyperbolische Gleichungen erster Ordnung sind die Maxwell Gleichungen, die (linearierten) Shallow-Water Gleichungen und die klassischen hydrodynamischen Gleichungen in der Halbleitersimulation (ohne Wärmeleitungsterm). Um eine eindeutige Lösung des Problems zu bestimmen, muß man eine Anfangswertverteilung vorschreiben, und im allgemeinen wird es auch notwendig sein, Randbedingungen an die Lösung zu stellen. Das Resultat ist ein *Anfangsrandwertproblem* (kurz: ARWP). Manchmal können die korrekten Randbedingungen leicht aus physikalischen Überlegungen heraus abgeleitet werden. Wird zum Beispiel eine Strömung durch eine feste Wand begrenzt, so muß sie dort tangential verlaufen, und man setzt somit am Rand die orthogonale Komponente der Strömungsgeschwindigkeit gleich Null. (Ist die Strömung reibungsbehaftet, so muß auch die Tangentialkomponente verschwinden.) Bei künstlichen Rändern, wie wir sie betrachten werden, ist die Wahl der Randbedingungen nicht so offensichtlich.

Im allgemeinen können Randbedingungen für eine hyperbolische Gleichung nicht willkürlich gewählt werden. Dies kann am einfachsten anhand von Systemen erster Ordnung in einer Ortsdimension, die wir im folgenden Abschnitt 1.1.1 betrachten, verdeutlicht werden. Dort werden wir zeigen, daß man einzelne Komponenten der Lösung als *fortschreitende Wellen* ansehen kann. Es wird sich herausstellen, daß jede akzeptable Randbedingung einerseits das Verhalten von einströmenden Komponenten vorschreiben muß und andererseits keine Bedingung an die ausströmenden Komponenten stellen darf. Die einströmenden (ausströmenden) Komponenten entsprechen dabei den in das Gebiet einlaufenden (auslaufenden) *Charakteristiken*.

In zwei Raumdimensionen ist die Situation komplizierter; man kann nicht mehr so leicht einströmende und ausströmende Komponenten einer Lösung identifizieren. Speziell können Wellen auftreten, die sich tangential zum Rand bewegen und es stellt sich die Frage, ob und wie man Randbedingungen an diese Komponenten stellen kann.

1.1.1. Systeme in einer Ortsdimension. Wir betrachten Systeme der Form

$$(1.1a) \quad U_t + A(x, t)U_x + C(x, t)U = F(x, t)$$

auf dem Streifen $0 \leq x \leq L$, $t \geq 0$ mit der Anfangsbedingung

$$(1.1b) \quad U(x, 0) = f(x), \quad 0 \leq x \leq L$$

und mit Randbedingungen an $x = 0$, $x = L$.

Die Koeffizienten $A(x, t), C(x, t) \in \mathbb{R}^{N \times N}$, der Quellterm $F(x, t) \in \mathbb{R}^N$ und die Anfangsverteilung $f(x) \in \mathbb{R}^N$ seien C^∞ -glatt. Wir machen die folgende Annahme.

ANNAHME 1.1 (HYPERBOLISCHES SYSTEM [38]). Das System (1.1a) ist *hyperbolisch*, d. h. $A(x, t)$ hat für $x \in [0, L]$, $t \geq 0$ reelle Eigenwerte $\lambda_j(x, t)$ und einen vollständigen Satz von reellen Eigenvektoren. Dies ermöglicht die Vereinfachung des Systems, da dann $A = A(x, t)$ *diagonalisierbar* ist, d. h. es gibt eine reguläre Matrix $T = T(x, t)$, so daß

$$(1.2) \quad T^{-1}AT = \Lambda = \text{diag}(\lambda_1, \dots, \lambda_N), \quad \lambda_j = \lambda_j(x, t), \quad j = 1, 2, \dots, N$$

ist. Ferner fordern wir bei Systemen mit *variablen Koeffizienten*, daß die Matrixnormen von $T(x, t)$, $T^{-1}(x, t)$ für $x \in [0, L]$, $t \geq 0$ beschränkt sind.

Die Matrix T enthält die zugehörigen Rechtseigenvektoren als Spalten und die Matrix T^{-1} die Linkseigenvektoren als Zeilen. Man erhält somit

$$\begin{aligned} (T^{-1}U)_t &= T^{-1}U_t + T_t^{-1}U \\ &\stackrel{(1.1a)}{=} -T^{-1}AU_x - T^{-1}CU + T^{-1}F + T_t^{-1}U \\ &= -T^{-1}AT(T^{-1}U)_x + T^{-1}ATT_x^{-1}U - T^{-1}CU + T^{-1}F + T_t^{-1}U \\ &= -\Lambda(T^{-1}U)_x - (T^{-1}C - T_t^{-1} - T^{-1}ATT_x^{-1})U + T^{-1}F. \end{aligned}$$

Mit Einführung der neuen Variablen (sog. *charakteristischen Variablen*)

$$V(x, t) := T^{-1}(x, t)U(x, t)$$

transformiert sich das System (1.1a) in die *charakteristische Form*

$$(1.3) \quad V_t + \Lambda(x, t)V_x + \tilde{C}(x, t)V = \tilde{F}(x, t),$$

wobei $\tilde{C} = (T^{-1}C - T_t^{-1} - T^{-1}ATT_x^{-1})T = T^{-1}CT - T_t^{-1}T + T^{-1}AT_x$ und $\tilde{F} = T^{-1}F$ ist.

BEMERKUNG 1.1. Ohne Einschränkung kann im folgenden angenommen werden, daß das System (1.1a) bereits in charakteristischen Variablen formuliert ist, d. h. die Systemmatrix A besitzt bereits Diagonalgestalt: $A = \Lambda$. Ansonsten geht man zu den transformierten Variablen $V(x, t) := T^{-1}(x, t)U(x, t)$ über.

ANNAHME 1.2 (KONSTANTE PARTITION AM RAND [23, Seite 254]).

$$(1.4) \quad \lambda_j(0, t) \quad \text{und} \quad \lambda_j(L, t), \quad j = 1, 2, \dots, N$$

wechseln als Funktionen der Zeit nicht ihr Vorzeichen, d. h. jede Funktion (1.4) ist entweder > 0 für alle t , $= 0$ für alle t oder < 0 für alle Zeiten t .

Der Fall von N skalaren Gleichungen. Um das System (1.1a) zu diskutieren, benutzen wir die *Methode der Charakteristiken* und beginnen mit dem *entkoppelten Fall* $C = 0$. Dann reduziert sich das System zu N unabhängigen skalaren hyperbolischen Gleichungen. Weiterhin nehmen wir $F = 0$ an. Es ergibt sich

$$(1.5) \quad \frac{\partial U_j}{\partial t} + \lambda_j \frac{\partial U_j}{\partial x} = 0, \quad j = 1, 2, \dots, N.$$

Diese Gleichungen entsprechen jeweils gewöhnlichen Differentialgleichungen für die U_j entlang der *charakteristischen Kurve* $(x(t), t)$, die durch $dx/dt = \lambda_j(x, t)$ definiert ist, denn (1.5) ist gleichbedeutend mit

$$\frac{d}{dt} (U_j(x(t), t)) = 0 \quad \text{für} \quad \frac{dx}{dt} = \lambda_j(x, t).$$

Die Komponenten U_j sind somit konstant entlang der zugehörigen Charakteristiken und können als Wellen, die sich mit den *charakteristischen Geschwindigkeiten* $dx/dt = \lambda_j$ fortbewegen, betrachtet werden.

i) Der Fall $\lambda_j = \text{konst.}$ Wir nehmen zunächst an, daß

$$\lambda_j(x, t) = \lambda_j = \text{konst.}, \quad j = 1, 2, \dots, N$$

gilt. In diesem Fall sind die Charakteristiken Geraden und damit gilt

$$U_j(x, t) = f_j(x - \lambda_j t), \quad 0 \leq x - \lambda_j t \leq L.$$

Wir werden nun folgende Notation verwenden:

NOTATION. U^+ , U^0 , U^- bestehen aus denjenigen Variablen $U_j = U_j(x, t)$ mit Index j , für die $\lambda_j > 0$, $\lambda_j = 0$, $\lambda_j < 0$ gilt. Dieses *Indizierungskonzept* der Komponenten werden wir auch bei anderen Größen, wie z. B. f , verwenden und auf die Indizierung von Untermatrizen über Zeilen- und Spaltenindizes ausdehnen.

Offensichtlich wird U^0 allein durch die Anfangsdaten bestimmt:

$$U^0(x, t) = f^0(x), \quad 0 \leq x \leq L, \quad t \geq 0.$$

Randbedingungen für die charakteristischen Variablen U^0 mit Geschwindigkeiten $\lambda_j = 0$ sind also weder notwendig noch erlaubt [37].

Ist $\lambda_j > 0$, so verlaufen die Charakteristiken von links nach rechts wie es in Abbildung 1.1 zu sehen ist und neben den Anfangsdaten braucht man auch Randdaten an $x = 0$, dem sog. *Einströmrand*. Es ist nicht erlaubt, an $x = L$, dem sog. *Ausströmrand* Randbedingungen zu stellen, denn diese könnten zur Anfangsbedingung im Widerspruch stehen und damit die Existenz einer Lösung verhindern. Für $\lambda_j < 0$ sind die Rollen und Bezeichnungen der beiden Ränder vertauscht. Daher formulieren wir

$$U^+(0, t) = g_0(t), \quad U^-(L, t) = g_L(t), \quad t > 0,$$

d. h. wir stellen eine Bedingung an die einfallenden charakteristischen Variablen an jedem Rand. Diese Randbedingungen kann man verallgemeinern zu

$$(1.6) \quad \begin{aligned} U^+(0, t) &= S_0(t)U^-(0, t) + g_0(t) \\ U^-(L, t) &= S_L(t)U^+(L, t) + g_L(t), \quad t > 0, \end{aligned}$$

wobei $S_0(t)$, $S_L(t)$ Matrizen geeigneter Dimension sind. Man sagt, daß die einströmenden charakteristischen Variablen durch Terme der ausströmenden ausgedrückt werden.

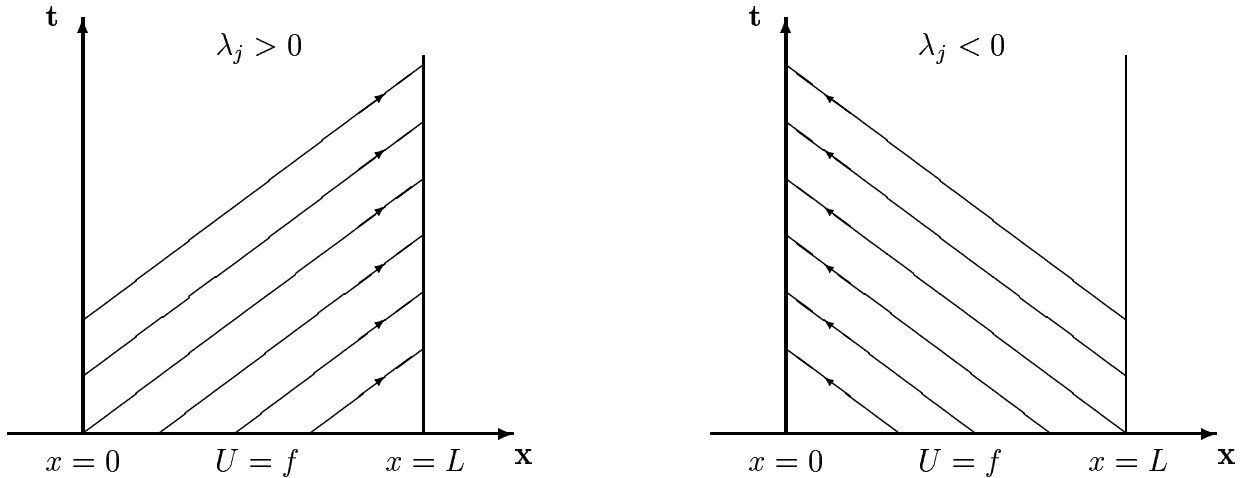


ABBILDUNG 1.1. Charakteristiken für Gleichung (1.5)

Entlang derjenigen Charakteristiken, die in den Eckpunkten $(x, t) = (0, 0)$, $(x, t) = (L, 0)$ beginnen, wird die Lösung eine Sprungunstetigkeit haben, falls die Anfangsvorgabe f mit den Randdaten g_0, g_L *nicht kompatibel*, d. h. $f(0) \neq g_0(0)$, $f(L) \neq g_L(0)$, ist. Allgemeiner gilt, wenn $f(x)$ eine Singularität (eine Unstetigkeit in f oder einer Ableitung von f) in einem Punkt x_0 hat, so wird die resultierende Lösung $U(x, t)$ eine Singularität der gleichen Ordnung entlang der charakteristischen Kurve durch x_0 besitzen, und entlang der Charakteristiken durch glatte Abschnitte der Anfangsverteilung wird die Lösung glatt bleiben. Dies ist eine fundamentale Eigenschaft von *linearen* hyperbolischen Gleichungen: Singularitäten bewegen sich nur entlang der Charakteristiken. Wir nehmen daher an, daß die Anfangsdaten mit den Randdaten *kompatibel* sind, worunter wir fortan nicht nur deren Übereinstimmung in den Funktionswerten, sondern auch in den Werten beliebig hoher Ableitungen verstehen wollen:

ANNAHME 1.3 (KOMPATIBILITÄT DER DATEN). Die Randdaten g_0, g_L seien C^∞ -glatt und mit der Anfangsverteilung f im Funktionswert, d. h. $f(0) = g_0(0)$, $f(L) = g_L(0)$, und auch in beliebig hohen Ableitungen kompatibel.

Die Annahme ist z. B. erfüllt, wenn die Daten f, g_0 und g_L in der Nähe der Eckpunkte $(x, t) = (0, 0)$, $(x, t) = (L, 0)$ identisch verschwinden.

ii) Der Fall $\lambda_j = \lambda_j(x, t)$. In diesem Fall sind die Charakteristiken i. a. keine Geraden. Wir nehmen an, daß alle Eigenwerte am Rand ungleich Null sind, d. h. der Rand ist *nicht charakteristisch*. Wechselt ein Eigenwert $\lambda_j(x, t)$ auf $0 < x < L$ sein Vorzeichen, so kann die Komponente U_j z. B. an $x = 0$ zu einer positiven und an $x = L$ zu einer negativen charakteristischen Geschwindigkeit gehören. Auch in diesem Fall werden wir die Notation von Seite 3 verwenden, um die Komponenten U_j an den Randpunkten zu referenzieren.

BEMERKUNG 1.2 (GEKOPPELTER FALL). Für $C \neq 0$ ist im allgemeinen das resultierende System von Gleichungen in (1.1a) *gekoppelt*, aber nur in den nichtdifferenzierten Termen. Dieser Term niedriger Ordnung, CU , bewirkt zwar Wachstum, Abnahme, Oszillationen oder Dämpfung in der Lösung, allerdings beeinträchtigt er nicht die *Partition* in einströmende und ausströmende Komponenten, die nur vom Term führender Ordnung abhängt [19], [38], d. h. die Randbedingungen müssen die gleichen Kriterien wie zuvor erfüllen.

Für ein hyperbolisches Problem mit den Randbedingungen (1.6) gilt folgender Satz.

SATZ 1.1 (EXISTENZ EINER C^∞ -LÖSUNG UND IHRE ABSCHÄTZUNG [23, Satz 7.6.4]). *Der Rand sei nicht charakteristisch und die Daten F, f, g_0, g_L seien kompatibel an $t = 0$. Dann hat das hyperbolische Anfangsrandwertproblem (1.1), (1.6) eine eindeutige Lösung. Die Lösung U ist eine C^∞ -Funktion und für jedes endliche Zeitintervall $0 \leq t \leq T$ existiert eine Konstante c_T , so daß gilt*

$$(1.7) \quad \|U(\cdot, t)\|_2^2 + \int_0^t (|U(0, \tau)|^2 + |U(L, \tau)|^2) d\tau \\ \leq c_T \left[\|f\|_2^2 + \int_0^t (|g_0(\tau)|^2 + |g_L(\tau)|^2 + \|F(\cdot, t)\|_2^2) d\tau \right]$$

für $0 \leq t \leq T$. Die Konstante c_T ist unabhängig von den Daten F, f, g_0, g_L .

BEMERKUNG 1.3 (WENIGER GLATTE DATEN). Die C^∞ -Glattheit der Daten wird nur der Einfachheit halber gefordert; alles geht mit sehr geringer Regularität der Daten (vgl. Fußnote auf Seite 9 und Bemerkung 1.10).

Damit ist dieses Problem *sachgemäß gestellt*. Die absorbierenden Randbedingungen, die wir in Abschnitt 1.2 formulieren werden, haben allerdings nicht die Form (1.6). Für die sachgemäße Gestelltheit des Problems (1.1) mit ARBen in 1D wird somit ein eigener Beweis nötig, den wir in Abschnitt 1.3 unter Verwendung von Satz 1.1 führen werden.

1.1.2. Systeme in zwei Ortsdimensionen. Wir betrachten Systeme der Form

$$(1.8a) \quad U_t + A(x, y, t)U_x + B(x, y, t)U_y + C(x, y, t)U = F(x, y, t)$$

auf $0 \leq x \leq L, -\infty < y < \infty, t \geq 0$ mit der Anfangsbedingung

$$(1.8b) \quad U(x, y, 0) = f(x, y), \quad 0 \leq x \leq L, -\infty < y < \infty$$

und mit Randbedingungen an $x = 0, x = L$.

Die auftretenden Koeffizienten und die Anfangsverteilung seien wie in einer Ortsdimension C^∞ -glatt.

In 2D ist es im allgemeinen nicht möglich, Lösungen mit Hilfe von Charakteristiken zu untersuchen.

Wir wollen hier nur zwei wichtige Begriffe einführen:

DEFINITION 1.1 (STRENG HYPERBOLISCHE SYSTEME [19]). Das System (1.8a) heißt *streng hyperbolisch*, wenn die Matrizen $k_1A + k_2B$ für alle $(k_1, k_2) \neq (0, 0)$ voneinander verschiedene reelle Eigenwerte besitzen.

DEFINITION 1.2 (SYMMETRISIERBARE HYPERBOLISCHE SYSTEME [19]). Das System (1.8a) heißt *symmetrisierbar hyperbolisch*, wenn die Matrizen A, B für alle Argumente $(x, y, t) \in [0, L] \times \mathbb{R} \times \mathbb{R}_0^+$ symmetrisch sind.

BEMERKUNG 1.4. Die obige Definition 1.2 läßt sich auf Systeme verallgemeinern, bei denen die Zeitableitung einen Koeffizienten ungleich der Identität besitzt. Zusätzlich muß dieser Koeffizient dann positiv definit sein.

1.2. Herleitung der absorbierenden Randbedingungen

In diesem Abschnitt wollen wir die Herleitung der absorbierenden Randbedingungen (kurz: ARB) an $x = 0$ skizzieren und orientieren uns dabei an der bereits klassischen Arbeit von Engquist und Majda [10, Abschnitt 2].

Wir betrachten *streng hyperbolische* Systeme erster Ordnung mit variablen Koeffizienten

$$(1.9) \quad U_t + \Lambda(x, y, t)U_x + B(x, y)U_y + C(x, y, t)U = 0$$

auf der Halbebene $x \geq 0$, $-\infty < y < \infty$, $t \geq 0$, wobei Λ, B symmetrische $N \times N$ -Matrizen sind. Wir nehmen an, daß Λ für alle $(x, y, t) \in \mathbb{R}_0^+ \times \mathbb{R} \times \mathbb{R}_0^+$ regulär ist und können wie in der Bemerkung 1.1 ohne Einschränkung Λ in der Form

$$\Lambda = \text{diag}(\lambda_1, \dots, \lambda_m, \lambda_{m+1}, \dots, \lambda_N)$$

mit $\lambda_j > 0$ für $1 \leq j \leq m$ und $\lambda_j < 0$ für $m+1 \leq j \leq N$ schreiben. Gemäß unserer Konvention von Seite 3 schreiben wir daher auch

$$\Lambda^+ = \text{diag}(\lambda_1, \dots, \lambda_m), \quad \Lambda^- = \text{diag}(\lambda_{m+1}, \dots, \lambda_N).$$

BEMERKUNG 1.5. Aufgrund der strengen Hyperbolizität ist $\lambda_j \neq \lambda_k$ für $j \neq k$.

Wir formulieren das System (1.9) um zu

$$U_x = -\Lambda^{-1}U_t - \Lambda^{-1}BU_y + \tilde{C}U$$

mit $\tilde{C} = -\Lambda^{-1}C$ und definieren $M = M(\xi, \omega)$ folgendermaßen

$$M(\xi, \omega) := -i\Lambda^{-1}\xi - i\Lambda^{-1}B\omega.$$

Dann ist

$$(1.10) \quad U_x = M \left(\frac{\partial_t}{i}, \frac{\partial_y}{i} \right) U + \tilde{C}U.$$

Im wesentlichen Schritt entkoppelt man die positiven λ von den negativen, denn nach einer *Konstruktion von Taylor* in [40, Kapitel 9] und [39] gibt es eine glatte Matrix $V = V(\xi, \omega, x, y, t)$, die für alle (ξ, ω) mit $|\omega/\xi| + |\omega| < c_0$ für ein $c_0 > 0$ invertierbar ist, so daß die Gleichung (1.10) mit Hilfe der Transformation

$$(1.11) \quad \begin{aligned} W &= V \left(\frac{\partial_t}{i}, \frac{\partial_y}{i}, x, y, t \right) U \quad \text{auf} \\ W_x &= \begin{pmatrix} \Omega_{11} & 0 \\ \Omega_{21} & \Omega_{22} \end{pmatrix} W \end{aligned}$$

transformiert werden kann. Dabei ist $\Omega_{11} = \Omega_{11}(\partial_t/i, \partial_y/i, x, y, t)$ ein $m \times m$ -Pseudodifferentialoperator der Ordnung 1, d. h. $\Omega_{11} \in \text{PS}(1)$. Ω_{21} und Ω_{22} sind $(N-m) \times m$ bzw. $(N-m) \times (N-m)$ -Pseudodifferentialoperatoren der Ordnung 1.

Besitzt die Differentialgleichung erst einmal die *entkoppelte Form* (1.11), so kann man die ARB leicht formulieren, denn Ω_{11} enthält gerade die positiven Eigenwerte, die zu den Einströmkomponenten an $x = 0$ gehören (man vergleiche dazu die Ausführungen in Abschnitt 1.1.1). Daher muß die ARB diese Komponenten an $x = 0$ eliminieren, also

$$(1.12) \quad W^+ \Big|_{x=0} = (VU)^+ \Big|_{x=0} = 0 \quad \text{mit} \quad W^+ = (W_1, \dots, W_m)$$

d. h. der reflektierte Anteil W^+ an $x = 0$ wird gleich Null gesetzt.

Nach Taylors Konstruktion läßt sich $V(\xi, \omega, x, y, t)$ wie folgt *asymptotisch entwickeln*

$$(1.13) \quad \begin{aligned} V(\xi, \omega, x, y, t) &\cong V_0(\xi, \omega, x, y, t) + \xi^{-1}V_{-1}(\xi, \omega, x, y, t) \\ &+ \xi^{-2}V_{-2}(\xi, \omega, x, y, t) + \dots, \end{aligned}$$

wobei jedes V_j homogen vom Grade Null in (ξ, ω) ist.

Aufgrund der strengen Hyperbolizität ist $M(1, 0)$ eine Diagonalmatrix mit paarweise verschiedenen Eigenwerten, und somit gibt es eine Konstante c_0 , so daß $M(\xi, \omega)$ in einer konischen Umgebung um $(1, 0)$, also für $|\omega/\xi| + |\omega| < c_0$, voneinander verschiedene Eigenwerte besitzt.

Man wählt nun V_0 derart, daß

$$(1.14) \quad V_0 M V_0^{-1} = \begin{pmatrix} {}_1\Omega_{11} & 0 \\ {}_1\Omega_{21} & {}_1\Omega_{22} \end{pmatrix}.$$

BEMERKUNG 1.6. V_0 ist damit nicht eindeutig bestimmt. In 1D ($B = 0$) oder falls B Diagonalgestalt besitzt, ist $M(\xi, \omega)$ Diagonalmatrix und man kann $V_0(\xi, \omega) = I$ wählen.

Durch Abbrechen der asymptotischen Entwicklung (1.13) nach jeweils endlich vielen Termen erhält man eine *Hierarchie von ARBen*. Wir benutzen die Approximation $V \cong V_0 + \xi^{-1}V_{-1} + O(\xi^{-2}) = (1 + K_1)V_0 + O(\xi^{-2})$ mit $K_1 V_0 = \xi^{-1}V_{-1}$, $(1 + K_1) \in \text{PS}(-1)$, $V_0 \in \text{PS}(0)$, um für $\tilde{W} := (1 + K_1)V_0 U$ eine Gleichung der Form (1.11) bis auf Fehler der Ordnung $O(\xi^{-1})$ auf Symbol Ebene zu schreiben. Anschließend kann man wie bei (1.12) vorgehen. Mit $V^{-1} \cong V_0^{-1}(1 - K_1) + O(\xi^{-2})$ erhalten wir

$$W_x = V_x U + V U_x \stackrel{(1.10)}{=} V_x U + V(MU + \tilde{C}U) = V_x V^{-1}W + V(M + \tilde{C})V^{-1}W.$$

Mit der *Kompositionsformel für Pseudodifferentialoperatoren* [36, Satz 3.4] ergibt sich

$$\begin{aligned} V_{0_x} V_0^{-1} + (1 + K_1)V_0(M + \tilde{C})V_0^{-1}(1 - K_1) \\ = V_{0_x} V_0^{-1} + V_0 M V_0^{-1} + K_1 V_0 M V_0^{-1} - V_0 M V_0^{-1} K_1 + V_0 \tilde{C} V_0^{-1} + O(\xi^{-1}). \end{aligned}$$

Man muß nun K_1 derart bestimmen, daß

$$D(x, y, t) := K_1 V_0 M V_0^{-1} - V_0 M V_0^{-1} K_1 + V_0 [V_0^{-1} V_{0_x} - \Lambda^{-1} C(x, y, t)] V_0^{-1}$$

eine untere Block-Dreiecksmatrix wird. Dabei wurde (1.14) verwendet. Für $B = \text{konst.}$ ist $D(x, y, t) = -\Lambda^{-1} C(x, y, t)$, und in diesem Fall kann eine untere Block-Dreiecksmatrix erreicht werden, wenn K_1 die Form

$$K_1(x, y, t) = \begin{pmatrix} 0 & K(x, y, t) \\ 0 & 0 \end{pmatrix}$$

hat und die Gleichung

$$-K(x, y, t) (\Lambda^-)^{-1} + (\Lambda^+)^{-1} K(x, y, t) - (\Lambda^+)^{-1} C^{+-}(x, y, t) = 0$$

erfüllt. Denn es ist ${}_1\Omega_{11} = -(\Lambda^+)^{-1}$ und ${}_1\Omega_{22} = -(\Lambda^-)^{-1}$. Somit ergibt sich

$$(1.15) \quad K(x, y, t) = (k_{j\ell})_{\substack{1 \leq j \leq m \\ m+1 \leq \ell \leq N}} \quad \text{mit} \quad k_{j\ell}(x, y, t) = \frac{\lambda_\ell}{\lambda_\ell - \lambda_j} c_{j\ell}(x, y, t).$$

BEMERKUNG 1.7. Es existiert ein eindeutiges K , falls $(\Lambda^+)^{-1}$ und $(\Lambda^-)^{-1}$ disjunkte Spektren haben [39, Abschnitt 1]. Dies ist hier offensichtlich erfüllt.

Aus einer *Störungsrechnung* in [10, Abschnitt 2C] folgt:

$$(1.16) \quad V_0(1, 0) = I \quad \text{und damit ist} \quad \xi^{-1}V_{-1}(1, 0, x, y, t) = K_1(x, y, t),$$

$$(1.17) \quad V_0(\xi, \omega) = V_0(1, \frac{\omega}{\xi}) = I + \frac{\omega}{\xi} \left(\frac{\partial}{\partial \omega} V_0 \right) (1, 0) + O \left(\left| \frac{\omega}{\xi} \right|^2 \right).$$

Ferner erhält man für den Ansatz

$$\frac{\partial}{\partial \omega} V_0(1, 0, x, y) = \begin{pmatrix} 0 & X(x, y) \\ 0 & 0 \end{pmatrix}$$

die Bedingung

$$-X(x, y) (\Lambda^-)^{-1} + (\Lambda^+)^{-1} X(x, y) - (\Lambda^+)^{-1} B^{+-}(x, y) = 0,$$

so daß sich folgendes ergibt:

$$(1.18) \quad X(x, y) = (\chi_{j\ell})_{\substack{1 \leq j \leq m \\ m+1 \leq \ell \leq N}} \quad \text{mit} \quad \chi_{j\ell}(x, y) = \frac{\lambda_\ell}{\lambda_\ell - \lambda_j} b_{j\ell}(x, y).$$

Aus der asymptotischen Entwicklung (1.13) und der Taylorentwicklung (1.17) um den *senkrechten Einfallswinkel* θ herum (es ist $\omega/\xi = \sin \theta$) ergeben sich nach einer Multiplikation mit ξ wie in [10] die ARBen:

ARB 0. Ordnung: Fehler $O(|\omega/\xi| + 1/|\xi|)$

(1.19)

$$U^+(0, y, t) = 0$$

ARB „ $\frac{1}{2}$. Ordnung“: Fehler $O(|\omega/\xi| + 1/|\xi|^2)$

(1.20)

$$U_t^+(0, y, t) + K(0, y, t)U^-(0, y, t) = 0$$

ARB 1. Ordnung: Fehler $O(|\omega/\xi|^2 + 1/|\xi|^2)$

(1.21)

$$U_t^+(0, y, t) + X(0, y)U_y^-(0, y, t) + K(0, y, t)U^-(0, y, t) = 0$$

BEMERKUNG 1.8. Ist B eine Diagonalmatrix, so ist (1.20) mit (1.21) identisch.

Inhomogene Randbedingungen. Bisher hatten wir angenommen, daß die Anfangsverteilung außerhalb des Rechengebietes verschwindet. Ist dies nicht der Fall, so muß man inhomogene Randbedingungen verwenden. Die *inhomogene ARB 1. Ordnung* hat in 1D die Form

$$(1.22) \quad U_t^+(0, t) + K(0, t)U^-(0, t) = g_t(t), \quad t \geq 0,$$

$$(1.23) \quad \text{bzw.} \quad U^+(0, t) = U^+(0, 0) - \int_0^t K(0, \tau)U^-(0, \tau) d\tau + g(t) - g(0), \quad t \geq 0.$$

Im folgenden setzen wir $K_0(\tau) := K(0, \tau)$.

1.3. Sachgemäße Gestelltheit in einer Ortsdimension

Wir betrachten das Problem (1.1) mit variablen Koeffizienten und inhomogenen absorbierenden Randbedingungen 1. Ordnung auf dem Streifen $0 \leq x \leq L$:

$$(1.24a) \quad U_t + \Lambda(x, t)U_x + C(x, t)U = F(x, t), \quad 0 \leq x \leq L, \quad t \geq 0$$

$$\Lambda = \text{diag}(\lambda_1, \dots, \lambda_N) \quad \text{mit} \quad \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m > 0 > \lambda_{m+1} \geq \dots \geq \lambda_N$$

$$(1.24b) \quad U(x, 0) = f(x), \quad 0 \leq x \leq L$$

$$(1.24c) \quad U^+(0, t) = U^+(0, 0) - \int_0^t K_0(\tau)U^-(0, \tau) d\tau + g_0(t) - g_0(0), \quad t \geq 0$$

$$U^-(L, t) = U^-(L, 0) - \int_0^t K_L(\tau)U^+(L, \tau) d\tau + g_L(t) - g_L(0), \quad t \geq 0$$

Die Koeffizienten $\Lambda(x, t)$, $C(x, t) \in \mathbb{R}^{N \times N}$ und $F(x, t)$, $f(x)$, $(g_0(t), g_L(t))^T \in \mathbb{R}^N$ seien C^∞ -glatt, und es gelten die in diesem Kapitel gemachten Annahmen. Speziell gilt mit der *Kompatibilitätsannahme* 1.3 in den obigen ARBen $U^+(0, 0) = g_0(0)$ und $U^-(L, 0) = g_L(0)$.

Grob gesagt ist das Anfangsrandwertproblem (1.24) *sachgemäß gestellt*, wenn es für alle glatten kompatiblen Daten F , f , g_0 , g_L eine eindeutige glatte Lösung U besitzt, die in jedem endlichen Zeitintervall $0 \leq t \leq T$ durch Terme der Daten abgeschätzt werden kann. Um zu einer Definition zu gelangen, muß man die Normen festlegen, die in der Abschätzung auftreten sollen¹. Neben der L^2 -Norm $\|\cdot\|_2$ auf $[0, L]$ verwenden wir die folgende Notation für *Normen am Rand*, d. h. für $x=0$, $x=L$:

euklidische Norm: $|\cdot|_\pm$

$$\text{Einströmanteil:} \quad |U(t)|_-^2 := |U^+(0, t)|^2 + |U^-(L, t)|^2$$

$$\text{Ausströmanteil:} \quad |U(t)|_+^2 := |U^+(L, t)|^2 + |U^-(0, t)|^2$$

gewichtete Norm: $\|\cdot\|_\pm$

$$\text{Einströmanteil:} \quad \|U(t)\|_-^2 := \sum_{\lambda_j(0,t) > 0} \lambda_j(0, t) |U_j(0, t)|^2 - \sum_{\lambda_j(L,t) < 0} \lambda_j(L, t) |U_j(L, t)|^2$$

$$\text{Ausströmanteil:} \quad \|U(t)\|_+^2 := \sum_{\lambda_j(L,t) > 0} \lambda_j(L, t) |U_j(L, t)|^2 - \sum_{\lambda_j(0,t) < 0} \lambda_j(0, t) |U_j(0, t)|^2$$

Bei Matrizen bedeutet $|\cdot|$ eine mit der euklidischen Vektornorm *verträgliche* Matrixnorm.

DEFINITION 1.3 (STARKE SACHGEMÄSSE GESTELLTHEIT [23, Def. 1, Seite 224]). Das Anfangsrandwertproblem (1.24) heißt *stark sachgemäß gestellt*, wenn es für alle glatten kompatiblen Daten F , f , g_0 , g_L eine eindeutige glatte Lösung U besitzt und für jedes endliche Zeitintervall $0 \leq t \leq T$ eine Konstante c_T existiert, so daß gilt

$$(1.25) \quad \|U(\cdot, t)\|_2^2 + \int_0^t \left(\|U(\tau)\|_-^2 + \|U(\tau)\|_+^2 \right) d\tau \\ \leq c_T \left[\|f\|_2^2 + \int_0^t \left(|g_0(\tau)|^2 + |g_L(\tau)|^2 + \|F(\cdot, \tau)\|_2^2 \right) d\tau \right]$$

für $0 \leq t \leq T$. Die Konstante c_T muß unabhängig von F , f , g_0 , g_L sein.

¹Es ist nicht nötig, die *Glattheitsvoraussetzungen* an F , f , g_0 und g_L zu präzisieren. Man kann hier „glatt“ stets durch „ C^∞ “ ersetzen. Hat man erstmal für diesen Fall eine Abschätzung hergeleitet, so kann man *weniger glatte Daten* durch eine Approximation behandeln, solange die Normen für diese Daten erklärt sind [23, Seite 223]. Dies gilt auch für Satz 1.1

1.3.1. Die Abschätzung der Lösung. Für das System mit konstanten Koeffizienten erhalten wir als Differenz von Ausström- und Einströmanteil:

$$\begin{aligned} \|U(t)\|_+^2 - \|U(t)\|_-^2 &= \sum_{m=1}^N \lambda_j |U_j(L, t)|^2 - \sum_{m=1}^N \lambda_j |U_j(0, t)|^2 \\ &= \langle \Lambda U(x, t), U(x, t) \rangle \Big|_{x=0}^{x=L} \\ &= 2 \int_0^L \langle \Lambda U_x(x, t), U(x, t) \rangle dx. \end{aligned}$$

Wir multiplizieren (1.24a) mit $2U$ und integrieren bezüglich x :

$$(1.26) \quad \partial_t \|U(\cdot, t)\|_2^2 + \|U(t)\|_+^2 - \|U(t)\|_-^2 + 2 \int_0^L \langle CU(x, t), U(x, t) \rangle dx \\ = 2 \int_0^L \langle F(x, t), U(x, t) \rangle dx.$$

Eine Integration bezüglich der Zeit liefert die folgende *Einström-Ausströmbilanz*:

$$\begin{aligned} \|U(\cdot, t)\|_2^2 &= \|f\|_2^2 + \int_0^t (\|U(\tau)\|_-^2 - \|U(\tau)\|_+^2) d\tau - 2 \int_0^t \int_0^L \langle CU(x, \tau), U(x, \tau) \rangle dx d\tau \\ &\quad + 2 \int_0^t \int_0^L \langle F(x, \tau), U(x, \tau) \rangle dx d\tau. \end{aligned}$$

BEMERKUNG 1.9. Falls C *schiefsymmetrisch* ist, gilt $\langle CU(x, \tau), U(x, \tau) \rangle = 0$.

Unser Ziel ist es, die Einströmdaten durch die Ausströmdaten abzuschätzen. Es gilt

$$(1.27) \quad \|U(t)\|_-^2 \leq \rho_1 \left(|U^+(0, t)|^2 + |U^-(L, t)|^2 \right), \quad \rho_1 := \rho(\Lambda) = \max_{p=1, \dots, N} |\lambda_p|.$$

Aus der Randbedingung (1.24c) an $x = 0$ erhält man mit der Cauchy-Schwarzschen Ungleichung

$$\left| \int_0^t K_0 U^-(0, \tau) d\tau \right| \leq |K_0| \sqrt{t} \left(\int_0^t |U^-(0, \tau)|^2 d\tau \right)^{\frac{1}{2}}$$

und damit ist

$$(1.28) \quad |U^+(0, t)|^2 \leq 2 |g_0(t)|^2 + 2 \left| \int_0^t K_0 U^-(0, \tau) d\tau \right|^2 \\ \leq 2 |g_0(t)|^2 + 2 |K_0|^2 t \int_0^t |U^-(0, \tau)|^2 d\tau.$$

Analog ergibt sich aus der Randbedingung an $x = L$ die Abschätzung

$$(1.29) \quad |U^-(L, t)|^2 \leq 2 |g_L(t)|^2 + 2 |K_L|^2 t \int_0^t |U^+(L, \tau)|^2 d\tau.$$

Somit gilt mit $\kappa := \max(|K_0|, |K_L|)$

$$(1.30) \quad \|U(t)\|_-^2 \leq 2\rho_1 \left(|g_0(t)|^2 + |g_L(t)|^2 + \kappa^2 t \int_0^t (|U^-(0, \tau)|^2 + |U^+(L, \tau)|^2) d\tau \right) \\ \leq 2\rho_1 \left(|g_0(t)|^2 + |g_L(t)|^2 + \kappa^2 t \int_0^t |U(\tau)|_+^2 d\tau \right).$$

Betrachten wir

$$(1.31) \quad z(t) := \|U(\cdot, t)\|_2^2 + \int_0^t \left(\|U(\tau)\|_-^2 + \|U(\tau)\|_+^2 \right) d\tau,$$

so gilt für $0 \leq t \leq T$:

$$\begin{aligned} z(t) &= \|f\|_2^2 + 2 \int_0^t \|U(\tau)\|_-^2 d\tau - 2 \int_0^t \int_0^L \left(\langle CU(x, \tau), U(x, \tau) \rangle - \langle F(x, \tau), U(x, \tau) \rangle \right) dx d\tau \\ &\leq \|f\|_2^2 + 4\rho_1 \left(\int_0^t \left(|g_0(\tau)|^2 + |g_L(\tau)|^2 \right) d\tau + \kappa^2 t \int_0^t \int_0^\tau |U(s)|_+^2 ds d\tau \right) \\ &\quad + 2|C| \int_0^t \int_0^L |U(x, \tau)|^2 dx d\tau + \int_0^t \int_0^L \left(|F(x, \tau)|^2 + |U(x, \tau)|^2 \right) dx d\tau \\ &\leq \|f\|_2^2 + \left(4\rho_1 \rho_2 \kappa^2 t + 2|C| + 1 \right) \int_0^t \left(\|U(\cdot, \tau)\|_2^2 + \int_0^\tau \|U(s)\|_+^2 ds \right) d\tau \\ &\quad + 4\rho_1 \int_0^t \left(|g_0(\tau)|^2 + |g_L(\tau)|^2 \right) d\tau + \int_0^t \|F(\cdot, \tau)\|_2^2 d\tau, \quad \rho_2 := \rho(\Lambda^{-1}) \\ &\leq z(0) + c_1 \int_0^t z(\tau) d\tau + c_2 \int_0^t \left(|g_0(\tau)|^2 + |g_L(\tau)|^2 + \|F(\cdot, \tau)\|_2^2 \right) d\tau \end{aligned}$$

mit den Konstanten $c_1 := 4\rho_1 \rho_2 \kappa^2 T + 2|C| + 1$, $c_2 := 4\rho_1 + 1$.

