

Pia Ruth LANGOTH

Modellierung des Grundwasserflusses in porösen Medien

DIPLOMARBEIT

zur Erlangung des akademischen Grades einer Diplom-Ingenieurin

Diplomstudium Technische Mathematik



Graz University of Technology

Technische Universität Graz

Betreuer:

Univ.-Prof. Dr. Olaf STEINBACH

Institut für Numerische Mathematik

Graz, im Juni 2011

EIDESSTATTLICHE ERKLÄRUNG

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Graz, am

.....

(Unterschrift)

Vorwort

Die vorliegende Diplomarbeit zum Thema „Modellierung des Grundwasserflusses in porösen Medien“ ist im Rahmen meines Studiums der Technischen Mathematik an der TU Graz entstanden.

Mein Dank gilt Herrn Professor Steinbach für die Bereitstellung dieses interessanten Themas, die aufgebrauchte Zeit und Geduld sowie die großartige Betreuung während der Verfassung dieser Arbeit.

Besonders bedanken möchte ich mich bei meiner Familie, insbesondere bei meiner Mutter, Frau Sigrid Langoth, deren Zuspruch und Unterstützung eine große Bereicherung und Hilfe während meiner gesamten Studienzeit waren.

Weiters gilt mein Dank meinem besten Freund und Partner Martin Weindlmayr, dessen Verständnis und Rückhalt grenzenlos waren.

Ich widme diese Arbeit meinem verstorbenen Vater, Herrn Ing. Hansjörg Langoth, der in mir die Freude an der Mathematik geweckt und damit meinen Lebensweg geprägt hat.

Inhaltsverzeichnis

Einleitung	9
1 Modellierung	13
1.1 Herleitung der Richards-Gleichung	13
1.2 Brooks-Corey-Parameterfunktionen	16
1.3 Kirchhoff-Transformation	17
1.4 Randwertprobleme für die Richards-Gleichung	18
2 Analysis	23
2.1 Sobolev-Räume	23
2.2 Herleitung der Variationsungleichung	28
2.3 Implizit-explicite Zeitdiskretisierung	37
3 Konvexe Minimierung	39
3.1 Eigenschaften konvexer Mengen und konvexer Funktionale	39
3.2 Konvexe Minimierungsprobleme	40
3.3 Eindeutige Lösbarkeit und Äquivalenz	45
4 Finite-Elemente-Methode	47
4.1 Diskretisierung	47
4.2 Konvergenz	52
5 Dirichlet-Neumann- und Robin-Methoden für nichtlineare Transmissionsprobleme	59
5.1 Allgemeine Problemstellung	59
5.2 Steklov-Poincaré-Operator-Formulierung	61
5.3 Nichtlineare Dirichlet-Neumann-Methode	63
5.4 Nichtlineare Robin-Methode	65
5.5 Konvergenz der Lösungsverfahren	67
6 Numerisches Beispiel	71
6.1 Modellproblem	71
6.2 Herleitung der Variationsformulierung	72
6.3 Aktive-Mengen-Strategie	73
6.4 Beispiel und Resultate	76

Einleitung

Die Grundlage dieser Arbeit über die Modellierung des Grundwasserflusses in porösen Medien bildet die Dissertation von H. Berninger [6], deren Kernstück die numerische Analyse der Richards-Gleichung ist. Diese dient als Modell für den gesättigten-ungesättigten Grundwasserfluss und wurde erstmals in der Dissertation von L. A. Richards [30] veröffentlicht. Die Richards-Gleichung kann durch Kombination zweier fundamentaler physikalischer Gesetze hergeleitet werden und zwar einerseits aus dem experimentellen Gesetz von Darcy, das als spezielle Lösung der Navier-Stokes-Gleichung die Grundlage für das Strömungsverhalten eines Fluids in porösen Medien bildet, und andererseits aus der Kontinuitätsgleichung, die den Massenerhalt eines Fluids beschreibt, siehe [5]. Für einen maximal gesättigten Boden, d.h. der Boden kann kein weiteres Wasser mehr aufnehmen, folgt, dass die Richards-Gleichung eine elliptische partielle Differentialgleichung ist, wohingegen die Richards-Gleichung für den ungesättigten Fall, d.h. für einen teilweise gesättigten Boden, parabolisch ist. Die Richards-Gleichung enthält außerdem zwei Nichtlinearitäten, wobei sich eine davon auf die Zeitableitung bezieht und die andere ein Faktor in der Ortsableitung ist.

Die Beziehungen zwischen der Sättigung und dem Druck bzw. zwischen der relativen Durchlässigkeit und der Sättigung werden aus physikalischen Experimenten bestimmt. Eine Möglichkeit der Wahl dieser Parameterfunktionen sind die auf van Genuchten zurückzuführenden Parameterfunktionen (siehe [34]), oder die Brooks-Corey-Funktionen, die in dieser Arbeit in Anlehnung an [6] verwendet werden.

Die Kirchhoff-Transformation ist ein nützliches Hilfsmittel, um Klassen von partiellen Differentialgleichungen zu vereinfachen, indem bestimmte Nichtlinearitäten beseitigt werden, siehe [3]. Im Falle der Richards-Gleichung in einem homogenen Boden kann durch Anwendung einer Kirchhoff-Transformation eine der beiden auftretenden Nichtlinearitäten eliminiert werden und es ergibt sich eine semilineare, elliptisch-parabolische transformierte Gleichung, siehe [6, 8].

Das Randwertproblem für die Richards-Gleichung bzw. für die transformierte Richards-Gleichung beinhaltet neben einer Dirichlet- und einer Neumann-Randbedingung auch eine Ungleichungsrandbedingung, die sogenannte Signorini-Randbedingung. In der Hydrologie treten Signorini-Randbedingungen üblicherweise bei Oberflächenwasserreservoirs oder bei sogenannten „Dammproblemen“ auf. Diese speziellen Randwertprobleme können unter Anwendung einer Verallgemeinerung der Green'schen Formel durch eine schwache Formulierung in Form einer Variationsungleichung beschrieben werden, siehe [24, 6]. Unter Durchführung der Zeitdiskretisierung der schwachen Variationsungleichung der transformierten Gleichung folgt die Äquivalenz der diskreten Variationsungleichung zu einem konvexen Minimierungsproblem. Die Vorgehensweise ist dabei dieselbe wie in [6] mit Verwendung von

Einleitung

Definitionen aus [25]. Das Äquivalenzresultat ergibt sich schließlich mit [15]. Die eindeutige Lösbarkeit der diskreten Variationsungleichung bzw. des konvexen Minimierungsproblems ergibt sich durch Anwendung von allgemeinen Existenz- und Eindeutigkeitsaussagen für konvexe Minimierungsprobleme in reflexiven Banachräumen, siehe [15].

Die Diskretisierung des Minimierungsproblems mit finiten Elementen wird unter der Annahme einer zulässigen Triangulierung und der Wahl des Ansatzraumes der stückweise linearen Basisfunktionen wie in [6] durchgeführt. Die eindeutige Lösbarkeit des Minimierungsproblems auf dem stetigen Level wird dabei auf das diskrete Level übertragen. Die Konvergenzanalyse wird unter der Annahme durchgeführt, dass die Maschenweiten für die zugehörige Folge von Triangulierungen abnehmen und dass die Folge von Triangulierungen formregulär ist. Die Vorgehensweise ist ähnlich wie in [6, 8] unter Anwendung von [17, 13]. Zunächst folgt damit die Konvergenz der Approximation des verallgemeinerten Drucks, siehe [6, 25]. Aus praktischer Sicht interessanter ist jedoch die Konvergenz für die Approximationen der Näherungslösungen des physikalischen Drucks und der Sättigung. Dazu werden grundlegende Eigenschaften von Superpositionsoperatoren aus [4] angewendet, die in [6] wiederholt werden.

Die Betrachtung der Richards-Gleichung in einem heterogenen Boden, d.h. dass die Bodenparameter in nicht überlappenden Teilgebieten Ω_i , $i = 1, 2$, des Gebietes $\Omega \subset \mathbb{R}^d$ konstant sind und sich unstetig entlang des Interfaces der beiden Teilgebiete ändern, stellt eine Verallgemeinerung des homogenen Falles dar. Als resultierendes Problem ergibt sich ein quasilineares, elliptisches Transmissionsproblem, bei dem sich die Nichtlinearität unstetig zwischen zwei Teilgebieten verändert, siehe [6, 9]. Das Gebietszerlegungsproblem, das aus der Kopplung der Richards-Gleichung im homogenen Boden auf jedem Teilgebiet mit der Stetigkeit des physikalischen Drucks und der des Normalenflusses folgt, wird analog zu [9] aufgestellt. Um die bereits erzielten Ergebnisse für den homogenen Fall auch auf die Richards-Gleichung in einem heterogenen Boden anwenden zu können, wird auf beiden Teilgebieten eine Kirchhoff-Transformation durchgeführt und ein Gebietszerlegungsproblem für den verallgemeinerten Druck formuliert. Für dieses Gebietszerlegungsproblem gilt, dass das Problem auf beiden Teilgebieten linear und die Nichtlinearität vollständig in der Interface-Gleichung enthalten ist. Diese Folgeerscheinung der Kirchhoff-Transformation macht einerseits die Analysis und andererseits die numerische Behandlung des Transmissionsproblems einfacher. Die Herleitung schwacher Formulierungen der Gebietszerlegungsprobleme erfolgt aus einer Verallgemeinerung der Ergebnisse für den linearen Fall (siehe [29]) und wird wie in [9] durchgeführt. Die Äquivalenz dieser schwachen Formulierungen für den physikalischen bzw. für den verallgemeinerten Druck folgt aus [7]. Die Darstellung der schwachen Formulierungen der Gebietszerlegungsprobleme mit Steklov-Poincaré-Operatoren und der Beweis der Äquivalenz der schwachen Formulierungen und der Steklov-Poincaré-Gleichungen werden wie in [6, 9] durchgeführt und ergeben sich wiederum aus einer Verallgemeinerung der linearen Theorie. Um geeignete iterative Lösungsverfahren für die Gebietszerlegungsprobleme zu erhalten, werden bereits bekannte lineare Verfahren erweitert, sodass schließlich ein nichtlinearer Dirichlet-Neumann- bzw. ein nichtlinearer Robin-Algorithmus, die ohne jegliche Linearisierung auskommen, angegeben werden können, siehe [9, 6, 29]. Unter Anwendung der Steklov-Poincaré-Operatoren wie in [6] folgt, dass der gedämpfte Dirichlet-

Neumann-Algorithmus ein vorkonditioniertes Richardson-Verfahren bzw. dass der nicht-lineare Robin-Algorithmus äquivalent zur iterativen Methode alternierender Richtungen (ADI-Methode) ist. Die Konvergenzanalyse der beiden Methoden wird durchgeführt, indem Aussagen, die für den linearen Fall gelten, erweitert und Bedingungen gefordert werden, sodass der Banachsche Fixpunktsatz angewendet werden kann, siehe [14, 9]. Damit folgen Konvergenzaussagen für im Ort eindimensionale Diskretisierungen, die sowohl auf dem stetigen als auch auf dem diskreten Level gelten. Weiters zeigen numerische Beispiele in 2D Konvergenz mit realistischen Konvergenzraten für eine geeignete Wahl der Parameter der Dirichlet-Neumann- bzw. Robin-Methode. Darüber hinaus zeigt sich Netzunabhängigkeit der Konvergenzraten und Parameter für die Dirichlet-Neumann-Methode. Da nicht bekannt ist, ob der Operator aus der Robin-Methode eine Kontraktion bildet, können für das Robin-Verfahren weder Konvergenzraten auf dem stetigen Level angegeben, noch die Netzunabhängigkeit der optimalen Konvergenzraten und Parameter auf dem diskreten Level gefolgert werden, siehe [9, 6].

Die numerische Behandlung eines Signorini-Randwertproblems für die Poisson-Gleichung als Vereinfachung des Randwertproblems für die Richards-Gleichung bildet schließlich den Abschluss dieser Arbeit. Als Modellgebiet wird dabei das Einheitsquadrat im \mathbb{R}^2 betrachtet. Die Diskretisierung wird unter der Annahme durchgeführt, dass die Folge von Unterteilungen in finite Dreieckselemente zulässig sei und dass die Dreieckselemente wiederum formregulär seien. Die Implementierung der Steifigkeitsmatrizen sowie das Einbauen der Dirichlet- und der Neumann-Randbedingung erfolgen dabei ähnlich wie in [23]. Für die Berücksichtigung der Signorini-Randbedingung ist der Einsatz der Aktive-Mengen-Strategie notwendig, siehe dazu [20] bzw. [22]. Die Fehlerberechnung der Näherungslösung in der L_2 -Norm wird unter der Annahme, dass die exakte Lösung bekannt ist und unter Anwendung einer 7-Punkt-Gauss-Integrationsformel für Dreiecksgebiete durchgeführt. Mit der Theorie übereinstimmend ergibt sich quadratische Konvergenz in der L_2 -Norm.

Ziel dieser Arbeit ist zunächst die Herleitung von Randwertproblemen für die Richards-Gleichung in homogenen Böden, die neben Dirichlet- und Neumann-Randbedingungen auch eine Signorini-Randbedingung berücksichtigen. Für diese Randwertprobleme sollen in der Folge Variationsformulierungen in Form von Variationsungleichungen auf konvexen Mengen angegeben und die Äquivalenz dieser Probleme zu einem konvexen Minimierungsproblem gezeigt werden. Unter Durchführung einer geeigneten Diskretisierung soll die Konvergenz des verallgemeinerten bzw. des physikalischen Drucks und der Sättigung unter bestimmten Voraussetzungen gezeigt werden. Durch Verallgemeinerung der linearen Theorie sollen schließlich geeignete iterative Verfahren zur Lösung der resultierenden Gebietszerlegungsprobleme hergeleitet und bezüglich ihrer Konvergenz untersucht werden. An einem konkreten Modellbeispiel soll abschließend gezeigt werden, wie eine Signorini-Randbedingung mittels einer Aktive-Mengen-Strategie numerisch behandelt werden kann.

Die vorliegende Arbeit ist wie folgt gegliedert:

Im ersten Kapitel soll die Richards-Gleichung zunächst durch Kombination des experimentellen Gesetzes von Darcy und der Kontinuitätsgleichung hergeleitet werden. Im An-

Einleitung

schluss werden die Brooks-Corey-Parameterfunktionen, die den Zusammenhang zwischen Sättigung und Druck bzw. zwischen relativer Durchlässigkeit und Sättigung beschreiben, angegeben. Um eine der beiden Nichtlinearitäten aus der Richards-Gleichung zu eliminieren, wird im Folgenden eine Kirchhoff-Transformation durchgeführt, die durch Anwendung der Kettenregel eine transformierte Version der Richards-Gleichung liefert. Den Abschluss dieses Kapitels bilden schließlich die Herleitung von Randwertproblemen sowohl für die Richards-Gleichung als auch für die zugehörige transformierte Gleichung und die physikalische Interpretation der Signorini-Randbedingung mithilfe eines Beispiels.

Im zweiten Kapitel werden eingangs die zugrundeliegenden Funktionenräume, die Sobolev-Räume, definiert und wichtige, in weiterer Folge oft verwendete Eigenschaften und Aussagen angegeben. Im Anschluss werden schwache Formulierungen der Variationsungleichung des Randwertproblems für die Richards-Gleichung bzw. für die zugehörige Kirchhoff-transformierte Gleichung auf konvexen Mengen betrachtet, die anschließend hinsichtlich ihrer Äquivalenz untersucht werden. In der Folge wird eine implizit-explicite Zeitdiskretisierung durchgeführt, die schlussendlich das Angeben einer diskreten Variationsungleichung ermöglicht.

Das dritte Kapitel widmet sich der konvexen Minimierung. Es wird gezeigt, dass die diskrete Variationsungleichung des Randwertproblems für die Kirchhoff-transformierte Richards-Gleichung äquivalent zu einem konvexen Minimierungsproblem ist und dass diese ferner durch Anwendung von allgemeinen Existenz- und Eindeutigkeitsaussagen für konvexe Minimierungsprobleme in reflexiven Banachräumen eine eindeutig Lösung besitzt.

Im vierten Kapitel wird zuerst eine geeignete Diskretisierung des konvexen Minimierungsproblems mit finiten Elementen angegeben, dessen eindeutige Lösbarkeit aus jener des konvexen Minimierungsproblems gefolgert werden kann. Im Anschluss daran wird das resultierende diskrete Minimierungsproblem für eine diskrete physikalische Variable umformuliert. Den Abschluss des Kapitels bildet einerseits die Konvergenzanalyse für den verallgemeinerten Druck und andererseits, die aus praktischer Sicht interessantere, Konvergenzanalyse für den physikalischen Druck und der Sättigung.

Im fünften Kapitel wird ein quasilineares, elliptisches Transmissionsproblem betrachtet, bei dem sich die Nichtlinearität unstetig zwischen zwei Teilgebieten verändert. Es wird zunächst ein Gebietszerlegungsproblem und dessen schwache Formulierung für den verallgemeinerten Druck angegeben. Anschließend wird eine Formulierung mit Steklov-Poincaré-Operatoren hergeleitet. In der Folge wird näher auf die nichtlineare Dirichlet-Neumann- und Robin-Methode eingegangen, deren Konvergenz schließlich am Ende des Kapitels analysiert wird.

Das sechste Kapitel beschäftigt sich mit der numerischen Behandlung eines Signorini-Randwertproblems für die Poisson-Gleichung. Dazu werden zunächst das zugrundeliegende Modellproblem und die zugehörige Diskretisierung mit finiten Dreieckselementen betrachtet. Im Anschluss wird eine diskrete Variationsformulierung in Form einer Variationsungleichung hergeleitet. Mit der Aktive-Mengen-Strategie wird danach eine Methode vorgestellt, mit der Signorini-Randbedingungen numerisch behandelt werden können. Den Abschluss des Kapitels bildet schließlich die Betrachtung eines konkreten Modellproblems, dessen Implementierung sowie die L_2 -Fehlerabschätzung der berechneten Näherungslösung.

1 Modellierung

In diesem Kapitel soll zunächst gezeigt werden, wie die Richards-Gleichung aus fundamentalen physikalischen Gesetzen hergeleitet werden kann. Anschließend werden konkrete analytische Versionen der benötigten Parameterfunktionen angegeben. Danach wird in Abschnitt 1.3 eine Kirchhoff-Transformation durchgeführt, um eine der beiden Nichtlinearitäten aus der Richards-Gleichung zu eliminieren. Zum Abschluss werden in Abschnitt 1.4 noch Randwertprobleme für die Richards-Gleichung angeführt und das Kapitel wird schließlich durch die Angabe eines praktischen Beispiels vervollständigt.

1.1 Herleitung der Richards-Gleichung

Ausgangspunkt für die Herleitung der Richards-Gleichung ist das Prinzip der Massenerhaltung für ein Kontrollvolumen $\Omega'(t)$. Dieses besagt, dass die zeitliche Änderung der Masse des Fluids in $\Omega'(t)$ zu einer bestimmten Zeit t gleich dem Massenfluss durch den Rand $\partial\Omega'(t)$ und einem eventuell vorhandenen Quellen- oder Senkenterm ist:

$$\int_{\Omega'(t)} \frac{\partial}{\partial t} (\theta(p(x, t)) n(x) \rho) dx + \int_{\partial\Omega'(t)} \theta(p(x, t)) n(x) \rho \bar{\mathbf{v}}(x, t) \cdot \mathbf{n}(x, t) ds_x = \int_{\Omega'(t)} \bar{f}(x, t) dx \quad (1.1)$$

mit

$\rho \dots$ Wasserdichte,	$\bar{f} \dots$ Quellenterm,
$n \dots$ Porosität,	$\bar{\mathbf{v}} \dots$ mikroskopische Geschwindigkeit,
$\theta \dots$ Sättigung.	

Es wird berücksichtigt, dass Wasser nicht zusammendrückbar ist, d.h. $\rho = \text{const.}$, dass die makroskopische Geschwindigkeit des Wassers $\mathbf{v}(x, t)$ durch $\mathbf{v}(x, t) = \theta(p(x, t)) n(x) \bar{\mathbf{v}}(x, t)$ gegeben ist und $f(x, t) = \bar{f}(x, t)/\rho$ gilt. Weiters sei die Porosität $n = n(x)$ nur vom Ort, aber nicht von der Zeit abhängig.

Um weitere Umformungen vornehmen zu können, wird der nachfolgende Satz von Gauss benötigt, der einen Zusammenhang zwischen einem Vektorfeld und dessen Divergenz angibt.

1 Modellierung

Satz 1.1 (Satz von Gauss). Sei $\Omega \subset \mathbb{R}^3$ ein beschränktes Lipschitz-Gebiet, d.h., dass der Rand $\Gamma = \partial\Omega$ bezüglich einer beliebigen Zerlegung stückweise durch eine Lipschitz-stetige Funktion dargestellt werden kann, und $F : \overline{\Omega} \rightarrow \mathbb{R}^3$ sei ein Vektorfeld mit den folgenden Eigenschaften:

1. F ist stetig auf $\overline{\Omega}$ und stetig differenzierbar auf Ω .
2. $\operatorname{div} F$ ist integrierbar auf Ω .
3. F ist integrierbar auf Ω .

Dann gilt:

$$\int_{\Omega} \operatorname{div} F \, dx = \int_{\partial\Omega} F \cdot \mathbf{n} \, ds_x.$$

Für den Beweis siehe zum Beispiel [24].

Unter Anwendung des Gauss'schen Satzes auf die Gleichung (1.1) und Division durch die konstante Dichte ρ ergibt sich

$$\int_{\Omega'(t)} \left(n(x) \frac{\partial}{\partial t} \theta(p(x, t)) + \operatorname{div} \mathbf{v}(x, t) \right) dx = \int_{\Omega'(t)} f(x, t) \, dx. \quad (1.2)$$

Mit der geforderten Stetigkeit der zu integrierenden Funktionen folgt schließlich die Kontinuitätsgleichung

$$n(x) \theta(p(x, t))_t + \operatorname{div} \mathbf{v}(x, t) = f(x, t). \quad (1.3)$$

Das experimentelle Gesetz von Darcy setzt den Wasserfluss zu jeder Zeit t folgendermaßen mit dem Wasserdruck in Beziehung:

$$\mathbf{v}(x, t) = -K_c(x, \theta) \nabla h(x). \quad (1.4)$$

K_c bezeichnet hierbei die hydraulische Leitfähigkeit und h die Piezometerhöhe, d.h. das Grundwasserlevel im Punkt $x = (x_1, x_2, z) \in \Omega \subset \mathbb{R}^3$, siehe [5, 30].

Der Druck ergibt sich aus dem Pascal'schen Gesetz

$$p(x, t) = \rho g (h(x) + z),$$

mit der Erdschwerebeschleunigung g und der Grundwasserhöhe z .

1.1 Herleitung der Richards-Gleichung

Für die mit dem Wasserdruck in Beziehung stehende Piezometerhöhe h folgt aus dem Pascal'schen Gesetz

$$h(x) = \frac{p(x, t)}{\rho g} - z. \quad (1.5)$$

Die z -Achse des Koordinatensystems wird hier, wie so oft in der Literatur, in Richtung Gravitation und gegen die Richtung der Piezometerhöhe angenommen, dazu siehe zum Beispiel [18, 12, 16].

Der erste Term auf der rechten Seite der Gleichung (1.5) wird weiters als Druckhöhe bezeichnet. Diese ergibt sich für $p > 0$ aus dem hydrostatischen Druck und für $p < 0$ aus dem Kapillardruck. Im ersten Fall befindet sich die Grundwasserhöhe über dem Punkt x , im zweiten Fall darunter.

Für die Definition der hydraulischen Leitfähigkeit $K_c(\cdot)$ wird im Folgenden zwischen vollkommen und nur teilweise gesättigten Böden unterschieden.

Ist die Sättigung mit Wasser im gesamten Medium maximal, so ist die hydraulische Leitfähigkeit nur vom Ort $x \in \Omega$ abhängig und es gilt der folgende Zusammenhang:

$$K_c(x) = K_h(x) = K(x)\mu^{-1}\rho g \quad \forall x \in \Omega, \quad (1.6)$$

mit der Viskosität des Wassers μ .

Die Funktion $K(\cdot)$ ist im Gegensatz zur hydraulischen Leitfähigkeit K_c , die maßgeblich von Dichte und Viskosität des durchströmenden Fluids beeinflusst wird, nicht mehr länger vom Fluid abhängig und wird nunmehr als Permeabilität des Bodens bezeichnet. $K_h(\cdot)$ kann als maximale hydraulische Leitfähigkeit betrachtet werden.

Tatsächlich bleibt das Gesetz von Darcy (1.4) auch für den allgemeineren Fall eines ungesättigten Bodens gültig. Jedoch ist zu beachten, dass in diesem Fall die hydraulische Leitfähigkeit $K_c(\cdot)$ auch von der Sättigung θ abhängt. Es gilt dann:

$$K_c(x, \theta) = K_h(x)k_r(\theta). \quad (1.7)$$

Dabei bezeichnet $k_r(\cdot)$ die relative und das Produkt $K_h(x)k_r(\theta)$ die effektive Permeabilität. Es kann also beobachtet werden, dass im Falle einer maximalen Sättigung $k_r(\theta) = 1$ gilt. Durch Einsetzen der Bedingung (1.6) in (1.7) folgt schließlich für die hydraulische Leitfähigkeit in teilweise gesättigten Böden:

$$K_c(x, \theta) = K(x)\mu^{-1}\rho g k_r(\theta).$$

Die relative Permeabilität $k_r(\cdot)$ ist eine monoton wachsende Funktion im Intervall $[0, 1]$ abhängig von der Sättigung θ . Zu bemerken bleibt, dass die konkrete Form der Funktion $k_r(\cdot)$ vom Boden abhängt.

Wiederum abhängig vom Boden ergibt sich eine weitere Zustandsgleichung, die die Sättigung θ mit dem Druck p in Beziehung setzt.

1 Modellierung

Die Funktion $p \mapsto \theta(p)$ ist ebenfalls monoton wachsend zwischen einer minimalen Sättigung $\theta_m > 0$ (aufgrund von Restwasser im Boden) und einer maximalen Sättigung $\theta_M < 1$ (aufgrund von Restluft im Boden).

Durch Einsetzen der soeben hergeleiteten Beziehungen in die Gleichung (1.3) und der Annahme $f = 0$ folgt die Richards-Gleichung in folgender Form:

Gesucht ist $p(x, t)$ mit $(x, t) \in \Omega \times (0, T)$ und $T > 0$, sodass

$$n(x)\theta(p(x, t))_t - \operatorname{div}(K(x)\mu^{-1}k_r(\theta(p(x, t)))\nabla(p(x, t) - \rho gz)) = 0. \quad (1.8)$$

Offensichtlich ist die Richards-Gleichung für den Fall eines maximal gesättigten Bodens vom elliptischen Typ, da in diesem Fall die Sättigung θ konstant ist und nicht mehr länger vom Druck p abhängt, wodurch die Zeitableitung wegfällt.

Für den ungesättigten Fall bleibt die Zeitableitung erhalten und die auftretende Differentialgleichung ist vom parabolischen Typ.

Die Richards-Gleichung enthält außerdem zwei Nichtlinearitäten, wobei sich eine davon auf die Zeitableitung bezieht und die andere ein Faktor in der Ortsableitung ist.

Von nun an werde $k_r(\theta(p))$ stets als positiv vorausgesetzt, da ansonsten für $k_r(\theta(p)) = 0$ die Richards-Gleichung vom hyperbolischen Typ wäre.

1.2 Brooks-Corey-Parameterfunktionen

Die Beziehungen zwischen der Sättigung θ und dem Druck p bzw. zwischen der relativen Durchlässigkeit k_r und der Sättigung θ sind aus physikalischen Experimenten bekannt.