Eine Anwendung der *Gronwallschen Ungleichung* [23, Lemma 3.1.1] ergibt

$$z(t) \leq e^{c_1 t} \|f\|_2^2 + c_2 e^{c_1 t} \int_0^t \left(|g_0(\tau)|^2 + |g_L(\tau)|^2 + \|F(\cdot, \tau)\|_2^2 \right) d\tau.$$

Damit haben wir mit $c_T := (1 + c_2)e^{c_1 T}$ eine Abschätzung der Form:

$$(1.32) \quad \|U(\cdot, t)\|_2^2 + \int_0^t \left(\|U(\tau)\|_-^2 + \|U(\tau)\|_+^2 \right) d\tau \\ \leq c_T \left[\|f\|_2^2 + \int_0^t \left(|g_0(\tau)|^2 + |g_L(\tau)|^2 + \|F(\cdot, \tau)\|_2^2 \right) d\tau \right].$$

Die obige Abschätzung läßt sich auch bei variablen Koeffizienten führen, wenn gilt:

ANNAHME 1.4. Die Koeffizientenfunktionen, Λ_x und die Randbedingungsmatrizen K_0 , K_L sind *gleichmäßig beschränkt*, d. h. $|\Lambda(x, t)|$, $|C(x, t)|$, $|\Lambda_x(x, t)| \leq \text{konst.}$ und $|K_0(t)| \leq \kappa_0$, $|K_L(t)| \leq \kappa_L$ für $0 \leq x \leq L$ und alle Zeiten $t \in [0, T]$. Ferner muß für alle $t \in [0, T]$ $|\lambda_j(0, t)|$, $|\lambda_j(L, t)| \geq \text{konst.} > 0$, $j = 1, \dots, N$ gelten, womit die auf Seite 9 definierten Normen äquivalent sind.

Diese Annahme gelte im folgenden.

1.3.2. Die Existenz und Eindeutigkeit der Lösung. Wir wollen hier den Beweis, der auf einem Fixpunktargument basiert, nur skizzieren. Bei vorgegebenen Einströmdaten wird die Gleichung mit variablen Koeffizienten gelöst und die Lösung dann am Ausströmrand ausgewertet. Mit den ARBen kommt man zurück zu Einströmdaten und die Iteration ist fertig. Der Banachraum für die Iteration ist als $L^2((0, T), \mathbb{R}^N)$ wählbar. Der *Fixpunktoperator* \mathbf{F} ist demnach wie folgt definiert:

Für $V = (V^+, V^-)^T \in L^2((0, T), \mathbb{R}^N)$ und festes, endliches T löse man das Problem

$$(1.33a) \quad Y_t + \Lambda(x, t)Y_x + C(x, t)Y = F(x, t), \quad 0 \leq x \leq L, \quad t \geq 0$$

$$(1.33b) \quad Y(x, 0) = f(x), \quad 0 \leq x \leq L$$

$$(1.33c) \quad Y^+(0, t) = V^+(t), \quad t \geq 0$$

$$Y^-(L, t) = V^-(t), \quad t \geq 0.$$

Dann ist $\mathbf{F}V = ((\mathbf{F}V)^+, (\mathbf{F}V)^-)^T$ definiert durch

$$(1.34) \quad (\mathbf{F}V)^+(t) := g_0(t) - \int_0^t K_0(\tau)Y^-(0, \tau) d\tau, \quad t \geq 0$$

$$(\mathbf{F}V)^-(t) := g_L(t) - \int_0^t K_L(\tau)Y^+(L, \tau) d\tau, \quad t \geq 0.$$

LEMMA 1.2. *Der Fixpunktoperator \mathbf{F} ist eine Selbstabbildung von $L^2((0, T), \mathbb{R}^N)$.*

BEWEIS. Nach Lösen von (1.33) gilt mit Satz 1.1:

$$\|(Y(\cdot, t))\|_2^2 + \int_0^t (|Y(0, \tau)|^2 + |Y(L, \tau)|^2) d\tau \leq c'_T \left[\|f\|_2^2 + \int_0^t (|V(\tau)|^2 + \|F(\cdot, \tau)\|_2^2) d\tau \right]$$

für $0 \leq t \leq T$ und wegen $V \in L^2((0, T), \mathbb{R}^N)$ hat man somit

$$\int_0^T (|Y^-(0, \tau)|^2 + |Y^+(L, \tau)|^2) d\tau \leq \text{konst.}$$

Also ist $(Y^-(0, \cdot), Y^+(L, \cdot))^T \in L^2((0, T), \mathbb{R}^N)$ und mit (1.34) gilt $\mathbf{F}V \in L^2((0, T), \mathbb{R}^N)$. \square

LEMMA 1.3. *\mathbf{F} ist zumindest auf einem Teilintervall $(0, T_1)$ von $(0, T)$ kontraktiv.*

BEWEIS. Sind zwei Einströmdaten V_1, V_2 gegeben, so kann nach Lösen von (1.33) die Differenz der zugehörigen Ausströmdaten mit Satz 1.1 abgeschätzt werden:

$$\begin{aligned} \|(Y_1 - Y_2)(\cdot, t)\|_2^2 + \int_0^t (|(Y_1 - Y_2)(0, \tau)|^2 + |(Y_1 - Y_2)(L, \tau)|^2) d\tau \\ \leq \tilde{c}_T \int_0^t (|(V_1^+ - V_2^+)(\tau)|^2 + |(V_1^- - V_2^-)(\tau)|^2) d\tau \end{aligned}$$

und damit

$$(1.35) \quad \int_0^t |Y_1(\tau) - Y_2(\tau)|_+^2 \leq \tilde{c}_T \int_0^t |V_1(\tau) - V_2(\tau)|^2 d\tau.$$

Wie bei (1.28) erhalten wir aus den Randbedingungen (1.34)

$$\begin{aligned} |\mathbf{F}V_1^+(t) - \mathbf{F}V_2^+(t)|^2 &\leq \kappa_0^2 t \int_0^t |Y_1^-(0, \tau) - Y_2^-(0, \tau)|^2 d\tau \\ |\mathbf{F}V_1^-(t) - \mathbf{F}V_2^-(t)|^2 &\leq \kappa_L^2 t \int_0^t |Y_1^+(L, \tau) - Y_2^+(L, \tau)|^2 d\tau \end{aligned}$$

und das liefert mit (1.35)

$$(1.36) \quad \begin{aligned} |\mathbf{F}V_1(t) - \mathbf{F}V_2(t)|^2 &\leq \kappa^2 t \int_0^t |Y_1(\tau) - Y_2(\tau)|_+^2 d\tau, \quad \kappa := \max(\kappa_0, \kappa_L) \\ &\leq \tilde{c}_T \kappa^2 t \int_0^t |V_1(\tau) - V_2(\tau)|^2 d\tau. \end{aligned}$$

Mit einer partiellen Integration bezüglich t erhält man schließlich aus (1.36)

$$\begin{aligned} \int_0^T |\mathbf{F}V_1(t) - \mathbf{F}V_2(t)|^2 dt &\leq \frac{\tilde{c}_T}{2} \kappa^2 \left(t^2 \int_0^t |V_1(\tau) - V_2(\tau)|^2 d\tau \Big|_0^T - \int_0^T t^2 |V_1(t) - V_2(t)|^2 dt \right) \\ &\leq \frac{\tilde{c}_T}{2} \kappa^2 T^2 \int_0^T |V_1(t) - V_2(t)|^2 dt. \end{aligned}$$

\mathbf{F} ist damit kontraktiv für $T_1 < \sqrt{2}/(\kappa\sqrt{\tilde{c}_T})$. \square

Das Kontraktionsintervall von \mathbf{F} hängt nur von $\kappa\sqrt{\tilde{c}_T}$ ab, und somit kann man die Iteration zumindest auf ein Teilintervall $(0, T_1)$ von $(0, T)$ anwenden. Da dies nur von der Länge des betrachteten Zeitintervalls abhängt, kann die resultierende lokale Lösung auf solchen Teilintervallen gleicher Länge bis zu T fortgesetzt werden.

BEMERKUNG 1.10 (GLATTHEIT DER DATEN). Die sachgemäße Gestelltheit von (1.24) läßt sich unter viel schwächeren Voraussetzungen an die Glattheit der Daten zeigen; es ist offensichtlich ausreichend, $F \in L^2((0, T), L^2((0, L), \mathbb{R}^N))$, $f \in L^2((0, L), \mathbb{R}^N)$ und $(g_0, g_L)^T \in L^2((0, T), \mathbb{R}^N)$ vorzusetzen.

BEMERKUNG 1.11 (NULLDURCHGANG EINES EIGENWERTES). Es stellt sich die Frage, was geschieht, wenn $\Lambda(x, t)$ auf $0 < x < L$ *singulär* wird, also ein Eigenwert $\lambda_j(x, t)$ im Inneren des Streifens Null wird. In der Annahme 1.4 reicht es aus, die gleichmäßige Beschränktheit von Λ nur an $x = 0$ und $x = L$ zu fordern. Somit hat man (neben der gleichmäßigen Beschränktheit von Λ_x) folgende Voraussetzung an die Eigenwerte λ_j :

$$\text{konst.} \geq |\lambda_j(0, t)|, |\lambda_j(L, t)| \geq \text{konst.} > 0, \quad j = 1, \dots, N.$$

Demnach darf ein Eigenwert λ_j auf $0 < x < L$ Null werden, solange Λ_x gleichmäßig beschränkt bleibt, z. B. $\lambda_j(x, t) = \alpha(x - x_0) + a(t)$, $a(t)$ geeignet gewählt.

KAPITEL 2

Numerische Betrachtung des linearen Falls in 1D

Wir betrachten finite Differenzenverfahren für lineare hyperbolische Systeme in 1D

$$(2.1) \quad U_t + AU_x + CU = F(x, t), \quad 0 \leq x \leq L, \quad t \geq 0,$$

wobei A, C konstante $N \times N$ -Matrizen sind. Der Fall von orts- und zeitabhängigen Koeffizienten läßt sich völlig analog betrachten. Mit einer *Ortschrittweite* $h := \Delta x$ und einer *Zeitschrittweite* $k := \Delta t$ diskretisieren wir den (x, t) -Streifen $[0, L] \times \mathbb{R}_0^+$ durch folgende *Gitterpunkte*

$$\begin{aligned} x_j &= jh, & j &= 0, 1, 2, \dots, J \\ t_n &= nk, & n &= 0, 1, 2, \dots \end{aligned}$$

Wir nehmen an, daß das *Schrittweitenverhältnis* $\lambda := k/h$ konstant ist und damit die Wahl einer Zeitschrittweite $k > 0$ ein äquidistantes Gitter definiert. Die finiten Differenzenmethoden liefern *Approximationen* $U_j^n \in \mathbb{R}^N$ der exakten Lösung $U(x_j, t_n)$ an den diskreten Gitterpunkten (x_j, t_n) .

2.1. Das numerische Schema für 1D

Das *Lax-Wendroff-Verfahren* [24] (kurz: LW-Verfahren) basiert auf der Entwicklung

$$(2.2) \quad U(x, t + k) = U(x, t) + kU_t(x, t) + \frac{1}{2} k^2 U_{tt}(x, t) + O(k^3),$$

wobei man U_{tt} aus (2.1) wie folgt bestimmt

$$\begin{aligned} U_{tt} &= (-AU_x - CU + F)_t \\ &= -AU_{tx} - CU_t + F_t \\ &= A(AU_x + CU - F)_x + C(AU_x + CU - F) + F_t \\ &= A^2U_{xx} + (AC + CA)U_x + C^2U + F_t - AF_x - CF, \end{aligned}$$

so daß sich (2.2) ergibt zu

$$(2.3) \quad \begin{aligned} U(x, t + k) &= U(x, t) - k(AU_x(x, t) + CU(x, t) - F(x, t)) \\ &+ \frac{1}{2} k^2 (A^2U_{xx}(x, t) + (AC + CA)U_x(x, t) + C^2U(x, t) \\ &+ F_t(x, t) - AF_x(x, t) - CF(x, t)) + O(k^3). \end{aligned}$$

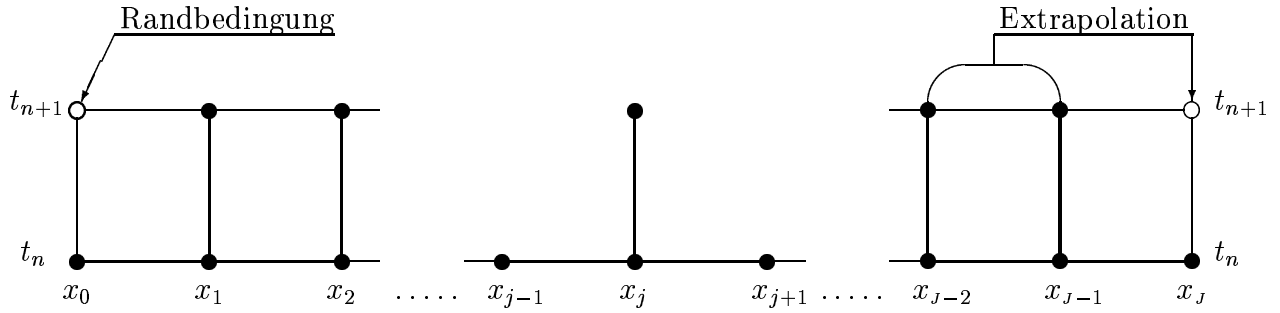


ABBILDUNG 2.1. Lax-Wendroff-Schema in 1D

Die LW-Methode verwendet nun zentrale Differenzenquotienten, um die in (2.3) auftretenden Ortsableitungen von U zu diskretisieren. Weiterhin werden die Ableitungen von F geeignet diskretisiert. Das resultierende Schema lautet damit:

$$\begin{aligned}
 (2.4) \quad U_j^{n+1} &= U_j^n - k \left(A \frac{U_{j+1}^n - U_{j-1}^n}{2h} + CU_j^n - F_j^n \right) \\
 &+ \frac{1}{2} k^2 \left(A^2 \frac{U_{j+1}^n - 2U_j^n + U_{j-1}^n}{h^2} + (AC + CA) \frac{U_{j+1}^n - U_{j-1}^n}{2h} + C^2 U_j^n \right. \\
 &\quad \left. + \frac{F_j^{n+1} - F_j^n}{k} - A \frac{F_{j+1}^n - F_{j-1}^n}{2h} - CF_j^n \right) \\
 &= U_j^n - \frac{1}{2} \lambda A (U_{j+1}^n - U_{j-1}^n) - kCU_j^n + \frac{1}{2} (\lambda A)^2 (U_{j+1}^n - 2U_j^n + U_{j-1}^n) \\
 &+ \frac{1}{4} \lambda k (AC + CA) (U_{j+1}^n - U_{j-1}^n) + \frac{1}{2} (kC)^2 U_j^n \\
 &+ \frac{1}{2} k (F_j^{n+1} + F_j^n) - \frac{1}{4} \lambda k A (F_{j+1}^n - F_{j-1}^n) - \frac{1}{2} k^2 CF_j^n.
 \end{aligned}$$

Neben einer Diskretisierung der analytischen Einströmrandbedingung benötigt man zusätzliche *numerische Randbedingungen* am Ausströmrand, um das Schema zu schließen. Wir werden dazu eine horizontale Extrapolation 1. Ordnung verwenden. Allgemein hat die *horizontale Extrapolation der Ordnung q* für die Ausströmdaten an $x = 0$ bzw. $x = L$ die Form

$$(2.5a) \quad (E - I)^{q+1} (U^-)^{n+1}_0 = 0, \quad q = 0, 1, 2, \dots$$

$$(2.5b) \quad \text{und} \quad (I - E^{-1})^{q+1} (U^+)^{n+1}_J = 0, \quad q = 0, 1, 2, \dots,$$

wobei $EU_j := U_{j+1}$ ein *Shift-Operator* ist. Speziell für $q = 1$ erhalten wir daraus

$$(2.6a) \quad (U^-)^{n+1}_0 = 2(U^-)^{n+1}_1 - (U^-)^{n+1}_2$$

$$(2.6b) \quad \text{bzw.} \quad (U^+)^{n+1}_J = 2(U^+)^{n+1}_{J-1} - (U^+)^{n+1}_{J-2}.$$

Um zur Formulierung der Randbedingungen die Ein- und Ausströmkomponenten identifizieren zu können, muß man, falls A keine Diagonalmatrix ist, die Lösung am Rand mit derjenigen Matrix transformieren, die A diagonalisiert. Die Abbildung 2.1 veranschaulicht dieses Vorgehen für den Fall, daß $x = 0$ Einströmrand und $x = L$ Ausströmrand ist, d. h. bei einem positiven Eigenwert von A .

2.2. Die Stabilität des Differenzenschemas in einer Ortsdimension

In diesem Abschnitt wollen wir eine *Stabilitätsanalyse* des Differenzenschemas zum Lösen des Anfangsrandwertproblems mit inhomogenen ARBen 1. Ordnung durchführen. Es wird die Stabilität der inneren Diskretisierung in Kombination mit der Randbedingung untersucht. Wir werden dabei die Stabilitätstheorie von Gustafsson, Kreiss und Sundström [18], [22] (kurz: *GKS-Stabilitätstheorie*) verwenden, die sowohl hinreichende als auch notwendige Bedingungen für die Stabilität eines diskreten ARWP in einer Raumdimension zur Verfügung stellt. Die komplizierten algebraischen Bedingungen der GKS-Theorie wurden in nachfolgenden Arbeiten von Goldberg und Tadmor [14], [15] vereinfacht. Wie bei der sachgemäßen Gestelltheit in Abschnitt 1.3 tritt das Problem auf, daß die ARBen nicht die Standardform der Randbedingungen in der GKS-Theorie haben und somit diese Stabilitätstheorie nicht unmittelbar anwendbar ist.

BEMERKUNG 2.1. Das *diskrete ARWP mit zwei Rändern*, d. h. das auf dem beschränkten Indexbereich $0 \leq j \leq J$ gegeben ist, ist stabil, wenn die zugehörigen Anfangsrandwertprobleme auf den halbunendlichen Indexbereichen $-\infty < j \leq J$ und $0 \leq j < \infty$ stabil sind [18, Satz 5.4]. Dies ist eine Konsequenz der *endlichen Fortpflanzungsgeschwindigkeit* von Information bei hyperbolischen Gleichungen.

Daher reicht es aus, das Problem auf der positiven Halbachse $x \geq 0$ zu betrachten:

$$(2.7a) \quad U_t + \Lambda U_x + CU = F(x, t), \quad \Lambda = \text{diag}(\lambda_1, \dots, \lambda_N), \quad x \geq 0, t \geq 0$$

$$\text{mit } \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m > 0 > \lambda_{m+1} \geq \dots \geq \lambda_N,$$

$$\text{d. h. } \Lambda^+ = \text{diag}(\lambda_1, \dots, \lambda_m), \quad \Lambda^- = \text{diag}(\lambda_{m+1}, \dots, \lambda_N)$$

$$(2.7b) \quad U(x, 0) = f(x), \quad x \geq 0$$

$$(2.7c) \quad U^+(0, t) = U^+(0, 0) - \int_0^t K_0 U^-(0, \tau) d\tau + g(t) - g(0), \quad t \geq 0.$$

Wegen der *Kompatibilitätsannahme* 1.3 gilt in der obigen ARB (2.7c) $U^+(0, 0) = g(0)$.

BEMERKUNG 2.2. Die Matrix $K_0 \in \mathbb{R}^{m \times (N-m)}$ stammt nicht notwendig von einer ARB.

Wir wollen das Problem (2.7) mit einer konsistenten Differenzenapproximation der Form

$$(2.8a) \quad U_j^{n+1} = QU_j^n + kb_j^n, \quad j \geq 1$$

$$\text{mit } Q = \sum_{\sigma=-1}^1 A_\sigma E^\sigma, \quad EU_j = U_{j+1}$$

lösen. Für das LW-Schema (2.4) ergibt sich dabei

$$A_0 = I - kC - (\lambda A)^2 + \frac{1}{2}(kC)^2$$

$$A_{\pm 1} = \mp \frac{1}{2} \lambda A + \frac{1}{2} (\lambda A)^2 \pm \frac{1}{4} \lambda k (AC + CA)$$

$$b_j^n = \frac{1}{2} (I - kC) F_j^n + \frac{1}{2} F_j^{n+1} - \frac{1}{4} \lambda A (F_{j+1}^n - F_{j-1}^n).$$

Damit die Lösung von (2.8a) eindeutig bestimmt ist, benötigt man eine *diskrete Anfangsbedingung*

$$(2.8b) \quad U_j^0 = f(jh), \quad j = 0, 1, 2, \dots$$

und *Randbedingungen* an $x = 0$ für die Einströmkomponenten U^+

$$(2.8c) \quad (U^+)_0^{n+1} = (U^+)_0^n - \frac{k}{2} K_0 [(U^-)_0^{n+1} + (U^-)_0^n] + g^{n+1} - g^n,$$

die man durch eine Diskretisierung von (2.7c) mit Hilfe der Trapezregel erhält. Für die Ausströmdaten U^- verwenden wir

$$(2.8d) \quad (U^-)_0^{n+1} = 2(U^-)_1^{n+1} - (U^-)_2^{n+1},$$

um das numerische Schema zu schließen. Dies ist die *horizontale Extrapolation* (2.6a). Die obigen Randbedingungen lassen sich in folgender Form schreiben:

$$(2.9) \quad U_0^{n+1} = S_{-1} U_0^{n+1} + S_0 U_0^n + \tilde{g}^n,$$

wobei $S_{-1} = \sum_{\sigma=0}^2 B_{\sigma,-1} E^\sigma, \quad S_0 = B_{0,0}.$

Dabei ist

$$B_{0,-1} = -\frac{k}{2} \begin{pmatrix} 0 & K_0 \\ 0 & 0 \end{pmatrix}, \quad B_{1,-1} = \begin{pmatrix} 0 & 0 \\ 0 & 2I_{N-m} \end{pmatrix}, \quad B_{2,-1} = \begin{pmatrix} 0 & 0 \\ 0 & -I_{N-m} \end{pmatrix}$$

$$B_{0,0} = \begin{pmatrix} I_m & -\frac{k}{2} K_0 \\ 0 & 0 \end{pmatrix}, \quad \tilde{g}^n = \begin{pmatrix} g^{n+1} - g^n \\ 0 \end{pmatrix}$$

und wir haben die *diskrete Kompatibilitätsbedingung*: $(U^+)_0^0 = g^0$.

BEMERKUNG 2.3. Die diskretisierte Randbedingung (2.8c) ist von 2. Ordnung in t um $t_{n+\frac{1}{2}} = (n + \frac{1}{2})k$, da der zentrale Differenzenquotient verwendet wurde.

Für die Stabilitätsanalyse der inneren Diskretisierung benötigen wir folgenden Begriff.

DEFINITION 2.1 (AMPLIFIKATIONSMATRIX [22]). Mit

$$(2.10) \quad \hat{Q}(\theta) := \sum_{\sigma=-1}^1 A_\sigma e^{i\sigma\theta}, \quad \theta \in \mathbb{R}$$

bezeichnen wir die *Amplifikationsmatrix* von Q .

BEMERKUNG 2.4. Praktisch kann man die Amplifikationsmatrix $\hat{Q}(\theta)$ bestimmen, indem man im Differenzenschema U_j^n durch $\hat{Q}^n e^{ij\theta}$ ersetzt und nach \hat{Q} auflöst [38, Seite 136].

In der Stabilitätstheorie wird gefordert, daß $\hat{Q}(\theta)$ die folgenden zwei Annahmen erfüllt.

ANNAHME 2.1 (STABILITÄT FÜR DISKRETES AWP [22]). $|\hat{Q}(\theta)| \leq 1$.

ANNAHME 2.2 (DISSIPATIVES SCHEMA [38]). Das Schema (ohne die Terme niedrigerer Ordnung, d. h. $C = 0$) ist *dissipativ von der Ordnung* $2r$, d. h. es gibt eine Konstante $c > 0$, so daß für die Eigenwerte $\mu(\theta)$ von $\hat{Q}(\theta)$ gilt

$$(2.11) \quad |\mu(\theta)| \leq 1 - c \sin^{2r} \frac{\theta}{2}.$$

Diese Bedingung (2.11) ist nach [38, Argumentation auf Seite 46] äquivalent zu

$$(2.12) \quad |\mu(\theta)|^2 \leq 1 - c' \sin^{2r} \frac{\theta}{2}.$$

In vielen Texten, z. B. [31, Seite 109], wird bei der Definition der Dissipation der Term $\sin^{2r}(\theta/2)$ in (2.11) durch $|\theta|^{2r}$ mit der Beschränkung $|\theta| \leq \pi$ ersetzt. Beide Definitionen sind nach [38, Kapitel 5.1] äquivalent; die Bedingung (2.11) tritt allerdings bei der Untersuchung der meisten Schemas auf und ist daher vorzuziehen.

BEMERKUNG 2.5. Annahme 2.1 garantiert, daß das Schema (2.8a), (2.8b) für das reine Anfangswertproblem stabil ist. Die *Stabilität des reinen AWP* ist neben der sachgemäßen Gestelltheit des zugrundeliegenden analytischen Problems (Abschnitt 1.3) eine *notwendige Bedingung* für die Stabilität des Differenzenverfahrens für das Anfangsrandwertproblem. Die Annahme 2.2 stellt sicher, daß hohe Frequenzen keinen Einfluß haben, d. h. die Größe der hochfrequenten Oszillationen nimmt in jedem Schritt ab [22, Seite 704].

Im folgenden wollen wir kurz auf diese zwei Annahmen eingehen.

Die Stabilität des reinen Anfangswertproblems. Das Lax-Wendroff-Verfahren zur Lösung des reinen AWP ist *stabil* (im Sinne von [34, Seite 16]) für

$$(2.13) \quad \nu := \max_{\ell=1, \dots, N} \left| \frac{\lambda_\ell k}{h} \right| \leq 1.$$

DEFINITION 2.2 (ABHÄNGIGKEITSBEREICH [34, Seite 16]). Der *Abhängigkeitsbereich* $\mathcal{D}(\bar{x}, \bar{t})$ ist die Menge aller Punkte (x, t) mit der folgenden Eigenschaft: Eine Störung der Lösung von (2.7a) hat einen Einfluß auf die Lösung an der Stelle (\bar{x}, \bar{t}) . Analog wird der *numerische Abhängigkeitsbereich* $\mathcal{D}_k(\bar{x}, \bar{t})$ für Differenzenverfahren definiert.

Es ist für die Stabilität einer numerischen Methode notwendig, daß asymptotisch für $k \rightarrow 0$ der numerische Abhängigkeitsbereich $\mathcal{D}_k(\bar{x}, \bar{t})$ den Abhängigkeitsbereich enthält. Dies ist die *CFL(Courant-Friedrichs-Levy)-Bedingung*. Für beliebige Verfahren der Form (2.8a) nimmt die CFL-Bedingung die obige Gestalt (2.13) an [34, Seite 16]. Die Zahl ν wird *Courantzahl* genannt.

Die Dissipation des Lax-Wendroff-Schemas. Nach Bemerkung 2.4 bestimmt man die Amplifikationsmatrix $\hat{Q}(\theta)$ (mit $C = 0, F = 0$) wie folgt:

$$\begin{aligned} \hat{Q}^{n+1} e^{ij\theta} &= \hat{Q}^n e^{ij\theta} \left[I - \frac{1}{2} \Lambda \lambda (e^{i\theta} - e^{-i\theta}) + \frac{1}{2} (\Lambda \lambda)^2 (e^{i\theta} - 2 + e^{-i\theta}) \right] \\ &= \hat{Q}^n e^{ij\theta} \left[I - i \Lambda \lambda \sin \theta - (\Lambda \lambda)^2 (1 - \cos \theta) \right]. \end{aligned}$$

Somit erhalten wir die Amplifikationsmatrix

$$(2.14) \quad \hat{Q}(\theta) = I - i \Lambda \lambda \sin \theta - (\Lambda \lambda)^2 (1 - \cos \theta),$$

die ein Polynom in Λ ist und damit die Eigenwerte

$$(2.15) \quad \mu_\ell(\theta) = 1 - i \lambda_\ell \lambda \sin \theta - (\lambda_\ell \lambda)^2 (1 - \cos \theta), \quad \ell = 1, \dots, N$$

besitzt. Dabei wurden die Eigenwerte von Λ mit λ_ℓ bezeichnet. Es ergibt sich

$$\begin{aligned} |\mu_\ell(\theta)|^2 &= \left[1 - (\lambda_\ell \lambda)^2 (1 - \cos \theta)\right]^2 + (\lambda_\ell \lambda)^2 \sin^2 \theta \\ &= 1 - (\lambda_\ell \lambda)^2 \left[2 - 2 \cos \theta - (\lambda_\ell \lambda)^2 (1 - \cos \theta)^2 - \sin^2 \theta\right] \\ &= 1 - (\lambda_\ell \lambda)^2 \left[(1 - \cos \theta)^2 - (\lambda_\ell \lambda)^2 (1 - \cos \theta)^2\right] \\ &= 1 - 4(\lambda_\ell \lambda)^2 \left[1 - (\lambda_\ell \lambda)^2\right] \sin^4 \frac{\theta}{2}, \quad \ell = 1, \dots, N, \end{aligned}$$

wobei wir die Identität $(1 - \cos \theta)^2 = 4 \sin^4(\theta/2)$ benutzt haben. Somit ist das Differenzenschema nach (2.12) *dissipativ von 4. Ordnung*, falls $0 < |\lambda_\ell \lambda| < 1$ für alle $\ell = 1, \dots, N$ gilt. Da Λ regulär ist, sind somit die Annahmen 2.1 und 2.2 erfüllt, wenn die CFL-Bedingung (2.13) gilt.

2.2.1. Abschätzung der Lösung der Differentialgleichung. Wir beweisen eine Abschätzung für die Lösung, die uns als Motivation für den zu verwendenden Stabilitätsbegriff dient. Außerdem werden wir später eine ähnliche Technik bei der Herleitung der Abschätzungen für die Lösungen der Differenzenapproximationen benutzen.

Für $U, V: \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathcal{C}^n$ definieren wir folgende *L^2 -innere Produkte*

$$\begin{aligned} \langle U, V \rangle_x &:= \int_0^\infty U^*(x, t) V(x, t) dx, & \langle U, V \rangle_t &:= \int_0^\infty U^*(x, t) V(x, t) dt, \\ \langle U, V \rangle_{x,t} &:= \int_0^\infty \int_0^\infty U^*(x, t) V(x, t) dx dt \end{aligned}$$

und definieren die dadurch *induzierten Normen* wie üblich durch $\|u\|_x^2 := \langle u, u \rangle_x$, usw. Wir bezeichnen die zugehörigen *L^2 -Räume* mit $L^2(x)$, $L^2(t)$ und $L^2(x, t)$.

DEFINITION 2.3 (LAPLACE-TRANSFORMATION [7]). Gibt es für ein $U: \mathbb{R}^+ \rightarrow \mathcal{C}$ ein $\alpha_0 \in \mathbb{R}$, so daß das Integral

$$\mathcal{L}\{U\}(s) := \hat{U}(s) := \int_0^\infty e^{-st} U(t) dt$$

für alle $s \in \mathcal{C}$ mit $\operatorname{Re} s > \alpha_0$ existiert, so heißt die so gebildete Zuordnung zwischen den Funktionen U und \hat{U} *Laplace-Transformation* (kurz: *\mathcal{L} -Transformation*). Die Funktion $\hat{U}(s)$ heißt dann *Laplace-Transformierte* von U .

Die \mathcal{L} -Transformation einer vektorwertigen Funktion geschieht dabei komponentenweise.

Weiterhin benötigen wir die *Parsevalsche Gleichung*, die es uns bei der folgenden Abschätzung erlaubt, anstatt der Lösung selbst die \mathcal{L} -transformierte Lösung zu betrachten.

LEMMA 2.1 (PARSEVALSCHE GLEICHUNG DER \mathcal{L} -TRANSFORMATION [23, L. 7.4.6]). Für eine stetige Funktion $U = U(x, t)$, $0 \leq x < \infty$, $t \geq 0$ mit Werten in \mathcal{C}^n gelte $U(\cdot, t) \in L^2(x)$ für jedes $t \geq 0$. Ferner habe man Abschätzungen der Form

$$\|U(\cdot, t)\|_x^2 \leq \text{konst.} \cdot e^{2\alpha_0 t} \quad \text{und} \quad |U(x, t)| \leq \text{konst.} \cdot e^{\alpha_0 t}, \quad 0 \leq x < \infty.$$

für alle $t \geq 0$. Dann gilt die Parsevalsche Gleichung

$$(2.16) \quad \int_0^\infty e^{-\alpha t} \|U(\cdot, t)\|_x^2 dt = \frac{1}{2\pi} \int_{-\infty}^\infty \|\hat{U}(\cdot, \alpha + i\xi)\|_x^2 d\xi, \quad \forall \alpha > \alpha_0.$$

LEMMA 2.2 (ABSCHÄTZUNG DER LÖSUNG). *Sei $f = 0$. Dann gibt es eine Konstante $\alpha_0 \geq 0$, so daß $e^{-\alpha t}U(x, t) \in L^2(x, t)$, $e^{-\alpha t}U(0, t) \in L^2(t)$ für $\alpha > \alpha_0$, und es existiert eine Konstante c_0 mit*

$$(2.17) \quad (\alpha - \alpha_0) \left\| e^{-\alpha t}U(0, t) \right\|_t^2 + (\alpha - \alpha_0)^2 \left\| e^{-\alpha t}U(x, t) \right\|_{x,t}^2 \\ \leq c_0^2 \left[(\alpha - \alpha_0) \left\| e^{-\alpha t}g(t) \right\|_t^2 + \left\| e^{-\alpha t}F(x, t) \right\|_{x,t}^2 \right].$$

BEWEIS. (der Abschätzung). Der Beweis folgt [18, Beweis von Satz 2.3]. Mit dem *Differentiationssatz* [7, Kapitel 9] erhält man für die Laplace-Transformation von (2.7a) unter Berücksichtigung der Anfangsbedingung (2.7b)

$$s\widehat{U}(x, s) - f(x) + \Lambda\widehat{U}_x(x, s) + B\widehat{U}(x, s) = \widehat{F}(x, s).$$

Zusammen mit der \mathcal{L} -transformierten Randbedingung (2.7c) hat man unter Verwendung des *Integrationssatzes* [7, Kapitel 8] somit

$$(2.18) \quad s\widehat{U}(x, s) = -\Lambda\widehat{U}_x(x, s) - B\widehat{U}(x, s) + \widehat{F}(x, s), \quad x > 0, \quad s = \alpha + i\xi, \quad \alpha > \alpha_0 \text{ fest,}$$

$$(2.19) \quad \widehat{U}^+(0, s) = \widehat{g}(s) - \frac{1}{s}K_0\widehat{U}^-(0, s),$$

da $f = 0$ ist. Wir definieren die Matrix $H \in \mathbb{R}^{N \times N}$ durch

$$H := \begin{pmatrix} \rho I_m & 0 \\ 0 & I_{N-m} \end{pmatrix}, \quad 0 < \rho \leq 1, \quad \rho = \text{konst.},$$

wobei ρ später geeignet gewählt wird. Wir erhalten mit $\alpha = \text{Re } s > 0$ und der Cauchy-Schwarzschen Ungleichung

$$(2.20) \quad \alpha \langle \widehat{U}, H\widehat{U} \rangle_x = \text{Re} \langle \widehat{U}, sH\widehat{U} \rangle_x \\ \stackrel{(2.18)}{=} -\text{Re} \langle \widehat{U}, H\Lambda\widehat{U}_x \rangle_x - \text{Re} \langle \widehat{U}, HB\widehat{U} \rangle_x + \text{Re} \langle \widehat{U}, H\widehat{F} \rangle_x \\ \leq R + |B| \left\| \widehat{U} \right\|_x^2 + \frac{1}{2} \alpha \langle \widehat{U}, H\widehat{U} \rangle_x + \frac{1}{2\alpha} \left\| \widehat{F} \right\|_x^2.$$

Zur Abschätzung des letzten Terms haben wir dabei eine Ungleichung der Form

$$0 \leq \langle a - b, a - b \rangle = \|a\|^2 + \|b\|^2 - 2 \text{Re} \langle a, b \rangle \\ \implies \text{Re} \langle a, b \rangle \leq \frac{1}{2} \|a\|^2 + \frac{1}{2} \|b\|^2 \\ \implies \text{Re} \langle \sqrt{\alpha} a, b/\sqrt{\alpha} \rangle \leq \frac{1}{2} \alpha \|a\|^2 + \frac{1}{2\alpha} \|b\|^2, \quad \alpha > 0,$$

sowie $\left\| H\widehat{U} \right\|_x^2 \leq \langle \widehat{U}, H\widehat{U} \rangle_x$ wegen $0 < \rho \leq 1$ benutzt, und der Term R ergibt sich mit partieller Integration zu

$$(2.21) \quad R = \frac{1}{2} \left(\widehat{U}^+(0, s) \right)^T \rho \Lambda^+ \widehat{U}^+(0, s) + \frac{1}{2} \left(\widehat{U}^-(0, s) \right)^T \Lambda^- \widehat{U}^-(0, s) \\ \leq \frac{\rho}{2} \lambda_1 \left| \widehat{U}^+(0, s) \right|^2 - \frac{|\lambda_{m+1}|}{2} \left| \widehat{U}^-(0, s) \right|^2 \\ = -\frac{\delta}{2} \left| \widehat{U}^+(0, s) \right|^2 + \frac{\rho\lambda_1 + \delta}{2} \left| \widehat{U}^+(0, s) \right|^2 - \frac{|\lambda_{m+1}|}{2} \left| \widehat{U}^-(0, s) \right|^2.$$

Es ist $\delta > 0$ eine Konstante, die später geeignet gewählt wird. Aus der \mathcal{L} -transformierten Randbedingung (2.19) erhält man die Abschätzung

$$\left| \widehat{U}^+(0, s) \right|^2 \leq 2 |\widehat{g}(s)|^2 + 2 \frac{|K_0|^2}{\alpha^2} \left| \widehat{U}^-(0, s) \right|^2, \quad \xi \in \mathbb{R}, \quad s = \alpha + i\xi,$$

die wir in (2.21) verwenden. Wir können nun δ, ρ so klein wählen, daß

$$\begin{aligned} R &\leq -\frac{\delta}{2} \left| \widehat{U}^+(0, s) \right|^2 - \underbrace{\left[\frac{|\lambda_{m+1}|}{2} - (\rho\lambda_1 + \delta) \frac{|K_0|^2}{\alpha^2} \right]}_{\geq \frac{\delta}{2}} \left| \widehat{U}^-(0, s) \right|^2 + (\rho\lambda_1 + \delta) |\widehat{g}(s)|^2 \\ &\leq -\frac{\delta}{2} \left| \widehat{U}(0, s) \right|^2 + (\rho\lambda_1 + \delta) |\widehat{g}(s)|^2 \end{aligned}$$

wird. Dies ist erfüllt, wenn

$$(2.22) \quad \left(1 + 2 \frac{|K_0|^2}{\alpha_0^2} \right) \delta \leq |\lambda_{m+1}| - 2\rho\lambda_1 \frac{|K_0|^2}{\alpha_0^2}$$

gilt. Somit wird (2.20) zu

$$\frac{1}{2} \alpha \langle \widehat{U}, H\widehat{U} \rangle_x \leq -\frac{\delta}{2} \left| \widehat{U}(0, s) \right|^2 + (\rho\lambda_1 + \delta) |\widehat{g}(s)|^2 + |B| \left\| \widehat{U} \right\|_x^2 + \frac{1}{2\alpha} \left\| \widehat{F} \right\|_x^2,$$

nach Multiplikation mit 2α ergibt sich

$$\alpha \left[\alpha \langle \widehat{U}, H\widehat{U} \rangle_x - 2|B| \left\| \widehat{U} \right\|_x^2 + \delta \left| \widehat{U}(0, s) \right|^2 \right] \leq \left\| \widehat{F} \right\|_x^2 + 2\alpha (\rho\lambda_1 + \delta) |\widehat{g}(s)|^2,$$

und mit $\rho \left\| \widehat{U} \right\|_x^2 \leq \langle \widehat{U}, H\widehat{U} \rangle_x$, da $\rho \leq 1$ ist, erhält man

$$\alpha \left[(\rho\alpha - 2|B|) \left\| \widehat{U} \right\|_x^2 + \delta \left| \widehat{U}(0, s) \right|^2 \right] \leq \left\| \widehat{F} \right\|_x^2 + 2\alpha (\rho\lambda_1 + \delta) |\widehat{g}(s)|^2.$$

Mit $\alpha_0 := (2|B| + \varepsilon)/\rho$ für beliebig kleines $\varepsilon > 0$ und einer Multiplikation mit $(\alpha - \alpha_0)/\alpha$ ergibt sich

$$(\alpha - \alpha_0)\delta \left| \widehat{U}(0, s) \right|^2 + (\alpha - \alpha_0)^2 \rho \left\| \widehat{U} \right\|_x^2 \leq \frac{\alpha - \alpha_0}{\alpha} \left\| \widehat{F} \right\|_x^2 + (\alpha - \alpha_0) 2 (\rho\lambda_1 + \delta) |\widehat{g}(s)|^2.$$

Die Parsevalsche Gleichung (2.16) liefert nun

$$\begin{aligned} (\alpha - \alpha_0)\delta \left\| e^{-\alpha t} U(0, t) \right\|_t^2 + (\alpha - \alpha_0)^2 \rho \left\| e^{-\alpha t} U(x, t) \right\|_{x,t}^2 \\ \leq (\alpha - \alpha_0) 2 (\rho\lambda_1 + \delta) \left\| e^{-\alpha t} g(t) \right\|_t^2 + \underbrace{\frac{\alpha - \alpha_0}{\alpha}}_{\leq 1} \left\| e^{-\alpha t} F(x, t) \right\|_{x,t}^2, \end{aligned}$$

und mit einer geeigneten Konstanten c_0 ergibt dies die Behauptung (2.17). \square

BEMERKUNG 2.6 (WAHL VON α_0). Bei Verwendung von ARBen (2.7c) kann man $\alpha_0 := 2|B|/\rho$ wählen, da aus $B = 0$ folgt $K_0 = 0$ und der kritische Term in (2.22) verschwindet.

FOLGERUNG 2.3. *Das Problem (2.7) ist stark sachgemäß gestellt im verallgemeinerten Sinne [23, Definition 3, Seite 227]*

2.2.2. Die Definition der Stabilität. Wir bezeichnen mit $\ell^2(x)$ den Raum aller Gitterfunktionen $U = \{U_j\}_{j=0}^\infty$, $U_j = U(x_j)$, $x_j = jh$, $j = 0, 1, 2, \dots$ mit $h \sum_{j=0}^\infty |U_j|^2 < \infty$ und definieren das innere Produkt und die Norm durch

$$(2.23) \quad \langle U, V \rangle_h := h \sum_{j=0}^\infty U_j^* V_j, \quad \|U\|_h^2 := \langle U, U \rangle_h.$$

Analog werden die Räume $\ell^2(t)$ und $\ell^2(x, t)$ definiert, und wir bezeichnen mit

$$(2.24) \quad \langle U, V \rangle_k := k \sum_{n=0}^\infty U^*(t_n) V(t_n), \quad \|U\|_k^2 := \langle U, U \rangle_k, \quad t_n = nk,$$

$$(2.25) \quad \langle U, V \rangle_{h,k} := hk \sum_{n=0}^\infty \sum_{j=0}^\infty U_j^*(t_n) V_j(t_n), \quad \|U\|_{h,k}^2 := \langle U, U \rangle_{h,k}$$

die zugehörigen inneren Produkte und Normen. Wir verwenden den folgenden *Stabilitätsbegriff*:

DEFINITION 2.4 (GKS-STABILITÄT [18, Def. 3.3]). Verschwindet die Anfangsvorgabe (2.8b), so heißt die Approximation (2.8) *GKS-stabil*, wenn es Konstanten $c_0 > 0$, $\alpha_0 \geq 0$ gibt, so daß für alle $t = t_n = nk$ und alle $\alpha > \alpha_0$ gilt

$$(2.26) \quad \frac{\alpha - \alpha_0}{\alpha k + 1} \|e^{-\alpha t} U_0\|_k^2 + \left(\frac{\alpha - \alpha_0}{\alpha k + 1} \right)^2 \|e^{-\alpha t} U\|_{h,k}^2 \\ \leq c_0^2 \left[\frac{\alpha - \alpha_0}{\alpha k + 1} \|e^{-\alpha(t+k)} g_0\|_k^2 + \|e^{-\alpha(t+k)} b\|_{h,k}^2 \right].$$

2.2.3. Das gestörte Schema. Wir werden zunächst die Stabilität des folgenden, *modifizierten* LW-Schemas analysieren, das wir erhalten, indem wir in Q und b (siehe (2.8a)) Terme der Ordnung $O(k)$, $O(k^2)$ weglassen. Es ergibt sich:

$$(2.27) \quad U_j^{n+1} = \tilde{Q} U_j^n + k d_j^n, \quad j \geq 1; \quad \tilde{Q} = \sum_{\sigma=-1}^1 \tilde{A}_\sigma E^\sigma, \\ \text{wobei} \quad \tilde{A}_0 = I - (\lambda A)^2, \quad \tilde{A}_{\pm 1} = \mp \frac{1}{2} \lambda A + \frac{1}{2} (\lambda A)^2, \\ d_j^n = \frac{1}{2} (F_j^n + F_j^{n+1}) - \frac{1}{4} \lambda A (F_{j+1}^n - F_{j-1}^n)$$

zusammen mit der Anfangsvorgabe (2.8b) und den Randbedingungen (2.8c), (2.8d).

Das Schema (2.8) ist eine Störung der Ordnung $O(k)$ von diesem Schema und zwar nur im Inneren des Gitters. Daher kann der *Störungssatz der GKS-Theorie* [18, Satz 4.3] angewendet werden, nach dem die GKS-Stabilität invariant gegenüber Störungen der Ordnung $O(k)$ ist und somit können wir bei der Stabilitätsanalyse anstelle des ursprünglichen LW-Schemas dieses *gestörte Schema* untersuchen. Dazu zerlegen wir die Matrix \tilde{Q} gemäß der Partitionierung von Λ (vgl. Notation von Seite 3) wie folgt:

$$\tilde{Q} = \begin{pmatrix} \tilde{Q}^{++} & \tilde{Q}^{+-} \\ \tilde{Q}^{-+} & \tilde{Q}^{--} \end{pmatrix}.$$

2.2.4. Die Stabilität des gestörten Schemas. Die Einström- und Ausströmprobleme sind bei dem gestörten Schema im Inneren des Gitters *entkoppelt* (\tilde{Q} enthält die „Kopplungsmatrix“ C nicht), d. h. man kann beide Probleme *sukzessive lösen*. Offensichtlich ist das gestörte Schema genau dann stabil, wenn beide Teilprobleme stabil sind. Das Ausströmproblem ist *in sich abgeschlossen* und die Lösung dieses Problems liefert die Ausströmdaten $(U^-)_{0}^{n+1}$, die man in der Einströmrandbedingung (2.8c) benötigt. Im folgenden betrachten wir die Stabilität der beiden Einzelprobleme und benutzen dabei die Z -Transformation der Lösung bezüglich des Zeitindex:

DEFINITION 2.5 (Z -TRANSFORMATION [8]). Der formale Zusammenhang zwischen einer Folge und einer komplexen Funktion, gegeben durch die Korrespondenz

$$(2.28) \quad \{u^n\} \longleftrightarrow \mathcal{Z}\{u^n\} := \hat{u}(z) := \sum_{n=0}^{\infty} z^{-n} u^n,$$

wird Z -Transformation genannt. Die Funktion $\hat{u}(z)$ heißt Z -Transformierte der Folge $\{u^n\}$, wobei $n \in \mathbb{N}_0$ ist.

Die Stabilität des Ausströmproblems. Es handelt sich um eine LW-Diskretisierung von $U_t^- + \Lambda^- U_x^- = F^-$ mit *horizontaler Ausströmextrapolation*. Dieses Schema ist nach [14, Beispiel 3.1] *GKS-stabil*, d. h. es existieren Konstanten $c_0 > 0$, $\alpha_0 \geq 0$, so daß für alle $t = t_n = nk$ und alle $\alpha > \alpha_0$ gilt

$$(2.29) \quad \frac{\alpha - \alpha_0}{\alpha k + 1} \|e^{-\alpha t} U_0^-\|_k^2 + \left(\frac{\alpha - \alpha_0}{\alpha k + 1}\right)^2 \|e^{-\alpha t} U^-\|_{h,k}^2 \leq c_0^2 \|e^{-\alpha(t+k)} d^-\|_{h,k}^2.$$

Eine Z -Transformation von

$$(U^-)_{j}^{n+1} = \tilde{Q}^{--} (U^-)_{j}^n + k (d^-)_{j}^n, \quad j = 1, 2, 3, \dots$$

und Division durch z ergibt die *Resolventengleichung*

$$(2.30) \quad \left(I - \frac{1}{z} \tilde{Q}^{--}\right) \widehat{U}_j^- = \widehat{d}_j^-$$

$$\text{mit } \widehat{U}_j^-(z) = k \sum_{n=0}^{\infty} z^{-n} (U^-)_{j}^n, \quad \widehat{d}_j^-(z) = k^2 \sum_{n=0}^{\infty} z^{-n-1} (d^-)_{j}^n, \quad z = e^{(\alpha+i\omega)k}.$$

Nach dem *Charakterisierungssatz der GKS-Stabilität* [18, Satz 4.2] gibt es Konstanten $\alpha_0^- \geq 0$, $c_1^- > 0$, so daß (2.30) für alle z mit $|z| > e^{\alpha_0^- k}$ und alle d^- eine eindeutige Lösung besitzt mit

$$(2.31) \quad k \left(\frac{|z| - e^{\alpha_0^- k}}{|z|}\right) |\widehat{U}_0^-(z)|^2 + \left(\frac{|z| - e^{\alpha_0^- k}}{|z|}\right)^2 \|\widehat{U}^-\|_h^2 \leq (c_1^-)^2 \|\widehat{d}^-\|_h^2$$

und dies liefert die Abschätzung

$$(2.32) \quad |\widehat{U}_0^-(z)|^2 \leq \frac{|z|}{k(|z| - e^{\alpha_0^- k})} (c_1^-)^2 \|\widehat{d}^-\|_h^2.$$

Die Stabilität des Einströmproblems. Eine Diskretisierung der inhomogenen absorbierenden Randbedingung (2.7c) ist gegeben durch

$$(2.33) \quad (U^+)_0^n = g^n - \frac{k}{2} K_0 \sum_{\ell=0}^{n-1} [(U^-)_0^\ell + (U^-)_0^{\ell+1}]$$

und eine Z-Transformation ergibt mit [8, Summationsregel]

$$(2.34) \quad \widehat{U}_0^+(z) = \widehat{g}(z) - \frac{k}{2} K_0 \frac{z+1}{z-1} \widehat{U}_0^-(z).$$

Das Einströmproblem ist eine LW-Diskretisierung von $U_t^+ + \Lambda^+ U_x^+ = F^+$ mit Einströmdaten $U_0^+ = \tilde{h}^+$. Es ist *stabil* [14, Lemma 2.3], und nach dem *Charakterisierungssatz* [18, Satz 4.2] existieren somit Konstanten $\alpha_0^+ \geq 0$, $c_1^+ > 0$, so daß die folgende Abschätzung

$$(2.35) \quad k \left(\frac{|z| - e^{\alpha_0^+ k}}{|z|} \right) |\widehat{U}_0^+(z)|^2 + \left(\frac{|z| - e^{\alpha_0^+ k}}{|z|} \right)^2 \|\widehat{U}^+\|_h^2 \leq (c_1^+)^2 \left[k \frac{|z| - e^{\alpha_0^+ k}}{|z|} |\widehat{h}^+|^2 + \|\widehat{d}^+\|_h^2 \right]$$

für alle z mit $|z| > e^{\alpha_0^+ k}$ gilt. Für die Stabilität brauchen wir allerdings eine Abschätzung mit Termen von \widehat{g} , \widehat{d} (und nicht \widehat{h}^+).

Wir leiten zunächst eine *Abschätzung der Einströmrandbedingung* her. Wir wählen $\alpha_0 > \max(\alpha_0^-, \alpha_0^+) \geq 0$ und mit $e^{\alpha_0 k} < |z| \leq |z-1| + 1$ gilt

$$k \left| \frac{z+1}{z-1} \right| \leq k \frac{|z-1|+2}{|z-1|} \leq k \left[1 + \frac{2}{|z-1|} \right] \leq k + \frac{2k}{e^{\alpha_0 k} - 1} \leq k + \frac{2}{\alpha_0},$$

denn $e^{\alpha_0 k} - 1 \geq \alpha_0 k$. Somit gilt mit einer Konstanten c_2

$$(2.36) \quad \frac{k^2}{2} \left| \frac{z+1}{z-1} \right|^2 \leq c_2,$$

da wir uns auf den interessanten Bereich $0 < k \leq 1$ beschränken können.

Die Z-transformierte Randbedingung (2.34) liefert zusammen mit (2.36) und (2.32)

$$(2.37) \quad \begin{aligned} |\widehat{U}_0^+(z)|^2 &\leq 2 |\widehat{g}(z)|^2 + \frac{k^2}{2} \left| \frac{z+1}{z-1} \right|^2 |K_0|^2 |\widehat{U}_0^-(z)|^2 \\ &\leq 2 |\widehat{g}(z)|^2 + c_2 |K_0|^2 |\widehat{U}_0^-(z)|^2 \\ &\leq 2 |\widehat{g}(z)|^2 + \frac{|z|}{k (|z| - e^{\alpha_0 k})} c_2 (c_1^-)^2 |K_0|^2 \|\widehat{d}^-\|_h^2, \quad \forall |z| > e^{\alpha_0 k}. \end{aligned}$$

BEMERKUNG 2.7. Gilt die Abschätzung (2.35) für ein $\alpha_0^+ \geq 0$, so gilt sie auch für alle $\tilde{\alpha}_0^+ > \alpha_0^+$. Dies folgt unmittelbar, wenn man diese Ungleichung durch $(|z| - e^{\alpha_0^+ k})/|z|$ dividiert. Eine analoge Aussage gilt für (2.31).

FOLGERUNG 2.4. Die Abschätzungen (2.31), (2.35) gelten auch für das oben gewählte $\alpha_0 > 0$ anstelle von α^- bzw. α^+ .

Mit $\widehat{h}^+ = \widehat{U}_0^+$ setzen wir (2.37) in (2.35) (mit α_0) ein und erhalten

$$(2.38) \quad k \left(\frac{|z| - e^{\alpha_0 k}}{|z|} \right) |\widehat{U}_0^+(z)|^2 + \left(\frac{|z| - e^{\alpha_0 k}}{|z|} \right)^2 \|\widehat{U}^+\|_h^2 \\ \leq (c_1^+)^2 \left[k \frac{|z| - e^{\alpha_0 k}}{|z|} 2 |\widehat{g}(z)|^2 + \|\widehat{d}^+\|_h^2 + c_2 (c_1^-)^2 |K_0|^2 \|\widehat{d}^-\|_h^2 \right],$$

womit die GKS–Stabilität des gestörten Schemas und damit auch die des ursprünglichen LW–Schemas (2.8) bewiesen ist. Denn es gilt:

BEMERKUNG 2.8. Man sieht leicht, daß die diskreten Einströmrandsbedingungen (2.8c) und (2.33) äquivalent sind und somit das Schema (2.8) GKS–stabil ist.

BEMERKUNG 2.9. Der obige Stabilitätsbeweis überträgt sich auf alle expliziten, dissipativen, zweistufigen, linearen Einschrittverfahren mit horizontaler Ausströmextrapolation (2.5), wobei wir unter *Stufen* die bei einem Verfahrensschritt benutzten Zeitebenen verstehen wollen.

2.3. Numerische Ergebnisse

In diesem Abschnitt wollen wir die in Abschnitt 2.2 gezeigte Stabilität des LW-Schemas anhand eines Beispiels numerisch verifizieren und anschließend die „Qualität“ der absorbierenden Randbedingungen untersuchen.

2.3.1. Das Testbeispiel. Wir betrachten als Beispiel das hyperbolische System

$$(2.39) \quad U_t + \underbrace{\begin{pmatrix} v+c & 0 & 0 \\ 0 & v & 0 \\ 0 & 0 & v-c \end{pmatrix}}_{=\Lambda} U_x + \underbrace{\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}}_{=C} U = 0, \quad 0 \leq x \leq 1, t \geq 0$$

mit $v = 0.2$, $c = 1$ und absorbierenden Randbedingungen 1. Ordnung. Es ist somit $m = 2$, $N = 3$. Als Anfangsverteilung wählen wir

$$(2.40) \quad U_\ell(x, 0) = f_\ell(x) := \begin{cases} \cos^2\left(\frac{\pi}{2} \frac{x-0.5}{0.45}\right) & |x - 0.5| < 0.45, \\ 0 & \text{sonst} \end{cases} \quad \ell = 1, 2, 3.$$

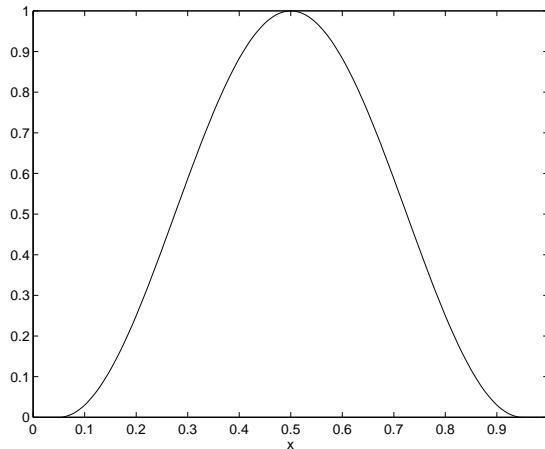


ABBILDUNG 2.2. Anfangsverteilung (2.40)

Die Matrizen $K_0 \in \mathbb{R}^{2 \times 1}$, $K_1 \in \mathbb{R}^{1 \times 2}$ in den absorbierenden Randbedingungen sind

$$(2.41) \quad K_0 = \begin{pmatrix} \frac{c-v}{c} c_{13} \\ \frac{2c}{c} v \\ \frac{c-v}{c} c_{23} \end{pmatrix} = \begin{pmatrix} 0.4 \\ 0 \\ 0 \end{pmatrix}, \quad K_1 = \left(\frac{v+c}{2c} c_{31} \quad \frac{v}{c} c_{32} \right) = (0.6 \quad 0).$$

Als zweites Beispiel wollen wir (2.39) mit dem schiefsymmetrischen Koeffizienten

$$(2.42) \quad C = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 2 \\ -1 & -2 & 0 \end{pmatrix}$$

betrachten. Die Matrizen K_0 , K_1 ergeben sich in diesem Fall zu

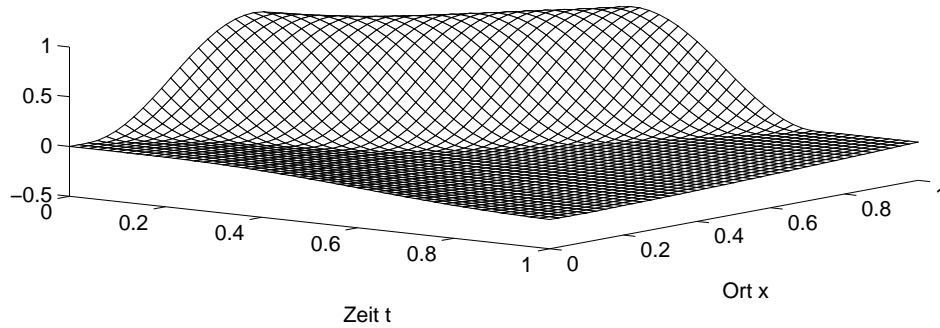
$$(2.43) \quad K_0 = \begin{pmatrix} 0.4 \\ 1.6 \end{pmatrix}, \quad K_1 = (-0.6 \quad -0.4).$$

Als CFL-Bedingung (2.13) erhält man für beide Testprobleme

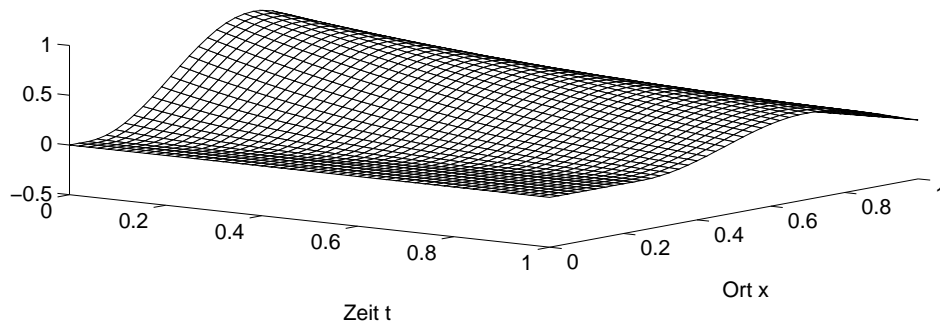
$$(2.44) \quad \lambda \leq \frac{5}{6} = 0.8\bar{3}.$$

Die folgenden zwei Abbildungen 2.3 und 2.4 zeigen die zeitliche Entwicklung der Lösung des ersten bzw. zweiten Testproblems bei ARBen 1. Ordnung. Man erkennt, wie sich die Komponenten U_ℓ mit ihren charakteristischen Geschwindigkeiten λ_ℓ fortbewegen. Dabei verwenden wir eine Zeitschrittweite $\Delta t = 0.01$ und ein Schrittweitenverhältnis $\lambda = 0.8$.

1. Komponente



2. Komponente



3. Komponente

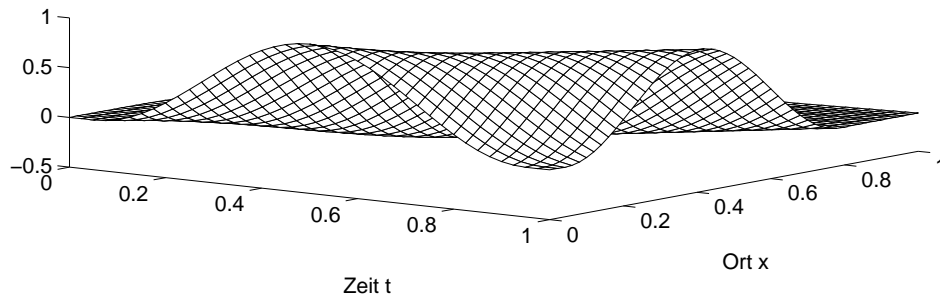


ABBILDUNG 2.3. Lösung des 1. Beispiels bei absorb. Randbedingung 1. Ordnung

BEMERKUNG 2.10. Der Unterschied zur jeweiligen Lösung bei homogenen Dirichlet-Randbedingungen ist zu gering, um ihn mit dieser Darstellungsweise zu verdeutlichen.

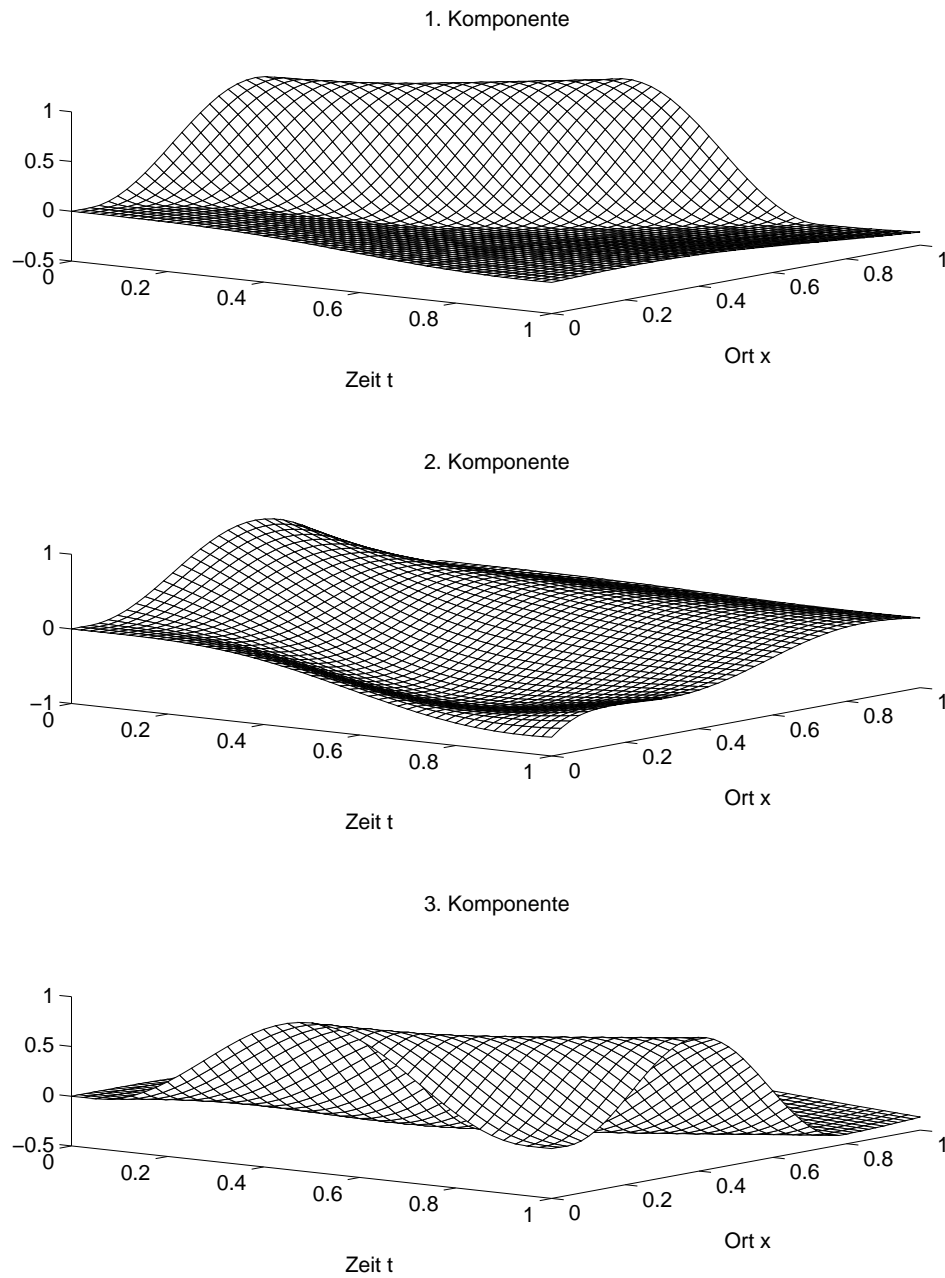


ABBILDUNG 2.4. Lösung des 2. Beispiels bei absorb. Randbedingung 1. Ordnung

2.3.2. Numerische Verifikation der Stabilität. Für diesen Abschnitt *modulieren* wir die Anfangsverteilung (2.40) mit Hilfe einer Sinusfunktion folgendermaßen:

$$(2.45) \quad U(x, 0) = \sin(2\pi px)f(x), \quad x \in [0, L].$$

Der Parameter p gibt dabei die Anzahl der Perioden des Sinustermes auf $0 \leq x \leq 1$ an. Wir wählen ein Schrittweitenverhältniss λ , das die CFL-Bedingung (2.44) erfüllt, womit das *innere Schema* stabil ist. Denn wir wollen nur die Auswirkungen der ARBen auf die Stabilität betrachten. Da etwaige Stabilitätsprobleme stark *frequenzabhängig* sein würden, benutzen wir verschiedene Modulationen zur numerischen Untersuchung.

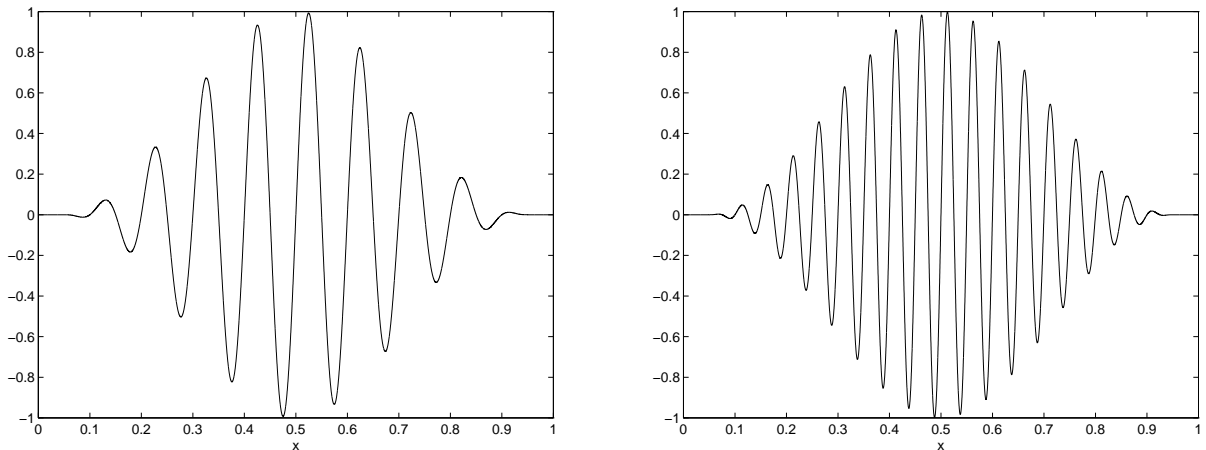


ABBILDUNG 2.5. modulierte Anfangsverteilung (2.45) für $p = 10$ bzw. $p = 20$

Neben der CFL-Bedingung muß man beachten, daß die Ortsschrittweite h klein genug ist, um die hohen Frequenzen der *Modulation* aufzulösen. Es sollen je Periode des Sinustermes mindestens 8 Gitterpunkte liegen. Nach Wahl des Schrittweitenverhältnisses λ ergibt diese Faustregel mit $L = 1$ folgende *Restriktion an die Zeitschrittweite* k :

$$(2.46) \quad h = \frac{k}{\lambda} \leq \frac{1}{8} \frac{L}{p} \quad \text{und somit} \quad k \leq \frac{\lambda}{8p}.$$

Die zeitliche Entwicklung der 3. Komponente bei modulierter Anfangsvorgabe zeigt die Abbildung 2.6 anhand des 1. Beispiels bei $p = 8$, $\lambda = 0.8$ und $\Delta t = 0.01$. Für die 3. Komponente ist $x = 1$ Einströmrand, und es läßt sich erkennen, wie sich dort die Modulation in Form von Wellen auswirkt und diese von der absorbierenden Randbedingung 1. Ordnung zumindest nicht wesentlich reflektiert werden.

Im folgenden geben wir den zeitlichen Verlauf der *diskreten L^2 -Norm*

$$(2.47) \quad \|U\|_h^2 := h \sum_{j=0}^J |U_j|^2$$

jeder einzelnen Komponente an, um zu verhindern, daß Schwankungen sich gegenseitig kompensieren oder Schwankungen in einer Komponente durch eine wesentlich größere L^2 -Norm einer anderen Komponente verdeckt werden. Hierbei ist die L^2 -Norm für verschiedene Zeitschrittweiten $k = \Delta t$ angegeben. In den jeweils darunterliegenden Abbildungen ist die *gesamte L^2 -Norm* bei gleichen Zeitschrittweiten wie zuvor dargestellt.

Die Abbildungen 2.7 und 2.8 zeigen die Resultate für $\lambda = 0.8$ und $p = 10$ bzw. $p = 20$. Die Ergebnisse für das 2. Beispiel sind in Abbildung 2.9 und Abbildung 2.10 zu sehen.

3. Komponente

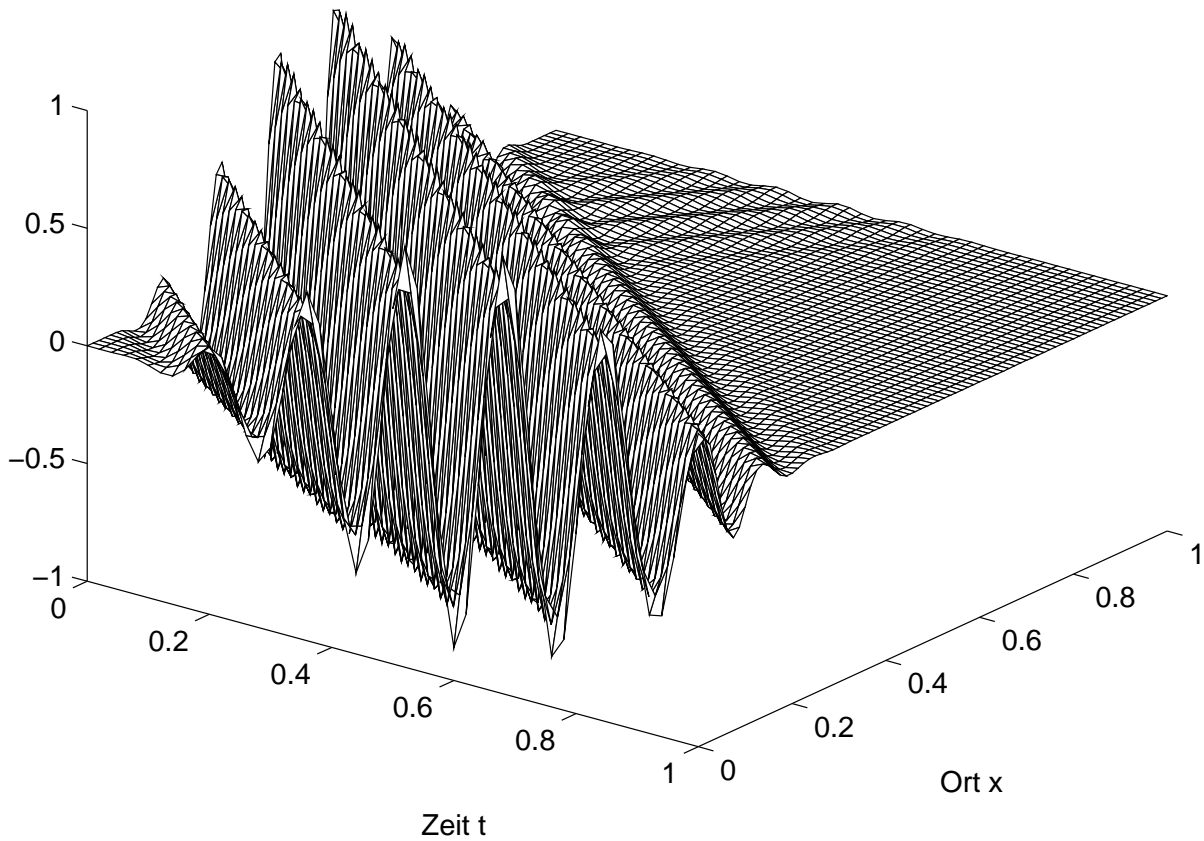


ABBILDUNG 2.6. 3. Komponente der Lösung des 1. Beispiels für $p = 8$

FAZIT (STABILITÄT). Der „gezackte“ Kurvenverlauf entsteht bei dem Passieren der hohen Frequenzen der Modulation durch den Rand. Die L^2 -Normen klingen rasch ab, was zu erwarten ist, da das „Wellenpaket“ in beide Richtungen aus dem Intervall $[0, 1]$ hinausläuft. Die ARBen verhindern dabei weitgehend Reflexionen. Letzteres ist hierbei der wesentliche Effekt und illustriert die in Abschnitt 2.2 gezeigte Stabilität des Schemas.