Es gibt verschiedene Methoden, um konkrete analytische Versionen dieser Parameterfunktionen $p \mapsto \theta(p)$ und $\theta \mapsto k_r(\theta)$ zu erhalten. Eine mögliche Wahl wären zum Beispiel die auf van Genuchten zurückzuführenden Parameterfunktionen, siehe [34], oder auch die Brooks-Corey-Funktionen, die in dieser Arbeit in Anlehnung an [6] verwendet werden.

Mit der Wahl der Brooks-Corey-Parameterfunktionen ergibt sich für die sogenannte „Retentionskurve“

$$\Theta(p) := \frac{\theta(p) - \theta_m}{\theta_M - \theta_m} = \left[\frac{p}{p_b} \right]^{-\lambda} := \begin{cases} \left(\frac{p}{p_b} \right)^{-\lambda} & \text{für } p \leq p_b, \\ 1 & \text{für } p \geq p_b \end{cases}$$

mit dem von der Art des Bodens abhängigen Porengrößenverteilungsfaktor $\lambda > 0$ und dem Lufteintrittspunkt $p_b < 0$, in dem die Brooks-Corey-Funktionen nicht differenzierbar sind. Für eine weiterführende Theorie zu solchen Parameterfunktionen sei auf [6] verwiesen. Insgesamt ergibt sich für die relative Durchlässigkeit $k_r(\cdot)$ mit $\Theta \in (0, 1]$ oder $\theta \in (\theta_m, \theta_M]$:

$$k_r(\theta) = \widehat{k}_r(\Theta) := \Theta^{3 + \frac{2}{\lambda}}.$$

Durch Einsetzen folgt

$$k_r(\theta(p)) = \widehat{k}_r(\Theta(p)) = \left[\frac{p}{p_b} \right]^{-3\lambda-2}.$$

1.3 Kirchhoff-Transformation

Wie bereits bekannt, liegt der Zweck einer Kirchhoff-Transformation darin, bestimmte Nichtlinearitäten zu beseitigen. Für eine allgemeinere Theorie zur Kirchhoff-Transformation siehe [3]. Im speziellen Fall der Richards-Gleichung soll die relative Durchlässigkeit $k_r(\cdot)$ in Gleichung (1.8) eliminiert werden.

Die Kirchhoff-Transformation $\kappa : \mathbb{R} \rightarrow \mathbb{R}$ für die Richards-Gleichung ist definiert als

$$\kappa : p \rightarrow u := \int_0^p k_r(\theta(q)) \, dq. \quad (1.9)$$

Die neue bzw. transformierte Variable u wird als verallgemeinerter Druck bezeichnet. Für die Sättigung als Funktion von u ergibt sich dann

$$M(u) := \theta(\kappa^{-1}(u)). \quad (1.10)$$

Bemerkung 1.1. *Für die Existenz der Umkehrabbildung*

$$\kappa^{-1}(u) = \int_0^u \frac{1}{k_r(M(s))} \, ds \quad \forall u \in \mathbb{R}$$

ist die Bedingung $k_r \geq c > 0$ hinreichend.

Unter Anwendung der Kettenregel gilt

$$\nabla u = \nabla(\kappa(p)) = \kappa'(p)\nabla p = k_r(\theta(p)) \nabla p. \quad (1.11)$$

Durch Einsetzen der Kirchhoff-Transformation (1.9) in die Richards-Gleichung (1.8), folgt mit (1.11) die transformierte Richards-Gleichung in der Form:

Gesucht ist $u(x, t)$ für $(x, t) \in \Omega \times (0, T)$ mit $T > 0$, sodass

$$n(x)M(u(x, t))_t - \operatorname{div}(K(x)\mu^{-1}(\nabla u(x, t) - k_r(M(u(x, t))))\rho g e_z) = 0, \quad (1.12)$$

wobei e_z den Einheitsvektor in Richtung der Gravitation angibt.

1 Modellierung

Diese transformierte Gleichung ist semilinear und die Nichtlinearität vor der Ortsableitung wurde eliminiert.

Bemerkung 1.2. *Ein Vorteil der Wahl der Brooks-Corey-Parameterfunktionen ist, dass sowohl die Kirchhoff-Transformation κ und ihre Inverse als auch die verallgemeinerte Sättigung M und die relative Durchlässigkeit k_r explizit in geschlossener Form angegeben werden können. Dabei gilt für den verallgemeinerten Druck*

$$u = \kappa(p) = \begin{cases} \frac{p_b}{-\lambda e(\lambda)+1} \left(\frac{p}{p_b}\right)^{-\lambda e(\lambda)+1} - u_c & \text{für } p \leq p_b, \\ p & \text{für } p \geq p_b \end{cases}$$

mit

$$e(\lambda) := 3 + \frac{2}{\lambda}, \quad u_c := \frac{\lambda e(\lambda)}{\lambda e(\lambda) - 1} p_b < p_b < 0.$$

u_c gibt dabei den sogenannten kritischen verallgemeinerten Druck an. Offensichtlich folgt für einen maximal gesättigten Boden, d.h. $p \geq p_b$ bzw. $k_r(\theta_M) = 1$, die Gleichheit des verallgemeinerten Drucks u und des physikalischen Drucks p . Für den ungesättigten Fall wird das Intervall $(-\infty, p_b)$ auf das beschränkte Intervall (u_c, p_b) abgebildet. Für die inverse Kirchhoff-Transformation folgt damit

$$p = \kappa^{-1}(u) = \begin{cases} p_b \left(\frac{u(-\lambda e(\lambda)+1)}{p_b} + \lambda e(\lambda) \right)^{-\frac{1}{-\lambda e(\lambda)+1}} & \text{für } u_c < u \leq p_b, \\ u & \text{für } u \geq p_b. \end{cases}$$

In ähnlicher Weise können geschlossene Darstellungen für die Sättigung und die relative Durchlässigkeit angegeben werden, siehe [6].

Bemerkung 1.3. *Ist die relative Permeabilität $k_r(\cdot)$ nicht nur vom Druck p , sondern auch explizit vom Ort $x \in \Omega$ abhängig, also $k_r = k_r(x, \theta(p(x, t)))$, so vereinfacht die Kirchhoff-Transformation die Richards-Gleichung nicht. So ein Fall tritt zum Beispiel für einen vollkommen heterogenen Boden auf, siehe [6].*

1.4 Randwertprobleme für die Richards-Gleichung

Ohne Beschränkung der Allgemeinheit gelte für die Porosität $n = 1$ und für die maximale hydraulische Leitfähigkeit $K_h = 1$.

Damit folgt die Richards-Gleichung in der Form

$$\theta(p(x, t))_t - \operatorname{div}(k_r(\theta(p(x, t))) \nabla(cp(x, t) - z)) = 0. \quad (1.13)$$

1.4 Randwertprobleme für die Richards-Gleichung

Die zugehörige transformierte Gleichung lautet

$$M(u(x, t))_t - \operatorname{div}(c \nabla u(x, t) - k_r(M(u(x, t))) e_z) = 0. \quad (1.14)$$

Es gilt $c = 1/\rho g = \text{const}$. Ohne Beschränkung der Allgemeinheit sei $c = 1$.

Als Anfangsbedingung wird $p_0(x) = p(x, 0) := \kappa^{-1}(u(x, 0)) = \kappa^{-1}(u_0(x)) \in H^1(\Omega)$ gewählt.

Die nächste Aufgabe ist die Beschreibung von Randbedingungen für die Richards-Gleichung (1.13) bzw. für die transformierte Richards-Gleichung (1.14).

Sei $\Omega \subset \mathbb{R}^3$, eine offene, beschränkte, zusammenhängende und nicht leere Menge mit einem Lipschitz-Rand $\partial\Omega$.

Für eine gegebene Zeit $t \in [0, T]$ werde $\partial\Omega$ zerlegt in Teilmengen $\Gamma_D(t)$, $\Gamma_N(t)$ und $\Gamma_S(t)$ mit verschiedenen Randbedingungen.

Weiters seien $p_D(x, t) := \kappa^{-1}(u_D(x, t)) \in H^{1/2}(\Gamma_D(t))$ und $f_N(x, t) \in L_2(\Gamma_N(t))$ gegeben.

Der unbekanntes physikalische Druck p und der unbekanntes Fluss

$$\mathbf{v}(x, t) = -(\nabla u(x, t) - k_r(M(u(x, t))) e_z) = -k_r(\theta(p(x, t))) \nabla(p(x, t) - z)$$

sollen nun die folgenden Randbedingungen erfüllen:

1. Dirichlet-Randbedingungen:

$$p(x, t) = p_D(x, t) \quad \text{auf } \Gamma_D(t).$$

Die Dirichlet-Randbedingungen können in äquivalenter Weise für den verallgemeinerten Druck u angegeben werden:

$$u(x, t) = u_D(x, t) \quad \text{auf } \Gamma_D(t).$$

Dirichlet-Randbedingungen treten für gewöhnlich als hydrostatischer Druck auf. Bei Seen und Flüssen ist dieser durch Oberflächenwasser gegeben, bei Flüssen und Meeren durch Wasser auf den Seitenrändern des Gebietes.

2. Neumann-Randbedingungen:

$$\mathbf{v}(x, t) \cdot \mathbf{n}(x, t) = f_N(x, t) \quad \text{auf } \Gamma_N(t).$$

Da sich die Neumann-Randbedingungen lediglich auf den physikalischen Wasserfluss \mathbf{v} beziehen, können sie sowohl durch p als auch durch u ausgedrückt werden.

1 Modellierung

Neumann-Randbedingungen bestimmen den Wasserfluss in Ω hinein oder aus Ω heraus aufgrund von Regen oder Wasserbewegungen auf $\partial\Omega$.

Sie treten oft als homogene Neumann-Randbedingungen auf, d.h. in diesem Fall entspricht der Neumann-Rand dem undurchlässigen Teil des Randes, durch den kein Herausströmen von Flüssigkeit erlaubt ist. Solche homogenen Neumann-Randbedingungen treten an trockenen Bereichen der Bodenoberfläche oder am Übergang zu undurchlässigen Böden auf.

3. Signorini-Randbedingungen:

$$p(x, t) \leq 0, \quad \mathbf{v}(x, t) \cdot \mathbf{n}(x, t) \geq 0, \quad p(x, t) (\mathbf{v}(x, t) \cdot \mathbf{n}(x, t)) = 0 \quad \text{auf } \Gamma_S(t).$$

Es gilt:

- $u = p$ für $p \geq 0$ oder $u \geq 0$,
- $p < 0 \Leftrightarrow u < 0$.

Somit können die Signorini-Randbedingungen äquivalent formuliert werden, indem p durch u ersetzt wird:

$$u(x, t) \leq 0, \quad \mathbf{v}(x, t) \cdot \mathbf{n}(x, t) \geq 0, \quad u(x, t) (\mathbf{v}(x, t) \cdot \mathbf{n}(x, t)) = 0 \quad \text{auf } \Gamma_S(t).$$

Damit diese Ungleichungsrandbedingung erfüllt ist, wird der Signorini-Rand in zwei Teilstücke zerlegt. Die sogenannte „seepage face“ ist dabei der Teil des Signorini-Randes, in dem der Wasserfluss positiv ist und der Wasserdruck verschwindet. Im anderen Teil des Signorini-Randes hingegen verschwindet der Wasserfluss, wobei der Wasserdruck negativ ist.

Signorini-Randbedingungen treten üblicherweise bei Oberflächenwasserreservoirs oder bei sogenannten „Dammproblemen“ auf. Das nachfolgende Beispiel 1.4.1 vertieft noch einmal die Herleitung der konkreten Form der Signorini-Randbedingungen und deren physikalische Interpretation, siehe [6].

Beispiel 1.4.1. *Das Modellgebiet aus Abbildung 1.1 kann als vertikaler Schnitt durch einen Boden betrachtet werden. Über dem Randstück Γ_3 , das zum Dirichlet-Rand gehört, sei außerdem ein Wasserreservoir gegeben. Obwohl die Oberfläche des Wasserreservoirs durch Γ_6 gegeben ist, kann das Gebiet Ω bis zu einer Kurve über Γ_6 , auf der der Druck verschwindet, vollkommen gesättigt sein. In diesem Beispiel ist diese Kurve gegeben durch $\Gamma_7 \cup \Gamma_2 \cup \Gamma_4 \cup \Gamma_8$. Das Randstück $\Gamma_2 \cup \Gamma_4$ wird als sogenannte „seepage face“ (z.B.: Ufer von Fließgewässern) bezeichnet. In dieser Zone kann Wasser aus dem Inneren von Ω durch $\partial\Omega$ hervortreten und dann wieder ins Reservoir hinunterfließen. Der Fluss ist hier also positiv, d.h. $\mathbf{v} \cdot \mathbf{n} \geq 0$, während der Wasserdruck $p = u = 0$ ist.*

Wird nun das Randstück $\Gamma_1 \cup \Gamma_5$ betrachtet, so ist zu bemerken, dass hier einerseits der Fluss verschwindet, d.h. $\mathbf{v} \cdot \mathbf{n} = 0$, und dass der Wasserdruck andererseits nicht positiv sein kann, d.h. $p \leq 0 \Leftrightarrow u \leq 0$.

1.4 Randwertprobleme für die Richards-Gleichung

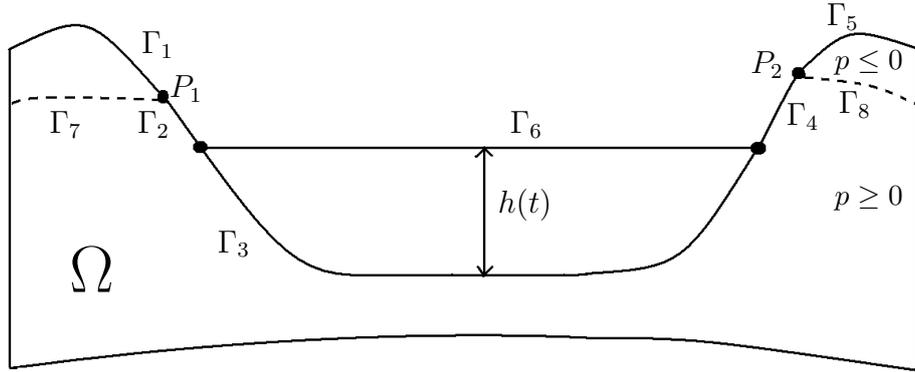


Abbildung 1.1: Modellgebiet Ω

Das ist der hydrologische Grund für die ergänzenden Bedingungen an u und den Fluss $\mathbf{v} \cdot \mathbf{n}$ bei der Signorini-Randbedingung

$$u \leq 0, \quad \mathbf{v} \cdot \mathbf{n} \geq 0, \quad u(\mathbf{v} \cdot \mathbf{n}) = 0 \quad \text{auf } \Gamma_S(t).$$

Es resultiert schließlich eine Zerlegung des Signorini-Randes in sogenannte „seepage faces“ (hier: $\Gamma_2 \cup \Gamma_4$) und in angrenzende Teilmengen von $\partial\Omega$ (hier: $\Gamma_1 \cup \Gamma_5$).

Weiters sei noch erwähnt, dass die Punkte P_1 und P_2 , die den Rand der „seepage faces“ innerhalb des Signorini-Randes bestimmen, üblicherweise unbekannt sind und sich erst als Teil der Lösung, die die Signorini-Randbedingungen erfüllt, ergeben.

Das Anfangs-Randwertproblem für die Richards-Gleichung (1.13) lautet:

Gesucht ist $p(x, t)$ mit $(x, t) \in \Omega \times (0, T)$, sodass

$$\theta(p(x, t))_t - \operatorname{div}(k_r(\theta(p(x, t))) \nabla(p(x, t) - z)) = 0 \quad \text{in } \Omega, \quad (1.15)$$

$$p(x, t) = p_D(x, t) \quad \text{auf } \Gamma_D(t), \quad (1.16)$$

$$\mathbf{v}(x, t) \cdot \mathbf{n}(x, t) = f_N(x, t) \quad \text{auf } \Gamma_N(t), \quad (1.17)$$

$$p(x, t) \leq 0, \quad \mathbf{v}(x, t) \cdot \mathbf{n}(x, t) \geq 0, \quad p(x, t)(\mathbf{v}(x, t) \cdot \mathbf{n}(x, t)) = 0 \quad \text{auf } \Gamma_S(t), \quad (1.18)$$

$$p(x, 0) = p_0(x) \quad \text{für } x \in \Omega. \quad (1.19)$$

Analog dazu lautet das Anfangs-Randwertproblem für die Kirchhoff-transformierte Richards-Gleichung (1.14):

1 Modellierung

Gesucht ist $u(x, t)$ mit $(x, t) \in \Omega \times (0, T)$, sodass

$$M(u(x, t))_t - \operatorname{div}(\nabla u(x, t) - k_r(M(u(x, t)))e_z) = 0 \quad \text{in } \Omega, \quad (1.20)$$

$$u(x, t) = u_D(x, t) \quad \text{auf } \Gamma_D(t), \quad (1.21)$$

$$\mathbf{v}(x, t) \cdot \mathbf{n}(x, t) = f_N(x, t) \quad \text{auf } \Gamma_N(t), \quad (1.22)$$

$$u(x, t) \leq 0, \quad \mathbf{v}(x, t) \cdot \mathbf{n}(x, t) \geq 0, \quad u(x, t)(\mathbf{v}(x, t) \cdot \mathbf{n}(x, t)) = 0 \quad \text{auf } \Gamma_S(t), \quad (1.23)$$

$$u(x, 0) = u_0(x) \quad \text{für } x \in \Omega. \quad (1.24)$$

Bemerkung 1.4. *Im Allgemeinen gilt für die Sättigung M als Funktion des verallgemeinerten Drucks u , dass $M : \mathbb{R} \rightarrow \mathbb{R}$. Aufgrund der Wahl der Brooks-Corey-Parameterfunktionen ist die verallgemeinerte Sättigung jedoch nur auf einem Intervall (u_c, ∞) mit $u_c < 0$ definiert, also $M : (u_c, \infty) \rightarrow \mathbb{R}$. Damit u weiterhin das Randwertproblem (1.20) - (1.23) erfüllt, muss gefordert werden, dass der Bildbereich von u eine Teilmenge von (u_c, ∞) ist, d.h. $u > u_c$. Analoges gilt für die Dirichlet-Randbedingung u_D , d.h. $u_D > u_c$.*

2 Analysis

Zunächst werden in Abschnitt 2.1 die grundlegenden Definitionen und Eigenschaften der Sobolev-Räume wiederholt. Es werden hier lediglich die Aussagen der Sätze angegeben, für die zugehörigen Beweise wird auf die jeweiligen Quellen verwiesen. Danach werden in Abschnitt 2.2 schwache Formulierungen des Randwertproblems (1.20) - (1.23) und des Randwertproblems (1.15) - (1.18) in Form von Variationsungleichungen auf konvexen Teilmengen des $H^1(\Omega)$ hergeleitet, welche anschließend hinsichtlich ihrer Äquivalenz näher betrachtet werden. In Abschnitt 2.3 wird schließlich genauer auf eine implizit-explicite Zeitdiskretisierung der zuvor hergeleiteten Variationsungleichungen eingegangen.

2.1 Sobolev-Räume

Für die nachfolgende Definition der Sobolev-Räume sei auf [33] verwiesen, wobei die Struktur und Notation hier größtenteils übernommen werden.

Für einen Multiindex $\alpha = (\alpha_1, \dots, \alpha_d)$, $\alpha_i \in \mathbb{N}_0$, mit

$$|\alpha| = \alpha_1 + \dots + \alpha_d, \quad \alpha! = \alpha_1! \dots \alpha_d!, \quad d \in \mathbb{N},$$

und für eine hinreichend oft differenzierbare Funktion $u(x)$ mit $x \in \mathbb{R}^d$ kann die partielle Ableitung wie folgt geschrieben werden:

$$D^\alpha u(x) = \left(\frac{\partial}{\partial x_1} \right)^{\alpha_1} \dots \left(\frac{\partial}{\partial x_d} \right)^{\alpha_d} u(x_1, \dots, x_d).$$

Definition 2.1 (C^k -Räume). Sei $\Omega \subseteq \mathbb{R}^d$ ein offenes Gebiet, $k \in \mathbb{N}_0$.

- $C^k(\Omega)$ ist definiert als der Raum der auf Ω beschränkten und k -mal stetig differenzierbaren Funktionen mit der zugehörigen Norm

$$\|u\|_{C^k(\Omega)} = \sum_{|\alpha| \leq k} \sup_{x \in \Omega} |D^\alpha u(x)|.$$

- $C^\infty(\Omega)$ ist der Raum der auf Ω beschränkten und unendlich oft stetig differenzierbaren Funktionen.
- Durch $C_0^\infty(\Omega) := \left\{ u \in C^\infty(\Omega) : \text{supp } u = \overline{\{x \in \Omega : u(x) \neq 0\}} \subset \Omega \right\}$ ist der Raum der unendlich oft stetig differenzierbaren Funktionen mit kompaktem Träger gegeben.

2 Analysis

- Für $\kappa \in (0, 1)$ wird durch $C^{k,\kappa}(\Omega)$ der Raum der auf Ω Hölder-stetigen Funktionen mit der Norm

$$\|u\|_{C^{k,\kappa}} := \|u\|_{C^k(\Omega)} + \sum_{|\alpha|=k} \sup_{x,y \in \Omega, x \neq y} \frac{|D^\alpha u(x) - D^\alpha u(y)|}{|x - y|^\kappa}$$

definiert. Insbesondere für $\kappa = 1$ ist $C^{k,1}(\Omega)$ der Raum der auf Ω Lipschitz-stetigen Funktionen.

Definition 2.2 (L_p-Räume). Sei $\Omega \subseteq \mathbb{R}^d$ ein offenes Gebiet.

- $L_p(\Omega)$ ist definiert als der Raum von Äquivalenzklassen aller auf Ω definierten messbaren Funktionen, deren p -te Potenz integrierbar ist. Die zugehörige Norm ist:

$$\|u\|_{L_p(\Omega)} = \left\{ \int_{\Omega} |u(x)|^p dx \right\}^{1/p} \quad \text{für } 1 \leq p < \infty.$$

- $L_\infty(\Omega)$ ist der Raum der auf Ω messbaren und fast überall beschränkten Funktionen mit der Norm

$$\|u\|_{L_\infty(\Omega)} := \operatorname{ess\,sup}_{x \in \Omega} \{|u(x)|\} := \inf_{K \subset \mathbb{R}, \mu(K)=0} \sup_{x \in \Omega \setminus K} |u(x)|.$$

Die $L_p(\Omega)$ -Räume sind bezüglich der angegebenen Norm vollständig, damit Banachräume. Für den Spezialfall $p = 2$ ist der $L_2(\Omega)$, der Raum der quadrat-integrierbaren Funktionen, mit dem Skalarprodukt

$$\langle u, v \rangle_{L_2(\Omega)} := \int_{\Omega} u(x)v(x) dx$$

und der Norm

$$\|u\|_{L_2(\Omega)}^2 = \langle u, u \rangle_{L_2(\Omega)}$$

ein Hilbertraum.

Mit dem Dualitätsprodukt

$$\langle u, v \rangle_{\Omega} = \int_{\Omega} u(x)v(x) dx$$

folgt:

$$\|u\|_{L_q(\Omega)} = \sup_{0 \neq v \in L_p(\Omega)} \frac{|\langle v, u \rangle_{\Omega}|}{\|v\|_{L_p(\Omega)}} \quad \text{für } 1 \leq p < \infty, \quad \frac{1}{p} + \frac{1}{q} = 1.$$

Für $1 \leq p < \infty$, $\frac{1}{p} + \frac{1}{q} = 1$, ist nun insbesondere $L_q(\Omega)$ der Dualraum zu $L_p(\Omega)$, $L_\infty(\Omega)$ ist der Dualraum zu $L_1(\Omega)$, jedoch ist $L_1(\Omega)$ nicht der Dualraum zu $L_\infty(\Omega)$. Weiters erfüllen die L_p -Räume die Hölder- sowie die Minkowskische-Ungleichung, siehe [33].

Definition 2.3 (Verallgemeinerte partielle Ableitung). *Existiert eine Funktion $w = D^\alpha u \in L_1^{\text{loc}}(\Omega) := \{v \in L_1(K) \mid \forall \text{ kompakten } K \subset \Omega\}$, sodass*

$$\int_{\Omega} w(x)\varphi(x) \, dx = (-1)^{|\alpha|} \int_{\Omega} u(x)D^\alpha\varphi(x) \, dx \quad \forall \varphi \in C_0^\infty(\Omega)$$

erfüllt ist, dann heißt w verallgemeinerte Ableitung der Ordnung α von $u \in L_1^{\text{loc}}(\Omega)$.

Für $k \in \mathbb{N}_0$ wird durch

$$\|u\|_{W_p^k(\Omega)} := \begin{cases} \left\{ \sum_{|\alpha| \leq k} \|D^\alpha u\|_{L_p(\Omega)}^p \right\}^{1/p} & \text{für } 1 \leq p < \infty \\ \max_{|\alpha| \leq k} \|D^\alpha u\|_{L_\infty(\Omega)} & \text{für } p = \infty \end{cases}$$

eine Norm definiert.

Definition 2.4 (Sobolev-Räume natürlicher Ordnung). *Durch Vervollständigung von $C^\infty(\Omega)$ bezüglich der Norm $\|\cdot\|_{W_p^k(\Omega)}$ wird der Sobolev-Raum $W_p^k(\Omega) = \overline{C^\infty(\Omega)}^{\|\cdot\|_{W_p^k(\Omega)}}$ definiert.*

Entsprechend definiert die Vervollständigung von $C_0^\infty(\Omega)$ bezüglich obiger Norm den Sobolev-Raum $\dot{W}_p^k(\Omega) = \overline{C_0^\infty(\Omega)}^{\|\cdot\|_{W_p^k(\Omega)}}$.

Definition 2.5 (Sobolev-Räume reeller Ordnung). *Für $0 < s \in \mathbb{R} \setminus \mathbb{N}$, $k \in \mathbb{N}_0$, $\kappa \in (0, 1)$ mit $s = k + \kappa$ wird durch*

$$\|u\|_{W_p^s(\Omega)} := \left\{ \|u\|_{W_p^k(\Omega)}^p + |u|_{W_p^\kappa(\Omega)}^p \right\}^{1/p}$$

die Sobolev-Slobodeckii-Norm mit der Halbnorm

2 Analysis

$$|u|_{W_p^k(\Omega)}^p = \sum_{|\alpha|=k} \int_{\Omega} \int_{\Omega} \frac{|D^\alpha u(x) - D^\alpha u(y)|^p}{|x-y|^{d+\kappa p}} dx dy$$

definiert.

Für den Spezialfall $p = 2$ ist der Sobolev-Raum $W_2^s(\Omega)$ ein Hilbertraum mit dem Skalarprodukt

$$\langle u, v \rangle_{W_2^s(\Omega)} := \langle u, v \rangle_{W_2^k(\Omega)} + \sum_{|\alpha|=k} \int_{\Omega} \int_{\Omega} \frac{(D^\alpha u(x) - D^\alpha u(y))(D^\alpha v(x) - D^\alpha v(y))}{|x-y|^{d+2\kappa}} dx dy$$

für $s = k + \kappa$, $k \in \mathbb{N}_0$, $\kappa \in (0, 1)$. Für $s < 0$, $1 < p < \infty$ und $\frac{1}{p} + \frac{1}{q} = 1$ wird der Sobolev-Raum $W_p^s(\Omega)$ als Dualraum von $\mathring{W}_q^{-s}(\Omega)$ mit der Norm

$$\|u\|_{W_p^s(\Omega)} := \sup_{0 \neq v \in \mathring{W}_q^{-s}(\Omega)} \frac{|\langle u, v \rangle_{\Omega}|}{\|v\|_{W_q^{-s}(\Omega)}}$$

definiert. Analog dazu gilt $\mathring{W}_p^s(\Omega) = (W_q^{-s}(\Omega))'$.

Für wichtige Eigenschaften von Sobolev-Räumen, wie den Sobolevschen Einbettungssatz, den Normierungssatz von Sobolev und das Bramble-Hilbert-Lemma, siehe [33].

Die Sobolev-Räume lassen sich auch mittels Distributionen und der Fourier-Transformierten definieren. In diesem Fall werden die Räume mit $H^s(\mathbb{R}^d)$ bezeichnet. Den Zusammenhang mit den hier definierten Sobolev-Räumen gibt der nachfolgende Satz an, siehe dazu [28, 35].

Satz 2.6. Für alle $s \in \mathbb{R}$ gilt:

$$H^s(\mathbb{R}^d) = W_2^s(\mathbb{R}^d).$$

Für ein Gebiet $\Omega \subset \mathbb{R}^d$ ist der Sobolev-Raum $H^s(\Omega)$ definiert durch

$$H^s(\Omega) := \{v = \tilde{v}|_{\Omega} : \tilde{v} \in H^s(\mathbb{R}^d)\}$$

mit der zugehörigen Norm

$$\|v\|_{H^s(\Omega)} := \inf_{\tilde{v} \in H^s(\mathbb{R}^d), \tilde{v}|_{\Omega} = v} \|\tilde{v}\|_{H^s(\mathbb{R}^d)}.$$

Falls Ω ein Lipschitz-Gebiet ist, gilt die Äquivalenz von $W_2^s(\Omega)$ und $H^s(\Omega)$ für alle $s > 0$, siehe [36].

Für den Rand $\Gamma = \partial\Omega$ eines Lipschitz-Gebietes $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, ist durch

$$\|u\|_{L_2(\Gamma)} = \left(\int_{\Gamma} |u(x)|^2 ds_x \right)^{1/2}$$

die $L_2(\Gamma)$ -Norm gegeben.

Definition 2.7 (Sobolev-Räume auf dem Rand). Für $s \in (0, 1)$ ist die Sobolev-Slobodeckii-Norm auf dem Rand wie folgt definiert:

$$\|u\|_{H^s(\Gamma)} := \left\{ \|u\|_{L_2(\Gamma)}^2 + \int_{\Gamma} \int_{\Gamma} \frac{|u(x) - u(y)|^2}{|x - y|^{d-1+2s}} ds_x ds_y \right\}^{1/2}.$$

Für $s < 0$ ist $H^s(\Gamma)$ wiederum der Dualraum von $H^{-s}(\Gamma)$, also $H^s(\Gamma) := [H^{-s}(\Gamma)]'$, mit der zugehörigen Norm

$$\|w\|_{H^s(\Gamma)} := \sup_{0 \neq v \in H^{-s}(\Gamma)} \frac{\langle w, v \rangle_{\Gamma}}{\|v\|_{H^{-s}(\Gamma)}}.$$

Dabei wird mit $\langle \cdot, \cdot \rangle_{\Gamma}$ das Dualitätsprodukt

$$\langle w, v \rangle_{\Gamma} := \int_{\Gamma} w(x)v(x) ds_x$$

bezeichnet.

Definition 2.8 (Sobolev-Räume auf einem Teil des Randes). Sei $\Gamma_0 \subset \Gamma$ eine offene Teilmenge eines hinreichend glatten Randes Γ . Dann definiere für $s \geq 0$:

$$H^s(\Gamma_0) := \{v = \tilde{v}|_{\Gamma_0} : \tilde{v} \in H^s(\Gamma)\}$$

mit der Norm

$$\|v\|_{H^s(\Gamma_0)} := \inf_{\tilde{v} \in H^s(\Gamma) : \tilde{v}|_{\Gamma_0} = v} \|\tilde{v}\|_{H^s(\Gamma)}$$

bzw.

$$\tilde{H}^s(\Gamma_0) = \{v = \tilde{v}|_{\Gamma_0} : \tilde{v} \in H^s(\Gamma), \text{supp } \tilde{v} \subset \Gamma_0\}.$$

Definition 2.9 (Innere Spur). Sei $\Omega \subset \mathbb{R}^d$ ein beschränktes, einfach zusammenhängendes Gebiet mit hinreichend glattem Rand $\Gamma = \partial\Omega$. Dann definiert

$$\gamma_0^{\text{int}} f(x) := \lim_{\Omega \ni \tilde{x} \rightarrow x \in \Gamma} f(\tilde{x}) \quad \text{für } x \in \Gamma$$

die innere Spur der in Ω gegebenen Funktion f .

Durch Definition 2.9 wird die innere Spur $\gamma_0^{\text{int}} u$ einer in Ω gegebenen Funktion u als Funktion auf dem Rand Γ erklärt. Damit folgt sofort der Zusammenhang zwischen den entsprechenden Funktionenräumen, siehe [35, 1, 28].

Satz 2.10 (Spursatz). Sei $\Omega \subset \mathbb{R}^d$ ein $C^{k-1,1}$ -Gebiet. Für $\frac{1}{2} < s \leq k$ ist

$$\gamma_0^{\text{int}} : H^s(\Omega) \rightarrow H^{s-1/2}(\Gamma)$$

ein beschränkter linearer Operator, d.h. es gilt

$$\|\gamma_0^{\text{int}} v\|_{H^{s-1/2}(\Gamma)} \leq c_T \|v\|_{H^s(\Omega)} \quad \text{für } v \in H^s(\Omega).$$

Satz 2.11 (Inverser Spursatz). Sei Ω ein $C^{k-1,1}$ -Gebiet. Dann besitzt der Spuroperator $\gamma_0^{\text{int}} : H^s(\Omega) \rightarrow H^{s-1/2}(\Gamma)$, $\frac{1}{2} < s \leq k$, eine stetige Rechtsinverse

$$\mathcal{E} : H^{s-1/2}(\Gamma) \rightarrow H^s(\Omega)$$

mit $\gamma_0^{\text{int}} \mathcal{E} w = w$ für alle $w \in H^{s-1/2}(\Gamma)$ und

$$\|\mathcal{E} w\|_{H^s(\Omega)} \leq c_{IT} \|w\|_{H^{s-1/2}(\Gamma)} \quad \forall w \in H^{s-1/2}(\Gamma).$$

2.2 Herleitung der Variationsungleichung

Die Herleitung einer schwachen Formulierung des Randwertproblems für die Richards-Gleichung (1.20) - (1.23) in Form einer Variationsungleichung folgt aus einer Verallgemeinerung der Green'schen Formel bzw. aus partieller Integration im \mathbb{R}^d .

Satz 2.12. Sei $\Omega \subset \mathbb{R}^d$ ein beschränktes Lipschitz-Gebiet mit messbarem Rand $\partial\Omega$, $G : \bar{\Omega} \rightarrow \mathbb{R}^d$ sei ein stetiges Vektorfeld und differenzierbar auf Ω und $v : \bar{\Omega} \rightarrow \mathbb{R}$ sei eine stetige skalare Funktion, die ebenfalls differenzierbar auf Ω ist. Weiters seien $\text{div} G$ und die partiellen Ableitungen von v beschränkt auf Ω . Dann gilt:

$$\int_{\Omega} \text{div}(G(x)) v(x) dx = - \int_{\Omega} G(x) \nabla v(x) dx + \int_{\partial\Omega} (G(x) \cdot \mathbf{n}(x)) v(x) ds_x.$$

Es sei $\Omega \subset \mathbb{R}^d$ ein Lipschitz-Gebiet und $\Gamma_D(t)$, $\Gamma_N(t)$ und $\Gamma_S(t)$ seien paarweise disjunkte, messbare Teilmengen von $\partial\Omega$ für alle $t \in [0, T]$ mit

$$\partial\Omega = \Gamma_D(t) \cup \Gamma_N(t) \cup \Gamma_S(t).$$

Weiters wird vorausgesetzt, dass die Sättigung $M : [u_c, \infty) \rightarrow \mathbb{R}$ und die relative Durchlässigkeit $k_r : M([u_c, \infty)) \rightarrow \mathbb{R}$ stetige, monoton wachsende und beschränkte Funktionen mit kritischem verallgemeinerten Druck $u_c < 0$ sind.

Wähle für $t \in [0, T]$ die Dirichlet-Randbedingung

$$u_D(t) \in \{v = \text{tr}_{\Gamma_D(t)} w : w \in H^1(\Omega) \wedge w \geq u_c \text{ f. ü.} \wedge \text{tr}_{\Gamma_S(t)} w \leq 0 \text{ f. ü. auf } \Gamma_S(t)\} \quad (2.1)$$

als Element von $H^{1/2}(\Gamma_D(t))$ mit dem Spuroperator

$$\text{tr}_{\Sigma} : H^1(\Omega) \rightarrow H^{1/2}(\Sigma) \quad \text{für } \Sigma \in \{\Gamma_D(t), \Gamma_S(t)\}.$$

Der Bildbereich der Dirichlet-Randbedingung $u_D(t)$ muss im Intervall (u_c, ∞) fast überall auf $\Gamma_D(t)$ enthalten sein, vergleiche dazu Bemerkung 1.4. Analog dazu muss auch die Fortsetzung von $u_D(t)$ im $H^1(\Omega)$ in diesem Intervall fast überall auf Ω enthalten sein. Weiters gilt, dass die Menge (2.1) nicht leer ist, da zumindest $u_D(t) = 0$ in der Menge (2.1) liegt.

Definiere die folgende Menge:

$$\mathcal{K}(t) := \{v \in H^1(\Omega) : v \geq u_c \text{ f.ü.} \wedge \text{tr}_{\Gamma_D(t)} v = u_D(t) \wedge \text{tr}_{\Gamma_S(t)} v \leq 0 \text{ f.ü. auf } \Gamma_S(t)\}. \quad (2.2)$$

Lemma 2.13. $\mathcal{K}(t)$ ist eine nichtleere, abgeschlossene und konvexe Teilmenge von $H^1(\Omega)$.

Der nachfolgende Beweis wird wie in [6] durchgeführt.

Beweis: Zunächst wird gezeigt, dass die Menge $\mathcal{K}(t)$ konvex ist. Dazu wähle $v, w \in \mathcal{K}(t)$ und $\lambda \in [0, 1]$. Zu zeigen ist, dass $\lambda v + (1 - \lambda) w$ auch in der Menge $\mathcal{K}(t)$ liegt.

Da $v \geq u_c$ bzw. $w \geq u_c$ gilt und sowohl $\lambda \geq 0$ als auch $(1 - \lambda) \geq 0$, folgt auch $\lambda v + (1 - \lambda) w \geq u_c$. Aufgrund der Linearität der Spurooperatoren $\text{tr}_{\Gamma_D(t)}$ und $\text{tr}_{\Gamma_S(t)}$ folgt schließlich

$$\text{tr}_{\Gamma_D(t)}(\lambda v + (1 - \lambda) w) = \lambda \underbrace{\text{tr}_{\Gamma_D(t)}(v)}_{=u_D(t)} + (1 - \lambda) \underbrace{\text{tr}_{\Gamma_D(t)}(w)}_{=u_D(t)} = u_D(t)$$

bzw.

$$\text{tr}_{\Gamma_S(t)}(\lambda v + (1 - \lambda) w) = \lambda \underbrace{\text{tr}_{\Gamma_S(t)}(v)}_{\leq 0} + (1 - \lambda) \underbrace{\text{tr}_{\Gamma_S(t)}(w)}_{\leq 0} \leq 0,$$

womit die Konvexität bewiesen wäre.

Die Menge $\mathcal{K}(t)$ ist weiters nicht leer, da die Dirichlet-Randdaten $u_D(t)$ aus der Menge (2.1) gewählt werden und laut dieser Definition eine Fortsetzung $w \in H^1(\Omega)$ von $u_D(t)$ mit $w \geq u_c$ fast überall auf Ω und $\text{tr}_{\Gamma_S(t)} w \leq 0$ fast überall auf $\Gamma_S(t)$ existiert. Somit gilt $w \in \mathcal{K}(t)$.

Zu zeigen bleibt noch die Abgeschlossenheit. Aufgrund der σ -Additivität des Lebesgue-Maßes und der Gültigkeit von

$$\bigcup_{n \in \mathbb{N}} \left\{ x \in \Omega' : w(x) < u_c - \frac{1}{n} \text{ f. ü.} \right\} = \{x \in \Omega' : w(x) < u_c \text{ f. ü.}\}$$

folgt zunächst, dass für jedes $w \in H^1(\Omega)$, das die Bedingung $w(x) < u_c$ fast überall auf einer Teilmenge $\Omega' \subset \Omega$ mit positivem Lebesgue-Maß erfüllt, auch eine Teilmenge $\Omega'' \subset \Omega'$ mit positivem Lebesgue-Maß gefunden werden kann, sodass $w(x) < u_c - \varepsilon$ für ein $\varepsilon > 0$ gilt. Es folgt die gleichmäßige Beschränktheit der L_2 -Norm $\|w - v\|_{L_2(\Omega)}$ und damit auch der H^1 -Norm $\|w - v\|_{H^1(\Omega)}$ von unten durch eine positive Konstante für alle $v \in \mathcal{K}(t)$.

Unter der gleichen Vorgehensweise und mit geeigneter Wahl der Konstanten $c, C > 0$ kann für ein $w \in H^1(\Omega)$, aber $w \notin \mathcal{K}(t)$, d.h. die Bedingung $\text{tr}_{\Gamma_D(t)} w = u_D(t)$ oder $\text{tr}_{\Gamma_S(t)} w \leq 0$ ist falsch, die nachfolgende Abschätzung angegeben werden. Diese folgt aus der Definition der Räume $H^{1/2}(\Sigma)$ für $\Sigma \in \{\Gamma_D(t), \Gamma_S(t)\}$ und dem Spursatz 2.10:

$$0 < c \leq \|\text{tr}_\Sigma w - \text{tr}_\Sigma v\|_{L_2(\Sigma)} \leq \|\text{tr}_\Sigma w - \text{tr}_\Sigma v\|_{H^{1/2}(\Sigma)} \leq C \|w - v\|_{H^1(\Omega)} \quad \forall v \in \mathcal{K}(t).$$

2 Analysis

Insgesamt ist damit $H^1(\Omega) \setminus \mathcal{K}(t)$ offen und daraus folgt die Abgeschlossenheit von $\mathcal{K}(t) \subset H^1(\Omega)$. ■

Definition 2.14. Sei X ein Banachraum. Dann ist durch

$$L_2(0, T; X) := \left\{ v : (0, T) \rightarrow X, \|v\|_{L_2(0, T; X)} = \left[\int_0^T \|v(t, \cdot)\|_X^2 dt \right]^{1/2} < \infty \right\}$$

ein Banachraum definiert.

Satz 2.15. Seien die Sättigung $M : [u_c, \infty) \rightarrow \mathbb{R}$ und die relative Durchlässigkeit $k_r : M([u_c, \infty)) \rightarrow \mathbb{R}$ stetige, monoton wachsende und beschränkte Funktionen mit $u_c < 0$. Weiters sei $u \in C^2(\Omega)$ und $f_N(t)$ sei stetig auf $\Gamma_N(t)$.

Dann heißt $u \in L_2(0, T; H^1(\Omega))$ unter der zusätzlichen Bedingung, dass $M(u)_t \in L_2(\Omega)$ fast überall auf dem Intervall $(0, T]$ gilt, eine schwache Lösung von (1.20) - (1.23) zum Zeitpunkt $t \in (0, T]$ genau dann, wenn $u(t) \in \mathcal{K}(t)$ und

$$\begin{aligned} \int_{\Omega} M(u(x, t))_t (v(x) - u(x, t)) dx + \int_{\Omega} \nabla u(x, t) \nabla (v(x) - u(x, t)) dx &\geq \quad (2.3) \\ \int_{\Omega} k_r(M(u(x, t))) e_z \nabla (v(x) - u(x, t)) dx - \int_{\Gamma_N(t)} f_N(t) (v(x) - u(x, t)) ds_x &\quad \forall v \in \mathcal{K}(t). \end{aligned}$$

Bemerkung 2.1. Die zusätzliche Bedingung $M(u)_t \in L_2(\Omega)$ fast überall in $(0, T]$ ist notwendig, da für die Lösung $u \in L_2(0, T; H^1(\Omega))$ eines parabolischen Problems in $\Omega \times (0, T)$, die partielle Ableitung nach der Zeit u_t im Allgemeinen nicht mehr als Funktion interpretiert werden kann.

Bemerkung 2.2. $k_r(M(u(\cdot)))$ bleibt beschränkt und monoton wachsend auf Ω , da sowohl die Sättigung M als auch die relative Durchlässigkeit k_r beschränkt sind. Da die Komposition einer monotonen und einer messbaren Funktion wieder messbar ist, folgt $k_r(M(u(\cdot))) \in L_{\infty}(\Omega)$ und damit macht jeder Term der Variationsungleichung (2.3) Sinn.

Beweis: Sei u eine schwache Lösung des Randwertproblems (1.20) - (1.23). Da u dieses Randwertproblem erfüllt, folgt laut Definition der konvexen Menge $\mathcal{K}(t)$ sofort $u \in \mathcal{K}(t)$. Wähle ein beliebiges $v \in \mathcal{K}(t)$, multipliziere (1.20) mit $(v - u)$ und integriere über das Gebiet Ω :

$$\int_{\Omega} M(u)_t (v - u) dx + \int_{\Omega} -\operatorname{div} (\nabla u - k_r(M(u)) e_z) (v - u) dx = 0.$$

2.2 Herleitung der Variationsungleichung

Da $u \in C^2(\Omega)$ vorausgesetzt wurde, kann Satz 2.12 angewendet werden und es ergibt sich:

$$\int_{\Omega} M(u)_t (v - u) \, dx + \int_{\Omega} \nabla u \nabla (v - u) \, dx - \int_{\Omega} k_r(M(u)) e_z \nabla (v - u) \, dx + \int_{\partial\Omega} \mathbf{v} \cdot \mathbf{n} (v - u) \, ds_x = 0.$$

Es gilt:

- $v - u = 0$ auf $\Gamma_D(t)$, da $u, v \in \mathcal{K}(t)$.
- $\mathbf{v} \cdot \mathbf{n} = f_N(t)$ auf $\Gamma_N(t)$, da u Lösung des Randwertproblems ist.

Mit $\partial\Omega = \overline{\Gamma_D(t)} \cup \overline{\Gamma_N(t)} \cup \overline{\Gamma_S(t)}$ folgt:

$$\int_{\partial\Omega} \mathbf{v} \cdot \mathbf{n} (v - u) \, ds_x = \int_{\Gamma_D(t)} \mathbf{v} \cdot \mathbf{n} (v - u) \, ds_x + \int_{\Gamma_N(t)} \mathbf{v} \cdot \mathbf{n} (v - u) \, ds_x + \int_{\Gamma_S(t)} \mathbf{v} \cdot \mathbf{n} (v - u) \, ds_x.$$

Betrachte zunächst das Integral über den Dirichlet-Rand:

$$\int_{\Gamma_D(t)} \mathbf{v} \cdot \mathbf{n} (v - u) \, ds_x = 0, \text{ da } u, v \in \mathcal{K}_c(t).$$

Für das Integral über den Neumann-Rand ergibt sich

$$\int_{\Gamma_N(t)} \mathbf{v} \cdot \mathbf{n} (v - u) \, ds_x = \int_{\Gamma_N(t)} f_N(t) (v - u) \, ds_x \text{ wegen (1.22).}$$

Um die Variationsungleichung (2.3) zu erhalten, bleibt noch zu zeigen, dass

$$\int_{\Gamma_S(t)} \mathbf{v} \cdot \mathbf{n} (v - u) \, ds_x \leq 0$$

gilt.

Da u der Signorini Randbedingung (1.23) genügt, folgt

$$\int_{\Gamma_S(t)} \mathbf{v} \cdot \mathbf{n} (v - u) \, ds_x = \int_{\Gamma_S(t)} \underbrace{(\mathbf{v} \cdot \mathbf{n})}_{\geq 0} \underbrace{v}_{\leq 0} \, ds_x - \int_{\Gamma_S(t)} \underbrace{(\mathbf{v} \cdot \mathbf{n})}_{=0} u \, ds_x \leq 0$$

und damit die Variationsungleichung (2.3).

2 Analysis

Umgekehrt sei $u \in \mathcal{K}(t)$ Lösung der Variationsungleichung (2.3) für $t \in (0, T]$.

Durch Anwendung von Satz 2.12 auf die Variationsungleichung (2.3) mit $G = \mathbf{v} = -(\nabla u - k_r(M(u))e_z)$ und der Testfunktion $(v - u)$ anstelle von v , ergibt sich:

$$\int_{\Omega} (M(u)_t + \operatorname{div} \mathbf{v})(v - u) dx + \int_{\Gamma_N(t)} f_N(t)(v - u) ds_x - \int_{\partial\Omega} (\mathbf{v} \cdot \mathbf{n})(v - u) ds_x \geq 0 \quad \forall v \in \mathcal{K}(t). \quad (2.4)$$

Zunächst soll gezeigt werden, dass $u \in \mathcal{K}(t)$ die Differentialgleichung (1.15) für ein t erfüllt. Dazu wird angenommen, dass für ein $x \in \Omega$ eine Kugel $\mathcal{B}_\varepsilon(x) \subset \Omega$ mit Radius ε um den Punkt x existiert, sodass ohne Beschränkung der Allgemeinheit

$$M(u(x, t))_t - \operatorname{div}(\nabla u(x, t) - k_r(M(u(x, t)))e_z) > 0 \quad (2.5)$$

gilt. Wähle $v \in \mathcal{K}(t)$, sodass $v = u$ auf $\overline{\Omega} \setminus \mathcal{B}_\varepsilon(x)$ und $(v - u) < 0$ auf $\mathcal{B}_\varepsilon(x)$. Mit dieser Wahl und der Behauptung (2.5) folgt mit der Stetigkeit der zu integrierenden Funktionen für die Variationsungleichung (2.4)

$$\underbrace{\int_{\mathcal{B}_\varepsilon(x)} \underbrace{(M(u)_t + \operatorname{div} \mathbf{v})}_{>0} \underbrace{(v - u)}_{<0} dx}_{<0} + \underbrace{\int_{\Gamma_N(t)} f_N(t)(v - u) ds_x}_{=0} - \underbrace{\int_{\partial\Omega} (\mathbf{v} \cdot \mathbf{n})(v - u) ds_x}_{=0} < 0.$$

Da obige Aussage einen Widerspruch zu (2.4) darstellt, erfüllt u die Differentialgleichung (1.15). Für eine genauere Theorie zur obigen Konstruktion von glatten Funktionen v bzw. der nachfolgenden Konstruktion von Testfunktionen mit speziellen Eigenschaften auf Teilen des Randes $\partial\Omega$, siehe zum Beispiel [6, 24].

Nächstes Ziel ist zu zeigen, dass die Randbedingungen (1.16) - (1.18) erfüllt sind.

Aufgrund der Tatsache, dass $u \in \mathcal{K}(t)$ gewählt wurde, ist die Dirichlet-Randbedingung (1.16) automatisch erfüllt.

Im Folgenden soll die Gültigkeit der Neumann-Randbedingung gezeigt werden.

Einerseits gilt aufgrund der Wahl von $u, v \in \mathcal{K}(t)$, dass $u = v$ auf $\Gamma_D(t)$ ist und andererseits wurde bereits gezeigt, dass u der Gleichung (1.15) genügt. Damit ergibt sich für die Variationsungleichung (2.4)

$$\int_{\Gamma_N(t)} (f_N(t) - \mathbf{v} \cdot \mathbf{n})(v - u) ds_x - \int_{\Gamma_S(t)} \mathbf{v} \cdot \mathbf{n}(v - u) ds_x \geq 0 \quad \forall v \in \mathcal{K}(t). \quad (2.6)$$

Zu zeigen bleibt, dass $f_N(t) = \mathbf{v} \cdot \mathbf{n}$ auf $\Gamma_N(t)$ gilt.

2.2 Herleitung der Variationsungleichung

Angenommen, die obige Aussage sei nicht erfüllt. Das heißt unter Berücksichtigung der Stetigkeit der auftretenden Funktionen kann eine Teilmenge $\tilde{\Gamma} \subset \Gamma_N(t)$ mit positivem Hausdorff-Maß und ein geeignetes $v \in \mathcal{K}(t)$ gefunden werden, sodass

$$(f_N(t) - \mathbf{v} \cdot \mathbf{n})(v - u) < 0 \quad \text{auf } \tilde{\Gamma}$$

gilt, während $v = u$ auf $\Gamma_S(t) \cup (\Gamma_N(t) \setminus \tilde{\Gamma})$ gilt. Mit dieser Wahl ergibt sich jedoch für (2.6) der Widerspruch

$$\int_{\tilde{\Gamma}} (f_N(t) - \mathbf{v} \cdot \mathbf{n})(v - u) ds_x < 0 \quad \forall v \in \mathcal{K}(t)$$

und es folgt die Gültigkeit der Neumann-Randbedingung (1.17).

Letztlich bleibt noch die Signorini-Randbedingung (1.18) zu beweisen.

Unter Berücksichtigung der eben gezeigten Gültigkeit der Neumann-Randbedingung, reduziert sich die Gleichung (2.6) auf:

$$\int_{\Gamma_S(t)} \mathbf{v} \cdot \mathbf{n}(v - u) ds_x \leq 0 \quad \forall v \in \mathcal{K}(t). \quad (2.7)$$

Zu zeigen ist, dass $\mathbf{v} \cdot \mathbf{n} = 0$ auf $\Gamma_S(t)$ gilt, wenn $u < 0$ auf $\Gamma_S(t)$ ist.

Sei $u(x, t) < 0$ für einen Punkt x in einer Teilmenge $\tilde{\Gamma}_\varepsilon = \mathcal{B}_\varepsilon(x) \cap \Gamma_S(t)$ mit positivem Hausdorff-Maß. Es wird angenommen, dass $\mathbf{v}(x, t) \cdot \mathbf{n} \neq 0$ und dass ε klein genug ist, sodass sowohl u als auch $\mathbf{v} \cdot \mathbf{n}$ ihr Vorzeichen auf $\tilde{\Gamma}_\varepsilon$ nicht ändern.

Konstruiere zunächst eine zulässige Testfunktion $v \in \mathcal{K}(t)$, sodass $u < v \leq 0$ auf $\tilde{\Gamma}_\varepsilon$, d.h. $(v - u) > 0$ auf $\tilde{\Gamma}_\varepsilon$. Überall sonst auf $\Gamma_S(t)$ sei $v = u$. Aus (2.7) folgt dann $\mathbf{v}(x, t) \cdot \mathbf{n} < 0$. Falls jedoch andererseits $v \in \mathcal{K}(t)$ mit $v < u$ auf $\tilde{\Gamma}_\varepsilon$, d.h. $(v - u) < 0$ auf $\tilde{\Gamma}_\varepsilon$, und $v = u$ überall sonst auf $\Gamma_S(t)$ gewählt wird, ergibt sich $\mathbf{v}(x, t) \cdot \mathbf{n} > 0$ und damit ein Widerspruch. Somit muss $\mathbf{v} \cdot \mathbf{n} = 0$ sein, wenn $u < 0$ ist. Da $u \in \mathcal{K}(t)$, gilt $u(\mathbf{v} \cdot \mathbf{n}) = 0$ auf $\Gamma_S(t)$. Wenn schließlich $\mathbf{v} \cdot \mathbf{n} < 0$ auf $\tilde{\Gamma}_\varepsilon \subset \Gamma_S(t)$ erfüllt ist, dann wähle in Anlehnung an die obige Vorgehensweise $v < 0$ auf $\tilde{\Gamma}_\varepsilon$ und $v = 0$ auf $\Gamma_S(t) \setminus \tilde{\Gamma}_\varepsilon$. Durch diese Wahl wird das Integral aus (2.7) positiv:

$$\int_{\tilde{\Gamma}_\varepsilon} \mathbf{v} \cdot \mathbf{n}(v - u) ds_x = \int_{\tilde{\Gamma}_\varepsilon} \underbrace{(\mathbf{v} \cdot \mathbf{n})}_{<0} \underbrace{v}_{<0} ds_x - \int_{\tilde{\Gamma}_\varepsilon} \underbrace{(\mathbf{v} \cdot \mathbf{n})}_{=0} u ds_x > 0.$$

Das führt neuerlich auf einen Widerspruch und auf die zusätzliche Bedingung $\mathbf{v} \cdot \mathbf{n} \geq 0$ auf $\Gamma_S(t)$, womit die Gültigkeit der Signorini-Randbedingung (1.18) bewiesen wäre. ■

2 Analysis

Auf ähnliche Weise kann eine entsprechende Variationsungleichung für den physikalischen Druck $p(t)$ aufgestellt werden.