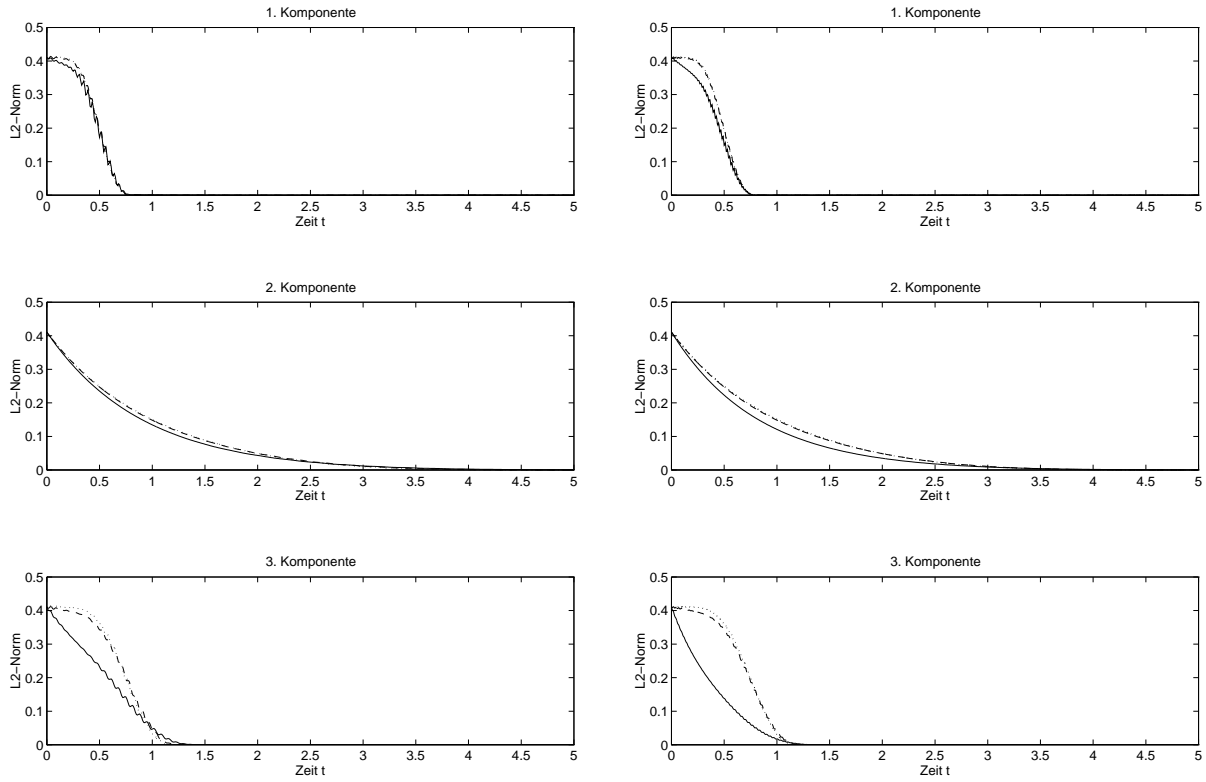


ABBILDUNG 2.7. L^2 -Normen jeder Komponente für $p = 10$ bzw. $p = 20$

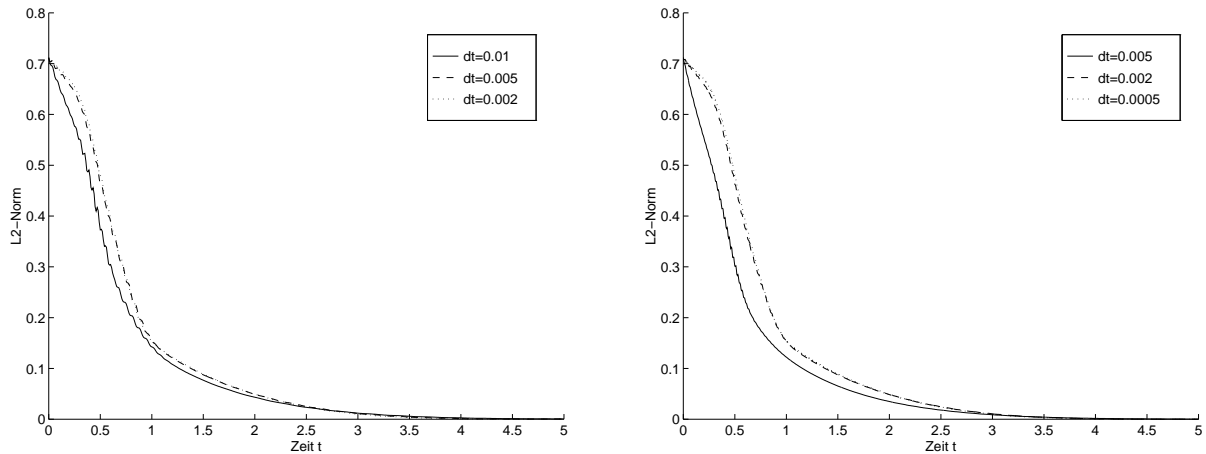


ABBILDUNG 2.8. L^2 -Normen 1. Beispiel für $p = 10$ bzw. $p = 20$

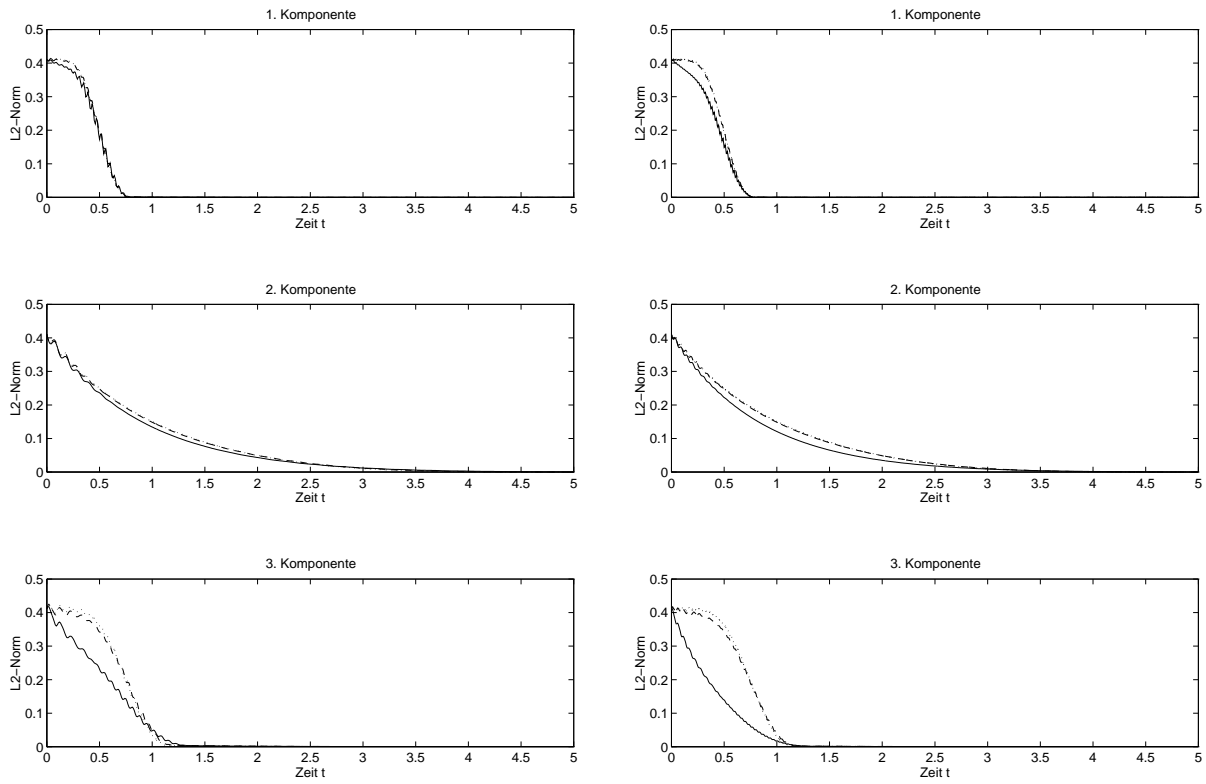


ABBILDUNG 2.9. L^2 -Normen jeder Komponente für $p = 10$ bzw. $p = 20$

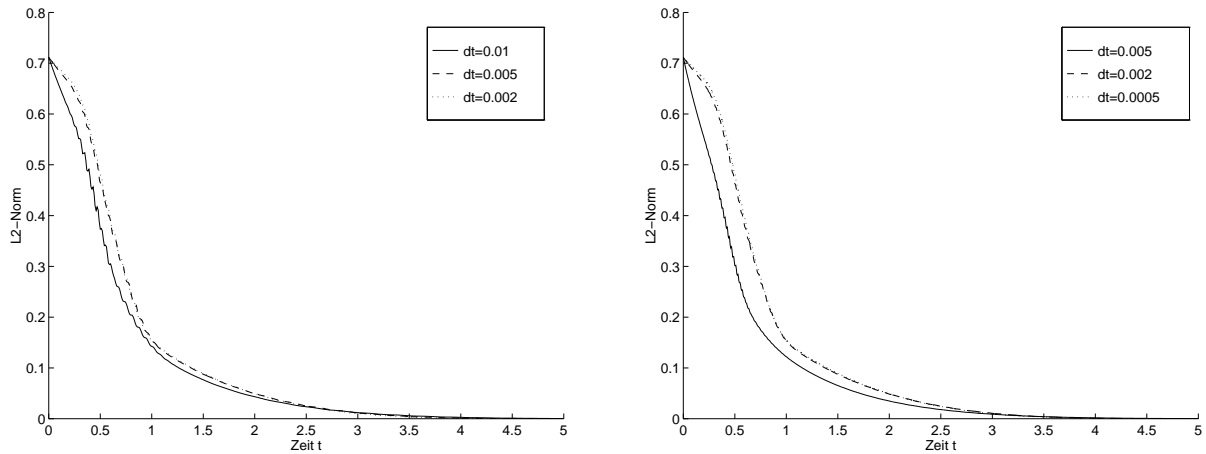


ABBILDUNG 2.10. L^2 -Normen 2. Beispiel für $p = 10$ bzw. $p = 20$

2.3.3. Die Absorptionsqualität der absorbierenden Randbedingungen. In diesem Abschnitt wollen wir die absorbierenden Randbedingungen 0. und 1. Ordnung miteinander vergleichen. Eine Diskretisierung der ARB 0. Ordnung (1.19) an $x = 0$ liefert

$$(2.48) \quad (U^+)_0^{n+1} = (U^+)_0^n.$$

Da die Anfangsverteilung (2.40) einen kompakten Träger besitzt, entspricht (2.48) einer *homogenen Dirichlet-Randbedingung*. Die diskrete ARB 1. Ordnung lautet

$$\frac{(U^+)_0^{n+1} - (U^+)_0^n}{k} + K_0(U^-)_0^{n+\frac{1}{2}} = 0$$

und damit ergibt sich

$$(2.49) \quad (U^+)_0^{n+1} = (U^+)_0^n - \frac{k}{2} K_0 \left((U^-)_0^n + (U^-)_0^{n+1} \right).$$

Wir bestimmen die durch Einführung der Ränder hervorgerufenen numerischen Reflexionen, den sog. *reflektierten Anteil*, indem wir an den Gitterpunkten x_j , $j = 0, 1, \dots, J$ die Differenz zwischen der Lösung des ARWP und der des reinen AWP (theoretisch auf $x \in \mathbb{R}$) bilden. Dabei wird letztere Lösung durch eine Berechnung auf einem wesentlich größeren x -Intervall simuliert, so daß die Lösung in der vorgegebenen Zeit den Rand dieses Gitters nicht erreicht. Den reflektierten Anteil, den wir mit dem Index „*refl*“ kennzeichnen, geben wir in den folgenden Abbildungen in der *relativen L^2 -Norm*

$$(2.50) \quad \|U\|_{h,rel} := \frac{\|U\|_h}{\|U^0\|_h}$$

für jede Komponente einzeln wieder. In den Abbildungen 2.11 und 2.12 wurden dabei die Parameterwerte $\lambda = 0.8$, $\Delta t = 0.001$ verwendet.

BEMERKUNG 2.11. Beim 1. Beispiel sind die Randbedingungen (2.48) und (2.49) für die 2. Komponente identisch.

FAZIT (ABSORPTIONSQUALITÄT). Für beide Testprobleme zeigt sich, daß die numerischen Reflexionen der ARBen 1. Ordnung deutlich geringer sind als die bei der homogenen Dirichlet-Randbedingung. Die Verbesserung liegt beim 1. Beispiel beim Faktor 50 bzw. 75 und beim 2. Beispiel zwischen Faktor 3 und 8.

BEMERKUNG 2.12. Dieses Verhalten gilt für kleine Endzeiten T . Die ARBen sind nur zeitlich lokal gute Approximationen; für größere Zeitintervalle können Dirichlet-Randbedingungen durchaus besser abschneiden. Das reine AWP konvergiert lokal (auf $0 < x < 1$) gegen Null. Daher ist anzunehmen, daß dieses *Langzeitverhalten* von der homogenen Dirichlet-Randbedingung besser widerspiegelt wird als von den ARBen 1. Ordnung (siehe starkes Anwachsen des reflektierten Anteils für $t \approx 1$).

In der folgenden Abbildung 2.13 für das 2. Beispiel auf dem Zeitintervall $[0, 10]$ mit den Parametern $\lambda = 0.8$, $\Delta t = 0.01$ sieht man, daß der reflektierte Anteil im 2. Beispiel für größere Zeiten bei beiden Randbedingungen praktisch gleich groß ist. Zwischenzeitlich ist der reflektierte Anteil der Dirichlet-RB sogar kleiner als der bei Verwendung der ARB 1. Ordnung .

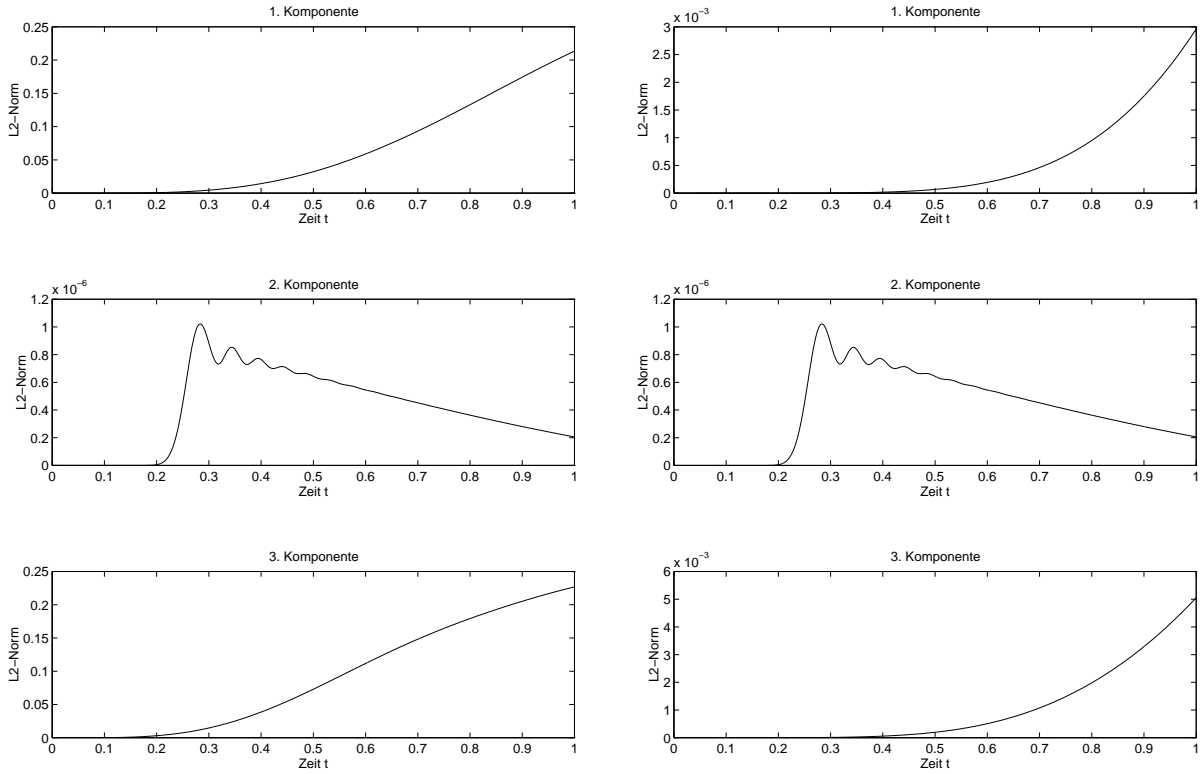


ABBILDUNG 2.11. Beispiel 1: Reflektierter Anteil für homogene Dirichlet-RB bzw. ARB 1. Ordnung bei $\lambda = 0.8$, $\Delta t = 0.001$

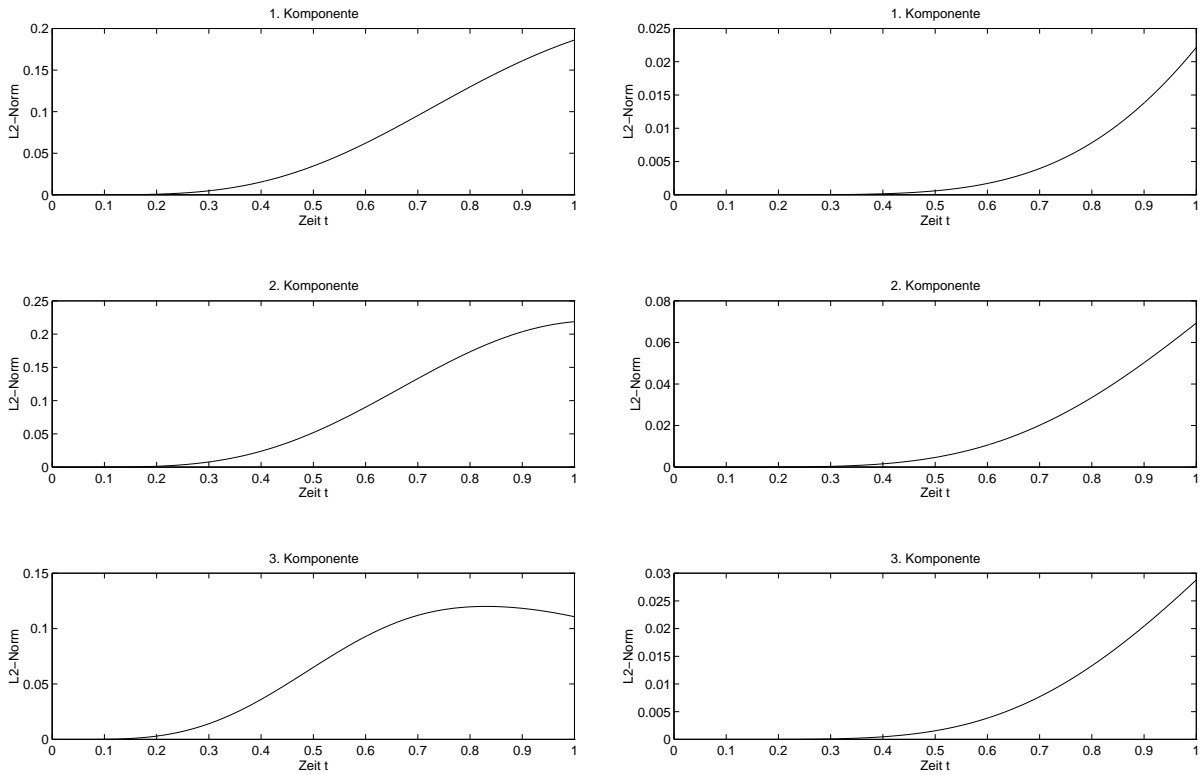
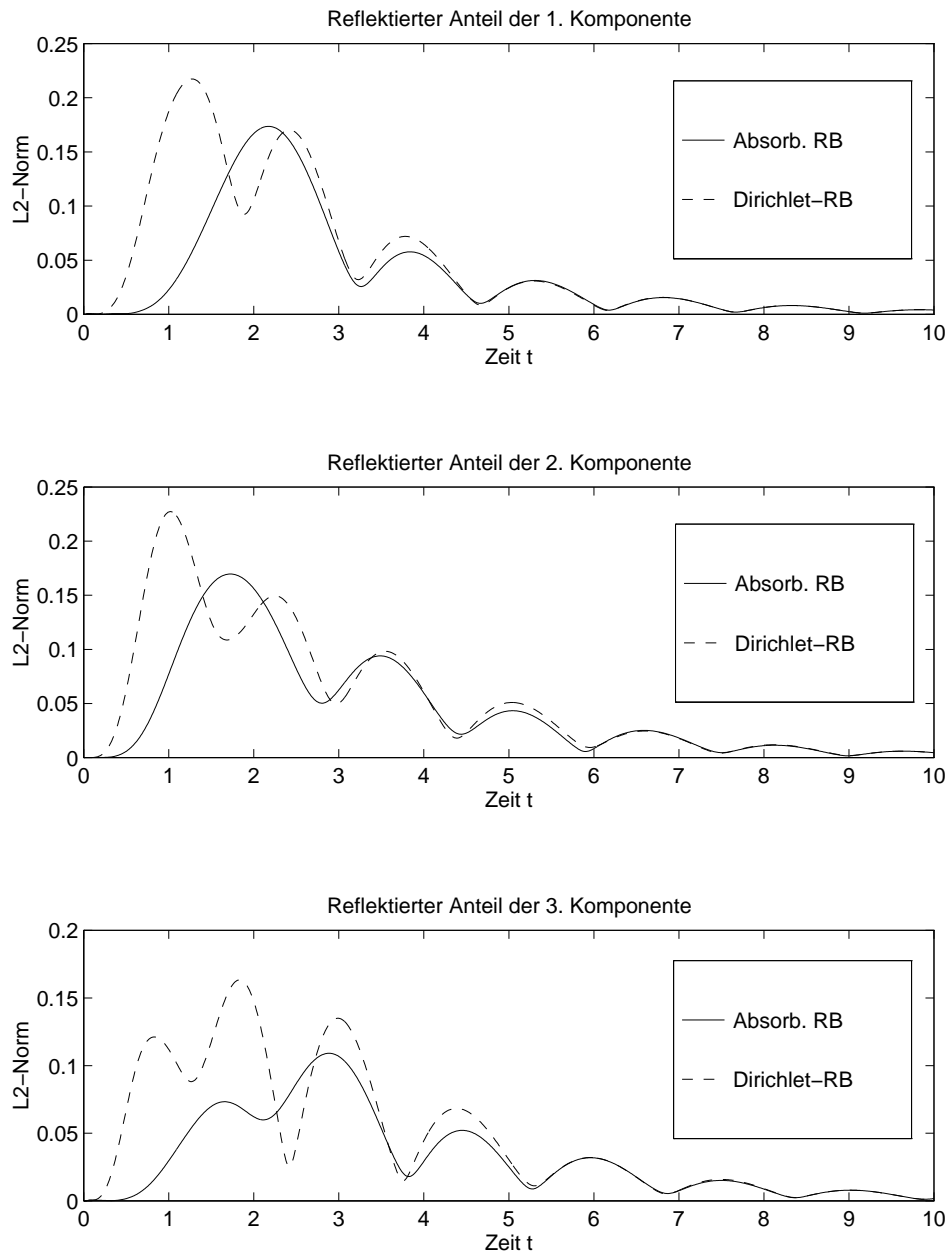


ABBILDUNG 2.12. Beispiel 2: Reflektierter Anteil für homogene Dirichlet-RB bzw. ARB 1. Ordnung bei $\lambda = 0.8$, $\Delta t = 0.001$

ABBILDUNG 2.13. Beispiel 2: Reflektierter Anteil für $t \in [0, 10]$, $\Delta t = 0.01$

KAPITEL 3

Numerische Betrachtung des linearen Falls in 2D

In diesem Kapitel wollen wir das *LW-Verfahren* für lineare hyperbolische Systeme in 2D

$$(3.1) \quad U_t + AU_x + BU_y + CU = F(x, y, t), \quad 0 \leq x \leq L, \quad -\infty < y < \infty, \quad t \geq 0,$$

wobei A, B, C konstante $N \times N$ -Matrizen sind, betrachten. Wie in einer Ortsdimension läßt sich der Fall von orts- und zeitabhängigen Koeffizienten analog untersuchen. Wir betrachten hier nur den Fall $\Delta x = \Delta y =: h$ mit einem konstanten *Schrittweitenverhältnis* $\lambda := k/h$. Ferner führen wir folgende *Gitterpunkte* ein:

$$\begin{aligned} x_j &= jh, & j &= 0, 1, 2, \dots, J \\ y_\ell &= \ell h, & \ell &= \dots, -1, 0, 1, 2, \dots \\ t_n &= nk, & n &= 0, 1, 2, \dots \end{aligned}$$

Die *Approximationen* der exakten Lösung $U(x_j, y_\ell, t_n)$ an den diskreten Gitterpunkten (x_j, y_ℓ) , die das LW-Verfahren zum Zeitpunkt t_n liefert, werden wir mit $U_{j,\ell}^n \in \mathbb{R}^N$ bezeichnen.

3.1. Das numerische Schema für 2D

Die Lax-Wendroff-Methode basiert wie in Abschnitt 2.1 auf der Taylorentwicklung

$$(3.2) \quad U(x, y, t + k) = U(x, y, t) + kU_t(x, y, t) + \frac{1}{2}k^2U_{tt}(x, y, t) + O(k^3),$$

wobei man U_{tt} aus (3.1) wie folgt bestimmt

$$\begin{aligned} U_{tt} &= (-AU_x - BU_y - CU + F)_t \\ &= -AU_{tx} - BU_{ty} - CU_t + F_t \\ &= A(AU_x + BU_y + CU - F)_x + B(AU_x + BU_y + CU - F)_y \\ &\quad + C(AU_x + BU_y + CU - F) + F_t \\ &= A^2U_{xx} + (AB + BA)U_{xy} + B^2U_{yy} + (AC + CA)U_x + (BC + CB)U_y \\ &\quad + C^2U + F_t - AF_x - BF_y - CF, \end{aligned}$$

so daß sich (3.2) ergibt zu

$$\begin{aligned} U(x, y, t + k) &= U(x, y, t) - k(AU_x(x, y, t) + BU_y(x, y, t) + CU(x, y, t) - F(x, y, t)) \\ &\quad + \frac{1}{2} k^2 (A^2 U_{xx}(x, y, t) + (AB + BA)U_{xy}(x, y, t) + B^2 U_{yy}(x, y, t) \\ &\quad + (AC + CA)U_x(x, y, t) + (BC + CB)U_y(x, y, t) + C^2 U(x, y, t) \\ &\quad + F_t(x, y, t) - AF_x(x, y, t) - BF_y(x, y, t) - CF(x, y, t)) + O(k^3). \end{aligned}$$

Es werden analog zu Abschnitt 2.1 zentrale Differenzenquotienten benutzt, um die oben auftretenden Ableitungen von U zu diskretisieren. Das *resultierende Schema* lautet damit

$$\begin{aligned} U_{j,\ell}^{n+1} &= U_{j,\ell}^n - k \left(A \frac{U_{j+1,\ell}^n - U_{j-1,\ell}^n}{2h} + B \frac{U_{j,\ell+1}^n - U_{j,\ell-1}^n}{2h} + CU_{j,\ell}^n - F_{j,\ell}^n \right) \\ &\quad + \frac{1}{2} k^2 \left(A^2 \frac{U_{j+1,\ell}^n - 2U_{j,\ell}^n + U_{j-1,\ell}^n}{h^2} + B^2 \frac{U_{j,\ell+1}^n - 2U_{j,\ell}^n + U_{j,\ell-1}^n}{h^2} \right. \\ &\quad + (AB + BA) \frac{U_{j+1,\ell+1}^n - U_{j+1,\ell-1}^n - U_{j-1,\ell+1}^n + U_{j-1,\ell-1}^n}{4h^2} + C^2 U_{j,\ell}^n \\ &\quad + (AC + CA) \frac{U_{j+1,\ell}^n - U_{j-1,\ell}^n}{2h} + (BC + CB) \frac{U_{j,\ell+1}^n - U_{j,\ell-1}^n}{2h} \\ &\quad \left. + \frac{F_{j,\ell}^{n+1} - F_{j,\ell}^n}{k} - A \frac{F_{j+1,\ell}^n - F_{j-1,\ell}^n}{2h} - B \frac{F_{j,\ell+1}^n - F_{j,\ell-1}^n}{2h} - CF_{j,\ell}^n \right) \\ &= U_{j,\ell}^n - \frac{1}{2} \lambda \left(A(U_{j+1,\ell}^n - U_{j-1,\ell}^n) + B(U_{j,\ell+1}^n - U_{j,\ell-1}^n) \right) - kCU_{j,\ell}^n \\ &\quad + \frac{1}{2} \lambda^2 \left(A^2 (U_{j+1,\ell}^n - 2U_{j,\ell}^n + U_{j-1,\ell}^n) + B^2 (U_{j,\ell+1}^n - 2U_{j,\ell}^n + U_{j,\ell-1}^n) \right) \\ &\quad + \frac{1}{8} \lambda^2 (AB + BA) (U_{j+1,\ell+1}^n - U_{j+1,\ell-1}^n - U_{j-1,\ell+1}^n + U_{j-1,\ell-1}^n) + \frac{1}{2} (kC)^2 U_{j,\ell}^n \\ &\quad + \frac{1}{4} \lambda k \left((AC + CA) (U_{j+1,\ell}^n - U_{j-1,\ell}^n) + (BC + CB) (U_{j,\ell+1}^n - U_{j,\ell-1}^n) \right) \\ &\quad + \frac{1}{2} k (F_{j,\ell}^{n+1} + F_{j,\ell}^n) - \frac{1}{4} \lambda k \left(A (F_{j+1,\ell}^n - F_{j-1,\ell}^n) + B (F_{j,\ell+1}^n - F_{j,\ell-1}^n) \right) - \frac{1}{2} k^2 C F_{j,\ell}^n. \end{aligned}$$

BEMERKUNG 3.1. Engquist und Majda stellten in [10, Seite 641] die Vermutung auf, daß das Problem (3.1) mit den ARBen (1.21) an $x = 0$ und $x = L$ *sachgemäß gestellt* ist. Für $N = 2$ kann man die sachgemäße Gestelltheit nicht allzu schwer zeigen [2]. Auch numerische Evidenz unterstützt diese Vermutung stark.

Wir werden die Stabilität dieses Schemas mit absorbierenden Randbedingungen nicht beweisen, sondern sie nur anhand eines Beispiels numerisch testen.

BEMERKUNG 3.2 (STABILITÄT FÜR DAS REINE AWP). Nach [29, Seite 182] ist dieses LW-Schema bei $C = 0$ für das reine AWP *stabil* (im Sinne von [24, Seite 383]), falls

$$(3.3) \quad \left| \frac{\lambda_{\max} k}{h} \right| \leq \frac{1}{2\sqrt{2}}$$

gilt. Dabei ist $\lambda_{\max} = \max(\rho(A), \rho(B))$, also der betragsmäßig größte Eigenwert von A und B . Dies ist eine stärkere Einschränkung als die CFL-Bedingung $|\lambda_{\max} k/h| \leq 1$. Eine Verallgemeinerung der Bedingung (3.3) auf $d \geq 1$ Raumdimensionen findet man in [42].

3.2. Die linearisierten Shallow-Water Gleichungen in 2D

Die Shallow-Water Gleichungen beschreiben die Bewegung eines inkompressiblen Fluids oberhalb einer festen Berandung, wenn die Tiefe klein ist im Vergleich zu einer typischen horizontalen Länge (engl. shallow $\hat{=}$ seicht). Die nichtlinearen Gleichungen lauten:

$$(3.4) \quad \begin{aligned} u_t + uu_x + vu_y + \varphi_x - fv &= 0 \\ v_t + uv_x + vv_y + \varphi_y + fu &= 0 \\ \varphi_t + (\varphi u)_x + (\varphi v)_y &= 0. \end{aligned}$$

Die linearisierte Form erhält man aus diesen Gleichungen, indem man um eine konstante Lösung $\bar{U} = (\bar{u}, \bar{v}, \bar{\varphi})^T$ herum linearisiert. Für eine Lösung U von (3.4) wird dazu der *Störungsansatz* $U = \bar{U} + U'$ mit einer kleinen Störung $U' = (u', v', \varphi')^T$ gemacht. Man setzt U in (3.4) ein und vernachlässigt Terme von quadratischer Ordnung in der Störung. Die *linearisierte Form* hat in 2 Ortsdimensionen die folgende Gestalt:

$$(3.5) \quad \begin{pmatrix} u' \\ v' \\ \varphi' \end{pmatrix}_t + \begin{pmatrix} \bar{u} & 0 & 1 \\ 0 & \bar{u} & 0 \\ \bar{\varphi} & 0 & \bar{u} \end{pmatrix} \begin{pmatrix} u' \\ v' \\ \varphi' \end{pmatrix}_x + \begin{pmatrix} \bar{v} & 0 & 0 \\ 0 & \bar{v} & 1 \\ 0 & \bar{\varphi} & \bar{v} \end{pmatrix} \begin{pmatrix} u' \\ v' \\ \varphi' \end{pmatrix}_y + \begin{pmatrix} 0 & -f & 0 \\ f & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} u' \\ v' \\ \varphi' \end{pmatrix} = 0.$$

Hier ist $(u, v)^T = (\bar{u} + u', \bar{v} + v')^T$ die *Strömungsgeschwindigkeit* und $\varphi = \bar{\varphi} + \varphi'$ das *Erdpotential*, das bis auf einen Faktor der Höhe der Fluidoberfläche entspricht. $f = f(x, y)$ ist der *Coriolis-Parameter*. Die Systemmatrizen der linearisierten Shallow-Water Gleichungen (3.5) lassen sich durch Multiplikation von links mit der wegen $\bar{\varphi} > 0$ positiv definiten Diagonalmatrix $\text{diag}(\bar{\varphi}, \bar{\varphi}, 1)$ simultan symmetrisieren [19], d. h. bei (3.5) handelt es sich um ein *symmetrisierbares hyperbolisches System* (vgl. Definition 1.2). Allerdings besitzt die Zeitableitung dann einen Koeffizienten ungleich der Identität. Um diesen Koeffizienten zu verhindern, nehmen wir eine Variablentransformation vor. Mit $S := \text{diag}(\sqrt{\bar{\varphi}}, \sqrt{\bar{\varphi}}, 1)$ ergibt sich

$$S \begin{pmatrix} \bar{u} & 0 & 1 \\ 0 & \bar{u} & 0 \\ \bar{\varphi} & 0 & \bar{u} \end{pmatrix} S^{-1} = \begin{pmatrix} \bar{u} & 0 & \sqrt{\bar{\varphi}} \\ 0 & \bar{u} & 0 \\ \sqrt{\bar{\varphi}} & 0 & \bar{u} \end{pmatrix} \quad \text{und} \quad S \begin{pmatrix} \bar{v} & 0 & 0 \\ 0 & \bar{v} & 1 \\ 0 & \bar{\varphi} & \bar{v} \end{pmatrix} S^{-1} = \begin{pmatrix} \bar{v} & 0 & 0 \\ 0 & \bar{v} & \sqrt{\bar{\varphi}} \\ 0 & \sqrt{\bar{\varphi}} & \bar{v} \end{pmatrix},$$

so daß sich das System (3.5) transformieren läßt zu

$$V_t + \begin{pmatrix} \bar{u} & 0 & c \\ 0 & \bar{u} & 0 \\ c & 0 & \bar{u} \end{pmatrix} V_x + \begin{pmatrix} \bar{v} & 0 & 0 \\ 0 & \bar{v} & c \\ 0 & c & \bar{v} \end{pmatrix} V_y + \begin{pmatrix} 0 & -f & 0 \\ f & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} V = 0,$$

wobei $V := SU' = (u'\sqrt{\bar{\varphi}}, v'\sqrt{\bar{\varphi}}, \varphi')^T = (cu', cv', \varphi')^T$ ist. Die Größe $c := \sqrt{\bar{\varphi}}$ kann man als „*Schallgeschwindigkeit*“ interpretieren. Zur Vereinfachung werden die Striche zur Bezeichnung der Störungsgrößen weggelassen. Mit $a := \bar{u}$, $b := \bar{v}$ erhalten wir

$$(3.6) \quad V_t + \begin{pmatrix} a & 0 & c \\ 0 & a & 0 \\ c & 0 & a \end{pmatrix} V_x + \begin{pmatrix} b & 0 & 0 \\ 0 & b & c \\ 0 & c & b \end{pmatrix} V_y + \begin{pmatrix} 0 & -f & 0 \\ f & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} V = 0$$

mit $V = (cu, cv, \varphi)^T$ und den *physikalischen Restriktionen* $0 < a^2 + b^2 < c^2$, $c > 0$. Weiterhin nehmen wir zur Vereinfachung an, daß f in (3.6) konstant ist.

Die Gleichungen in der Form (3.6) wurden in [10] als Anwendungsbeispiel für die Theorie der absorbierenden Randbedingungen aus Abschnitt 1.2 verwendet.

3.2.1. Die Diagonalisierung der Systemmatrix. Um die Theorie aus Abschnitt 1.2 anwenden zu können, müssen wir den Koeffizienten der x -Ableitung diagonalisieren. Dies geschieht mittels der orthogonalen Matrix T mit

$$(3.7) \quad T = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{2}} \end{pmatrix}.$$

Somit erhalten wir aus (3.6) eine neue Gleichung für $W := T^{-1}V$ der Form (1.9):

$$W_t + \underbrace{\begin{pmatrix} a+c & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a-c \end{pmatrix}}_{=A} W_x + \underbrace{\begin{pmatrix} b & \frac{c}{\sqrt{2}} & 0 \\ \frac{c}{\sqrt{2}} & b & \frac{-c}{\sqrt{2}} \\ 0 & \frac{-c}{\sqrt{2}} & b \end{pmatrix}}_{=B} W_y + \underbrace{\begin{pmatrix} 0 & \frac{-f}{\sqrt{2}} & 0 \\ \frac{f}{\sqrt{2}} & 0 & \frac{f}{\sqrt{2}} \\ 0 & \frac{-f}{\sqrt{2}} & 0 \end{pmatrix}}_{=C} W = 0$$

3.2.2. Die absorbierenden Randbedingungen. Für den Fall $a > 0$ (d. h. Einströmsituation an $x = 0$) ergeben sich die Matrizen $K_0 \in \mathbb{R}^{2 \times 1}$, $K_L \in \mathbb{R}^{1 \times 2}$ in den absorbierenden Randbedingungen (1.21) zu

$$(3.8) \quad K_0 = \begin{pmatrix} \frac{c-a}{2c} c_{13} \\ \frac{c-a}{c} c_{23} \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{(c-a)f}{c\sqrt{2}} \end{pmatrix}, \quad K_L = \begin{pmatrix} \frac{a+c}{2c} c_{31} & \frac{a}{c} c_{32} \end{pmatrix} = \begin{pmatrix} 0 & \frac{-af}{c\sqrt{2}} \end{pmatrix},$$

und die Matrizen $X_0 \in \mathbb{R}^{2 \times 1}$, $X_L \in \mathbb{R}^{1 \times 2}$ sind

$$(3.9) \quad X_0 = \begin{pmatrix} \frac{c-a}{2c} b_{13} \\ \frac{c-a}{c} b_{23} \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{(a-c)}{\sqrt{2}} \end{pmatrix}, \quad X_L = \begin{pmatrix} \frac{a+c}{2c} b_{31} & \frac{a}{c} b_{32} \end{pmatrix} = \begin{pmatrix} 0 & \frac{-a}{\sqrt{2}} \end{pmatrix}.$$

Analog ergibt sich bei $a < 0$ (d. h. Ausströmsituation an $x = 0$) für $K_0, X_0 \in \mathbb{R}^{1 \times 2}$ und $K_L, X_L \in \mathbb{R}^{2 \times 1}$:

$$(3.10) \quad K_0 = \begin{pmatrix} \frac{-a}{c} c_{12} & \frac{c-a}{2c} c_{13} \end{pmatrix} = \begin{pmatrix} \frac{af}{c\sqrt{2}} & 0 \end{pmatrix}, \quad K_L = \begin{pmatrix} \frac{a+c}{c} c_{21} \\ \frac{a+c}{2c} c_{31} \end{pmatrix} = \begin{pmatrix} \frac{(a+c)f}{c\sqrt{2}} \\ 0 \end{pmatrix},$$

$$(3.11) \quad X_0 = \begin{pmatrix} \frac{-a}{c} b_{12} & \frac{c-a}{2c} b_{13} \end{pmatrix} = \begin{pmatrix} \frac{-a}{\sqrt{2}} & 0 \end{pmatrix}, \quad X_L = \begin{pmatrix} \frac{a+c}{c} b_{21} \\ \frac{a+c}{2c} b_{31} \end{pmatrix} = \begin{pmatrix} \frac{(a+c)}{\sqrt{2}} \\ 0 \end{pmatrix}.$$

Eine Diskretisierung der ARB 1. Ordnung (1.21) an $x = 0$ ist dann gegeben durch

$$\frac{(U^+)_{0,\ell}^{n+1} - (U^+)_{0,\ell}^n}{k} + X_0 \frac{(U^-)_{0,\ell+1}^{n+\frac{1}{2}} - (U^-)_{0,\ell-1}^{n+\frac{1}{2}}}{2h} + K_0 (U^-)_{0,\ell}^{n+\frac{1}{2}} = 0.$$

Eine derartige Diskretisierung (speziell der Zeitableitung) erscheint für ein Verfahren mit zwei Zeitniveaus t_n, t_{n+1} sinnvoll. Damit ergibt sich die *diskrete Einstömrandbedingung*

$$(3.12) \quad (U^+)_{0,\ell}^{n+1} = (U^+)_{0,\ell}^n - \frac{\lambda}{2} X_0 \left((U^-)_{0,\ell+1}^{n+\frac{1}{2}} - (U^-)_{0,\ell-1}^{n+\frac{1}{2}} \right) - k K_0 (U^-)_{0,\ell}^{n+\frac{1}{2}},$$

wobei man Größen mit halbzahligen Zeitindizes durch arithmetische Mittelung definiert

$$(U^-)_{0,\ell}^{n+\frac{1}{2}} := \frac{1}{2} \left((U^-)_{0,\ell}^n + (U^-)_{0,\ell}^{n+1} \right).$$

Für die Ausströmdaten $(U^-)_{0,\ell}^{n+1}$ verwenden wir wie in 1D eine *horizontale Extrapolation*:

$$(3.13) \quad (U^-)_{0,\ell}^{n+1} = 2 (U^-)_{1,\ell}^{n+1} - (U^-)_{2,\ell}^{n+1}.$$

An $x = L$ verfährt man analog.