Analog zu (2.1) und (2.2) werden die folgenden beiden Definitionen eingeführt:

$$p_D(t) \in \{v = \text{tr}_{\Gamma_D(t)} w : w \in H^1(\Omega) \wedge \text{tr}_{\Gamma_S(t)} w \leq 0 \text{ f.ü. auf } \Gamma_S(t)\} \quad (2.8)$$

und die nichtleere, abgeschlossene und konvexe Menge

$$\mathcal{K}_0(t) := \{v \in H^1(\Omega) : \text{tr}_{\Gamma_D(t)} v = p_D(t) \wedge \text{tr}_{\Gamma_S(t)} v \leq 0 \text{ f.ü. auf } \Gamma_S(t)\}. \quad (2.9)$$

Dann heißt $p(x, t)$ für $(x, t) \in \Omega \times (0, T]$ schwache Lösung des Randwertproblems (1.15) - (1.18) genau dann, wenn $p(t) \in \mathcal{K}_0(t)$ und

$$\int_{\Omega} \theta(p(x, t))_t (v(x) - p(x, t)) \, dx + \int_{\Omega} k_r(\theta(p(x, t))) \nabla p(x, t) \nabla (v(x) - p(x, t)) \, dx \geq \quad (2.10)$$

$$\int_{\Omega} k_r(\theta(p(x, t))) e_z \nabla (v(x) - p(x, t)) \, dx - \int_{\Gamma_N(t)} f_N(t) (v(x) - p(x, t)) \, ds_x \quad \forall v \in \mathcal{K}_0(t).$$

Bemerkung 2.3. *Damit die schwache Variationsungleichung (2.10) Sinn macht, muss vorausgesetzt werden, dass eine Lösung $p \in L_2(0, T; H^1(\Omega))$ regulär genug ist, sodass $\theta(p)_t \in L_2(\Omega)$ fast überall auf $(0, T]$ gilt, siehe Bemerkung 2.1.*

Die Äquivalenz der Variationsungleichungen (2.3) und (2.10) liefert der folgende Satz, siehe auch [6].

Satz 2.16. *Es seien $\theta : \mathbb{R} \rightarrow \mathbb{R}$ und $k_r : \theta(\mathbb{R}) \rightarrow (0, 1]$ beschränkt und monoton wachsend. Die Kirchhoff-Transformation $\kappa : \mathbb{R} \rightarrow \mathbb{R}$ sei definiert wie in (1.9) und es sei $u_D(t) := \kappa(p_D(t))$.*

Dann löst $u(t) = \kappa(p(t))$ die Variationsungleichung (2.3), wenn $p(t)$ die Variationsungleichung (2.10) löst.

Wenn darüberhinaus $k_r \geq c$ für ein $c > 0$ (nicht-entarteter Fall) und $\Gamma_S(t) = \emptyset$ gilt, dann sind die Variationsungleichungen (2.3) und (2.10) äquivalent. Das heißt:

$u(t)$ erfüllt die Variationsungleichung (2.3) genau dann, wenn $p(t) = \kappa^{-1}(u(t))$ die Variationsungleichung (2.10) erfüllt.

Für den Beweis wird eine schwache Formulierung der Kettenregel und die Interpretation der Kirchhoff-Transformation (1.9) als Superpositionsoperator im $H^1(\Omega)$ und in Spurräumen benötigt. Dazu werden im Anschluss die Definition sowie einige Eigenschaften eines Superpositionsoperators angegeben, für eine allgemeinere Theorie sei auf [6, 7, 4] verwiesen.

2.2 Herleitung der Variationsungleichung

Definition 2.17 (Superpositionsoperatoren). Sei p eine reellwertige Funktion, die auf einer Teilmenge $S \subset \mathbb{R}^d$ fast überall bezüglich einem geeigneten Maß definiert ist. $\kappa : \mathbb{R} \rightarrow \mathbb{R}$ sei eine reelle Funktion. Durch die punktweise Anwendung von κ auf p

$$(\kappa_S(p))(x) := \kappa(p(x))$$

für x fast überall auf S , wird der Superpositionsoperator

$$\kappa_S : p \rightarrow \kappa(p)$$

definiert. Sei X ein normierter Raum, der aus einer Teilmenge aller messbaren Funktionen auf der offenen Menge S besteht. Wenn der Superpositionsoperator die Bedingung $\kappa_S(p) \in X$ für alle $p \in X$ erfüllt, dann wirkt der Superpositionsoperator auf dem Raum X . In diesem Fall bezeichnet

$$\kappa_X : X \rightarrow X$$

die Einschränkung von κ_S auf den Raum X und κ_X heißt Superpositionsoperator auf X . Analog dazu werden Superpositionsoperatoren, die zwischen zwei Räumen X_1 und X_2 wirken, definiert.

Bemerkung 2.4. Das in der obigen Definition geforderte „geeignete“ Maß ist in diesem Kontext durch das Lebesgue-Maß auf einem Lipschitz-Gebiet $S = \Omega \subset \mathbb{R}^d$ bzw. durch das Hausdorff-Maß auf einer Teilmenge $S = \Sigma \subset \partial\Omega$ des Randes gegeben.

Lemma 2.18. Sei $\Omega \subset \mathbb{R}^d$ eine offene Menge, $(k_r \circ \theta) \in L_\infty(\mathbb{R})$ und die Kirchhoff-Transformation $\kappa : \mathbb{R} \rightarrow \mathbb{R}$ sei definiert wie in (1.9). Dann folgt für $p \in H^1(\Omega)$, dass $u = \kappa_\Omega(p) \in H^1(\Omega)$ und dass die Kettenregel (1.11) in $L_2((\Omega))^d$ gilt. Weiters folgt die Bedingung:

$$\|u\|_{H^1(\Omega)} \leq \|k_r \circ \theta\|_{L_\infty(\Omega)} \|p\|_{H^1(\Omega)}.$$

Umgekehrt, falls $(k_r \circ \theta) \in L_\infty(\mathbb{R})$ und zusätzlich eine Konstante $c > 0$ existiert, sodass $k_r(s) \geq c$ für fast alle $s \in \mathbb{R}$, dann gilt auch

$$c \|p\|_{H^1(\Omega)} \leq \|u\|_{H^1(\Omega)}. \tag{2.11}$$

Mit Bedingung (2.11) folgt für $u \in H^1(\Omega)$ schließlich $p = (\kappa^{-1})_\Omega(u) \in H^1(\Omega)$.

Für den Beweis von Lemma 2.18 siehe [6].

Beweis von Satz 2.16: Es sei $p(t) \in L_2(0, T; H^1(\Omega))$ eine Lösung von (2.10) für alle $t \in [0, T]$.

Mit Lemma 2.18 folgt, dass $u(t) = \kappa(p(t)) \in H^1(\Omega)$ für fast alle $t \in (0, T]$ gilt und außerdem ist

$$\|u(t)\|_{H^1(\Omega)} \leq \|k_r \circ \theta\|_{L_\infty(\Omega)} \|p(t)\|_{H^1(\Omega)} < \infty \quad \text{für fast alle } t.$$

2 Analysis

Damit kann sofort $u \in L_2(0, T; H^1(\Omega))$ gefolgert werden.
Weiters gilt

$$\theta(p(t)) = \theta(\kappa^{-1}(\kappa(p(t)))) = \theta(\kappa^{-1}(u(t))) = M(u(t)) \in L_\infty(\Omega) \quad \text{für fast alle } t$$

bzw.

$$k_r(\theta(p(t))) = k_r(M(u(t))) \in L_\infty(\Omega) \quad \text{für fast alle } t.$$

Die Messbarkeit der Sättigung θ und der relativen Durchlässigkeit k_r wird auf die zusammengesetzten Funktionen übertragen. Mit der Kettenregel folgt schließlich eine zur Variationsungleichung (2.10) äquivalente Formulierung bezüglich u :

$$\begin{aligned} \int_{\Omega} M(u(t))_t (v - p(t)) \, dx + \int_{\Omega} \nabla u(t) \nabla (v - p(t)) \, dx &\geq \\ &\geq \int_{\Omega} k_r(M(u(t))) e_z \nabla (v - p(t)) \, dx - \int_{\Gamma_N(t)} f_N(t) (v - p(t)) \, ds_x \quad \forall v \in \mathcal{K}_0(t). \end{aligned}$$

Obige Formulierung entspricht der Variationsungleichung (2.3), bis auf die unterschiedliche Wahl der Testfunktionen $(v - p)$ mit $v \in \mathcal{K}_0(t)$. Jedoch kann in der Folge gezeigt werden, dass die Menge der in (2.3) angewendeten Testfunktionen

$$\mathcal{K}(t) - u(t) = \{v \in H^1(\Omega) : v \geq u_c - u \wedge \text{tr}_{\Gamma_D(t)} v = 0 \wedge \text{tr}_{\Gamma_S(t)} v \leq -\text{tr}_{\Gamma_S(t)} u(t)\}$$

eine Teilmenge der Menge der in (2.10) eingesetzten Testfunktionen

$$\mathcal{K}_0(t) - p(t) = \{v \in H^1(\Omega) : \text{tr}_{\Gamma_D(t)} v = 0 \wedge \text{tr}_{\Gamma_S(t)} v \leq -\text{tr}_{\Gamma_S(t)} p(t)\}$$

ist. Unter Anwendung der Kommutativität von Superpositionsoperatoren, der punktweisen Definition der Kirchhoff-Transformation und der Definition von $u(t) := \kappa(p(t))$ ergibt sich:

$$\text{tr}_{\Gamma_S(t)} p(t) \leq \kappa(\text{tr}_{\Gamma_S(t)} p(t)) = \text{tr}_{\Gamma_S(t)} \kappa(p(t)) = \text{tr}_{\Gamma_S(t)} u(t) \leq 0 \quad \text{fast überall auf } \Gamma_S(t)$$

und damit die Behauptung; siehe auch Definition 2.17 bzw. [6].

Schlussendlich bleibt zu zeigen, dass aus $p(t) \in \mathcal{K}_0(t)$ die Bedingung $u(t) \in \mathcal{K}(t)$ folgt.

Betrachte zunächst $p_D(t)$, definiert wie in (2.8). Die Funktion $u_D(t) := \kappa(p_D(t))$ fast überall auf $\Gamma_D(t)$ ist folglich eine zulässige Dirichlet-Bedingung für die Menge $\mathcal{K}(t)$, da in (2.1) enthalten.

Mit der Kommutativität von Superpositionsoperatoren folgt aus $\text{tr}_{\Gamma_D(t)}p(t) = p_D(t)$ die geforderte Bedingung:

$$\text{tr}_{\Gamma_D(t)}u(t) = \text{tr}_{\Gamma_D(t)}\kappa(p(t)) = \kappa(\text{tr}_{\Gamma_D(t)}p(t)) = \kappa(p_D(t)) = u_D(t).$$

Im Hinblick auf den Beweis der Äquivalenz kann beobachtet werden, dass für den Fall $\Gamma_S(t) = \emptyset$, die Mengen der Testfunktionen $\mathcal{K}(t) - u(t)$ und $\mathcal{K}_0(t) - p(t)$ mit dem Raum $H_{\Gamma_D}^1(\Omega)$ übereinstimmen, siehe [6].

■

2.3 Implizit-explizite Zeitdiskretisierung

Sei $0 = t_0 < t_1 < \dots < t_N = T$ eine Zerlegung des Intervalls $[0, T]$ mit den Zeitschrittweiten $\tau_n := t_n - t_{n-1}$, $n \in \{1, \dots, N\}$.

Für die Diskretisierung von $M(u(t_n))_t$ aus der Variationsungleichung (2.3) wird der rückwärtige Euler-Differenzenquotient

$$M(u(t_n))_t \approx \frac{M(u(t_n)) - M(u(t_{n-1}))}{\tau_n}$$

verwendet.

$\nabla u(t)$ wird auf der linken Seite der Variationsungleichung (2.3) implizit $\nabla u(t_n)$ und $k_r(M(u(t)))$ wird auf der rechten Seite explizit $k_r(M(u(t_{n-1})))$ gesetzt.

Die zugehörige Approximation von $u(t_n)$ wird mit u^n für $n \in \{1, \dots, N\}$ bezeichnet. Als Anfangsbedingung wird $u^0 = u(0) \in H^1(\Omega)$ gesetzt.

Dann lautet die diskrete Variationsungleichung:

Gesucht sind $u^n \in \mathcal{K}(t_n)$ für $n \in \{1, \dots, N\}$, sodass

$$\begin{aligned} \int_{\Omega} M(u^n)(v - u^n) dx + \tau_n \int_{\Omega} \nabla u^n \nabla(v - u^n) dx &\geq \\ &\geq \int_{\Omega} M(u^{n-1})(v - u^n) dx + \tau_n \int_{\Omega} k_r(M(u^{n-1}))e_z \nabla(v - u^n) dx \quad \forall v \in \mathcal{K}(t_n). \end{aligned} \tag{2.12}$$

Da im Anschluss nur räumliche Probleme betrachtet werden, werden folgende Abkürzungen eingeführt:

Für ein gegebenes $n \in \{1, \dots, N\}$ seien:

2 Analysis

$$\begin{aligned}\mathcal{K} &:= \mathcal{K}(t_n), & u_D &:= u_D(t_n), \\ \Gamma_D &:= \Gamma_D(t_n), & \Gamma_S &:= \Gamma_S(t_n), & \Gamma_N &:= \Gamma_N(t_n).\end{aligned}$$

Die rechte Seite der diskreten Variationsungleichung (2.12) ist durch ein lineares Funktional l auf $\mathcal{K} \subset H^1(\Omega)$ gegeben, das wie folgt definiert ist:

$$l(v) := \int_{\Omega} M(u^{n-1})v \, dx + \tau_n \int_{\Omega} k_r(M(u^{n-1}))e_z \nabla v \, dx \quad \forall v \in H^1(\Omega). \quad (2.13)$$

Die Linearform l ist stetig, d.h. ein Element von $(H^1(\Omega))'$, wenn $M : [u_c, \infty) \rightarrow \mathbb{R}$ und $k_r : M([u_c, \infty)) \rightarrow \mathbb{R}$ beschränkte Funktionen sind.

Die stetige, symmetrische Bilinearform $a(\cdot, \cdot)$, die gegeben ist durch

$$a(v, w) := \tau_n \int_{\Omega} \nabla v(x) \cdot \nabla w(x) \, dx \quad \forall v, w \in H^1(\Omega) \quad (2.14)$$

ist elliptisch im Teilraum $H_{\Gamma_D}^1(\Omega) := \{v \in H^1(\Omega) : \text{tr}_{\Gamma_D} v = 0\}$, d.h. es existiert eine Konstante $c > 0$, sodass

$$a(v, v) \geq c \|v\|_{H^1(\Omega)}^2 \quad \forall v \in H_{\Gamma_D}^1(\Omega).$$

Mit der obigen Notation und durch das Ersetzen von u^n durch u kann die diskrete Variationsungleichung (2.12) kompakter dargestellt werden in der Form:

$$u \in \mathcal{K} : \int_{\Omega} M(u)(v - u) \, dx + a(u, v - u) - l(v - u) \geq 0 \quad \forall v \in \mathcal{K}. \quad (2.15)$$

3 Konvexe Minimierung

In diesem Kapitel werden zunächst in Abschnitt 3.1 einige Definitionen und Aussagen über konvexe Mengen und konvexe Funktionale wiederholt, die im Anschluss benötigt werden, siehe dazu auch [24] oder [31]. Ziel dieses Kapitels und Aufgabe von Abschnitt 3.2 und Abschnitt 3.3 ist schließlich, zu zeigen, dass die Variationsungleichung (2.15) äquivalent zu einem konvexen Minimierungsproblem ist.

3.1 Eigenschaften konvexer Mengen und konvexer Funktionale

Definition 3.1 (Konvexe Menge). *Eine Teilmenge K eines reellen Vektorraumes V heißt konvex, wenn für alle $y, z \in K$ stets auch*

$$(1 - \lambda)y + \lambda z \in K \quad \forall \lambda \in \mathbb{R} \quad \text{mit } 0 \leq \lambda \leq 1$$

gilt.

Definition 3.2 (Konvexes Funktional). *Sei V ein reeller Vektorraum und $K \subset V$ eine konvexe Menge. Das Funktional $F : K \rightarrow \mathbb{R}$ heißt konvex, wenn*

$$F((1 - \lambda)y + \lambda z) \leq (1 - \lambda)F(y) + \lambda F(z) \quad \forall y, z \in K, \lambda \in (0, 1).$$

Falls die obige Ungleichung für alle $y \neq z$ mit „ $<$ “ erfüllt ist, dann heißt das Funktional F streng konvex.

Lemma 3.3. *Sei V ein reeller Vektorraum und $K \subset V$ eine konvexe Menge. Sei weiters $F : K \rightarrow \mathbb{R}$ ein konvexes Funktional und $u, v \in K$. Dann ist durch die reelle Funktion $g : [0, 1] \rightarrow \mathbb{R}$, definiert durch*

$$g(\lambda) = F(u + \lambda(v - u)) \quad \forall \lambda \in [0, 1],$$

eine konvexe Funktion gegeben.

Für den Beweis siehe zum Beispiel [24].

Das nachfolgende Lemma gibt schließlich Bedingungen für die Konvexität von reellen Funktionen und eine fundamentale Bedingung für konvexe Funktionen an. Diese können auf einfache Art und Weise aus der Definition abgeleitet werden, siehe [24].

3 Konvexe Minimierung

Lemma 3.4. Sei $I \subset \mathbb{R}$ ein Intervall. Dann gelten für die Funktion $f : I \rightarrow \mathbb{R}$ die folgenden Aussagen:

1. Die Funktion f ist genau dann konvex, wenn die Ungleichung

$$\frac{f(z) - f(z_1)}{z - z_1} \leq \frac{f(z_2) - f(z)}{z_2 - z}$$

für $z_1, z, z_2 \in I$ mit $z_1 < z < z_2$ erfüllt ist.

2. Wenn die Funktion f konvex ist, dann ist für ein beliebiges $z \in I$ der Differenzenquotient

$$\frac{f(y) - f(z)}{y - z}$$

eine monoton wachsende Funktion in $y \in I \setminus \{z\}$.

3.2 Konvexe Minimierungsprobleme

In diesem Abschnitt soll die Äquivalenz der diskreten Variationsungleichung (2.15) zu einem konvexen Minimierungsproblem gezeigt werden.

Ausgangspunkt dafür ist die Stammfunktion $\Phi : [u_c, \infty) \rightarrow \mathbb{R}$ von M , die definiert ist durch

$$\Phi(z) := \int_0^z M(s) ds \quad \forall z \in [u_c, \infty). \quad (3.1)$$

Obige Stammfunktion führt auf das Funktional $\phi : \mathcal{K} \rightarrow \mathbb{R}$:

$$\phi(v) := \int_{\Omega} \Phi(v(x)) dx \quad \forall v \in \mathcal{K}, \quad (3.2)$$

siehe auch [6, 25]. Zur Erinnerung:

Definition 3.5 (Lipschitz-Stetigkeit). Eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ heißt Lipschitz-stetig, wenn eine Konstante $L \geq 0$ existiert, sodass

$$|f(x) - f(y)| \leq L |x - y| \quad \forall x, y \in \mathbb{R}.$$

Ist die Lipschitz-Konstante $L < 1$, so nennt man die zugehörige Funktion eine Kontraktion.

Da M monoton wachsend und beschränkt ist, ist Φ konvex und Lipschitz-stetig, siehe [6].

3.2 Konvexe Minimierungsprobleme

Lemma 3.6. *Es sei Φ eine konvexe Funktion. Dann ist durch $\phi : \mathcal{K} \rightarrow \mathbb{R}$ ein konvexes Funktional gegeben. Wenn Φ weiters Lipschitz-stetig ist, so ist auch das Funktional ϕ Lipschitz-stetig und mit $C > 0$ gilt die Bedingung*

$$|\phi(v)| \leq C \|v\|_{H^1(\Omega)} \quad \forall v \in \mathcal{K}.$$

Beweis: Die Konvexität des Funktionals ϕ folgt direkt aus jener der Funktion Φ . Auch die Lipschitz-Stetigkeit von ϕ kann aus der Lipschitz-Stetigkeit von Φ bzgl. der Lipschitz-Konstanten L und Anwendung der Cauchy-Schwarz-Ungleichung gewonnen werden:

$$\begin{aligned} |\phi(v) - \phi(w)| &= \left| \int_{\Omega} (\Phi(v(x)) - \Phi(w(x))) dx \right| \leq \int_{\Omega} |\Phi(v(x)) - \Phi(w(x))| dx \leq \\ &\leq L \int_{\Omega} |v(x) - w(x)| dx \leq L |\Omega|^{1/2} \|v - w\|_{L^2(\Omega)} \leq C \|v - w\|_{H^1(\Omega)} \quad \forall v, w \in \mathcal{K} \end{aligned} \quad (3.3)$$

mit der Konstanten $C = L |\Omega|^{1/2}$. Durch Einsetzen von $w = 0 \in \mathcal{K}$ und $\phi(0) = 0$ in (3.3) folgt schließlich die geforderte Bedingung. ■

Aufgrund der Stetigkeit von M folgt überdies hinaus die Differenzierbarkeit von Φ mit $\Phi' = M$, siehe [6].

Definition 3.7 (Richtungsableitung). *Sei V ein normierter Raum, $S \subset V$, $F : S \rightarrow \mathbb{R}$ und $u \in S$, $v \in V$.*

1. *Falls ein $\varepsilon > 0$ existiert, sodass $u + hv \in S$ für alle $h \in [0, \varepsilon]$, dann heißt*

$$\partial_v F(u) := \lim_{h \downarrow 0} \frac{F(u + hv) - F(u)}{h} \quad (3.4)$$

für den Fall, dass dieser einseitige Grenzwert existiert, Richtungsableitung von F an der Stelle u in Richtung v .

2. *Wenn zusätzlich zu (3.4) ein $u' \in V'$ existiert, sodass*

$$\partial_v F(u) = \langle u', v \rangle_V \quad \forall v \in V \quad (3.5)$$

gilt, dann heißt F Gâteaux-differenzierbar im Punkt u mit der Gâteaux-Ableitung $F'(u) := u'$.

3 Konvexe Minimierung

3. Wenn u ein innerer Punkt aus S ist und zusätzlich zu (3.5) die Konvergenz in (3.4) gleichmäßig bezüglich der Elemente aus der Einheitskugel in V ist, dann heißt F Fréchet-differenzierbar und $F'(u)$ ist die Fréchet-Ableitung von F im Punkt u .

Da Φ konvex und differenzierbar ist, können Differentiation und Integration vertauscht werden und es folgt mit der obigen Definition der Richtungsableitung:

Lemma 3.8. *Es existiert für jedes $u, v \in \mathcal{K}$ die Richtungsableitung $\partial_{v-u}\phi(u)$, die in folgender Form geschrieben werden kann:*

$$\partial_{v-u}\phi(u) = \int_{\Omega} \Phi'(u(x)) (v(x) - u(x)) \, dx = \int_{\Omega} M(u(x)) (v(x) - u(x)) \, dx. \quad (3.6)$$

Beweis: Es seien $u, v \in \mathcal{K}$. Dann folgt aus der Konvexität der Menge \mathcal{K} , dass $u + \lambda(v - u) \in \mathcal{K}$ für $\lambda \in [0, 1]$. Setze $w := v - u$ und betrachte den folgenden Differenzenquotient:

$$\frac{\phi(u + \lambda w) - \phi(u)}{\lambda} = \int_{\Omega} \frac{\Phi(u(x) + \lambda w(x)) - \Phi(u(x))}{\lambda} \, dx \quad \text{für } \lambda \downarrow 0. \quad (3.7)$$

Im Weiteren wird nun Aussage (3.5) aus Lemma 3.4 verwendet. In Anbetracht dessen folgt die Abschätzung des Integranden aus (3.7)

$$\frac{\Phi(u(x) + \lambda w(x)) - \Phi(u(x))}{u(x) + \lambda w(x) - u(x)} \leq \frac{\Phi(u(x) + w(x)) - \Phi(u(x))}{u(x) + w(x) - u(x)}$$

bzw.

$$\frac{\Phi(u(x) + \lambda w(x)) - \Phi(u(x))}{\lambda} \leq \frac{\Phi(u(x) + w(x)) - \Phi(u(x))}{1} =: G(x)$$

für $w(x) \geq 0$, $\lambda \in (0, 1]$ und weiters

$$H(x) := \frac{\Phi(u(x) - w(x)) - \Phi(u(x))}{1} \leq \frac{\Phi(u(x) + \lambda w(x)) - \Phi(u(x))}{\lambda}$$

für $w(x) \geq 0$ und $\lambda \in (0, 1]$. Alles in allem ist der Integrand aus (3.7) damit durch die integrierbare Funktion $\max(|H(\cdot)|, |G(\cdot)|)$ auf Ω beschränkt, unabhängig von $\lambda \in (0, 1]$. Aufgrund der Differenzierbarkeit von Φ konvergieren die punktweisen Werte des Integranden

$$\frac{\Phi(u(x) + \lambda w(x)) - \Phi(u(x))}{\lambda}$$

gegen $\Phi'(u(x)) w(x)$ fast überall in Ω für $\lambda \downarrow 0$.

Mit dem Satz von der monotonen Konvergenz bzw. dem Satz von der dominierten Konvergenz folgt die Konvergenz des Integrals aus (3.7) und weiters die Behauptung des Lemmas mit $\Phi' = M$, siehe [6].

■

Mit Hilfe von Lemma 3.8 kann die Variationsungleichung (2.15) wie folgt umgeschrieben werden:

$$u \in \mathcal{K} : \quad \partial_{v-u}\phi(u) + a(u, v - u) - l(v - u) \geq 0 \quad \forall v \in \mathcal{K}. \quad (3.8)$$

Das quadratische Funktional $\mathcal{J} : H^1(\Omega) \rightarrow \mathbb{R}$, definiert durch

$$\mathcal{J}(v) := \frac{1}{2}a(v, v) - l(v) \quad \forall v \in H^1(\Omega) \quad (3.9)$$

ist streng konvex und stetig, siehe dazu [15].

Zusätzlich ist \mathcal{J} Fréchet-differenzierbar im Punkt $u \in H^1(\Omega)$ (siehe Definition 3.6) mit der Ableitung

$$\begin{aligned} \mathcal{J}'(u)(v) &= \partial_v \mathcal{J}(u) = \lim_{h \downarrow 0} \frac{\frac{1}{2}a(u + hv, u + hv) - l(u + hv) - \frac{1}{2}a(u, u) + l(u)}{h} = \\ &= \lim_{h \downarrow 0} \frac{\frac{1}{2}h^2 a(v, v) + h a(u, v) - h l(v)}{h} = a(u, v) - l(v) \quad \forall v \in H^1(\Omega). \end{aligned}$$

Folglich ist das Funktional $F : \mathcal{K} \rightarrow \mathbb{R}$, definiert durch

$$F(v) := \phi(v) + \mathcal{J}(v) \quad \forall v \in \mathcal{K}, \quad (3.10)$$

streng konvex und es existiert die Ableitung $\partial_{v-u}F(u)$ für alle $u, v \in \mathcal{K}$.

Insgesamt folgt die Darstellung der Variationsungleichung (2.15) in der folgenden Form:

Gesucht ist $u \in \mathcal{K}$, sodass

$$\partial_{v-u}F(u) \geq 0 \quad \forall v \in \mathcal{K}. \quad (3.11)$$

Lemma 3.9 gibt schließlich die Verbindung zur konvexen Minimierung an.

Lemma 3.9. *Sei V ein reeller Vektorraum, $K \subset V$ eine konvexe Menge und $F : K \rightarrow \mathbb{R}$ ein konvexes Funktional, dessen Ableitung $\partial_{v-u}F(u)$ für alle $u, v \in K$ existiert. Dann ist*

3 Konvexe Minimierung

$$u \in K : \quad \partial_{v-u}F(u) \geq 0 \quad \forall v \in K \quad (3.12)$$

äquivalent zu

$$u \in K : \quad F(u) \leq F(v) \quad \forall v \in K. \quad (3.13)$$

Beweis: Angenommen, es gelte (3.12). Dann gilt unter Anwendung von Lemma 3.3 und Aussage (3.5) von Lemma 3.4 für ein beliebiges $v \in \mathcal{K}$ und $\lambda \in (0, 1]$ die Abschätzung

$$F(v) - F(u) = \frac{F(u + (v - u)) - F(u)}{1} \geq \frac{F(u + \lambda(v - u)) - F(u)}{\lambda}.$$

Mit der Definition der Richtungsableitung und $\lambda \downarrow 0$ folgt dann sofort

$$F(v) - F(u) \geq \partial_{v-u}F(u) \geq 0 \quad \Leftrightarrow \quad F(v) \geq F(u) \quad \forall v \in \mathcal{K}$$

und damit die Gültigkeit des Minimierungsproblems (3.13).

Umgekehrt, wenn (3.13) erfüllt ist, dann ergibt sich mit $\lambda \in (0, 1]$

$$F(u + \lambda(v - u)) \geq F(u) \quad \forall v \in \mathcal{K}$$

und schließlich

$$\frac{F(u + \lambda(v - u)) - F(u)}{\lambda} \geq 0 \quad \forall v \in \mathcal{K}.$$

Mit $\lambda \downarrow 0$ folgt schlussendlich die Variationsungleichung (3.12). ■

Zur Wiederholung:

Definition 3.10. Sei V ein topologischer Raum und $K \subset V$ eine nicht leere, abgeschlossene und konvexe Menge.

1. Ein Funktional $F : K \rightarrow \mathbb{R} \cup \{+\infty\}$ heißt unterhalbstetig, wenn

$$\liminf_{w \rightarrow v} F(w) \geq F(v) \quad \forall v \in K$$

gilt.

2. Ein konvexes Funktional $F : K \rightarrow \mathbb{R} \cup \{+\infty\}$ heißt eigentlich, wenn es nicht identisch $+\infty$ ist. In diesem Fall heißt

$$\text{dom } F := \{v \in K : F(v) < \infty\}$$

Definitionsbereich von F .

3.3 Eindeutige Lösbarkeit und Äquivalenz

Es kann gezeigt werden, dass das in (3.10) definierte Funktional F konvex, eigentlich, stetig und koerziv ist, d.h., dass für jede Folge $(u_n) \subset \mathcal{K}$ mit $\|u_n\| \rightarrow \infty$, auch $F(u_n) \rightarrow \infty$ gilt. Aus diesem Grund können bereits bekannte allgemeine Existenz- und Eindeutigkeitsaussagen für konvexe Minimierungsprobleme in reflexiven Banachräumen angewendet werden, siehe zum Beispiel [6] oder für eine allgemeinere Theorie auch [15].

Es kann nun der folgende Satz angegeben werden, dessen Beweis in [15] zu finden ist.

Satz 3.11. *Sei V ein reflexiver Banachraum, $K \subset V$ sei eine nicht leere, abgeschlossene und konvexe Menge und das Funktional $F : K \rightarrow \mathbb{R} \cup \{+\infty\}$ sei konvex, unterhalbstetig, eigentlich und koerziv.*

Dann besitzt das Minimierungsproblem

$$u \in K : \quad F(u) \leq F(v) \quad \forall v \in K$$

eine Lösung. Diese Lösung ist eindeutig, wenn das Funktional F streng konvex ist.

3.3 Eindeutige Lösbarkeit und Äquivalenz

Ziel dieses Abschnitts ist einerseits, die eindeutige Lösbarkeit der diskreten Variationsungleichung (2.15) und andererseits die Äquivalenz zu einem konvexen Minimierungsproblem zu zeigen.

Unter Verwendung von Satz 3.11 gilt die folgende Aussage:

Satz 3.12. *Seien die Menge $\mathcal{K} \subset H^1(\Omega)$, die Bilinearform $a(\cdot, \cdot)$ und die Linearform $l(\cdot)$ definiert wie in (2.2), (2.14) bzw. (2.13).*

Die Sättigung $M : [u_c, \infty) \rightarrow \mathbb{R}$ sei beschränkt, monoton wachsend, stetig und die relative Durchlässigkeit $k_r : M(\mathbb{R}) \rightarrow \mathbb{R}$ sei beschränkt und monoton wachsend.

Dann besitzt die Variationsungleichung (2.15) eine eindeutige Lösung. Genauer gesagt ist die diskrete Variationsungleichung (2.15) äquivalent zum Minimierungsproblem

$$u \in \mathcal{K} : \quad \mathcal{J}(u) + \phi(u) \leq \mathcal{J}(v) + \phi(v) \quad \forall v \in \mathcal{K} \tag{3.14}$$

mit den beiden Funktionalen \mathcal{J} und ϕ definiert wie in (3.9) bzw. (3.2).

Beweis: Es genügt zu zeigen, dass das Funktional $F = \mathcal{J} + \phi$ auf \mathcal{K} die Eigenschaften aus Satz 3.11 erfüllt. Der Hilbertraum $H^1(\Omega)$ ist reflexiv und wie bereits bewiesen ist die Menge $\mathcal{K} \subset H^1(\Omega)$ nicht leer, abgeschlossen und konvex. Weiters ist das Funktional $F : \mathcal{K} \rightarrow \mathbb{R}$ streng konvex aufgrund der strengen Konvexität des Funktionals \mathcal{J} und der Konvexität von ϕ , und F ist eigentlich. Da \mathcal{J} und ϕ stetig sind, ist auch F stetig und damit insbesondere auch unterhalbstetig. Mit Satz 3.11, der H^1 -Elliptizität der Bilinearform $a(\cdot, \cdot)$, der

3 Konvexe Minimierung

Beschränktheit des Funktionals $l(\cdot)$ und Lemma 3.6 folgt schließlich die Koerzivität des Funktionals F :

$$\begin{aligned} |\mathcal{J}(v) + \phi(v)| &= \left| \frac{1}{2}a(v, v) - l(v) + \phi(v) \right| \geq \frac{1}{2}a(v, v) - |l(v)| - |\phi(v)| \geq \\ &\geq \frac{1}{2}c \|v\|_{H^1(\Omega)}^2 - \left(\|l\|_{(H^1(\Omega))'} + C \right) \|v\|_{H^1(\Omega)} \rightarrow \infty \end{aligned}$$

für $\|v\|_{H^1(\Omega)} \rightarrow \infty$, $v \in \mathcal{K}$. Somit sind alle Eigenschaften aus Satz 3.11 erfüllt und es folgt die eindeutige Lösbarkeit des Minimierungsproblems (3.14). Die Äquivalenz zur Variationsungleichung ergibt sich aus Lemma 3.9. ■

Bemerkung 3.1. Satz 3.12 kann für den Fall einer ortsabhängigen Porosität $n(\cdot)$ und einer ortsabhängigen hydraulischen Leitfähigkeit $K_h(\cdot)$ erweitert werden. Dazu muss zusätzlich gelten, dass $n(\cdot)$ nicht-negativ und beschränkt ist und $K_h(\cdot)$ muss die Bedingung

$$c \leq K_h(\cdot) \leq C \quad \text{mit } c, C > 0$$

erfüllen, damit Stetigkeit und Elliptizität der Bilinearform $a(\cdot, \cdot)$ erhalten bleiben.

4 Finite-Elemente-Methode

In diesem Kapitel soll zunächst eine geeignete Diskretisierung mit finiten Elementen für das Minimierungsproblem (3.14) angegeben werden, wobei das zugrundeliegende Modellgebiet Ω als Teilmenge des \mathbb{R}^2 angenommen und die Notation zum größten Teil aus [33] übernommen wird. Im Anschluss daran wird dieses resultierende diskrete Minimierungsproblem für eine diskrete physikalische Variable umformuliert. Abschnitt 4.2 behandelt schließlich einerseits die Konvergenz der Approximation des verallgemeinerten Drucks und andererseits, die aus praktischer Sicht interessantere, Konvergenz der Näherungslösungen des physikalischen Drucks und der Sättigung.

4.1 Diskretisierung

Sei $\Omega \subset \mathbb{R}^2$ ein gegebenes polygonales Gebiet. $(\mathcal{T}_j)_{j \geq 0}$ sei eine Folge von Unterteilungen von Ω

$$\bar{\Omega} = \bar{\mathcal{T}}_j = \bigcup_{l=1}^{N_j} \bar{\tau}_l \quad (4.1)$$

in finite Dreieckselemente τ_l .

Es wird dabei von der Annahme ausgegangen, dass die Unterteilung \mathcal{T}_j zulässig ist. Das heißt, dass der Durchschnitt zweier Dreiecke τ_l, τ_k mit $l \neq k$ in \mathcal{T}_j entweder leer, ein gemeinsamer Knoten oder eine gemeinsame Kante ist.

Die Menge aller Dreiecksknoten werde mit $\mathcal{M} := \{x_k\}_{k=1}^M$ bezeichnet. Weiters ist für jedes finite Element τ_l durch

$$\Delta_l := \int_{\tau_l} dx$$

das Volumen, durch

$$h_l := \Delta_l^{1/2}$$

die lokale Maschenweite, sowie durch

$$d_l := \sup_{x,y \in \tau_l} |x - y|$$

der Durchmesser, der durch die längste Kante des jeweiligen finiten Elements angenommen wird, definiert. Darüber hinaus wird der Radius des größten im finiten Element τ_l enthaltenen Kreises mit r_l bezeichnet.

4 Finite-Elemente-Methode

Die Folge von Unterteilungen (4.1) sei außerdem formregulär. Das bedeutet, dass die Durchmesser d_l der finiten Elemente τ_l der Unterteilung \mathcal{T}_j durch ein festes Vielfaches der Radien r_l gleichmäßig abgeschätzt werden können. Damit gilt

$$d_l \leq c \cdot r_l \quad \forall l = 1, \dots, N_j \quad (4.2)$$

mit einer von \mathcal{T}_j unabhängigen Konstanten c .

Für eine konsistente Diskretisierung muss vorausgesetzt werden, dass jeder Schnittpunkt von zwei Abschlüssen der Teilmengen Γ_D , Γ_N und Γ_S des Randes $\partial\Omega$ in der Menge aller Dreiecksknoten \mathcal{M} enthalten ist. Außerdem wird angenommen, dass Γ_D und $\Gamma_S \cup \Gamma_D$ abgeschlossen sind. Weiters sei $\mathcal{M}_D := \mathcal{M} \cap \Gamma_D$ bzw. $\mathcal{M}_S := \mathcal{M} \cap \Gamma_S$.

Als Ansatzraum wird $\mathcal{S}_j := \mathcal{S}_h^1(\mathcal{T}_j) \subset H^1(\Omega)$, der Raum der in $H^1(\Omega)$ stetigen Funktionen, die linear auf jedem Dreieck $\tau \in \mathcal{T}_j$ sind, gewählt. Analog dazu wird auch der Raum $\mathcal{S}_j^D \subset H_{\Gamma_D}^1(\Omega) = \{v \in H^1(\Omega) : v|_{\Gamma_D} = 0\}$ definiert.

Die Räume \mathcal{S}_j und \mathcal{S}_j^D werden jeweils durch die nodale Basis

$$\Lambda_j := \{\lambda_p^{(j)} : p \in \mathcal{M}\} \quad \text{bzw.} \quad \Lambda_j^D := \{\lambda_p^{(j)} : p \in \mathcal{M} \setminus \mathcal{M}_D\}$$

aufgespannt.

Die Definition des endlich dimensionalen Analogons zur konvexen Menge \mathcal{K} , definiert in (2.2), wird unter der Annahme durchgeführt, dass die Dirichlet-Randbedingung u_D in jedem Knoten $p \in \mathcal{M}_D$ stetig ist.

Die konvexe Menge $\mathcal{K}_j \subset \mathcal{S}_j$ wird wie folgt definiert:

$$\mathcal{K}_j := \{v \in \mathcal{S}_j : v(p) \geq u_c \forall p \in \mathcal{M} \wedge v(p) = u_D(p) \forall p \in \mathcal{M}_D \wedge v(p) \leq 0 \forall p \in \mathcal{M}_S\}. \quad (4.3)$$

\mathcal{K}_j ist als Teilmenge des endlich dimensionalen Raumes \mathcal{S}_j wiederum nicht leer und abgeschlossen.

Bemerkung 4.1. *Die Folgerung $\mathcal{K}_j \subset \mathcal{K}$ ist im Allgemeinen falsch, da sich die Dirichlet-Randbedingungen in \mathcal{K}_j im Allgemeinen von denen in \mathcal{K} unterscheiden.*

In weiterer Folge wird das Integral in der Definition des Funktionals ϕ in (3.2) durch eine aus der \mathcal{S}_j -Interpolation des Integranden $\Phi(v)$ hervorgehenden Quadraturformel approximiert. Es ergibt sich damit das diskrete Funktional $\phi_j : \mathcal{S}_j \rightarrow \mathbb{R} \cup \{+\infty\}$, das definiert ist durch

$$\phi_j(v) := \sum_{p \in \mathcal{M}} \Phi(v(p)) h_p \quad \forall v \in \mathcal{S}_j \quad (4.4)$$

mit den positiven Gewichten

$$h_p := \int_{\Omega} \lambda_p^{(j)}(x) dx. \quad (4.5)$$

Die Eigenschaften der Funktionale im Kontinuierlichen, also die Konvexität und die Lipschitz-Stetigkeit, sollen auch für die diskreten Funktionale aufrechterhalten bleiben, siehe Kapitel 3. Im Hinblick auf die spätere Konvergenzanalyse gibt das nachfolgende Lemma die entscheidenden Eigenschaften der diskreten Funktionale ϕ_j an. Der zugehörige Beweis ist in [6] zu finden.

Lemma 4.1. *Die in (3.1) definierte Funktion M sei monoton wachsend und beschränkt. Dann ist das Funktional ϕ_j aus (4.4) konvex und Lipschitz-stetig im Gebiet*

$$\text{dom } \phi_j = \{v \in \mathcal{S}_j : v(p) \geq u_c \forall p \in \mathcal{M}\}$$

mit einer von $j \geq 0$ unabhängigen Lipschitz-Konstanten.

ϕ_j ist außerdem unterhalbstetig, eigentlich und es erfüllt die folgende Abschätzung

$$\phi_j(v) \geq C \|v\|_{H^1(\Omega)} \quad \forall v \in \mathcal{S}_j$$

mit einer von $j \geq 0$ unabhängigen Konstanten $C \in \mathbb{R}$. Ferner gilt für $v_j \in \mathcal{S}_j$, $j \geq 0$, und $v \in H^1(\Omega)$:

$$v_j \rightharpoonup v, j \rightarrow \infty \implies \liminf_{j \rightarrow \infty} \phi_j(v_j) \geq \phi(v),$$

wobei $v_j \rightharpoonup v$ die schwache Konvergenz von v_j gegen v im $H^1(\Omega)$ bezeichnet.

An dieser Stelle kann die diskrete Form des Minimierungsproblems (3.14) wie folgt angegeben werden:

$$u_j \in \mathcal{K}_j : \quad \mathcal{J}(u_j) + \phi_j(u_j) \leq \mathcal{J}(v) + \phi_j(v) \quad \forall v \in \mathcal{K}_j. \quad (4.6)$$

Da \mathcal{K}_j , \mathcal{J} und ϕ_j die gleichen Eigenschaften wie \mathcal{K} , \mathcal{J} und ϕ aus Satz 3.11 besitzen, kann die eindeutige Lösbarkeit des Minimierungsproblems auf den Teilraum \mathcal{S}_j des Hilbertraumes $H^1(\Omega)$ übertragen werden.

Satz 4.2. *Das diskrete Minimierungsproblem (4.6) besitzt eine eindeutige Lösung.*

Das nächste Ziel ist eine Neuinterpretation des diskreten Minimierungsproblems (4.6) bezüglich diskreter physikalischer Variabler.

4 Finite-Elemente-Methode

Unter Anwendung einer speziellen Quadraturformel mit von der Kirchhoff-Transformation κ abhängigen Stützstellen auf den Term $k_r(\theta(p))$, kann das diskrete Minimierungsproblem (4.6) als eine Diskretisierung mit finiten Elementen der Richards-Gleichung (1.15) in physikalischen Variablen betrachtet werden.

Unter Anwendung von Lemma 3.8 und unter Voraussetzung der Stetigkeit der Sättigung $M : [u_c, \infty) \rightarrow \mathbb{R}$, folgt die Äquivalenz des diskreten Minimierungsproblems (4.6) mit der folgenden Variationsungleichung:

$$u_j \in \mathcal{K}_j : \sum_{q \in \mathcal{M}} M(u_j(q)) (v(q) - u_j(q)) h_q + a(u_j, v - u_j) - l(v - u_j) \geq 0 \quad \forall v \in \mathcal{K}_j \quad (4.7)$$

mit den positiven Gewichten h_q aus (4.5) und der konvexen Menge \mathcal{K}_j , definiert wie in (4.3). Die obige Variationsungleichung (4.7) kann damit als Diskretisierung der ursprünglichen Variationsungleichung (2.15) betrachtet werden.

Bemerkung 4.2. Für den Fall, dass $u_j(q) = u_c$ für $q \in \mathcal{M}$ gilt, folgt die physikalisch unrealistische Situation $\kappa^{-1}(u_j(q)) = -\infty$. Diese ergibt sich aus der Bedingung $v \geq u_c$ anstelle von $v > u_c$ in den Definitionen der konvexen Menge \mathcal{K} durch (2.2) bzw. der konvexen Menge \mathcal{K}_j durch (4.3). Die Abgeschlossenheit dieser beiden konvexen Mengen ist jedoch notwendig, um die Existenz einer Lösung des Minimierungsproblems zu garantieren, die Gleichheit in der obigen Bedingung tritt aber im ursprünglichen physikalischen Problem nicht ein.

Im Folgenden wird daher die Annahme getroffen, dass

$$u_j(q) > u_c \quad \forall q \in \mathcal{M}, \quad (4.8)$$

woraus die Reellwertigkeit der inversen Kirchhoff Transformation $\kappa^{-1}(u_j)$ gefolgert und die nachfolgende Definition angegeben werden kann.

Definition 4.3. Es sei $\mathcal{I}_{\mathcal{S}_j} : H^1(\Omega) \cap C(\bar{\Omega}) \rightarrow \mathcal{S}_j$ der stückweise lineare Interpolationsoperator, definiert durch

$$(\mathcal{I}_{\mathcal{S}_j} v)(q) = v(q) \quad \forall q \in \mathcal{M}, v \in H^1(\Omega) \cap C(\bar{\Omega}).$$

Unter der Annahme, dass die Bedingung (4.8) erfüllt ist, heißt $\mathcal{I}_{\mathcal{S}_j} \kappa : \mathcal{S}_j \rightarrow \mathcal{S}_j$ diskrete Kirchhoff-Transformation auf \mathcal{S}_j und $p_j := \mathcal{I}_{\mathcal{S}_j} \kappa^{-1}(u_j)$ diskreter physikalischer Druck bzgl. des diskreten Minimierungsproblems (4.6).

Um in der Folge eine diskrete Variationsungleichung für die physikalische Variable p_j angeben zu können, wird die Forderung $\kappa \in C^1(\mathbb{R})$ an die Kirchhoff-Transformation gestellt,

d.h. $k_r \circ \theta$ sei stetig. Im Falle der gewählten Brooks-Corey-Parameterfunktionen ist diese Bedingung erfüllt.

Aufgrund der Definition der transformierten Sättigung M in (1.10) gilt zunächst

$$M(u_j(q)) = \theta(p_j(q)).$$

Dementsprechend kann die Linearform $l(\cdot)$ aus (2.13) für eine diskrete physikalische Variable zurücktransformiert werden. Die verbleibende Schwierigkeit ist die Transformation der Bilinearform

$$a(u_j, w) = \int_{\Omega} \nabla u_j \nabla w \, dx, \quad w = v - u_j, \quad v \in \mathcal{K}_j \quad (4.9)$$

auf die physikalische Variable p_j .

Zur Erinnerung: Für das stetige Problem (1.15) wurde die Transformation unter Anwendung einer schwachen Kettenregel durchgeführt. Mit der Vorgehensweise wie in [8] und Verwendung des Mittelwertsatzes ergibt sich

$$\nabla u_j = D_l(p_j) \nabla p_j \quad \text{auf dem Dreieck } \tau_l. \quad (4.10)$$

Dabei ist die Transformation vom Referenzdreieck

$$\tau = \{\xi \in \mathbb{R}^2 : 0 \leq \xi_1 \leq 1, 0 \leq \xi_2 \leq 1 - \xi_1\} \subset \mathbb{R}^2$$

auf das Dreieck $\tau_l \in \mathcal{T}_j$ durch die affine Abbildung im \mathbb{R}^2

$$F_l : \xi \rightarrow x = J_l \xi + b_l$$

mit einer nicht singulären Matrix $J_l \in \mathbb{R}^{2 \times 2}$ und dem Vektor $b_l \in \mathbb{R}^2$ gegeben. Weiters sei D_l eine Diagonalmatrix, die definiert ist durch

$$D_l(p_j) := \begin{pmatrix} k_r(\theta(p_j(\bar{x}_1))) & 0 \\ 0 & k_r(\theta(p_j(\bar{x}_2))) \end{pmatrix}$$

mit den Punkten $\bar{x}_i = F_l(\bar{\xi}_i)$, $i = 1, 2$, wobei $\bar{\xi}_i$ durch den Mittelwertsatz eindeutig bestimmte Punkte auf den Kanten des Referenzdreiecks τ sind.

(4.10) kann nun als diskrete Kettenregel bzgl. der diskreten Kirchhoff-Transformation in \mathcal{S}_j betrachtet werden.

Definiere die nichtlineare Form

$$b(p_j, v) := \sum_{\tau_l \in \mathcal{T}_j} \int_{\tau_l} D_l(p_j) \nabla p_j \nabla v \, dx, \quad p_j, v \in \mathcal{S}_j. \quad (4.11)$$

4 Finite-Elemente-Methode

Weiters seien p_D die diskreten Dirichlet-Randwerte auf \mathcal{M}_D und die zugehörige diskrete abgeschlossene und konvexe Menge \mathcal{K}_j^0 sei wie folgt definiert:

$$\mathcal{K}_j^0 := \{v \in \mathcal{S}_j : v(q) = p_D(q) \forall q \in \mathcal{M}_D \wedge v(q) \leq 0 \forall q \in \mathcal{M}_S\}. \quad (4.12)$$

Schließlich kann das diskrete Problem für den diskreten physikalischen Druck angegeben werden:

$$p_j \in \mathcal{K}_j^0 : \sum_{q \in \mathcal{M}} \theta(p_j(q)) (v(q) - p_j(q)) h_q + b(p_j, v - p_j) - l(v - p_j) \geq 0 \quad \forall v \in \mathcal{K}_j^0. \quad (4.13)$$

Der nachfolgende Satz stellt das diskrete Gegenstück zu Satz 2.15 dar, siehe [6].

Satz 4.4. *Seien $\theta : \mathbb{R} \rightarrow \mathbb{R}$ und $k_r : \theta(\mathbb{R}) \rightarrow (0, 1]$ beschränkt, monoton wachsend und stetig. Die Kirchhoff-Transformation $\kappa : \mathbb{R} \rightarrow \mathbb{R}$ sei definiert wie in (1.9). Außerdem gilt $p_D = \kappa^{-1}(u_D)$ auf \mathcal{M}_D .*

Dann löst $u_j = \mathcal{I}_{\mathcal{S}_j} \kappa(p_j)$ die Variationsungleichung (4.7), wenn p_j die Variationsungleichung (4.13) löst. Umgekehrt löst $p_j = \mathcal{I}_{\mathcal{S}_j} \kappa^{-1}(u_j)$ die Variationsungleichung (4.13), wenn u_j die Variationsungleichung (4.7) unter der Bedingung (4.8) löst und $\Gamma_S(t) = \emptyset$.

Falls $k_r \geq c > 0$ und $\Gamma_S(t) = \emptyset$ gilt, dann sind die Variationsungleichungen (4.7) und (4.13) äquivalent, d.h., u_j erfüllt (4.7) genau dann, wenn $p_j = \mathcal{I}_{\mathcal{S}_j} \kappa^{-1}(u_j)$ der Variationsungleichung (4.13) genügt.

4.2 Konvergenz

In diesem Abschnitt soll zunächst eine Konvergenzaussage für den verallgemeinerten Druck angegeben werden. Die Aufgabe besteht also darin, die Konvergenz der Lösungen der diskreten Probleme (4.6) gegen die Lösung des stetigen Problems (3.14) zu zeigen.

Aus hydrologischer Sicht interessanter ist jedoch die Konvergenz der Näherungslösungen für die Sättigung und den physikalischen Druck, die im Anschluss näher betrachtet werden.

Die Konvergenzanalyse beruht auf der Annahme, dass die lokalen Maschenweiten h_j für die zugehörige Folge von Triangulierungen $(\mathcal{T}_j)_{j \geq 0}$ abnehmen. Zusätzlich sei die Folge von Triangulierungen $(\mathcal{T}_j)_{j \geq 0}$ formregulär, das heißt es gelte (4.2).

Es wird sich später herausstellen, dass Konvergenz nur dann garantiert werden kann, wenn die Dirichlet-Randwerte u_D als Spur einer gleichmäßig stetigen Funktion $w \in H^1(\Omega)$ auf Γ_D betrachtet werden können, siehe Satz 2.10. Es gelte also:

$$u_D = \text{tr}_{\Gamma_D} w \quad \text{für ein } w \in H^1(\Omega) \cap C(\overline{\Omega}). \quad (4.14)$$

Falls u_D stetig auf Γ_D ist, dann kann u_D zu einer stetigen Funktion auf $\overline{\Omega}$ fortgesetzt

werden, siehe [6, 17]. Angenommen es existiert eine solche Fortsetzung $w \in H^1(\Omega)$. Für den Konvergenzbeweis wird darüber hinaus benötigt, dass

$$w_j := \mathcal{I}_{\mathcal{S}_j} w \rightarrow w \quad \text{für } j \rightarrow \infty \text{ in } H^1(\Omega) \quad (4.15)$$

gilt, d.h. die stückweise linear Interpolierenden w_j von w in \mathcal{S}_j approximieren w in $H^1(\Omega)$. Um die obige Aussage zu garantieren, muss jedoch w als regulär genug vorausgesetzt werden. Für eine weiterführende Interpolationstheorie siehe zum Beispiel [13].

Für den hier betrachteten Fall, also $\tau_l \in \mathcal{T}_j$ und $d = 2$, folgt (4.15) für $w \in H^2(\Omega)$ unter der Voraussetzung, dass die lokalen Maschenweiten h_j für die zugehörige Folge von Triangulierungen abnehmen und dass die Folge von Triangulierungen $(\mathcal{T}_j)_{j \geq 0}$ formregulär ist. Somit können (4.14) und (4.15) auch durch die stärkere Bedingung $u_D = \text{tr}_{\Gamma_D} w$ mit $w \in H^2(\Omega)$ ersetzt werden.