BEMERKUNG 3.3. Da die ARBen die Ausströmdaten auf dem neuen Zeitniveau benutzen, muß man zunächst extrapolieren, bevor man die ARBen verwendet (analog zu 1D). Zur Vorgehensweise am Rand bzgl. der y -Koordinate sei auf Bemerkung 3.5 verwiesen.

Für die folgenden numerischen Untersuchungen vereinfachen wir das Problem, indem wir $f = 0$ setzen.

3.3. Numerische Ergebnisse

Wir wollen die Stabilität des LW-Schemas numerisch untersuchen und die absorbierenden Randbedingungen 0. und 1. Ordnung hinsichtlich ihrer Absorptionsqualität mit der reflektierenden Randbedingung vergleichen.

3.3.1. Das Testbeispiel. Wir verwenden als Beispiel die *linearisierten Shallow-Water Gleichungen* aus Abschnitt 3.2 mit $a = 0.1$, $b = -0.2$, $c = 1.2$, $f = 0$ und $L = 1$. Es ist somit $m = 2$, $N = 3$. Als Anfangsbedingung haben wir $u = v = 0$ und für φ die Verteilung

$$(3.14) \quad \varphi(x, y, 0) := \begin{cases} \cos^2\left(\frac{\pi}{2} \frac{|(x-0.5, y)^T|}{0.45}\right) & |(x-0.5, y)^T| < 0.45, \\ 0 & \text{sonst,} \end{cases}$$

die in folgender Abbildung 3.1 auch mittels Niveaulinien dargestellt ist.

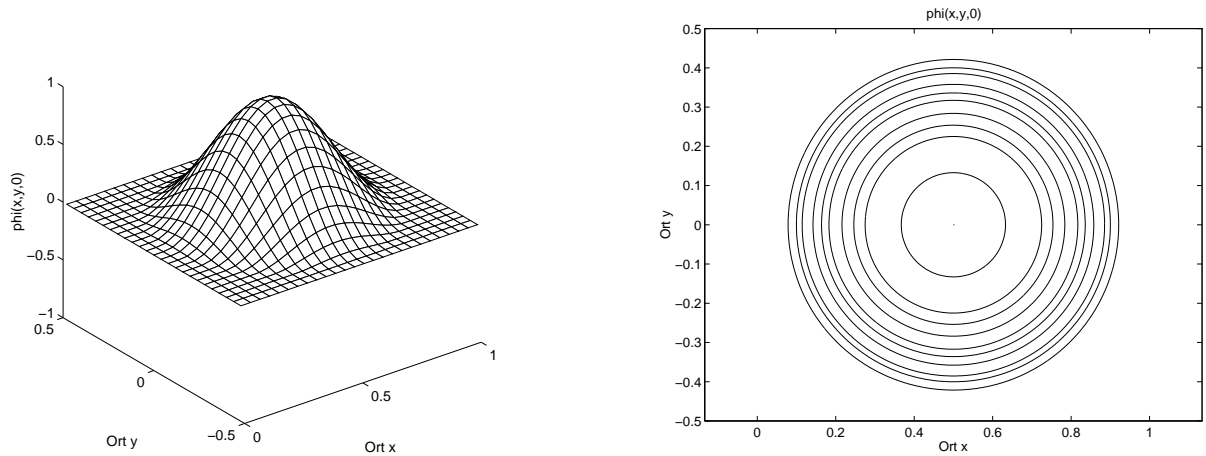


ABBILDUNG 3.1. Anfangsverteilung für φ

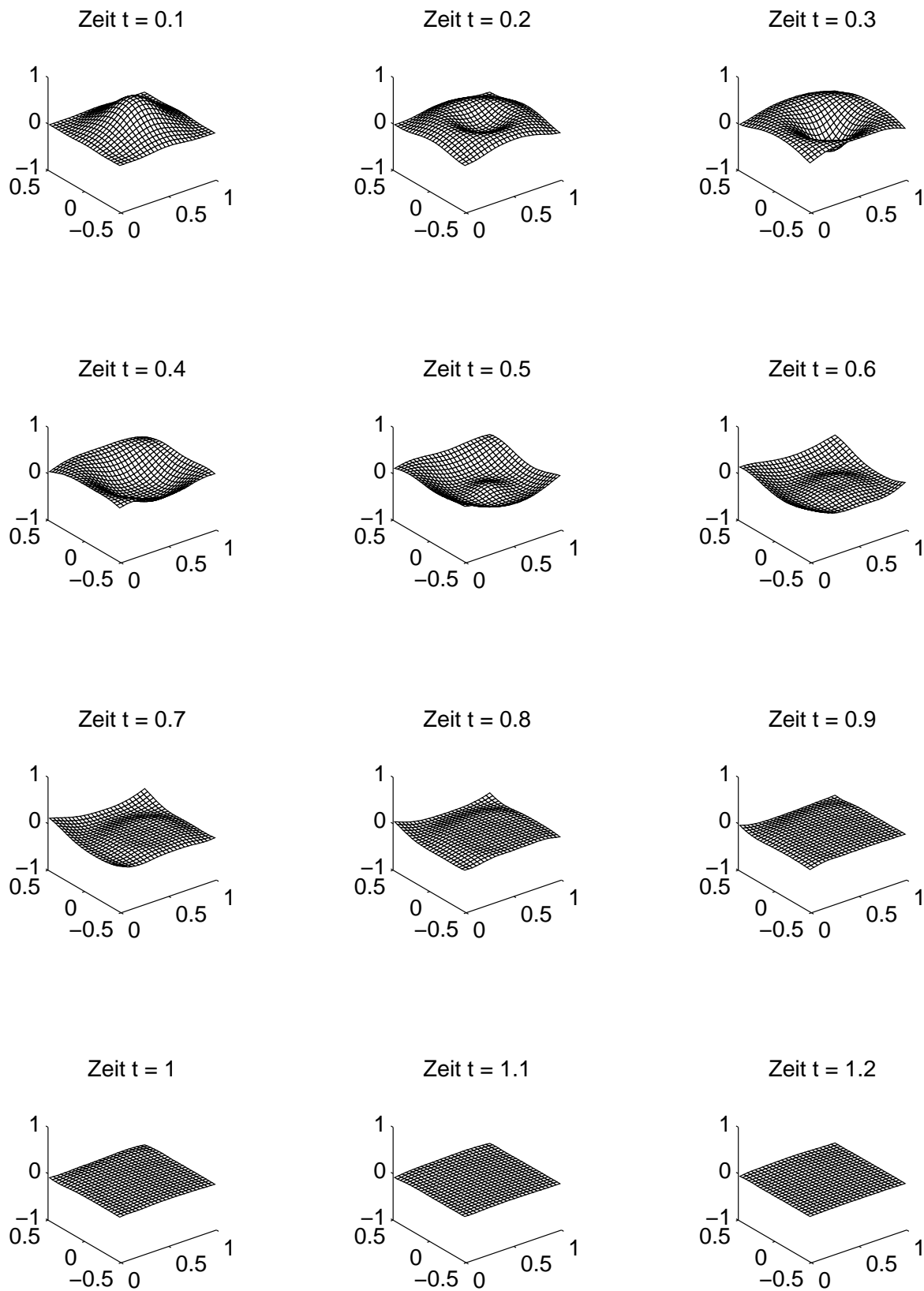
Für unser Beispiel erhalten wir $\lambda_{\max} = \max(|a|, |a \pm c|, |b|, |b \pm c|) = |b - c| = 1.4$ und somit ergibt sich als *notwendige Stabilitätsbedingung* (3.3) für diese Testgleichung

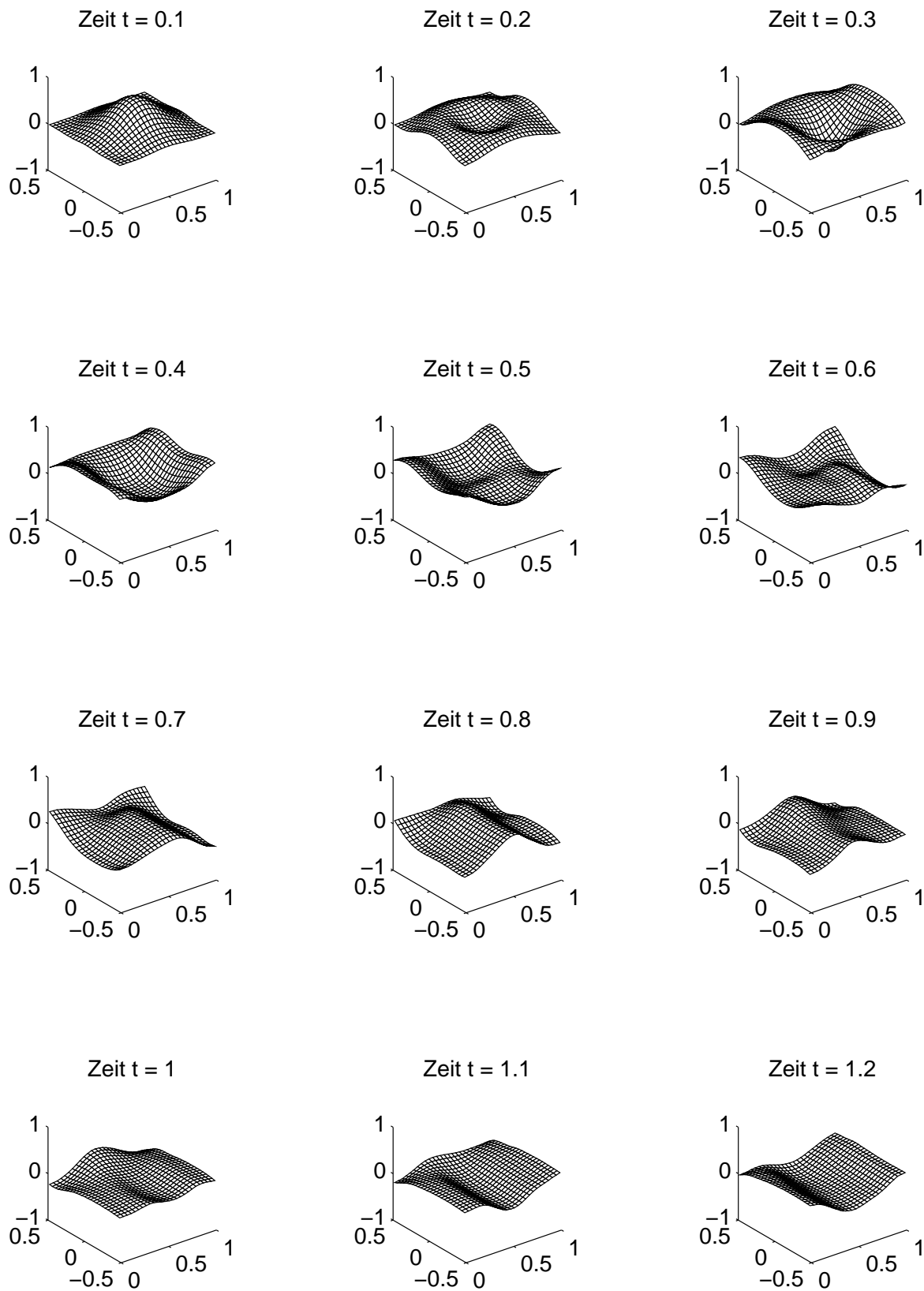
$$(3.15) \quad \lambda \leq \frac{5}{14\sqrt{2}} \approx 0.2525.$$

Wir werden für alle folgenden Rechnungen das Schrittweitenverhältnis $\lambda = 0.25$ nehmen und auf dem Gebiet $0 \leq x \leq 1$, $-3 \leq y \leq 3$ rechnen.

Zunächst wollen wir einen Eindruck vom Verlauf der Lösung gewinnen und verwenden eine Zeitschrittweite $\Delta t = 0.01$. Abbildung 3.2 und Abbildung 3.3 zeigen die zeitliche Entwicklung des Potentials φ bei ARBen 1. Ordnung bzw. *reflektierender Randbedingung*, die man erhält, indem man am Rand (bei $x = 0, x = 1$) die Reflektionsbedingung $u = v = 0$ benutzt und φ horizontal extrapoliert, um das numerische Schema zu schließen. Abbildung 3.4 und Abbildung 3.5 zeigen das Entsprechende mittels Niveaulinien.

BEMERKUNG 3.4. Die Abbildungen sind (im Gegensatz zu denen in Abschnitt 5.3.1) nicht symmetrisch um $x = 0.5$, $y = 0$; es gibt eine Drift in positive x -Richtung und negative y -Richtung. Dies erklärt sich aus den gewählten Werten von a, b , denn diese entsprechen den mittleren Horizontalgeschwindigkeiten \bar{u} , \bar{v} der Strömung in x - bzw. y -Richtung. Zur Illustration werden wir (nur für die Abbildung 3.6) die Werte $a = -0.2$ und $b = 0.1$ verwenden.

ABBILDUNG 3.2. φ für $t = 0.1, 0.2, \dots, 1.2$ bei absorb. Randbedingung 1. Ordnung

ABBILDUNG 3.3. φ für $t = 0.1, 0.2, \dots, 1.2$ bei reflektierender Randbedingung

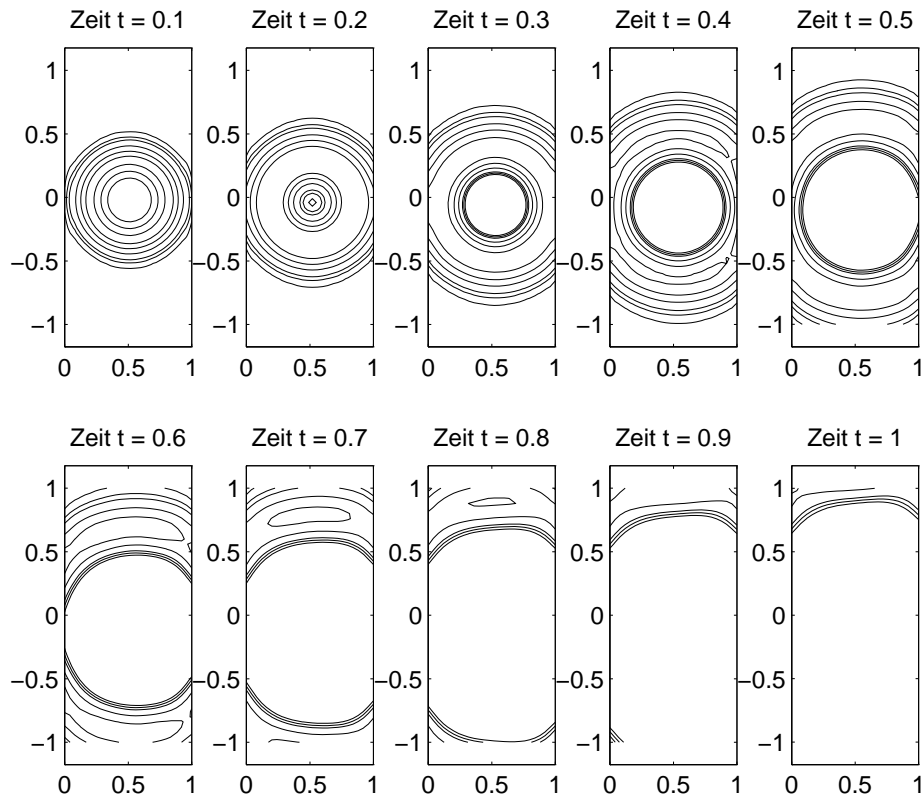


ABBILDUNG 3.4. φ für $t = 0.1, 0.2, \dots, 1$ bei absorb. Randbedingung 1. Ordnung

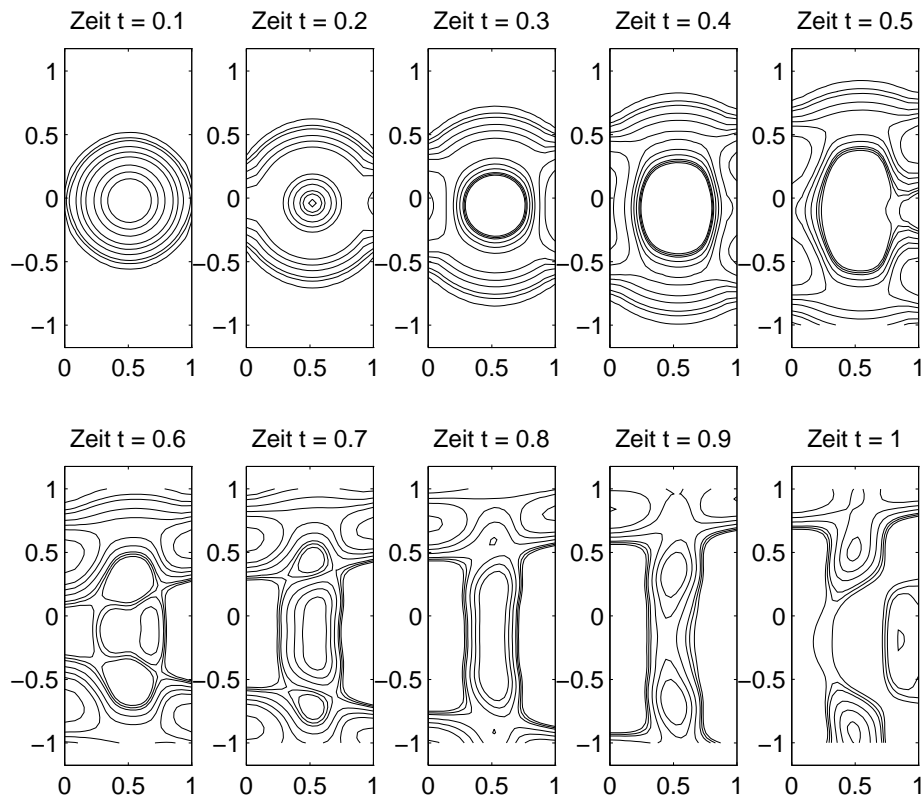


ABBILDUNG 3.5. φ für $t = 0.1, 0.2, \dots, 1$ bei reflektierender Randbedingung

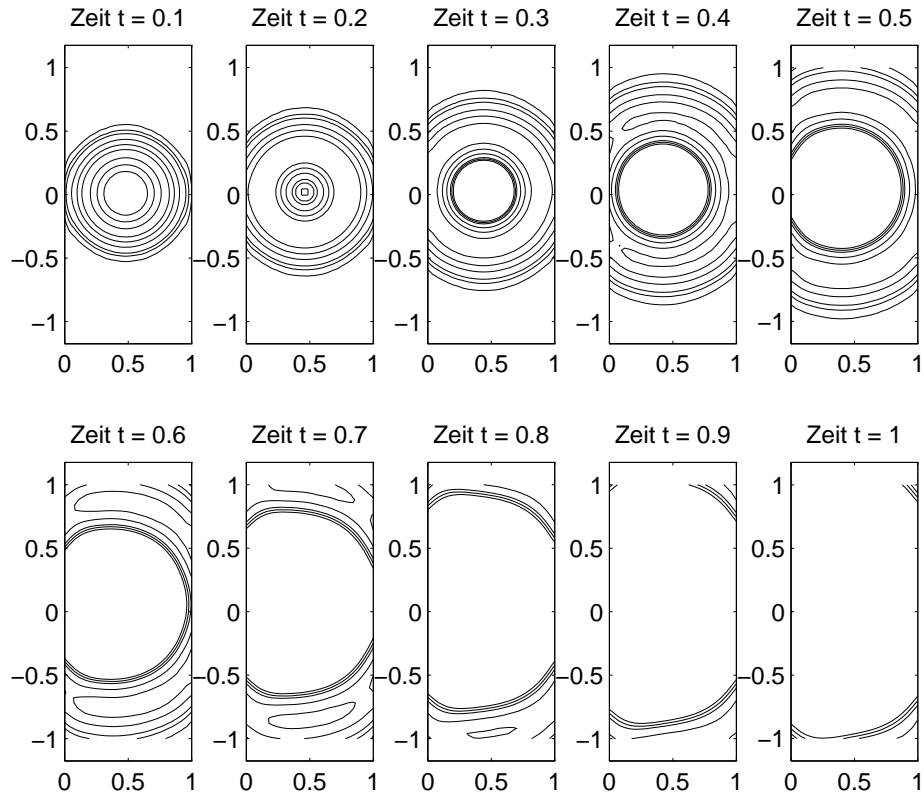


ABBILDUNG 3.6. φ für $t = 0.1, 0.2, \dots, 1$ bei absorb. Randbedingung 1. Ordnung, sowie $a = -0.2$, $b = 0.1$ zur Illustration der Drift

3.3.2. Numerische Untersuchung der Stabilität. Analog zu dem Vorgehen in Abschnitt 2.3.2 verwenden wir die *modulierte Anfangsverteilung*

$$\varphi(x, y, 0) := \begin{cases} \sin\left(2\pi p|(x - 0.5, y)^T|\right) \cos^2\left(\frac{\pi}{2} \frac{|(x - 0.5, y)^T|}{0.45}\right) & |(x - 0.5, y)^T| < 0.45, \\ 0 & \text{sonst,} \end{cases}$$

wobei der Parameter p ein Maß für die Frequenzen ist.

Das *innere Schema* ist nach Bemerkung 3.2 stabil, falls das Schrittweitenverhältnis λ die Bedingung (3.15) erfüllt. Wir wählen nun zunächst ein derartiges λ , da nur der Effekt der ARBen getestet werden soll. Für kleine Zeiten t verhält sich das Testproblem sehr ähnlich wie die Situation in 1D: die Wellen treffen anfänglich orthogonal auf den Rändern $x = 0$, $x = 1$ auf, d. h. Probleme – falls überhaupt – sind erst bei wesentlich „nicht-orthogonalen“ Einfallswinkeln zu erwarten. Die etwaigen Stabilitätsprobleme würden stark *frequenzabhängig* sein, und daher benutzen wir verschiedene Modulationen.

Wir geben hier den zeitlichen Verlauf der *diskreten L^2 -Norm*

$$(3.16) \quad \|U\|_{h,h}^2 := h^2 \sum_{\ell=-\infty}^{\infty} \sum_{j=0}^J |U_{j,\ell}|^2$$

jeder einzelnen Komponente (u, v, φ) an. Da wir uns nur für das *Abklingverhalten* bezüglich der Randbedingungen an $x = 0$ und $x = 1$ interessieren, wird die L^2 -Norm für den gesamten y -Bereich berechnet.

BEMERKUNG 3.5. Die formal unendliche Summe in (3.16) bzgl. der y -Koordinate wird in der Implementierung nur über ein hinreichend großen Bereich erstreckt, den man wegen der *endlichen Fortpflanzungsgeschwindigkeit* stets so groß wählen kann, daß die Lösung in der vorgegebenen Zeitspanne den Rand bzgl. der y -Koordinate nicht erreicht. An diesem Rand schließen wir das Schema mit einer *Extrapolation 1. Ordnung*, und zwar für *alle* Komponenten von U .

Die Resultate für $\lambda = 0.25$, $\Delta t = 0.001$ und dem Verfahrensparameter $s = 0.5$ sind in der Abbildung 3.7 dargestellt.

Um den Einfluß des Schrittweitenverhältnisses λ auf die Stabilität zu untersuchen, lassen wir bei den folgenden Rechnungen λ bei $p = 10$, $\Delta t = 0.001$ variieren.

FAZIT (STABILITÄT). Die L^2 -Normen klingen rasch ab, was zu erwarten ist, da sich die Wellen ausbreiten. Das *oszillatorische Verhalten* der L^2 -Normen entsteht beim Durchgang der hohen Frequenzen der Modulation durch den Rand bei $x = 0$ bzw. $x = 1$. Das Schema verhält sich bei diesen Beispielen mit Schrittweitenverhältnissen λ , die (3.15) erfüllen, *stabil*. Bei einem größeren λ wird in unserem Beispiel das Schema für $\lambda \geq 0.36$ *instabil*. Dies ist in den Abbildungen 3.9 und 3.10 zu sehen.

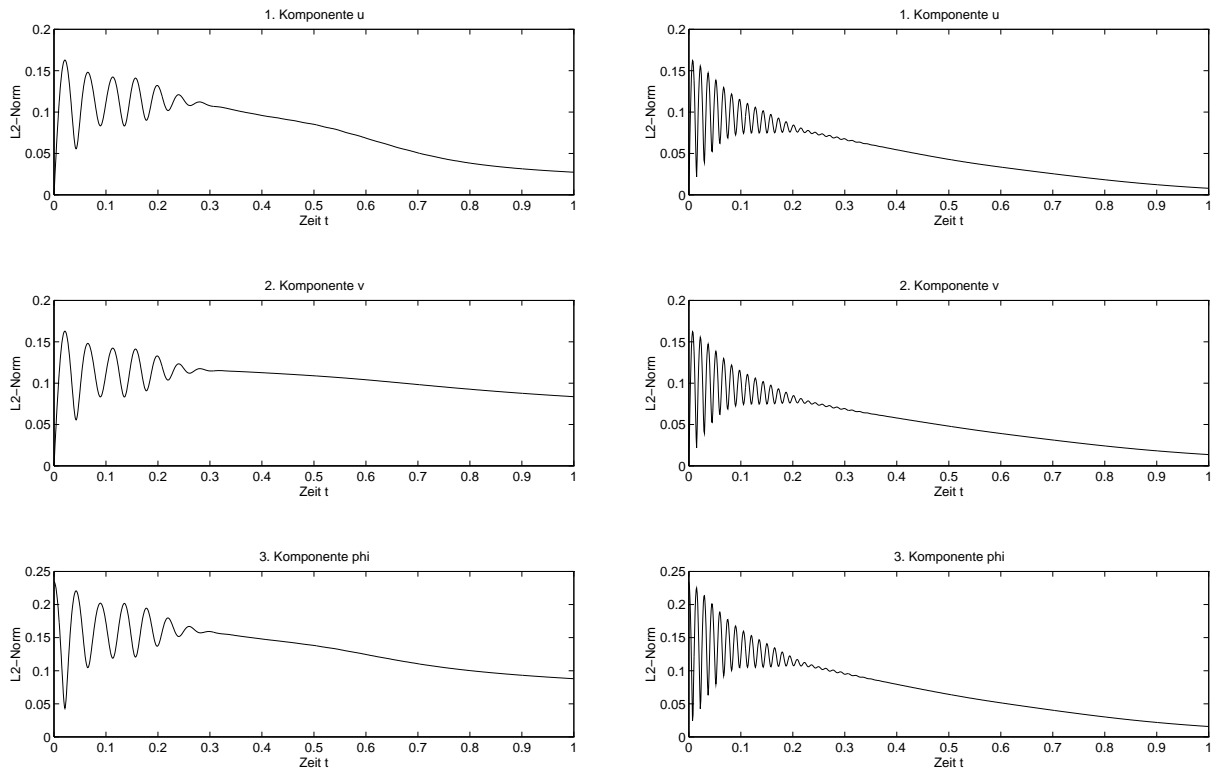


ABBILDUNG 3.7. L^2 -Normen jeder Komponente für $p = 10$ bzw. $p = 30$

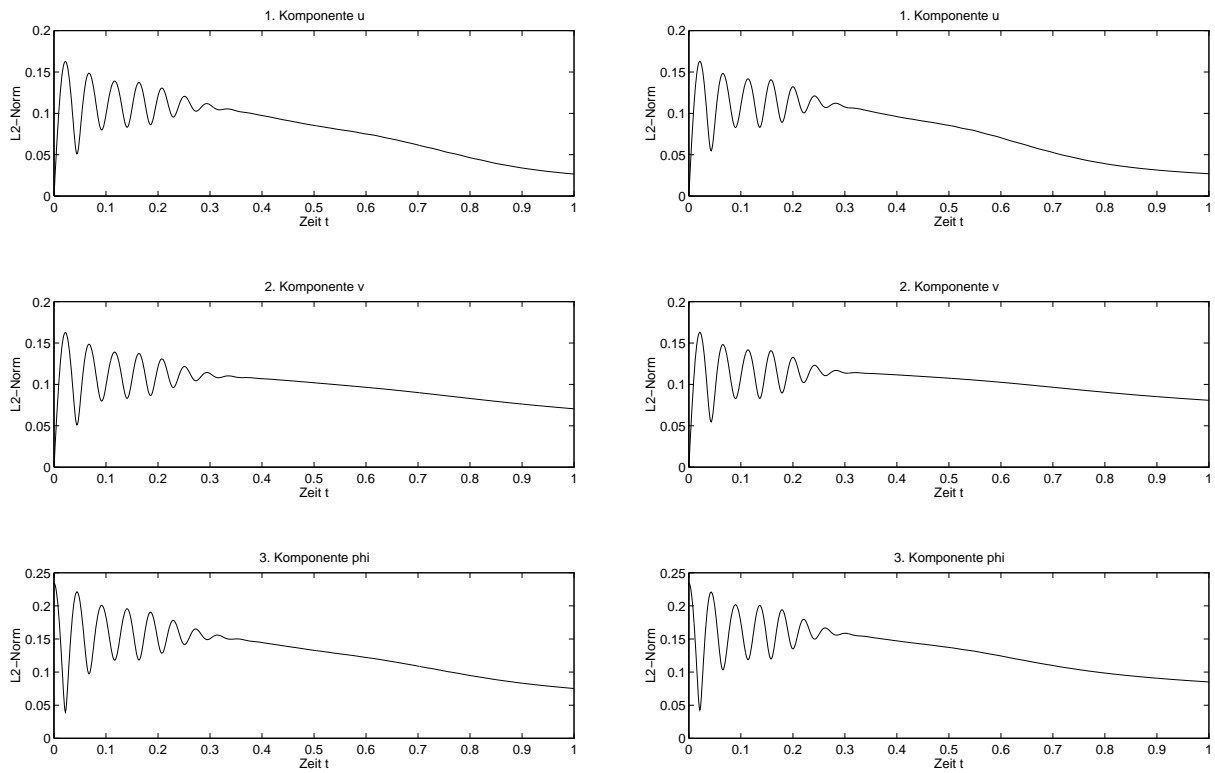


ABBILDUNG 3.8. L^2 -Normen jeder Komponente für $\lambda = 0.1$ bzw. $\lambda = 0.18$

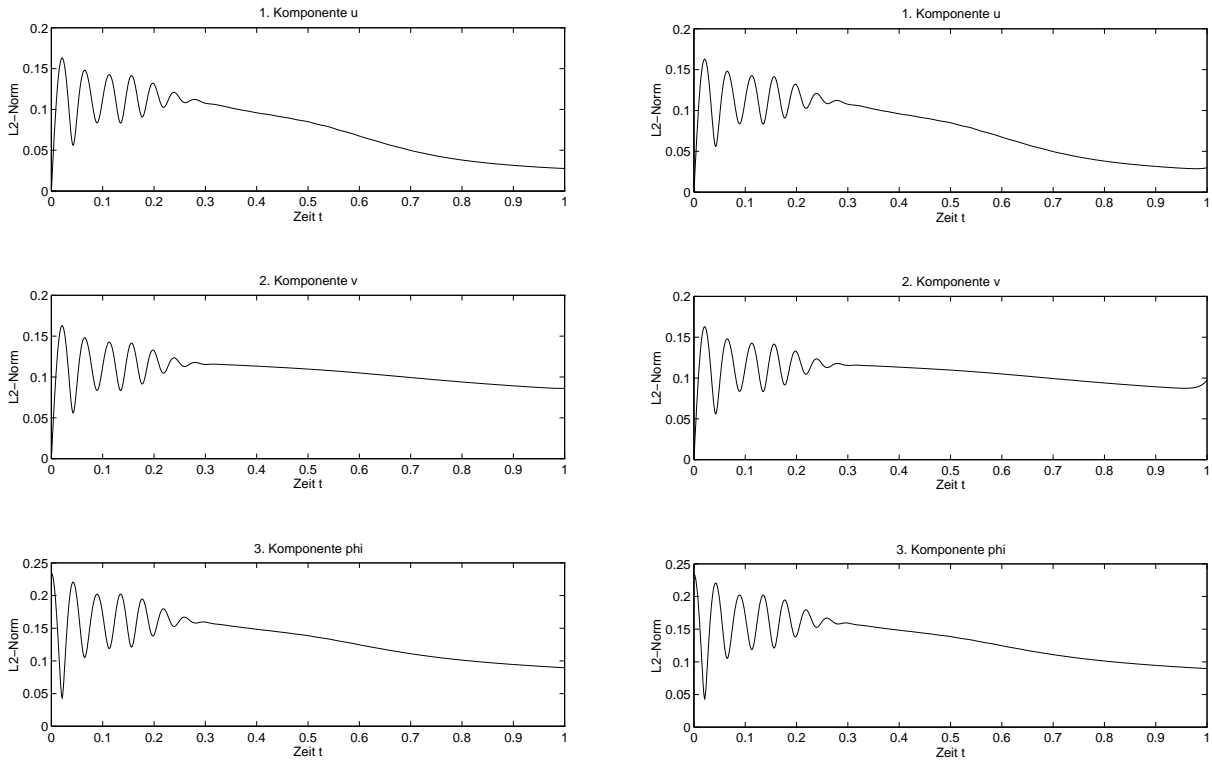


ABBILDUNG 3.9. L^2 -Normen jeder Komponente für $\lambda = 0.35$ bzw. $\lambda = 0.36$

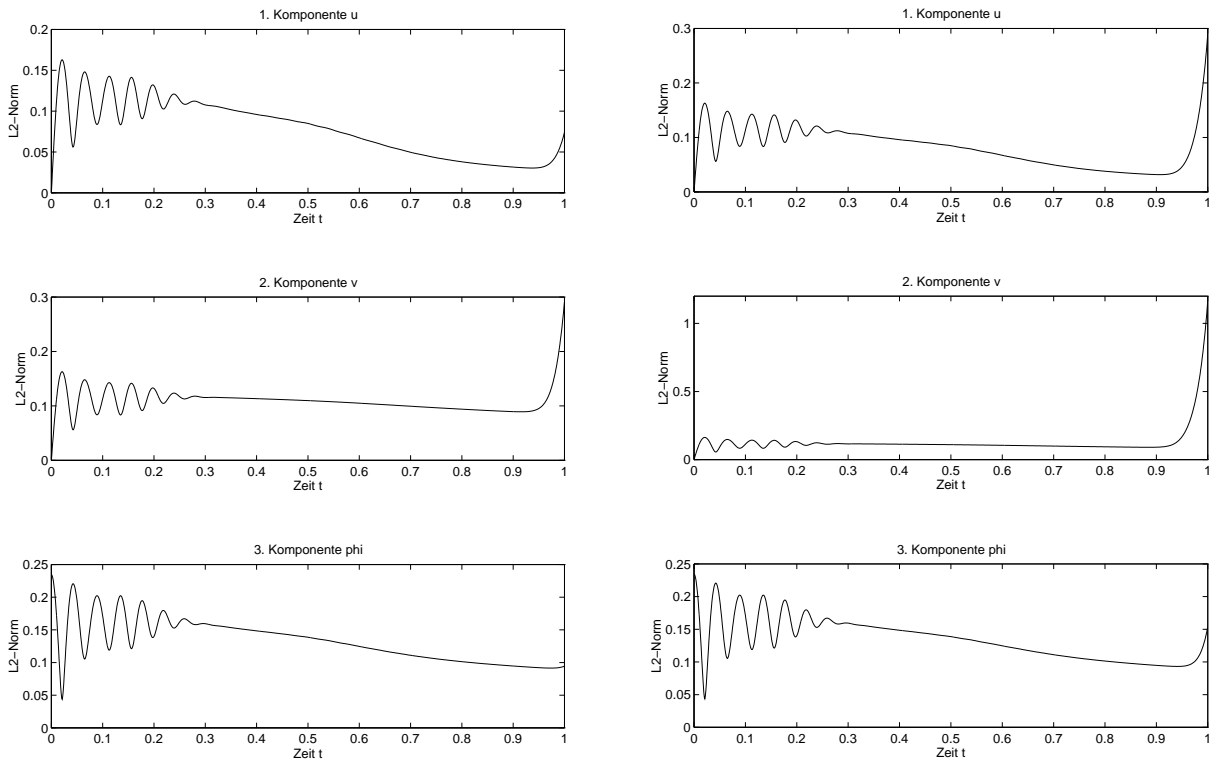


ABBILDUNG 3.10. L^2 -Normen jeder Komponente für $\lambda = 0.37$ bzw. $\lambda = 0.38$

3.3.3. Die Absorptionsqualität der absorbierenden Randbedingungen. Mit zunehmender Zeit trifft die Wellenfront flacher an Stellen mit wachsenden $|y|$ auf den Rand bei $x = 0$ und $x = 1$, und es ist zu erwarten, daß die numerischen Reflexionen der ARBen anwachsen, da bei der Herleitung der ARBen in Abschnitt 1.2 die Taylorentwicklung (1.17) um den senkrechten Einfallswinkel herum benutzt wurde.

Wir betrachten die numerischen Reflexionen bei den ARBen 0. und 1. Ordnung und relativieren sie an denen bei der reflektierenden Randbedingung. Der *reflektierte Anteil* wird wie in Abschnitt 2.3.3 berechnet, indem man von der Lösung des ARWP die Lösung des reinen AWP subtrahiert.

Die folgende Abbildung zeigt die zeitliche Entwicklung des reflektierten Anteils im Potential φ_{refl} , und zwar in der ersten Zeile bei reflektierenden Randbedingungen und in der zweiten Zeile bei absorbierenden Randbedingungen 1. Ordnung. Bei den reflektierenden Randbedingungen wurden die gleichen Niveauewerte wie auf Seite 44 verwendet und bei den ARBen wurden diese mit 0.5 multipliziert.

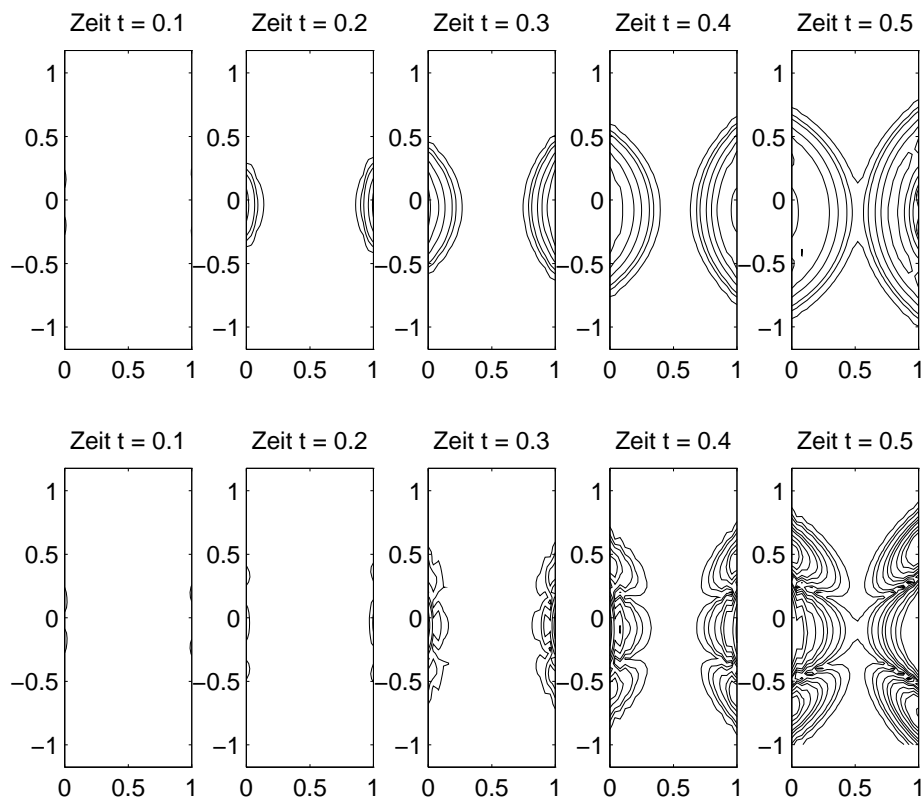


ABBILDUNG 3.11. Reflektierter Anteil von φ für $\lambda = 0.25$, $\Delta t = 0.01$.