Das nachfolgende Lemma gibt nun eine notwendige Bedingung für die Konvergenz der Lösung des diskreten Minimierungsproblems (4.6) zur Lösung des Minimierungsproblems (3.14) an, und zwar, die Konsistenz der diskreten Funktionale ϕ_j .

Lemma 4.5. *Die Sättigung M aus (3.1) sei beschränkt und monoton wachsend. Außerdem gelte, dass die lokalen Maschenweiten h_j abnehmen und dass die Folge von Unterteilungen formregulär ist.*

Falls $v \in C^\infty(\bar{\Omega})$ und $v_j = \mathcal{I}_{\mathcal{S}_j} v \in \mathcal{S}_j$ für $j \geq 0$, dann ergibt sich

$$v_j \rightarrow v \text{ in } H^1(\Omega) \quad \text{und} \quad \phi_j(v_j) \rightarrow \phi(v) \quad \text{für } j \rightarrow \infty. \quad (4.16)$$

Seien weiters (4.14) und (4.15) erfüllt, dann bleibt (4.16) auch für $v = w + \tilde{v} \in w + C^\infty(\bar{\Omega})$ und $v_j = \mathcal{I}_{\mathcal{S}_j} w + \mathcal{I}_{\mathcal{S}_j} \tilde{v} = w_j + \tilde{v}_j$, $j \geq 0$ gültig.

Für den Beweis siehe [6].

Mit Lemma 4.5 und den weiteren Eigenschaften der diskreten Funktionale ϕ_j aus Lemma 4.1 kann die Konvergenz der Näherungslösungen gezeigt werden.

Satz 4.6. *Außer der Stetigkeit von M seien alle Voraussetzungen aus Satz 3.11 erfüllt. Definiere die konvexe Menge*

$$\mathcal{K}_{\Gamma_D} := \mathcal{K} - w = \{v \in H_{\Gamma_D}^1(\Omega) : v \geq u_c - w \wedge \text{tr}_{\Gamma_S} v \leq -\text{tr}_{\Gamma_S} w\}.$$

Weiters sei $C_{\Gamma_D}^\infty(\bar{\Omega}) \cap \mathcal{K}_{\Gamma_D}$ dicht in \mathcal{K}_{Γ_D} , wobei

$$C_{\Gamma_D}^\infty(\bar{\Omega}) := \{v \in C^\infty(\bar{\Omega}) : v = 0 \text{ in einer Umgebung von } \Gamma_D\}.$$

4 Finite-Elemente-Methode

Dann konvergieren die Lösungen u_j des diskreten Minimierungsproblems (4.6) gegen die Lösung u des Minimierungsproblems (3.14) im Sinne von

$$u_j \rightarrow u \text{ in } H^1(\Omega) \quad \text{und} \quad \phi_j(u_j) \rightarrow \phi(u) \quad \text{für } j \rightarrow \infty.$$

Für den Beweis siehe zum Beispiel [25, 6].

Bemerkung 4.3. Der obige Satz lässt sich auch für den Fall einer ortsabhängigen hydraulischen Leitfähigkeit $K_h(\cdot)$ mit

$$c \leq K_h(\cdot) \leq C \quad \text{mit } c, C > 0$$

und für $M : \mathbb{R} \rightarrow \mathbb{R}$ erweitern.

Falls zusätzlich eine positive und beschränkte Porosität $n(\cdot)$ gegeben ist und eine Diskretisierung des dementsprechend geänderten Funktionals ϕ durch Anpassen der zugehörigen Gewichte in (4.4) gemacht wird, bleibt der obige Satz ebenfalls gültig, siehe auch [6].

Die nächste Aufgabe besteht nun darin, die Konvergenz der Sättigung $M(u_j) \rightarrow M(u)$ und die des physikalischen Drucks $p_j := \kappa^{-1}(u_j) \rightarrow p := \kappa^{-1}(u)$ zu zeigen. Dazu wird zunächst der Begriff der Hölder-Stetigkeit wiederholt.

Definition 4.7 (Hölder-Stetigkeit). Sei $U \subset \mathbb{R}$ offen und $0 < \alpha \leq 1$. Eine Abbildung $f : U \rightarrow \mathbb{R}$ heißt Hölder-stetig, wenn eine positive reelle Zahl C existiert, sodass

$$|f(x) - f(y)| \leq C |x - y|^\alpha \quad \forall x, y \in U.$$

Weiters werden einige Aussagen über Superpositionsoperatoren benötigt, die im Anschluss angegeben werden.

Seien $M : \mathbb{R} \rightarrow \mathbb{R}$ und $\kappa^{-1} : \mathbb{R} \rightarrow \mathbb{R}$. Dann werden durch $M \circ u$ und durch $\kappa^{-1} \circ u$ Superpositionsoperatoren induziert, siehe auch [4].

Lemma 4.8. Sei $\Omega \subset \mathbb{R}^2$ beschränkt und die Sättigung $M : \mathbb{R} \rightarrow \mathbb{R}$ sei gleichmäßig stetig und beschränkt. Dann wirkt der zugehörige Superpositionsoperator M_Ω im $L_2(\Omega)$ und ist stetig.

Lemma 4.9. Sei $\Omega \subset \mathbb{R}^2$ beschränkt. Wenn die Sättigung $M : \mathbb{R} \rightarrow \mathbb{R}$ Hölder-stetig bzgl. dem Exponenten $\alpha \in (0, 1]$ ist, dann induziert M einen Hölder-stetigen Superpositionsoperator bzgl. α :

$$M_\alpha : L_2(\Omega) \rightarrow L_{2/\alpha}(\Omega).$$

Sei $\Omega \subset \mathbb{R}^2$ beschränkt. Mit der stetigen Einbettung $i : L_{2/\alpha}(\Omega) \rightarrow L_2(\Omega)$ folgt die Hölder-Stetigkeit von $i \circ M_\alpha : L_2(\Omega) \rightarrow L_2(\Omega)$ bzgl. $\alpha \in (0, 1]$. Das wiederum erweitert die Stetigkeitsaussage aus Lemma 4.8 für eine Hölder-stetige Sättigung M . Es bleibt zu bemerken, dass die Sättigung M im Falle der Brooks-Corey-Parameterfunktionen Hölder-stetig ist.

Um in der Folge die Konvergenz der Sättigung $M(u_j) \rightarrow M(u)$ in $L_2(\Omega)$ für $j \rightarrow \infty$ aus der Konvergenz des verallgemeinerten Drucks $u_j \rightarrow u$ in $L_2(\Omega)$ für $j \rightarrow \infty$ mit Hilfe von Satz 4.6 abzuleiten, genügt es, die Eigenschaften der Sättigung M aus Lemma 4.8 und Lemma 4.9 auf der Vereinigung der Bildbereiche der Funktionen u_j , $j \geq 0$, und u anzunehmen.

Zunächst werde der nicht-entartete Fall betrachtet, d.h.

$$k_r(\cdot) \geq c \quad \text{für eine Konstante } c > 0, \quad (4.17)$$

siehe dazu auch Kapitel 1. Dann wird die Konvergenz des verallgemeinerten Drucks auf die Sättigung und den zurücktransformierten Druck vererbt und es ergibt sich der folgende Satz:

Satz 4.10. *Es seien die Voraussetzungen aus Satz 4.6 erfüllt. Für den nicht-entarteten Fall (4.17) folgt die Konvergenz der Sättigung*

$$M(u_j) \rightarrow M(u) \quad \text{in } H^1(\Omega) \quad \text{für } j \rightarrow \infty$$

und die des zurücktransformierten Drucks

$$\kappa^{-1}(u_j) \rightarrow \kappa^{-1}(u) \quad \text{in } H^1(\Omega) \quad \text{für } j \rightarrow \infty. \quad (4.18)$$

Für den Beweis siehe zum Beispiel [8, 27].

In der Praxis ist jedoch das Konvergenzverhalten der \mathcal{S}_j -Interpolationen von $M(u_j)$ und $\kappa^{-1}(u_j)$ von größerem Interesse, da insbesondere die \mathcal{S}_j -Interpolation von $\kappa^{-1}(u_j)$ den diskreten physikalischen Druck der Diskretisierung (4.13) angibt.

Für den Beweis der Konvergenz der \mathcal{S}_j -Interpolationen wird das nachfolgende Lemma benötigt:

Lemma 4.11. *Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ Hölder-stetig bzgl. dem Exponenten $\alpha \in (0, 1]$.*

$u_j \in \mathcal{S}_j$, $j \geq 0$, seien stückweise lineare Funktionen und es gelte $u_j \rightarrow u$ in $H^1(\Omega)$ für $j \rightarrow \infty$. Dann gilt:

$$f(u_j) - \mathcal{I}_{\mathcal{S}_j} f(u_j) \rightarrow 0 \quad \text{in } H^1(\Omega) \quad \text{für } j \rightarrow \infty. \quad (4.19)$$

4 Finite-Elemente-Methode

Beweis: Betrachte einen beliebigen Punkt x , der in einem Dreieck $\tau_l \in \mathcal{T}_j$ mit den Knoten q_1, q_2 und q_3 enthalten ist. Dann existieren $\lambda_i \in [0, 1], i = 1, 2, 3$, mit $\sum_{i=1}^3 \lambda_i = 1$, sodass

$$\mathcal{I}_{\mathcal{S}_j} f(u_j)(x) = \sum_{i=1}^3 \lambda_i f(u_j(q_i)).$$

Unter Anwendung von Binomischen Formeln und der Hölder-Stetigkeit von f mit der Hölder-Konstanten C_α , folgt die Abschätzung

$$\begin{aligned} |f(u_j(x)) - \mathcal{I}_{\mathcal{S}_j} f(u_j)(x)|^2 &= \left| f(u_j(x)) - \sum_{i=1}^3 \lambda_i f(u_j(q_i)) \right|^2 = \left| \sum_{i=1}^3 \lambda_i (f(u_j(x)) - f(u_j(q_i))) \right|^2 \leq \\ &3 \sum_{i=1}^3 |f(u_j(x)) - f(u_j(q_i))|^2 \leq 3 C_\alpha^2 \sum_{i=1}^3 |u_j(x) - u_j(q_i)|^{2\alpha}. \end{aligned} \quad (4.20)$$

Mit dem Mittelwertsatz ergibt sich auf dem Dreieck τ_l

$$|u_j(x) - u_j(q_i)| \leq |\nabla u_j(\eta_j)| |x - q_i|$$

mit der Zwischenstelle $\eta_j \in (x, q_i)$ und der Euklidischen Norm $|\cdot|$ in \mathbb{R}^2 .

Unter Berücksichtigung, dass $|\nabla u_j|$ konstant auf dem Dreieck τ_l ist, kann der letzte Term aus (4.20) weiter abgeschätzt werden durch

$$|f(u_j(x)) - \mathcal{I}_{\mathcal{S}_j} f(u_j)(x)|^2 \leq 3 C_\alpha^2 \sum_{i=1}^3 |\nabla u_j|^{2\alpha} |x - q_i|^{2\alpha} \leq 9 C_\alpha^2 |\nabla u_j|^{2\alpha} h_j^{2\alpha}$$

mit der lokalen Maschenweite h_j . Integration über das Gebiet Ω liefert schließlich

$$\begin{aligned} \int_{\Omega} |f(u_j(x)) - \mathcal{I}_{\mathcal{S}_j} f(u_j)(x)|^2 dx &\leq \sum_{\tau_l \in \mathcal{T}_j} \int_{\tau_l} |f(u_j(x)) - \mathcal{I}_{\mathcal{S}_j} f(u_j)(x)|^2 dx \leq \\ &\leq 9 C_\alpha^2 h_j^{2\alpha} \int_{\Omega} (|\nabla u_j|^2 + 1) dx. \end{aligned}$$

Da $(u_j)_{j \geq 0}$ in $H^1(\Omega)$ konvergiert, ist das letzte Integral in der obigen Abschätzung gleichmäßig beschränkt und somit geht dieser gesamte letzte Term für $j \rightarrow \infty$ gegen 0 aufgrund der Annahme, dass die lokalen Maschenweiten h_j für die zugehörige Folge von Unterteilungen abnehmen. ■

Als Folgerung von Lemma 4.11 ergibt sich schließlich die Konvergenz der Näherungslösungen der diskreten Sättigung und des diskreten physikalischen Drucks.

Satz 4.12. *Es seien die Voraussetzungen aus Satz 4.6 erfüllt. Dann folgt für die Konvergenz der diskreten Sättigung*

$$\theta_j(p_j) := \mathcal{I}_{S_j} M(u_j) \rightarrow \theta(p) \text{ in } L_2(\Omega) \text{ für } j \rightarrow \infty,$$

wenn $M : [u_c, \infty) \rightarrow \mathbb{R}$ (oder $M : \mathbb{R} \rightarrow \mathbb{R}$) Hölder-stetig, oder beschränkt und gleichmäßig stetig ist.

Für den nicht-entarteten Fall (4.17) folgt außerdem die Konvergenz des diskreten physikalischen Drucks

$$p_j = \mathcal{I}_{S_j} \kappa^{-1}(u_j) \rightarrow p = \kappa^{-1}(u) \text{ in } L_2(\Omega) \text{ für } j \rightarrow \infty. \quad (4.21)$$

Bemerkung 4.4. *Wenn die relative Permeabilität $k_r(\cdot)$ beliebig klein wird, führt das auf eine Singularität der inversen Kirchhoff-Transformation κ^{-1} . In diesem Fall kann nicht garantiert werden, dass $\kappa^{-1}(u_j), \kappa^{-1}(u) \in L_1(\Omega)$, $j \geq 0$, gilt.*

Wenn jedoch (4.7) und (2.15) auf physikalisch realistische Lösungen p_j bzw. $p = \kappa^{-1}(u)$ führen, dann sollten diese Lösungen, aufgrund der Tatsache, dass das Auftreten eines beliebig großen physikalischen Drucks in porösen Medien unnatürlich ist, gleichmäßig beschränkt sein. Dann sind u_j , $j \geq 0$, und u gleichmäßig beschränkt fernab vom kritischen verallgemeinerten Druck u_c . Damit ergibt sich die gleiche Situation wie für den nicht-entarteten Fall und es kann sowohl (4.18) als auch (4.21) gefolgert werden.

Aus dem Beweis von Lemma 4.11 folgt bereits die Konvergenzordnung $\mathcal{O}(h_j^\alpha)$ von (4.19) für $\alpha \in (0, 1]$. Insgesamt ergibt sich damit unter Annahme der Voraussetzungen von Satz 4.10:

$$\|\mathcal{I}_{S_j} M(u_j) - M(u)\|_{L_2(\Omega)} \leq C_\alpha \left(h_j^\alpha + \|u_j - u\|_{H^1(\Omega)}^\alpha \right)$$

und

$$\|\mathcal{I}_{S_j} \kappa^{-1}(u_j) - \kappa^{-1}(u)\|_{L_2(\Omega)} \leq C_1 \left(h_j + \|u_j - u\|_{H^1(\Omega)} \right).$$

Das heißt die Konvergenz des diskreten physikalischen Drucks $p_j \rightarrow p$ in der L_2 -Norm ist von der Ordnung $\mathcal{O}(h_j)$, wenn die Konvergenz des diskreten verallgemeinerten Drucks $u_j \rightarrow u$ in der H^1 -Norm von der selben Ordnung ist.

5 Dirichlet-Neumann- und Robin-Methoden für nichtlineare Transmissionsprobleme

Thema dieses Kapitels ist ein quasilineares, elliptisches Transmissionsproblem, bei dem sich die Nichtlinearität unstetig zwischen zwei Teilgebieten verändert. Dieses Problem ergibt sich aus der Betrachtung der Richards-Gleichung in einem heterogenen Boden. Heterogen bedeutet in diesem Zusammenhang, dass die Bodenparameter auf nicht überlappenden Teilgebieten Ω_i , $i = 1, 2$, des Gebietes $\Omega \subset \mathbb{R}^2$ konstant sind und sich unstetig entlang des Interface der beiden Teilgebiete ändern. Das resultierende Problem ist somit ein Problem mit springenden Nichtlinearitäten oder anders gesagt mit nichtlinearen Parameterfunktionen, die zwar auf beiden Teilgebieten fest gewählt, aber auf jeder Seite des Interface unterschiedlich sind.

In Abschnitt 5.1 wird zunächst ein Gebietszerlegungsproblem und dessen schwache Formulierung für den verallgemeinerten Druck u angegeben. Danach wird in Abschnitt 5.2 eine zugehörige Steklov-Poincaré-Formulierung hergeleitet. Auf geeignete Lösungsverfahren wird schließlich in Abschnitt 5.3 bzw. in Abschnitt 5.4 näher eingegangen. Ergänzt wird dieses Kapitel letztendlich durch die Konvergenzanalyse der Dirichlet-Neumann- und Robin-Methode in Abschnitt 5.5.

5.1 Allgemeine Problemstellung

Sei $\Omega \subset \mathbb{R}^2$ ein beschränktes Lipschitz-Gebiet, unterteilt in zwei nicht überlappende Teilgebiete Ω_1 und Ω_2 mit dem Interface $\Gamma = \overline{\Omega_1} \cap \overline{\Omega_2}$; siehe Abbildung 5.1.

Dabei sei \mathbf{n} der äußere Normalenvektor bezüglich Ω_1 . Weiters seien $f \in L_2(\Omega)$ und $k_1, k_2 \in L_\infty(\mathbb{R})$ mit $k_i \geq \alpha > 0$ für $i = 1, 2$, das heißt die relative Durchlässigkeit k_i , $i = 1, 2$, ist von unten durch eine positive Konstante beschränkt.

Um ein Gebietszerlegungsproblem für den verallgemeinerten physikalischen Druck u angeben und die Ergebnisse, die im Zusammenhang mit der Betrachtung der Richards-Gleichung in homogenen Böden erzielt wurden, anwenden zu können, wird auf beiden Teilgebieten eine Kirchhoff-Transformation durchgeführt, wobei neue Variable u_i , $i = 1, 2$,

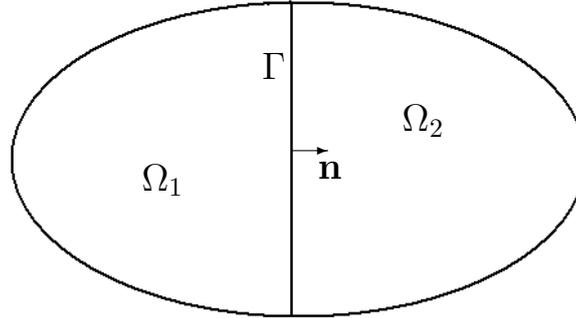


Abbildung 5.1: Nicht überlappende Zerlegung von Ω

gemäß (1.9) eingeführt werden:

$$u_i(x) := \kappa_i(p_i(x)) = \int_0^{p_i(x)} k_i(q) dq \quad \text{fast überall in } \Omega_i. \quad (5.1)$$

Mit $k_i(p_i)\nabla p_i = \nabla u_i$ folgt das Gebietszerlegungsproblem für den verallgemeinerten Druck in der Form:

Gesucht ist eine Funktion u in Ω mit $u|_{\Omega_i} = u_i \in H^1(\Omega_i)$, $i = 1, 2$, und $u|_{\partial\Omega} = 0$, sodass

$$-\Delta u_i = f \quad \text{in } \Omega_i, i = 1, 2, \quad (5.2)$$

$$\kappa_1^{-1}(u_1) = \kappa_2^{-1}(u_2) \quad \text{auf } \Gamma, \quad (5.3)$$

$$\frac{\partial}{\partial \mathbf{n}} u_1 = \frac{\partial}{\partial \mathbf{n}} u_2 \quad \text{auf } \Gamma. \quad (5.4)$$

Bemerkung 5.1. Die Probleme (5.2) sind linear auf Ω_1 und Ω_2 , die Nichtlinearität des Problems ist vollständig in der Interface-Gleichung (5.3) enthalten. Diese Folgeerscheinung der Kirchhoff-Transformation macht einerseits die Analysis und andererseits die numerische Behandlung des Transmissionsproblems einfacher.

Das nächste Ziel ist die Herleitung einer schwachen Formulierung des Gebietszerlegungsproblems (5.2) - (5.4).

Dazu werden folgende Räume definiert:

$$V_i := \{v_i \in H^1(\Omega_i) \mid v_i|_{\partial\Omega \cap \partial\Omega_i} = 0\}, \quad V_i^0 := H_0^1(\Omega_i), \quad \Lambda := \tilde{H}^{1/2}(\Gamma).$$

5.2 Steklov-Poincaré-Operator-Formulierung

Es wird die folgende Bilinearform eingeführt:

$$a_i(w_i, v_i) := (\nabla w_i, \nabla v_i)_{\Omega_i} = \int_{\Omega_i} \nabla w_i(x) \nabla v_i(x) dx \quad \forall w_i, v_i \in V_i.$$

Wenn die Kirchhoff-Transformationen κ_i im $H^1(\Omega_i)$ aus (5.1) als punktweise fast überall definierte Funktionen in Ω_i , $i = 1, 2$, aufgefasst werden, so ergibt sich unter Anwendung der Green'schen Formel folgende schwache Formulierung für das Problem (5.2) - (5.4):

Gesucht sind $u_i \in V_i$, $i = 1, 2$, sodass

$$a_i(u_i, v_i) = (f, v_i)_{\Omega_i} \quad \forall v_i \in V_i^0, \quad (5.5)$$

$$\kappa_1^{-1}(u_{1|\Gamma}) = \kappa_2^{-1}(u_{2|\Gamma}) \quad \text{in } \Lambda, \quad (5.6)$$

$$a_1(u_1, R_1\mu) - (f, R_1\mu)_{\Omega_1} = -a_2(u_2, R_2\mu) + (f, R_2\mu)_{\Omega_2} \quad \forall \mu \in \Lambda, \quad (5.7)$$

wobei $R_i : \Lambda \rightarrow V_i$, $i = 1, 2$, lineare Fortsetzungsoperatoren bezeichnen.

Bemerkung 5.2. *Durch Kopplung der Richards-Gleichung in einem homogenen Boden auf jedem Teilgebiet Ω_i , $i = 1, 2$, mit der Stetigkeit des physikalischen Drucks und der des Normalenflusses kann auf ähnliche Weise ein Gebietszerlegungsproblem sowie eine zugehörige schwache Formulierung für den physikalischen Druck p angegeben werden. Die entsprechende schwache Formulierung lautet dann:*

Gesucht sind $p_i \in V_i$, $i = 1, 2$, sodass

$$b_i(p_i, v_i) = (f, v_i)_{\Omega_i} \quad \forall v_i \in V_i^0, \quad (5.8)$$

$$p_{1|\Gamma} = p_{2|\Gamma} \quad \text{in } \Lambda, \quad (5.9)$$

$$b_1(p_1, R_1\mu) - (f, R_1\mu)_{\Omega_1} = -b_2(p_2, R_2\mu) + (f, R_2\mu)_{\Omega_2} \quad \forall \mu \in \Lambda, \quad (5.10)$$

mit

$$b_i(w_i, v_i) := (k_i(w_i) \nabla w_i, \nabla v_i)_{\Omega_i} = \int_{\Omega_i} k_i(w_i(x)) \nabla w_i(x) \nabla v_i(x) dx \quad \forall w_i, v_i \in V_i.$$

Für den Beweis der Äquivalenz der schwachen Formulierung des Gebietszerlegungsproblems für den physikalischen Druck (5.8) - (5.10) und der schwachen Formulierung (5.5) - (5.7), der stark von der Interpretation der Kirchhoff-Transformation als Superpositionsoperator abhängt, siehe [7].

5.2 Steklov-Poincaré-Operator-Formulierung

Im Anschluss soll nun eine Steklov-Poincaré-Operator-Formulierung der Probleme (5.8) - (5.10) bzw. (5.5) - (5.7) angegeben werden.

5 Dirichlet-Neumann- und Robin-Methoden für nichtlineare Transmissionsprobleme

Dazu werden für ein gegebenes $\lambda \in \Lambda$ die harmonischen Fortsetzungen $u_i^0 = H_i(\kappa_i(\lambda)) \in V_i$, $i = 1, 2$, der Dirichlet-Randbedingungen $\kappa_i(\lambda)$ auf Γ betrachtet.

$u_i^* = \mathcal{G}_i f$ seien die Lösungen der Teilprobleme (5.5) mit homogenen Dirichlet-Daten $u_i^*|_{\partial\Omega_i} = 0$.

Aufgrund der Linearität der lokalen Probleme (5.5), erfüllen die Funktionen

$$u_i = H_i(\kappa_i(\lambda)) + \mathcal{G}_i f, \quad i = 1, 2,$$

das Problem (5.5) - (5.7) genau dann, wenn

$$\begin{aligned} a_1(H_1(\kappa_1(\lambda)), R_1\mu) + a_2(H_2(\kappa_2(\lambda)), R_2\mu) = \\ = (f, R_1\mu)_{\Omega_1} - a_1(\mathcal{G}_1 f, R_1\mu) + (f, R_2\mu)_{\Omega_2} - a_2(\mathcal{G}_2 f, R_2\mu) \quad \forall \mu \in \Lambda. \end{aligned} \quad (5.11)$$

Die Fortsetzungsoperatoren R_i , $i = 1, 2$, können beliebig gewählt werden, müssen jedoch stetig sein. Setze: $R_i = H_i$.

Die Steklov-Poincaré-Operatoren $S_i : \Lambda \rightarrow \Lambda'$ sind wie folgt definiert:

$$\langle S_i \eta, \mu \rangle_{\Gamma} = a_i(H_i \eta, H_i \mu) \quad \forall \eta, \mu \in \Lambda, \quad i = 1, 2, \quad S := S_1 + S_2, \quad (5.12)$$

wobei $\langle \cdot, \cdot \rangle_{\Gamma}$ das Dualitätsprodukt zwischen Λ' und Λ bezeichnet.

Weiters sei das Funktional $\chi = \chi_1 + \chi_2 \in \Lambda'$ durch

$$\langle \chi_i, \mu \rangle_{\Gamma} = (f, H_i \mu)_{\Omega_i} - a_i(\mathcal{G}_i f, H_i \mu) \quad \forall \mu \in \Lambda, \quad i = 1, 2, \quad (5.13)$$

definiert, siehe auch [29].

Die Bedingung (5.11) kann an dieser Stelle als nichtlineare Steklov-Poincaré-Interface-Gleichung umformuliert werden:

$$\lambda \in \Lambda : \quad (S_1 \kappa_1 + S_2 \kappa_2) \lambda = \chi \quad (5.14)$$

oder äquivalent für $\lambda_2 = \kappa_2 \lambda$:

$$\lambda_2 \in \Lambda : \quad (S_1 \kappa_1 \kappa_2^{-1} + S_2) \lambda_2 = \chi. \quad (5.15)$$

Satz 5.1. *Aufgrund der Definition der Kirchhoff-Transformation (5.1) und der Beziehung zwischen λ und u_i*

$$u_i = H_i \kappa_i \lambda + \mathcal{G}_i f, \quad i = 1, 2,$$

sind die schwachen Formulierungen der Gebietszerlegungsprobleme (5.8) - (5.10) bzw. (5.5) - (5.7) äquivalent zur nichtlinearen Steklov-Poincaré-Gleichung (5.14) bzw. (5.15).