In den folgenden Tabellen sieht man, daß bei φ_{refl} der Unterschied zwischen den ARBen nullter und erster Ordnung nicht sehr groß ist. Die Prozentzahlen beziehen sich dabei auf den jeweiligen Wert der L^2 -Norm des reflektierten Anteils bei Verwendung der Reflektionsrandbedingung.

Reflektierter Anteil bei $T = 0.3$

Randbedingung an $x = 0$ bzw. $x = 1$	$\ u_{refl}\ _{h,h}$	$\ v_{refl}\ _{h,h}$	$\ \varphi_{refl}\ _{h,h}$
totale Reflektion $u = v = 0$	0.07866	0.02605	0.07864
absorbierende Randbedingung 0. Ordnung (1.19)	0.00369 [4.7%]	0.01063 [40.8%]	0.00402 [5.1%]
absorbierende Randbedingung 1. Ordnung (1.21)	0.00318 [4.0%]	0.00202 [7.8%]	0.00355 [4.5%]

TABELLE 3.1. Ergebnisse für $\lambda = 0.25$ und $\Delta t = 0.01$.**Reflektierter Anteil bei $T = 0.3$**

Randbedingung an $x = 0$ bzw. $x = 1$	$\ u_{refl}\ _{h,h}$	$\ v_{refl}\ _{h,h}$	$\ \varphi_{refl}\ _{h,h}$
totale Reflektion $u = v = 0$	0.07142	0.02358	0.06944
absorbierende Randbedingung 0. Ordnung (1.19)	0.00322 [4.5%]	0.00457 [19.4%]	0.00294 [4.2%]
absorbierende Randbedingung 1. Ordnung (1.21)	0.00272 [3.8%]	0.00150 [6.4%]	0.00245 [3.5%]

TABELLE 3.2. Ergebnisse für $\lambda = 0.25$ und $\Delta t = 0.001$.

In der folgenden Abbildung 3.12 geben wir die L^2 -Norm des reflektierten Anteils wieder. Dabei ist $\lambda = 0.25$, $\Delta t = 0.001$. Die gestrichelte Linie ist der Verlauf der L^2 -Norm bei ARB 0. Ordnung und die durchgezogene bei ARB 1. Ordnung.

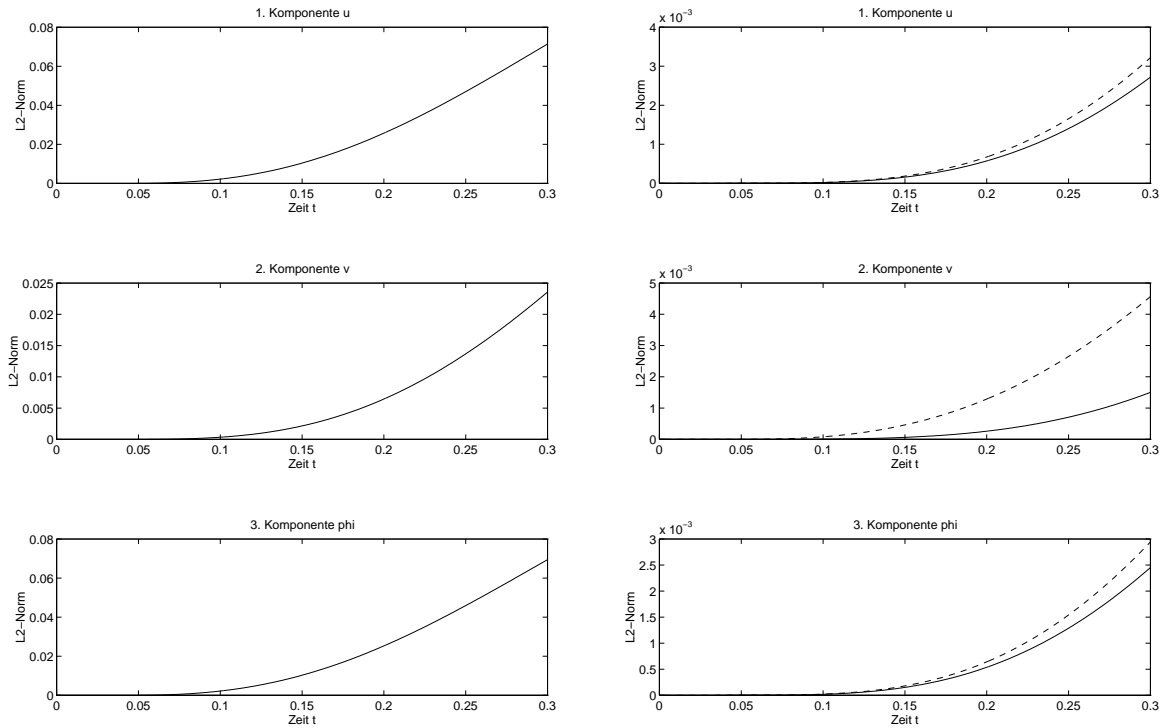


ABBILDUNG 3.12. Reflektierter Anteil bei totaler Reflexion bzw. ARBen: Parameter $\lambda = 0.25$, $\Delta t = 0.001$.

Reflektierter Anteil bei $T = 0.3$

Randbedingung an $x = 0$ bzw. $x = 1$	$\ u_{refl}\ _{h,h}$	$\ v_{refl}\ _{h,h}$	$\ \varphi_{refl}\ _{h,h}$
totale Reflexion $u = v = 0$	0.07922	0.02661	0.08045
absorbierende Randbedingung 0. Ordnung (1.19)	0.00407 [5.1%]	0.01161 [43.6%]	0.00442 [5.5%]
absorbierende Randbedingung 1. Ordnung (1.21)	0.00368 [4.6%]	0.00206 [7.7%]	0.00401 [5.0%]

TABELLE 3.3. Ergebnisse für $\lambda = 0.1$ und $\Delta t = 0.005$.

In Abbildung 3.13 geben wir die L^2 -Norm des reflektierten Anteils wieder. Dabei ist $\lambda = 0.1$, $\Delta t = 0.0005$.

Reflektierter Anteil bei $T = 0.3$

Randbedingung an $x = 0$ bzw. $x = 1$	$\ u_{refl}\ _{h,h}$	$\ v_{refl}\ _{h,h}$	$\ \varphi_{refl}\ _{h,h}$
totale Reflexion $u = v = 0$	0.07158	0.02371	0.06960
absorbierende Randbedingung 0. Ordnung (1.19)	0.00323 [4.5%]	0.00476 [20.1%]	0.00296 [4.3%]
absorbierende Randbedingung 1. Ordnung (1.21)	0.00274 [3.8%]	0.00152 [6.4%]	0.00247 [3.5%]

TABELLE 3.4. Ergebnisse für $\lambda = 0.1$ und $\Delta t = 0.0005$.

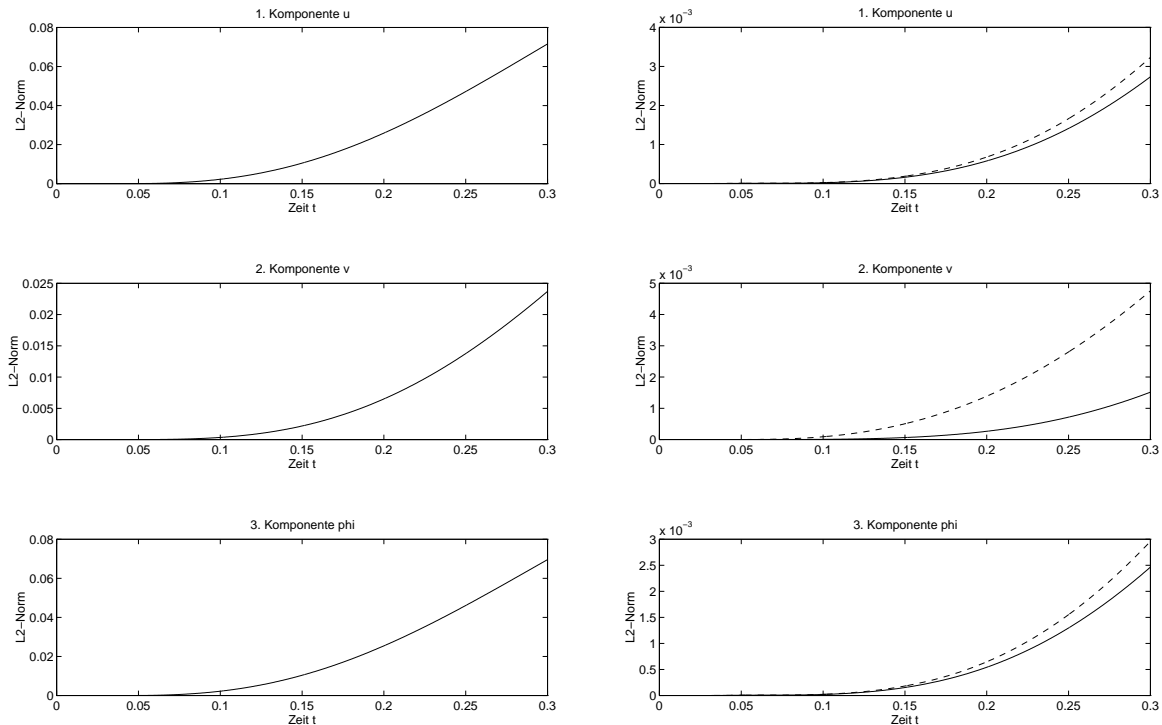


ABBILDUNG 3.13. Reflektierter Anteil bei totaler Reflexion bzw. ARBen:
Parameter $\lambda = 0.1$, $\Delta t = 0.0005$.

FAZIT (ABSORPTIONSQUALITÄT). Bei allen Testläufen ist erwartungsgemäß die L^2 -Norm des reflektierten Anteils bei der ARB 1. Ordnung geringer als bei der ARB nullter Ordnung. Der Fortschritt, den man macht, indem man von der ARB nullter Ordnung zu der ARB erster Ordnung übergeht, ist besonders deutlich in der 2. Komponente v und liegt dort beim Faktor 3–5.7 bzgl. der L^2 -Norm von v_{refl} . Bei $\lambda = 0.25$ und $\lambda = 0.1$ ergeben sich ähnliche Ergebnisse. Die Resultate bei Verwendung der ARB 1. Ordnung sind vergleichbar mit denen in [10].

KAPITEL 4

Numerische Betrachtung des nichtlinearen Falls in 1D

Wir betrachten quasilineare hyperbolische Systeme in 1D in der Erhaltungsform

$$(4.1) \quad U_t + F(U)_x = H(U), \quad 0 \leq x \leq L, \quad t \geq 0$$

und wollen dafür eine Klasse von Lax-Wendroff-Verfahren wie in [28] herleiten.

Aus Aufwandsgründen ist es vorteilhaft, für nichtlineare Systeme die LW-Methode als sog. *Zweischrittverfahren* zu formulieren. Die Diskretisierung des (x, t) -Raumes $[0, L] \times \mathbb{R}_0^+$ und die Bezeichnungsweise der Approximationen geschehen dabei wie in Kapitel 2. Weiterhin sind F_j^n , H_j^n abkürzende Schreibweisen für $F(U_j^n)$ bzw. $H(U_j^n)$.

4.1. Das numerische Schema für 1D

Eine Taylorentwicklung von $U(x, t + k)$ in der Zeit liefert

$$U(x, t + k) = U(x, t) + kU_t(x, t) + \frac{1}{2} k^2 U_{tt}(x, t) + O(k^3),$$

wobei U Lösung von (4.1) ist. Nun ersetzt man die auftretenden Zeitableitungen durch Ortsableitungen, indem man die Gleichung (4.1) benutzt. Dabei sollte man es vermeiden, die Jacobimatrix $A(U) := \partial F / \partial U$ zu verwenden, da sonst ein ineffizientes Verfahren wie etwa das ursprüngliche LW-Verfahren [31, Seite 302] daraus resultieren würde.

Wir verallgemeinern nun die *Richtmyer-Variante* des Zweischritt-LW-Verfahrens [31], indem wir einen Zwischenschritt zur Approximation von $U(x, t + sk)$, $s \in]0, 1]$ folgendermaßen einführen:

$$\begin{aligned} U(x, t + k) &= U(x, t) + k \left(1 - \frac{1}{2s}\right) U_t(x, t) + \frac{k}{2s} \left[\frac{\partial}{\partial t} (U(x, t) + skU_t(x, t)) \right] + O(k^3) \\ &= U(x, t) + k \left(1 - \frac{1}{2s}\right) U_t(x, t) + \frac{k}{2s} U_t(x, t + sk) + O(k^3) \\ &\stackrel{(4.1)}{=} U(x, t) - k \left(1 - \frac{1}{2s}\right) [F(U(x, t))_x - H(U(x, t))] \\ &\quad - \frac{k}{2s} [F(U(x, t + sk))_x - H(U(x, t + sk))] + O(k^3) \end{aligned}$$

mit $U(x, t + sk) = U(x, t) + skU_t(x, t) + O(k^2)$

$$\stackrel{(4.1)}{=} U(x, t) - sk [F(U(x, t))_x - H(U(x, t))] + O(k^2)$$

Eine *Diskretisierung* führt zu folgender Iteration:

$$\begin{aligned}
 (4.2a) \quad U_j^{n+s} &= U_j^n - sk \left(\frac{F_{j+\frac{1}{2}}^n - F_{j-\frac{1}{2}}^n}{h} - H_j^n \right) \\
 &= U_j^n - s\lambda(F_{j+\frac{1}{2}}^n - F_{j-\frac{1}{2}}^n - hH_j^n) \\
 (4.2b) \quad U_j^{n+1} &= U_j^n - k \left(1 - \frac{1}{2s} \right) \left(\frac{F_{j+1}^n - F_{j-1}^n}{2h} - H_j^n \right) \\
 &\quad - \frac{k}{2s} \left(\frac{F_{j+\frac{1}{2}}^{n+s} - F_{j-\frac{1}{2}}^{n+s}}{h} - H_j^{n+s} \right) \\
 &= U_j^n - \frac{\lambda}{2} \left[\left(1 - \frac{1}{2s} \right) (F_{j+1}^n - F_{j-1}^n - 2hH_j^n) \right. \\
 &\quad \left. + \frac{1}{s} (F_{j+\frac{1}{2}}^{n+s} - F_{j-\frac{1}{2}}^{n+s} - hH_j^{n+s}) \right].
 \end{aligned}$$

VORGEHENSWEISE (IMPLEMENTIERUNG). Aus Effizienzgründen wollen wir auf den Zeitstufen t_n, t_{n+1} nur Approximationen mit ganzzahligen Ortsindizes explizit berechnen; tritt ein halbzahlgiger Ortsindex auf, so wird zwischen den benachbarten Werten arithmetisch gemittelt. Auf der Zwischenstufe t_{n+s} werden dagegen nur Größen mit halbzahligen Ortsindizes explizit berechnet; ansonsten wird analog verfahren. Demnach führt eine Kombination von geradzahligen Zeitindizes mit halbzahligen Ortsindizes oder umgekehrt zu einer Mittelung, die wir mit „ $\hat{}$ “ bezeichnen werden.

Diese Vorgehensweise ergibt folgende *Klasse von Zweischritt-Verfahren*:

Erster Schritt:

$$\begin{aligned}
 (4.3a) \quad U_{j+\frac{1}{2}}^{n+s} &= \hat{U}_{j+\frac{1}{2}}^n - s\lambda(F_{j+1}^n - F_j^n - h\hat{H}_{j+\frac{1}{2}}^n), \\
 \text{wobei} \quad \hat{U}_{j+\frac{1}{2}}^n &:= \frac{1}{2}(U_j^n + U_{j+1}^n) \\
 \text{und} \quad \hat{H}_{j+\frac{1}{2}}^n &:= H(\hat{U}_{j+\frac{1}{2}}^n) \quad \text{ist.}
 \end{aligned}$$

Zweiter Schritt:

$$\begin{aligned}
 (4.3b) \quad U_j^{n+1} &= U_j^n - \frac{\lambda}{2} \left[\left(1 - \frac{1}{2s} \right) (F_{j+1}^n - F_{j-1}^n - 2hH_j^n) \right. \\
 &\quad \left. + \frac{1}{s} (F_{j+\frac{1}{2}}^{n+s} - F_{j-\frac{1}{2}}^{n+s} - h\hat{H}_j^{n+s}) \right], \\
 \text{wobei} \quad \hat{U}_j^{n+s} &:= \frac{1}{2}(U_{j-\frac{1}{2}}^{n+s} + U_{j+\frac{1}{2}}^{n+s}) \\
 \text{und} \quad \hat{H}_j^{n+s} &:= H(\hat{U}_j^{n+s}) \quad \text{ist.}
 \end{aligned}$$

BEMERKUNG 4.1 (SPEZIELLE VERFAHREN). Für eine Wahl des Verfahrensparameters $s = 1/2$ erhält man die *Richtmyer-Variante* und für $s = 1$ die *Variante von Rubin und Burstein* [33].

4.2. Das hydrodynamische Modell der Halbleitergleichungen

Die hydrodynamischen Gleichungen bestehen aus den 3 Erhaltungssätzen für Partikeldichte, Impuls und Energie. Die nichtlinearen Erhaltungssätze sind die *Eulergleichungen der Gasdynamik* ($\gamma = 5/3$) für ein Gas von geladenen Partikeln in einem elektrischen Feld mit zusätzlichen Relaxationszeittermen. Den Wärmeleitungsterm vernachlässigen wir, damit das System *hyperbolisch* ist.

Erhaltung der Anzahl der Partikel (Elektronen):

$$(4.4) \quad n_t + (nv)_x = C_n,$$

wobei n die Elektronenkonzentration, v die mittlere Geschwindigkeit der Elektronen und C_n ein Maß für die Rekombinations- und Regenerationsprozesse ist. Wir setzen $C_n = 0$, da es sich hier um ein *unpolares Modell* (d. h. nur Elektronen, keine Löcher) handelt.

Impulserhaltung:

$$(4.5) \quad (nv)_t + (nv^2 + P)_x = fn + \frac{C_p}{m},$$

$f = f(x, t)$ ist eine *äußere Kraft* aufgrund des elektrischen Feldes, m ist die effektive Masse eines Elektrons und C_p ein Maß für die Elektron-Elektron- und Elektron-Gitter-Kollisionen. Der Druck P sei gegeben durch $P = k_B n T / m$.

Erhaltung der Energie:

$$(4.6) \quad \left[n \left(\frac{1}{2} v^2 + e_I \right) \right]_t + \left[nv \left(\frac{1}{2} v^2 + e_I \right) + Pv \right]_x = fnv + \frac{C_w}{m}.$$

Dabei ist e_I die innere Energie und C_w ein Kollisionsterm.

Wir nehmen an, daß die Energiebänder parabolisch sind und mit der *Energiedichte* $w = mne_I + mnv^2/2 = nk_B T / (\gamma - 1) + mnv^2/2$ und der *Impulsdichte* $p = mnv$ lassen sich die obigen Gleichungen schreiben als

$$\begin{aligned} \rho_t + p_x &= 0 \\ p_t + (pv + k_B n T)_x &= f\rho + C_p \\ w_t + [v(w + k_B n T)]_x &= fp + C_w, \end{aligned}$$

wobei wir die Gleichungen mit m multipliziert und die *Dichte* $\rho = mn$ eingeführt haben. Unter Verwendung von $k_B n T = (\gamma - 1)(w - pv/2)$ ergeben die beiden letzten Gleichungen

$$\begin{aligned} p_t + \left[(\gamma - 1)w - \frac{1}{2}(\gamma - 3)pv \right]_x &= f\rho + C_p \\ w_t + \left[v \left(\gamma w - \frac{1}{2}(\gamma - 1)pv \right) \right]_x &= fp + C_w. \end{aligned}$$

Für unsere Anwendung sind die *Kollisionsterme* C_p, C_w gegeben durch [11]:

$$(4.7) \quad C_p = -\frac{p}{\tau_p} \quad \text{und} \quad C_w = -\frac{w - \frac{1}{\gamma-1}nk_B T_0}{\tau_w} = -\frac{w}{\tau_w} + \frac{1}{\gamma-1} \frac{\rho k_B T_0}{m\tau_w}$$

mit *Relaxationszeiten* τ_p, τ_w , für die die folgenden empirischen Relationen [13] gelten:

$$(4.8) \quad \tau_p = \tau_p(T) = \frac{m\mu_0 T_0}{eT}, \quad \tau_w = \tau_w(T) = \frac{\tau_p}{2} + \frac{3\mu_0 k_B T T_0}{2ev_s^2(T + T_0)}$$

Physikalische Parameter für Silizium bei Gittertemperatur $T_0 = 300$ K [13]:

effektive Masse eines Elektrons	$m = 0.26 \cdot m_e$
Sättigungsgeschwindigkeit	$v_s = 1.035 \cdot 10^5 \text{ m s}^{-1}$
Beweglichkeit der Elektronen	$\mu_0 = 0.145 \text{ m}^2 \text{ s}^{-1} \text{ V}^{-1}$

Die auftretenden Parameter werden für unser Beispiel wie oben gewählt [12]. Zur Vereinfachung verwenden wir konstante Relaxationszeiten, und zwar $\tau_p = \tau_p(T_0)$, $\tau_w = \tau_w(T_0)$. Mit $v = p/\rho$ kann man diese Gleichungen nun in der Form (4.1) formulieren, wobei

$$U = \begin{pmatrix} \rho \\ p \\ w \end{pmatrix}, \quad F(U) = \begin{pmatrix} p \\ (\gamma - 1)w - \frac{1}{2}(\gamma - 3)\frac{p^2}{\rho} \\ \gamma \frac{pw}{\rho} - \frac{1}{2}(\gamma - 1)\frac{p^3}{\rho^2} \end{pmatrix}, \quad H(U) = \begin{pmatrix} 0 \\ f\rho - \frac{p}{\tau_p} \\ fp - \frac{w}{\tau_w} + \frac{1}{\gamma - 1} \frac{\rho k_B T_0}{m\tau_w} \end{pmatrix}$$

ist. Der Anfangszustand wird festgelegt durch [11]:

$$n(x, 0) = n_D(x), \quad v(x, 0) = 0, \quad T(x, 0) = T_0,$$

wobei n_D das *Dotierungsprofil* ist. Somit ergibt sich als *Anfangsbedingung*:

$$(4.9) \quad \rho(x, 0) = mn_D(x), \quad p(x, 0) = 0, \quad w(x, 0) = \frac{1}{\gamma - 1} k_B T_0 n_D(x).$$

4.2.1. Die Diagonalisierung der Systemmatrix. Zur Formulierung der Randbedingungen betrachten wir für $U = (\rho, p, w)^T$ anstatt (4.1) die Gleichungen in der Form

$$(4.10) \quad U_t + AU_x + CU = 0, \quad 0 \leq x \leq L, \quad t \geq 0$$

mit den Koeffizienten

$$(4.11) \quad A = \frac{\partial F}{\partial U} = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{2}(\gamma - 3)v^2 & -(\gamma - 3)v & \gamma - 1 \\ -\gamma \frac{wv}{\rho} + (\gamma - 1)v^3 & \gamma \frac{w}{\rho} - \frac{3}{2}(\gamma - 1)v^2 & \gamma v \end{pmatrix}, \quad v = \frac{p}{\rho}$$

und

$$(4.12) \quad C = \begin{pmatrix} 0 & 0 & 0 \\ -f & \tau_p^{-1} & 0 \\ \frac{-k_B T_0}{(\gamma - 1)m\tau_w} & -f & \tau_w^{-1} \end{pmatrix}.$$

Für die ARBen an $x = 0$ ($j = 0$) frieren wir die Koeffizienten des hyperbolischen Systems (4.10) ein und transformieren es in Randnähe ($j = 0, 1, 2$) in charakteristische Variablen.

Mit der *Schallgeschwindigkeit* c , $c^2 = \gamma k_B T/m$ ergibt sich aus der Gleichung für die Energiedichte w :

$$\gamma \frac{w}{\rho} = \frac{c^2}{\gamma - 1} + \frac{1}{2} \gamma v^2.$$

Die Systemmatrix $A = A(\rho, p, w)$ läßt sich damit schreiben als

$$A = A(v, c) = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{2}(\gamma - 3)v^2 & -(\gamma - 3)v & \gamma - 1 \\ -v \left[\frac{c^2}{\gamma - 1} + \left(1 - \frac{1}{2}\gamma\right)v^2 \right] & \frac{c^2}{\gamma - 1} - \left(\gamma - \frac{3}{2}\right)v^2 & \gamma v \end{pmatrix}.$$

A besitzt die Eigenwerte

$$\xi_2 = v, \quad \xi_{1,3} = v \pm c$$

mit den zugehörigen Eigenvektoren

$$\zeta_2 = \begin{pmatrix} 1 \\ v \\ \frac{v^2}{2} \end{pmatrix}, \quad \zeta_{1,3} = \begin{pmatrix} 1 \\ v \pm c \\ \frac{v^2}{2} \pm vc + \frac{c^2}{\gamma - 1} \end{pmatrix}.$$

Wir frieren die Matrix A lokal an $x = 0$, t fest, ein, d. h. wir betrachten $A_0 := A(v_0, c_0)$ mit $v_0 = v(x=0)$, $c_0 = c(x=0)$, sowie $C_0 := C(x=0)$ und diagonalisieren A_0 mittels:

$$T_0 := T(v_0, c_0) = \begin{pmatrix} 1 & 1 & 1 \\ v_0 + c_0 & v_0 & v_0 - c_0 \\ \frac{v_0^2}{2} + v_0 c_0 + \frac{1}{b_0} & \frac{v_0^2}{2} & \frac{v_0^2}{2} - v_0 c_0 + \frac{1}{b_0} \end{pmatrix},$$

$$T_0^{-1} := T^{-1}(v_0, c_0) = \begin{pmatrix} \frac{1}{2} \left(a_0 - \frac{v_0}{c_0} \right) & \frac{1}{2} \left(\frac{1}{c_0} - v_0 b_0 \right) & \frac{b_0}{2} \\ 1 - a_0 & v_0 b_0 & -b_0 \\ \frac{1}{2} \left(a_0 + \frac{v_0}{c_0} \right) & -\frac{1}{2} \left(\frac{1}{c_0} + v_0 b_0 \right) & \frac{b_0}{2} \end{pmatrix},$$

wobei $a_0 := (\gamma - 1)v_0^2/(2c_0^2)$, $b_0 := (\gamma - 1)/c_0^2$ ist. Somit erhalten wir aus (4.10) eine neue Gleichung für die charakteristische Variable $V := T_0^{-1}U$ der Form:

$$V_t + \Lambda_0 V_x + \tilde{C}_0 V = 0$$

$$\text{mit } \Lambda_0 = \text{diag}(v_0 + c_0, v_0, v_0 - c_0), \quad \tilde{C}_0 = T_0^{-1}C_0 T_0.$$

An $x = L$ verfährt man analog.

4.2.2. Die Randbedingungen. Für die Gleichung in der Form (4.2.1) kann man die ARBen aus Abschnitt 1.2 formulieren. In einer *Überschallsituation*, d. h. $|v_0| > c_0$, hat man an $x = 0$ für $v_0 > 0$ nur Einströmkomponenten und für $v_0 < 0$ nur Ausströmkomponenten. Bei *Unterschall*, d. h. $|v_0| < c_0$, hängt die Partition vom Vorzeichen von v_0 ab: bei $v_0 > 0$ hat man zwei Einströmkomponenten und eine Ausströmkomponente und bei $v_0 < 0$ ist

es umgekehrt. Komponenten zum Eigenwert Null werden wie eine Ausströmkomponente horizontal extrapoliert.

Wir werden die ARB 1. Ordnung mit der (inhomogenen) Dirichlet–Randbedingung

$$(4.13) \quad (V^+)_0^{n+1} = (V^+)_0^n$$

hinsichtlich der numerischen Reflektionen vergleichen. Bei der Implementierung ist die Bedingung (4.13) äquivalent zur absorbierenden Randbedingung 0. Ordnung

$$(V^+)_0^{n+1} = (V^+)_0^n.$$

Somit hat man die folgende

VORGEHENSWEISE (NUMERISCHE RANDBEDINGUNGEN). An den inneren Punkten berechnet man U_j^{n+1} , $j = 1, 2, \dots, J - 1$ mit Hilfe des Zweischritt–LW–Verfahrens (4.3). An $j = 0$ implementiert man die Randbedingungen wie folgt:

$$\text{Gegeben} \quad : \quad U_0^n = (\rho_0^n, p_0^n, w_0^n)^T, \quad U_1^{n+1}, U_2^{n+1}.$$

$$\text{Berechnung} \quad : \quad v_0^n = \frac{p_0^n}{\rho_0^n}, \quad c_0^n = \sqrt{\gamma(\gamma - 1) \left(\frac{w_0^n}{\rho_0^n} - \frac{(v_0^n)^2}{2} \right)}.$$

$$\text{Transformation} \quad : \quad V_0^n = T_0^{-1} U_0^n, \quad V_1^{n+1} = T_0^{-1} U_1^{n+1}, \quad V_2^{n+1} = T_0^{-1} U_2^{n+1},$$

wobei $T_0^{-1} = T^{-1}(v_0^n, c_0^n)$.

$$\text{Ausströmextrapolation} \quad : \quad (V^-)_0^{n+1} = 2(V^-)_1^{n+1} - (V^-)_2^{n+1}.$$

$$\text{i) Dirichlet–Randbedingung} \quad : \quad (V^+)_0^{n+1} = (V^+)_0^n$$

bzw.

$$\text{ii) Absorb. Randbedingung} \quad : \quad (V^+)_0^{n+1} = (V^+)_0^n - \frac{k}{2} K_0 \left((V^-)_0^n + (V^-)_0^{n+1} \right)$$

$$\text{bzgl.} \quad \Lambda_0 = \text{diag}(v_0^n + c_0^n, v_0^n, v_0^n - c_0^n), \quad \tilde{C}_0(v_0^n, c_0^n) = T_0^{-1} C_0 T_0.$$

$$\text{Rücktransformation} \quad : \quad U_0^{n+1} = T_0 V_0^{n+1}, \quad \text{wobei} \quad T_0 = T(v_0^n, c_0^n).$$

Für $j = J$ ergibt sich eine analoge Vorgehensweise.

4.3. Numerische Ergebnisse

Wir wollen hier die numerischen Ergebnisse des Zweischritt-LW-Schemas (4.3) für ein Anwendungsbeispiel betrachten.

4.3.1. Das Testbeispiel. Für unser Beispiel benutzen wir das Einheitensystem [11]:

Länge	$10^{-6} \text{ m} = 1 \mu\text{m}$
Zeit	$10^{-12} \text{ s} = 1 \text{ ps}$
Masse	10^{-30} kg
Spannung	1 V
Temperatur	1 K
Energie	10^{-18} J
Ladung	10^{-18} C

Die Anfangsbedingung (4.9) für $0 \leq x \leq 1$ ist durch eine Gittertemperatur $T_0 = 300$ und das folgende Dotierungsprofil [11] gegeben:

$$(4.14) \quad n_D(x) := \begin{cases} 5 \cdot 10^5 & \text{falls } 0 \leq x \leq 0.25 \text{ oder } 0.75 \leq x \leq 1, \\ 2 \cdot 10^3 & \text{falls } 0.35 \leq x \leq 0.65. \end{cases}$$

Dabei verbinden wir die beiden Sprungstellen durch eine geeignet skalierte Form von

$$\begin{cases} \cos^2\left(\frac{\pi}{2} \frac{x-0.25}{0.1}\right) & \text{falls } 0.25 \leq x \leq 0.35, \\ \cos^2\left(\frac{\pi}{2} \frac{x-0.75}{0.1}\right) & \text{falls } 0.65 \leq x \leq 0.75. \end{cases}$$

In der Abbildung 4.1 ist das resultierende Dotierungsprofil dargestellt.

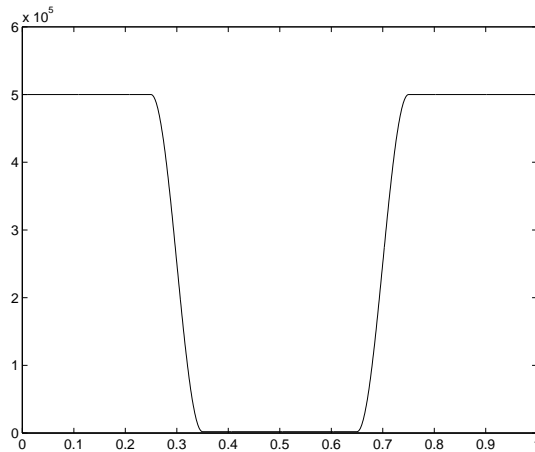


ABBILDUNG 4.1. Dotierungsprofil n_D

Hier ist $f = f(x, t)$ fest vorgegeben. Der Einfachheit halber werden wir für die numerischen Berechnungen $f = 0$ setzen. Die stationäre Verteilung von n , die sich bei $f = 0$ ergibt, wurde in [11] als Anfangsverteilung für folgende Rechnungen benutzt, in denen f selbstkonsistent mittels $f_x = -e^2/(m\varepsilon)(n - n_D)$ ermittelt wurde. Dabei ist $\varepsilon = 11.7 \cdot \varepsilon_0$ die Dielektrizitätskonstante.

Die *notwendige Stabilitätsbedingung* für das Testbeispiel ist die CFL-Bedingung (2.13)

$$(4.15) \quad \lambda \leq \frac{1}{\lambda_{\max}}$$

mit $\lambda_{\max} = \max_{(x,t) \in [0,1] \times [0,T]} (c(x,t) + |v(x,t)|)$.

Wir wollen zunächst die zeitliche Entwicklung der Geschwindigkeit und der Temperatur bei Verwendung der Dirichlet-Randbedingung bzw. der absorbierenden Randbedingungen 1. Ordnung betrachten. Dabei ist wie in [11] $t \in [0, 3]$, $\lambda = 0.2$, $\Delta t = 0.002$, sowie $s = 0.5$ gewählt. Die Resultate sind in den Abbildungen 4.3 bis 4.6 zu sehen.

In Abbildung 4.2 vergleichen wir die Geschwindigkeit v an $t = 3$ bei der Dirichlet-Randbedingung und der ARB 1. Ordnung. Die gestrichelte Kurve in Abbildung 4.2 ist vergleichbar mit dem Resultat in [11], bei dem allerdings die Lösung starke Oszillationen besitzt.