5.3 Nichtlineare Dirichlet-Neumann-Methode

Ziel dieses Abschnittes ist die Herleitung eines geeigneten iterativen Verfahrens zur Lösung des Gebietszerlegungsproblems (5.8) - (5.10). Für den linearen Fall, d.h., dass k_i , $i = 1, 2$, konstante Funktionen sind, ist die Dirichlet-Neumann-Methode ein wohlbekanntes Iterationsverfahren zur Lösung von nicht überlappenden, elliptischen Gebietszerlegungsproblemen. Die Aufgabe dieses Abschnittes besteht nun darin, durch Verallgemeinerung des linearen Falles einen nichtlinearen Dirichlet-Neumann-Algorithmus abzuleiten, der in jedem Iterationsschritt die Lösung von zwei linearen Problemen und eine nichtlineare Transformation auf dem Interface, aber keine weitere Linearisierung, erfordert.

Da für eine vollständige Analysis die Dämpfung in den transformierten Variablen durchgeführt werden muss, wird der Dirichlet-Neumann-Algorithmus für das transformierte Problem (5.2) - (5.4) angegeben. Dieser lautet wie folgt:

Für ein gegebenes $\lambda_2^0 \in \Lambda$ suche sukzessive $u_1^{k+1} \in V_1$ und $u_2^{k+1} \in V_2$ für jedes $k \geq 0$, sodass

$$-\Delta u_1^{k+1} = f \quad \text{in } \Omega_1, \quad (5.16)$$

$$u_1^{k+1} = 0 \quad \text{auf } \partial\Omega_1, \quad (5.17)$$

$$u_1^{k+1} = \kappa_1 \kappa_2^{-1} \lambda_2^k \quad \text{auf } \Gamma, \quad (5.18)$$

und dann

$$-\Delta u_2^{k+1} = f \quad \text{in } \Omega_2, \quad (5.19)$$

$$u_2 = 0 \quad \text{auf } \partial\Omega_2, \quad (5.20)$$

$$\frac{\partial}{\partial \mathbf{n}} u_1^{k+1} - \frac{\partial}{\partial \mathbf{n}} u_2^{k+1} = 0 \quad \text{auf } \Gamma. \quad (5.21)$$

Der gesuchte Dirichlet-Neumann-Algorithmus angewandt auf die schwache Formulierung (5.5) - (5.7) besagt schließlich:

Für ein gegebenes $\lambda_2^0 \in \Lambda$ suche sukzessive $u_1^{k+1} \in V_1$ und $u_2^{k+1} \in V_2$ für jedes $k \geq 0$, sodass

$$a_1(u_1^{k+1}, v_1) = (f, v_1)_{\Omega_1} \quad \forall v_1 \in V_1^0 \quad (5.22)$$

$$u_{1|\Gamma}^{k+1} = \kappa_1 \kappa_2^{-1} \lambda_2^k \quad \text{in } \Lambda \quad (5.23)$$

und dann

$$a_2(u_2^{k+1}, v_2) = (f, v_2)_{\Omega_2} \quad \forall v_2 \in V_2^0 \quad (5.24)$$

$$a_2(u_2^{k+1}, H_2 \mu) - (f, H_2 \mu)_{\Omega_2} = -a_1(u_1^{k+1}, H_1 \mu) + (f, H_1 \mu)_{\Omega_1} \quad \forall \mu \in \Lambda. \quad (5.25)$$

Mit einem Dämpfungsparameter $\vartheta \in (0, 1)$ wird die neue Iterierte definiert durch

$$\lambda_2^{k+1} = \vartheta u_{2|\Gamma}^{k+1} + (1 - \vartheta) \lambda_2^k. \quad (5.26)$$

Es soll nun wie im vorherigen Abschnitt 5.2 eine Formulierung mit Steklov-Poincaré-Operatoren hergeleitet werden.

Abermals können die Iterierten u_i^{k+1} , $i = 1, 2$, wie folgt dargestellt werden:

$$u_i^{k+1} = H_i(u_{i|\Gamma}^{k+1}) + G_i f \quad i = 1, 2.$$

Dabei seien $\widehat{u}_i^{k+1} := H_i(u_{i|\Gamma}^{k+1})$ die harmonischen Fortsetzungen der Dirichlet-Randbedingungen auf Γ und $\widetilde{u}_i^{k+1} = G_i f$ seien die Lösungen der Probleme (5.5) mit homogenen Dirichlet-Randdaten für $i = 1, 2$.

Mit (5.23) ergibt sich schließlich

$$u_1^{k+1} = H_1(\kappa_1 \kappa_2^{-1} \lambda_2^k) + G_1 f \quad \text{und} \quad u_2^{k+1} = H_2(u_{2|\Gamma}^{k+1}) + G_2 f.$$

Durch Einsetzen der Definition der Steklov-Poincaré-Operatoren (5.12) und des Funktionals χ aus (5.13) in (5.25), folgt

$$\left\langle S_1 \kappa_1 \kappa_2^{-1} \lambda_2^k + S_2 u_{2|\Gamma}^{k+1}, \mu \right\rangle_{\Gamma} = \langle \chi, \mu \rangle_{\Gamma} \quad \forall \mu \in \Lambda. \quad (5.27)$$

Unter Anwendung von (5.26) ergibt sich schließlich

$$S_2 (\lambda_2^{k+1} - \lambda_2^k) = \vartheta (\chi - (S_1 \kappa_1 \kappa_2^{-1} + S_2) \lambda_2^k) \quad \text{in } \Lambda'$$

bzw.

$$\begin{aligned} \lambda_2^{k+1} &= \lambda_2^k - \vartheta S_2^{-1} ((S_1 \kappa_1 \kappa_2^{-1} + S_2) \lambda_2^k - \chi) = \\ &= (I - \vartheta S_2^{-1} (S_1 \kappa_1 \kappa_2^{-1} + S_2)) \lambda_2^k + \vartheta S_2^{-1} \chi \quad \text{in } \Lambda'. \end{aligned} \quad (5.28)$$

Offensichtlich ist der gedämpfte Dirichlet-Neumann-Algorithmus angewandt auf (5.5) - (5.7) ein vorkonditioniertes Richardson-Verfahren für die nichtlineare Steklov-Poincaré-Operator-Formulierung (5.15) mit S_2 als Vorkonditionierer. Genauer gesagt ist die Iteration durch den Operator $T_{\vartheta} : \Lambda \rightarrow \Lambda$ gegeben, der wie folgt definiert ist:

$$T_{\vartheta} : \lambda_2^k \rightarrow \lambda_2^{k+1} = \lambda_2^k + \vartheta S_2^{-1} (\chi - (S_1 \kappa_1 \kappa_2^{-1} + S_2) \lambda_2^k). \quad (5.29)$$

Bemerkung 5.3. Dieses Richardson-Verfahren kann verallgemeinert werden indem statt S_2 ein beliebiger Operator $B : H^{-1/2}(\Gamma) \rightarrow \tilde{H}^{1/2}(\Gamma)$ zur Vorkonditionierung verwendet wird. Eine Möglichkeit für die Wahl von B wäre das Einfachschichtpotential V .

Bemerkung 5.4. Eine analoge Formulierung für die Interface-Gleichung (5.14) kann aufgrund der Nichtlinearität von $S_2\kappa_2$ nicht gewonnen werden, siehe [6].

5.4 Nichtlineare Robin-Methode

Analog zur Herleitung der nichtlinearen Dirichlet-Neumann-Methode soll in diesem Abschnitt durch Verallgemeinerung der Robin-Methode, die für den linearen Fall zur Behandlung von nicht überlappenden, elliptischen Gebietszerlegungsproblemen angewendet werden kann, ein nichtlinearer Robin-Algorithmus für das transformierte Problem (5.2) - (5.4), der ohne Linearisierung der Transmissionsbedingungen auskommt, angegeben werden.

Die grundlegende Idee dafür, ist die Umformung der Transmissionsbedingungen

$$\kappa_1^{-1}(u_1) = \kappa_2^{-1}(u_2) \quad \text{auf } \Gamma \quad (5.30)$$

$$\frac{\partial}{\partial \mathbf{n}} u_1 = \frac{\partial}{\partial \mathbf{n}} u_2 \quad \text{auf } \Gamma \quad (5.31)$$

in äquivalente Bedingungen der Form

$$B\kappa_1^{-1}(u_1) + \frac{\partial}{\partial \mathbf{n}} u_1 = B\kappa_2^{-1}(u_2) + \frac{\partial}{\partial \mathbf{n}} u_2 \quad \text{auf } \Gamma$$

bzw.

$$B\kappa_1^{-1}(u_1) - \frac{\partial}{\partial \mathbf{n}} u_1 = B\kappa_2^{-1}(u_2) - \frac{\partial}{\partial \mathbf{n}} u_2 \quad \text{auf } \Gamma$$

mit einem geeignet gewählten Operator B .

Der Robin Algorithmus für das Gebietszerlegungsproblem (5.2) - (5.4) lautet:

Gegeben seien der Anfangswert $u_2^0 \in V_2$ und positive Parameter γ_1, γ_2 . Gesucht sind sukzessive $u_1^{k+1} \in V_1$ und $u_2^{k+1} \in V_2$ für $k \geq 0$, sodass

$$-\Delta u_1^{k+1} = f \quad \text{in } \Omega_1, \quad (5.32)$$

$$u = 0 \quad \text{auf } \partial\Omega_1, \quad (5.33)$$

$$\left[\frac{\partial}{\partial \mathbf{n}} u_1^{k+1} - \frac{\partial}{\partial \mathbf{n}} u_2^k \right] + \gamma_1 B [\kappa_1^{-1} u_1^{k+1} - \kappa_2^{-1} u_2^k] = 0 \quad \text{auf } \Gamma, \quad (5.34)$$

und dann

$$-\Delta u_2^{k+1} = f \quad \text{in } \Omega_2, \quad (5.35)$$

$$u = 0 \quad \text{auf } \partial\Omega_2, \quad (5.36)$$

$$\left[\frac{\partial}{\partial \mathbf{n}} u_1^{k+1} - \frac{\partial}{\partial \mathbf{n}} u_2^{k+1} \right] - \gamma_2 B [\kappa_1^{-1} u_1^{k+1} - \kappa_2^{-1} u_2^{k+1}] = 0 \quad \text{auf } \Gamma, \quad (5.37)$$

mit einem Operator $B : \tilde{H}^{1/2}(\Gamma) \rightarrow H^{-1/2}(\Gamma)$, der als beschränkt und $\tilde{H}^{1/2}(\Gamma)$ -elliptisch vorausgesetzt wird, das heißt es existiert eine Konstante $c_1^B > 0$, sodass

$$\langle Bv, v \rangle_\Gamma \geq c_1^B \|v\|_{\tilde{H}^{1/2}(\Gamma)}^2 \quad \forall v \in \tilde{H}^{1/2}(\Gamma).$$

Der gesuchte Robin-Algorithmus angewandt auf die schwache Formulierung (5.5) - (5.7) besagt schließlich:

Gegeben sei der Anfangswert $u_2^0 \in V_2$ und positive Parameter γ_1, γ_2 . Gesucht sind sukzessive $u_1^{k+1} \in V_1$ und $u_2^{k+1} \in V_2$ für $k \geq 0$, sodass

$$a_1(u_1^{k+1}, v_1) = (f, v_1)_{\Omega_1} \quad \forall v_1 \in V_1^0 \quad (5.38)$$

$$\begin{aligned} & (a_1(u_1^{k+1}, R_1\mu) - (f, R_1\mu)_{\Omega_1} + a_2(u_2^k, R_2\mu) - (f, R_2\mu)_{\Omega_2}) + \\ & + \gamma_1 (B(\kappa_1^{-1} u_1^{k+1} - \kappa_2^{-1} u_2^k), \mu)_\Gamma = 0 \quad \forall \mu \in \Lambda \quad (5.39) \end{aligned}$$

und dann

$$a_2(u_2^{k+1}, v_2) = (f, v_2)_{\Omega_2} \quad \forall v_2 \in V_2^0 \quad (5.40)$$

$$\begin{aligned} & (a_1(u_1^{k+1}, R_1\mu) - (f, R_1\mu)_{\Omega_1} + a_2(u_2^{k+1}, R_2\mu) - (f, R_2\mu)_{\Omega_2}) - \\ & - \gamma_2 (B(\kappa_1^{-1} u_1^{k+1} - \kappa_2^{-1} u_2^{k+1}), \mu)_\Gamma = 0 \quad \forall \mu \in \Lambda. \quad (5.41) \end{aligned}$$

Nächstes Ziel ist die Herleitung einer Steklov-Poincaré-Operator-Formulierung der Robin-Methode (5.38) - (5.41).

Der nachfolgende Satz gibt nun eine Verallgemeinerung der linearen Theorie aus [14] an.

Satz 5.2. *Das Robin-Iterationsverfahren (5.38) - (5.41) angewandt auf (5.5) - (5.7) ist äquivalent zur ADI-Methode (**A**lternating-**D**irection-**I**terative-Methode) angewandt auf die Steklov-Poincaré-Operator-Gleichung (5.14).*

Mit einem gegebenen $\lambda_2^0 \in \Lambda$ lautet die ADI-Methode

$$\langle (\gamma_1 B + S_1 \kappa_1) \lambda_1^{k+1}, \mu \rangle_\Gamma = \langle \chi + (\gamma_1 B - S_2 \kappa_2) \lambda_2^k, \mu \rangle_\Gamma \quad (5.42)$$

$$\langle (\gamma_2 B + S_2 \kappa_2) \lambda_2^{k+1}, \mu \rangle_\Gamma = \langle \chi + (\gamma_2 B - S_1 \kappa_1) \lambda_1^{k+1}, \mu \rangle_\Gamma \quad (5.43)$$

sukzessive für $k \geq 0$.

Die Äquivalenz zur Robin-Methode ist im Sinne von Satz 5.1 gegeben, das heißt durch

$$u_i^k = H_i(\kappa_i(\lambda_i^k)) + G_i f \Leftrightarrow \lambda_i^k = \kappa_i^{-1}(u_{i|\Gamma}^k) \quad \forall k \geq 0, i = 1, 2, \quad (5.44)$$

mit den Iterierten u_i^k aus (5.38) - (5.41). Der Operator der ADI-Methode

$$T_{\gamma_1, \gamma_2} : \Lambda \rightarrow \Lambda, \quad T_{\gamma_1, \gamma_2} : \lambda_2^k \rightarrow \lambda_2^{k+1}, \quad (5.45)$$

und äquivalent dazu die Robin-Iteration sind durch

$$\lambda_2^{k+1} = (\gamma_2 B + S_2 \kappa_2)^{-1} \left(\chi + (\gamma_2 B - S_1 \kappa_1) (\gamma_1 B + S_1 \kappa_1)^{-1} \left(\chi + (\gamma_1 B - S_2 \kappa_2) \lambda_2^k \right) \right) \quad (5.46)$$

gegeben.

5.5 Konvergenz der Lösungsverfahren

Um die Konvergenz der Dirichlet-Neumann- bzw. der Robin-Methode zu zeigen, werden die Operatoren T_ϑ bzw. T_{γ_1, γ_2} , die in (5.29) bzw. (5.45) definiert wurden, näher betrachtet. Es ist leicht zu sehen, dass ein Fixpunkt λ von (5.29) bzw. von (5.46) eine Lösung der Steklov-Poincaré-Operator-Gleichung (5.15) bzw. (5.14) ist.

Weiters ist es möglich, Aussagen, die für den linearen Fall gelten, zu erweitern und Bedingungen an T_ϑ und T_{γ_1, γ_2} zu stellen, sodass der Banach'sche Fixpunktsatz angewendet werden kann.

Die hinreichende Bedingung, damit Konvergenz garantiert werden kann, beinhaltet für die Dirichlet-Neumann Methode, dass $T_\vartheta : \Lambda \rightarrow \Lambda$ eine Kontraktion ist, wenn nur ϑ klein genug ist.

Mit Hilfe von Satz 5.3 kann eine Konvergenzaussage für die nichtlineare Dirichlet-Neumann-Methode (5.22) - (5.25) gefolgert werden, die schließlich in Satz 5.4 zusammengefasst wird.

Satz 5.3. *Sei X ein Hilbertraum mit dem Dualraum X' . Der lineare und invertierbare Operator $Q_2 : X \rightarrow X'$ sei stetig und koerziv, das heißt es existieren positive Konstante β_2 und α_2 , sodass*

$$\langle Q_2 \eta, \mu \rangle_\Gamma \leq \beta_2 \|\eta\|_X \|\mu\|_X \quad \forall \eta, \mu \in X, \quad \langle Q_2 \eta, \eta \rangle_\Gamma \geq \alpha_2 \|\eta\|_X^2.$$

5 Dirichlet-Neumann- und Robin-Methoden für nichtlineare Transmissionsprobleme

Der Operator $Q_1 : X \rightarrow X'$ sei Lipschitz-stetig, das heißt es existiert eine Konstante $\beta_1 > 0$, sodass

$$\langle Q_1\eta - Q_1\mu, \lambda \rangle_\Gamma \leq \beta_1 \|\eta - \mu\|_X \|\lambda\|_X \quad \forall \eta, \mu, \lambda \in X.$$

Angenommen es existiert eine Konstante $k > 0$, sodass

$$\langle Q_2(\eta - \mu), Q_2^{-1}(Q\eta - Q\mu) \rangle_\Gamma + \langle Q\eta - Q\mu, \eta - \mu \rangle_\Gamma \geq k \|\eta - \mu\|_X^2 \quad \forall \eta, \mu \in X$$

mit $Q := Q_1 + Q_2$. Dann besitzt das Problem

$$\lambda \in X : \quad Q\lambda = G$$

mit einem gegebenen $G \in X'$ eine eindeutige Lösung $\lambda \in X$. Zusätzlich konvergiert die Folge, gegeben durch

$$\lambda^{k+1} = \lambda^k + \vartheta Q_2^{-1}(G - Q\lambda^k),$$

gegen die eindeutige Lösung $\lambda \in X$ für ein gegebenes $\lambda^0 \in X$ und einen Dämpfungsparameter $\vartheta \in (0, \vartheta_{\max})$ mit

$$\vartheta_{\max} := \frac{k\alpha_2^2}{\beta_2(\beta_1 + \beta_2)^2}.$$

Der Beweis wird unter Anwendung des Banach'schen Fixpunktsatzes durchgeführt, siehe [6, 29]. Es folgt nun die Konvergenzaussage für den nichtlinearen Dirichlet-Neumann-Algorithmus:

Satz 5.4. Sei S_2 beschränkt, das heißt

$$\langle S_2\eta, \mu \rangle_\Gamma \leq \beta_2 \|\eta\|_\Lambda \|\mu\|_\Lambda \quad \forall \eta, \mu \in \Lambda$$

mit einer positiven Konstanten β_2 und Λ -elliptisch, das heißt

$$\langle S_2\eta, \eta \rangle_\Gamma \geq \alpha_2 \|\eta\|_\Lambda^2 \quad \forall \eta \in \Lambda$$

mit einer positiven Konstanten α_2 . Weiters sei $\kappa_1\kappa_2^{-1} : \Lambda \rightarrow \Lambda$ und damit $S_1\kappa_1\kappa_2^{-1}$ Lipschitz-stetig, das heißt es existiert eine Konstante $\beta_1 > 0$, sodass

$$\langle S_1\kappa_1\kappa_2^{-1}\eta - S_1\kappa_1\kappa_2^{-1}\mu, \lambda \rangle_\Gamma \leq \beta_1 \|\eta - \mu\|_\Lambda \|\lambda\|_\Lambda \quad \forall \eta, \mu, \lambda \in \Lambda.$$

5.5 Konvergenz der Lösungsverfahren

Angenommen es existiert zusätzlich eine Konstante $k > 0$, sodass

$$\langle S_1 \kappa_1 \kappa_2^{-1} \eta - S_1 \kappa_1 \kappa_2^{-1} \mu, \eta - \mu \rangle_\Gamma \geq \left(\frac{k}{2} - \alpha_2 \right) \|\eta - \mu\|_\Lambda^2 \quad \forall \eta, \mu \in \Lambda.$$

Dann besitzt die Steklov-Poincaré-Operator-Gleichung (5.15) eine eindeutige Lösung $\lambda_2 \in \Lambda$. Außerdem konvergiert die Folge (5.29) gegen λ_2 in Λ für ein beliebiges $\lambda_2^0 \in \Lambda$ und $\vartheta \in (0, \vartheta_{\max})$ mit

$$\vartheta_{\max} := \frac{k\alpha_2^2}{\beta_2 (\beta_1 + \beta_2)^2}.$$

Die theoretisch optimalen Konvergenzraten ρ_{opt} für die zugehörigen optimalen Dämpfungsparameter ϑ_{opt} sind

$$\vartheta_{opt} = \frac{\vartheta_{\max}}{2} \quad \text{und} \quad \rho_{opt} = 1 - \frac{k}{2\beta_2} \vartheta_{opt}.$$

Bemerkung 5.5. Für den eindimensionalen Fall ist $\kappa_1 \kappa_2^{-1} : \Lambda \rightarrow \Lambda$ Lipschitz-stetig und $S_1 \kappa_1 \kappa_2^{-1} : \Lambda \rightarrow \Lambda'$ ist ein streng monotoner Operator. Damit sind die Voraussetzungen aus Satz 5.4 erfüllt. Für den Beweis siehe [6].

Es ist nicht bekannt, ob die Voraussetzungen aus Satz 5.4 auch in höheren Dimensionen notwendig sind. Numerische Gegenbeispiele in 2D mit nicht monotonen Operatoren $S_1 \kappa_1 \kappa_2^{-1} : \Lambda \rightarrow \Lambda'$ haben gezeigt, dass in diesem Fall trotzdem Konvergenz mit realistischen Konvergenzraten für den nichtlinearen Dirichlet-Neumann-Algorithmus mit geeignetem Dämpfungsparameter erzielt wird, siehe [9, 6].

Falls Satz 5.4 jedoch angewendet werden kann, ergibt sich das folgende Existenz- und Eindeutigkeitsresultat für das nichtlineare, heterogene Problem (5.8) - (5.10).

Satz 5.5. Falls die Bedingungen aus Satz 5.4 erfüllt sind, ist das Problem (5.8) - (5.10) wohldefiniert. Außerdem konvergieren die Iterierten p_i^k , die gegeben sind durch

$$p_i^k = \kappa_i^{-1} (H_i(\kappa_i(\lambda^k)) + G_i f) \in V_i, \quad i = 1, 2,$$

gegen p_i in V_i für $k \rightarrow \infty$. Der Zusammenhang zu den Iterierten $(\lambda_2^k)_{k \geq 0}$ aus dem Dirichlet-Neumann-Verfahren (5.22) - (5.26) ist durch $\lambda^k = \kappa_2^{-1} \lambda_2^k$, $k \geq 0$, gegeben.

Für den Beweis siehe [6].

Unter der zusätzlichen Annahme, dass die Probleme aus (5.5) und (5.7) mit stückweise linearen finiten Elementen diskretisiert werden und dass auf die Funktionen aus (5.6) nach erfolgter Kirchhoff-Transformation eine stückweise lineare Interpolation in den Knoten des Interface angewendet wird, kann Satz 6.4 auch auf diese Diskretisierung mit den gleichen Konstanten angewendet werden. Man erhält dadurch netzunabhängige optimale Konvergenzraten und optimale Dämpfungsparameter.

Bemerkung 5.6. *Zu bemerken bleibt, dass aufgrund der Nichtlinearitäten $S_i \kappa_i$, $i = 1, 2$, in der symmetrischen Gleichung (5.14), nicht bekannt ist, ob die Dirichlet-Neumann-Methode angewendet auf (5.8) - (5.10) für den „physikalischen Druck“ $\lambda = p_{i|\Gamma}$ für hinreichend kleinen Dämpfungsparameter konvergiert.*

Im Gegensatz zur Dirichlet-Neumann-Methode, kann für die Robin-Methode nur gezeigt werden, dass die Komposition von $T_{\gamma_1, \gamma_2} : \Lambda \rightarrow \Lambda$ mit einer stetigen Transformation von Λ eine Kontraktion bildet für beliebig gewählte, aber gleiche γ_1 und γ_2 . Um die Konvergenz der Robin-Methode zu beweisen, wird Lipschitz-Stetigkeit und die strenge Monotonie von $S_1 \kappa_1$, $S_2 \kappa_2 : \Lambda \rightarrow \Lambda'$ benötigt. Aufgrund von Bemerkung 5.4 sind diese Bedingungen im Eindimensionalen erfüllt. Wiederum gibt es numerische Gegenbeispiele von nicht monotonen Operatoren, für die die nichtlineare Robin-Methode trotz allem mit realistischen Konvergenzraten und geeignet gewählten Parametern konvergiert. Im Gegensatz zur nichtlinearen Dirichlet-Neumann-Methode wird im Falle der nichtlinearen Robin-Methode, angewandt auf das diskrete Problem, keine Netzunabhängigkeit der optimalen Konvergenzraten und Parameter erzielt, das heißt die Konvergenzraten verschlechtern sich für feinere Vernetzungen.

Satz 5.6. *Sei $\gamma_1 = \gamma_2 = \gamma > 0$ und $\Omega \subset \mathbb{R}$. Für einen beliebigen Anfangswert $\lambda_2^0 \in \Lambda$ liefert der Operator $T_\gamma = T_{\gamma_1, \gamma_2}$, der definiert ist durch (5.45) bzw. (5.46), eine Folge $(\lambda_2^k)_{k \geq 0}$, die in Λ gegen den Fixpunkt von T_γ konvergiert. Weiters ergibt sich die Konvergenz der Folgen $(u_i^k)_{k \geq 1}$ und $(\kappa_i^{-1} u_i^k)_{k \geq 1}$ für $i = 1, 2$ definiert in (5.44) gegen die Lösungen von (5.5) - (5.7) bzw. (5.8) - (5.10).*

Unter Durchführung der Diskretisierung des Problems (5.5) - (5.7) in der Art, wie sie im Anschluss an Satz 5.5 angegeben ist, konvergiert die zugehörige diskrete Version der Robin-Methode gegen das diskrete Problem.

Insgesamt folgen damit für die nichtlineare Dirichlet-Neumann-Methode im Kontinuierlichen theoretische Konvergenzraten in $[0, 1)$, die, wie sich herausgestellt hat, netzunabhängig sind, also auch für die zugehörige diskrete Version zutreffen. Für die nichtlineare Robin-Methode können hingegen weder Konvergenzraten auf dem stetigen Level, noch die Netzunabhängigkeit der optimalen Konvergenzraten oder der optimalen Parameter für das diskrete Problem angegeben werden.

6 Numerisches Beispiel

Dieses Kapitel beschäftigt sich mit der numerischen Behandlung von Randwertproblemen, die neben Dirichlet- und Neumann-Randbedingungen zusätzlich auch eine Signorini-Randbedingung berücksichtigen. In Abschnitt 6.1 wird zunächst das zugrundeliegende Modellproblem betrachtet und eine Diskretisierung mit finiten Elementen angegeben. Im Anschluss widmet sich Abschnitt 6.2 der Herleitung einer diskreten Variationsungleichung, für dessen numerische Behandlung eine Aktive-Mengen-Strategie angewendet wird, auf die in Abschnitt 6.3 näher eingegangen wird. Abschnitt 6.4 beschäftigt sich abschließend mit der Implementierung eines konkreten Modellproblems sowie der Berechnung des Fehlers der erzielten Näherungslösung.

6.1 Modellproblem

Sei $\Omega \subset \mathbb{R}^2$ ein beschränktes Gebiet mit polygonalem Rand $\partial\Omega$. Der Rand $\partial\Omega$ sei die disjunkte Vereinigung der Teilmengen Γ_D , Γ_N und Γ_S :

$$\partial\Omega = \bar{\Gamma}_D \cup \bar{\Gamma}_N \cup \bar{\Gamma}_S.$$

Unter der Annahme, dass

$$f := M(u)_t - \operatorname{div}(k_r(M(u)) e_z)$$

für einen gegebenen Zeitpunkt t bekannt ist, kann das Randwertproblem für den verallgemeinerten Druck (1.20) - (1.24) in das folgende vereinfachte Modellproblem umformuliert werden:

Gesucht ist $u \in H^1(\Omega)$, sodass

$$-\Delta u = f \quad \text{in } \Omega, \quad (6.1)$$

$$u = g_D \quad \text{auf } \Gamma_D, \quad (6.2)$$

$$\frac{\partial}{\partial n} u = g_N \quad \text{auf } \Gamma_N, \quad (6.3)$$

$$u \leq g_S, \quad \frac{\partial}{\partial n} u \geq 0, \quad \frac{\partial}{\partial n} u (u - g_S) = 0 \quad \text{auf } \Gamma_S, \quad (6.4)$$

wobei $f \in H^{-1}(\Omega)$, $g_D \in H^{1/2}(\Gamma_D)$, $g_N \in H^{-1/2}(\Gamma_N)$ und $g_S \in H^{1/2}(\Gamma_S)$.