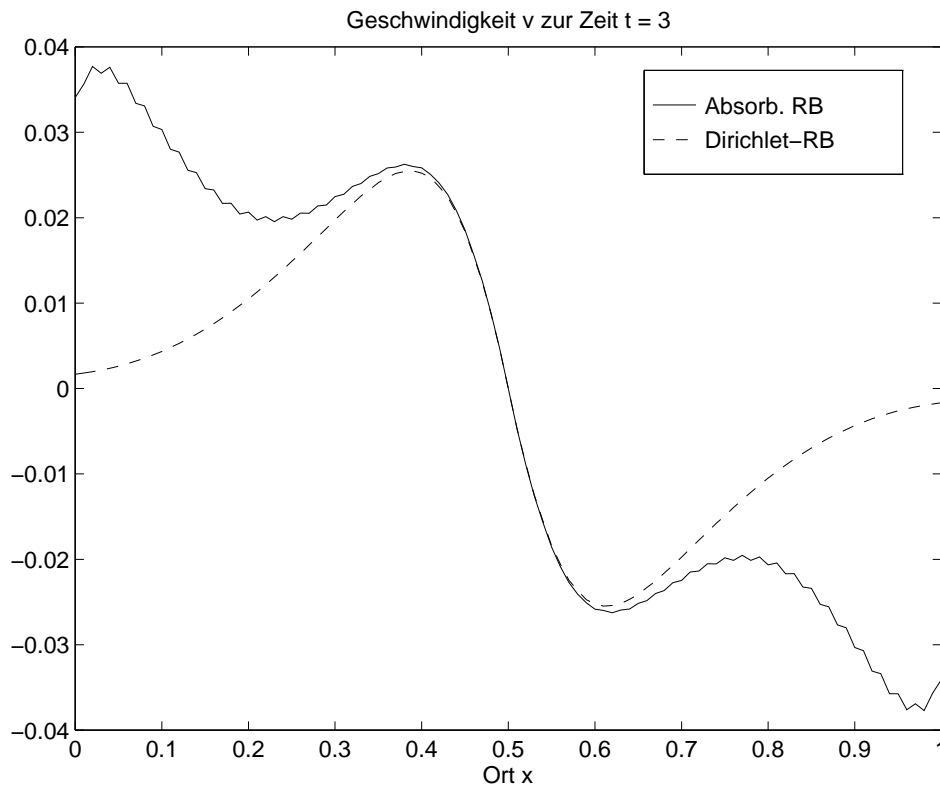
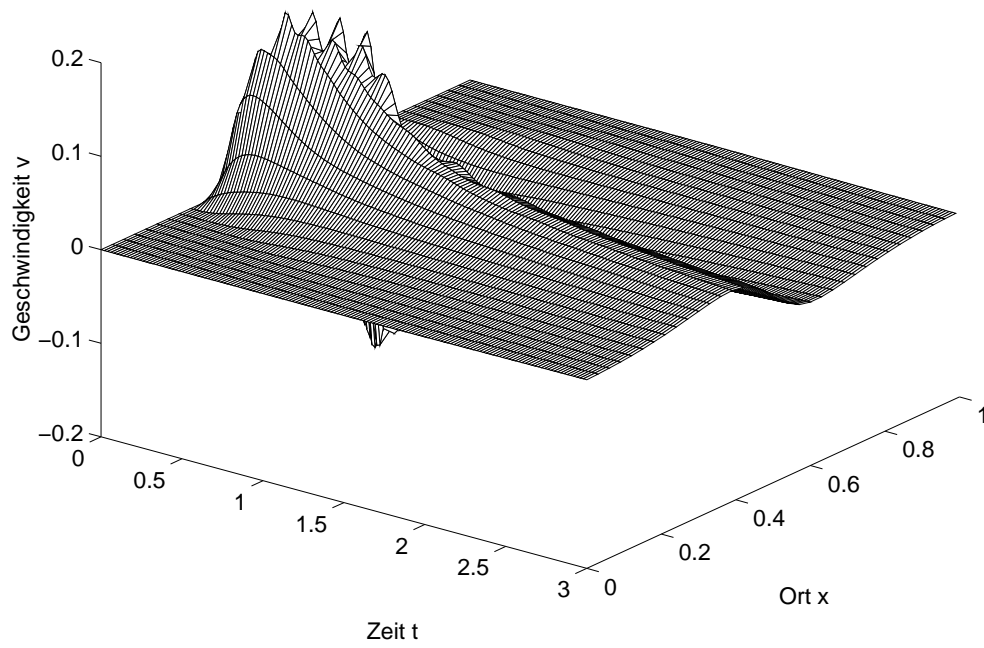
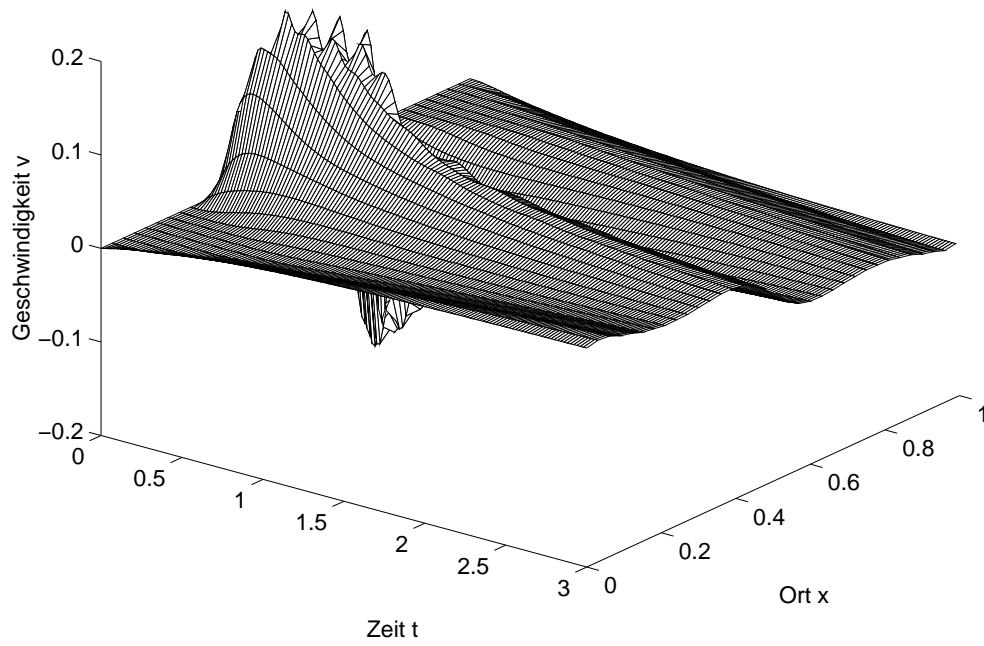
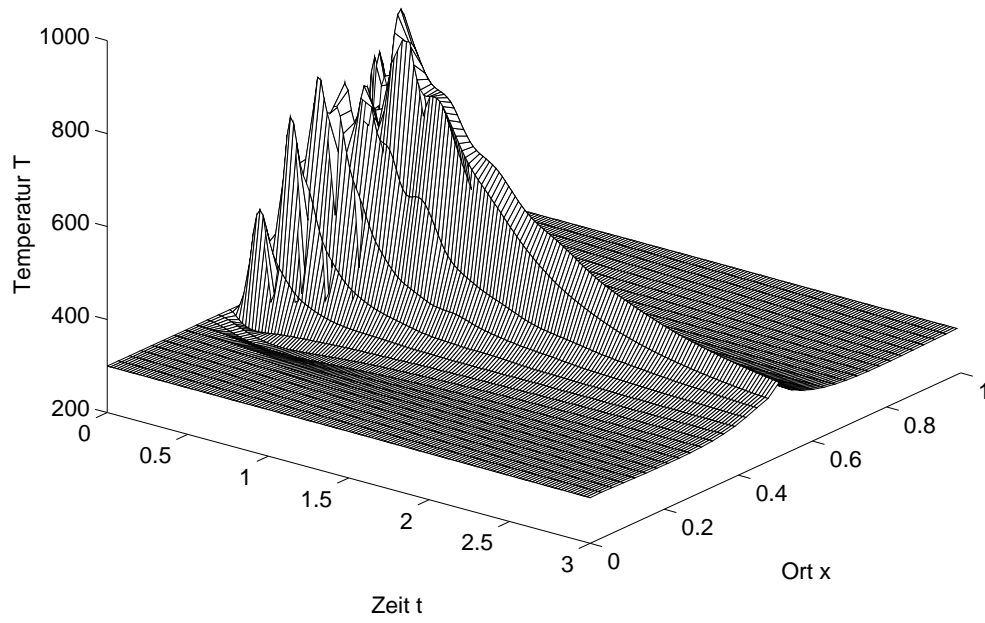
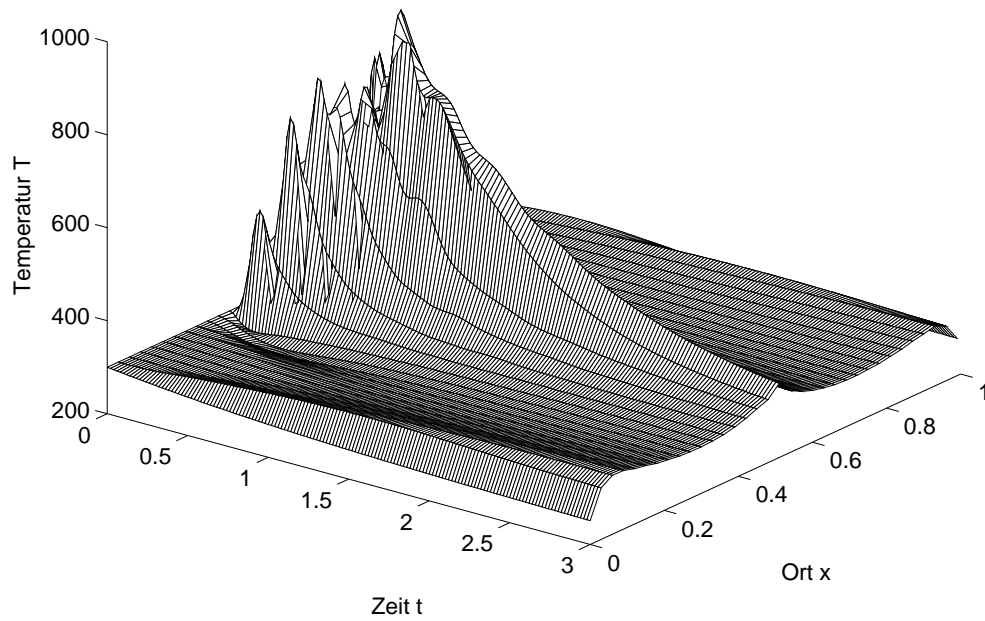


ABBILDUNG 4.2. Geschwindigkeit v bei $t = 3$.

BEMERKUNG 4.2 (STABILITÄT). Wir werden hier die *Stabilität* des Schemas nicht numerisch testen, da in der Lösung unseres Beispiels keine „Wellen“ von wesentlicher Größe den Rand passieren. Aus dem gleichen Grund verzichten wir hier auf eine numerische Untersuchung der *Absorptionsqualität* der Randbedingungen, weil keine sinnvolle Vergleichssituation vorliegt. Die Oszillationen in der Nähe des Randes deuten auf gewisse Probleme bei den ARBen hin; eventuell ist die Diskretisierung der ARBen hier nicht sehr günstig oder sogar *instabil*.

ABBILDUNG 4.3. Geschwindigkeit v bei Dirichlet-Randbedingung.ABBILDUNG 4.4. Geschwindigkeit v bei absorb. Randbedingung 1. Ordnung.

ABBILDUNG 4.5. Temperatur T bei Dirichlet-Randbedingung.ABBILDUNG 4.6. Temperatur T bei absorb. Randbedingung 1. Ordnung.

KAPITEL 5

Numerische Betrachtung des nichtlinearen Falls in 2D

Wir betrachten quasilineare hyperbolische Systeme in 2D in der Erhaltungsform

$$(5.1) \quad U_t + F(U)_x + G(U)_y = H(U), \quad 0 \leq x \leq L, \quad -\infty < y < \infty, \quad t \geq 0$$

und wollen hierfür die Klasse von Lax-Wendroff-Verfahren aus Abschnitt 4.1 auf zwei Ortsdimensionen verallgemeinern. Die Diskretisierung des (x, y, t) -Raumes $[0, L] \times \mathbb{R} \times \mathbb{R}_0^+$ geschieht wie in Kapitel 3.

Wiederum wird mit $U_{j,\ell}^n \in \mathbb{R}^N$ die *Approximation* der exakten Lösung $U(x_j, y_\ell, t_n)$ zum Zeitpunkt t_n an den diskreten Gitterpunkten (x_j, y_ℓ) bezeichnet. Desweiteren sind $F_{j,\ell}^n$, $G_{j,\ell}^n$, $H_{j,\ell}^n$ Abkürzungen für $F(U_{j,\ell}^n)$, $G(U_{j,\ell}^n)$ bzw. $H(U_{j,\ell}^n)$.

5.1. Das numerische Schema für 2D

Der Ausgangspunkt ist wie in 1D eine Taylorentwicklung von $U(x, y, t + k)$ in der Zeit:

$$U(x, y, t + k) = U(x, y, t) + kU_t(x, y, t) + \frac{1}{2}k^2U_{tt}(x, y, t) + O(k^3).$$

Dabei ist U Lösung von (5.1). Wir gehen analog zum eindimensionalen Fall vor, um die dortige Vorgehensweise zu verallgemeinern und ersetzen mit Hilfe der Gleichung (5.1) die auftretenden Zeitableitungen durch Ortsableitungen. Mit einem *Verfahrensparameter* $s \in]0, 1]$ erhalten wir:

$$\begin{aligned} U(x, y, t + k) &= U(x, y, t) + kU_t(x, y, t) + \frac{1}{2}k^2U_{tt}(x, y, t) + O(k^3) \\ &= U(x, y, t) + k \left(1 - \frac{1}{2s}\right) U_t(x, y, t) \\ &\quad + \frac{k}{2s} \left[\frac{\partial}{\partial t} (U(x, y, t) + skU_t(x, y, t)) \right] + O(k^3) \\ &= U(x, y, t) + k \left(1 - \frac{1}{2s}\right) U_t(x, y, t) + \frac{k}{2s} U_t(x, y, t + sk) + O(k^3) \\ &\stackrel{(5.1)}{=} U(x, y, t) - k \left(1 - \frac{1}{2s}\right) [F(U(x, y, t))_x + G(U(x, y, t))_y - H(U(x, y, t))] \\ &\quad - \frac{k}{2s} [F(U(x, y, t + sk))_x + G(U(x, y, t + sk))_y - H(U(x, y, t + sk))] \\ &\quad + O(k^3), \end{aligned}$$

wobei wir zuvor

$$\begin{aligned} U(x, y, t + sk) &= U(x, y, t) + skU_t(x, y, t) + O(k^2) \\ &\stackrel{(5.1)}{=} U(x, y, t) - sk [F(U(x, y, t))_x + G(U(x, y, t))_y - H(U(x, y, t))] + O(k^2) \end{aligned}$$

benutzt haben. Eine *Diskretisierung* der zwei Schritte sieht wie folgt aus:

$$\begin{aligned} (5.2a) \quad U_{j,\ell}^{n+s} &= U_{j,\ell}^n - sk \left(\frac{F_{j+\frac{1}{2},\ell}^n - F_{j-\frac{1}{2},\ell}^n}{h} + \frac{G_{j,\ell+\frac{1}{2}}^n - G_{j,\ell-\frac{1}{2}}^n}{h} - H_{j,\ell}^n \right) \\ &= U_{j,\ell}^n - s\lambda (F_{j+\frac{1}{2},\ell}^n - F_{j-\frac{1}{2},\ell}^n + G_{j,\ell+\frac{1}{2}}^n - G_{j,\ell-\frac{1}{2}}^n - hH_{j,\ell}^n) \end{aligned}$$

$$\begin{aligned} (5.2b) \quad U_{j,\ell}^{n+1} &= U_{j,\ell}^n - k \left(1 - \frac{1}{2s} \right) \left(\frac{F_{j+1,\ell}^n - F_{j-1,\ell}^n}{2h} + \frac{G_{j,\ell+1}^n - G_{j,\ell-1}^n}{2h} - H_{j,\ell}^n \right) \\ &\quad - \frac{k}{2s} \left(\frac{F_{j+\frac{1}{2},\ell}^{n+s} - F_{j-\frac{1}{2},\ell}^{n+s}}{h} + \frac{G_{j,\ell+\frac{1}{2}}^{n+s} - G_{j,\ell-\frac{1}{2}}^{n+s}}{h} - H_{j,\ell}^{n+s} \right) \\ &= U_{j,\ell}^n - \frac{\lambda}{2} \left[\left(1 - \frac{1}{2s} \right) (F_{j+1,\ell}^n - F_{j-1,\ell}^n + G_{j,\ell+1}^n - G_{j,\ell-1}^n - 2hH_{j,\ell}^n) \right. \\ &\quad \left. + \frac{1}{s} (F_{j+\frac{1}{2},\ell}^{n+s} - F_{j-\frac{1}{2},\ell}^{n+s} + G_{j,\ell+\frac{1}{2}}^{n+s} - G_{j,\ell-\frac{1}{2}}^{n+s} - hH_{j,\ell}^{n+s}) \right] \end{aligned}$$

VORGEHENSWEISE (IMPLEMENTIERUNG). Wir gehen analog zu Abschnitt 4.1 vor und berechnen auf den Zeitstufen t_n, t_{n+1} nur Größen mit ganzzahligen x - und y -Ortsindizes. Andererseits werden auf der Zwischenstufe t_{n+s} nur Approximationen mit halbzahligen x - und y -Ortsindizes explizit berechnet. Weitere auftretende Größen werden durch arithmetische Mittelung bestimmt und mit „ $\hat{}$ “ bzw. „ $\tilde{}$ “ bezeichnet.

Somit ergibt sich die folgende *Klasse von Zweischritt-Verfahren*:

Erster Schritt:

$$\begin{aligned} (5.3a) \quad U_{j+\frac{1}{2},\ell+\frac{1}{2}}^{n+s} &= \hat{U}_{j+\frac{1}{2},\ell+\frac{1}{2}}^n - s\lambda \left(\tilde{F}_{j+1,\ell+\frac{1}{2}}^n - \tilde{F}_{j,\ell+\frac{1}{2}}^n + \tilde{G}_{j+\frac{1}{2},\ell+1}^n - \tilde{G}_{j+\frac{1}{2},\ell}^n - h\hat{H}_{j+\frac{1}{2},\ell+\frac{1}{2}}^n \right), \\ \text{wobei} \quad \hat{U}_{j+\frac{1}{2},\ell+\frac{1}{2}}^n &:= \frac{1}{4} (U_{j,\ell}^n + U_{j+1,\ell}^n + U_{j,\ell+1}^n + U_{j+1,\ell+1}^n) \\ \text{und} \quad \tilde{U}_{j,\ell+\frac{1}{2}}^n &:= \frac{1}{2} (U_{j,\ell}^n + U_{j,\ell+1}^n) \\ \tilde{U}_{j+\frac{1}{2},\ell}^n &:= \frac{1}{2} (U_{j,\ell}^n + U_{j+1,\ell}^n) \quad \text{usw.} \end{aligned}$$

Zweiter Schritt:

$$(5.3b) \quad U_{j,\ell}^{n+1} = U_{j,\ell}^n - \frac{\lambda}{2} \left[\left(1 - \frac{1}{2s}\right) (F_{j+1,\ell}^n - F_{j-1,\ell}^n + G_{j,\ell+1}^n - G_{j,\ell-1}^n - 2hH_{j,\ell}^n) \right. \\ \left. + \frac{1}{s} \left(\tilde{F}_{j+\frac{1}{2},\ell}^{n+s} - \tilde{F}_{j-\frac{1}{2},\ell}^{n+s} + \tilde{G}_{j,\ell+\frac{1}{2}}^{n+s} - \tilde{G}_{j,\ell-\frac{1}{2}}^{n+s} - h\hat{H}_{j,\ell}^{n+s} \right) \right],$$

$$\text{wobei} \quad \hat{U}_{j,\ell}^{n+s} := \frac{1}{4} \left(U_{j-\frac{1}{2},\ell-\frac{1}{2}}^{n+s} + U_{j+\frac{1}{2},\ell-\frac{1}{2}}^{n+s} + U_{j-\frac{1}{2},\ell+\frac{1}{2}}^{n+s} + U_{j+\frac{1}{2},\ell+\frac{1}{2}}^{n+s} \right)$$

$$\text{und} \quad \tilde{U}_{j+\frac{1}{2},\ell}^{n+s} := \frac{1}{2} \left(U_{j+\frac{1}{2},\ell-\frac{1}{2}}^{n+s} + U_{j+\frac{1}{2},\ell+\frac{1}{2}}^{n+s} \right)$$

$$\tilde{U}_{j,\ell+\frac{1}{2}}^{n+s} := \frac{1}{2} \left(U_{j-\frac{1}{2},\ell+\frac{1}{2}}^{n+s} + U_{j+\frac{1}{2},\ell+\frac{1}{2}}^{n+s} \right) \quad \text{usw.}$$

BEMERKUNG 5.1 (SPEZIELLE VERFAHREN). Bei einem Verfahrensparameter $s = 1/2$ ergibt sich die *Variante von Zwas* [42].

5.2. Die Shallow-Water Gleichungen in zwei Ortsdimensionen

Die Shallow-Water Gleichungen in zwei Ortsdimensionen haben die folgende Gestalt:

$$(5.4) \quad \begin{aligned} u_t + uu_x + vu_y + \varphi_x - fv &= 0 \\ v_t + uv_x + vv_y + \varphi_y + fu &= 0 \\ \varphi_t + (\varphi u)_x + (\varphi v)_y &= 0. \end{aligned}$$

Dabei sind $u = u(x, y, t)$, $v = v(x, y, t)$ die *Horizontalkomponenten der Geschwindigkeit*, $f = f(x, y)$ ist der *Coriolis-Parameter* und $\varphi = gh(x, y, t)$ ist das *Erdpotential*, wobei g die Fallbeschleunigung und h die Höhe der Oberfläche ist. Für eine ausführliche Herleitung sei auf das Buch von Whitham [41] verwiesen. Das System (5.4) besitzt noch nicht die Erhaltungsform (5.1). Um dies zu erreichen, multiplizieren wir die erste Zeile von (5.4) mit φ und addieren dazu das u -fache der dritten Zeile:

$$\begin{aligned} \varphi u_t + u\varphi_t + \varphi uu_x + u(\varphi u)_x + \varphi\varphi_x + \varphi vu_y + u(\varphi v)_y - f\varphi v &= 0 \\ \iff (\varphi u)_t + (\varphi u^2 + \frac{1}{2}\varphi^2)_x + (\varphi uv)_y - f\varphi v &= 0. \end{aligned}$$

Ein analoges Vorgehen bei der zweiten Zeile von (5.4) liefert

$$\begin{aligned} \varphi v_t + v\varphi_t + \varphi uv_x + v(\varphi u)_x + \varphi vv_y + v(\varphi v)_y + \varphi\varphi_y + f\varphi u &= 0 \\ \iff (\varphi v)_t + (\varphi uv)_x + (\varphi v^2 + \frac{1}{2}\varphi^2)_y + f\varphi u &= 0. \end{aligned}$$

Damit ergibt sich

$$(5.5) \quad \begin{pmatrix} \varphi u \\ \varphi v \\ \varphi \end{pmatrix}_t + \begin{pmatrix} \varphi u^2 + \frac{1}{2}\varphi^2 \\ \varphi uv \\ \varphi u \end{pmatrix}_x + \begin{pmatrix} \varphi uv \\ \varphi v^2 + \frac{1}{2}\varphi^2 \\ \varphi v \end{pmatrix}_y = \begin{pmatrix} f\varphi v \\ -f\varphi u \\ 0 \end{pmatrix}$$

und mit $U := (\tilde{u}, \tilde{v}, \varphi)^T$, $\tilde{u} := \varphi u$, $\tilde{v} := \varphi v$ erhalten wir die *Erhaltungsform* (5.1), wobei

$$(5.6) \quad F(U) = \begin{pmatrix} \tilde{u}^2/\varphi + \frac{1}{2}\varphi^2 \\ \tilde{u}\tilde{v}/\varphi \\ \tilde{u} \end{pmatrix}, \quad G(U) = \begin{pmatrix} \tilde{u}\tilde{v}/\varphi \\ \tilde{v}^2/\varphi + \frac{1}{2}\varphi^2 \\ \tilde{v} \end{pmatrix}, \quad H(U) = \begin{pmatrix} f\tilde{v} \\ -f\tilde{u} \\ 0 \end{pmatrix}$$

ist.

5.2.1. Die Randbedingungen. Zur Formulierung der Randbedingungen betrachten wir die Shallow-Water Gleichungen in der Form

$$(5.7) \quad \begin{pmatrix} u \\ v \\ \varphi \end{pmatrix}_t + \begin{pmatrix} u & 0 & 1 \\ 0 & u & 0 \\ \varphi & 0 & u \end{pmatrix} \begin{pmatrix} u \\ v \\ \varphi \end{pmatrix}_x + \begin{pmatrix} v & 0 & 0 \\ 0 & v & 1 \\ 0 & \varphi & v \end{pmatrix} \begin{pmatrix} u \\ v \\ \varphi \end{pmatrix}_y + \begin{pmatrix} 0 & -f & 0 \\ f & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} u \\ v \\ \varphi \end{pmatrix} = 0.$$

Zu einer festen Zeit t und für jedes beliebige, feste y frieren wir die Koeffizientenmatrizen lokal an $x = 0$ ein und kennzeichnen dies mit einem Index Null. Damit können wir wie

bei den linearisierten Shallow-Water Gleichungen in Abschnitt 3.2 vorgehen, d. h. wir symmetrisieren die Koeffizienten simultan mit Hilfe von $S_0 := \text{diag}(c_0, c_0, 1)$, $c_0 = \sqrt{\varphi_0}$ zu

$$\begin{pmatrix} c_0 u \\ c_0 v \\ \varphi \end{pmatrix}_t + \underbrace{\begin{pmatrix} u_0 & 0 & c_0 \\ 0 & u_0 & 0 \\ c_0 & 0 & u_0 \end{pmatrix}}_{=A_0(u_0, c_0)} \begin{pmatrix} c_0 u \\ c_0 v \\ \varphi \end{pmatrix}_x + \underbrace{\begin{pmatrix} v_0 & 0 & 0 \\ 0 & v_0 & c_0 \\ 0 & c_0 & v_0 \end{pmatrix}}_{=B_0(v_0, c_0)} \begin{pmatrix} c_0 u \\ c_0 v \\ \varphi \end{pmatrix}_y + \underbrace{\begin{pmatrix} 0 & -f_0 & 0 \\ f_0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{=C_0(f_0)} \begin{pmatrix} c_0 u \\ c_0 v \\ \varphi \end{pmatrix} = 0$$

und diagonalisieren die Systemmatrix A_0 wie in Abschnitt 3.2.1. Dann hängt die zu benutzende ARB vom Vorzeichen von u_0 ab. Die ARBen für $u_0 > 0$ bzw. $u_0 < 0$ sind in Abschnitt 3.2.2 angegeben. Bei $u_0 = 0$ wird die 2. Komponente (zum Eigenwert Null) wie eine Ausströmkomponente mit (3.13) horizontal extrapoliert. An $x = L$ verfährt man analog.

VORGEHENSWEISE (NUMERISCHE RANDBEDINGUNGEN). Im Inneren berechnet man $U_{j,\ell}^{n+1}$, $j = 1, 2, \dots, J-1$, $\ell = \dots, -1, 0, 1, 2, \dots$ mit Hilfe des Zweischritt-Lax-Wendroff-Verfahrens (5.3). An $j = 0$ implementiert man die *absorbierenden Randbedingungen* für jedes $\ell = \dots, -1, 0, 1, 2, \dots$ folgendermaßen:

$$\begin{aligned} \text{Gegeben} & : U_{0,\ell}^n = (\varphi_{0,\ell}^n u_{0,\ell}^n, \varphi_{0,\ell}^n v_{0,\ell}^n, \varphi_{0,\ell}^n)^T, U_{1,\ell}^{n+1}, U_{2,\ell}^{n+1} \\ \text{Damit} & : \tilde{U}_{0,\ell}^n = (u_{0,\ell}^n, v_{0,\ell}^n, \varphi_{0,\ell}^n)^T, \tilde{U}_{1,\ell}^{n+1}, \tilde{U}_{2,\ell}^{n+1} \quad \text{analog.} \\ \text{Berechnung} & : c_{0,\ell}^n = \sqrt{\varphi_{0,\ell}^n} \quad (\text{„Schallgeschwindigkeit“}). \\ \text{Transformation} & : V_{0,\ell}^n = T^{-1} S_{0,\ell} \tilde{U}_{0,\ell}^n, V_{j,\ell}^{n+1} = T^{-1} S_{0,\ell} \tilde{U}_{j,\ell}^{n+1}, \quad j = 1, 2 \\ & \text{mit } T^{-1} = T \text{ aus (3.7) und } S_{0,\ell} := \text{diag}(c_{0,\ell}^n, c_{0,\ell}^n, 1). \\ \text{Ausströmextrapolation} & : (V^-)_{0,\ell}^{n+1} = 2(V^-)_{1,\ell}^{n+1} - (V^-)_{2,\ell}^{n+1}. \\ \text{Absorb. Randbedingung} & : (V^+)_{0,\ell}^{n+1} = (V^+)_{0,\ell}^n - \frac{\lambda}{2} X_0 \left((V^-)_{0,\ell+1}^{n+\frac{1}{2}} - (V^-)_{0,\ell-1}^{n+\frac{1}{2}} \right) \\ & \quad - k K_0 (V^-)_{0,\ell}^{n+\frac{1}{2}} \\ \text{bzgl.} & \Lambda_{0,\ell} = \text{diag}(u_{0,\ell}^n + c_{0,\ell}^n, u_{0,\ell}^n, u_{0,\ell}^n - c_{0,\ell}^n), \\ & \tilde{B}_{0,\ell} = T^{-1} B_0(v_{0,\ell}^n, c_{0,\ell}^n) T, \tilde{C}_{0,\ell} = T^{-1} C_0(f_{0,\ell}^n) T, \\ & \text{wobei } (V^-)_{0,\ell}^{n+\frac{1}{2}} := \frac{1}{2} \left((V^-)_{0,\ell}^{n+1} + (V^-)_{0,\ell}^n \right) \text{ usw. ist.} \\ \text{Rücktransformation} & : U_{0,\ell}^{n+1} = S_{0,\ell} T V_{0,\ell}^{n+1}. \end{aligned}$$

Für die *reflektierende Randbedingung* $u_0^n = v_0^n = 0$ sind diese Transformationen nicht nötig. Daneben muß noch das Potential extrapoliert werden

$$\varphi_{0,\ell}^{n+1} = 2\varphi_{1,\ell}^{n+1} - \varphi_{2,\ell}^{n+1},$$

um das numerische Schema zu schließen.

An $j = J$ geht man analog vor. Im folgenden setzen wir für die numerischen Untersuchungen $f = 0$, womit die Matrizen K_0 und K_L in den ARBen verschwinden.

5.3. Numerische Ergebnisse

Es wird die Stabilität des LW-Schemas aus Abschnitt 5.1 und die Absorptionsqualität der Randbedingungen 0. und 1. Ordnung numerisch untersucht.

5.3.1. Das Testbeispiel. Wir verwenden als Beispiel die *Shallow-Water Gleichungen* aus Abschnitt 5.2 auf $0 \leq x \leq 1$. Als Anfangsvorgabe haben wir wie in Abschnitt 3.3.1 die Bedingung $u = v = 0$. Für φ benutzen wir die Anfangsverteilung

$$(5.8) \quad \varphi(x, y, 0) := \begin{cases} 0.5 + 0.4 \cdot \cos^2\left(\frac{\pi}{2} \frac{|(x-0.5, y)^T|}{0.45}\right) & |(x-0.5, y)^T| < 0.45, \\ 0.5 & \text{sonst,} \end{cases}$$

die in Abbildung 5.1 auch mit Hilfe von Niveaulinien dargestellt ist.

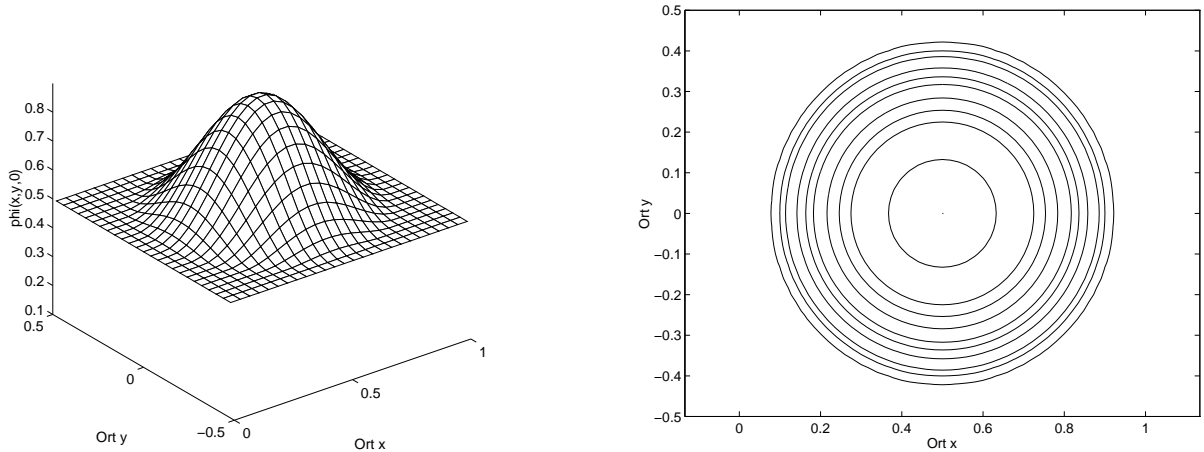


ABBILDUNG 5.1. Anfangsverteilung für φ

Die Anfangsverteilung (5.8) für φ wurde gegenüber (3.14) um 0.5 verschoben. Dies ist sinnvoll, da hier φ im wesentlichen gleich der (positiven) *Wassertiefe* ist, während φ in Abschnitt 3.3 nur die *kleine Störung* einer positiven Tiefe ist und als solche positiv, negativ oder Null sein kann. Auch wegen der Transformation von Seite 68 muß φ zumindest am Rand positiv sein.

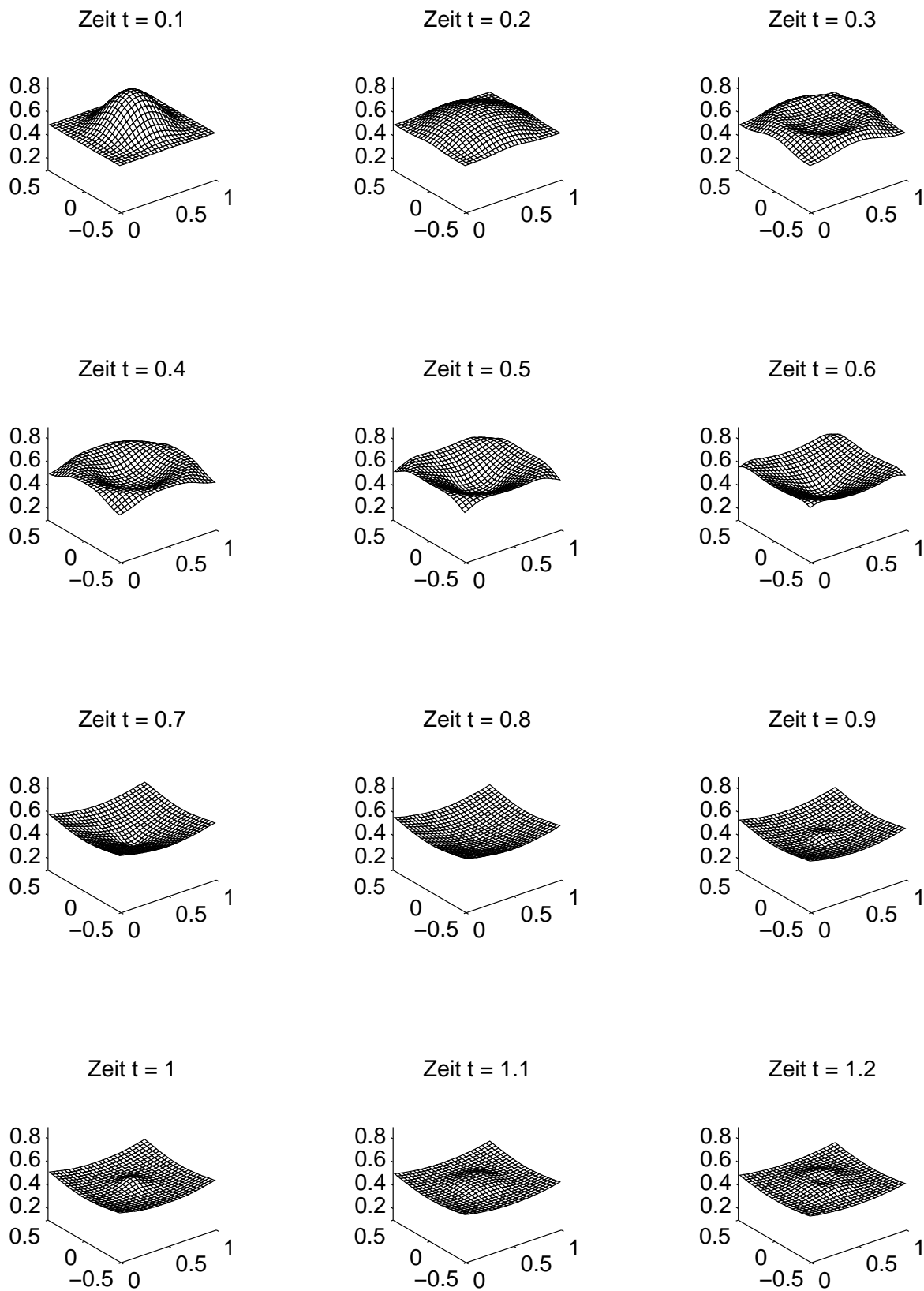
Wegen Bemerkung 3.2 ist eine *notwendige Stabilitätsbedingung* für diese Gleichung (5.5)

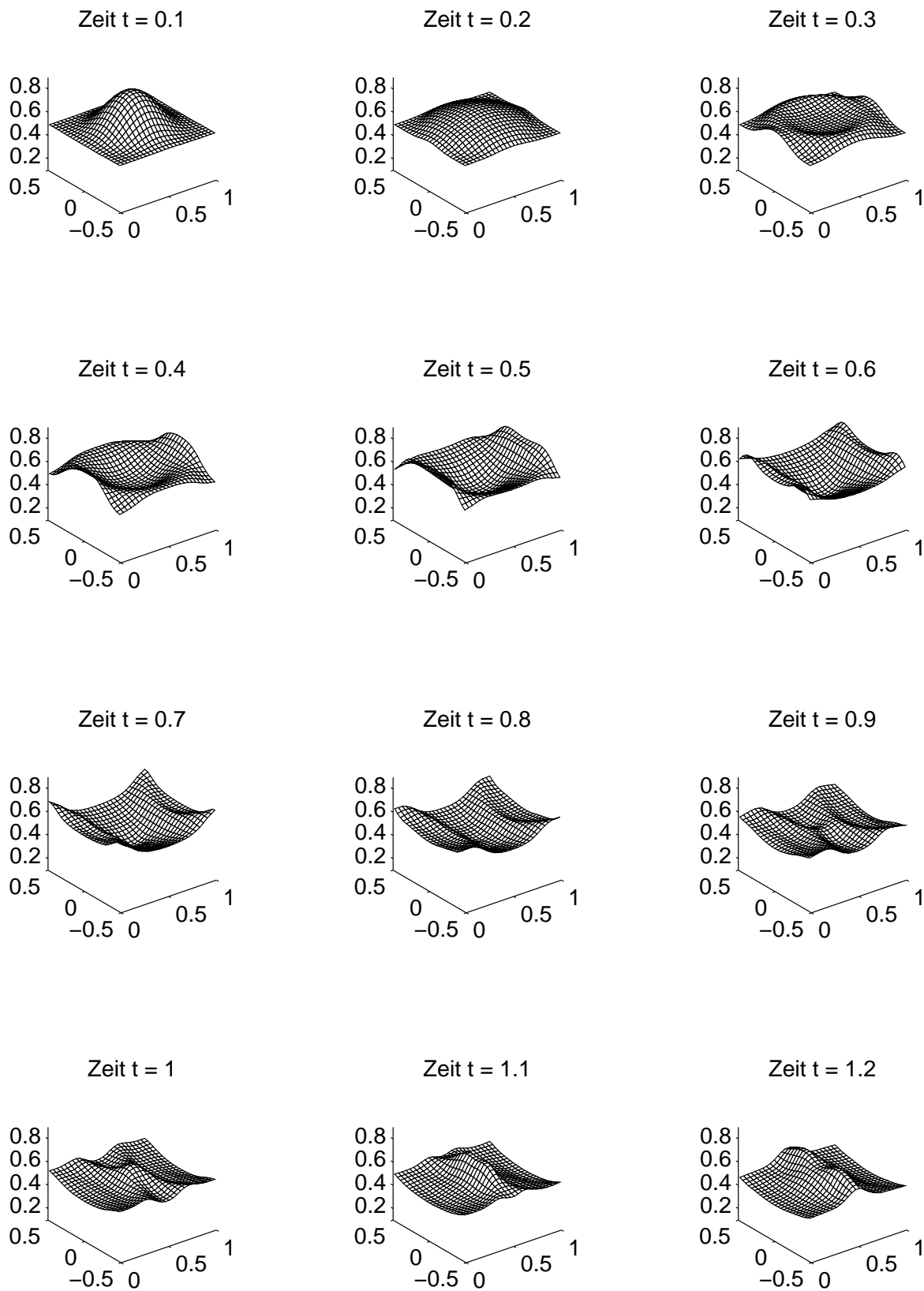
$$(5.9) \quad \lambda \leq \frac{1}{2\sqrt{2}} \frac{1}{\lambda_{\max}}$$

$$\text{mit } \lambda_{\max} = \max_{(x,y,t) \in [0,1] \times (-\infty, \infty) \times [0,T]} \left(\sqrt{\varphi(x, y, t)} + |u(x, y, t)|, \sqrt{\varphi(x, y, t)} + |v(x, y, t)| \right).$$

Damit diese Restriktion an λ nicht zu ungünstig ist, haben wir in (5.8) den Faktor 0.4 eingeführt. Für die folgenden Rechnungen wählen wir zunächst das Schrittweitenverhältnis $\lambda = 0.25$.

Zunächst wollen wir die zeitliche Entwicklung des Potentials φ bei einer Zeitschrittweite $\Delta t = 0.01$ und einem Verfahrensparameter $s = 0.5$ (*Verfahren von Zwas*) betrachten. Dies zeigen Abbildung 5.2 und Abbildung 5.3 bei ARBen 1. Ordnung bzw. reflektierender Randbedingung. Abbildung 5.4 und Abbildung 5.5 zeigen das Entsprechende mittels Niveaulinien.