Die Signorini-Randbedingung (6.4) stellt dabei eine Verallgemeinerung der Signorini-Randbedingung (1.23), bei der $g_S = 0$ vorausgesetzt wurde, dar.

6 Numerisches Beispiel

$(\mathcal{T}_j)_{j \geq 0}$ sei eine Folge von zulässigen Unterteilungen von Ω in finite Dreieckselemente τ_l definiert wie in (4.1). Weiters wird vorausgesetzt, dass die finiten Elemente τ_l der Unterteilung formregulär sind, d.h. es gelte (4.2). Für die Definition des Volumens Δ_l , der lokalen Maschenweite h_l und des Durchmesser d_l eines finiten Elements siehe ebenfalls Kapitel 4.

6.2 Herleitung der Variationsformulierung

Für die Herleitung einer Variationsformulierung von (6.1) - (6.4) wird die erste Green'sche Formel benötigt:

$$a(u, v) = \int_{\Omega} (-\Delta u) v \, dx + \int_{\Gamma} \frac{\partial}{\partial n} u v \, ds_x \quad (6.5)$$

mit der symmetrischen Bilinearform

$$a(u, v) = \int_{\Omega} \nabla u \cdot \nabla v \, dx$$

für hinreichend oft differenzierbare Funktionen u und v .

Die Variationsformulierung für das Signorini-Randwertproblem (6.1) - (6.4) lautet somit:

Gesucht ist $u \in \mathcal{K} := \{w \in H^1(\Omega) : w = g_D \text{ auf } \Gamma_D, w - g_S \leq 0 \text{ auf } \Gamma_S\}$, sodass

$$a(u, v - u) = \int_{\Omega} f(v - u) \, dx + \int_{\Gamma_N} g_N(v - u) \, ds_x + \int_{\Gamma_S} \frac{\partial}{\partial n} u(v - u) \, ds_x \quad \forall v \in \mathcal{K}. \quad (6.6)$$

Mit der Signorini-Randbedingung (6.4) kann das Integral über den Signorini-Rand wie folgt abgeschätzt werden:

$$\int_{\Gamma_S} \frac{\partial}{\partial n} u(v - u) \, ds_x = \underbrace{\int_{\Gamma_S} \underbrace{\frac{\partial}{\partial n} u}_{\geq 0} \underbrace{(v - g_S)}_{\leq 0} \, ds_x}_{\leq 0} - \int_{\Gamma_S} \underbrace{\frac{\partial}{\partial n} u(u - g_S)}_{=0} \, ds_x \leq 0. \quad (6.7)$$

Damit ergibt sich die Variationsungleichung:

Gesucht ist $u \in \mathcal{K}$, sodass

$$a(u, v - u) \leq \int_{\Omega} f(v - u) \, dx + \int_{\Gamma_N} g_N(v - u) \, ds_x \quad \forall v \in \mathcal{K}. \quad (6.8)$$

Das nächste Ziel ist die Herleitung einer diskreten Variationsungleichung.

Als Ansatzraum wird $S_h^1(\mathcal{T}_j) \subset H^1(\Omega)$, der Raum der in $H^1(\Omega)$ stetigen Funktionen, die linear auf jedem Dreieck $\tau_l \in \mathcal{T}_j$ sind, gewählt. Das endlich dimensionale Analogon zur Menge \mathcal{K} wird wie folgt definiert:

$$\mathcal{K}_h := \{v_h \in H^1(\Omega) : v_h(x_l) = g_D(x_l) \text{ für } x_l \in \Gamma_D \text{ bzw. } v_h(x_l) - g_S(x_l) \leq 0 \text{ für } x_l \in \Gamma_S\} \quad (6.9)$$

mit den Dreiecksknoten x_l der Unterteilung \mathcal{T}_j und dem Ansatz

$$v_h(x) = \sum_{i=1}^M v_i \varphi_i^1(x),$$

wobei φ_i^1 die stückweise linearen Basisfunktionen bezeichnen.

Weiters sei ohne Beschränkung der Allgemeinheit $g_N = 0$. Damit ergibt sich die diskrete Variationsungleichung für das Signorini-Randwertproblem:

Gesucht ist $u_h \in \mathcal{K}_h$, sodass

$$a(u_h, v_h - u_h) \leq \int_{\Omega} f(v_h - u_h) dx \quad \forall v_h \in \mathcal{K}_h. \quad (6.10)$$

Äquivalent dazu ergibt sich:

Gesucht ist $\underline{u} \in \mathbb{R}^M$, sodass

$$(K_h \underline{u}, \underline{v} - \underline{u}) \leq (\underline{f}, \underline{v} - \underline{u}) \quad (6.11)$$

mit

$$K_h[l, k] := a(\varphi_k^1, \varphi_l^1), \quad f[l] := \langle f, \varphi_l^1 \rangle \quad k, l = 1, \dots, M$$

für alle $\underline{v} \in \mathbb{R}^M \leftrightarrow v_h \in \mathcal{K}_h$.

6.3 Aktive-Mengen-Strategie

In diesem Abschnitt wird näher auf die Aktive-Mengen-Strategie eingegangen, die für die numerische Behandlung der Signorini-Randbedingungen benötigt wird. Für eine weiterführende Theorie siehe zum Beispiel [20], [22].

6 Numerisches Beispiel

Es werde das folgende reine Signorini-Randwertproblem betrachtet:

$$-\Delta u = f \quad \text{in } \Omega \quad (6.12)$$

$$u \leq g_S, \quad \frac{\partial}{\partial n} u \geq 0, \quad \frac{\partial}{\partial n} u (u - g_S) = 0 \quad \text{auf } \Gamma = \partial\Omega. \quad (6.13)$$

Die grundlegende Idee der Aktiven-Mengen-Strategie zur Lösung des Randwertproblems (6.12) - (6.13) lautet nun wie folgt:

Es werde die Menge der Randknoten in eine aktive Menge und eine inaktive Menge unterteilt. Zu Beginn seien alle Randknoten aktiv. Weiters seien I_A bzw. I_I die Indexmengen der aktiven bzw. der inaktiven Knoten. Durch so eine Aufteilung der Randknoten in aktive und inaktive Knoten kann das betrachtete reine Signorini-Randwertproblem (6.12) - (6.13) in ein Randwertproblem umformuliert werden, bei dem die aktiven Knoten einer Dirichlet-Randbedingung und die inaktiven Knoten einer Neumann-Randbedingung genügen müssen. Nach Berechnung der Näherungslösung ist schließlich zu überprüfen, ob diese Lösung zulässig ist. Das heißt es muss überprüft werden, ob die Signorini-Randbedingung (6.13) in jedem Randknoten erfüllt ist. Ist diese in mindestens einem Knoten verletzt, so muss eine neue Aufteilung des Randes in aktive und inaktive Knoten vorgenommen und eine neue Näherungslösung berechnet werden. Dieses Vorgehen wird solange iteriert, bis die Signorini-Randbedingung in jedem Randknoten erfüllt ist.

Im Folgenden wird die Aktive-Mengen-Strategie im Diskreten hergeleitet.

Analog zu Abschnitt 6.2 kann nun unter Anwendung der ersten Green'schen Formel (6.5) eine Variationsformulierung für das Randwertproblem (6.12) - (6.13) hergeleitet werden. Diese lautet wie folgt:

Gesucht ist $u \in \mathcal{K} := \{w \in H^1(\Omega) : w - g_S \leq 0 \text{ auf } \Gamma\}$, sodass

$$a(u, v - u) = \int_{\Omega} f(v - u) dx + \int_{\Gamma} \frac{\partial}{\partial n} u (v - u) ds_x \quad \forall v \in \mathcal{K}. \quad (6.14)$$

Unter Berücksichtigung der Signorini-Randbedingung (6.13) kann das Randintegral aus (6.14) abgeschätzt werden, sodass sich die folgende Variationsungleichung ergibt:

Gesucht ist $u \in \mathcal{K}$, sodass

$$a(u, v - u) \leq \int_{\Omega} f(v - u) dx \quad \forall v \in \mathcal{K}.$$

Für die Herleitung einer diskreten Variationsungleichung definiere die endlich dimensionale

Menge

$$\mathcal{K}_h := \{v_h \in H^1(\Omega) : v_h(x_l) - g_S(x_l) \leq 0 \text{ für } x_l \in \Gamma\}.$$

Die diskrete Variationsungleichung für das reine Signorini-Randwertproblem (6.12) - (6.13) stimmt mit der in Abschnitt 6.2 hergeleiteten diskreten Variationsungleichung (6.10) für das Randwertproblem (6.1) -(6.4) bis auf die endlich dimensionale Menge \mathcal{K}_h überein. Äquivalent zur diskreten Variationsungleichung ergibt sich:

Gesucht ist $\underline{u} \in \mathbb{R}^M$, sodass

$$(K_h \underline{u}, \underline{v} - \underline{u}) \leq (\underline{f}, \underline{v} - \underline{u}) \quad \forall \underline{v} \in \mathbb{R}^M \leftrightarrow v_h \in K_h. \quad (6.15)$$

Definiere

$$\underline{\lambda} := K_h \underline{u} - \underline{f}.$$

Dann folgt mit (6.15)

$$0 \geq (K_h \underline{u} - \underline{f}, \underline{v} - \underline{u}) = (\underline{\lambda}, \underline{v} - \underline{u}) = \sum_{i=1}^M \lambda_i (v_i - u_i). \quad (6.16)$$

Betrachte nun für $j = 1, \dots, M$

$$v_j < u_j \leq g_{S_j}, \quad v_i = u_i \quad \text{für } i \neq j.$$

Unter Berücksichtigung von (6.16) ergibt sich für diese Wahl von \underline{v} :

$$0 \geq \sum_{i=1}^M \lambda_i (v_i - u_i) = \lambda_j \underbrace{(v_j - u_j)}_{<0}$$

und damit die Bedingung $\lambda_j \geq 0$. Andererseits folgt für $\underline{v} = \underline{g}_S$:

$$0 \geq \sum_{i=1}^M \lambda_i (v_i - u_i) = \underbrace{\sum_{i=1}^M (v_i - g_{S_i})}_{=0} - \underbrace{\sum_{i=1}^M \lambda_i (u_i - g_{S_i})}_{\substack{\geq 0 \\ \leq 0}} \geq 0$$

und damit die Bedingungen

$$u_i \leq g_{S_i}, \quad \lambda_i \geq 0, \quad \lambda_i (u_i - g_{S_i}) = 0. \quad (6.17)$$

Diese Bedingungen sind äquivalent zu

$$\lambda_i = \max \{0, \lambda_i + c(u_i - g_{S_i})\} \quad \forall c > 0.$$

Der Aktive-Mengen-Algorithmus kann damit wie folgt zusammengefasst werden:

6 Numerisches Beispiel

Zu lösen ist die diskrete Variationsungleichung:

Gesucht ist $u_h \in V := \{w_h \in H^1(\Omega) : w_h(x_l) = g_S(x_l) \text{ für } x_l \in \Gamma_D\}$, sodass

$$a(u_h, v_h) = \int_{\Omega} f v_h dx \quad (6.18)$$

für alle $v_h \in \{w_h \in H^1(\Omega) : w_h(x_l) = 0 \text{ für } x_l \in \Gamma_D\}$. Äquivalent dazu ist das Gleichungssystem:

Gesucht ist $\underline{u} \in \mathbb{R}^M$, sodass

$$(K_h \underline{u}, \underline{v}) = (\underline{f}, \underline{v}) \quad \forall \underline{v} \in \mathbb{R}^M \leftrightarrow v_h \in \mathcal{K}_h. \quad (6.19)$$

Es sei \underline{u}^k die Lösung von (6.19) im k-ten Schritt.

1. Berechne den Lösungsvektor \underline{u}^k von (6.19) sowie $\underline{\lambda}^k = K_h \underline{u}^k - \underline{f}$.

2. Überprüfe:

- $u_i^k - g_{S_i} \leq 0$ auf $\Gamma_N^k, i \in I_I$
- $\lambda_i^k \geq 0$ auf $\Gamma_D^k, i \in I_A$

3. Falls die Bedingungen erfüllt sind, dann STOP. Andernfalls berechne die neue Aufteilung des Randes Γ durch

- $\Gamma_D^{k+1} := \{x_i \in \Gamma : \lambda_i^k + c(u_i^k - g_{S_i}) > 0\}$
- $\Gamma_N^{k+1} := \{x_i \in \Gamma : \lambda_i^k + c(u_i^k - g_{S_i}) \leq 0\}$

mit einer beliebigen Konstanten $c > 0$ und gehe zurück zu 1.

6.4 Beispiel und Resultate

Es werde das Modellproblem (6.1) - (6.4) aus Abschnitt 6.1 mit $g_N = 0$ betrachtet.

Die diskrete Variationsformulierung sowie die dazu äquivalente Matrix-Vektor-Darstellung wurden bereits in Abschnitt 6.2 hergeleitet, siehe (6.10) bzw. (6.11).

Bei der Implementierung werden zunächst die lokalen Steifigkeitsmatrizen und die lokalen rechten Seiten ohne Berücksichtigung der Randbedingungen aufgestellt. Das heißt es werden für jedes Element τ_l $K_{h,l}$ und $\langle f, \varphi_l^1 \rangle_{\tau_l}$ berechnet und anschließend zu einer globalen Steifigkeitsmatrix und einem globalen Lastvektor assembliert.

Erst danach wird die Dirichlet-Randbedingung berücksichtigt. Dabei behilft man sich der Tatsache, dass die gesuchte Näherungslösung \underline{u} aufgespalten werden kann in $\underline{u} = \underline{u}_0 + \underline{u}_g$ mit

$$u_0[i] = \begin{cases} 0 & \text{für } i \in \gamma_h, \\ u[i] & \text{sonst.} \end{cases} \quad u_g[i] = \begin{cases} g_D(x_i) & \text{für } i \in \gamma_h, \\ 0 & \text{sonst.} \end{cases} \quad (6.20)$$

für $i = 1, \dots, M$, wobei γ_h die Indexmenge der Dirichlet-Randknoten bezeichne. Da die Dirichlet-Randbedingung in den zugehörigen Knoten bekannt ist, entspricht diese Aufspaltung der Näherungslösung einer Aufteilung in bereits bekannte und in unbekannte Einträge des gesuchten Lösungsvektors \underline{u} . Damit kann (6.11) wie folgt umgeschrieben werden:

Gesucht ist $\underline{u}_0 \in \mathbb{R}^M$, sodass

$$(K_h \underline{u}_0, \underline{v} - \underline{u}) \geq (\underline{f}, \underline{v} - \underline{u}) - (K_h \underline{u}_g, \underline{v} - \underline{u}). \quad (6.21)$$

Im Anschluss werden die Zeilen bzw. Spalten, die zu Dirichlet-Knoten gehören, aus der globalen Steifigkeitsmatrix K_h und aus dem Vektor der rechten Seite von (6.21) gestrichen. Die resultierende Matrix \tilde{K}_h und der resultierende Lastvektor $\tilde{\underline{f}}$ sind von der Dimension (Anzahl Knoten – Anzahl Dirichlet-Knoten). Die Matrix \tilde{K}_h ist außerdem symmetrisch und positiv definit, woraus die Existenz ihrer Inversen folgt, die bei der Berücksichtigung der Signorini-Randbedingung notwendig sein wird.

In Abschnitt 6.3 wurde bereits eine Aktive-Mengen-Strategie vorgestellt, die in weiterer Folge auf das Randwertproblem (6.24) - (6.27) bzw. auf die zugehörige diskrete Variationsungleichung angewendet werden soll. Wie zuvor gezeigt, können die Signorini-Knoten aus der aktiven Menge als Dirichlet-Knoten, die der Randbedingung

$$u[i] = g_S(x_i) \quad \forall i \in I_A \quad (6.22)$$

genügen, betrachtet werden. Die grundlegende Idee ist jetzt, diese Dirichlet-Randbedingungen (6.22) als Nebenbedingung eines Sattelpunktproblems zu berücksichtigen, wodurch insbesondere eine direkte Berechnung von $\underline{\lambda} := \tilde{K}_h \underline{u} - \tilde{\underline{f}}$ gewährleistet wird.

Das zu lösende lineare Gleichungssystem lautet schließlich:

$$\begin{pmatrix} \tilde{K}_h & -B_h^T \\ B_h & 0 \end{pmatrix} \begin{pmatrix} \underline{u} \\ \underline{\lambda} \end{pmatrix} = \begin{pmatrix} \tilde{\underline{f}} \\ \underline{g} \end{pmatrix}$$

wobei $g[i] = g_S(x_i)$ für $i \in I_A$. $B_h \in \mathbb{R}^{M_A \times M}$ mit der Anzahl der aktiven Knoten M_A ist eine Matrix, die in jeder Zeile genau eine Eins enthält und zwar in jener Spalte, die der globalen Knotennummer des betrachteten aktiven Knotens entspricht (nach Streichung der

6 Numerisches Beispiel

Dirichlet-Knoten).

Da die Matrix \tilde{K}_h invertierbar ist, kann zunächst die Inverse iterativ mit dem CG-Verfahren mit einer relativen Genauigkeit $\varepsilon = 10^{-8}$ bestimmt und im Anschluss das Gleichungssystem in das nachfolgende Schur-Komplement-System umgeformt werden:

$$B_h \tilde{K}_h^{-1} B_h^T \underline{\lambda} = \underline{g} - B_h \tilde{K}_h^{-1} \underline{f}.$$

Obiges Schur-Komplement-System wird jetzt mit dem CG-Verfahren mit diagonaler Vorkonditionierung nach $\underline{\lambda}$ aufgelöst. Die Lösung \underline{u} folgt schließlich aus

$$\underline{u} = \tilde{K}_h^{-1} B_h^T \underline{\lambda} + \tilde{K}_h^{-1} \underline{f}.$$

Laut Schritt 2 der Aktive-Mengen-Strategie müssen nun die Bedingungen

$$\lambda[i] \geq 0 \quad \forall i \in I_A, \quad u[i] - g_S(x_i) \leq 0 \quad \forall i \in I_I \quad (6.23)$$

überprüft werden.

Falls alle Bedingungen erfüllt sind, muss nur noch der Lösungsvektor \underline{u} auf die volle Dimension M (=Anzahl der Knoten) gebracht werden. Dies erfolgt, indem an die Stellen des Lösungsvektors, die zu Dirichlet-Knoten gehören, der Dirichlet-Randwert g_D geschrieben wird.

Falls die Bedingung jedoch für mindestens einen Signorini-Knoten verletzt ist, so werden die aktive und die inaktive Menge laut Schritt 3 der Aktive-Mengen-Strategie neu bestimmt, wobei hier die Konstante $c = 1$ gewählt wird. Konkret bedeutet das, dass ein aktiver Knoten für den die zugehörige Bedingung $\lambda \geq 0$ nicht erfüllt ist, im nächsten Iterationsschritt inaktiv wird und umgekehrt.

Diese Vorgehensweise wird so lange iteriert, bis die Bedingungen für alle Signorini-Knoten erfüllt sind und somit eine Näherungslösung des Randwertproblems gefunden ist.

Die nächste Aufgabe besteht nun darin, den Fehler der berechneten Näherungslösung des nachfolgenden Modellproblems in der L_2 -Norm anzugeben.

Gesucht ist u , sodass

$$-\Delta u(x, y) = -2 \quad \text{in } \Omega \subset \mathbb{R}^2, \quad (6.24)$$

$$u(x, y) = x^2 \quad \text{auf } \Gamma_D, \quad (6.25)$$

$$\frac{\partial}{\partial n} u(x, y) = 0 \quad \text{auf } \Gamma_N, \quad (6.26)$$

$$u(x, y) \leq \frac{1}{4}, \quad \frac{\partial}{\partial n} u(x, y) \geq 0, \quad \frac{\partial}{\partial n} u(x, y) \left[u(x, y) - \frac{1}{4} \right] = 0 \quad \text{auf } \Gamma_S, \quad (6.27)$$

mit $\Omega = (0, 1) \times (0, 1)$.

Die Fehlerberechnung wird dabei unter der Annahme, dass die exakte Lösung $u = x^2$ bekannt ist, durchgeführt.

Zu bestimmen ist das nachfolgende Integral

$$\|u - u_h\|_{L^2(\Omega)}^2 = \int_{\Omega} |u(x) - u_h(x)|^2 dx = \sum_{l=1}^N \int_{\tau_l} |u(x) - u_h(x)|^2 dx.$$

Dieses Integral kann durch Transformation auf das Referenzelement τ und anschließender numerischer Integration durch Anwendung einer 7-Punkt-Gauss-Formel für Dreiecksbereiche berechnet werden.

Die nachstehende Tabelle gibt den berechneten Fehler der Näherungslösung des betrachteten Modellproblems (6.24) - (6.27) in der L_2 -Norm an.

N	M	DoF	It	$\ u - u_h\ _{L_2(\Omega)}$	eoc
16	13	5	6	$3.077 - 2$	
64	41	25	7	$7.769 - 3$	1.986
256	145	113	8	$1.949 - 3$	1.995
1024	545	481	9	$4.856 - 4$	2.005
4096	2113	1985	10	$1.213 - 4$	2.001
16384	8321	8065	11	$3.031 - 5$	2.000

Dabei steht N für die Anzahl der Elemente, M für die Anzahl der Knoten, DoF für die Anzahl der Freiheitsgrade, It für die Anzahl der Iterationen der Aktiven-Mengen-Strategie und eoc bezeichnet die „experimental order of convergence“, also die Konvergenzordnung. Mit der Theorie übereinstimmend ergibt sich für den L_2 -Fehler der Näherungslösung eine quadratische Konvergenzordnung.

Literaturverzeichnis

- [1] R. A. Adams. *Sobolev Spaces*. Academic Press, New York, London, 1975.
- [2] H. W. Alt. *Lineare Funktionalanalysis*. Springer, Berlin, Heidelberg, 2006.
- [3] H. W. Alt and S. Luckhaus. Quasilinear elliptic-parabolic differential equations. *Math. Z.*, 183:311–341, 1983.
- [4] J. Appell and P. P. Zabrejko. Nonlinear superposition operators. *Cambridge University Press*, 1990.
- [5] J. Bear. *Dynamics in porous media*. Dover Publications, New York, 1988.
- [6] H. Berninger. *Domain decomposition methods for elliptic problems with jumping nonlinearities and application to the Richards equation*. PhD thesis, FU Berlin, 2007.
- [7] H. Berninger. Non-overlapping domain decomposition for the Richards equation via superposition operators. *Domain Decomposition Methods in Science and Engineering XVIII, LNCSE*, pages 169–176, 2009.
- [8] H. Berninger, R. Kornhuber, and O. Sander. Fast and robust numerical solution of the Richards equation in homogeneous soil. 2010.
- [9] H. Berninger, R. Kornhuber, and O. Sander. Convergence behaviour of Dirichlet-Neumann and Robin methods for a nonlinear transmission problem. *Domain Decomposition Methods in Science and Engineering XIX*, pages 87–100, 2011.
- [10] F. Bonani and G. Ghione. On the application of the Kirchhoff transformation to the steady-state thermal analysis of semiconductor devices with temperature-dependent and piecewise inhomogeneous thermal conductivity. *Solid-St. Electron.*, 38(7):1409–1412, 1995.
- [11] S. Brenner and R. L. Scott. *The mathematical theory of finite element methods*. Springer, New York, 1994.
- [12] G. Chavent and J. Jaffré. *Dynamics of fluids in porous media*. Elsevier Science, 1986.
- [13] P. G. Ciarlet. *The finite element method for elliptic problems*. North-Holland, 1978.
- [14] M. Discacciati. *Domain decomposition methods for the coupling of surface and ground-water flows*. PhD thesis, Ecole Polytechnique Fédérale de Lausanne (EPFL), 2004.

- [15] I. Ekeland and R. Temam. *Convex analysis and variational problems*. North Holland Publications, Amsterdam-Oxford, 1976.
- [16] R. Eymard, T. Gallouet, R. Herbin, M. Gutnic, and D. Hilhorst. Approximation by the finite volume method of an elliptic-parabolic equation arising in environmental studies. *Math. Models Methods Appl. Sci.*, 11(9):1505–1528, 2001.
- [17] J. Fuhrmann. *Zur Verwendung von Mehrgitterverfahren bei der numerischen Behandlung elliptischer partieller Differentialgleichungen mit variablen Koeffizienten*. PhD thesis, TU Chemnitz-Zwickau, 1994.
- [18] J. Fuhrmann. On numerical solution methods for nonlinear parabolic problems. *Modeling and Computation in Environmental Sciences, First GAMM-Seminar at ICA Stuttgart*, pages 170–180, 1997.
- [19] M. J. Gander. Optimized Schwarz methods. *SIAM J. Numer. Anal.*, 44(2):699–731, 2006.
- [20] R. Glowinski. *Numerical methods for nonlinear variational problems*. Scientific Computation. Springer, Berlin, 2008. Reprint of the 1984 original.
- [21] H. Heuser. *Lehrbuch der Analysis*. Teubner, Stuttgart, 2000.
- [22] M. Hintermüller, K. Ito, and K. Kunisch. The primal-dual active set strategy as a semismooth Newton method. *SIAM J. Optim.*, 13(3):865–888, 2002.
- [23] M. Jung and U. Langer. *Methode der finiten Elemente für Ingenieure*. Teubner, Stuttgart-Leipzig-Wiesbaden, 2001.
- [24] K. Königsberger. *Analysis 1*. Springer, 1990.
- [25] R. Kornhuber. *Adaptive monoton multigrid methods for nonlinear variational problems*. Teubner, Stuttgart, 1997.
- [26] P. L. Lions. On the Schwarz alternating method. III: A variant for nonoverlapping subdomains. In *Dom. Decom. Meth. for Part. Diff. Eq., Proc. 3rd Int. Symp.*, pages 202–223. SIAM, 1990.
- [27] M. Marcus and V. J. Mizel. Every superposition operator mapping one Sobolev space into another is continuous. *J. Funct. Anal.*, 33:217–229, 1979.
- [28] W. McLean. *Strongly elliptic systems and boundary integral equations*. Cambridge University Press, 2000.
- [29] A. Quarteroni and A. Valli. *Domain decomposition methods for partial differential equations*. Oxford University Press, USA, 1999.

- [30] L. A. Richards. The usefulness of capillary potential to soil-moisture and plant investigators. *Jour. Agr. Res.*, 37(12):719–742, 1928.
- [31] R. T. Rockafellar and R. J.-B. Wets. *Variational Analysis*. Springer, 1998.
- [32] H. R. Schwarz. *Methode der finiten Elemente*. Teubner, Stuttgart, 1980.
- [33] O. Steinbach. *Numerische Näherungsverfahren für elliptische Randwertprobleme. Finite Elemente und Randelemente*. Teubner, Stuttgart-Leipzig-Wiesbaden, 2003.
- [34] M. T. van Genuchten. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci. Soc. Am. J.*, 44:892–898, 1980.
- [35] J. Wloka. *Funktionalanalysis und Anwendungen*. Walter de Gruyter, Berlin, 1971.
- [36] J. Wloka. *Partielle Differentialgleichungen*. Teubner, Stuttgart, 1982.