ABBILDUNG 5.2. φ für $t = 0.1, 0.2, \dots, 1.2$ bei absorb. Randbedingung 1. Ordnung

ABBILDUNG 5.3. φ für $t = 0.1, 0.2, \dots, 1.2$ bei reflektierender Randbedingung

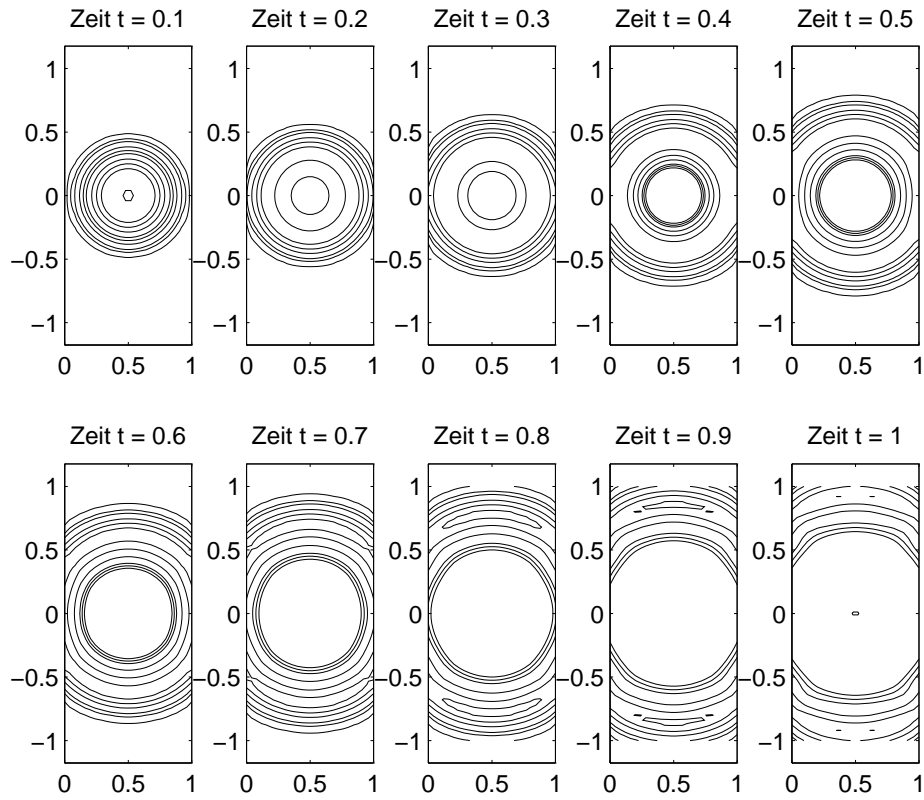


ABBILDUNG 5.4. φ für $t = 0.1, 0.2, \dots, 1$ bei absorb. Randbedingung 1. Ordnung

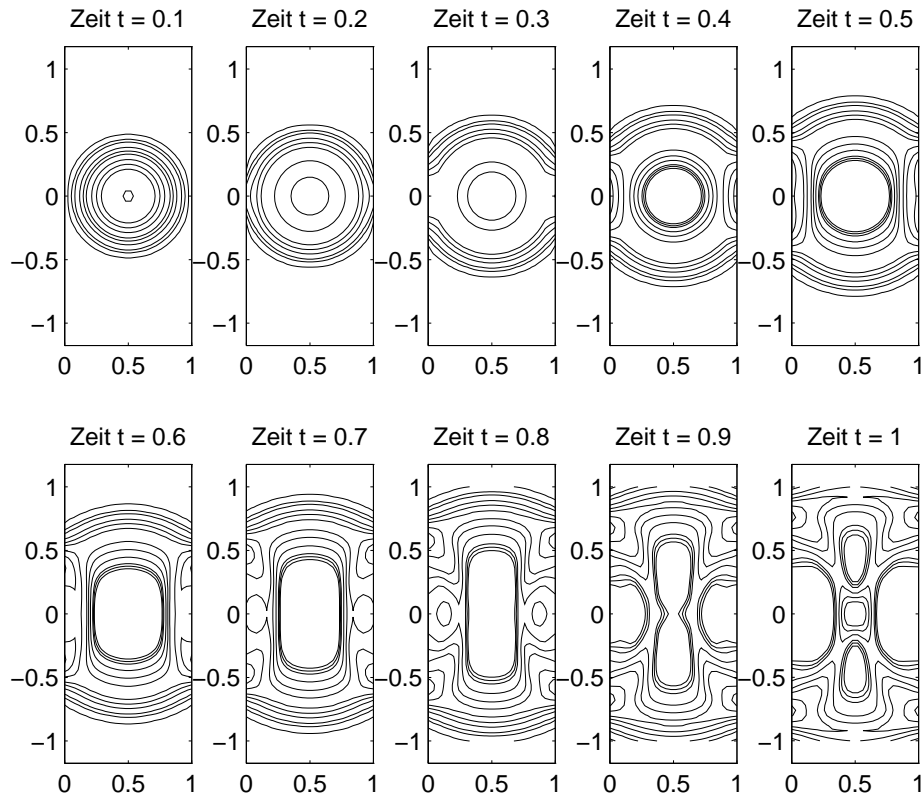


ABBILDUNG 5.5. φ für $t = 0.1, 0.2, \dots, 1$ bei reflektierender Randbedingung

5.3.2. Numerische Untersuchung der Stabilität. Analog zu Abschnitt 3.3.2 verwenden wir neben $u = v = 0$ für φ die modulierte Anfangsverteilung

$$\varphi(x, y, 0) := \begin{cases} 0.5 + 0.4 \cdot \sin\left(2\pi p|(x - 0.5, y)^T|\right) \cos^2\left(\frac{\pi}{2} \frac{|(x-0.5, y)^T|}{0.45}\right) & |(x - 0.5, y)^T| < 0.45, \\ 0.5 & \text{sonst} \end{cases}$$

mit dem *Modulationsparameter* p , um die Stabilität des LW-Verfahrens (5.3) in Kombination mit absorbierenden Randbedingungen 1. Ordnung zu untersuchen. Dabei geben wir im folgenden den zeitlichen Verlauf der diskreten L^2 -Norm (3.16) von u , v und $\varphi - 0.5$ an.

Die Ergebnisse für $\lambda = 0.25$, $s = 0.5$ und $\Delta t = 0.001$, die wir bei $p = 10$ bzw. $p = 30$ erhalten, sind in Abbildung 5.6 zu sehen.

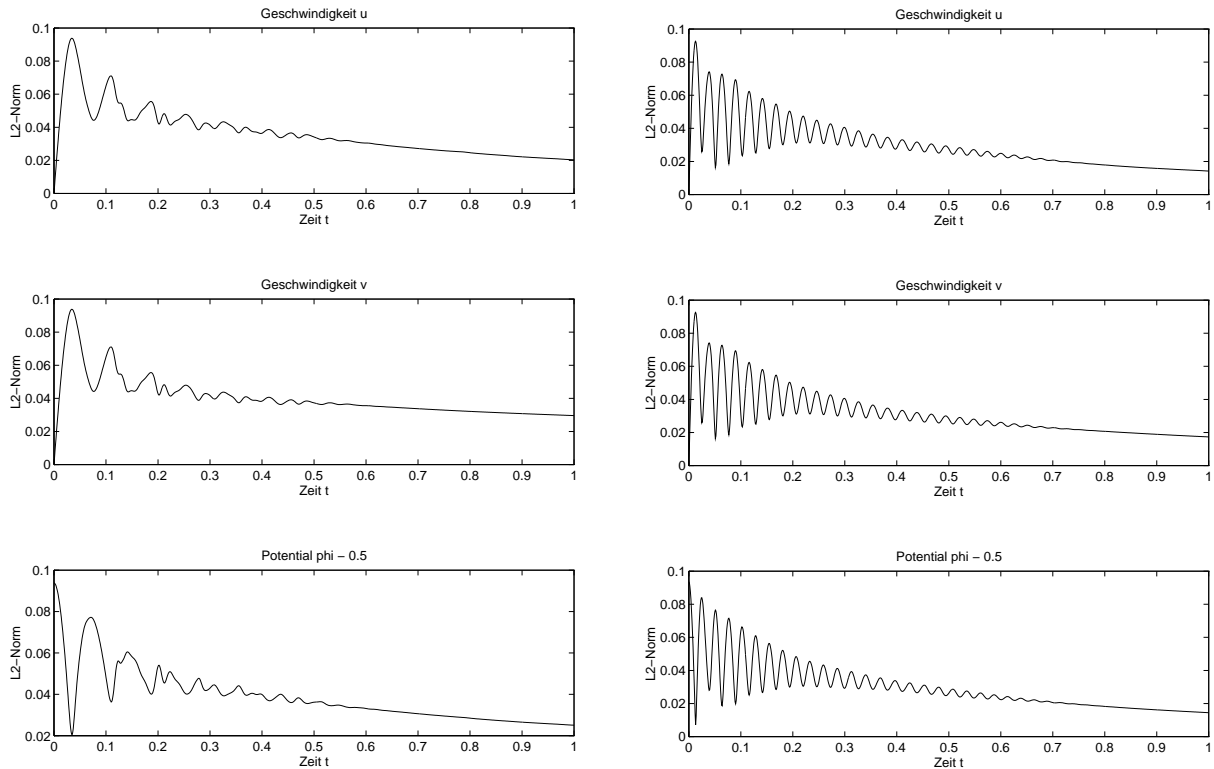


ABBILDUNG 5.6. L^2 -Normen von u , v und $\varphi - 0.5$ für $p = 10$ bzw. $p = 30$

Wir wollen nun den Einfluß von λ auf die Stabilität betrachten und verwenden im folgenden unterschiedliche Werte für das Schrittweitenverhältnis λ bei einem Modulationsparameter $p = 10$ und einer Zeitschrittweite $\Delta t = 0.001$. Die Resultate sind in den Abbildungen 5.7 und 5.8 dargestellt.

FAZIT (STABILITÄT). Erwartungsgemäß klingen die L^2 -Normen schnell ab, da die Wellen sich ausbreiten. Das *oszillatorische Verhalten* der L^2 -Normen entsteht wie in Abschnitt 3.3.2 beim Durchgang der hohen Frequenzen der Modulation durch den Rand bei $x = 0$ bzw. $x = 1$. Das Zweischritt-LW-Schema verhält sich bei diesen Beispielen mit einem Schrittweitenverhältnis $\lambda = 0.25$ (und bei kleineren Werten $\lambda = 0.1$, $\lambda = 0.18$) *stabil*. Auch bei $\lambda = 0.3$ verhält sich das Schema stabil.

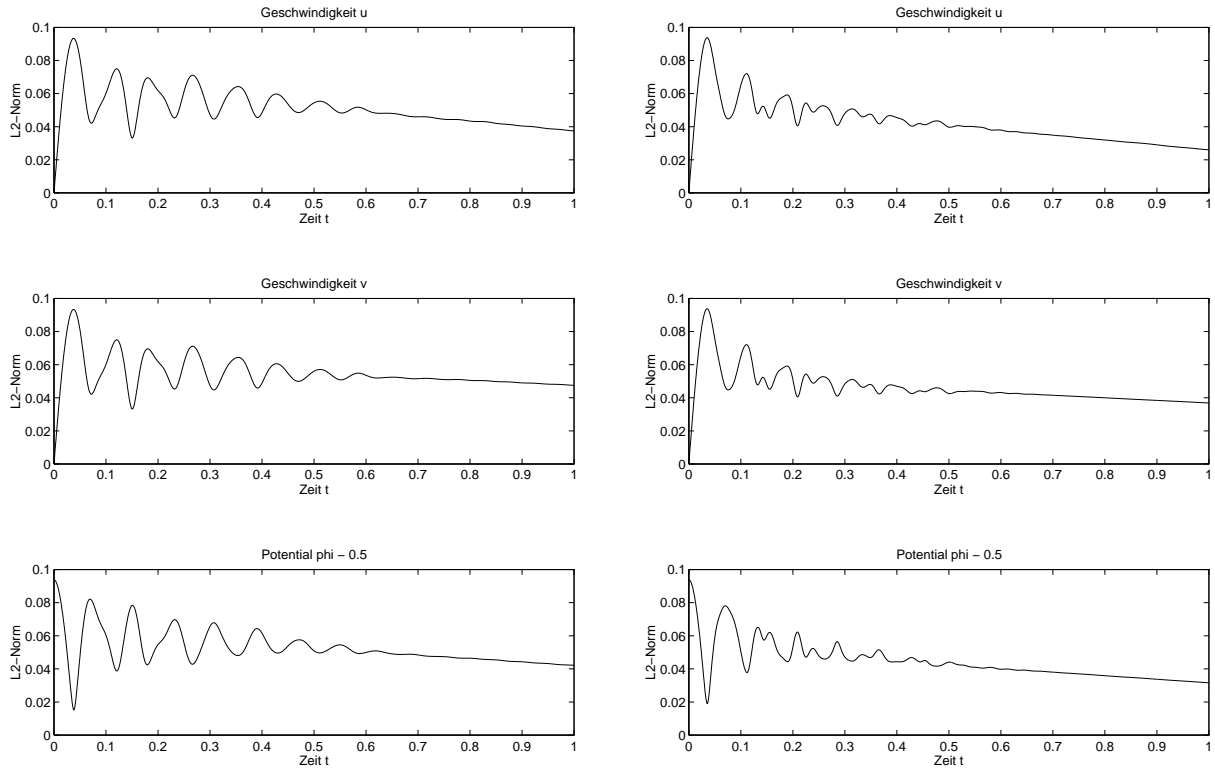


ABBILDUNG 5.7. L^2 -Normen von u , v und $\varphi - 0.5$ für $\lambda = 0.1$ bzw. $\lambda = 0.18$

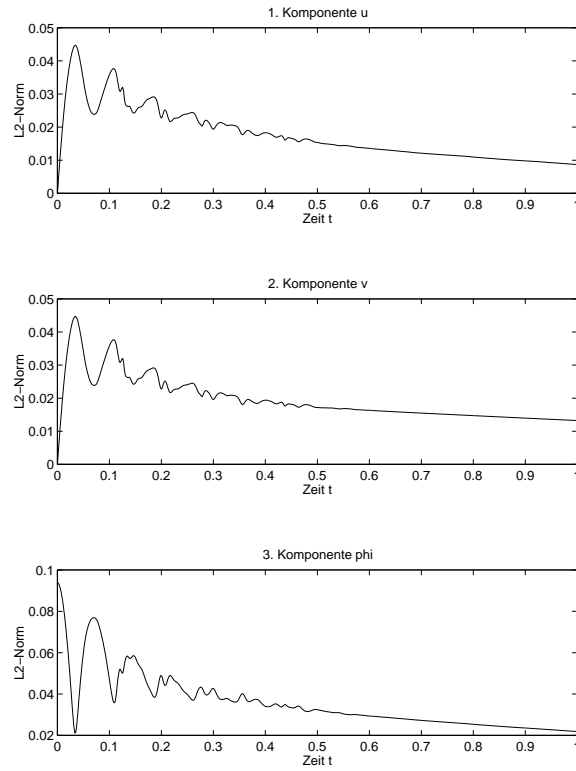


ABBILDUNG 5.8. L^2 -Normen von u , v und $\varphi - 0.5$ für $\lambda = 0.3$

5.3.3. Die Absorptionsqualität der absorbierenden Randbedingungen. Wie in Abschnitt 3.3.3 betrachten wir die numerischen Reflexionen bei den ARBen 0. und 1. Ordnung und relativieren sie an denen bei der reflektierenden Randbedingung. Dabei wird der *reflektierte Anteil* analog zu Abschnitt 3.3.3 berechnet.

Die folgenden Abbildungen 5.9 und 5.10 geben die zeitliche Entwicklung des reflektierten Anteils im Potential φ_{refl} wieder. Bei den reflektierenden Randbedingungen wurden die Niveauwerte von Seite 72 (ohne die Verschiebung um 0.5, vgl. (5.8)) und bei den ARBen wurden diese Werte mit 0.5 multipliziert. Zu den Rechnungen wurde $\lambda = 0.25$, $\Delta t = 0.01$ verwendet.

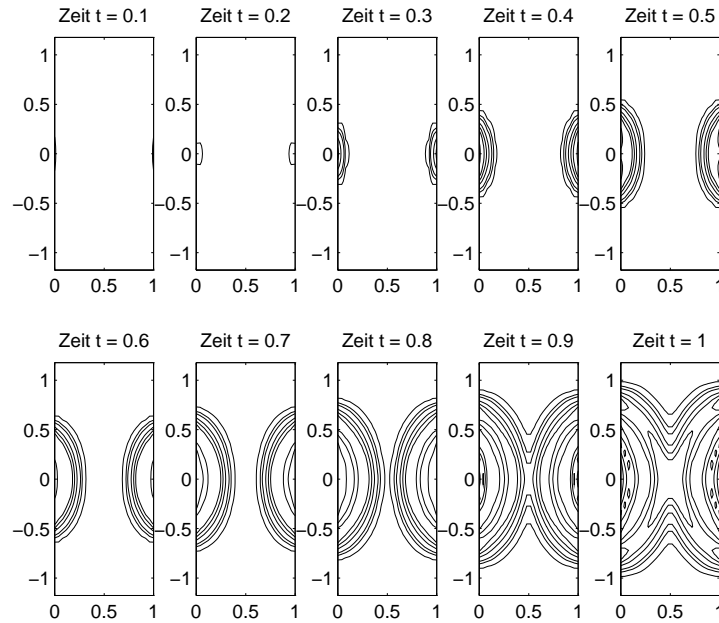


ABBILDUNG 5.9. reflektierter Anteil von φ bei reflektierender Randbedingung

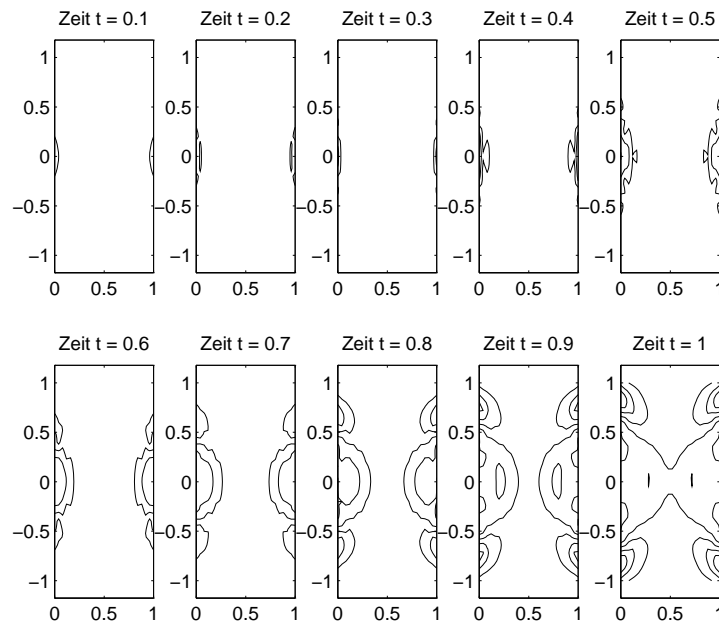


ABBILDUNG 5.10. reflektierter Anteil von φ bei absorb. Randbedingung 1. Ordnung

Wie auch die folgenden Tabellen zeigen, ist bei φ_{refl} der Unterschied zwischen den ARBen nullter und erster Ordnung gering; dies ist auch in den Abbildungen 5.11 und 5.12 zu erkennen. Die Prozentangaben beziehen sich dabei auf den jeweiligen Wert der L^2 -Norm des reflektierten Anteils bei Verwendung der reflektierenden Randbedingung.

Reflektierter Anteil bei $T = 0.6$

Randbedingung an $x = 0$ bzw. $x = 1$	$\ u_{refl}\ _{h,h}$	$\ v_{refl}\ _{h,h}$	$\ \varphi_{refl}\ _{h,h}$
totale Reflexion $u = v = 0$	0.07318	0.02850	0.05587
absorbierende Randbedingung 0. Ordnung (1.19)	0.00592 [8.1%]	0.00321 [11.3%]	0.00405 [7.3%]
absorbierende Randbedingung 1. Ordnung (1.21)	0.00478 [6.5%]	0.00276 [9.7%]	0.00324 [5.8%]

TABELLE 5.5. Ergebnisse für $s = 0.5$, $\lambda = 0.25$ und $\Delta t = 0.01$.

Reflektierter Anteil bei $T = 0.6$

Randbedingung an $x = 0$ bzw. $x = 1$	$\ u_{refl}\ _{h,h}$	$\ v_{refl}\ _{h,h}$	$\ \varphi_{refl}\ _{h,h}$
totale Reflexion $u = v = 0$	0.06743	0.02500	0.05098
absorbierende Randbedingung 0. Ordnung (1.19)	0.00546 [8.1%]	0.00276 [11.0%]	0.00380 [7.5%]
absorbierende Randbedingung 1. Ordnung (1.21)	0.00439 [6.5%]	0.00230 [9.2%]	0.00307 [6.0%]

TABELLE 5.6. Ergebnisse für $s = 0.5$, $\lambda = 0.25$ und $\Delta t = 0.002$.

In der folgenden Abbildung 5.11 geben wir die L^2 -Norm des reflektierten Anteils wieder. Dabei ist $\lambda = 0.25$, $\Delta t = 0.002$. Die gestrichelte Linie ist der Verlauf der L^2 -Norm bei ARB 0. Ordnung und die durchgezogene bei ARB 1. Ordnung.

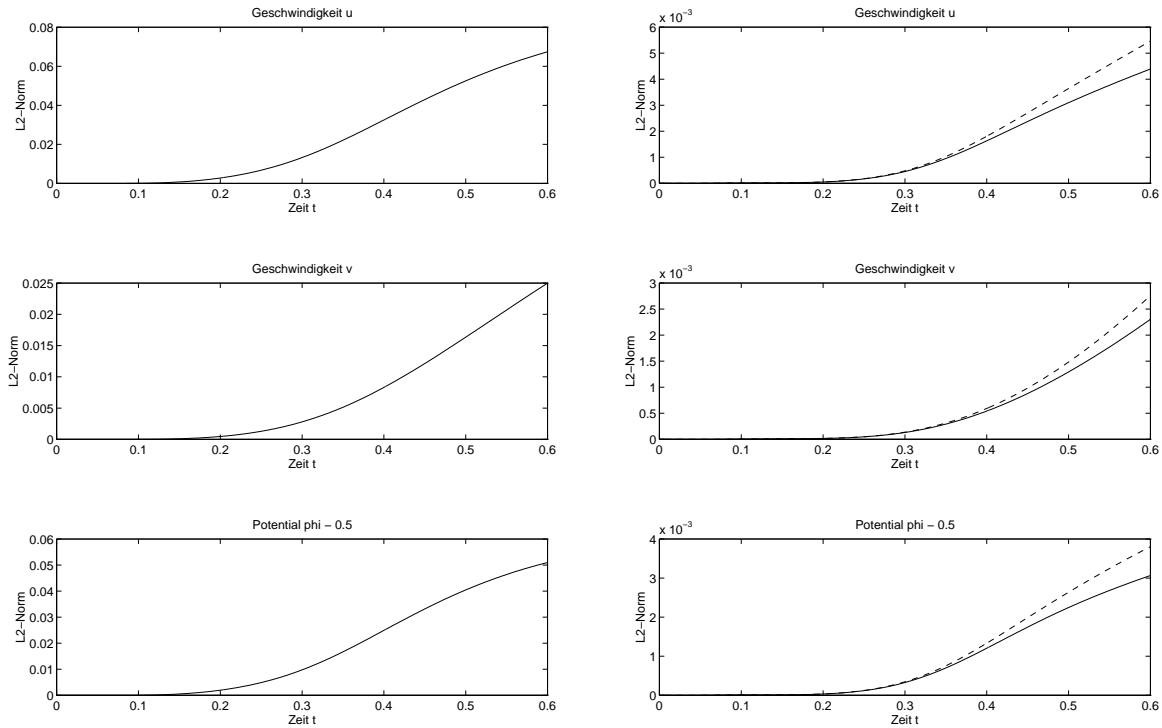


ABBILDUNG 5.11. Reflektierter Anteil bei totaler Reflexion bzw. ARBen: Parameter $\lambda = 0.25$, $\Delta t = 0.002$.

Reflektierter Anteil bei $T = 0.6$

Randbedingung an $x = 0$ bzw. $x = 1$	$\ u_{refl}\ _{h,h}$	$\ v_{refl}\ _{h,h}$	$\ \varphi_{refl}\ _{h,h}$
totale Reflexion $u = v = 0$	0.07462	0.02869	0.05733
absorbierende Randbedingung 0. Ordnung (1.19)	0.00611 [8.2%]	0.00348 [12.1%]	0.00420 [7.3%]
absorbierende Randbedingung 1. Ordnung (1.21)	0.00507 [6.8%]	0.00311 [10.8%]	0.00342 [6.0%]

TABELLE 5.7. Ergebnisse für $s = 0.5$, $\lambda = 0.1$ und $\Delta t = 0.005$.

In Abbildung 5.12 geben wir die L^2 -Norm des reflektierten Anteils wieder. Dabei ist $\lambda = 0.1$, $\Delta t = 0.001$.

Reflektierter Anteil bei $T = 0.6$

Randbedingung an $x = 0$ bzw. $x = 1$	$\ u_{refl}\ _{h,h}$	$\ v_{refl}\ _{h,h}$	$\ \varphi_{refl}\ _{h,h}$
totale Reflexion $u = v = 0$	0.06770	0.02527	0.05119
absorbierende Randbedingung 0. Ordnung (1.19)	0.00549 [8.1%]	0.00278 [11.0%]	0.00382 [7.5%]
absorbierende Randbedingung 1. Ordnung (1.21)	0.00440 [6.5%]	0.00231 [9.2%]	0.00307 [6.0%]

TABELLE 5.8. Ergebnisse für $s = 0.5$, $\lambda = 0.1$ und $\Delta t = 0.001$.

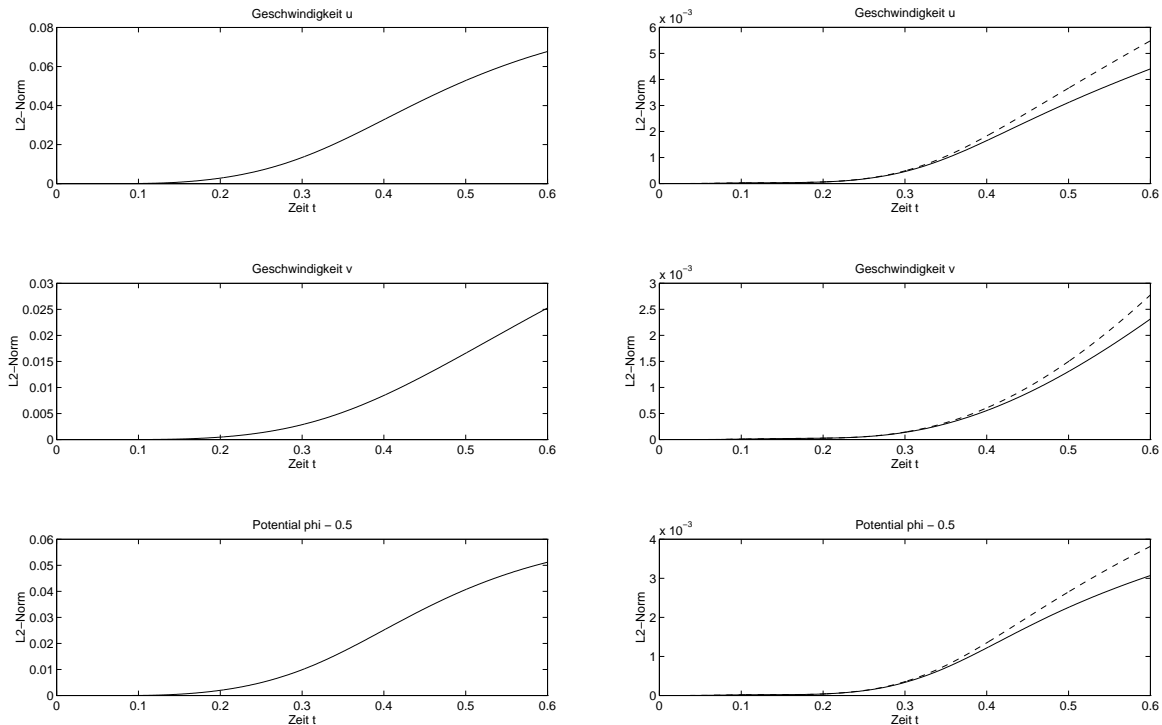


ABBILDUNG 5.12. Reflektierter Anteil bei totaler Reflexion bzw. ARBen:
Parameter $\lambda = 0.1$, $\Delta t = 0.001$.

Wir variieren nun den Verfahrensparameter $s \in]0, 1]$. Für den reflektierten Anteil ergeben sich die folgenden Resultate:

Reflektierter Anteil bei $T = 0.6$

Randbedingung an $x = 0$ bzw. $x = 1$	$\ u_{refl}\ _{h,h}$	$\ v_{refl}\ _{h,h}$	$\ \varphi_{refl}\ _{h,h}$
totale Reflexion $u = v = 0$	0.07455	0.02988	0.05707
absorbierende Randbedingung 0. Ordnung (1.19)	0.00577 [7.7%]	0.00340 [11.4%]	0.00388 [6.8%]
absorbierende Randbedingung 1. Ordnung (1.21)	0.00471 [6.3%]	0.00307 [10.2%]	0.00313 [5.5%]

TABELLE 5.9. Ergebnisse für $s = 0.25$, $\lambda = 0.25$ und $\Delta t = 0.01$.

Reflektierter Anteil bei $T = 0.6$

Randbedingung an $x = 0$ bzw. $x = 1$	$\ u_{refl}\ _{h,h}$	$\ v_{refl}\ _{h,h}$	$\ \varphi_{refl}\ _{h,h}$
totale Reflexion $u = v = 0$	0.07248	0.02801	0.05526
absorbierende Randbedingung 0. Ordnung (1.19)	0.00560 [8.3%]	0.00322 [11.5%]	0.00414 [7.5%]
absorbierende Randbedingung 1. Ordnung (1.21)	0.00483 [6.7%]	0.00273 [9.7%]	0.00331 [6.0%]

TABELLE 5.10. Ergebnisse für $s = 1$, $\lambda = 0.25$ und $\Delta t = 0.01$.

BEMERKUNG 5.2. Für einen Verfahrensparameter $s \leq 0.13$ ergaben sich bei unseren Testläufen Instabilitäten.

FAZIT (ABSORPTIONSQUALITÄT). Bei allen Programmläufen ist erwartungsgemäß die L^2 -Norm des reflektierten Anteils bei der ARB 1. Ordnung geringer als bei der ARB nullter Ordnung. Der Übergang von der ARB nullter Ordnung zu der ARB erster Ordnung liefert eine Reduktion der numerischen Reflektionen um einen Faktor 1.1–1.25 bezogen auf die L^2 -Norm des reflektierten Anteils der einzelnen Komponenten. Dies entspricht einem Unterschied von 1–2% in den Prozentangaben der vorangegangenen Tabellen. Beide Schrittweitenverhältnisse $\lambda = 0.25$ und $\lambda = 0.1$ ergeben ähnliche Resultate. Bei $s = 0.25$ bzw. $s = 1$ ist der reflektierte Anteil praktisch gleich groß wie bei $s = 0.5$, was zu erwarten war, da nur das innere Schema und nicht die Randbedingungen vom Verfahrensparameter s abhängen.

ANHANG A

Glossar

1.1. Notation

Einige Symbole werden mehrfach verwendet; dies ist jedoch auf wenige Fälle beschränkt und dürfte nicht zu Verwechslungen führen. Wenn nicht anders definiert, bedeutet T die Transposition, $\bar{}$ die komplexe Konjugation und $* $ die Adjunktion. Um das Argument einer Funktion zu betonen, wird z. B. $f(x)$ anstatt f benutzt, ohne daß ein bestimmter Funktionswert an dieser Stelle gemeint ist. Die Einschränkung einer von zwei Variablen abhängigen Funktion wie $U(x, t)$ auf die Variable x bei festgehaltenen t wird mit $U(., t)$ bezeichnet.

1.2. Mathematische Größen

f	Anfangsverteilung
g	Inhomogenität in ARB
h	Ortsschrittweite Δx
i	imaginäre Einheit
j	Ortsindex für x
k	Zeitschrittweite Δt
ℓ	Ortsindex für y in 2D
m	Partitionsindex
n	Zeitindex
p	Modulationsparameter
s	Verfahrensparameter
E	Shift-Operator: $EU_j := U_{j+1}$
K	Matrix in ARB
L	Ortsintervall $[0, L]$ für x
N	Dimension von Vektoren
\hat{Q}	Amplifikationsmatrix von Q
X	Matrix in ARB
T	Zeitintervall $[0, T]$ für t
$\lambda = k/h$	Schrittweitenverhältnis
μ	Eigenwerte von Q
ν	Courantzahl

1.3. Physikalische Größen

$e = 1.6021773 \cdot 10^{-19} \text{ C}$	Elementarladung
$m_e = 9.1093897 \cdot 10^{-31} \text{ kg}$	Ruhemasse eines Elektrons
$k_B = 1.3806581 \cdot 10^{-23} \text{ J K}^{-1}$	Boltzmann-Konstante
$\varepsilon_0 = 8.85419 \cdot 10^{-12} \text{ A s V}^{-1} \text{ m}^{-1}$	Dielektrizitätskonstante für Vakuum
n	Partikeldichte der Elektronen
m	effektive Masse eines Elektrons
$\rho = mn$	Dichte der Elektronen
v	mittlere Geschwindigkeit der Elektronen
v_s	Sättigungsgeschwindigkeit
T	Temperatur der Elektronen
T_0	Gittertemperatur
$P = n \frac{k_B T}{m}$	Druck
γ	Isentropen-Exponent
$e_I = \frac{1}{\gamma-1} \frac{k_B T}{m}$	innere Energie
C_n, C_p, C_w	Kollisionsterme
τ_p, τ_w	Relaxationszeiten
μ_0	Beweglichkeit der Elektronen
$p = \rho v = mnv$	Impulsdichte der Elektronen
$w = \rho e_I + \frac{1}{2} \rho v^2 = \rho e_I + \frac{1}{2} p v$	Energiedichte der Elektronen

Literaturverzeichnis

- [1] A. Arnold, *On Absorbing Boundary Conditions for Quantum Transport Equations*, erscheint in: Math. Meth. Num. Anal. (1994).
- [2] A. Arnold, *private Kommunikation*.
- [3] A. Bayliss und E. Turkel, *Radiation Boundary Conditions for Wave-like Equations*, Comm. Pure Appl. Math. **33** (1980), 707-725.
- [4] J. G. Blaschak und G. A. Kriegsmann, *A Comparative Study of Absorbing Boundary Conditions*, J. Comput. Physics **77** (1988), 109-139.
- [5] D. M. Boore, *Finite Difference Methods for Seismic Wave Propagation in Heterogeneous Materials*, Methods of Comp. Physics (Seismology) **11** (1972), 1-37.
- [6] R. Clayton und B. Engquist, *Absorbing Boundary Conditions for Acoustic and Elastic Wave Calculations*, Bull. Seismological Sec. Amer. **67** (1977), 1529-1540.
- [7] G. Doetsch, *Einführung in die Theorie und Anwendung der Laplace-Transformation*, Birkhäuser Verlag, Basel, 1958.
- [8] G. Doetsch, *Anleitung zum praktischen Gebrauch der Laplace-Transformation und der Z-Transformation*, R. Oldenburg Verlag München, Wien, 3. Auflage 1967.
- [9] T. Elvius und A. Sundström, *Computationally efficient schemes and boundary conditions for a fine-mesh barotropic model based on the shallow-water equations*, Tellus **25** (1973), 132-156.
- [10] B. Engquist und A. Majda, *Absorbing Boundary Conditions for the Numerical Simulation of Waves*, Math. Comput. **31** (1977), 629-651.
- [11] E. Fatemi, J. W. Jerome und S. Osher, *Solution of the Hydrodynamic Device Model Using High-Order Nonoscillatory Shock Capturing Algorithms*, IEEE Trans. Computer-Aided Design **10** (1991), 232-243.
- [12] C. L. Gardner, *Numerical Simulation of a Steady-State Electron Shock Wave in a Submicron Semiconductor Device*, IEEE Trans. Electron Devices **38** (1991), 392-398.
- [13] C. L. Gardner, J. W. Jerome und D. J. Rose, *Numerical Methods for the Hydrodynamic Device Model: Subsonic Flow*, IEEE Trans. Computer-Aided Design **8** (1989), 501-507.
- [14] M. Goldberg und E. Tadmor, *Scheme-Independent Stability Criteria for Difference Approximations of Hyperbolic Initial-Boundary Value Problems. II*, Math. Comput. **36** (1981), 603-626.
- [15] M. Goldberg und E. Tadmor, *Convenient Stability Criteria for Difference Approximations of Hyperbolic Initial-Boundary Value Problems II*, Math. Comput. **48** (1987), 503-520.
- [16] C. E. Grosch und S. A. Orszag, *Numerical Solution of Problems in Unbounded Regions: Coordinate Transforms*, J. Comput. Physics **25** (1977), 273-296.
- [17] B. Gustafsson und H.-O. Kreiss, *Boundary Conditions for Time Dependent Problems with an Artificial Boundary*, J. Comput. Physics **30** (1979), 333-351.
- [18] B. Gustafsson, H.-O. Kreiss und A. Sundström, *Stability Theory for Difference Approximations of Mixed Initial Boundary Value Problems. II*, Math. Comp. **26** (1972), 649-686.
- [19] R. L. Higdon, *Initial Boundary Value Problems for Linear Hyperbolic Systems*, SIAM Review **28** (1986), 177-217.
- [20] F. John, *Partial Differential Equations*, Springer-Verlag, New York, 1982.

- [21] K. R. Kelly, R. M. Alford, S. Treitel und R. W. Ward, *Applications of Finite Difference Methods to Exploration Seismology*, Proc. Roy. Irish Acad. Conf. on Numerical Analysis, Academic Press, London and New York (1974), 57-76.
- [22] H.-O. Kreiss, *Stability Theory for Difference Approximations of Mixed Initial Boundary Value Problems. I*, Math. Comp. **22** (1968), 703-714.
- [23] H.-O. Kreiss und J. Lorenz, *Initial-Boundary Value Problems and the Navier-Stokes Equations*, Academic Press, 1989.
- [24] P. D. Lax und B. Wendroff, *Difference Schemes for Hyperbolic Equations with High Order of Accuracy*, Comm. Pure Appl. Math. **17** (1964), 381-398.
- [25] R. J. LeVeque, *Numerical Methods for Conservation Laws*, Birkhäuser Verlag, Basel, 1990.
- [26] G. Lill, *Diskrete Randbedingungen an künstlichen Rändern*, Dissertation an der Technischen Universität Darmstadt, 1992.
- [27] P. Markowich, C. Ringhofer und C. Schmeiser, *Semiconductor Equations*, Springer-Verlag, Wien, New York, 1990.
- [28] G. R. McGuire und J. L. Morris, *A Class of Second-Order Accurate Methods for the Solution of Systems of Conservation Laws*, J. Comput. Physics **11** (1973), 531-549.
- [29] A. R. Mitchell und D. F. Griffiths, *The Finite Difference Method in Partial Differential Equations*, Wiley, New York, 1980.
- [30] A. C. Reynolds, *Boundary Conditions for the Numerical Solution of Wave Propagation Problems*, Geophysics **63** (1978), 1099-1110.
- [31] R. D. Richtmyer und K. W. Morton, *Difference Methods for Initial-Value Problems*, Wiley & Sons, New York, 2. Auflage 1967.
- [32] P. J. Roache, *Computational Fluid Dynamics*, Hermosa Press, Albuquerque, N. M., 1972.
- [33] E. L. Rubin und S. Z. Burstein, *Difference Methods for the Inviscid and Viscous Equations of a Compressible Gas*, J. Comput. Physics **2** (1967), 178-196.
- [34] C. Schmeiser, *Numerische Methoden der Strömungsmechanik*, Lernbehelf Technische Universität Wien.
- [35] D.-W. Sheen, *Absorbing Boundary Conditions for Wave Transmissions*, Technical Report **159** (1991), Purdue University, West Lafayette, IN 47907.
- [36] A. Shubin, *Pseudodifferential Operators and Spectral Theory*, Springer-Verlag, Berlin, New York, 1986.
- [37] G. Sod, *Numerical Methods for Fluid Dynamics*, Cambridge University Press, 1985.
- [38] J. C. Strikwerda, *Finite Difference Schemes and Partial Differential Equations*, Wadsworth & Brooks/Cole, 1989.
- [39] M. E. Taylor, *Reflection of Singularities of Solutions to Systems of Differential Equations*, Comm. Pure Appl. Math. **27** (1975), 457-478.
- [40] M. E. Taylor, *Pseudodifferential Operators*, Princeton University Press, Princeton N.J., 1981.
- [41] G. B. Whitham, *Linear and Nonlinear Waves*, Wiley-Interscience, New York, 1974.
- [42] G. Zwas, *On Two Step Lax-Wendroff Methods in Several Dimensions*, Numer. Math. **20** (1973), 350-355.